3 Thermodynamics

Thermodynamics describes the macroscopic behavior of matter, which is an ensemble of atoms or molecules. In this chapter, we review the thermodynamics necessary for the understanding of cloud formation. We start by defining some fundamental terms (Section 3.1) that will be used throughout the textbook. We discuss the physical laws that describe processes in dry air and introduce the concept of *air parcels* in Section 3.2. Since thermodynamic processes that involve adiabatic changes can best be visualized with the help of thermodynamic charts, the use of tephigrams is introduced in Section 3.3. Water is a unique substance because it exists in all three phases (ice, liquid water, and water vapor) in Earth's atmosphere. The transitions between its different phases are crucial for the understanding of liquid and ice cloud formation and microphysical processes inside clouds. They are the subject of Section 3.4. We end the chapter by extending the laws introduced for dry air to moist air (Section 3.5), to analyze how water vapor influences certain processes and how this relates to cloud formation.

3.1 Basic Definitions

3.1.1 Thermodynamic States and State Variables

The aim of thermodynamics is to describe the spatially and temporally averaged properties of macroscopic systems of gases, liquids, or solids. The concepts of thermodynamics are suitable for systems containing a huge number of molecules on the order of the Avogadro constant ($N_A = 6.022 \times 10^{23}\,\mathrm{mol}^{-1}$) and that are slowly changing in time. We start by considering a large, macroscopic quantity of a substance and neglecting surface effects. Surface effects will become important when we discuss cloud droplet formation in Chapter 7.

Let us consider the quantity X, which contains a large number of molecules N. Each molecule can be described by six coordinates: three for its spatial position x, y, z, and three for its momentum components (mass times velocity) mu, mv, mw. A system of N molecules is called a *microstate* of the system. For most practical purposes, to describe a system in terms of its microstates is too complicated and cumbersome, as it implies solving the equations of motion with $6N$ degrees of freedom. Instead, in thermodynamics, we describe a system via macroscopic, observable quantities like atmospheric pressure p, temperature T, volume V, number of moles n, entropy S, and internal energy U. The main advantage of a thermodynamic, macroscopic description of a system, as opposed to a microscopic (atomistic) description, is that a macrostate given, for example, by p, T, V, and n summarizes an infinite number of microstates leading to the same macrostate. For example, the microstates of n moles of a gas contained in a

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

cylinder with volume V at pressure p and temperature T change very fast due to the Brownian motion of the molecules, which is averaged out when considering the macrostate. Hence, the macrostate remains constant.

The macroscopic, observable quantities describing a macrostate are termed *state variables*. A set of state variables unambiguously describes a state of thermodynamic equilibrium of a system. For example, the temperature of an air parcel is unambiguously defined when its pressure, density, and gas constant are known (eq. 3.5). How the system reached that state is of no relevance for the value of the state variables. In contrast, work W is an example of a quantity that depends on the process connecting two states of a system and therefore cannot serve as a state variable. The number of state variables required depends on the system under consideration. Because different relations exist between state variables, they depend on each other. For example, the ideal gas law (eq. 3.7) relates the state variables p, T, and ρ, leaving two out of three of these variables to be varied independently.

3.1.2 Extensive and Intensive Variables

An *extensive variables* is a state variables whose value is proportional to the size, for example, volume, of a system. For example, the state variables internal energy U, volume V, number of moles n, and entropy S scale with the amount of matter the system contains. Hence, they are extensive state variables.

Intensive variables are normalized such that they are independent of the size of the system. Examples of intensive variables are pressure p, temperature T, density ρ, and the chemical potential μ. We can assign a conjugate intensive variable a_i to every extensive variable A_i to get rid of the dependence of the size of the system. For this, we take the partial derivative with respect to any variable x_i while holding all other extensive variables constant:

$$a_i = \left(\frac{\partial A}{\partial x_i}\right)_{x_{j\neq i}}. \tag{3.1}$$

An example is a gas described by the state variables p, V, T, and S. The first law of thermodynamics states that:

$$dU = dW + dQ = -pdV + TdS. \tag{3.2}$$

If we consider an infinitesimal change dV of the volume, while all other extensive variables are held constant, so that in this case $dS = 0$, we have $dU = -pdV$, or

$$p = -\left(\frac{\partial U}{\partial V}\right)_S. \tag{3.3}$$

Hence, p is the conjugate intensive variable of V. This implies that at higher pressures, a change in the gas volume dV leads to a greater decrease of internal energy dU since more work is associated with a given change in volume. Note that we adopt the usual notation from thermodynamics, where dX denotes an infinitesimal change in the quantity X. Infinitesimal variations allow us to study the behavior of a system without moving it out of *thermodynamic equilibrium* (TDE).

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

Intensive variables can also be obtained by dividing extensive variables by the mass contained in the system. These quantities are called *specific quantities*. Specific quantities are usually denoted with lowercase letters, for example, we write the specific internal energy as u, the specific entropy as s, and the specific volume as v.

3.1.3 Open, Closed, and Isolated Systems

The boundaries of an *open system* are defined to be permeable to both mass and energy (Figure 3.1a). A *closed system*, on the other hand, can be thought of as having "rigid" walls that do not allow the exchange of mass with its surroundings (Figure 3.1b). However, a closed system may exchange energy with its surroundings, which is the case when it is in contact with a heat reservoir or through radiative heating/cooling. If a system is *isolated*, neither mass nor energy can be exchanged via the boundaries of the system. This does not exclude any interaction with the outside world, however. Long-range forces, like gravity $\vec{F}_g$, although originating outside an isolated system, can nevertheless have an effect on it (Figure 3.1c).

3.1.4 Thermodynamic Equilibrium

The state of TDE of an isolated system is characterized by the absence of any net macroscopic flow of matter or energy within it. This implies that all observable quantities, the state variables, are constant in time once the system has reached TDE. This does not apply to the microstate of the system, which will continuously change by microscopic fluctuations even when in TDE. In thermodynamics, it is assumed that a system starting off in any non-equilibrium state spontaneously moves toward its state of TDE, which is reached after a sufficiently long time. TDE is the only thermodynamically well-defined state of a system, being the only state which is unambiguously determined by the values of the system's state variables. Spatial invariance, that is, homogeneity, of the intensive variables is not equivalent to the

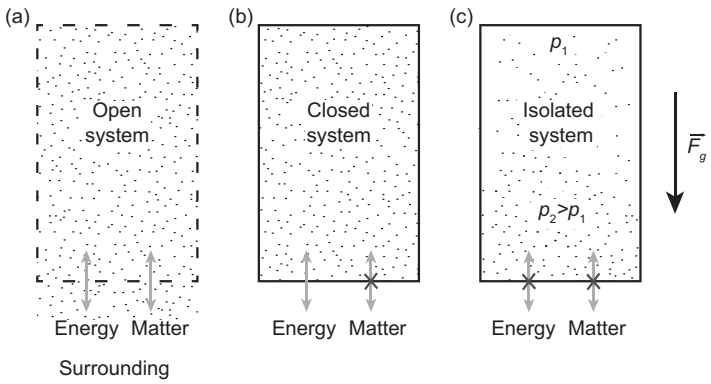


Figure 3.1 Schematic of (a) an open system, (b) a closed system, and (c) an isolated system, and the associated absence and/or presence of a possibility to exchange energy and mass with the surrounding. The isolated system in (c) is shown in the presence of a downward-oriented gravity force $\vec{F}_g$, where the pressure p increases downward.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

system being in TDE: If more than one phase of a substance co-exists in TDE, some intensive variables, for example, the density, may be inhomogeneous throughout the system. As an example, an isolated system containing liquid water and water vapor in TDE shows a jump in density upon phase transition from the gas to the liquid phase or vice versa.

If the system is subject to long-range forces like gravity, these may give rise to gradients in some internal variables. Consider an isolated system consisting of a gas in a container subject to an external constant gravitational force (Figure 3.1c). In TDE, a pressure gradient exists within the gas along the direction of the gravitational force. Temperature is the only intensive variable which is always homogeneous throughout a system in TDE, because any temperature gradient is inevitably associated with a net flux of heat, which is in contradiction with TDE.

Note that a system in TDE is always in a stationary state, being in a state in which all observables are independent of time. However, a system in a stationary state is not necessarily in TDE. As a simple example, consider a metal rod with its two ends kept at different temperatures (Figure 3.2a). As long as those temperatures do not change, the system is in a stationary state. However, the inhomogeneity of temperature along the metal rod and the consequent net heat flux through it from the end with higher temperature T_2 toward the end with lower temperature T_1 are incompatible with the definition of TDE. Only after the heat reservoirs have been decoupled and the temperature along the metal rod is the same everywhere, the rod is in TDE (Figure 3.2b).

The concept of TDE can also be applied to the interaction of two or more systems with each other, called *mutual TDE* (Figure 3.3). The TDE-defining conditions of the absence of any net flux of matter or energy still holds, but are no longer sufficient. Consider two gas containers at different pressures with one mutual wall, in which a shutter is implemented. If the shutter is closed, and both containers are at the same temperature, no flux of matter or energy exists, and each container is in TDE. Nevertheless, the system comprising both containers is not in

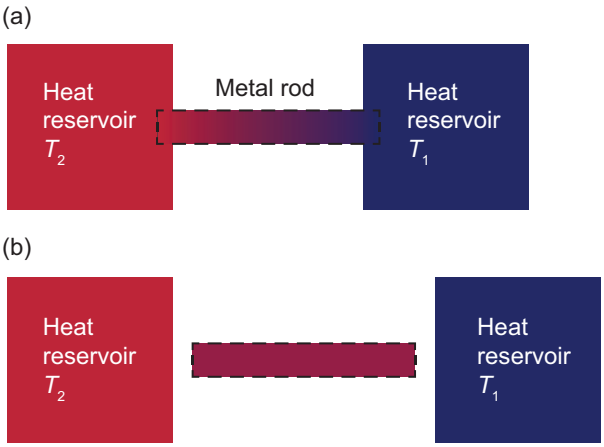


Figure 3.2 Schematic of (a) the stationary state of a metal rod coupled to two heat reservoirs at both ends with $T_2 > T_1$ and (b) the situation after the heat reservoirs have been decoupled.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

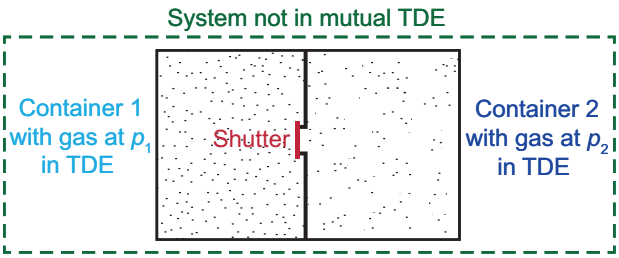


Figure 3.3 Schematic of a system with two gas containers at different pressures p_1 and p_2 being separated by a shutter.

TDE, since the shutter acts as a mechanical constraint between the two containers. This constraint artificially prevents the systems from exchanging gas, even though the two containers are at different pressures. In fact, TDE requires all kinds of equilibria between systems to hold simultaneously. In other words, the systems have to be simultaneously in mechanical ($p_1 = p_2$), thermal ($T_1 = T_2$), chemical ($\mu_1 = \mu_2$), and radiative ($F_1 = F_2$) equilibria. In the example in Figure 3.3, the requirement of mechanical equilibrium is not met. Hence, Earth's atmosphere as a total is not in TDE, because differences in outgoing longwave and incoming shortwave radiation lead to a radiative imbalance, causing temperature gradients.

3.1.5 Reversible and Irreversible Processes

A *reversible process* connects two states of TDE with infinitesimal steps on a path where every intermediate state is also a state of TDE. In turn, an *irreversible process* contains at least some intermediate states along its path that are not states of TDE. Usually, an irreversible process moves a system from a state that is out of TDE into a state that is in TDE.

An example of a reversible process is the isothermal expansion of a gas, contained in a cylinder with a movable piston that is in contact with a heat reservoir, as shown in Figure 3.4a. During the isothermal expansion, the gas (or system) does work against the external force applied by the piston. This work dW is balanced by heat delivered to the gas from the heat reservoir dQ. Reversing the process is possible by simply moving the piston back and doing the necessary work, allowing the released heat to be transferred back to the reservoir. However, if a gas is allowed to escape from a container, there is no simple way to reverse this process. Its irreversibility is intuitively plausible because it is highly unlikely that the random motions of the escaped gas molecules would bring all of them simultaneously back to their original position within the container.

Another example of an irreversible process is heat conduction. Consider the metal rod in Figure 3.2, which is initially in contact at each end with heat reservoirs having different temperatures. When the rod is then isolated from the heat reservoirs, heat will flow in such a way that the temperature equilibrates along the rod, moving the rod from a state of non-equilibrium into a state of TDE. The opposite process will never occur, as heat never flows from colder to warmer temperatures. The heat engine (or Carnot engine) is the only way to reversibly transfer heat between two reservoirs at different temperatures (Section 3.2.5).

A powerful, yet less intuitive, criterion for the reversibility of a process is its entropy balance, as will be discussed in Section 3.2.7. In a reversible process, no net entropy is generated by the system and its surroundings, that is, $dS = 0$. In the earlier example of reversible expansion of

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

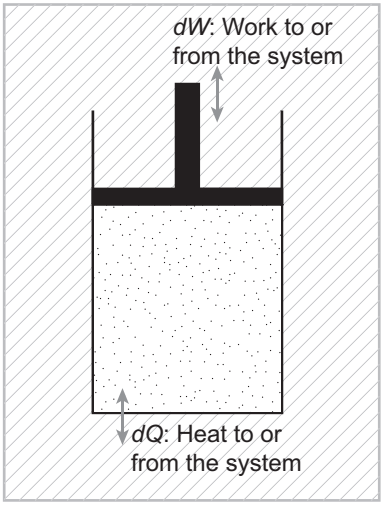


Figure 3.4 Schematic of a reversible isothermal process.

a gas in a cylinder with a piston, the entropy of the expanding gas increases, but at the same time, the entropy of the connected heat reservoir decreases by the same amount, leading to zero change in the total entropy. During the irreversible free expansion of a gas in a container, however, there is no heat transfer from a reservoir to balance the increase of entropy in the expanding gas, and so net entropy is generated by this irreversible process. This is stated in the second law of thermodynamics:

$$dS \geq 0, \tag{3.4}$$

where the equality applies if the process is reversible. This expression is also referred to as the *Clausius*[1] *inequality* (Atkins et al., 2022).

Most processes in the real world are irreversible. The reversibility of the expansion of a gas in a container relies on the assumption that the piston moves without any friction, which is not realistic. Friction irreversibly converts some work into heat. As soon as such losses are encountered, the process under consideration cannot be considered reversible; the more important the losses are, the larger is the deviation of the process from reversibility. Processes that come very close to reversibility in the real world are phase transitions. Take the example of evaporation of a liquid, for example, water, which co-exists in TDE with its vapor phase at its boiling point. The addition of some quantity of heat to the system will evaporate some liquid. As long as the addition of heat occurs slowly enough, the temperature of the system will not move from the boiling point (assuming that the pressure of the gas phase stays constant and that its volume is allowed to increase) and the liquid stays in co-existence with its gas phase in TDE at every intermediate point along the process. In practice, mechanical action inevitably entails some frictional losses, causing some irreversibility.

[1] www.britannica.com/biography/Rudolf-Clausius

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

3.2 Dry Air

The goal of this section is to derive the law that relates the changes in the intensive state variables pressure p and temperature T of an air parcel in dry air, and thus explain the processes occurring in dry air. In volume mixing ratios (volume of one substance per volume of dry air, often expressed in percent), dry air is mainly composed of nitrogen, N_2 (78.08 %) and oxygen, O_2 (20.95 %). Argon (Ar) follows in third place with 0.93 %. Together, these three gases comprise about 99.96 % and are well mixed throughout the troposphere. There are various other gases that collectively make up the remaining 0.04 %. These are called *trace gases*, such as carbon dioxide, CO_2, and methane, CH_4 (Masson-Delmotte et al., 2021).

Here, we assume that the atmosphere can be treated as an *ideal gas*. Ideal gases are defined as gases in which the volume of the individual molecules can be neglected and where no interactions between the molecules exist besides collisions, that is, there are no forces acting between the point-like molecules except when they collide. From this, Dalton's law (Dalton, 1802) follows, as will be discussed in Section 3.2.2.

3.2.1 Ideal Gas Law

The ideal gas law describes the relationship between the state variables for an ideal gas. In its most general form, the ideal gas law can be written as

$$pV = nR^*T, \tag{3.5}$$

where p is the pressure in Pa, V is the volume in m^3, T is the absolute temperature in K, $R^* = 8.314\,\mathrm{J\,mol^{-1}\,K^{-1}}$ is the universal gas constant, which does not depend on the type of gas, and n is the number of moles of the gas. n can be expressed as $n = m/M$ being the ratio of the gas mass, m to the molar mass of the gas, M. For dry air, we can replace n in eq. (3.5) by $n = m_d/M_d$:

$$\frac{pV_dM_d}{m_d} = \frac{pM_d}{\rho_d} = R^*T, \tag{3.6}$$

where the subscript "d" denotes dry air, m_d is the mass of dry air, $\rho_d = m_d/V_d$ the density of dry air, and $M_d = 28.96\,\mathrm{g\,mol^{-1}}$ is the molecular weight of dry air, given by the sum of the mole fractions of each gas component in dry air multiplied by their corresponding molar mass. Note that the value of M_d takes into account trace gases and thus is slightly larger than the sum of the molar mass weighted by mole fraction of the three major components, N_2, O_2, and Ar. By replacing V_d, we have an expression to describe the state of an ideal gas with only intensive variables. Dividing both sides of eq. (3.6) by M_d, we arrive at the ideal gas law for dry air given by

$$\frac{p}{\rho_d} = R_dT, \tag{3.7}$$

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

where R_d is the specific gas constant of dry air:

$$R_d = \frac{R^*}{M_d} = \frac{8.314\,\mathrm{J\,mol^{-1}\,K^{-1}}}{28.96\,\mathrm{g\,mol^{-1}}} = 287.1\,\mathrm{J\,kg^{-1}\,K^{-1}}. \tag{3.8}$$

The ideal gas law is also referred to as the *equation of state* for an ideal gas, which means that it describes the state of an ideal gas in TDE by defining a relationship between the state variables T, p, and ρ_d.

The ideal gas law is an approximation to reality as it neglects molecular size and intermolecular forces. In a real gas such as air, molecules with different sizes co-exist and at least van der Waals[2] forces exist. These are rather weak electric attractive forces between neutral atoms or molecules that decay with the sixth power of their distance. How well the ideal gas law applies to a given gas also depends on its density. Neglecting molecular size becomes less important for lower gas densities, because the average distance between neighbouring molecules becomes much larger than the molecular size. The assumptions of an ideal gas works well for air, because it is a low-density gas and consists mainly of neutral molecules. For air, the deviation of $p/(\rho_d R_d T)$ from 1, being the difference between the real density of air and ρ_d, is less than 0.2 % (Iribane and Godson, 1981). Above an altitude of about 100 km (~50 Pa), the air molecules are charged, which initiates ionic interactions between the air molecules. Thus, the assumption of the ideal gas law that intermolecular forces are absent is no longer valid here. As the ideal gas law assumes TDE, the assumption of TDE also breaks down above 100 km.

3.2.2 Partial Pressure

For gas mixtures, the most practical quantity to describe the relative amounts of each component in the mixture is their partial pressure. The partial pressure of a gas component is defined as the pressure that component i would exert on the walls of a volume V if the volume V was occupied by only component i, with all other components removed. The ideal gas law directly links the partial pressure p_i with the number concentration n_i of gas molecules of component i in the mixture:

$$n_i = \frac{p_i}{k_B T} = \frac{p_i N_A}{R^* T}, \tag{3.9}$$

where $k_B = 1.38 \times 10^{-23}\,\mathrm{J\,K^{-1}}$ is the Boltzmann constant. The sum of the partial pressures of all components constituting the system is equal to the total pressure of the gas mixture:

$$\sum_i p_i = p. \tag{3.10}$$

This is known as Dalton's law[3] (Dalton, 1802). In moist air, the partial pressure of water vapor, e, is often used to characterize the water vapor content in the atmosphere. Alternatively, the density of water vapor, ρ_v, also referred to as the *absolute humidity*, can be used. As moist air also approximately behaves as an ideal gas, and because the ideal gas law applies to the partial

[2] www.britannica.com/biography/Johannes-Diederik-van-der-Waals
[3] www.britannica.com/biography/John-Dalton

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

pressures of the individual components in a gas mixture, ρ_v can be linked to e via the ideal gas law:

$$e = R_v \rho_v T, \tag{3.11}$$

where $R_v = R^*/M_w = 461.5\,\mathrm{J\,kg^{-1}\,K^{-1}}$ is the specific gas constant of water vapor and $M_w = 18\,\mathrm{g\,mol^{-1}}$ is the molecular weight of water.

3.2.3 First Law of Thermodynamics

The first law of thermodynamics states that, for any process, energy is conserved; and that heat is a form of energy. As stated in Section 3.1.2, we use specific quantities in atmospheric science, which are quantities normalized to unit mass and are denoted by lowercase letters. We start by defining the specific internal energy u. From here on, the term "specific" is dropped for brevity. A change in internal energy can result from a change in the kinetic energy of the molecules in at least one of the available degrees of freedom (translation, rotation, vibration) or by a change in the forces of attraction between the molecules (Curry and Webster, 1999). Any of these changes will affect the temperature of the system. For ideal gases, where intermolecular forces are neglected, u is proportional to T because it consists only of the kinetic energy of the molecules.

As the first law of thermodynamics states that energy is conserved, any change in the internal energy du has to be balanced by changes in other quantities of energy. These could be a quantity of heat dq supplied to or subtracted from a unit mass of air or a quantity of work dw done by or exerted on the same unit mass of air, or both. The basic form of the first law of thermodynamics can thus be described as:

$$du = dq + dw. \tag{3.12}$$

Here, we apply the convention that all net transfer rates of energy in the form of heat or work to the system are positive ($q > 0$ or $w > 0$) and all net transfer rates of energy away from the system are negative ($q < 0$ or $w < 0$), following Atkins et al. (2022). Equation (3.12), even if already written in specific quantities, is still not very practical for applications in atmospheric sciences because changes in u, q, and w cannot readily be measured. The goal is thus to rewrite the first law in such a way that temperature and pressure become the independent variables, as their changes can be easily measured in the atmosphere.

Let us again consider a gas ("a system") in a cylinder with a movable piston (Figure 3.5). The work dW done by the gas when expanding against an external pressure p and increasing its volume is given by:

$$dW = \vec{F} \cdot \vec{ds} = -Fds = -pAds = -pdV \cdot \tag{3.13}$$

Here, $\vec{F}$ is the external pressure force, $\vec{ds}$ is the displacement of the piston of area A, and Ads equals the change in volume, dV, during the expansion. The minus sign arises because the system resists the external pressure force upon expansion and thus performs work on the

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

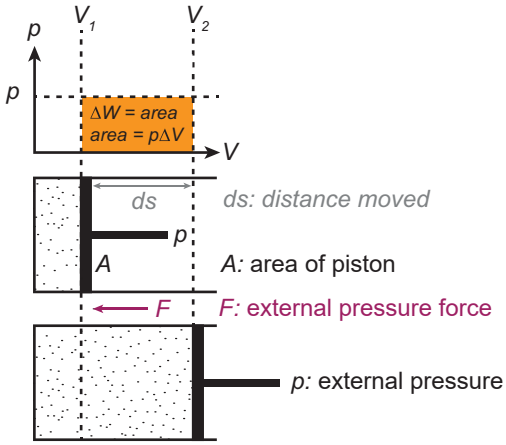


Figure 3.5 Schematic of a gas enclosed within a cylinder with a movable piston that expands against a constant external (environmental) pressure, p.

environment. More intuitively, we can imagine that when the gas expands, dV is positive. In this case, dW is negative and the gas performs work on the environment, hence the internal energy of the system (gas) decreases. By contrast, if the gas gets compressed, dV is negative. In this case, dW is positive, as work is performed by the environment on the gas.

Normalizing eq. (3.13) by the mass leads to an expression for the work in specific quantities:

$$dw = -pd\alpha, \tag{3.14}$$

where α is the volume per unit mass, called the *specific volume*. It is the inverse of the density: $\alpha = \rho^{-1}$. Note that p denotes both the pressure within the system and the external pressure, since we are dealing with reversible processes where every state along the process is a state of TDE. A state with $p_{int} \neq p_{ext}$ would violate the definition of TDE. In real-world processes, the equality $p_{int} = p_{ext} = p$ is rarely fulfilled.

If we add a quantity of heat dq to a unit mass of material, then T increases by dT. In the absence of phase transitions, we have

$$dq = cdT, \tag{3.15}$$

where c is the specific heat capacity ($\mathrm{J\,kg^{-1}\,K^{-1}}$). For gases, c depends on the process and whether work is done while heat is added.

If no work is done on or by the system, the volume remains constant, $d\alpha = 0$. In our example with the gas cylinder with the movable piston, $d\alpha = 0$ only if the pressure in the system (cylinder) equals the external pressure. Then, the first law reduces to

$$du = dq. \tag{3.16}$$

We now rewrite eq. (3.15) in terms of derivatives, indicating that volume is kept constant ($d\alpha = 0$) by the use of the subscript v. We obtain:

$$c_v = \left(\frac{dq}{dT}\right)_v = \left(\frac{du}{dT}\right)_v. \tag{3.17}$$

where we made use of eq. (3.16).

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

This expression relates the specific heat capacity at constant volume, c_v, to the change in heat or internal energy with temperature, respectively. We can also define the specific heat capacity at constant pressure, c_p, which is given as

$$c_p = \left(\frac{dq}{dT}\right)_p . \tag{3.18}$$

In an isobaric process ($dp = 0$), the volume is no longer constant ($d\alpha \neq 0$). Thus, if heat is supplied to an air mass, it expands, and part of the supplied heat dq is transferred to the work dw required to expand a unit mass of air against the external, environmental (ambient) pressure. Here, work is done by the system on the environment. In an isochoric process ($d\alpha = 0$), the entire amount of supplied heat is transferred to internal energy. Consequently, the temperature change is larger in an isochoric process than in an isobaric process for the same dq. Therefore, the heat capacity at constant volume is smaller than that at constant pressure. At 273.15 K, c_p and c_v take the following values for dry air (Curry and Webster, 1999):

$$c_p = 1005\,\mathrm{J\,kg^{-1}\,K^{-1}}, \tag{3.19}$$

$$c_v = 718\,\mathrm{J\,kg^{-1}\,K^{-1}}. \tag{3.20}$$

c_p of dry air depends weakly on temperature as given in eq. (A.6) and Table 3.1 (Tsilingiris, 2008).

In (most) real situations, neither the pressure nor the volume is constant, and there are contributions to both the internal energy and to the work exerted by the system. When adding a total amount of heat dq to a system, the amount that goes into the internal energy is given by

$$du = c_v dT. \tag{3.21}$$

The remainder of the added heat goes into the work term in order to conserve the total energy of the system. We can then rewrite the first law of thermodynamics by replacing dw and du with eqs. (3.14) and (3.21) and obtain

$$dq = c_v dT + p d\alpha. \tag{3.22}$$

Adiabatic processes are those in which no heat or mass exchange with the environment occurs. They are good approximations for ascending and descending air parcels (Section 3.2.6) or air masses in the atmosphere. To obtain the equation describing adiabatic processes in dry air in terms of measurable quantities, we need to combine the first law of thermodynamics with the ideal gas law. Then, the change in the specific volume $d\alpha$ in the first law of thermodynamics can be related to the change in pressure. For that, we take the total differential of the ideal gas law (eq. 3.7) using $\alpha = \rho^{-1}$:

$$p d\alpha + \alpha dp = R_d dT, \tag{3.23}$$

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

and insert the expression obtained for $pd\alpha$ into eq. (3.22):

$$dq = (c_v + R_d)dT - \alpha dp. \tag{3.24}$$

Note that the difference between the specific heat capacity of an ideal gas at constant pressure and its specific heat capacity at constant volume is given by the specific gas constant. For dry air,

$$R_d = c_p - c_v. \tag{3.25}$$

This results from differentiating eq. (3.24) with respect to temperature at constant pressure and making use of the definition of c_p (eq. 3.18). Inserting eq. (3.25) into eq. (3.24) results in:

$$dq = c_p dT - \alpha dp. \tag{3.26}$$

This is the preferred form of the first law of thermodynamics in atmospheric science because the independent variables on the right side (T, p) can easily be measured.

3.2.4 Special Processes

The first law of thermodynamics can be simplified for some special processes, shown schematically in Figure 3.6. For an isobaric process ($dp = 0$, line AE in Figure 3.6), the first law (eq. 3.26) reduces to

$$dq = c_p dT. \tag{3.27}$$

For an isothermal process ($dT = 0$, line AD in Figure 3.6), we obtain

$$dq = -\alpha dp. \tag{3.28}$$

For an isochoric process ($d\alpha = 0$, line AB in Figure 3.6), it reduces to

$$dq = c_v dT = du. \tag{3.29}$$

For an adiabatic process ($dq = 0$, line AC in Figure 3.6), we obtain

$$c_p dT = \alpha dp. \tag{3.30}$$

In an isothermal process, the system can be thought of as connected to a heat reservoir keeping the temperature of the system constant. For example, the work done by the gas upon expansion is taken entirely from the heat reservoir. In such an isothermal process, an increase in volume leads to only a moderate decrease in pressure because the temperature remains constant (line AD). Since for an infinitesimal change dV the work done by the system on the surrounding is defined as $-pdV$, the total work done is given by the gray shaded area in Figure 3.6. Recall that following the convention of the first law, all net transfers of work away from the system (i.e., done by the system) are negative.

In an adiabatic process, the system is isolated from its surroundings so that no transfer of energy in the form of heat can occur. As the gas expands adiabatically from point A to point C,

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

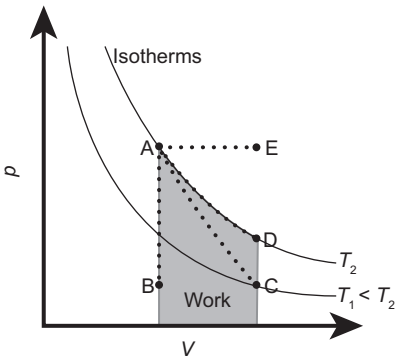


Figure 3.6 Schematic of the various thermodynamic processes in a p-V diagram: isobaric process (AE), isothermal process (AD), isochoric process (AB), adiabatic process (AC). The gray area denotes the work done by the gas in the isothermal expansion from A to D.

it does work at the expense of its internal energy and thus its temperature decreases. According to the ideal gas law, this leads to an additional decrease in pressure for the same volume change (line AC) in comparison with the isothermal case.

3.2.5 The Carnot Process

We have seen earlier that thermodynamic processes can transform heat into mechanical work, and vice versa. For instance, in a heat engine, energy in the form of heat is extracted from a hot reservoir and discarded into a cold reservoir, with some of the energy being converted into work (Figure 3.7a). As another example, during a reversible isothermal expansion of a gas at the higher temperature T_2 from volume V_1 to volume V_2, where $V_2 > V_1$, the first law of thermodynamics implies that an amount $\Delta Q_{1\to 2}$ of the heat from the reservoir is transformed into mechanical work $-\Delta W$ exerted by the gas on the surroundings, as given by:

$$\Delta Q_{1\to 2} = -\Delta W = \int_{V_1}^{V_2} p dV. \tag{3.31}$$

Invoking the ideal gas law in the form $pV = nR^*T$, and noting that we are evaluating this energy conservation for the higher temperature T_2, upon integration we get

$$\Delta Q_{1\to 2} = nR^* T_2 \ln\left(\frac{V_2}{V_1}\right). \tag{3.32}$$

During this isothermal expansion, the entropy increases $\Delta S > 0$. According to the thermodynamic definition of entropy changes, an increase in S is related to the addition of heat dQ to a unit mass of a gas at temperature T (Clausius, 1854):

$$dS = \frac{dQ}{T}. \tag{3.33}$$

For our isothermal expansion, this yields:

$$\Delta Q_{1\to 2} = T_2(S_2 - S_1). \tag{3.34}$$

Note that here we use ΔQ, ΔS, and ΔW instead of dQ, dS, and dW, because we deal with discrete changes of heat, entropy, and work. In a cyclic process, a thermodynamic system (termed an *engine* in the following) performs a closed cycle in the phase space of its state

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

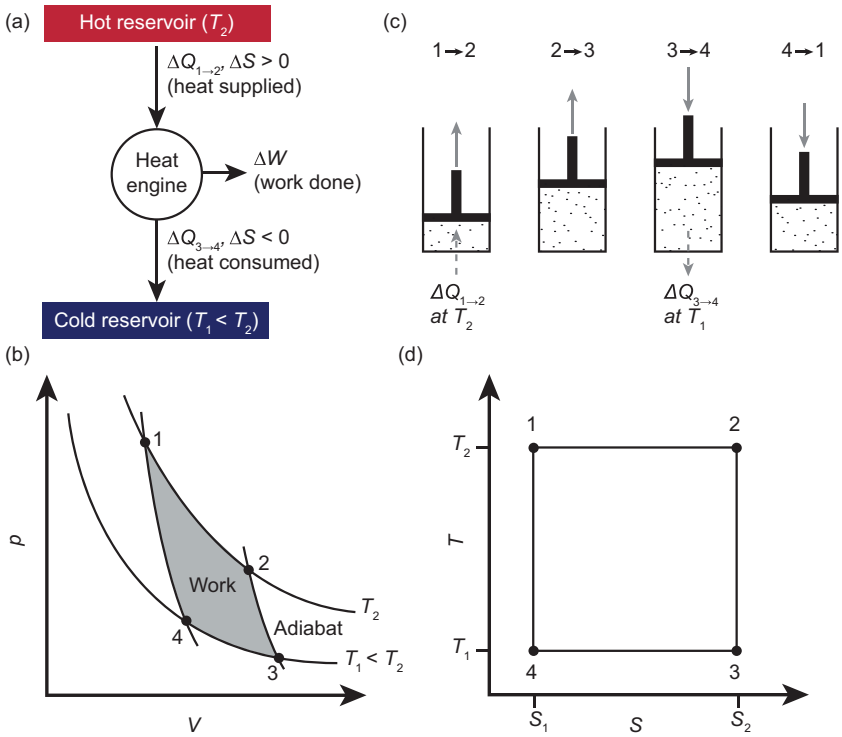


Figure 3.7 (a) Illustration of a heat engine as a Carnot process, where energy in the form of heat is extracted from a hot reservoir. Some heat is converted into work and the remaining heat is discarded into a cold reservoir. (b) Schematic of the different thermodynamic processes in a Carnot cycle for an ideal gas in a p-V diagram: isothermal expansion ($1 \to 2$), adiabatic expansion ($2 \to 3$), isothermal compression ($3 \to 4$), adiabatic compression ($4 \to 1$). (c) A piston model of the Carnot process, showing the associated changes in ΔQ. (d) Same as in (b) but viewed in a $S - T$ diagram.

variables, in $p - V$ or $T - S$ space (Figure 3.7b,d). This means that the engine returns to the state where it started from. We obtain such a periodic cycle by combining isothermal expansion with its reverse, isothermal compression. However, this type of engine can only achieve a net transformation of heat to mechanical work if the isothermal compression is done at a lower temperature T_1.

The *Carnot*[4] *cycle* or Carnot process, is the idealized concept of a heat engine, as it has the maximum theoretically attainable efficiency to convert heat energy into mechanical work. The Carnot process consists of four subprocesses, taking place at two different temperatures, as shown in Figure 3.7:

[4] www.britannica.com/biography/Sadi-Carnot-French-scientist

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

1 → 2 *Isothermal expansion* at the higher temperature T_2: $\Delta Q_{1\to2} = nR^* T_2 \ln(V_2/V_1) = -\Delta W_1$ of heat from the hot reservoir is transformed to mechanical work done by the system (gas). The entropy increases from S_1 to S_2.

2 → 3 *Adiabatic expansion:* At point 2 in Figure 3.7b, the system is decoupled from the hot reservoir and expands adiabatically until the lower temperature T_1 is reached at point 3. Since for an adiabatic process the entropy remains constant, $\Delta S = \frac{\Delta Q}{T} = 0$, the work done by the system during the expansion reduces the internal energy U of the system by $\Delta U = U(T_2) - U(T_1)$.

3 → 4 *Isothermal compression* at the lower temperature, T_1: At point 3, the system is coupled to the cold reservoir. Isothermal compression at the lower temperature T_1 proceeds until point 4. An amount $\Delta W_2 = nR^* T_2 \ln(V_3/V_4) = -\Delta Q_{3\to4}$ of mechanical work done on the system is transformed to heat, which is absorbed by the cold reservoir and the entropy is decreased back to its lower value S_1.

4 → 1 *Adiabatic compression:* At point 4, the system is decoupled from the cold reservoir, and compressed adiabatically until the higher temperature T_2 is reached back at point 1. The work done on the system to achieve this compression increases U by $\Delta U = U(T_2) - U(T_1)$.

The total work done by the system during the Carnot cycle is given by the area enclosed by the cycle. The mechanical work done by the system and during the adiabatic subprocesses cancel as does the entropy increase and decrease in this idealized, reversible Carnot cycle. The total heat absorbed by the system and transformed to mechanical work done by the system is therefore

$$\Delta Q = -\Delta W = \Delta Q_{1\to2} + \Delta Q_{3\to4} = nR^* T_2 \ln\left(\frac{V_2}{V_1}\right) - nR^* T_1 \ln\left(\frac{V_3}{V_4}\right). \tag{3.35}$$

Using the properties of adiabatic processes, no mass and energy exchange with the environment, one can show that $\ln(V_2/V_1) = \ln(V_3/V_4)$:

$$\Delta Q = -\Delta W = nR^*(T_2 - T_1) \ln\left(\frac{V_2}{V_1}\right) = (T_2 - T_1)(S_2 - S_1). \tag{3.36}$$

Unlike the individual subprocesses, the Carnot cycle cannot convert the entire heat absorbed from the hot reservoir to mechanical work: part of it, $\Delta Q_{3\to4}$, is lost to the cold reservoir. The efficiency, also termed the *Carnot efficiency*, η, of the Carnot process is the ratio of the heat ΔQ converted to mechanical work and the added heat $\Delta Q_{1\to2}$:

$$\eta = \frac{\Delta Q}{\Delta Q_{1\to2}} = \frac{T_2 - T_1}{T_2}. \tag{3.37}$$

Whereas it is always possible to fully convert mechanical work to heat (e.g., by friction), eq. (3.37) shows that even the most idealized heat engine works at an efficiency of $\eta < 1$. Real heat engines do not reach the efficiency given by eq. (3.37), as the Carnot process assumes all subprocesses to be fully reversible, which is not possible in the real world. From a different point of view, one can say that the Carnot process transports heat from the hot to the cold reservoir, while some of the heat is lost as mechanical work. Heat never flows from colder to warmer temperatures by itself. However, the inverted Carnot process can transport heat from

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

the colder to the warmer reservoir upon the investment of mechanical work, as is done for example in refrigerators.

Such periodic transformations as in the Carnot process are used in, for example, steam or combustion engines. In atmospheric science, the Carnot process can be used to estimate the destructive power of TCs being the fraction of the heat energy that is converted into kinetic energy, as expressed in terms of their maximum windspeeds (Section 11.5).

3.2.6 Adiabatic Processes

Adiabatic processes are of special relevance for the atmosphere, as the rising of any air parcel is an adiabatic process, at least initially. An *air parcel* is an imaginary volume of air to which the basic dynamic and thermodynamic properties of air are assigned. On the one hand, an air parcel must be large enough to contain a very large number of molecules. On the other hand, it must be small enough that the macroscopic properties assigned to it are approximately uniform within it and that its motion with respect to the surrounding atmosphere does not induce noticeable compensatory movements in the environment. One could visualize such an air parcel as having dimensions of some tens of centimeters in each direction, like a party balloon. An air parcel is assumed to be thermally insulated from its environment, so that its temperature only changes adiabatically as it rises or sinks. It is further assumed that the air parcel is always at the same pressure as the environmental air at the same altitude.

To describe the changes of state of an adiabatically changing air parcel, we need to establish a relation between T and p. We start from eq. (3.30) and replace α using the ideal gas law of dry air eq. (3.7). We then obtain

$$c_p dT = R_d T \frac{dp}{p}. \tag{3.38}$$

Now we rearrange this equation by separating the variables:

$$\frac{dT}{T} = \frac{R_d}{c_p} \frac{dp}{p}. \tag{3.39}$$

Upon integration of eq. (3.39) from a reference point (p_0, T_{p_0}) to (p, T), we arrive at

$$\ln\left(\frac{T}{T_{p_0}}\right) = \frac{R_d}{c_p} \ln\left(\frac{p}{p_0}\right) = \ln\left[\left(\frac{p}{p_0}\right)^{R_d/c_p}\right], \tag{3.40}$$

$$\frac{T}{T_{p_0}} = \left(\frac{p}{p_0}\right)^{\kappa}. \tag{3.41}$$

Here κ denotes the fraction R_d/c_p (eq. 3.8 and 3.19), which is 0.286 for dry air. Usually, the reference pressure p_0 is defined as $p_0 = 1000\,\text{hPa}$ for adiabatic processes and T_{p_0} refers to the temperature that an air parcel would attain if it was brought dry adiabatically to the reference pressure p_0, where "dry adiabatically" refers to an adiabatic process in the absence of phase changes. Hence, eq. (3.41) describes how a dry adiabatically ascending air parcel cools with decreasing pressure and warms when descending dry adiabatically.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

The temperature T_{p_0} is also called the *potential temperature*, θ, and defined as:

$$\theta = T\left(\frac{p_0}{p}\right)^{\kappa}. \tag{3.42}$$

Because the definition of θ arises from the first law of thermodynamics, it follows that during adiabatic transformations, the right hand side of eq. (3.42) remains constant and so does θ. Hence, θ is a conserved quantity for dry adiabatic transformations. θ can be regarded as a measure of the energy content of an air parcel, being a characteristic property of an air parcel, as long as the parcel undergoes only dry adiabatic transformations. In addition, θ is also used to describe the vertical temperature profile in a well-mixed atmospheric boundary layer (Section 5.2.4).

3.2.7 Specific Entropy

In thermodynamics, introducing a different thermodynamic quantity sometimes helps to show the relationships between variables more clearly, to simplify the description of certain processes, or to change the point of view, for instance, from temperature changes to energy changes. To point out an important property of dry adiabatic processes, we introduce the specific entropy s, the entropy per unit mass, and its relation to θ. According to the thermodynamic definition of entropy changes (Clausius, 1854), an increase in s is related to the addition of heat dq to a unit mass of a gas at temperature T, as follows:

$$ds = \frac{dq}{T}. \tag{3.43}$$

In an adiabatic process, where $dq = 0$ and no heat exchange with the environment takes place, the change in specific entropy is also zero ($ds = 0$). Now let us replace dq in the definition of entropy using the first law of thermodynamics (eq. 3.26):

$$ds = \frac{1}{T}(c_p dT - \alpha dp). \tag{3.44}$$

Again, we replace α with the help of the ideal gas law:

$$ds = c_p \frac{dT}{T} - R_d \frac{dp}{p}. \tag{3.45}$$

Now we insert κ into the second term on the right side so that c_p can be placed in front of the parentheses:

$$ds = c_p\left(\frac{dT}{T} - \kappa\frac{dp}{p}\right). \tag{3.46}$$

The term in parentheses turns out to be the same as $d\theta/\theta$. This becomes clear if we start from taking the total derivative of the potential temperature (eq. 3.42),

$$d\theta = \frac{\partial\theta}{\partial T}dT + \frac{\partial\theta}{\partial p}dp. \tag{3.47}$$

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

Then, we arrive at:

$$\begin{aligned} d\theta &= dT\left(\frac{p_0}{p}\right)^{\kappa} - T\kappa\left(\frac{p_0^{\kappa}}{p^{\kappa+1}}\right)dp \\ &= dT\left(\frac{p_0}{p}\right)^{\kappa} - \kappa\left(\frac{p_0}{p}\right)^{\kappa}\frac{T}{p}dp \\ &= \left(\frac{p_0}{p}\right)^{\kappa}\left(dT - \kappa\frac{T}{p}dp\right) \\ &= \frac{\theta}{T}\left(dT - \kappa\frac{T}{p}dp\right) \\ &= \theta\left(\frac{dT}{T} - \kappa\frac{dp}{p}\right). \end{aligned} \tag{3.48}$$

With this, eq. (3.46) reduces to

$$ds = c_p\frac{d\theta}{\theta}. \tag{3.49}$$

Upon integration, we obtain:

$$s = c_p \ln\theta + c, \tag{3.50}$$

where c is a constant. This means that a dry adiabatic process (θ = constant) is also an *isentropic* process, a process in which the entropy remains constant. Note that as soon as phase transitions occur (saturated adiabatic processes, Section 3.5), the simple relation (eq. 3.50) between s and θ does not hold any longer, since latent heat needs to be taken into account in the first law of thermodynamics (eq. 3.26).

3.3 Thermodynamic Charts

In atmospheric sciences, especially in weather analysis and forecasting, thermodynamic charts are used to diagnose weather situations (Alvord and Smith, 1929). They are used to visualize the vertical profiles of temperature, wind, and moisture of the environmental air as obtained from a given radiosonde ascent. Moreover, thermodynamic charts are used to visualize adiabatic processes occurring in the vertical direction, for example, the ascents and descents of air parcels and for determining the bases and tops of clouds. Isothermal or isobaric processes can also be studied, for instance, to forecast the conditions at which fog could form (Section 5). Various types of thermodynamic charts exist. The most common ones are the Stüve diagram, the Skew-T-log-p diagram, and the *tephigram*. Each thermodynamic chart has lines of constant p, T, θ, saturation specific humidity, and saturated adiabats. They only differ in the way in which these lines are plotted. The thermodynamic chart used in this book is the tephigram (Shaw, 1927), as shown in Figure 3.8 and provided in the Appendix (Figure A.1).

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

The name tephigram evolved from the letters T and ϕ, the latter being a symbol of the entropy. Tephigrams are thus also called T-S-diagrams. In the tephigram, the defining coordinates are θ and T, from which all other quantities can be calculated.

A key property of the tephigram is that the lines of constant θ (dry adiabats) or constant entropy (isentropes) are perpendicular to the lines of constant temperature (isotherms) (Iribane and Godson, 1981) and at $\pm 45°$ from the vertical. This ensures that equal areas in a tephigram correspond to equal amounts of energies. As we will see later, this is advantageous for estimating the energy released in a convective cloud (Section 5.2.5).

An example of a tephigram where the various lines are highlighted is shown in Figure 3.8. The lines of constant pressure are nearly parallel to the x-axis (highlighted orange medium-dashed line) and pressure can be calculated according to eq. (3.42).

Lines of constant temperature rise from the lower left to the upper right (highlighted red solid line with positive slope), perpendicular to the lines of constant θ, which rise from the lower right to the upper left (highlighted brown solid line with negative slope). Thus, if an air parcel follows a dry adiabatic ascent, its potential temperature remains constant, that is, it rises from the lower right to the upper left of the diagram. Once the air parcel is saturated with respect to water, it follows a saturated adiabat (highlighted blue, large-dashed curved line). Whether or not an air parcel is saturated can be determined from the saturation specific humidity q_s (highlighted green short-dashed line). Note that the lines of constant q_s are uniquely determined from T and P at every point.

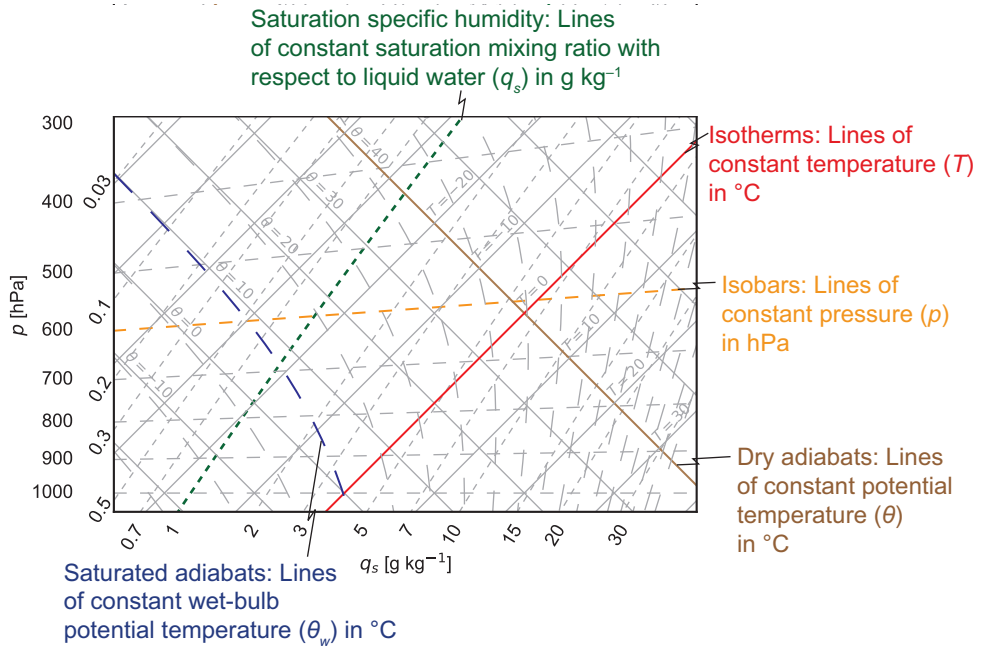


Figure 3.8 Tephigram, in which examples of the various types of lines are highlighted and described.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

3.4 Thermodynamics of Phase Transitions

As water can exist in all three phases in the atmosphere (ice, liquid water, and water vapor), phase transitions, also called *phase changes*, between these states are crucial for the understanding of microphysical processes in clouds. The phase transitions occurring in clouds are condensation (vapor → liquid), evaporation (liquid → vapor), deposition (vapor → solid), sublimation (solid → vapor), freezing (liquid → solid), and melting (solid → liquid).

3.4.1 The Role of the Gibbs Free Energy in Phase Transitions

In this section, the role of the *Gibbs*[5] *free energy* G, also referred to as the Gibbs potential, in phase transitions will be discussed. G is related to the chemical potential μ. Since every system aims to minimize G, μ also wants to be minimized, as we will see later. The Gibbs free energy is used to derive the Clausius–Clapeyron equation, which describes the co-existence of two phases in TDE (Section 3.4.2). Furthermore, G can be used to retrieve information about the direction of phase transitions under given conditions and about kinetic inhibitions of phase transitions. Therefore, G can be seen as an energy barrier that needs to be overcome for nucleation processes, as will be discussed in Chapters 7 and 9.

In thermodynamics, various quantities are used to gather information about the state of a system. Which state variable is most suitable to describe the state of the system depends on the boundary conditions. The right choice of state variables often allows us to formulate a minimum principle for that quantity, which can be derived from the second law of thermodynamics and can be used to describe the TDE of the system under consideration.

For a system at constant T and p, as shown in Figure 3.9, a minimum principle for TDE can be formulated with the help of the Gibbs free energy, which is given by

$$G = U - TS + pV. \tag{3.51}$$

Since all quantities on the right side are state variables, the Gibbs free energy is a state variable itself. In our example system with the movable piston, the ratio of the amounts of liquid and gas (i.e., pressure) and the volume will arrange themselves to minimize G.

Let us now illustrate the role of the Gibbs free energy for a closed system containing a substance in two different phases, namely the gas and the liquid phase, which can exchange matter and energy within the system. Thus, we need to consider the infinitesimal change of G in both phases. Furthermore, the internal energy of each of the two phases, which are not closed with respect to the number of moles n_j, has to be modified compared to eq. (3.2). This is done by adding a term to the right side of eq. (3.2) that takes the change in particle number of the phase j into account:

$$dU_j = TdS_j - pdV_j + \mu_j dn_j. \tag{3.52}$$

Here, μ_j denotes the *chemical potential* of phase j in J mol^{-1} (Gibbs, 1875).

[5] www.britannica.com/biography/J-Willard-Gibbs

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

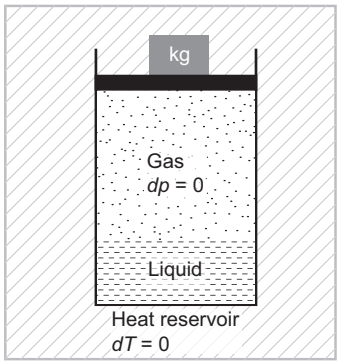


Figure 3.9 Schematic of an isothermal and isobaric system with a movable piston in which a gas and liquid phase of the same substance co-exist.

As a simple approximation, one can think about the chemical potential as a more general version of the concentration gradient. It is the driving force toward physicochemical equilibrium, like what a temperature gradient is to heat transfer, and a pressure gradient is to momentum transfer. If there is a difference in chemical potential for a specific substance, then there will be a chemical or physical change in response (e.g., a chemical reaction, diffusion, or evaporation) until the chemical potential of the substance is the same in each phase (and is the same at all points in each phase).

Starting from eq. (3.52) and considering eq. (3.51), we arrive at:

$$\begin{aligned} dG &= \sum_{j=1}^{2} (dU_j - TdS_j - S_j dT + pdV_j + V_j dp) \\ &= \sum_{j=1}^{2} (-S_j dT + V_j dp + \mu_j dn_j) \\ &= \sum_{j=1}^{2} \mu_j dn_j, \end{aligned} \tag{3.53}$$

since T and p are constant for our system. For any system, this expression can be reformulated to the general definition of μ_j:

$$\mu_j = \left(\frac{\partial G}{\partial n_j}\right)_{T,p,n_i \neq j}. \tag{3.54}$$

Hence, the chemical potential of a given species j is defined as the change in Gibbs free energy of the system with respect to the number of moles of species j under constant T, p, and constant number of moles for all species except species j. Equation (3.54) holds for each chemical component of a homogeneous, multicomponent system, such as a solution. As the combination of the two phases in our example is closed with respect to the number of moles, we have

$$dn_1 = -dn_2, \tag{3.55}$$

or, from eq. (3.53),

$$dG = dn_1(\mu_1 - \mu_2). \tag{3.56}$$

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

Because the (closed) system always tends toward a state with minimum G, it will increase phase 2 at the expense of phase 1 (i.e., $dn_1 < 0$) if $\mu_2 < \mu_1$. Therefore, the minimum principle of G yields information about the direction in which a transition between two co-existing phases occurs under isothermal and isobaric conditions in order to reach TDE, if the difference in chemical potential between the two phases ($\Delta\mu = \mu_2 - \mu_1$) is known. For water vapor and liquid water, $\Delta\mu$ can be written as (Seinfeld and Pandis, 2016):

$$\Delta\mu = -R^* T \ln S_w. \tag{3.57}$$

We note that S_w in eq. (3.57) is not the entropy but rather the *saturation ratio* with respect to liquid water. In our example system with water and water vapor depicted in Figure 3.9, the saturation ratio is given by $S_w = e/e_{s,w}$, where e is given by eq. (3.11) and $e_{s,w}$ is its value at saturation (Section 3.4.4). For a supersatured water vapor phase, $S_w > 1$, so that the logarithm in eq. (3.57) is positive, and hence $\Delta\mu < 0$. That is, the chemical potential of the vapor phase is higher than that of the liquid phase, so condensation will take place for the system to minimize G.

Generally, the chemical potential μ_j can be viewed as the Gibbs free energy per mole in phase j. For example, in reversible chemical reactions, differences in chemical potential can be regarded as related to different compounds (reactants and products) rather than to one compound in different phases. The differences in chemical potential indicate in which direction the reaction tends to take place. Note that the sign of a difference in chemical potential gives no information about whether the reaction releases heat into the surroundings or takes heat from it. For instance, for a reaction which converts a large part of the reactants to products, the chemical potential of the products must be lower than that of the reactants (Figure 3.10). It does not matter whether the reaction is exothermic (in which energy is released) or endothermic (in which energy is consumed).

The Gibbs free energy G has the important property that in an infinitesimal, reversible process, it is not sensitive to the work done by the system nor to the heat exchanged between the system and a connected heat reservoir. Here, the corresponding terms pdV and TdS in the calculation of dG (eq. 3.53) cancel out.

Generally, the thermodynamic potential which measures the consumed or released heat during phase transitions in TDE is the *enthalpy* H:

$$H = U + pV. \tag{3.58}$$

Enthalpy is also a thermodynamic state variable, since it only depends on other state variables, namely U, p, and V. Taking the total differential of eq. (3.58), we arrive at

$$dH = dU + pdV + Vdp. \tag{3.59}$$

Considering isobaric conditions ($dp = 0$) and making use of eq. (3.2), an infinitesimal change is given by:

$$dH = dU + pdV = TdS - pdV + pdV = TdS = dQ. \tag{3.60}$$

Here the non-specific form of eq. (3.43) was used for the last step.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

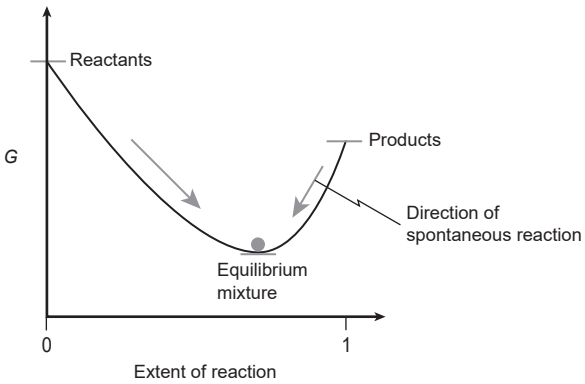


Figure 3.10 Schematic of the Gibbs free energy G for a reversible chemical reaction that reaches an equilibrium state. The gray disk denotes the equilibrium point, corresponding to a minimum in the Gibbs free energy. At this point, the formation rates of products from reactants (forward reaction) and of reactants from products (backward reaction) are equal.

Depending on the conditions (i.e., on T and p), the TDE of our two-phase system containing the gas and liquid phase of a substance, is reached either at a state where only phase 1, only phase 2, or both phases (co-)exist. In the latter case, p and T are related by the Clausius–Clapeyron equation, as will be explained in the following.

3.4.2 Phase Transitions in Thermodynamic Equilibrium: The Clausius–Clapeyron Equation

If two phases co-exist in TDE, their chemical potentials must be equal, that is, $\mu_1 = \mu_2$, and therefore $dG = 0$ (eq. 3.56). Thus, also the specific Gibbs free energies are equal, that is, $g_1 = g_2$, during the transition of an infinitesimal amount of a substance from one phase to the other. The Gibbs free energy is constant during a phase transition as long as the system remains in TDE. Note, however, that heat is required for the phase to change from liquid to vapor because energy is required to break the intermolecular hydrogen bonds. This agrees with experiences in everyday life: stepping out of the water after a bath feels chilly, even on a hot summer day, because the water droplets on the skin withdraw energy from their environment during evaporation.

We will now derive an equation to describe the equilibrium between two phases of a substance in TDE, for example, the equilibrium between liquid water and water vapor in the *phase diagram* of bulk water. Bulk water refers to such a large volume that the number of molecules does not matter and that its thermodynamic properties do not change if the volume is doubled. The equation describing this equilibrium is generally referred to as the *Clausius–Clapeyron*[6]

[6] https://en.wikipedia.org/wiki/Emile_Clapeyron and https://gallica.bnf.fr/ark:/12148/bpt6k3414331n.image#

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

equation (Clausius, 1850). To derive it, we use the fact that the specific Gibbs free energies of the individual phases are equal in TDE:

$$g_1 = g_2. \tag{3.61}$$

The dependence of g on p and T can be obtained by taking the total derivative of g (eq. 3.53):

$$dg = du + pd\alpha + \alpha dp - Tds - sdT. \tag{3.62}$$

Since $du + pd\alpha = dq = Tds$, according to the first law of thermodynamics, eq. (3.62) reduces to

$$dg = \alpha dp - sdT. \tag{3.63}$$

For phase transitions in TDE, the change in the specific Gibbs free energy in the first phase equals that in the second phase, $dg_1 = dg_2$, so that

$$\alpha_1 dp - s_1 dT = \alpha_2 dp - s_2 dT. \tag{3.64}$$

We recall from earlier that the energy required or released during a phase transition in TDE is the enthalpy, H. A closely related quantity is the *latent heat* L, which denotes the integral over the enthalpy H or the heat energy Q consumed or released during a phase change at constant temperature and pressure. For a transition from phase 1 to phase 2, and using specific quantities, it can be expressed as:

$$L_{2,1} = \int_{h_1}^{h_2} dh = T \int_{h_1}^{h_2} \frac{dh}{T} = T \int_{q_1}^{q_2} \frac{dq}{T} = T(s_2 - s_1), \tag{3.65}$$

where h_1, h_2 and q_1, q_2 are the specific enthalpies and heat energies of the initial and final state, respectively. We further made use of the definition of the specific entropy s (eq. 3.43) and that T can be taken into or out of the integral because it is constant during a phase change. Combining eqs. (3.64) and (3.65) yields:

$$\frac{dp}{dT} = \frac{s_2 - s_1}{\alpha_2 - \alpha_1} = \frac{L_{2,1}}{T(\alpha_2 - \alpha_1)}, \tag{3.66}$$

where $L_{2,1}$ denotes the latent heat ($\mathrm{J\,kg^{-1}}$) involved in going from phase 1 to phase 2. Equation (3.66) is the general form of the Clausius–Clapeyron equation (Clausius, 1850). It states the condition of T and p for the co-existence of two phases of a substance in TDE. In particular, it describes the dependence of the saturation vapor pressure of a pure substance on temperature.

The saturation vapor pressure is the vapor pressure at which a liquid or solid phase is in equilibrium with its vapor phase. Here, equilibrium means that, on average, the same number of molecules go from the liquid (or solid) to the vapor phase as go from the vapor to the liquid (or solid) phase per unit time interval.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

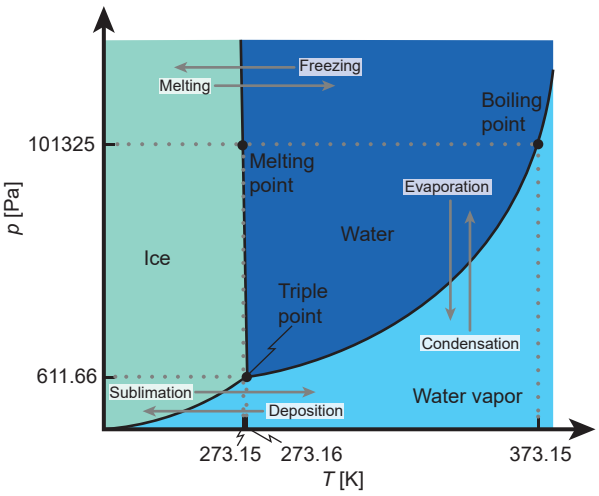


Figure 3.11 Schematic phase diagram of bulk water as a function of temperature and pressure. The temperatures and pressures at the triple point and at the melting and boiling point at atmospheric pressure are indicated by dotted lines.

3.4.3 The Phase Diagram of Bulk Water

For our purposes here, we are mainly concerned with phase transitions of water. The dependence of the saturation water vapor pressure above a flat, pure liquid or solid water surface on temperature is illustrated by the curves in Figure 3.11. This figure is called a *phase diagram* because it shows in which phase water would exist in TDE as a function of pressure and temperature. The specification of TDE is the key phrase here, as water may be present in metastable phases in the real world, like (supercooled) liquid water existing at temperatures below 0 °C. The lines in the phase diagram are phase equilibria or phase transition curves, describing the conditions where two phases can co-exist in TDE. Crossing one of these lines is accompanied by a transformation from one phase to the other.

Co-existence in TDE of all three phases of water, in a single-component system, is only possible for a single combination of pressure and temperature, called the *triple point*. In a single component system (water-only, no air), the temperature at the triple point is $T_{tr} = 273.16\,\mathrm{K} = 0.01\,^{\circ}\mathrm{C}$, and the saturation water vapor pressures over liquid water and ice are equal ($e_{s,w} = e_{s,i} = 611.66\,\mathrm{Pa}$).

The *melting point temperature* T_0, where liquid water and ice can co-exist under atmospheric pressure, is 0.01 K less than the triple point. The melting point temperature is $T_0 = 273.15\,\mathrm{K}$ at 101 325 Pa and varies with pressure. One anomaly of water is the negative slope of the melting curve, indicating that water expands when it freezes, that is, $\alpha_l - \alpha_i < 0$. Note that this negative slope is hardly visible in Figure 3.11 because only a small section of the whole phase diagram of bulk water is shown here. Most other substances have a positive slope, meaning that their specific volume is reduced upon crystallization. The negative slope of the melting curve of water is caused by the hydrogen-bonded structure in ice, which leads to a lower density than the one of liquid water and explains why ice forms on top of lakes and not at the bottom. The negative slope of the melting curve further explains why ice can be melted by increasing the pressure. Against common belief, this is not the explanation of why an ice skater glides on the ice, because the weight of a human being is not sufficient to melt

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

ice. Also, the quasi-liquid layer (QLL) (Section 7.3) that exists on ice at temperatures higher than −25 °C (Li and Somorjai, 2007) is too thin to explain the good gliding of the ice skater. Instead, ice skaters glide because friction creates heat, which causes the ice to melt (Rosenberg, 2005).

The Clausius–Clapeyron equation (3.66) is valid for every phase boundary in the phase diagram of bulk water shown in Figure 3.11. For the equilibrium curve describing the co-existence between the liquid ($\alpha_2 = \alpha_l$) and solid phases ($\alpha_1 = \alpha_i$), we obtain

$$\frac{dp}{dT} = \frac{L_f}{T(\alpha_l - \alpha_i)}, \tag{3.67}$$

where L_f denotes the *latent heat of fusion*, also called the enthalpy of fusion. It is the specific enthalpy of liquid water minus that of ice (eq. 3.65). Thus, its value is also temperature dependent, but this temperature dependency is often neglected. Evaluating eq. (3.67) at 273.15 K, we obtain:

$$\frac{dp}{dT} = \frac{3.33 \times 10^5\,\mathrm{J\,kg^{-1}}}{273.15\,\mathrm{K}(10^{-3}\,\mathrm{m^3\,kg^{-1}} - 1.091 \times 10^{-3}\,\mathrm{m^3\,kg^{-1}})} = -1.34 \times 10^7\,\mathrm{Pa\,K^{-1}}. \tag{3.68}$$

Equation (3.68) shows that large increases in pressure are needed for a small decrease in the freezing temperature.

3.4.4 Saturation Water Vapor Pressure Over Liquid Water

The general form of the Clausius–Clapeyron equation (eq. 3.66) was derived for a one-component system, in which two different phases of the component are in equilibrium. It also holds for partial pressures p_i, in multi-component systems. In moist air, e refers to the partial pressure of the water vapor. Recall that in a one-component system, the total pressure p equals the partial pressure p_i (eq. 3.10), so that in the case of a one-component system consisting only of water, we have $p = e$.

For the phase transition between water vapor ($\alpha_2 = \alpha_v$) and liquid water ($\alpha_1 = \alpha_l$) in a one-component system, we can simplify eq. (3.66) by neglecting the specific volume of liquid water. This simplification is justified because the specific volume of water vapor at 273.16 K and 100 % relative humidity is

$$\alpha_v = \frac{R_v T}{e_{s,w}} = \frac{461.5\,\mathrm{J\,kg^{-1}\,K^{-1}} \times 273.16\,\mathrm{K}}{611.66\,\mathrm{Pa}} = 206.1\,\mathrm{m^3\,kg^{-1}}. \tag{3.69}$$

α_v is five orders of magnitude larger than α_l, which is $10^{-3}\,\mathrm{m^3\,kg^{-1}}$ at the same temperature. Neglecting α_l in eq. (3.66) and using the saturation water vapor pressure with respect to liquid water $e_{s,w}$ yields the typical approximation of the Clausius–Clapeyron equation:

$$\frac{de_{s,w}}{dT} = \frac{L_v}{T\alpha_v} = \frac{L_v e_{s,w}}{R_v T^2}. \tag{3.70}$$

Here, L_v denotes the *latent heat of vaporization* or enthalpy of vaporization. It is the specific enthalpy of water vapor minus the specific enthalpy of liquid water (eq. 3.65). Thus, L_v denotes

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

the heat required to convert a unit mass of liquid water into water vapor, at constant temperature. Here, we have made use of the ideal gas law for water vapor under saturated conditions ($e = e_{s,w}$). In fact, the Clausius–Clapeyron equation (3.70) describes the local slope of the $e_{s,w}$ curve shown in Figure 3.11 in a one-component system that purely consists of liquid water and water vapor (i.e., no air). Thus, it describes the local change of the equilibrium saturation water vapor pressure with respect to liquid water for a change in temperature for $T > T_{tr}$. Note that this slope is not constant. Nevertheless, to a first approximation, the Clausius–Clapeyron equation can be used to infer the saturation water vapor pressures as a function of temperature by integrating eq. (3.70) and assuming that L_v shows no temperature dependence. Taking L_v and $e_{s,w}$ at the triple point ($L_{v,tr}$), this yields:

$$\ln \frac{e_{s,w}}{e_{s,tr}} = \frac{L_{v,tr}}{R_v}\left(\frac{1}{T_{tr}} - \frac{1}{T}\right), \tag{3.71}$$

where $e_{s,tr} = 611.7\,\text{Pa}$ and $L_{v,tr} = 2.501 \times 10^6\,\text{J kg}^{-1}$ at $T_{tr} = 273.16\,\text{K}$ (Table 3.1). The accuracy of this approximation is highest for temperatures near the triple point and deteriorates when the temperature difference increases due to the temperature dependence of L_v.

3.4.5 Saturation Water Vapor Pressure Over Ice

The phase diagram of bulk water, as shown in Figure 3.11, includes only the thermodynamically stable phases (solid, liquid, gaseous). Nevertheless, thermodynamically metastable phases are of high atmospheric relevance. They occur frequently, for example, when water vapor becomes supersaturated in a cooling parcel of moist air or when small (non-bulk) volumes of liquid water, such as cloud droplets, get supercooled prior to freezing in a cloud that has cooled below the melting point of ice. Indeed, the identity $g_1 = g_2$ is still valid for the two phases of supercooled water and water vapor, even though the former is metastable. However, this identity is only true in the absence of an ice phase, which is the thermodynamically stable phase of water at $T < 273.16\,\text{K}$. If ice is absent, the Clausius–Clapeyron equation can be used to describe the equilibrium between (supercooled) liquid water and water vapor even at temperatures below 273.16 K.

At temperatures below 273.16 K, we therefore have two saturation water vapor pressure curves, one describing the equilibrium water vapor pressure over supercooled water, given by eq. (3.70) and one describing the equilibrium water vapor pressure over ice, given by:

$$\frac{de_{s,i}}{dT} = \frac{L_s}{T\alpha_v} = \frac{L_s e_{s,i}}{R_v T^2}. \tag{3.72}$$

Here, L_s is the *latent heat of sublimation*, and $e_{s,i}$ is the saturation water vapor pressure with respect to ice. Again, neglecting its temperature dependency and assuming the latent heat of sublimation to be constant, eq. (3.72) can be integrated to yield an analytical solution for the Clausius–Clapeyron equation describing the equilibrium between ice and the vapor phase:

$$\ln \frac{e_{s,i}}{e_{s,tr}} = \frac{L_{s,tr}}{R_v}\left(\frac{1}{T_{tr}} - \frac{1}{T}\right), \tag{3.73}$$

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

Table 3.1 As a function of temperature T in °C: Saturation water vapor pressures with respect to liquid water $e_{s,w}$ and ice $e_{s,i}$ in Pa (eqs. 3.80 and 3.81), saturation specific humidity q_s in g kg^{-1} at 1000 hPa (eq. 3.88), latent heat of vaporization L_v and sublimation L_s in 10^6 J kg^{-1} (eqs. 3.78 and 3.85), and specific heat capacities of dry air c_p at 1013 hPa (Tsilingiris, 2008) (eq. A.6), of water vapor c_{pv} at 1013 hPa (Tsilingiris, 2008) (eq. A.7), of ice c_i in 10^3 J kg^{-1} K^{-1} (Murphy and Koop, 2005) (eq. A.9), and of liquid water c_l in 10^3 J kg^{-1} K^{-1} (Millero et al., 1973) (eq. A.8).

T	$e_{s,w}$	$e_{s,i}$	q_s	L_v	L_s	c_p	c_{pv}	c_i	c_l
−70	0.48	0.26	0.003		2.833	1.004		1.587	6.583
−65	0.96	0.54	0.006		2.834	1.004		1.622	6.159
−60	1.86	1.08	0.011		2.836	1.004		1.658	5.794
−55	3.49	2.10	0.021		2.837	1.004		1.695	5.483
−50	6.32	3.94	0.039		2.837	1.004		1.731	5.219
−45	11.09	7.21	0.068		2.838	1.004		1.768	4.999
−40	18.91	12.84	0.116		2.838	1.004		1.806	4.816
−35	31.41	22.35	0.193	2.591	2.839	1.004		1.843	4.666
−30	50.94	38.01	0.233	2.576	2.839	1.004		1.881	4.545
−25	80.78	63.28	0.389	2.562	2.838	1.004		1.920	4.448
−20	125.50	103.25	0.634	2.549	2.838	1.004		1.958	4.373
−15	191.31	165.29	1.015	2.537	2.837	1.005		1.997	4.315
−10	286.45	259.89	1.597	2.525	2.837	1.005		2.036	4.272
−5	421.76	401.76	2.470	2.513	2.836	1.005		2.075	4.240
0	611.21	611.15	3.760	2.500	2.834	1.005	1.869	2.114	4.217
0.01	611.66	611.66	3.763	2.501	2.834	1.005	1.869	2.114	4.217
5	872.60		5.374	2.488		1.005	1.868		4.202
10	1228.26		7.575	2.476		1.006	1.868		4.192
15	1705.88		10.539	2.464		1.006	1.870		4.185
20	2339.40		14.487	2.453		1.006	1.872		4.182
25	3169.94		19.692	2.441		1.006	1.875		4.179
30	4246.81		26.489	2.429		1.007	1.879		4.178
35	5628.62		35.293	2.417		1.007	1.884		4.178
40	7384.31		46.614	2.405		1.006	1.890		4.178
100	101325			2.263		1.013	2.037		

with $L_{s,tr} = L_s(T_{tr}) = 2.834 \times 10^6$ J kg^{-1} at $T_{tr} = 273.16$ K (Table 3.1). We will see in Chapter 9 that small volumes of water, such as cloud droplets, can be supercooled down to temperatures ∼235 K before they freeze and form ice.

3.4.6 Temperature Dependencies of Latent Heat and Empirical Parameterizations of Saturation Water Vapor Pressure Over Liquid Water and Ice

As stated earlier, the temperature dependence of latent heat makes an exact analytical solution of eqs. (3.70) and (3.72) impossible.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

For example, L_v depends weakly on temperature, changing by about 6 % from −30 to 30 °C (Table 3.1). We can deduce the temperature dependence of L_v by noting that $L_{2,1}$ corresponds to a difference in the enthalpy $H = U + pV$ (eq. 3.65) of the two phases:

$$L = L_{2,1} = u_2 - u_1 + p(\alpha_2 - \alpha_1). \tag{3.74}$$

For now, we again neglect $\alpha_1 = \alpha_l$ compared to $\alpha_2 = \alpha_v$ and replace $\alpha_v e_{s,w}$ with $R_v T$ according to the ideal gas law for water vapor (eq. 3.11). Upon differentiation with respect to temperature, we obtain

$$\frac{dL_v}{dT} = \frac{du_v}{dT} - \frac{du_l}{dT} + R_v = c_{vv} - c_l + R_v = c_{pv} - c_l. \tag{3.75}$$

Here, we used eqs. (3.17) and (3.25) for water vapor instead of for dry air: c_{vv} refers to the specific heat capacity of water vapor at constant volume, c_l is the specific heat capacity of liquid water, and c_{pv} is the specific heat capacity of water vapor at constant pressure. Note that the specific heat capacities also vary slightly with temperature (Appendix A.3), but less than the latent heat. Therefore, we regard them as constant. The typical values used for the specific heat capacities at the triple point are $c_{vv}(T_{tr}) = 1418.4\,\mathrm{J\,kg^{-1}\,K^{-1}}$, $c_{pv}(T_{tr}) = 1884.4\,\mathrm{J\,kg^{-1}\,K^{-1}}$ (Lemmon, 2015), and $c_l(T_{tr}) = 4217\,\mathrm{J\,kg^{-1}\,K^{-1}}$ (Millero et al., 1973, Table 3.1). If the temperature dependence of the specific heat capacities is neglected, eq. (3.75) can be integrated to yield an approximation for $L_v(T)$:

$$L_v = (c_{pv} - c_l)(T - T_{tr}) + L_{v,tr}. \tag{3.76}$$

The constant of integration can be taken as $L_{v,tr} = L_v(T_{tr}) = 2.501 \times 10^6\,\mathrm{J\,kg^{-1}}$ (Table 3.1).

The expression for L_v can now be introduced into the Clausius–Clapeyron equation (3.70), which can then be integrated more precisely. For the saturation water vapor pressure with respect to water, this yields the following result:

$$\ln \frac{e_{s,w}}{e_{s0}} = \frac{L_{v0} - (c_{pv} - c_l)T_{tr}}{R_v}\left(\frac{1}{T_{tr}} - \frac{1}{T}\right) + \frac{c_{pv} - c_l}{R_v}\ln\frac{T}{T_{tr}}. \tag{3.77}$$

The temperature dependence of L_v can also be parameterized, for instance, following Murphy and Koop (2005):

$$L_v = \frac{56579 - 42.212T + \exp\left(0.1149(281.6 - T)\right)}{M_w}, \tag{3.78}$$

with L_v in $\mathrm{J\,kg^{-1}}$ for T in K and M_w in $\mathrm{kg\,mol^{-1}}$. Equation (3.78) is valid between 236 and 273.16 K, that is, for supercooled water. We apply it for temperatures below 273.15 K in Table 3.1. For temperatures > 273.15 K, we use the parameterization by Hess (1959) in Table 3.1:

$$L_v = 2.5 \times 10^6 - (2.37 \times 10^3)(T - 273.15). \tag{3.79}$$

To solve eq. (3.70) analytically, different empirical formulas for the saturation water vapor pressure as a function of temperature have been developed. Empirical formulas for the saturation vapor pressures over water and ice are provided by Murphy and Koop (2005):

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

$$e_{s,w}(T) = \exp\left\{54.842763 - \frac{6763.22}{T} - 4.210\ln T + 0.000367T + \tanh\left[0.0415(T - 218.8)\right]\left(53.878 - \frac{1331.22}{T} - 9.44523\ln T + 0.014025T\right)\right\}, \tag{3.80}$$

$$e_{s,i}(T) = \exp\left(9.550426 - \frac{5723.265}{T} + 3.53068\ln T - 0.00728332T\right). \tag{3.81}$$

The results for $e_{s,w}$ are valid for temperatures between 123 and 332 K, and those for $e_{s,i}$ between 110 and 273.16 K. They are shown in Figure 3.12 for temperatures below 273.16 K. When water vapor deposits at very low temperatures ($T < 160$ K), it can form amorphous solid water. The parameterization of $e_{s,w}$ from Murphy and Koop (2005) denotes a unified description of the equilibrium water vapor pressure over supercooled liquid water and amorphous solid water. Recent observations, however, called this unified approach into question, and suggested independent parameterization for amorphous solid water, that is, for the saturation water vapor pressure over water at $T < 235$ K (Nachbar et al., 2019).

A simpler approximation and empirical expression for $e_{s,w}$ is given in Rogers and Yau (1989):

$$e_{s,w}(T) = A_w \exp\left(-\frac{B_w}{T}\right), \tag{3.82}$$

where $A_w = 2.53 \times 10^{11}$ Pa and $B_w = 5420$ K. A similar approximation to $e_{s,i}$ is given by (Rogers and Yau, 1989):

$$e_{s,i}(T) = A_i \exp\left(-\frac{B_i}{T}\right), \tag{3.83}$$

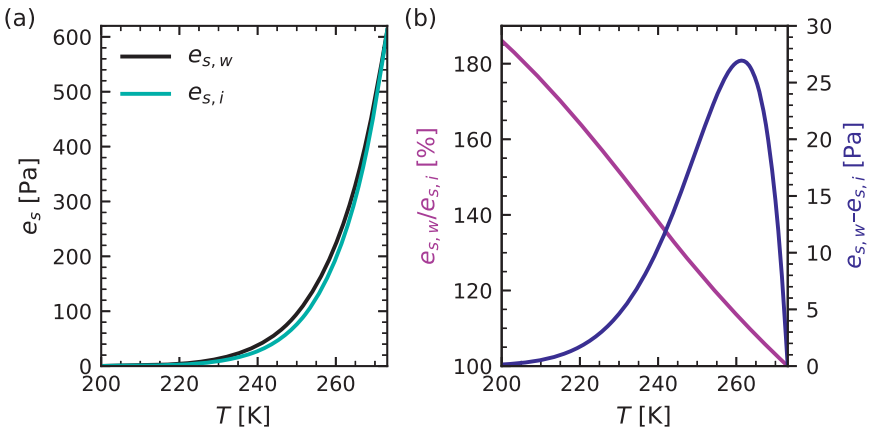


Figure 3.12 (a) Saturation vapor pressure in Pa over water ($e_{s,w}$) and ice ($e_{s,i}$) at temperatures below 273.16 K, according to eqs. (3.80) and (3.81). (b) Difference in saturation vapor pressures in Pa (right axis) and the ratio $e_{s,w}/e_{s,i}$ in % (left axis).

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

where $A_i = 3.64 \times 10^{12}$ Pa and $B_i = 6148$ K. Equations (3.82) and (3.83) can easily be inverted, which is useful for determining the dew point temperature (Section 3.5.5).

The latent heat of sublimation can be obtained analogously to eq. (3.76) by using the specific heat capacity of ice:

$$L_s = (c_{pv} - c_i)(T - T_{tr}) + L_{s,tr}. \tag{3.84}$$

Similarly, an equation for the temperature dependency of the latent heat of sublimation is given by Murphy and Koop (2005):

$$L_s = \frac{46782.5 + 35.8925T - 0.07414T^2 + 541.5 \exp\left\{-\left(\frac{T}{123.75}\right)^2\right\}}{M_w}, \tag{3.85}$$

which is valid between 30 and 273.16 K. For more detailed descriptions of the thermodynamic properties of water and water vapor, we refer the reader to the IAPWS[7] and Wagner and Pruss (2002).

3.5 Moist Air

3.5.1 Water in the Atmosphere

Tropospheric air contains between 0 % and 4 % water vapor by volume. The highest amount of water vapor is found near the surface over the tropical oceans, decreasing with increasing altitude to nearly 0 % at $p \leq 300$ hPa (Sodemann, 2020). The average water vapor concentration declines rapidly with height because (i) its major sources and sinks are near Earth's surface and (ii) because the saturation water vapor pressure decays exponentially with decreasing temperature (eq. 3.71).

Water vapor is also the most important natural greenhouse gas (Masson-Delmotte et al., 2021). Because Earth has an atmosphere, its global average surface temperature is 33.4 K higher than it would be without an atmosphere (Section 2.2.1). Water vapor is responsible for 60 % or 20 K (Kiehl and Trenberth, 1997) of this natural greenhouse effect. It has a large latent heat of condensation and therefore dominates the energy fluxes in the Earth-atmosphere system (Section 2.2).

The presence of water vapor influences, on the one hand, basic properties of air such as its density. On the other hand, phase transitions between the different phases of water alter atmospheric processes, for example, the dry vs. saturated adiabatic ascent of air parcels. In contrast to the dry adiabatic ascent, where the potential temperature θ remains constant, θ increases in a saturated adiabatic ascent once the air is saturated with water vapor and starts to condense. Therefore, the temperature T decreases more slowly in a saturated adiabatic ascent, due to the release of latent heat of condensation, as compared with a dry adiabatic ascent. This can be seen from the difference in the slopes of the dry adiabats and the saturated adiabats in Figure 3.8. At very low temperatures, the saturated adiabats approach the dry adiabats

[7] www.iapws.org

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

because of the small moisture content in cold air. The calculation of the saturated adiabats will be outlined in Section 4.4.

Moist air can be considered a two-component system containing dry air and water vapor. From now on, we use T_0 to distinguish the liquid and the ice phase, being the relevant temperature in our two-component system, the atmosphere. The water vapor content of a moist air mass can be described by different quantities, including the partial water vapor pressure e, water vapor mass mixing ratio w_v, or specific humidity q_v. Depending on the application, expressing the water vapor content of air by means of a temperature, such as the virtual temperature T_v or the dew point temperature T_d can be advantageous. We will discuss some of these quantities in the following.

3.5.2 Water Vapor Mass Mixing Ratio and Specific Humidity

It is often convenient to express the moisture content of the atmosphere in terms of the *water vapor mass mixing ratio* w_v, which is the mass of water vapor m_v per unit mass of *dry* air m_d. w_v can be obtained from

$$w_v = \frac{m_v}{m_d} = \frac{\rho_v V}{\rho_d V} = \frac{\rho_v}{\rho_d} = \frac{e/R_v T}{(p-e)/R_d T} = \frac{\epsilon e}{p-e} \approx \epsilon \frac{e}{p}, \tag{3.86}$$

where

$$\epsilon = \frac{R_d}{R_v} = \frac{M_w}{M_d} = 0.622. \tag{3.87}$$

The approximation in eq. (3.86) is justified because the water vapor pressure is much smaller than the atmospheric pressure: $e \ll p$.

The water vapor mass mixing ratio is similar to the *specific humidity* q_v, which is the mass of water vapor m_v per unit mass of moist air m_m:

$$q_v = \frac{m_v}{m_m} = \frac{\rho_v}{\rho_a} = \frac{\rho_v}{\rho_d + \rho_v} = \frac{e/R_v T}{(p-e)/R_d T + e/R_v T} = \frac{\epsilon e}{p - e + \epsilon e} \approx \epsilon \frac{e}{p}, \tag{3.88}$$

where again we have made use of $e \ll p$. The saturation specific humidity q_s is obtained when $e = e_s$.

3.5.3 Relative Humidity

Knowing the saturation water vapor pressure allows us to define the relative humidity (RH), which is a measure of how much water vapor a quantity of air contains in comparison with its "capacity" for water vapor at that temperature. RH is normally expressed in units of % and is defined with respect to a reference phase (liquid water or ice). For example, relative humidity with respect to liquid water (RH_w) is given by:

$$\mathrm{RH}_w = \frac{e}{e_{s,w}} \times 100 \approx \frac{q_v}{q_s} \times 100. \tag{3.89}$$

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

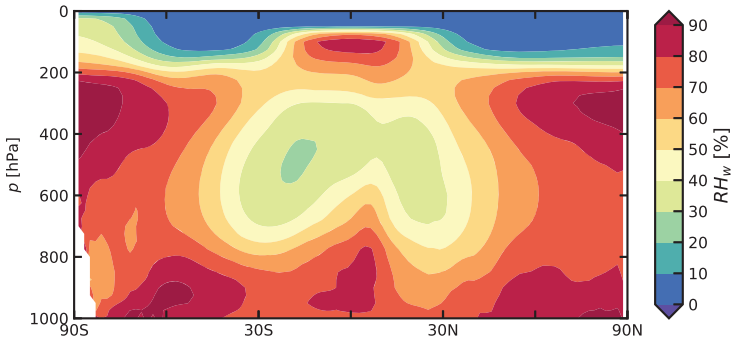


Figure 3.13 Annual and zonal mean latitude versus pressure distribution of the relative humidity with respect to water, based on 10-year climatology (2003–2012) simulated with the ECHAM6-HAM2 global climate model (Neubauer et al., 2019).

Figure 3.13 shows the annual mean distribution of the RH_w as a function of latitude and pressure. In the troposphere, RH_w is highly variable, with typically lower values in the mid troposphere. Close to the Earth's surface, RH_w often exceeds 70 %. The stratosphere (at altitudes above 100 hPa) is very dry, with values $\mathrm{RH}_w < 10\,\%$.

At temperatures below 273.15 K, the relative humidity with respect to ice is defined analogously:

$$\mathrm{RH}_i = \frac{e}{e_{s,i}} \times 100. \tag{3.90}$$

The reference to a phase is important, because for example, air with a partial water vapor pressure $e = 125.5$ Pa at $-20\,^{\circ}\mathrm{C}$ is saturated with respect to supercooled liquid water (i.e., $\mathrm{RH}_w = 100\,\%$) but supersaturated by 21.5 % with respect to ice (i.e., $\mathrm{RH}_i = 121.5\,\%$). RH_i is an important quantity for ice crystal formation and growth (Chapter 9). For air that is saturated with respect to liquid water at temperatures below 273.15 K, $e = e_{s,w}$, the relative humidity with respect to ice is the ratio of $e_{s,w}$ and $e_{s,i}$, which can be obtained by combining eqs. (3.71) and (3.73):

$$\frac{e_{s,w}}{e_{s,i}} = \exp\left[\frac{L_f}{R_v T_0}\left(\frac{T_0}{T} - 1\right)\right]. \tag{3.91}$$

Here, L_f is the latent heat of fusion, which is the difference between L_s and L_v. At 273.16 K, L_f amounts to $L_{f0} = 0.333 \times 10^6\,\mathrm{J\,kg^{-1}}$. For $T < T_{tr}$, it follows that the ratio $e_{s,w}/e_{s,i}$ always exceeds 1 and that $e_{s,w}/e_{s,i}$ steadily increases as T decreases, as seen in Figure 3.12b, where $e_{s,w}/e_{s,i}$ is plotted as a function of temperature. Any atmosphere saturated with respect to liquid water is therefore supersaturated with respect to ice. The maximum difference between $e_{s,w}$ and $e_{s,i}$ occurs at $-12\,^{\circ}\mathrm{C}$ (Figure 3.12b). This has important implications for ice crystal growth and precipitation formation via the ice phase (Mülmenstädt et al., 2015), as will be discussed in Section 9.3 and Chapter 10.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

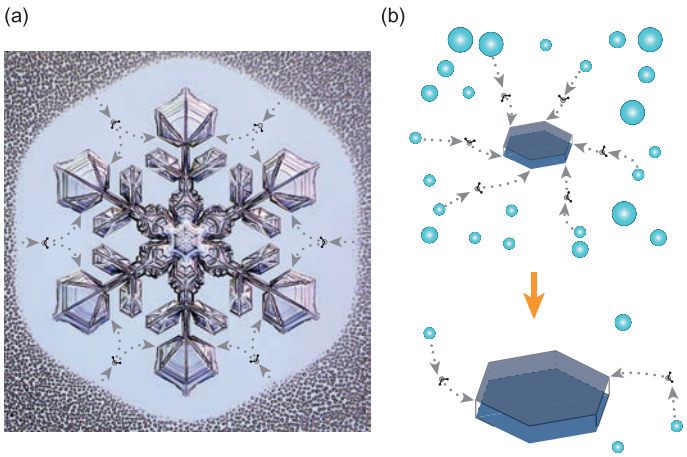


Figure 3.14 Wegener–Bergeron–Findeisen shown by (a) a photograph and (b) a schematic, illustrating how the ice crystals grow at the expense of the evaporating cloud droplets. Photograph in (a) taken from www.snowcrystals.com and used by permission and courtesy of Kenneth Libbrecht.

If a cloud consists of supercooled cloud droplets at temperatures below 0 °C, then the air surrounding the droplets is in equilibrium with the droplets, at $RH_w \approx 100\,\%$. It is thus supersaturated with respect to ice, as shown in Figure 3.12. As a rule of thumb, per 10 K supercooling, the supersaturation with respect to ice increases by 12 % when $RH_w = 100\,\%$. If some ice crystals form in this supercooled liquid water cloud, they experience an environment that is supersaturated with respect to ice. This supersaturation causes a flux of water vapor onto the ice crystals, leading to a rapid growth of these crystals. Over time, the growth of the ice crystals will decrease the vapor pressure to values below saturation with respect to liquid water, causing the cloud droplets to evaporate, that is, the ice crystals grow at the expense of the cloud droplets, as shown in Figure 3.14. This process is referred to as the *Wegener–Bergeron–Findeisen* process (Wegener, 1911; Bergeron, 1935; Findeisen, 1938).

3.5.4 Virtual Temperature

Instead of using moisture quantities to describe the amount of moisture in the atmosphere, there are also various temperature measures that can be used for this purpose.

The *virtual temperature* T_v is a temperature that accounts for the water vapor content of air. The virtual temperature is a useful concept for converting pressure levels to altitude levels, and vice versa, using the hypsometric equation (Section 4.2.2) or when calculating convective available potential energy (CAPE, Section 5.2.5).

While the major atmospheric gas components remain nearly constant, water vapor is highly variable depending on the location (Sodemann, 2020). This variation of q_v would require formulating a "variable gas constant" for moist air. As a variable gas constant is a contradiction in itself and would be rather cumbersome, the virtual temperature T_v is introduced instead.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

This allows us to use the ideal gas law of dry air for moist air by replacing T with T_v, as we will see in the following.

The density of moist air is lower than that of dry air because the molecular weight of water vapor is smaller than that of dry air. As an example, if we have a mixture of 4% (per volume) water vapor and 96% dry air, the molecular weight of the moist air M_m is $M_m = 0.96M_d + 0.04M_w = 28.526\,\mathrm{g\,mol^{-1}}$ as compared to the molecular weight of dry air of $28.96\,\mathrm{g\,mol^{-1}}$. Even though this difference is small, it matters for buoyancy of air (Section 4.2.3). To obtain the density of moist air, we divide the molecular weight of moist air by the volume v that one mole of an ideal gas would occupy, given as $v = 22.71\,\mathrm{mol^{-1}}$ at $T_0 = 273.15\,\mathrm{K}$ and $p_0 = 1000\,\mathrm{hPa}$, independent of the gas composition. Hence, we obtain:

$$\rho_d(q_v = 0) = \frac{M_d}{v} = \frac{28.96\,\mathrm{g\,mol^{-1}}}{22.71\,\mathrm{mol^{-1}}} = 1.27\,\mathrm{kg\,m^{-3}}, \tag{3.92}$$

$$\rho_a = \frac{M_m}{v} = \frac{28.53\,\mathrm{g\,mol^{-1}}}{22.71\,\mathrm{mol^{-1}}} = 1.25\,\mathrm{kg\,m^{-3}}. \tag{3.93}$$

If air does not contain any condensate but does contain water vapor, then it has the density of moist air given by ρ_a, which will be used from here on to denote the atmospheric density.

Applying the ideal gas law for dry air to moist air using the gas constant for dry air R_d requires replacing the temperature by the virtual temperature T_v, yielding:

$$p = R_d \rho_a T_v. \tag{3.94}$$

Here, T_v is the temperature that dry air would have if its density and pressure were equal to a given sample of moist air: if we compare eq. (3.7) with eq. (3.94), we see that the product $\rho_d T$ has to equal $\rho_a T_v$ at the same pressure. Because moist air has a lower density than dry air, as we have shown earlier, $\rho_d T = \rho_a T_v$ implies that the temperature of moist air must be higher. Thus, T_v is higher than the normal temperature and increases with increasing water vapor mass mixing ratio.

To derive T_v, we start with the gas constant of moist air R_a, which can be obtained from the mass-weighted gas constants of dry air R_d and water vapor R_v:

$$R_a = \frac{m_d R_d + m_v R_v}{m_m} = (1 - q_v)R_d + q_v R_v, \tag{3.95}$$

where $m_m = m_d + m_v$ is the total mass, that is, the mass of moist air. Applying the relationship $R_d = \epsilon R_v$ and rearranging terms yields

$$R_a = \left[1 + \left(\frac{1}{\epsilon} - 1\right) q_v\right] R_d = (1 + 0.608 q_v) R_d. \tag{3.96}$$

Instead of using R_a, the term in parenthesis on the right side is now multiplied by the temperature, leading to the definition of the virtual temperature T_v:

$$T_v = T(1 + 0.608 q_v). \tag{3.97}$$

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

3.5.5 Dew Point Temperature

The *dew point temperature* is the relevant temperature for predicting fog formation, especially for radiation fog: The emission of longwave radiation from Earth's surface at night causes an isobaric cooling. As the saturation water vapor pressure $e_{s,w}$ decreases exponentially with decreasing temperature, sufficient cooling of any air parcel with constant specific humidity q_v (or mixing ratio w_v) brings it to condensation, that is, to saturation with respect to liquid water. The temperature at which saturation with respect to liquid water is reached if the air cools isobarically and q_v remains constant is called the dew point temperature, T_d. Thus, T_d is the temperature at which the phase boundary curve between water vapor and liquid water is reached (Figure 3.11), that is, the temperature at which the actual specific humidity equals the saturation specific humidity q_s ($\mathrm{RH}_w = 100\,\%$). Mathematically, T_d can be obtained by inverting the following equation:

$$q_v = q_s(p, T_d). \tag{3.98}$$

This isobaric cooling of moist air can lead to fog formation. In the atmosphere, the presence of so-called CCN (Section 7.6) ensures that the phase change to liquid fog droplets is initiated in close proximity to T_d.

T_d can be measured with a chilled-mirror dew point hygrometer (Figure 3.15), which is one of the oldest and most precise instruments to measure the moisture content of air (Wexler and Brombacher, 1951). Within this device, the moist air is passed over a small mirror that is thermally regulated using a cooling and heating assembly, usually a Peltier device. A light beam is directed onto the mirror, and the intensity of the reflected light beam is measured by a photodetector that also regulates the cooling and heating assembly. When the mirror is cooled down sufficiently, condensation sets in, which is detected optically by a change in reflection from the mirror. Besides direct measurements, T_d can be calculated iteratively or numerically from p and q_v (eq. 3.98). An analytical approximation is obtained by inverting eq. (3.82) to obtain T_d:

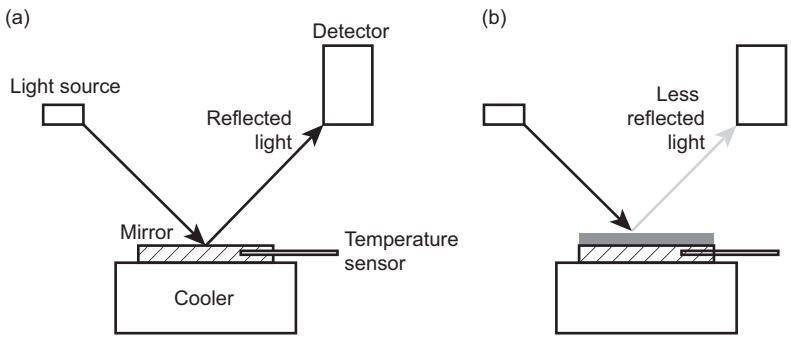


Figure 3.15 Schematic illustration of the working principle of a chilled-mirror dew point hygrometer (a) before and (b) after water vapor condenses on the mirror. The gray layer in (b) indicates dew.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

$$T_d = T_d(q_v, p) = -\frac{B_w}{\ln(e/A_w)} = -\frac{B_w}{\ln(q_v p/A_w \epsilon)}, \tag{3.99}$$

where ϵ is given by eq. (3.87) and A_w and B_w by eq. (3.82).

The analog of the dew point temperature for saturation with respect to ice is called the *frost point temperature* T_f. When moist air cools down below the melting point of bulk water (0 °C at 1013.25 hPa), saturation with respect to ice is reached prior to saturation with respect to liquid water, because $e_{s,i} < e_{s,w}$ (Figure 3.12). Therefore, at the surface of a cooled mirror, deposition of ice starts first, preventing the air from reaching saturation with respect to supercooled water. Thus, the chilled-mirror hygrometer detects T_f at temperatures below 0 °C because it detects bulk properties. T_f lies between the environmental air temperature and T_d due to the lower saturation vapor pressure over ice than over liquid water at sub-zero temperatures (Figure 3.12). Even though a chilled-mirror dew point hygrometer cannot measure T_d below 0 °C, dew point temperatures below 0 °C exist and are relevant for cloud and fog formation (Figure 3.16), as discussed later. The larger the difference between the environmental temperature and the dew point temperature, the drier the air mass. The dew point temperature is one indicator for human discomfort in warm, humid weather. A dew point temperature above 20 °C is often felt as uncomfortable and muggy[8].

If the temperature T, the specific humidity q_v, and the pressure of an air parcel p are known, its dew point temperature can easily be determined graphically on a tephigram. The dotted lines in the tephigram (highlighted green short-dashed line in Figure 3.8) that rise more steeply from the lower left to the upper right than the isotherms are lines of constant saturation specific humidity q_s. They are not parallel to the isotherms because q_s depends not only on the saturation vapor pressure $e_{s,w}$ but also on atmospheric pressure p (eq. 3.88).

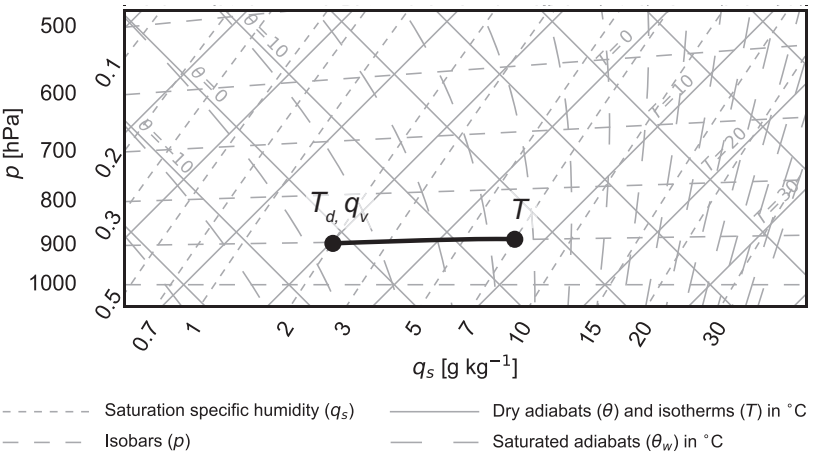


Figure 3.16 Tephigram showing the dew point temperature T_d for an air parcel at 900 hPa (bold solid line) with temperature 7 °C and specific humidity $q_v = 2\,\mathrm{g\,kg^{-1}}$.

[8] https://www.meteoswiss.admin.ch/weather/weather-and-climate-from-a-to-z/dew-point.html

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

Table 3.2 Estimate of stratiform clouds based on the observed dew point depression near Earth's surface ("UK-Met-Office", 2011).

Dew-point depression	Cloud type and amount
0–2 °C	Solid stratus deck
3–5 °C	Broken thick clouds
6–10 °C	Scattered thin clouds
> 10 °C	No stratus clouds

Let us consider an air parcel that starts at $T = 7\,°\mathrm{C}$ with a specific humidity $q_v = 2\,\mathrm{g\,kg^{-1}}$ at 900 hPa (Figure 3.16). To cool the air parcel isobarically until saturation with respect to liquid water is reached, we follow the bold solid line to the left until we reach the temperature, for which $q_v = 2\,\mathrm{g\,kg^{-1}}$ corresponds to the saturation specific humidity q_s. This is then the dew point temperature. In our example, $T_d = -10\,°\mathrm{C}$. The difference between the environmental temperature and T_d is called the *dew-point depression*. It can be used as a rough empirical rule for estimating the amount of layer clouds as indicated in Table 3.2: If the dew-point depression as measured near Earth's surface is less than 2 °C, an overcast sky can be expected. If the dew-point depression exceeds 10 °C, no layer clouds are expected to form on this day. At in-between dew-point depressions, layer clouds can be expected, but they will be broken or scattered.

3.5.6 Lifting Condensation Level and Isentropic Condensation Temperature

Clouds form once RH_w exceeds 100 %. In an adiabatic air parcel ascent, the altitude at which saturation is reached is called the *lifting condensation level (LCL)*. Thus, the LCL marks the cloud base. Let us consider an air parcel at 900 hPa with a specific humidity $q_v = 2\,\mathrm{g\,kg^{-1}}$ and starting temperature of $T = 5\,°\mathrm{C}$, being the environmental temperature, as shown in Figure 3.17. To determine the LCL on a tephigram, we follow the dry adiabat that goes through the starting temperature of our air parcel ($\theta = 13.5\,°\mathrm{C}$) upwards until it intersects with the water vapor line where $q_s = q_v = 2\,\mathrm{g\,kg^{-1}}$ (light blue line). Note that for an air parcel that ascends dry-adiabatically, q_v remains constant and is a conserved quantity and not q_s, which is an intrinsic property of water vapor depending on temperature and pressure. Hence, the LCL is purely determined by the parcel properties at the point where the parcel starts rising. In our example in Figure 3.17, the LCL is reached at 714 hPa.

The temperature at the LCL is called the *isentropic condensation temperature* T_c. According to eq. (3.41), it is given as:

$$T_c = \theta \left(\frac{p_c}{p_0} \right)^{\kappa}, \tag{3.100}$$

where p_c is the pressure at the LCL and p_0 is the reference pressure. In the example in Figure 3.17 with $p_c = 714\,\mathrm{hPa}$, T_c is $-12.8\,°\mathrm{C}$. Above the LCL, water vapor will condense and latent heat will be released. Any further ascent of the air parcel will then follow a saturated adiabat (the red long-dashed line in Figure 3.8).

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

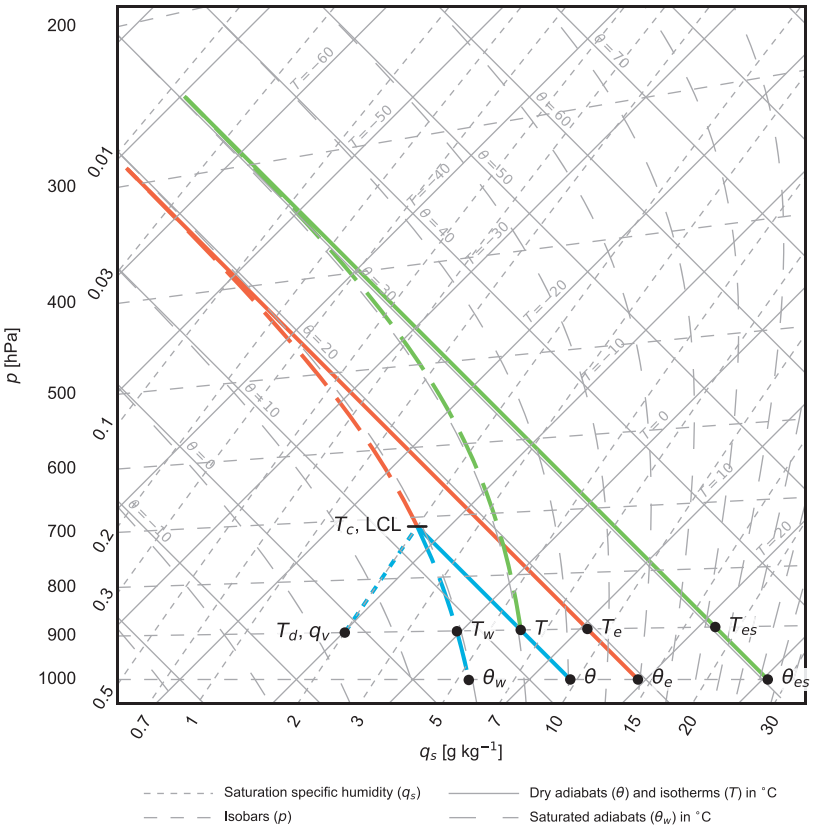


Figure 3.17 A tephigram depicting an air parcel starting to rise at 900 hPa with $T = 5\,°\text{C}$ and $q_v = 2\,\text{g kg}^{-1}$. It depicts the following temperatures: the dew point temperature T_d, the adiabatic wet-bulb temperature T_w, the isentropic condensation temperature T_c (i.e., the temperature at the lifting condensation level (LCL)), the equivalent temperature T_e (Section 3.5.8), the saturated equivalent temperature T_{es} (Section 3.5.9), the potential temperature θ, the wet-bulb potential temperature θ_w, the equivalent potential temperature θ_e (Section 3.5.8), and the saturated equivalent potential temperature θ_{es} (Section 3.5.9). See Table 3.3 for a summary of these temperatures.

3.5.7 Wet-bulb Temperature and Wet-bulb Potential Temperature

Another quantity to describe the relative humidity or the moisture content of the air is the *wet-bulb temperature* T_w. It is used for two things: (i) T_w is used to denote stability criteria in condensing air, as will be discussed in Section 4.5. (ii) T_w can be used to make statements about the combination of temperature and humidity that we humans can tolerate. This is because T_w is related to evaporation, our ability to cool our body by producing sweat that withdraws heat from the body upon evaporation, which is our most efficient method to remove excess heat. Performing physical labor becomes difficult when $T_w > 31\,°\text{C}$ (Buzan and Huber, 2020), because dissipation of heat becomes increasingly difficult. $T_w = 35\,°\text{C}$ has been thought to be the upper physiological limit that humans can withstand for 6 h (Sherwood and Huber, 2010;

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

Raymond et al., 2020), but this depends on the age of humans, the exposure to Sun, and the relative humidity (Vanos et al., 2023). Therefore, T_w is also used to characterize the moist heat stress and how it changes with global warming (Buzan and Huber, 2020).

In atmospheric science, the wet-bulb temperature is defined as the temperature of an air parcel after it is cooled to saturation by evaporation of water within the air parcel, while the latent heat is supplied by the parcel itself. We will further elaborate on this definition later. In general, the wet-bulb temperature can be obtained in two different ways: (i) it can be measured at the environmental pressure with a *psychrometer*, in which case it is referred to as the *isobaric wet-bulb temperature* T_{iw}, or (ii) it can be obtained from a tephigram, in which case it is referred to as the *adiabatic wet-bulb temperature* T_w. The values of these two wet-bulb temperatures are very similar.

T_{iw} is regarded as the "measured" definition of T_w, as T_{iw} is often measured in weather stations with a psychrometer. In fact, the usage of psychrometers to estimate environmental humidity dates back to experiments in the late eighteenth century (Cullen, 1756; Ivory, 1822; August, 1825). A psychrometer consists of two thermometers, where one is surrounded by a wet cloth (wet bulb) that is continuously ventilated and the other is used for a normal temperature measurement, necessary to obtain RH_w from the psychrometer. The ventilation of the wet thermometer ensures that sufficient water evaporates from the wet bulb to establish an equilibrium between the water vapor surrounding the thermometer and the water in the wet cloth, while keeping the system at constant pressure. As evaporation consumes energy, the air in the vicinity of the wet bulb cools until the air is saturated. At this time, the temperature of the wet bulb reaches a steady value, the isobaric wet-bulb temperature T_{iw}. Because of evaporative cooling, $T_{iw} < T$.

In the example discussed earlier, the subsaturated air approaching the wet bulb has specific humidity q_v and evaporation from the wet bulb causes q_v in its vicinity to increase to $q_v' > q_v$. Hence, saturation is reached at a higher T upon isobaric cooling in the vicinity of the wet bulb compared to approaching the dew point T_d, because q_v is a conserved quantity, being constant for the dew point calculation. Thus, $T_{iw} \geq T_d$. T_{iw} can be converted to T_d or RH_w at a given environmental pressure and temperature, by using a so-called *psychrometer chart* (Figure A.2). For the example in Figure 3.17 with the temperatures at 1000 hPa of $\theta = T = 13.5\,^{\circ}\mathrm{C}$ and $\theta_w = T_w = 4.8\,^{\circ}\mathrm{C}$, we obtain a $\mathrm{RH}_w = 20\,\%$. This corresponds very well to the calculated value of RH_w at 1000 hPa given by $q_v/q_s(T = 13.5\,^{\circ}\mathrm{C}) = 2\,\mathrm{g\,kg^{-1}}/9.6\,\mathrm{g\,kg^{-1}} = 20.8\,\%$. It can be shown that the ratio of the latent heat flux away from and the sensible heat flux toward the wet bulb stays the same for changing ventilation speeds, causing T_{iw} to be independent of the ventilation speed.

T_w, as inferred from the tephigram, is the temperature a parcel of air would have if cooled dry adiabatically to saturation at the LCL, and then compressed moist adiabatically to the original pressure level in a wet adiabatic process (blue long-dashed line in Figure 3.17). The LCL is also referred to as Normand's point and the way of graphically determining the wet-bulb temperature is known as *Normand's rule* (Normand, 1921). As the temperature increases due to compression along the saturated adiabatic descent, the specific humidity q_v increases too. This can be understood by recalling that RH_w remains at 100 % during a saturated adiabatic descent, and that according to the Clausius–Clapeyron equation (eq. 3.66),

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

a warmer air parcel can hold more water vapor. For the air parcel to remain saturated at $q_v = q_s(T)$, water must be constantly evaporated into it as the temperature increases due to adiabatic compression.

In Figure 3.17, the wet-bulb temperature T_w at 900 hPa is $-0.4\,°C$ corresponding to a specific humidity $q_v = 4.1\,g\,kg^{-1}$ at this pressure level. Note that q_v at T_w is higher than q_v at the starting temperature T, reflecting the fact that water vapor was added to the air parcel through evaporation. As noted above, T_w is lower than the environmental temperature ($T_w < T$) because water had to be evaporated during the descent, which consumed latent heat.

In addition, a wet-bulb potential temperature θ_w exists, which can be obtained on a tephigram by following the saturated adiabat from T_w to 1000 hPa. In fact, the saturated adiabats in tephigrams are lines of constant θ_w. In the example in Figure 3.17, θ_w is 4.8 °C. It is a useful quantity to analyze when discussing the stability of air layers upon lifting, because it is a conserved quantity during this process, as we will see in Section 4.5.

Comparing the isobaric wet-bulb temperature T_{iw} with the adiabatic wet-bulb temperature T_w, we realize that in the former case, all the water evaporated into the air was initially at the temperature of the wet bulb. By contrast, in the case of T_w, the added water occurred at different pressures and hence different temperatures during the saturated adiabatic descent. Because the latent heat of vaporization increases with decreasing temperature (Table 3.1), the largest latent heat of vaporization is needed for evaporation of condensed water at the lowest temperature (T_c at the LCL) of the process. Thus, more heat is subtracted from the air parcel during the wet-adiabatic process than during the isobaric process (Iribane and Godson, 1981). Therefore, T_w is smaller than T_{iw}, though the difference is usually less than 0.5 °C (Rogers and Yau, 1989). Thus, the relation between T_d, T_w, T_{iw}, and T is given as:

$$T_d \leq T_w \leq T_{iw} \leq T, \tag{3.101}$$

where the equal signs apply only under saturated conditions.

3.5.8 Equivalent Potential Temperature and Equivalent Temperature

Unlike in a dry adiabatic process, θ is no longer conserved if phase transitions occur, and the ascent of an air parcel follows a saturated adiabat θ_w instead of θ (Figure 3.8). Because of the latent heat release, the decrease in temperature with decreasing pressure is less along saturated adiabats than along dry adiabats (Chapter 4).

In the presence of phase transitions, the (pseudo-) adiabatic *equivalent potential temperature* θ_e defines a second conserved quantity in addition to θ_w (Bindon, 1940; Davies-Jones, 2008). θ_e is the temperature that an air parcel attains if all water vapor contained in the parcel condensed and is removed from it before bringing it to 1000 hPa. Due to the latent heat release during condensation, θ_e is higher than θ. Unlike the dry adiabatic process, in which θ can be found by bringing the air parcel directly to 1000 hPa, there is no atmospheric process in which the air parcel reaches 1000 hPa while condensing out all its water vapor. Thus, θ_e is a hypothetical moisture quantity. The value of θ_e can be obtained with the help of a tephigram: (i) first lift the air parcel dry-adiabatically to the LCL, (ii) further lift it along the saturated adiabat until the point where all water vapor is condensed and precipitated out, that is, where θ_w can

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

be approximated by θ, and (iii) bring the air parcel back down to 1000 hPa along the corresponding dry adiabat. In Figure 3.17, we lift the air parcel dry-adiabatically from $T = 5\,°C$ and $q_v = 2\,g\,kg^{-1}$ at 900 hPa to the LCL at 714 hPa. Thereafter, we follow the air parcel along the saturated adiabat (red long-dashed line) up to approximately 300 hPa where the saturated adiabat meets the dry adiabat and then follow it along the dry adiabat (solid line) down to 1000 hPa. The value of θ_e is 19.3 °C. Whereas θ_e remains constant in this process, θ increases from 13.5 °C to 19.3 °C.

To derive θ_e mathematically, we start from eq. (3.43), $ds = dq/T$. In a process involving condensation, eq. (3.26) for dq has to be modified to

$$dq = c_p dT + L_v dq_v - \alpha\, dp, \tag{3.102}$$

where $q_v = q_s$ denotes the saturation specific humidity in case of a saturated adiabatic ascent. Equation (3.102) now accounts for the latent heat release due to condensation. Since in an adiabatic process $dq = 0$, we have

$$-L_v dq_s = c_p dT - \alpha\, dp. \tag{3.103}$$

It was shown in Section 3.2.7 that $(c_p dT - \alpha dp)/T = c_p d\theta/\theta$, so that

$$-\frac{L_v dq_s}{T} = \frac{c_p d\theta}{\theta}. \tag{3.104}$$

Note that some authors prefer to express θ_e in terms of the saturation water vapor mass mixing ratio w_s instead of q_s, because w_s is the mass of water vapor at saturation per unit mass of *dry* air. Thus, the denominator in eq. (3.86) (unit mass of dry air) remains constant during phase changes, whereas the unit mass of *moist* air that is used in the denominator of q_s (eq. 3.88) changes during phase transitions involving the vapor phase. Because q_s and w_s are approximately the same, we stick to q_s to be consistent with the moisture quantity in the tephigram. Following North and Erukhimova (2009), the left side can be approximated using

$$\frac{dq_s}{T} \approx d\left(\frac{q_s}{T}\right), \tag{3.105}$$

so that we can integrate eq. (3.104) to obtain

$$-\frac{L_v q_s}{T} = c_p \ln\theta + c. \tag{3.106}$$

If we identify the constant with $-c_p \ln\theta_e$, take the exponential of this equation, and re-arrange the terms, we obtain

$$\theta_e = \theta \exp\left(\frac{L_v q_s}{c_p T}\right). \tag{3.107}$$

If the air parcel is initially not saturated, then T and q_s must be evaluated at the LCL. Since q_s at LCL is the same as q_v at the starting level of the air parcel, θ_e can be calculated from

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

$$\theta_e = \theta \exp\left(\frac{L_v q_v}{c_p T_c}\right), \tag{3.108}$$

where T_c is the isentropic condensation temperature (eq. 3.100). An example for using θ_e will be discussed in Section 11.5.5 when we discuss tropical cyclones in terms of a Carnot heat engine, because the air ascending in the eyewall of a tropical cyclone is a good example for a pseudoadiabatic ascent (Emanuel, 1986). In the special case that $q_v = 0\,\mathrm{g\,kg^{-1}}$, θ_e reduces to θ.

Analogously, we can define a pseudo-adiabatic *equivalent temperature* T_e:

$$T_e = T \exp\left(\frac{L_v q_v}{c_p T_c}\right). \tag{3.109}$$

In Figure 3.17, the value of T_e at 900 hPa is 10.6 °C. T_e is the temperature an air parcel would have if it were brought dry-adiabatically to the LCL, followed by a saturated adiabatic expansion until all moisture content is removed by precipitation, and then brought back to its original pressure level by dry adiabatic compression. Similarly to the case of the wet bulb temperature, an isobaric equivalent temperature can be defined (Iribane and Godson, 1981), but this temperature is of little practical use and thus not discussed here.

3.5.9 Saturation Equivalent Potential Temperature and Saturation Equivalent Temperature

Another characteristic of moist air that is useful when analyzing the stability of the atmosphere (Chapter 4), is the *saturation equivalent potential temperature* θ_{es}. θ_{es} is a special case of θ_e, which can be derived by assuming that the air parcel is saturated throughout the ascent. Thus, θ_{es} represents a theoretical maximum value of θ_e. Related to θ_{es} is the *saturation equivalent temperature* T_{es}. Both temperatures can be obtained by assuming a saturated air parcel at the starting level ($q_v = q_s$) in eqs. (3.108) and (3.109), leading to

$$\theta_{es} = \theta \exp\left(\frac{L_v q_s}{c_p T}\right), \tag{3.110}$$

$$T_{es} = T \exp\left(\frac{L_v q_s}{c_p T}\right). \tag{3.111}$$

On a tephigram, the temperatures θ_{es} and T_{es} are obtained by following the saturated adiabat, that goes through the starting temperature T at 900 hPa (green long-dashed line in Figure 3.17), upward until all moisture is condensed and precipitated out, and then following the dry adiabat (green solid line) downward to the starting level to 1000 hPa. In Figure 3.17, the saturation specific humidity at the starting temperature is $5.4\,\mathrm{g\,kg^{-1}}$. This yields a T_{es} value of 21.5 °C and a θ_{es} value of 30.5 °C. Note that θ_{es} is only conserved during a fully saturated process, in which the air parcel is already saturated at the starting point, that is, $T_d = T$ and $q_v = q_s$. In such a case, $\theta_e = \theta_{es}$.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

Table 3.3 Summary of different temperatures commonly used in atmospheric science with their formular and typical application. See also Figure 3.17 for a graphical summary.

Temp.	Name	Formula	Application
T (or T_{env})	Temperature or environmental (ambient) temperature		
T_c	Isentropic condensation temperature	Temperature at LCL (Figure 3.17)	Cloud base temperature
T_d	Dew point temperature	$q_v = q_s(p, T_d)$ (eq. 3.98)	Fog formation
T_e	Equivalent temperature	$T_e = T \exp\left(\frac{L_v q_v}{c_p T_c}\right)$ (eq. 3.109)	Pseudoadiabatic ascent
T_{es}	Saturated equivalent temperature	$T_{es} = T \exp\left(\frac{L_v q_s}{c_p T}\right)$ (eq. 3.111)	Equivalent temperature at the starting altitude of an entirely saturated ascent
T_v	Virtual temperature	$T_v = T(1 + 0.608 q_v)$ (eq. 3.97)	Ideal gas law for moist air
T_w	Wet bulb temperature	From tephigram (Fig 3.17)	Human discomfort
T_{iw}	Isobaric wet bulb temperature	From psychrometer chart (Figure A.2)	Human discomfort
θ	Potential temperature	$\theta = T\left(\frac{p_0}{p}\right)^{\kappa}$ (eq. 3.42)	Stability in dry air
θ_e	Equivalent potential temperature	$\theta_e = \theta \exp\left(\frac{L_v q_v}{c_p T_c}\right)$ (eq. 3.108)	Conserved in a pseudo-adiabatic ascent, for example, in the eyewall of a tropical cyclone
θ_{es}	Saturated equivalent potential temperature	$\theta_{es} = \theta \exp\left(\frac{L_v q_s}{c_p T}\right)$ (eq. 3.110)	Conserved in an entirely saturated ascent
θ_v	Virtual potential temperature	$\theta_v = T_v\left(\frac{p_0}{p}\right)^{\kappa}$	Stability in moist air
θ_w	Wet bulb potential temperature	From tephigram (Figure 3.17)	Stability in condensing air, also conserved in a pseudo-adiabatic ascent

The different temperatures with their typical applications are summarized in Table 3.3. The temperatures can be ranked as follows:

$$T_d \leq T_w \leq T_{iw} \leq T \leq T_v \leq T_e \leq T_{es}, \tag{3.112}$$

where the first three and the sixth equal signs apply only under saturated conditions and the fourth and fifth equal signs apply only for completely dry conditions. Similarly, the potential temperatures can be ranked as follows:

$$\theta_w \leq \theta \leq \theta_v \leq \theta_e \leq \theta_{es}. \tag{3.113}$$

Likewise, here the first and east equal signs apply only under saturated conditions and the second and third equal signs apply only for completely dry conditions.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

3.5.10 Wet Adiabatic Processes

For wet adiabatic processes, we can distinguish reversible and irreversible processes. In both processes, the release or consumption of latent heat during condensation or evaporation of condensate needs to be taken into account when calculating the temperature of an air parcel with changing pressure. Thus, θ is no longer conserved in any type of reversible or irreversible wet adiabatic process.

In a reversible wet adiabatic process, the assumption is made that all condensed water remains in the air parcel. The process is reversible because upon reversal of its direction, an air parcel will descend along the same line of constant θ_w, while the formerly released latent heat is consumed again to evaporate the condensed liquid. Note that in the tephigram, the difference between latent heat of vaporization and sublimation is neglected and phase changes are reduced to phase changes between water vapor and liquid water for simplicity. The reversible wet adiabatic process is cumbersome for calculations because one would need to consider the heat content of the condensate (Rogers and Yau, 1989).

Instead, one can define a pseudo-adiabatic process in which the assumption is made that all condensate is liquid water that immediately leaves the parcel by falling out as precipitation. In this case, the adiabatic assumption is violated, because the precipitation leaving the air parcel is an exchange of mass with the environment. Therefore, this process is called a *pseudo-adiabatic process*, which is equivalent to an irreversible wet adiabatic process. Adiabatic is kept in the name because the air parcel's ascent and descent follow saturated and dry adiabats, respectively.

In a pseudo-adiabatic process, both θ_w and θ_e are conserved. To calculate the change in the air parcel's temperature in a pseudo-adiabatic process, we start from eq. (3.102) using $q_v = q_s$. Since the process is adiabatic, $dq = 0$, and we obtain the following expression for the change in the air parcel's temperature in a pseudo-adiabatic process:

$$dT = \frac{1}{\rho_a c_p} dp - \frac{L_v}{c_p} dq_s. \tag{3.114}$$

We will revisit this equation in Chapter 4 when we calculate the saturated adiabatic lapse rate, the rate at which an air parcel cools with increasing altitude in a pseudo-adiabatic ascent.

The cloud water content formed during a (reversible) wet adiabatic ascent is called the *adiabatic liquid water mixing ratio*, w_l, where

$$dw_l = -dq_s. \tag{3.115}$$

To be precise, the increase in w_l is given by the decrease in w_s rather than in q_s, but as mentioned in Section 3.5.8, we neglect this difference to be able to use the same moisture quantity that is used in a tephigram. In a reversible wet adiabatic process, w_l is the maximum cloud water mixing ratio that can be contained in the air parcel in a wet adiabatic ascent because it neglects entrainment of dry air (Chapter 5) and precipitation leaving the air parcel. In a pseudo-adiabatic expansion, w_l thus corresponds to the amount of condensate that is assumed to leave the air parcel as precipitation.

In the atmosphere, irreversible wet adiabatic or pseudo-adiabatic processes are more common than reversible wet adiabatic processes. Pseudo-adiabatic processes can be responsible for

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press

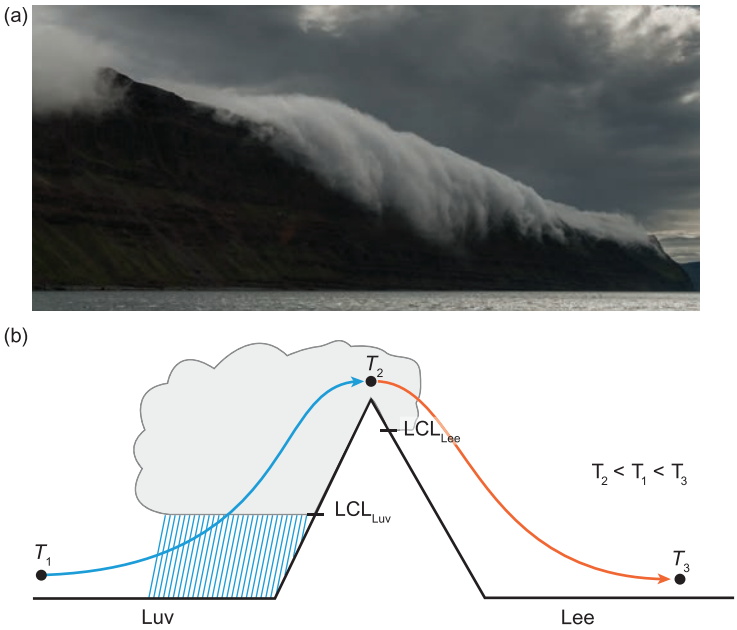


Figure 3.18 (a) Photograph of a Foehn wall cloud by Laurent Barbe. (b) Schematic illustration of the pseudo-adiabatic process leading to a Foehn wall cloud, along with different temperatures along the air flow.

warm downslope winds, such as the Foehn wind in the Alps, the Chinook wind at the eastern slopes of the Rocky Mountains, and the Santa Ana desert wind in California. After the descent on the downslope (lee) side of a mountain, the air is warmer than it was bevor the ascent on the upslope (windward or luv) side, if condensation occurred on the luv side. In this case, the air has been cooled saturated-adiabatically above the LCL at a rate of less than 1 K per 100 m on the upslope side and the cloud water has been lost by precipitation. The air is warmed by $\approx$ 1 K per 100 m dry-adiabatically on the downslope side (Figure 3.18b). Nowadays, however, it is recognized that this thermodynamic *Foehn* theory cannot explain all Foehn events. In addition, dynamical aspects such as the cold air in the luv, which needs to be eroded by the descending air, have to be taken into account (Drobinski et al., 2007, Section 4.3).

A spectacular visible example of the Foehn flow are so-called *Foehn wall clouds*, as shown in Figure 3.18a. Here the clouds are restricted to the luv of the mountain. If the clouds precipitate, then the air that is blown over the mountain crest contains less moisture. Consequently, when this air sinks and is warmed by adiabatic compression, the remaining condensate evaporates (sublimates). If some precipitation fell on the luv side, then the LCL is higher on the lee side (Figure 3.18b). Below the LCL on the lee side, the air is warmed dry adiabatically on its way to the surface. This Foehn process can lead to a distinct boundary between the clouds and the cloud-free air, in which case the clouds form a Foehn wall.

https://doi.org/10.1017/9781009485654.004 Published online by Cambridge University Press