

---

## G TEMPERATURE

### G.1 *Thermal energy*

Thermodynamics is a very abstract physical theory which defines state variables and establishes relationships between them. Most importantly, it provides a definition of temperature for systems in a state of thermal equilibrium. But thermodynamics does not make any assumption about the internal structure of systems and how they would give rise to relations between their state variables: that particular relation is provided by statistical mechanics, i.e. the mechanical theory of systems with many degrees of freedom. One can think of thermodynamics as an effective theory of these systems (typically, the number of degrees of freedom is of the order of Avogadro's number or higher) in thermal equilibrium, where on average each mode carries the same share of the total thermal energy: this is meant by equipartition.

It is only sensible to speak about thermal energy in thermal equilibrium: While microscopically there is a continuous reshuffling of energy between all degrees of freedom of a system, macroscopically there is no discernible dynamics at all. The amount  $Q$  of thermal energy that an object of mass  $m$  stores at temperature  $T$  is given by

$$Q = c(T)mT \tag{G.332}$$

with the specific energy  $c(T)$ , possibly a function of  $T$ . It summarises how much thermal energy can be stored in a system, effectively by counting the degrees of freedom the substance provides if the total mass is  $m$ .

### G.2 *Axioms of thermodynamics*

Thermodynamics is a sensibly defined theory *even if* one does not have the slightest clue about the internal structure of matter. But we can turn this around to our advantage: There are sensible thermodynamical properties and definitions of temperature for almost every physical system, even if we do not know a priori how it works internally. For instance, it is quite reasonable to consider the electrodynamic field in thermal equilibrium or to imagine thermodynamical properties of black holes, i.e. of gravity.

Thermodynamics is defined axiomatically, and it's really a good place to appreciate the abstractness of it:

0. Heat flows from hot to cold until a thermal equilibrium is established, characterised by a common temperature  $T$ .
1. The energy content  $U$  of a system can be changed by  $dU$ , either by performing mechanical work  $\delta W$  on it or by changing its thermal energy by  $\delta Q$ : Effectively, the law of energy conservation encompasses thermal energy, too:  $dU = \delta W + \delta Q$
2. A system can not perform mechanical work out of thermal equilibrium; you need a non-equilibrium to generate work out of a heat flow, and that is done imperfectly.
3. Entropy approaches zero at absolute zero in temperature.

With reference to the first law I'd like to clarify that changing the internal energy content can be done in very different ways:

$$dU = TdS - pdV + \mu dN + \Phi dQ + \mathbf{B} \cdot d\mathbf{M} + \dots \tag{G.333}$$

where one sees typically the combination of an extensive (i.e. proportional to the size of the system or the amount of substance) state variable with an intensive (independent of the size of the system) one: work can be performed by a change  $dV$  in volume against pressure  $p$ , or by changing the charge  $dQ$  against an external electric potential  $\Phi$ , by changing the magnetisation  $d\mathbf{M}$  against an external magnetic field  $\mathbf{B}$ , by changing the number of particles in the system  $dN$  against the chemical potential  $\mu$  and lastly, by changing the entropy  $dS$  against temperature  $T$ .

The separation between intensive and extensive state variables is sometimes washed out, and please be very careful in these cases of "non-extensive thermodynamics": If the potential  $\Phi$  is in fact sourced by  $Q$  itself, it would become extensive.

### G.3 Measuring temperature

Temperature is perhaps the most abstract concept in theoretical physics, despite the fact that everyone has the feeling of intuitiveness about temperature: This is because the fundamental definition of temperature  $T$  is

$$\frac{\partial S}{\partial E} = \frac{1}{k_B T} \quad (\text{G.334})$$

so one needs to have an intuition about entropy  $S$  first and how it depends on  $E$ , before thinking about temperature  $T$ . When measuring temperature, one can go about in very different ways: Firstly, one can use some empirical relationship between an easily observable quantity, for instance the length of a column of mercury as it is determined by the [thermal expansion coefficient](#) or the volume of a gas as an expression of [Gay-Lussac's law](#) to estimate  $T$ . Specifically, for an ideal gas we have a constant  $V/T$  at constant pressure  $p$ , so

$$\frac{V_1}{T_1} = \frac{V_2}{T_2} \quad \rightarrow \quad T_2 = \frac{V_2}{V_1} T_1 \quad (\text{G.335})$$

with a reference  $V_1$ ,  $T_1$  and a measurement of  $V_2$ , where for instance, the reference could be defined through the [molar volume](#) of 22.41396954 litres at [standard conditions](#)  $p = 101.325$  kPa and  $T = 273.15$  Kelvin.

But conceptually, one would like to measure temperature mechanically: By conversion of thermal energy to mechanical energy, the world of thermodynamics is linked to the world of mechanics, and mechanical energies can be measured unambiguously and in accordance with the laws of Galilean or Lorentzian relativity, for instance by accelerating an object of known mass. Surely, this can not be achieved in thermal equilibrium, as clarified by the second law of thermodynamics, but if there is a disequilibrium one can employ a Carnot-engine to convert thermal energy  $\delta Q$  to mechanical energy  $\delta W$

$$\delta W = \eta \delta Q \quad (\text{G.336})$$

at a known and unique efficiency  $\eta$ , which is only a function of temperatures

$$\eta = 1 - \frac{T_2}{T_1} < 1 \quad (\text{G.337})$$

with the temperature  $T_1$  of the hot reservoir and  $T_2$  for the cold reservoir: With a Carnot-engine one can determine at least temperature differences relative to a

reference temperature that needs to be fixed. For this one chooses the [triple point of water](#) at  $T = 273.1600$  Kelvin at  $p = 611.657$  Pa. At the triple point, the phases of ice, water and vapour exist simultaneously which is easy to observe.

#### G.4 Carnot-engines

Carnot-engines are thermal engines: They can convert thermal energy back to mechanical energy (or pump thermal energy from the cold to the hot reservoir against the natural tendency of heat to flow from hot to cold). The conversion from thermal to mechanical energy is not ideal but happens at an ideal efficiency  $\eta = 1 - T_2/T_1$ , which is identical for all Carnot-engines irrespective of how they are built: All that matters is perfect reversibility in their working principle: If the temperatures are changed, the efficiency changes without any delay. And Carnot-engines work in a cyclic fashion: There is no energy stored internally after one sequence is completed.

A traditional construction is the steam-engine-type, where one proceeds in four phases through (i) isothermal expansion at  $T_1$ , sucking in the heat  $Q_1$ , followed by (ii) adiabatic expansion to bring the temperature from  $T_1$  down to  $T_2$  with no heat exchange, (iii) isothermal compression at  $T_2$ , squeezing out the heat  $Q_2$ , and completed by (iv) adiabatic compression to get the temperature from  $T_2$  up to  $T_1$  again. The amounts of work gained in step (ii) is equal to the work to be invested in step (iv), so all that matters are the energies in step (i) and (iii):

$$Q_1 = \int dS T = T_1 \Delta S > 0 \quad \text{and} \quad Q_2 = \int dS T = T_2 \Delta S < 0 \quad (\text{G.338})$$

After one complete cycle of the engine there is no change  $dU$  in internal energy as it works cyclically, so one can conclude

$$dU = \delta W + \delta Q = 0 \quad \rightarrow \quad W = -Q = -(Q_1 + Q_2) = (T_1 - T_2)\Delta S \quad (\text{G.339})$$

so we can find for the Carnot-efficiency  $\eta$  the iconic result

$$\eta = \frac{W}{Q_1} = 1 - \frac{T_2}{T_1} \quad (\text{G.340})$$

But actually this is only one way of constructing a Carnot-engine. A completely different engine would be a propeller on an axis submerged in a gas at temperature  $T_1$  on which there is a ratched and pawl-mechanism that only allows the engine to turn into one direction, lifting a weight in the process and performing mechanical work. In forward motion, one needs an energy  $\epsilon$  to disengage the ratched such that an amount  $W$  of work can be gained, and this happens at a rate  $\exp(-(\epsilon + W)/(k_B T_1))$ . Motion in the opposite direction frees work  $W$ , but nevertheless needs an activation energy  $\epsilon$  of the ratched.  $W$  is lost, or more exactly, transferred to the gas and dissipated there, while a random thermal fluctuation provides  $\epsilon$  to the ratched, at a rate  $\exp(-\epsilon/(k_B T_2))$ . In a reversible engine, the two rates are equal,

$$\exp\left(-\frac{\epsilon + W}{k_B T_1}\right) = \exp\left(-\frac{\epsilon}{k_B T_2}\right) \quad (\text{G.341})$$

from which we can define the efficiency  $\eta$  as

$$\eta = \frac{W}{\epsilon} = 1 - \frac{T_2}{T_1} \quad (\text{G.342})$$

as before! It is perhaps a funny image to combine the steam engine with the ratchet and pawl-machine, one operating as a thermal engine and the other one as a heat pump: They would exactly cancel each other out. And in fulfilment of the second law of thermodynamics: In thermal equilibrium,  $T_1 = T_2$ , the efficiency drops to zero  $\eta = 0$  and no mechanical work can be performed.

### G.5 Thermal wavelength and quantum statistics

We can take this idea of determining temperature with a mechanical measurement one step further, specifically by bringing in quantum mechanics and focusing on kinetic systems, i.e. systems where the thermal energy is present in the form of kinetic energy in the motion of the particles: The de Broglie-wavelength  $\lambda$  of a particle at momentum  $p$  is given by the relation

$$p = \frac{h}{\lambda} \quad (\text{G.343})$$

with Planck's constant  $h$ . On the other hand, the particle's typical energy would be

$$E = k_B T \quad (\text{G.344})$$

as a consequence of equipartition. The two ideas are linked through the dispersion relation

$$E = \sqrt{(cp)^2 + (mc^2)^2} = \begin{cases} cp, & \text{for large momenta } cp \gg mc^2 \\ \frac{p^2}{2m} & \text{for small momenta } cp \ll mc^2 \end{cases} \quad (\text{G.345})$$

Then, we can define the thermal wavelength for relativistic particles,

$$E = k_B T = cp = \frac{ch}{\lambda} \quad \rightarrow \quad \lambda = \frac{ch}{k_B T} \quad (\text{G.346})$$

and analogously for non-relativistic particles,

$$E = k_B T = \frac{p^2}{2m} = \frac{1}{2m} \left( \frac{h}{\lambda} \right)^2 \quad \rightarrow \quad \lambda = \frac{h}{\sqrt{2mk_B T}} \quad (\text{G.347})$$

In both cases, the thermal wavelength becomes shorter with increasing temperature, as a reflection of the particle's higher momenta. Measuring  $\lambda$  spectroscopically by means of a diffraction grid is a perfectly valid determination of temperature: From the fact that the light of the Sun is visible and yellow in colour we can conclude that the surface temperature of the Sun must be around 6000 Kelvin.

Thermal wavelength as a scale matters physically as it is intricately linked to the particle's being indistinguishable: If the separation of two particles is small compared to the thermal wavelength, their wave functions, which have a typical extension of the order  $\lambda$  overlap heavily and a localisation of the particles is not able to determine which particle is which!

Additionally, thermal wavelength  $\lambda$  and the associated volume  $\lambda^3$  of a wave packet provide a scale for the volume of the system, as can be seen for instance in the canonical partition  $Z(T, V, N)$  for an ideal, non-relativistic gas:

$$Z(T, V, N) = \frac{1}{N!} \int \prod_i \frac{d^3 p_i d^3 q_i}{h^3} \exp\left(-\frac{1}{k_B T} \sum_i \frac{p_i^2}{2m}\right) \quad (\text{G.348})$$

which, due to the non-interaction of particles, separates:

$$Z(T, V, N) = Z(T, V, 1)^N \quad (\text{G.349})$$

Evaluating the partition sum shows that  $\int d^3 q$  is just the volume  $V$  of the system and substituting the thermal wavelength  $\lambda$  gives

$$Z(T, V, N) = \frac{1}{N!} \left(\frac{V}{\lambda^3}\right)^N \quad (\text{G.350})$$

Funnily, exactly the same expression is valid for a relativistic ideal gas! Writing

$$Z(T, V, N) = \frac{1}{N!} \int \prod_i \frac{d^3 p_i d^3 q_i}{h^3} \exp\left(-\frac{1}{k_B T} \sum_i c p_i\right) \quad (\text{G.351})$$

by substitution of the definition of  $\lambda$  for the relativistic case. Clearly,  $\lambda^3$  is a scale for  $V$  and we would expect some kind of quantum-mechanical interference effect when the wave functions are extended and become comparable to the size of the system,  $V \approx N\lambda^3$ . And the example of these two classical gases show that  $h$  plays a role even in apparently non-quantum mechanical systems.

### G.6 Boltzmann-factor and the fundamental postulate

To imagine a system in thermal equilibrium is not straightforward: On the macroscopic level, nothing at all is happening as the system does not evolve mechanically, and there are no heat fluxes in or out either. But on the microscopic level, there is a lot going on! All degrees of freedom follow their dynamics defined in Hamiltonian mechanics and are continuously reshuffling energy, but maintaining equipartition on average, with a typical energy of  $k_B T$  present in every degree of freedom. In addition, there is the fundamental postulate to be fulfilled, that finding the system in any of the microstates is equally probable, and that observing a degree of freedom acquiring an amount of energy  $\epsilon$  by a thermal fluctuation is given by the Boltzmann-probability  $\exp(-\epsilon/(k_B T))$ .

There are various ways which suggest the Boltzmann-factor convincingly, but it is clearly one way to enforce transitivity: The probability  $p(\epsilon, T)$  to find a fluctuation of  $\epsilon$  at temperature  $T$  should be a function of energy difference, i.e.

$$\frac{p(\epsilon_2, T)}{p(\epsilon_1, T)} = g(\epsilon_2 - \epsilon_1, T) \quad (\text{G.352})$$

where we can introduce an intermediate step,

$$g(\epsilon_3 - \epsilon_1, T) = \frac{p(\epsilon_3, T)}{p(\epsilon_1, T)} = \frac{p(\epsilon_3, T)}{p(\epsilon_2, T)} \frac{p(\epsilon_2, T)}{p(\epsilon_1, T)} = g(\epsilon_3 - \epsilon_2, T)g(\epsilon_2 - \epsilon_1, T) \quad (\text{G.353})$$

This suggests a functional equation for the unknown function  $g(\epsilon, T)$  which is uniquely solved by  $g(\epsilon, T) = \exp(-\beta(T)\epsilon)$ . The dependence of  $\beta$  on temperature  $T$  is heuristically given by

$$\beta = \frac{1}{k_B T} \quad (\text{G.354})$$

with the Boltzmann-constant  $k_B$ . Heuristically, this is very sensible, as higher temperatures make large thermal fluctuations more likely, and the minus-sign is a reflection of stability if the system is energetically bounded from below and if the temperature is positive. This would be the opposite in systems with negative absolute temperature: Please refer to Sect. G.9 for this.

Let's see whether thermal fluctuations are real and how they would enter in a discrete picture of matter versus a continuum picture. A molecule in the atmosphere experiences continuous collisions with the other molecules maintaining thermal equilibrium, and by interactions with more than one particle the energy  $\epsilon$  fluctuates. The particle can invest this energy to rise up in the gravitational field of the Earth to the height  $h$  determined by the potential energy,  $\epsilon = mgh$ . This process would take place at the Boltzmann-probability

$$p = \exp\left(-\frac{E}{k_B T}\right) = \exp\left(-\frac{mgh}{k_B T}\right) \quad (\text{G.355})$$

such that the fraction of molecules that can reach the height  $h$ , i.e. the density  $\rho$  becomes proportional to  $\exp(-h)$ : This is just the barometric formula.

In a continuum picture the same result has to be explained by this: The atmosphere as a continuum is described by density  $\rho$ , velocity  $v^i$  and pressure  $p$  by the Euler-equation of ideal fluid mechanics,

$$\partial_t v^i + (v^j \partial_j) v^i = -\frac{\partial^i p}{\rho} - \partial^i \Phi \quad (\text{G.356})$$

from which we derive the hydrostatic equation

$$\frac{\partial^i p}{\rho} = -\partial^i \Phi \quad (\text{G.357})$$

if the velocities vanish,  $v^i = 0$  and if there are no accelerations  $\partial_t v^i = 0$ . To continue, we need to assume a relationship between pressure and density, for instance that  $p \propto \rho$  at fixed temperature as predicted by the law by Boyle and Mariotte for ideal gases. Then,

$$\frac{\partial^i p}{\rho} = \partial^i \ln \rho \propto -\partial^i \Phi \quad \rightarrow \quad \rho \propto \exp(-\Phi) \quad (\text{G.358})$$

with the scaling  $\rho \propto \exp(-h)$  for a homogeneous gravitational potential, again leading to the barometric formula. You see that these two pictures have almost nothing in common yet lead to the same result, and that the equation of state  $p \propto \rho$  at fixed  $T$

brushes over a lot of physics but establishes the equivalence between the two pictures. It is clear that on the level of molecules a fluctuation of  $k_B T \approx 10^{-21}$  Joules matters a lot ( $k_B = 1.380649 \times 10^{-23}$  J/K and room temperature is about 300 Kelvin), but that it is completely irrelevant for macroscopic objects.

### G.7 Ultra-relativistic Bose gases and the Planck-spectrum

The Planck-spectrum of a thermal gas of photons was one of the decisive systems which established quantum mechanics along with the hydrogen atom: As abstract as it may seem, the electromagnetic field can be in a state of thermal equilibrium! As Maxwell-electrodynamics is perfectly linear, the field can not reach equilibrium by itself, it can only do so through the interaction with matter. This is important because otherwise the superposition principle would apply to the modes of the field, and it would not be possible to transfer energy from one mode to another. In Planck's original works he conjures up the picture of a container with perfectly mirrored walls to contain the electromagnetic field and a grain of coal as a means of interaction and thermal equilibration. The grain of coal is able to absorb energy from the field and re-emit it in another mode, called photon, and there is really no resistance of the system to change the photon number, expressed by the chemical potential  $\mu = 0$ . The situation would be different if we're dealing with particles with a finite rest mass. Then, one would need to invest at least  $mc^2$  to change the particle number and the chemical potential would be consequently nonzero.

As the particle number is not fixed but controlled by the chemical potential  $\mu$  with the specific value  $\mu = 0$ , corresponding to the fugacity  $z = \exp(\beta) = 1$ , along with a fixed volume  $V$  and a temperature  $T$ , we have to work with a macrocanonical partition  $\mathcal{Z}(T, V, \mu)$ ,

$$\ln \mathcal{Z}(T, V, \mu) = -\frac{4\pi V}{(hc)^3} \int_0^\infty \epsilon^2 d\epsilon \ln [1 - \exp(\beta\epsilon)] = \frac{4\pi V}{(hc)^3} \frac{\beta}{3} \int_0^\infty d\epsilon \frac{\epsilon^3}{\exp(\beta\epsilon) - 1} \quad (\text{G.359})$$

where we use the linear dispersion  $\epsilon = cp$  valid for photons and  $\beta = 1/(k_B T)$ . Integrals like the one in eqn. G.359 involving a monomial  $\epsilon^n$  and the Bose-factor are typical for calculations with bosons: Substituting  $x = \beta\epsilon$  and  $dx = \beta d\epsilon$  gives

$$\ln \mathcal{Z}(T, V, \mu) = \frac{4\pi V}{3(hc)^3} \frac{1}{\beta^3} \int_0^\infty dx \frac{x^3}{\exp(x) - 1} = \frac{4\pi^5 V}{90(hc)^3} (k_B T)^3 \quad (\text{G.360})$$

using

$$\int_0^\infty dx \frac{x^{n-1}}{\exp(x) - 1} = \zeta(n)\Gamma(n) = \zeta(n)(n-1)! \quad (\text{G.361})$$

with Riemann's  $\zeta$ -function and the  $\Gamma$ -function as a generalisation of the factorial. Here, we need the specific value  $\zeta(4)3! = \pi^4/90$ . With these results, one finds for the macrocanonical potential  $J(T, V, \mu)$  the expression

$$J(T, V, \mu) = -k_B T \ln \mathcal{Z}(T, V, \mu) = \frac{8\pi^5 V}{90(hc)^3} (k_B T)^4 \quad (\text{G.362})$$

where we include an additional pre-factor of 2 to take care of the two possible spin states of photons. From this result derives many properties of the Planck-spectrum automatically: Entropy  $S$  and photon number  $N$  are both proportional to  $T^3$ , pressure is proportional to  $T^4$  as well as total energy, as an expression of the Stefan-Boltzmann-law.

One of the decisive properties is the appearance of the Bose factor

$$\frac{1}{\exp(\beta\epsilon) - 1} \rightarrow \exp(-\beta\epsilon) \quad (\text{G.363})$$

which falls back on the familiar Boltzmann-factor for  $\beta\epsilon \gg 1$ . Primarily the consequences are slight numerical differences to a classical computation involving the Boltzmann-factor only, as carried out by Wien originally, who found puzzling pre-factors that he could not make much sense of. If we isolate the spectral energy density  $S(\omega)$  from eqn.G.360 and rewrite it in terms of frequency  $\omega$ ,  $x = \bar{\omega}/(k_B T)$  we get

$$S(\omega) = \frac{\hbar}{4\pi^2 c^2} \frac{\omega^3}{\exp(\beta\hbar\omega) - 1} \quad (\text{G.364})$$

which transitions for high frequencies  $\hbar\omega \gg k_B T$  into the classical Wien-limit as the Bose-factor can be replaced by the Boltzmann-factor

$$S(\omega) = \frac{\hbar}{4\pi^2 c^2} \omega^3 \exp(-\beta\hbar\omega) \quad (\text{G.365})$$

while for small frequencies  $\hbar\omega \ll k_B T$  one recovers the Rayleigh-Jeans limit,

$$S(\omega) = \frac{\omega^2}{4\pi^2 c^2} k_B T \quad (\text{G.366})$$

which of course does not yield a finite result when integrating over all frequencies: That's the ultraviolet catastrophe. The key result in this context is that at high energies, the system behaves classically and at low energies quantum mechanically with an overabundance of photons at low energies. The two regimes are separated roughly by the peak of the Planck-spectrum as the Wien displacement law shows:

$$\frac{dS(\omega)}{d\omega} = 0 \rightarrow \hbar\omega_{\max} \simeq 2.8k_B T \quad (\text{G.367})$$

This overabundance of photons at low energies leads to a super-Poissonian counting statistic experimentally verified by the Hanbury-Brown and Twiss experiment.

### G.8 Entropy

The concept of entropy is mysterious and perhaps as complex to understand as temperature, so let's go through different aspects of entropy:

- Weirdly, the most straightforward view on entropy is a system of ultra-relativistic bosons. As we've shown in Sect. G.7, entropy  $S \propto T^3$  as well as particle number  $N \propto T^3$ , so that both  $S$  and  $N$  increase in proportion with increasing temperature  $T$ , as the system generates new photons. Therefore, entropy is just the number of photons in the system and is perfectly extensive, and clearly one



does not have independent control over  $S$  and  $N$ : As a macrocanonical system, the ultra-relativistic photon gas has  $T$ ,  $V$  and  $\mu$  as independent state variables, with the particular property  $\mu = 0$  as a reflection of the masslessness of photons.

- In a classical ideal gas the particle number is fixed (and the state variable in the canonical ensemble are  $T$ ,  $V$  and  $N$ ): Entropy is rather a reflection of the volume of the energetically allowed phase space.
- The Carnot-engine offers a completely different view on entropy: The Carnot-engine absorbs  $Q_1$  from one reservoir at  $T_1$  and dumps  $Q_2 = T_2/T_1 Q_1$  onto the reservoir at  $T_2$ , because  $Q_i/T_i = S_i$ , which is equal: Entropy controls the conversion from thermal to mechanical energy.  $W/Q_1 = \eta = 1 - T_2/T_1$  is then the efficiency.
- The first law of thermodynamics states that  $dU = TdS - pdV \pm \dots$ : Entropy is an extensive quantity. If one changes it, one performs work against temperature, similarly to changes in volume perform work against pressure. Vice versa,  $\partial S/\partial E = 1/T$  is the formal definition of temperature as a derived quantity.

A nice example of entropy in a discrete system a polymer chain: Please assume that a polymer string is made from  $N$  monomers, which can be built into the chain in the long configuration with length  $a$  and the short configuration with length  $b < a$ . From the outside, one controls temperature  $T$  (through a heat bath), string tension  $\sigma$  (effectively as an analogue to pressure  $p$ ) while  $N$  is fixed; With  $T$ ,  $\sigma$  and  $N$  the suitable state function is the enthalpy  $G(T, \sigma, N)$ , and as  $N$  is fixed, we'll use a canonical description, with a replacement of  $l$  (or  $V$ ) by  $\sigma$  (or  $p$ ): The canonical partition is given by

$$Z_G(T, \sigma, N) = \sum_i \binom{N}{i} \exp\left(-\frac{\sigma l(i)}{k_B T}\right) \quad (G.368)$$

summing over all possible states weighted by the Boltzmann-factor, with the chain length  $l(i)$

$$l(i) = ai + (n - i)b \quad (G.369)$$

The canonical partition can be summed out to yield

$$Z_G(T, \sigma, N) = \left( \exp\left(-\frac{\sigma a}{k_B T}\right) + \exp\left(-\frac{\sigma b}{k_B T}\right) \right)^N \quad (G.370)$$

which factorises, which is typical for non-interacting states,  $Z_G(T, \sigma, N) = Z_G(T, \sigma, 1)^N$ . The enthalpy  $G$  is given by

$$G(T, \sigma, N) = -k_B T \ln Z(T, \sigma, N) = -TS + \sigma l + \mu N \quad (G.371)$$

along with the differential  $dG$

$$dG = -SdT + l d\sigma + \mu dN \quad (G.372)$$

such that we get

$$S = -\frac{\partial G}{\partial T} = Nk_B \ln Z_G(T, \sigma, 1) + \frac{N}{Z(T, \sigma, 1)} \left( a \exp\left(-\frac{\sigma a}{k_B T}\right) + b \exp\left(-\frac{\sigma b}{k_B T}\right) \right) \frac{\sigma}{T} \quad (\text{G.373})$$

as well as

$$l = \frac{\partial G}{\partial \sigma} = N \times \frac{a \exp\left(-\frac{\sigma a}{k_B T}\right) + b \exp\left(-\frac{\sigma b}{k_B T}\right)}{\exp\left(-\frac{\sigma a}{k_B T}\right) + \exp\left(-\frac{\sigma b}{k_B T}\right)} \quad (\text{G.374})$$

The equation of state  $l(T, \sigma, N)$  has a curious property, as the length of the chain decreases with increasing temperature: Higher temperatures enable the system to transition from elements in the long configuration to the energetically disfavoured short configuration by providing thermal fluctuations more easily. This unusual behaviour is an example of an entropic force, as the shortening of the chain comes with an increase in entropy. In addition, the proportionality of  $S$  with the number  $N$  of chain elements underlines the extensivity of  $S$ .

### G.9 Negative absolute temperatures

To make things even weirder, it's perfectly valid to construct systems of negative absolute temperature! Imagine a system that is energetically bounded from above, with fewer and fewer possibilities to realise states of increasing energy. Then, the derivative  $\partial S/\partial E$  that defines temperature, would be negative, and consequently,  $T$  would be smaller than zero as well, according to

$$\frac{\partial S}{\partial E} = \frac{1}{k_B T} \quad (\text{G.375})$$

Of course, this could never be realised in a gas-dynamical system! There, the energy is bounded from below, and  $S$  increases as a function of  $E$ , as there is a larger phase space at higher energies, and  $T$  is necessarily positive.

It gets even weirder when powering a Carnot-engine with two reservoirs, one at positive and one at negative  $T$ : Then, the Carnot-efficiency  $\eta$  becomes larger than one! In some sense, thermal energy is in this case the more useful form of energy compared to mechanical energy and while the first law of thermodynamics formulating energy conservation for mechanical and thermal energy combined is of course valid, it becomes more attractive to store energy in these systems combining a reservoir with  $T > 0$  with one where  $T' < 0$ . By powering the Carnot-engine with mechanical energy one makes  $T$  more positive and  $T'$  more negative, and using the Carnot-engine as a thermal engine one decreases  $T$  and make  $T'$  less negative, and gets mechanical energy back at the efficiency:

$$\eta = 1 - \frac{T'}{T} > 1 \quad (\text{G.376})$$

so in some sense, thermal energy is the more useful energy form compared to mechanical energy.

The efficiency  $\eta$  would be strictly lower than one for thermal engines operating between only positive absolute temperatures or only negative absolute temperatures. Examples of systems with negative absolute temperature are for instance spin-systems, where the state of highest energy (all spins aligned) corresponds to a state of low entropy because of the high degree of order, and in approaching the state of highest energy would find fewer and fewer realisations compatible with the energy,

such that  $\partial S/\partial E$  becomes negative: Imagine a grid of spins, where the state of highest alignment (all spins pointing into the same direction) would be the state of highest energy. Then, having one spin point into the other direction could be realised in  $n$  different ways, having two spins point into the other direction already by  $n(n-1)$  possibilities, so the number of possible realisations increases with decreasing energy, or decreases with increasing energy: That would imply a negative temperature.

### G.10 Reversibility

It seems that statistics of systems with many degrees of freedom brings in something new: While all fundamental laws of Nature are perfectly time reversible (due to the second derivatives in the field equations or the equations of motion), we see that there are irreversible processes like mixing of liquids accompanied by entropy increase.

But this difference is quantitative and not fundamental: Imagine a ball pit filled with balls; 1000 balls of each of 10 different colours, and a group of children mixing the balls continuously. The state where the ball pit is unmixed is only a single one out of  $10^{1000}$  different realisations! (Unmixed means that each ball is situated in one corner of the ball pit assigned to its colour.) If the children playing in the ball pit stir the balls continuously and if there is a new realisation every second, one can expect a spontaneous unmixing in  $10^{1000}$  seconds. As the Universe is "only"  $10^{17}$  seconds old, one would need to wait  $10^{983}$  times the [age of the Universe](#) for this to occur. And a system of  $10^4$  particles is really nothing compared to [Avogadro's number](#)  $N_A = 6.02214076 \times 10^{23}/\text{mol}$ !

### G.11 Information entropies

The fundamental postulate of statistical mechanics, namely that in thermal equilibrium all states of a given energy are equally probable to be assumed by the system and that states of different energy are weighted relative to each other with the Boltzmann-factor  $\exp(-\beta(\epsilon_2 - \epsilon_1))$  opened the way to a fundamental microscopic theory behind thermodynamics. We can ask the question whether the fundamental postulate can be motivated. You might have already guessed that this is the case: Thermal equilibrium could be characterised by making the least assumption about the system in the sense that the random process that distributes the system among its possible states is as random and non-committal as possible.

For a given discrete random distribution  $p_i$  one can define the Shannon-entropy  $S$

$$S = - \sum_i p_i \ln p_i \quad (\text{G.377})$$

as a measure of randomness. Shannon's entropy  $S$  has these properties:

- $S \geq 0$ , because the overall minus-sign takes care of the fact that  $p_i \leq 1$ , and the entropy is bounded from below.
- $S = 0$  if one of the  $p_i = 1$ . Because  $\sum_i p_i = 1$ , the other  $p_i$  need to be zero, and there is no randomness involved.
- For equally probable outcomes,  $p_i = 1/n$  and consequently  $S = \ln n$ , such that the entropy increases with the number of possible outcomes.

- S is additive for statistically independent events,  $p_{ij} = p_i p_j$  implies

$$S = - \sum_{ij} p_{ij} \ln p_{ij} = - \sum_i \sum_j p_i p_j \ln(p_i p_j) = - \sum_j p_j \sum_i p_i \ln p_i - \sum_i p_i \sum_j p_j \ln p_j = S_i + S_j \quad (\text{G.378})$$

- Shannon’s entropy is maximal for equally probable outcomes. Think of S as a functional dependent on the set of probabilities  $p_i$ , so that variation of S would determine  $p_i$ , under the condition that  $\sum_i p_i = 1$  is met, which can be incorporated with a Lagrange multiplier  $\lambda$ :

$$S = - \sum_i p_i \ln p_i - \lambda \left( \sum_i p_i - 1 \right) \quad (\text{G.379})$$

such that the variation  $\delta S$  with respect to  $p_i$  becomes

$$\delta S = - \sum_i \delta p_i \ln p_i + p_i \frac{1}{p_i} \delta p_i - \lambda \sum_i \delta p_i = - \sum_i (\ln p_i + 1 + \lambda) \delta p_i = 0 \quad (\text{G.380})$$

such that  $p_i$  is constant with the value  $\exp(-(1 + \lambda))$ , where  $\lambda$  can be fixed with the boundary condition  $\sum_i p_i = 1$ .

In particular the last point suggests already now a uniform distribution in maximisation of the information entropy S.

The generalisation of Shannon’s entropy to a continuum of outcomes is very interesting but contains a few dangerous spots: For a probability density  $p(x)dx$  one can define

$$S = - \int dx p(x) \ln p(x) \quad (\text{G.381})$$

which shares with the definition eqn. G.377 for the discrete case the value  $S = 0$  for the certain outcome, is additive for independent events and is proportional to the logarithm of the interval length  $b - a$  for a uniform distribution, which incidentally maximises S, too. But it is not bounded from below by zero, which can be seen in the example of the uniform distribution:  $S = \ln(b - a)$  can assume arbitrarily negative values if the interval size  $b - a$  is small enough.

Perhaps even more importantly, S changes under transforms of the random variable:  $p(x)dx = p(y)dy$  as a transformation law is suggested by integration by substitution,

$$\int dx p(x) = \int dy \frac{dx}{dy} p(x(y)) \quad (\text{G.382})$$

so that  $p(x)dx$  is invariant, but  $\ln p(x)$  is not as it becomes  $\ln p(y) + \ln dx/dy$ . This, however, does not play a role in relative information entropies like the Kullback-Leibler-divergence

$$\Delta S = - \int dx p(x) \ln \frac{p(x)}{q(x)} \quad (\text{G.383})$$

which quantify the relative amount of randomness between the distributions  $p(x)$  and  $q(x)$ : The transformation Jacobian drops out in the ratio  $p(x)/q(x)$ .



## ACKNOWLEDGEMENTS

It is a pleasure to thank my students Eileen Sophie Giesel, Robert Reischke, Jonah Cedric Strauss and Hannes Keppler for making me think about physics, and about teaching physics. I thank Martin Feix who supported the lecture so skilfully as a tutor: Let's meet for an espresso at Venice Airport. I thank the students Heinrich Campe, Paul Christians, Ege Görgün, Colin Kühner, Tobias Johannes Metzger, Luca Anakin Prechtel, Amelie Sophia Schulze and Güleser Selin Üstündağ in my lecture for spotting and reporting typos. I thank Matthias Bartelmann for giving this lecture with me in the first iteration.

I would like to acknowledge the lecture  $\LaTeX$ -class [vhbelvadi.com/latex-lecture-notes-class/](http://vhbelvadi.com/latex-lecture-notes-class/) by V.H. Belvadi.