



UNIVERSITY OF TRENTO

DEPARTMENT OF PHYSICS

THESIS

for the degree in

BACHELOR OF SCIENCE

Physics of negative absolute temperatures

SUPERVISOR:

Prof. Raffaello Potestio

CANDIDATE:

Matteo Zortea

ACADEMIC YEAR 2020/2021

Introduction

Everyone has an intuitive idea of what temperature is in everyday life: it is a quantity that tells us how *hot* or *cold* an object is.

Temperature is an intensive property of a system, which means that it does not depend on the object's dimensions nor on the quantity of matter contained in it.

Temperature is measured via thermometers and, since it is a physical quantity, it must be expressed in terms of physical units. All thermometers are expected to give the same number when used to measure the temperature of the same object, provided that the number is expressed always in the same physical unit. This can be possible if the thermometers are accurately calibrated according to some criteria that make the calibration process reproducible. A roughly good calibration of a thermometer in the so called Celsius scale can be done by assigning the instrument a value of 0°C (0 degrees Celsius) at the freezing point of water, and 100°C at the boiling point of water. A more sophisticated and precise way to calibrate a thermometer on the Celsius scale consists in assigning the value of -273.15°C to the *coldest* temperature admitted according to nowadays' physics, and 0.01°C to the triple point of water, that is the point in which water coexists in liquid, solid and gaseous form. One can then extract a measure of all the other values of temperature by assuming a linear relation between temperature and the property in the material used as a thermometer, such as an elastic deformation.

Another interesting temperature scale is the one measured in Kelvin (K), called the *absolute temperature*. This scale of temperature is defined in a way such that it gives a values of 0 K at the coldest possible situation admitted by physics, the so called *absolute zero* point: this means that no physical systems can be cooled more than a system whose absolute temperature is 0 K.

The purpose of this thesis is to introduce the concept of *negative absolute temperatures*. This does not constitute a contradiction to what told before because negative absolute temperatures should not be searched below the absolute zero, but rather above infinity: negative absolute temperatures are hotter than all the positive ones. The world of negative absolute temperatures is often accompanied by strange phenomena: for example, when a system at negative temperature is put into contact with one at positive temperature, the former gives heat to the latter, leading the temperature of the former to be lowered and the temperature of the latter to be raised.

In the first chapter of this thesis I will introduce the concept of temperature in a

rigorous way according to physical laws, justifying the existence of negative absolute temperatures basing on thermodynamical arguments.

A simple system that admits negative temperatures, namely the two level system, will then be presented in chapter 2.

In chapter 3 I will show that negative absolute temperatures were experimentally observed by Purcell and Pound in 1950. After that discovery, some criticisms were moved against the definition of the entropy used to define the absolute scale, putting in doubt the existence of negative temperatures: this will be discussed in chapter 4.

Finally some simulations of a systems at negative temperature will be presented in the last chapter.

Contents

Introduction	iii
1 A systematic definition of temperature	1
1.1 Temperature as an equilibrium property	1
1.2 Thermodynamic temperature	3
1.3 Insight on the thermodynamic temperature	5
2 Two levels system	7
2.1 Two-levels systems admit negative temperatures	7
2.2 Ramsey’s criteria	10
3 Purcell and Pound experiment	13
3.1 Experimental evidence for negative temperature	13
3.2 The experiment carried by Purcell and Pound	13
4 Issues on the definition of entropy	17
4.1 Boltzmann’s framework	17
4.2 Gibbs’ framework	18
4.3 Consistent thermodynamics forbids negative temperatures	19
5 Simulations at negative temperatures	25
5.1 The Ising model	25
5.1.1 Ferromagnetic systems	26
5.1.2 Antiferromagnetic systems	27
5.2 Markov Chain Monte Carlo	28
5.3 Single system at negative temperature	30
5.4 Two systems into contact	33
6 Conclusions	37

Chapter 1

A systematic definition of temperature

The purpose of this section is to provide a rigorous definition of temperature. The definition subsequently given is general and holds for every thermodynamical system.

1.1 Temperature as an equilibrium property

Temperature can be more systematically defined starting from the concept of *equilibrium*. We consider a thermodynamic system to be in equilibrium when no changes are observed in any macroscopic observable and when the system does not exchange heat nor work with the environment. In particular we say that the system is at *thermal equilibrium* when no heat is exchanged by the system, and at *mechanical equilibrium* when no work is done by or on the system.

Let us consider three systems A, B, C whose equilibrium properties are described by the variables $\{X_1, X_2, \dots, X_p\}$, $\{Y_1, Y_2, \dots, Y_p\}$ and $\{Z_1, Z_2, \dots, Z_p\}$ respectively.

If A and B are in equilibrium for some precise values of the coordinates $\{X_1, X_2, \dots, X_p\}$ and $\{Y_1, Y_2, \dots, Y_p\}$, then there must be a relation that connects the two set of values uniquely. Let us express this relation as

$$f_{AB}(X_1, X_2, \dots, X_p, Y_1, Y_2, \dots, Y_p) = 0 \quad (1.1)$$

In an analogous manner, if B and C are in equilibrium for some values $\{Y'_1, Y'_2, \dots, Y'_p\}$ and $\{Z_1, Z_2, \dots, Z_p\}$ of the coordinates, then there must be a constraint on the values $\{Y'_1, Y'_2, \dots, Y'_p\}$ and $\{Z_1, Z_2, \dots, Z_p\}$, which we express as

$$f_{BC}(Y'_1, Y'_2, \dots, Y'_p, Z_1, Z_2, \dots, Z_p) = 0 \quad (1.2)$$

The equations above can be inverted to express one thermodynamic coordinate as a function of the others, or in other words, the above expressions may be written as

$$\begin{aligned} Y_1 &= g_{AB}(X_1, X_2, \dots, X_p, Y_2, \dots) \\ Y'_1 &= g_{BC}(Y'_2, \dots, Y'_p, Z_1, Z_2, \dots, Z_p) \end{aligned}$$

Now let us bring the system B in the same state in both cases, which means imposing $Y_1 = Y'_1, Y_2 = Y'_2, \dots, Y_p = Y'_p$. The last two equations imply that

$$g_{AB}(X_1, X_2, \dots, X_p, Y_2, \dots, Y_p) = g_{BC}(Y_2, \dots, Y_p, Z_1, Z_2, \dots, Z_p) \quad (1.3)$$

or

$$G_{ABC}(X_1, X_2, \dots, X_p, Y_2, \dots, Y_p, Z_1, Z_2, \dots, Z_p) = 0$$

One can use this relation to express X_1 as

$$X_1 = h_{ABC}(X_2, \dots, X_p, Y_2, \dots, Y_p, Z_1, Z_2) \quad (1.4)$$

According to the zeroth principle of thermodynamics if A and B are in equilibrium and B and C are in equilibrium then A and C are in equilibrium¹. This means that the constraints 1.1, 1.2, lead to a new constraint between A and C there must be a constrain on the values $X_1, X_2, \dots, X_p, Z_1, Z_2, \dots, Z_p$ which can be expressed as

$$f_{AC}(X_1, X_2, \dots, X_p, Z_1, Z_2, \dots, Z_p) = 0$$

which means that X_1 can be expressed as

$$X_1 = g_{AC}(X_2, \dots, X_p, Z_1, Z_2, \dots, Z_p) \quad (1.5)$$

and imposing the equality between 1.4 and 1.5

$$g_{AC}(X_2, \dots, X_p, Z_1, Z_2, \dots, Z_p) = h_{ABC}(X_2, \dots, X_p, Y_2, \dots, Y_p, Z_1, Z_2, \dots, Z_p)$$

The term on the left does not depend on the coordinates of B . This means that both g functions in equation 1.3 must be of the type

$$\begin{aligned} g_{AB}(X_1, X_2, \dots, X_p, Y_2, \dots, Y_p) &= \Theta(X_1, X_2, \dots, X_p) + \phi(Y_2, \dots, Y_p) \\ g_{BC}(Y_2, \dots, Y_p, Z_1, Z_2, \dots, Z_p) &= \Theta(Z_1, Z_2, \dots, Z_p) + \phi(Y_2, \dots, Y_p) \end{aligned}$$

so that the dependence on $\{Y\}$ gets cancelled out when equating the two functions leading to

$$\Theta_A(X_1, X_2, \dots, X_p) = \Theta_C(Z_1, Z_2, \dots, Z_p)$$

We started this reasoning by assuming equilibrium between $A - B$ and $B - C$, but one could repeat this reasoning by assuming for example equilibrium between $A - C$ and $B - C$ obtaining an analogous result in terms of $\{Y_1, Y_2, \dots, Y_p\}$. Because of the properties of the equivalence relation, one can extend the reasoning to an arbitrary number of systems. This means that if N systems are in equilibrium, then there must be a function Θ such that

$$\Theta_A(X_1, X_2, \dots, X_p) = \Theta_B(Y_1, Y_2, \dots, Y_p) = \Theta_C(Z_1, Z_2, \dots, Z_p) = \dots$$

¹or, alternatively, equilibrium is an equivalence relation.

Let us call this function *empirical temperature*, and its value on a set of coordinates identifies a particular equivalence class of systems at equilibrium.

What just proven shows that systems at equilibrium are identified by the same value of a certain function Θ . By the way no specifications are given about the origin of this function and which precise value it has for a given set of systems at equilibrium. In fact there are multiple ways to define the values of such function, leading to many *temperature scales*.

An example of a possible way to define a scale of temperature is the one that concerns ideal gases. Practically it consists in assigning a value $\Theta = 273.16$ degrees Kelvin (K) at the triple point of water (coexistence of ice-water-gas) and then other values of temperature for ideal gases are defined via the relation

$$T(K) = \lim_{P \rightarrow 0} 273.16 \times \frac{(PV)_{system}}{(PV)_{ice-water-gas}}$$

because for an ideal gas $T \propto PV$.

Another possible definition of the function Θ , the one relevant for what follows, will be presented later in the next section.

1.2 Thermodynamic temperature

It is known from the first principle of thermodynamics that

$$dE = TdS - pdV$$

Once introducing an entropy as a function of the energy $S = S(E)$ it is possible to define a so called *thermodynamic temperature* via the above equation, yielding the relation

$$\frac{1}{T} = \frac{\partial S}{\partial E}$$

To see why this is coherent with what told before let us look at this example.

First let us consider a system isolated from the environment, so that it cannot exchange heat or work (energy fixed). Let us indicate a generic state of the system by the microscopic coordinates $\mathbf{x} = (q_1, \dots, q_n, p_1, \dots, p_n)$ where (q_i, p_i) is a pair of canonical coordinates. If $\mathcal{H}(\mathbf{x})$ denotes the hamiltonian of the system, the condition

$$\mathcal{H}(\mathbf{x}) = E \tag{1.6}$$

for a certain value of energy E , defines the microcanonical ensemble.

The central postulate of a priori probability in statistical mechanics states that all the microstates satisfying 1.6 are equally probable. In other words, one can define a probability density function

$$p(E, \mathbf{x}) = \frac{1}{\Omega(E)} \delta(H(\mathbf{x}) - E)$$

where $\Omega(E)$ denotes the volume of the phase space satisfying equation 1.6. We also assume the Boltzmann definition of entropy ²

$$S(E) = k_B \log(\Omega(E)) \quad (1.7)$$

Let us now consider two separated systems with fixed energies E_1, E_2 . By putting them into contact and allowing them to exchange energy, one can create another system with fixed energy $E = E_1 + E_2$ which can be studied in the microcanonical ensemble. For fixed values E_1 and E_2 the phase space volume allowed for the system is

$$\Omega_{E_1}(E, \mathbf{x}) = \Omega_1(E_1) \cdot \Omega_2(E_2)$$

but E_1 (and as a consequence $E_2 = E - E_1$) is free to move between 0 and E , hence the total phase space volume is given by an integral sum over E_1 and E_2 of the volumes at fixed E_1, E_2 with the constrain $E_2 = E - E_1$

$$\begin{aligned} \Omega(E) &= \int_0^E dE_1 \int_0^E dE_2 \Omega_1(E_1) \Omega_2(E_2) \delta(E_1 + E_2 - E) = \\ &= \int_0^E dE_1 \Omega_1(E_1) \Omega_2(E - E_1) \end{aligned}$$

By using now equation 4.1 the last equation can be written as

$$\Omega(E) = \int_0^E dE_1 e^{(S_1(E_1) + S_2(E - E_1))/k_B} \quad (1.8)$$

The extensivity of the entropy suggests that S_1 and S_2 are proportional to the number of particles in the corresponding systems. It is also legitimate to expect the entropy to have a global maximum, either inside or at the boundaries of the energy interval. Under this assumption, in the limit of a large number of particles, the integral can be evaluated using the Laplace's method. Hence, said E_1^* the maximum point of the exponential inside the integral in equation 1.8, the number of microstates can be approximated to

$$\Omega(E) \approx C e^{(S_1(E_1^*) + S_2(E - E_1^*))/k_B}$$

The energy value E_1^* is the one that is represented by the largest number of microstates, hence the most probable or, in other words, the one that it is most likely at equilibrium. This value can then be found by imposing the derivative to 0

$$0 = \frac{\partial}{\partial E_1} (S_1(E_1) + S_2(E - E_1)) = \frac{\partial S_1(E_1)}{\partial E_1} - \frac{\partial S_2(E_2)}{\partial E_2}$$

which leads to the condition

$$\frac{\partial S_1(E_1)}{\partial E_1} = \frac{\partial S_2(E_2)}{\partial E_2}$$

²This assumption is non trivial and will be further discussed in chapter 4.

Any two systems at equilibrium satisfy this last equation. For what told in the previous section the function $\frac{\partial S}{\partial E}$ must be an empirical temperature or, better, because of dimensional arguments, an inverse of a temperature. Hence the condition can be read as

$$T_1 = T_2$$

This justifies the definition given at the beginning of this section

$$\frac{1}{T} = \frac{\partial S(E)}{\partial E} \quad (1.9)$$

The derivation of the last equation made use only of the zeroth principle of thermodynamics and physical consideration about equilibrium: in this sense the temperature defined via 1.9 sometimes is also called *absolute temperature*.

1.3 Insight on the thermodynamic temperature

Here I want to provide an insight into the meaning of the formal definition $\frac{1}{T} = \frac{\partial S(E)}{\partial E}$. According to this definition, the temperature is a measure of the tendency of a system to increase/decrease entropy when an amount of energy is added. When the temperature is positive, adding energy to the system always increases its entropy, giving it access to more states. A negative temperature, on the other side, implies the fact that the system decreases its entropy when an amount of energy is added: in other words this means that the number of accessible states for the system decreases as the energy increases. One can imagine that if the maximum energy state of a system is realised only by one or few microstates, the system might admit a decreasing entropy as a function of energy, hence admitting negative temperatures. This idea will be formalized in *Ramsey's* criteria which provide the conditions under which a system admits negative temperature.

A rather counterintuitive result at negative absolute temperatures concerns the direction of heat flow. Consider two systems \mathcal{S}_1 and \mathcal{S}_2 isolated and individually in thermal equilibrium at temperatures T_1 and T_2 respectively. We say that \mathcal{S}_1 is *hotter* and \mathcal{S}_2 is *colder* if, when \mathcal{S}_1 and \mathcal{S}_2 are put into contact, the heat flows from \mathcal{S}_1 to \mathcal{S}_2 . When \mathcal{S}_1 and \mathcal{S}_2 are put into contact the change in \mathcal{S}_1 's entropy is $\delta S_1 = \frac{\delta Q}{T_1}$ and in \mathcal{S}_2 's is $\delta S_2 = -\frac{\delta Q}{T_2}$. If the process happens keeping the system $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$ thermally isolated from the environment, then we know that $\delta S = \delta S_1 + \delta S_2 > 0$ because of the second law of thermodynamics, which implies

$$\delta Q \left(\frac{1}{T_1} - \frac{1}{T_2} \right) > 0$$

One can now note that

1. If $T_1, T_2 > 0$ and $T_1 > T_2$ then it must be $\delta Q < 0$

2. If $T_1, T_2 < 0$ and $T_1 > T_2$ then it must be $\delta Q > 0$
3. If $T_1 > 0$ and $T_2 < 0$ then it must be $\delta Q > 0$

According to the convention that $\delta Q > 0$ means that the system (S_1) absorbs a heat δQ , condition 3. means that heat flows from the system at negative temperature to the one at positive temperature or, in other words, that systems at negative temperatures are hotter than those at positive ones. Instead, condition 2. means that when two systems at negative temperatures are put into contact, the system at lower temperature (or more negative temperature, that is S_2) gives heat to the one at higher temperature (or less negative temperature, that is S_1). If one would make a hierarchy of hotness of systems depending on the temperature, it would be

$$0^+ < +\infty < -\infty < 0^-$$

This fact will be further explored in [chapter 2](#).

Chapter 2

Two levels system

2.1 Two-levels systems admit negative temperatures

The simplest system that can exhibit negative temperatures is the two levels system (TLS).

A TLS is a system, for example a particle, for which only two values of energy are admitted, say E_1 and E_2 . Let us denote the corresponding eigenstates by $|1\rangle$ and $|2\rangle$. An example of such a system, which we may consider in the following derivations, can be a spin-1/2 fermion in a uniform magnetic field. Let us now consider a system composed of N TLS. It is convenient to introduce the occupation numbers n_1, n_2 which denote, respectively, the number of TLS at energy E_1 and E_2 . If we set $E_1 = \epsilon$ and $E_2 = 0$ for simplicity, the energy of the system is

$$E = n_1 E_1 + n_2 E_2 = n_1 \epsilon \quad (2.1)$$

where $n_1 + n_2 = N$.

One macrostate of the system is thus identified by its energy and the total number of particles. The number of microstates corresponding to one given macrostate is the number of ways in which one can rearrange the particles in a way such that the total energy remains fixed, that is

$$\Omega(E, N) = \frac{N!}{n_1! n_2!} = \frac{N!}{n_1! (N - n_1)!}$$

which corresponds to the Boltzmann entropy

$$S(E, N) = k_B \ln \left(\frac{N!}{n_1! (N - n_1)!} \right) \quad (2.2)$$

In the limit of large N the last expression can be expanded by using the Stirling's formula $\ln(N!) \approx N \ln N - N$ which yields

$$S(E, N) \approx N \ln \left(\frac{N}{N - n_1} \right) + n_1 \ln \left(\frac{N - n_1}{n_1} \right) \quad (2.3)$$

By using relation 2.1

$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{\partial S}{\partial n_1} \frac{\partial n_1}{\partial E} = \frac{k_B}{\epsilon} \ln \left(\frac{N - n_1}{n_1} \right) = -\frac{k_B}{\epsilon} \ln \left(\frac{E}{N\epsilon - E} \right) \quad (2.4)$$

where in the last step I used equation 2.1 again inside the logarithm's argument.

A plot of the temperature as a function of the system's energy is reported in figure 2.1. Negative temperatures occur in the region in which $E > \frac{N\epsilon}{2}$, which corresponds to the states in which there are more particles in the excited state than in the lower one. As mentioned at the end of chapter 1, negative temperatures occur above $+\infty$ rather than below the absolute zero. When a TLS is cooled down to 0°K all the spins lie in the ground state, i.e. the state of minimum energy. As soon as one adds energy to the system, the temperature starts increasing and more and more particles get excited to the higher energy state until the point in which, on average, we expect the particles to be half excited and half in the ground state, that is we have the same probability of having an excited particle or a particle in the ground state. This happens when the temperature is ideally infinite, either positive or negative, and this constitutes the state of maximum entropy. By keeping adding energy to the system one ends up having more particles in the excited state rather than in the lower one, experiencing the *population inversion*.

Equation 2.4 can be rewritten as

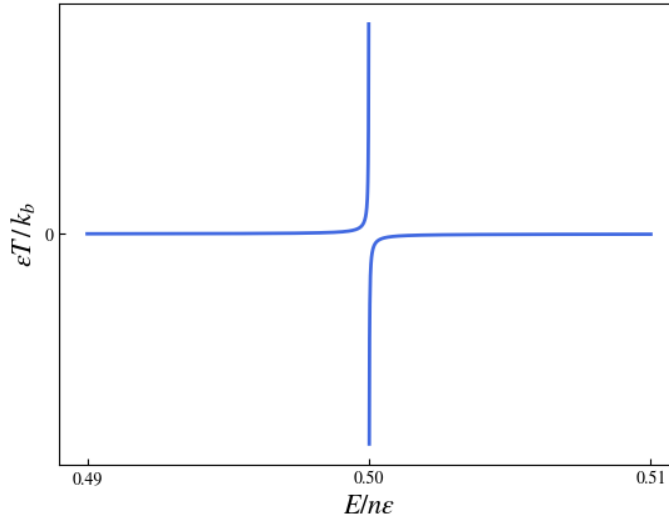


Figure 2.1: The plot reports the temperature as a function of the energy ($E/N\epsilon$) in a two-levels system. When there are more excited particles than those in the lower state (energy > 0.5), the system exhibits negative absolute temperatures.

$$\frac{1}{T} = \frac{k_B}{\epsilon} \ln \left(\frac{n_2}{n_1} \right)$$

which leads to

$$\frac{n_2}{n_1} = e^{-\epsilon/k_B T} \quad (2.5)$$

This formula allows us to introduce the so called *spin temperature*.

In the derivation of the results for the two level system we implicitly assume that the spin system is isolated from the environment and that the spins do not interact each other or, in other words, that each spin is isolated from the rest of the thermodynamic universe. In practice this is not the case and in general such a system is described by a Hamiltonian of the type

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{ss} + \mathcal{H}_{sl} \quad (2.6)$$

where H_0 represents the single particle Hamiltonian implicitly assumed in the previous derivation, H_{ss} represents the spin-spin interaction and H_{sl} represents the spin-lattice interaction.

The term H_{ss} can be conveniently considered negligible but cannot be zero: indeed this term plays a fundamental role for guaranteeing the *ergodicity* of the system.

The spin-lattice interaction on the other side can be characterized by a typical interaction time τ_L . The time τ_L characterizes the speed under which the spins adapt to a change in the lattice temperature. When the observation time of the experiment τ_s is much smaller than the typical interaction time $\tau_s \ll \tau_L$ the change in the system due to the spin-lattice interaction is negligible and equation 2.5 defines a temperature which is dependent only on the spins' configuration.

We call this temperature *spin temperature* [1]: Note that for $t \ll \tau_L$ this temperature may differ by the one of the lattice while for $t \gg \tau_L$ the two temperatures coincide.

Let us now recall what we mentioned at the end of chapter 1: a system whose maximum energy state is allowed by only one or few microstates may exhibit a decreasing entropy as a function of the energy, hence admitting negative temperatures. This is exactly the case of a TLS for which the maximum energy state corresponds to exactly one precise microstate, that is when all the particles are in the excited state. This of course corresponds to a null entropy. Analogously, the same happens at the minimum energy for which there is only one corresponding microstate and the entropy is null. For all the other states the entropy is non-zero and is given by formula 2.3. The whole expressions as a function of the energy can be easily obtained by 2.2 by multiplying and dividing by ϵ both inside and outside the logarithm

$$S(E, N)/k_B = N \ln \left(\frac{N\epsilon}{N\epsilon - E} \right) + \frac{E}{\epsilon} \ln \left(\frac{N\epsilon - E}{\epsilon} \right) \quad (2.7)$$

and is reported in figure 2.2.

This insight can be formalized into *Ramsey's criteria*.

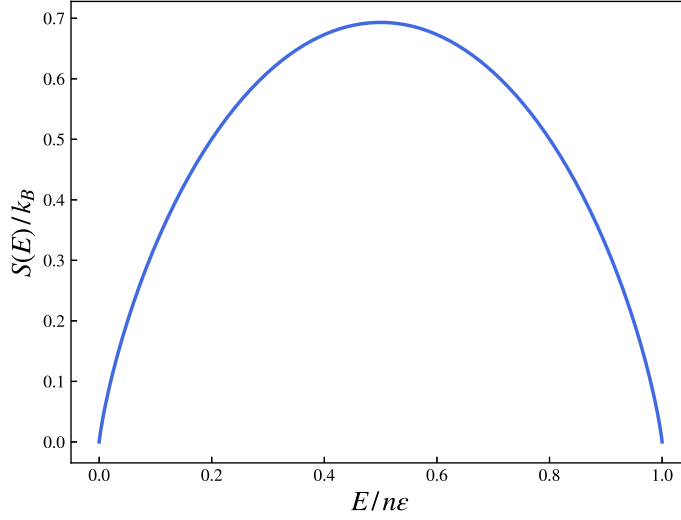


Figure 2.2: Entropy of a TLS as a function of the energy.

2.2 Ramsey's criteria

Ramsey [2] provided 3 conditions under which a thermodynamic system exhibits negative temperatures

1. *The various elements of the thermodynamic system under analysis must be at equilibrium with each other.*

This condition must be verified in order to define a temperature for the whole system.

2. *There must be an upper bound on the energy of the system*

In fact it is known from statistical mechanics that the probability (or probability density) for the system to be in a state of energy E is

$$p(E) \propto e^{-\beta E} \quad (2.8)$$

where $\beta = 1/k_B T$. If negative temperatures are admitted by the system and the latter does not admit an upper bound on the energy, the exponential factor in equation 2.8 becomes infinitely large for increasing energy. This means that if the system admits negative temperature, the higher the energy, the higher the probability of the system to be in that state. The most probable state would then be the one at infinite energy making the other states negligible in probability. Clearly, an infinite amount of energy cannot be put into the system, meaning that such a system cannot exist.

3. *The thermodynamic system under analysis must be thermally isolated from other systems that do not satisfy the above conditions.*

In good approximation one can think that the time required to reach equilibrium among the elements of the system is small compared to the time of interaction between the system and the environment or another system.

The above conditions are all satisfied by the TLS described above. Indeed it admits negative temperatures.

Chapter 3

Purcell and Pound experiment

3.1 Experimental evidence for negative temperature

The existence of negative absolute temperatures was first predicted by Lars Onsager in 1949 [3] in the context of 2D confined turbulence vortices.

The confinement of the vortices caused the phase space of the system to be bounded and Onsager showed that this results in a peak of the entropy as a function of the energy, in accordance to what stated by Ramsey's criteria.

The first experimental evidence of negative temperatures was the experiment carried by Purcell and Pound [4] in 1951, who managed to bring the spins of a LiF crystal in a negative absolute temperature state for several minutes. In this section I want to illustrate the experimental procedure followed by Purcell and Pound to realize such a state.

3.2 The experiment carried by Purcell and Pound

The experiment used a system of nuclear spins on a LiF crystal lattice immersed in a uniform magnetic field \mathbf{h} . The energy of the system is given by the analogous of equation 2.6

$$E = -\mathbf{h} \cdot \mathbf{M} + \mathcal{H}_{ss} + \mathcal{H}_{sl}$$

where \mathbf{M} is the magnetic moment vector $\mathbf{M} = g \frac{q}{2m} \mathbf{S} = \mu \mathbf{S}$, \mathbf{S} is the spin and g is the Landé factor.

As before we consider negligible the spin-spin interaction term but still considering it non-zero since it is fundamental to keep the system in internal equilibrium [5].

The spin-lattice interaction is characterized by the typical relaxation time τ_L which was found to be around 5 minutes for the considered system. For times much smaller than τ_L , the system can be considered as in a transient equilibrium, and the particles' state distribution defines a spin temperature according to equation 2.5. This temperature is precisely the one that was deduced to be negative. The effective energy for the system

in the experiment is then

$$E \approx -\mathbf{h} \cdot \mathbf{M} \quad (3.1)$$

Initially the system was brought into equilibrium aligning all the spins with the magnetic field in either a parallel or antiparallel way: for positive temperatures, for what said in chapter 2, we expect the majority of the spins to be parallel causing the energy given by equation 3.1 to be negative.

Then, suddenly, the magnetic field was switched in direction, causing the majority of the spins to be antiparallel to the magnetic field, hence giving raise to a positive energy. This state is characterized by a negative spin temperature. Figure 3.1 illustrates schematically the states before and after the change in the magnetic field. The temperature is the inverse of the slope of the tangent of the plotted function in the the desired point.

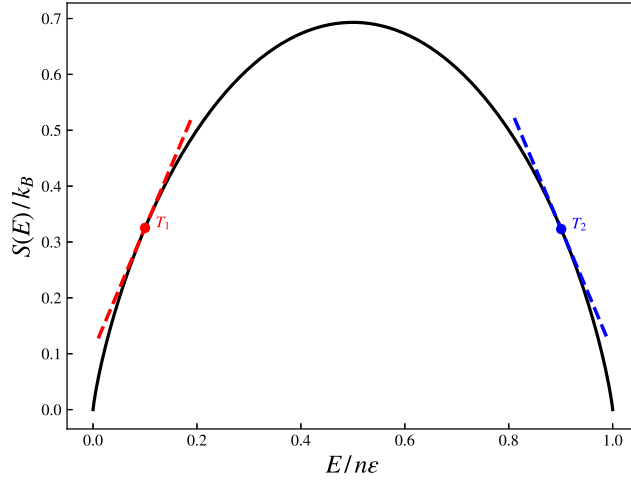


Figure 3.1: The figure reports the relation between entropy and energy in a two levels system as found in chapter 2. The system before the experiment was at equilibrium at a positive temperature T_1 , while after the magnetic field inversion, the system was found to be in an inverted population state described by a negative temperature T_2 . The temperature is the reciprocal of the slope of the tangent of the plotted function in the two points.

Experimentally it was known that the spins were aligned antiparallel with respect to the magnetic field by looking at the magnetic susceptibility of the system. To define such a quantity, let us first define the magnetization of the system as

$$M = \sum_i M_i \quad (3.2)$$

where M_i is the projection of the magnetic moment vector \mathbf{M} of the i -th particle along the magnetic field axis, or, in other words,

$$M_i = \mu S_i$$

where S_i is the spin of the i -th particle.

The magnetic susceptibility χ is then defined as the constant such that

$$M = \chi H$$

A negative susceptibility then implies that the system is in an inverted population state, which means at negative temperatures. The susceptibility as a function of time in the experiment is reported in figure 3.1, directly exported from the original article by Purcell and Pound [4].

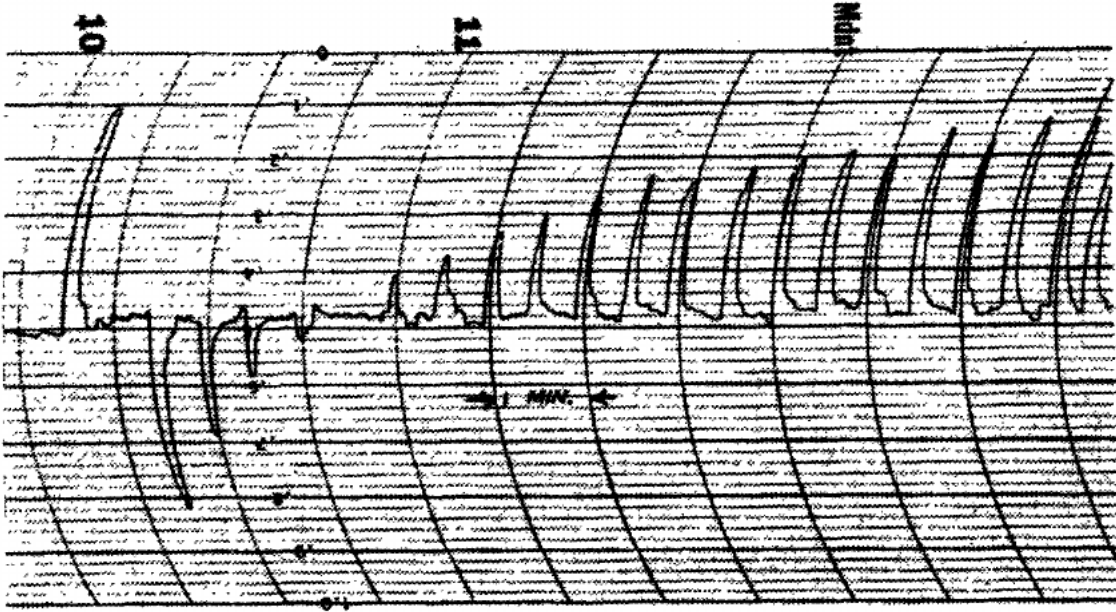


Figure 3.2: The plot (exported from [4]) reports the record of the magnetic susceptibility during the experiment. The first peak denotes the initial equilibrium state in which the spins are aligned with the magnetic field in the lowest energy state. As the magnetic field is instantaneously reversed, the spins remained aligned in the opposite direction of the field, remaining in the highest energy state and causing the second (negative) peak. After a relaxation time τ_L the spin system returns at equilibrium in the lowest energy state.

Chapter 4

Issues on the definition of entropy

The definition of temperature via

$$\frac{1}{T} = \frac{\partial S(E)}{\partial E}$$

is strictly dependent on the definition of the entropy and the existence of negative temperatures is mainly a consequence of this definition. It was already discussed in chapter 1 why it is legitimate to define the temperature as the derivative of the entropy with respect to the energy. The purpose of this section, on the other side, is to discuss the definition of the entropy with particular regards to the definitions given by Gibbs and Boltzmann since they have direct consequences on the existence of negative temperatures.

After a brief introduction over the two perspectives, the focus of the discussion will be on the comparison between the two definitions with particular emphasis on the direct consequences on the negative temperatures.

4.1 Boltzmann's framework

The conceptual basis for Boltzmann's approach to statistical mechanics presented in Boltzmann's paper (1877b) relies on the attempt to explain the Second Law of thermodynamics via probability calculus.

To introduce the idea let us consider a system of N particles and let us work in a $6N$ dimensional phase space with coordinates $q_1, \dots, q_{3N}, p_1, \dots, p_{3N}$ and let us consider also the μ -space associated to each particle in the system, i.e. the phase space in which each single particle moves.

Let us partition each μ -space into m disjoint rectangular cells of volume $\Delta\omega$ so that $\mu = \omega_1 \cup \dots \cup \omega_m$ and each cell ω_i is characterized by an energy value ϵ_i . Once specifying the mechanical state of the system, a point $x \in \Gamma$, one can associate a collection of N points in the μ -spaces, one for each particle (the spaces are the same for each particle). For each x , also called *microstate* of the system, one can define a *macrostate* Z for the system by specifying the number of particles n_i included in each cell ω_i in the μ -space.

Formally $Z = (n_1, \dots, n_m)$ where n_i is the number of particles in the cell $\omega_i \subset \mu$.

By this definition it is clear that more than one microstate can describe the same macrostate of the system. One can associate a phase space volume to a macrostate Z_0 , that is the set of corresponding microstates

$$\Gamma_{Z_0} \equiv \{x \in \Gamma : Z(x) = Z_0\}$$

The Boltzmann's entropy of a system in a macrostate Z is then defined as

$$S = k_B \ln(\text{Vol}(\Gamma_Z)) \quad (4.1)$$

In the case of a particle moving in the phase space, the microstates are counted by dividing the volume available for the particle in the phase space by the equivalent volume of one microstate, which, after quantum mechanics, it is known to be h^3 . In formulas, equation 4.1 reduces to

$$S(E) = k_B \ln \omega(E)$$

where

$$\omega(E) = \frac{1}{h^3} \int d^3q d^3p \delta(E - H(q, p))$$

4.2 Gibbs' framework

Gibbs' approach to statistical mechanics is based on the idea of *statistical ensemble*. To introduce this concept let us work again in a phase space Γ and describe the system via $3N$ canonical coordinates, so that Γ is a $6N$ dimensional space. A point in Γ denotes a precise configuration of the system and it is referred to as a *representative point*.

A given macroscopic configuration of the system can correspond to multiple microscopic configurations, that is multiple points in the Γ space might correspond to the same macroscopic state.

In other words, when specifying a precise macroscopic configuration, we are not referring to one system, but rather to a collection of systems which we call an *ensemble*.

An ensemble is conveniently described by means of a *density function* $\rho(q, p, t)$ such that $\rho(q, p, t) d^{3N}q d^{3N}p$ is the number of representative points in a phase space volume $d^{3N}q d^{3N}p$.

Given the value of ρ at time $t = 0$, the evolution of the function is completely determined by means of the Hamilton equation

$$\begin{aligned} \frac{dp_i}{dt} &= -\frac{\partial \mathcal{H}}{\partial q_i} \\ \frac{dq_i}{dt} &= \frac{\partial \mathcal{H}}{\partial p_i} \end{aligned}$$

More precisely the evolution of ρ is determined by the *Liouville's theorem* which states that

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial p_i} \dot{p}_i + \frac{\partial \rho}{\partial q_i} \dot{q}_i \right) = 0 \quad (4.2)$$

or

$$\frac{\partial \rho}{\partial t} = \{H, \rho\}$$

In developing his theory, Gibbs' main goal was to produce a rational fundation for thermodynamics. Hence, Gibbs's work was guided by analogies between his theory and thermodynamics.

In his book [6] Gibbs derived a relation in the canonical ensemble that is

$$d\langle H \rangle = \theta d\sigma - \sum_i \langle A_i \rangle da_i \quad (4.3)$$

which is formally analogous to the fundamental equation of thermodynamics

$$dU = TdS + \sum_i F_i da_i$$

where $\langle H \rangle$ in equation 4.3 denotes the expectation value of the hamiltonian in the canonical ensemble. The analogy suggests that θ , the so called *modulus of the ensemble*, can be identified as the temperature of the system and σ , which was defined as

$$\sigma[\rho_\theta] = - \int \rho_\theta(x) \ln \rho_\theta(x) dx$$

can be identified as the entropy of the system, namely the *Gibbs entropy*.

The important point is that, in the Gibbs' approach, the entropy is not a function on the phase space but rather a functional on the ensemble density ρ_θ . In principle one can propose a couple (θ, σ) arbitrarily chosen and check if relation 4.3 is satisfied.

The next step is to understand whether an equation such 4.3 can be obtained in the microcanonical ensemble. Gibbs showed (see [6] page 124-128, 169,171) that the relation is satisfied with the following definitions

$$\begin{aligned} T &\longleftrightarrow \left(\frac{\partial \ln \Omega(E)}{\partial E} \right)^{-1} \\ S &\longleftrightarrow \ln \Omega(E) \end{aligned}$$

where

$$\Omega(E) \equiv \int_{H(x) \leq E} dx \quad (4.4)$$

is called *integrated density of states*.

4.3 Consistent thermodynamics forbids negative temperatures

The difference between the Gibbs entropy ¹

$$S_G = k_B \ln \Omega(E) \quad (4.5)$$

¹Boltzmann's constant is added for dimensional arguments.

and the Boltzmann's entropy

$$S_B = k_B \ln \omega(E) \quad (4.6)$$

has a direct and important consequence on the existence of negative temperatures. Indeed it is easy to see that the integrated density of states defined in equation 4.4, that is the number of states whose energy is less than or equal to E , is a monotonically increasing function of the energy, hence the temperature is always positive. On the other side, the density of states $\omega(E)$ that enters in the Boltzmann's entropy denotes the number of states in the range $(E, E + dE)$ and there is no reason to believe that is a monotonically increasing function of the energy. In fact we have already seen that in the case of a non interacting two levels system the Boltzmann's entropy leads to negative temperatures. The choice of the correct microcanonical entropy is not a straightforward process since different definitions of the entropy can be used until they are able to reproduce thermodynamics.

In 1991 Berdichevsky *et al.* [7] pointed out some arguments in favour of Gibbs' entropy, which have then been developed by many authors. The main arguments were proposed by Dunkel and Hillbert in 2014 [8] introducing the so called *thermostatistical consistency condition*.

Let us suppose that the system is described by some control variables $\{E, V, A_i\}$ so that $S = S(E, V, A_i)$. By differentiating S with respect to the control variables one obtains

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial E} \right) dE + \left(\frac{\partial S}{\partial V} \right) dV + \sum_i \left(\frac{\partial S}{\partial A_i} \right) dA_i \\ &\equiv \frac{1}{T} dE + \frac{p}{T} dV + \sum_i \frac{a_i}{T} dA_i \end{aligned}$$

If one now considers an isentropic process for which only energy and one control variable are allowed to change, one gets that

$$T \left(\frac{\partial S}{\partial A_\mu} \right)_E = - \left(\frac{\partial E}{\partial A_\mu} \right)_S = - \left\langle \frac{\partial H}{\partial A_\mu} \right\rangle \quad (4.7)$$

Now, since

$$\begin{aligned} T_B(E, V, A) &= \left(\frac{\partial S_B}{\partial E} \right)^{-1} = \frac{1}{k_B} \frac{\Omega'}{\Omega''} \\ T_G(E, V, A) &= \left(\frac{\partial S_G}{\partial E} \right)^{-1} = \frac{1}{k_B} \frac{\Omega}{\Omega'} \end{aligned} \quad (4.8)$$

where the prime indicates the differentiation with respect to the energy of the system, one has that

$$\begin{aligned} T_B \left(\frac{\partial S_B}{\partial A_\mu} \right) &= \frac{1}{\Omega''} \frac{\partial \Omega'}{\partial A_\mu} \\ T_G \left(\frac{\partial S_G}{\partial A_\mu} \right) &= \frac{1}{\Omega'} \frac{\partial \Omega}{\partial A_\mu} \end{aligned}$$

so in general we expect that if one of the two definitions of entropy satisfies condition 4.7, the other one does not, and it is easy to see that the Gibbs' entropy does

$$\begin{aligned} T_G \left(\frac{\partial S_G}{\partial A_\mu} \right) &= \left(\frac{\partial S_G}{\partial E} \right)^{-1} \left(\frac{\partial S_G}{\partial A_\mu} \right) = \frac{1}{\omega} \frac{\partial}{\partial A_\mu} \int dx \theta(E - \mathcal{H}(x)) \\ &= \frac{1}{\omega} \int dx \frac{\partial \theta(E - \mathcal{H}(x))}{\partial A_\mu} = -\frac{1}{\omega} \int dx \delta(E - \mathcal{H}(x)) \frac{\partial \mathcal{H}(x)}{\partial A_\mu} \\ &= - \int dx \frac{\delta(E - \mathcal{H})}{\omega} \frac{\partial \mathcal{H}}{\partial A_\mu} = - \int dx \rho \frac{\partial \mathcal{H}}{\partial A_\mu} = - \left\langle \frac{\partial \mathcal{H}}{\partial A_\mu} \right\rangle \end{aligned}$$

Another argument in favour of Gibbs' entropy, still proposed by Dunkel and Hillbert [8], is that it satisfies the generalized equipartition theorem

$$\left\langle \xi_i \frac{\partial \mathcal{H}}{\partial \xi_j} \right\rangle = \delta_{ij} T \quad (4.9)$$

while Boltzmann's does not.

To see this one can consider the simple example of an ideal gas confined in a cubic box of side L . In this case the hamiltonian reads

$$\mathcal{H}(t, \mathbf{q}, \mathbf{p}) = \sum_i \frac{\mathbf{p}_i^2}{2m} + V(\mathbf{q}_i)$$

where

$$V(\mathbf{q}_i) = \begin{cases} 0 & \text{if } \mathbf{q}_i \in \text{box} \\ +\infty & \text{otherwise} \end{cases}$$

The integrated density of states, that is the number of states with energy less than or equal to E , can be computed as

$$\begin{aligned} \Omega(E) &= \frac{1}{h^{3N}} \int d^{3N} p \int d^{3N} q \theta(E - \mathcal{H}(t, \mathbf{q}, \mathbf{p})) = \\ &= \frac{V^N}{h^{3N}} \int d^{3N} p \theta \left(E - \sum_i \frac{\mathbf{p}_i^2}{2m} \right) = \frac{V^N (2\pi m E)^{3N/2}}{h^{3N} \Gamma(\frac{3N}{2} + 1)} \end{aligned}$$

so that the density of states is

$$\omega(E) = \frac{\partial \Omega}{\partial E} = \frac{3NV^N (2\pi m)^{3N/2}}{2h^3 \Gamma(\frac{3N}{2} + 1)} E^{\frac{3N}{2}-1}$$

One can now calculate the two temperatures using 4.8

$$\begin{aligned} \frac{1}{T_G} &= \frac{\partial S_G}{\partial E} = \frac{\partial \ln \Omega}{\partial E} = \frac{h^3 \Gamma(\frac{3N}{2} + 1)}{V^N (2\pi m E)^{3N/2}} \frac{3NV^N (2\pi m)^{3N/2}}{2h^3 \Gamma(\frac{3N}{2} + 1)} E^{\frac{3N}{2}-1} \\ \frac{1}{T_B} &= \frac{\partial S_B}{\partial E} = \frac{\partial \ln \omega}{\partial E} = \left(\frac{3NV^N (2\pi m)^{3N/2}}{2h^3 \Gamma(\frac{3N}{2} + 1)} E^{\frac{3N}{2}-1} \right)^{-1} \frac{3N}{2} \left(\frac{3N}{2} - 1 \right) E^{\frac{3N}{2}-2} \end{aligned}$$

and rearranging terms

$$\begin{aligned} E &= \frac{3N}{2} T_G \\ E &= \left(\frac{3N}{2} - 1 \right) T_B \end{aligned} \tag{4.10}$$

showing that only the Gibbs' entropy satisfies the equipartition theorem.

The fact that Gibbs' entropy was "the correct entropy" created some disturb. This was mainly because of the unphysical meaning of Gibbs' entropy. Why should one take into account the states with a different energy of the system when calculating the number of microstates available for a given energy E ?

Another position on the issue was indeed argued by Frenkel and Warren [9] who not only showed that Boltzmann's entropy satisfies the thermodynamic condition in the thermodynamic limit, but also that Gibbs' entropy fails to meet a basic thermodynamic criterion which is the equality of the temperatures of two systems in reciprocal equilibrium.

Let us consider a two level system \mathcal{S}_1 with bounded phase space (e.g. a two levels system) in an inverted population state. According to what seen in chapter 2, the system is described by a negative temperature, say $T_{1,B}$, using the Boltzmann's definition. Instead if one makes use of the Gibbs' definition of temperature, the latter is certainly positive since T_G does not admit negative values: let us call this value $T_{1,G}$.

It is then possible to create another system \mathcal{S}_2 with an unbounded phase space, for example a free ideal gas, such that $T_{2,G} = T_{1,G}$, but obviously $T_{2,B} \neq T_{1,B}$, since the free ideal gas does not admit negative Boltzmann temperature. According to what told in chapter 1, $T_{2,G} = T_{1,G}$ implies that the two systems are in equilibrium, while $T_{2,B} \neq T_{1,B}$ predicts them to be out of equilibrium. When the two systems are put into contact, they are allowed to exchange heat keeping the total energy of the collective system constant. The configuration that is most probable at equilibrium is the one that maximizes the entropy of the whole system compatibly with the available energy.

Since the system \mathcal{S}_1 is not in its maximum entropy state, when allowed to exchange heat to \mathcal{S}_2 , it lowers its energy releasing heat increasing also \mathcal{S}_2 's entropy. This means that \mathcal{S}_1 and \mathcal{S}_2 were out of equilibrium implying Gibbs's temperature to be wrong.

The only issue that needs to be solved now is to understand why Boltzmann's entropy does not satisfy the thermodynamic consistency relation or the equipartition theorem as previously shown. Frenkel and Warren argued that the conditions are instead satisfied also by the Boltzmann's definition of entropy in the thermodynamic limit, that is for $N \gg 1$, that is the limit in which statistical mechanics is expected to connect to thermodynamics. This is obvious for the particular case of the ideal gas which leads to 4.10 and it is immediate to check that the two expressions coincide in the limit of large N . To prove that the consistency relation is respected in the general case, let us first start by proving that, in the canonical ensemble, the following relation holds

$$-\left. \frac{\partial F}{\partial A_\mu} \right|_T = - \left\langle \frac{\partial H}{\partial A_\mu} \right\rangle_T \tag{4.11}$$

Since $F = -\frac{1}{\beta} \ln Z$, where

$$Z = \sum_x e^{-\beta H(x)}$$

then

$$e^{-\beta F} = \sum_x e^{-\beta H(x)} \quad (4.12)$$

Differentiating both members with respect to A_μ gives

$$-\beta \frac{\partial F}{\partial A_\mu} \Big|_T e^{-\beta F} = -\beta \sum \frac{\partial H}{\partial A_\mu} e^{-\beta H} \quad (4.13)$$

and by taking the ratio of equations 4.12 and 4.13, the result is proved.

Now, to prove the consistency relation 4.7, we need to move to the microcanonical ensemble. In this case the averages are performed as

$$\langle \dots \rangle_E = \frac{\int (\dots) \delta(E - H)}{\int \delta(E - H)}$$

and the equivalent of 4.12 for the entropy is

$$\int \delta(E - H) = e^{S_B/k_B}$$

One can write that

$$\begin{aligned} \int dx (...) e^{-\beta H(x)} &= \int dE \int dx (...) \delta(E - H(x)) e^{-\beta E} = \\ &= \int dE e^{-\beta E} \int dx (...) \delta(E - H(x)) \end{aligned}$$

Let us multiply and divide the left member by $Z = e^{-\beta F}$ and the right member by $e^{-\beta E + S_B/k_B}$ obtaining

$$e^{-\beta F} \langle \dots \rangle_T = \int_0^\infty dE e^{-\beta E + S_B/k_B} \langle \dots \rangle_E \quad (4.14)$$

where $\langle \dots \rangle_T$ and $\langle \dots \rangle_E$ indicate the averages in the canonical and microcanonical ensemble respectively. The last equation, in the special case of $(...) = 1$ yields

$$e^{-\beta F} = \int e^{-\beta E + S_B/k_B}$$

Now, the last integral is peaked around a value E^* and the higher the number of degrees of freedom in the system, the more peaked the integrand. In the thermodynamic limit, let us then evaluate this last integral by the saddle point method. We obtain that

$$-\beta F = -\beta E + \frac{S_B}{k_B}$$

where $-\beta + \frac{\partial}{\partial E} \left(\frac{S_B}{k_B} \right) = 0$ (condition for the extremum). Recalling that $\beta = \frac{1}{k_B T}$, the condition for the extremum states that $\frac{1}{T} = \frac{\partial S_B}{\partial E} = \frac{1}{T_B}$. Hence

$$F = E - TS \quad (4.15)$$

where $T = T_B = \frac{\partial S_B}{\partial E}$ which is the definition of the Legendre transform of the energy with respect to the couple $T \leftrightarrow S$.

In an analogous way it is possible to apply the same method to 4.14 obtaining as a result that

$$\langle \cdots \rangle_T = \langle \cdots \rangle_E \quad (4.16)$$

Now, by differentiating equation 4.15 with respect to A_μ , one gets that

$$\left. \frac{\partial F}{\partial A_\mu} \right|_T = \left(1 - T \frac{\partial S_B}{\partial E} \right) \frac{\partial E}{\partial A_\mu} - T \left. \frac{\partial S_B}{\partial A_\mu} \right|_E = -T \left. \frac{\partial S_B}{\partial A_\mu} \right|_E$$

combining this result with 4.11, 4.16 and the fact that $T = T_B = \frac{\partial S}{\partial E}$, we get the desired result

$$T_B \left. \frac{\partial S_B}{\partial A_\mu} \right|_E = - \left\langle \frac{\partial H}{\partial A_\mu} \right\rangle_E$$

Chapter 5

Simulations at negative temperatures

In this last section I move the focus towards the study of a more complex model, namely the Ising model [10]. The higher complexity of the problem requires a numerical solution via a Monte Carlo simulation.

After a brief introduction on the 2D Ising model, I will briefly recap the main points of a Markov Chain Monte Carlo algorithm, moving then towards the numerical experiments. The results of the simulations are presented and discussed according to what said in the previous chapters.

5.1 The Ising model

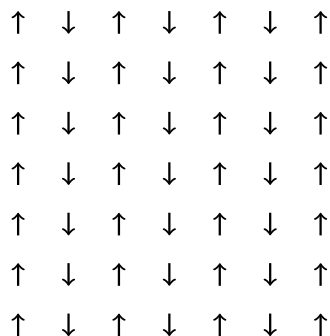


Figure 5.1: 2D Ising model. The spins are organised in a 2D squared lattice and can assume only two values, either up or down. In the most general way the system is described by the Hamiltonian given by equation 5.1

The Ising model is a mathematical model initially used to describe magnetic properties of materials in statistical mechanics. It consists of a collection of spins organised in a lattice that can assume two values, either 1 or -1.

The spins are allowed to interact with their nearest neighbors and in the most general form the Hamiltonian assumes the form

$$\mathcal{H}(\{\sigma_k\}) = - \sum_{\langle ij \rangle} J_{ij} \sigma_i \sigma_j - \sum_j B_j \sigma_j \quad (5.1)$$

where B_j indicates the intensity of the magnetic field on the spin j and J_{ij} indicates the strength of the interaction between the spin i and the spin j . In the following analysis I will assume $J_{ij} = J, B_j = B \forall i, j$.

The sign of the interaction parameter J determines whether we are dealing with ferromagnetic materials ($J > 0$) or antiferromagnetic materials ($J < 0$). In the case of $J = 0$ we return to the two levels system discussed in chapter 2.

Let us first restrict the analysis to the purely interacting case setting $B = 0$. In this case the Hamiltonian initially given by 5.1 becomes

$$\mathcal{H}(\{\sigma_k\}) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j$$

At fixed temperature T the probability of a state is given by the Boltzmann factor

$$p(\{\sigma_k\}) = \frac{e^{-\beta \mathcal{H}(\{\sigma_k\})}}{Z} \quad (5.2)$$

where $\beta = \frac{1}{k_B T}$ and Z is the canonical partition function

$$Z(\beta, B, N) = \sum_{\{\sigma_k\}} e^{-\beta \mathcal{H}(\{\sigma_k\})}$$

Let us consider the probability factor 5.2, which explicitly becomes

$$p(\{\sigma_k\}) \propto e^{\beta J \sum_{\langle ij \rangle} \sigma_i \sigma_j}$$

and let us define a new parameter $\xi = \beta J$ so that

$$p(\{\sigma_k\}) \propto e^{\xi \sum_{\langle ij \rangle} \sigma_i \sigma_j} \quad (5.3)$$

5.1.1 Ferromagnetic systems

In the ferromagnetic case $J > 0$, hence $\xi > 0$.

- For $\xi \rightarrow 0$, which can be either a high temperature or a weak coupling, one can see that $p(\{\sigma_k\}) \rightarrow 1$ independently of the values σ_k so that every distribution $\{\sigma_k\}$ is equally probable. Thus, macroscopically speaking, we expect the system to be found in a *disordered* state.
- When $\xi \rightarrow +\infty$ the exponential factor in 5.3 grows indefinitely if $\sum_{\langle ij \rangle} \sigma_i \sigma_j > 0$, that is when the spins are aligned. We then expect the system to be in a *highly ordered* configuration organised into clusters of spins aligned in the same direction. Thus no preference is given to the configuration in which spins are aligned both up or down, and we expect it to depend on the history of the system.

5.1.2 Antiferromagnetic systems

In this case $J < 0$, hence $\xi < 0$.

- For $\xi \rightarrow 0$ the result is the same of ferromagnetic systems.
- On the other side for $\xi \rightarrow -\infty$, which can be either a low temperature or a strong coupling, one can see that $p(\{\sigma_k\}) \rightarrow 0$ except when $\sum_{\langle ij \rangle} \sigma_i \sigma_j = 0$, that is when neighbour spins are disaligned. We then expect the system to be found in an *ordered configuration* in which neighbor spins are disaligned.

The study is carried in a 2D lattice which is of the shape reported in figure 5.1. Let us focus the attention on a 2×2 subset of the lattice and let us consider $J > 0$ as an example. The possible energies are

$$\begin{array}{ll}
 \begin{pmatrix} \uparrow & \uparrow \\ \uparrow & \uparrow \end{pmatrix}, \begin{pmatrix} \downarrow & \downarrow \\ \downarrow & \downarrow \end{pmatrix} & E = -4J \\
 \begin{pmatrix} \downarrow & \uparrow \\ \uparrow & \uparrow \end{pmatrix}, \begin{pmatrix} \uparrow & \downarrow \\ \uparrow & \uparrow \end{pmatrix}, \begin{pmatrix} \uparrow & \uparrow \\ \downarrow & \uparrow \end{pmatrix}, \begin{pmatrix} \uparrow & \uparrow \\ \uparrow & \downarrow \end{pmatrix} & E = 0J \\
 \begin{pmatrix} \downarrow & \downarrow \\ \uparrow & \uparrow \end{pmatrix}, \begin{pmatrix} \downarrow & \uparrow \\ \downarrow & \uparrow \end{pmatrix}, \begin{pmatrix} \uparrow & \downarrow \\ \uparrow & \downarrow \end{pmatrix}, \begin{pmatrix} \uparrow & \uparrow \\ \downarrow & \downarrow \end{pmatrix} & E = 0J \\
 \begin{pmatrix} \uparrow & \downarrow \\ \downarrow & \downarrow \end{pmatrix}, \begin{pmatrix} \downarrow & \uparrow \\ \downarrow & \downarrow \end{pmatrix}, \begin{pmatrix} \downarrow & \downarrow \\ \uparrow & \downarrow \end{pmatrix}, \begin{pmatrix} \downarrow & \downarrow \\ \downarrow & \uparrow \end{pmatrix} & E = 0J \\
 \begin{pmatrix} \downarrow & \uparrow \\ \uparrow & \downarrow \end{pmatrix}, \begin{pmatrix} \uparrow & \downarrow \\ \downarrow & \uparrow \end{pmatrix}, & E = 4J
 \end{array}$$

The whole system is then a repetition of these blocks and the total energy of the system is the sum of the energies of each block. One can observe that

1. The lowest energy configuration is the one with all the spins aligned either up or down as predicted previously.
2. The highest energy configuration corresponds to the lowest energy configuration of the antiferromagnetic case according to what predicted previously. I will show later that this is explained if the ferromagnetic system is described by a negative temperature, more precisely $T = -0 \text{ K}$.
3. The number of states at energy given energy E is equal to the number of states at energy $-E$. This means that the density of states function $\omega(E)$ is an even function of the energy, that is $\omega(E) = \omega(-E)$. Remembering that the derivative of a continuous even function is an odd function, one has that the temperature function

$$\frac{1}{T} = \frac{\partial S}{\partial E} = k_B \frac{\omega'(E)}{\omega(E)}$$

is an odd function of the energy, hence proving that the described model admits negative temperatures. Since the maximum point for $\omega(E)$ is expected for $E = 0$, $\omega'(E)$ is negative for $E > 0$, suggesting that negative temperatures describe high energy configurations of the system.

In what follows I will adopt the so called *free boundary conditions*, which means not putting any constraints on the spins at the boundary of the system.

5.2 Markov Chain Monte Carlo

To study the physical properties of the system at equilibrium in a certain macrostate X , one needs to perform ensemble averages (or expectation values). The expectation value of a physical quantity A is computed as

$$\bar{A} = \frac{\int A e^{-\beta H} d\tau}{\int e^{-\beta H} d\tau} \quad (5.4)$$

where the variable τ runs over all the possible microstates compatible with the given macrostate X .

Practically speaking, to perform such averages during the simulation, one needs to

1. Bring the system in a configuration compatible with the specified macrostate
2. Calculate the value of the desired quantity in the specific microstate and store the value
3. Move the system to another microstate compatible with the specified macrostate and repeat from point 2 until desired
4. Perform the average

Of course, in order to perform an accurate average, one needs to

1. Guarantee that the system visits every possible microstate compatible with the macroscopical configuration so that all the possible values of the physical quantity are taken into account when performing the average
2. Guarantee that every possible state is visited with a frequency that is equal to the probability of the state itself

This can be accomplished with a Monte Carlo method, precisely by means of the *Metropolis-Hastings algorithm*, which allows one to perform a random walk through the configurations of the system and respect the above stated conditions.

The main points of the Metropolis-Hastings' [11] algorithm are now briefly reported, inviting the reader to consult more appropriate sources for further information.

Let us suppose that we want to sample from a probability distribution $p(x)$. The system is initially prepared in a certain configuration x which can be arbitrarily chosen. Each Monte-Carlo cycle then consists in the following steps

1. A new configuration of the system x^* is proposed sampling from a proposal distribution $J(x^*|x_i)$. The proposal distribution can be arbitrarily chosen, provided the fact that it is symmetric $J(x^*|x_i) = J(x_i|x^*)$ ¹ and that in an infinite number of steps, it guarantees that every state is visited.
2. The ratio of the probabilities of the two states of the system is then calculated

$$w = \frac{p(x^*)}{p(x_i)}$$

3. if $w \geq 1$ the proposal change is accepted and $x_{i+1} = x^*$
4. otherwise a number r is extracted from a uniform distribution in $[0, 1]$. If $w > r$ the proposal change is still accepted and $x_{i+1} = x^*$. Otherwise the proposal moved is refused and $x_{i+1} = x_i$

The higher the number of Monte Carlo cycles, the closer the distribution of the visited states to the probability distribution $p(x)$, the more precise the expectation values calculated as 5.4.

The simplest choice for the proposal distribution $J(x^*|x)$, even if not the most efficient, is the uniform distribution. The distribution $p(x)$, on the other side, is related to the ensemble in which we describe the system: in the canonical ensemble $p(x)$ is given by the Maxwell-Boltzmann distribution given by equation 5.2.

¹Strictly speaking this is the simple Metropolis algorithm. The more general Metropolis-Hastings algorithm does not need the proposal distribution to be symmetric, provided the probability ratio w is modified to $w = \frac{J(x_i|x^*)p(x^*)}{J(x^*|x_i)p(x_i)}$.

5.3 Single system at negative temperature

The first case I want to study is the equilibrium configuration of one single system prepared at a given temperature T . After fixing the temperature, the system is prepared in a random configuration: then the system goes under 10^6 Monte Carlo cycles sampling from the probability distribution given by 5.2. An example of initial configuration is reported in figure 5.2.

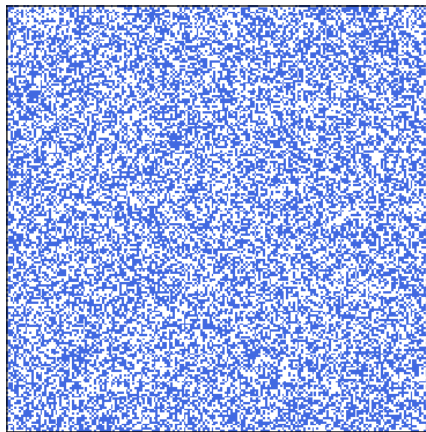


Figure 5.2: Initial configuration for the system. Each spin value was extracted randomly between $+1$ or -1 . A blue tile indicates an up spin, while a white one indicates a down spin.

The simulations are run for four different values of temperature, that is

$$T_1 = 0.01 \text{ K} \quad T_2 = -0.01 \text{ K} \quad T_3 = 100 \text{ K} \quad T_4 = -100 \text{ K}$$

Each of them is run both for the ferromagnetic and antiferromagnetic cases.

Figure 5.3 reports the final states of a ferromagnetic system, while figure 5.4 reports those of an antiferromagnetic system.

According to what predicted in paragraph 5.1.1, at low positive temperature the system tends to organize in an ordered structure grouping the spins into bigger clusters. As the temperature increases, the system starts forming imperfections increasing its disorder. This can be understood also in terms of the minimal free energy principle. In the canonical ensemble the most probable configuration is the one that minimizes the free energy

$$F(T, V) = E - TS \tag{5.5}$$

which is the Legendre transform of the internal energy $E(S, V) = TS - PV$ with respect to the conjugate variables $T \leftrightarrow S$. The free energy minimization at fixed temperature, according to the definition 5.5, can be driven either by the minimization of the internal

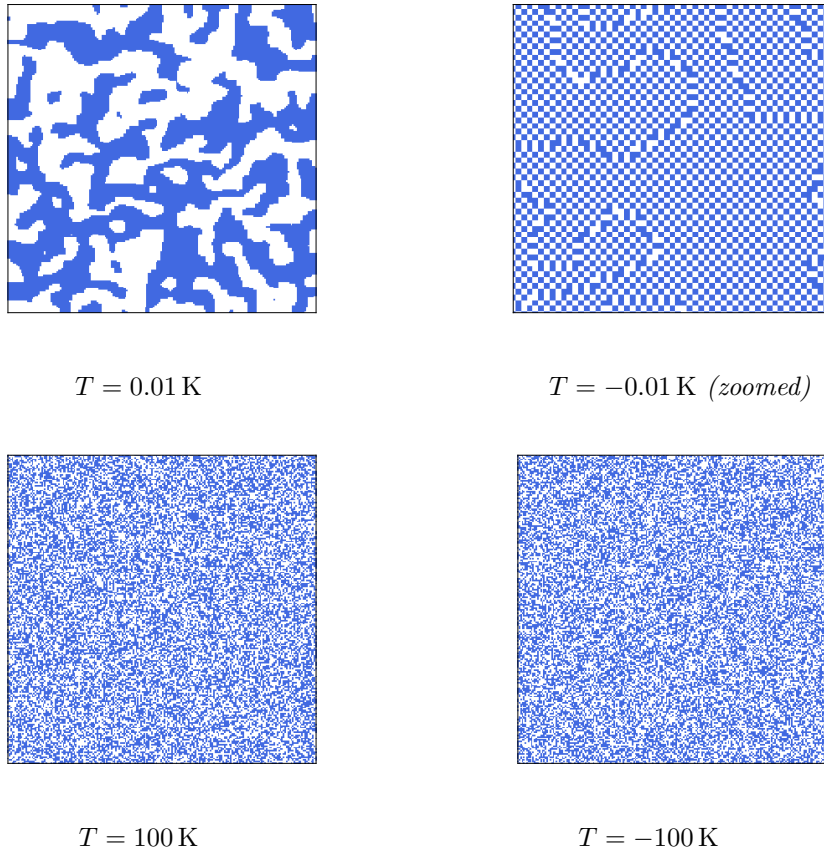


Figure 5.3: Ferromagnetic system

energy or by an increase in the entropy. The higher the temperature, the higher the entropic contribution, hence explaining why we find a disordered configuration at high temperature.

As for the ferromagnetic case, the results for the antiferromagnetic system agree with what predicted in section 5.1.2. For high temperature the free energy minimization is driven by the entropic contribute, leading to a high disordered configuration. For low positive temperature the minimization of the energy

$$\mathcal{H}(\{\sigma_k\}) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j$$

requires the spins to be disaligned in couples, thus leading to a configuration in which neighbor spins alternate between up and down.

The result for negative temperatures can be easilly interpreted by looking at the definition of $\xi = \beta J$. All the result I discussed in this chapter made use of the value and

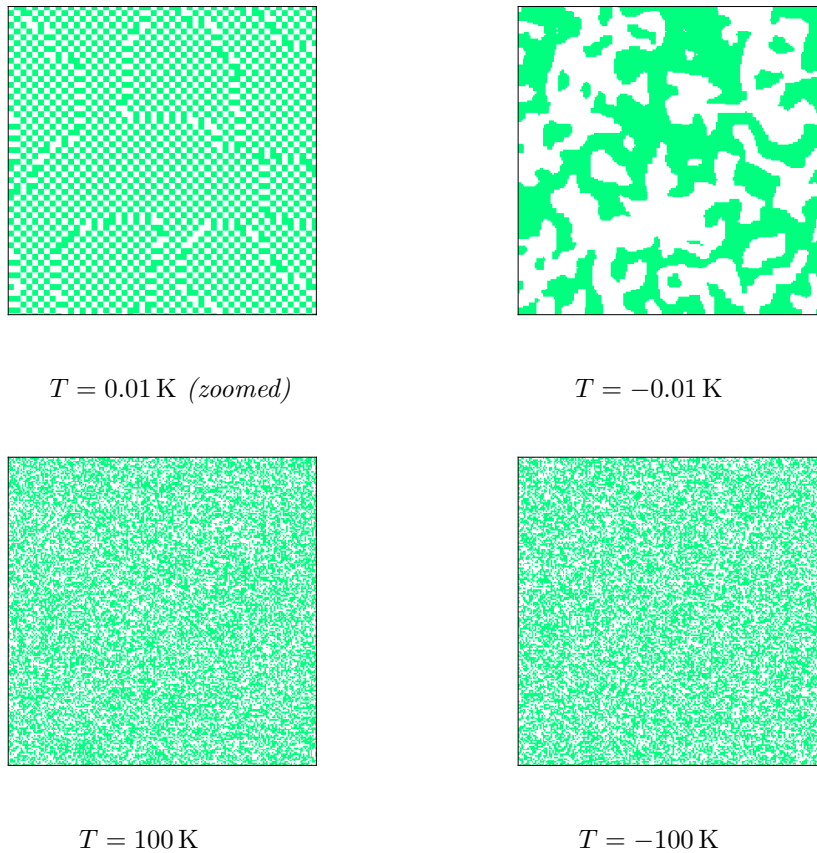


Figure 5.4: Antiferromagnetic system

the sign of ξ but no specifications were made on whether the global sign came from β or J (see sections 5.1.1 and 5.1.2). A negative sign on ξ can in fact be attributed either to a negative β and a positive J , or vice versa. For example for $J > 0$ (ferromagnetic) and temperature $T < 0$ one has that $\xi < 0$ but the sign can be "moved" to J suggesting a formal switch to the antiferromagnetic case with temperature $-T$. In formulas, for $\beta < 0$ and $J > 0$

$$0 > \xi = \beta J = -|\beta|J = |\beta|(-J)$$

Indeed it can be seen comparing figure 5.3 with 5.4 that the result for $T = 0.01 \text{ K}$ (-0.01 K) in the ferromagnetic case corresponds to the one for $T = -0.01 \text{ K}$ (0.01 K) in the antiferromagnetic case.

One can explain the similarity between the cases $T = 100 \text{ K}$ and $T = -100 \text{ K}$ by noting that the limit $T \rightarrow \pm\infty$ corresponds to $\beta \rightarrow 0$ independently of the sign of the temperature, and β is the parameter that determines the physical behaviour since it is the one that enters in the probability factor 5.2.

5.4 Two systems into contact

The second simulation consists in preparing two systems at different temperatures, one of which is negative, and putting them into contact, observing the behaviour of the two systems at equilibrium.

The physical content of the experiment is better appreciated if we return to the non-interacting case of the two levels system and we use the Hamiltonian

$$\mathcal{H}(\{\sigma_k\}) = -B \sum_i \sigma_i$$

The first system is prepared at temperature $T_1 = -0.01$ K, and since $\frac{1}{T} = \frac{\partial S}{\partial E}$, the system lies on the right edge of the plot in figure 2.1, that is the configuration with maximal energy with all the spins aligned antiparallel with respect to the field. This can also be seen by looking at the Boltzmann probability factor.

$$p(\{\sigma_k\}) = e^{-|\beta|B \sum_i \sigma_i}$$

which, for small β , is maximal when $\sum_i \sigma_i$ is minimal (most negative), that is when all the spins are aligned antiparallel with respect to the field. The second system is prepared at temperature $T = 100$ K and corresponds to a highly disordered configuration, since for $\beta \rightarrow 0$ the Boltzmann factor goes to 1 independently of the values σ_i , meaning that all the microscopic configurations are equally probable.

Each system is a 200×200 squared lattice. Again, I choose the free boundary conditions. Each system goes under 10^6 Monte Carlo steps in order to bring them in thermal equilibrium. The configurations of the systems before contact are reported in figure 5.5.

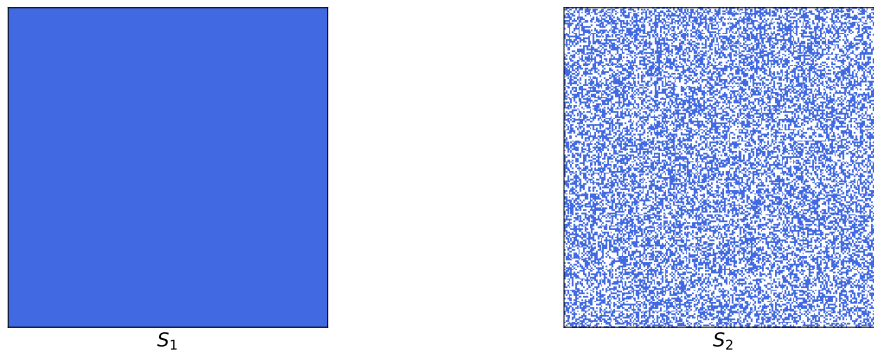


Figure 5.5: The figure reports the configurations of the systems \mathcal{S}_1 and \mathcal{S}_2 after thermalization. The systems \mathcal{S}_1 and \mathcal{S}_2 are prepared respectively at temperatures $T_1 = -0.01$ K and $T_2 = 100$ K. A blue tile indicates an up spin, while a white one indicates a down spin.

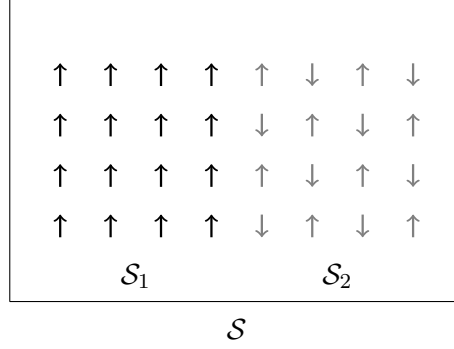


Figure 5.6: The systems \mathcal{S}_1 and \mathcal{S}_2 , prepared at temperatures T_1 and T_2 , are put into contact forming a new system \mathcal{S} of $2N \times N$ sites (the figure is shown transposed).

The systems are then brought into contact: this is done practically by creating a 400×200 lattice and copying the values of the two separated systems into the new one. The system is initialized with energy $E = E_1 + E_2$ where the two energies refers to the energies of the systems \mathcal{S}_1 and \mathcal{S}_2 before putting them into contact. The study then proceeds in the microcanonical ensemble with energy E , trying to search the configuration that maximizes the entropy given the available energy. The algorithm used for the simulation is not exactly the one introduced in section 5.2 but the so called *demon Monte Carlo*, which can be summarized in the following steps.

At first, a new degree of freedom is introduced in the system, called the demon. The demon energy must be positive but there are no restrictions on the upper value. Then

1. Propose the new state for the system as explained in section 5.2
2. Calculate the energy difference between the two states
3. If $\Delta E < 0$ the proposal is accepted and the demon takes the energy excess
4. If $\Delta E > 0$ but it is smaller than the demon's energy, the demon releases his energy and the move is accepted. Otherwise the move is rejected.
5. Repeat from point 1

The higher the number of degrees of freedom in the system, the less the demon contribution to the energy of the system, the better energy can be considered conserved in the true degrees of freedom of the system.

After 10^7 Monte Carlo cycles, the configuration is reported in figure 5.7, for convenience splitted again into the two subsystems. It can be seen that in the final state of the two systems there is a prevalence of up spins rather than downs due to energy conservation in the process of putting \mathcal{S}_1 and \mathcal{S}_2 into contact. The system \mathcal{S}_1 , initially prepared at temperature $T_1 = -0.01$ K with almost all the spins up, at equilibrium has reduced his number of up spins, still beeing in a configuration of inverted population: this means

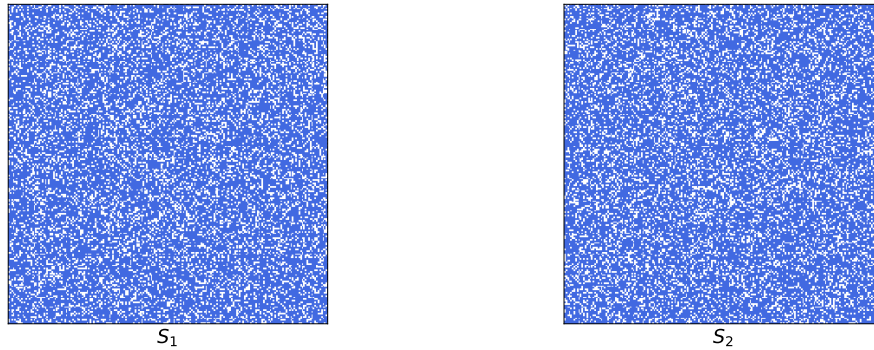


Figure 5.7: The figure reports the configurations of the systems \mathcal{S}_1 and \mathcal{S}_2 after being put into contact. After a thermalization of 10^7 Monte Carlo steps, the two systems must share the same temperature as by definition of equilibrium.

As before, a blue tile indicates an up spin, while a white one indicates a down spin.

that its temperature has decreased passing to a more negative one. The heat was transferred to the system \mathcal{S}_2 , who passed from a high positive temperature to a negative one. The total energy of the system $E = E_1 + E_2$ is then equally splitted between \mathcal{S}_1 and \mathcal{S}_2 . The situation is resumed in figure 5.8

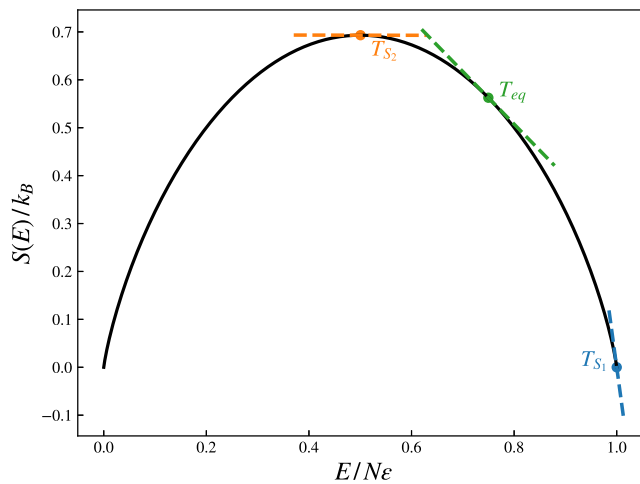


Figure 5.8: The plot resumes the configurations of the systems \mathcal{S}_1 and \mathcal{S}_2 before and after putting them into contact. The system \mathcal{S}_1 is prepared at $T = -0.01$ K in a high energy configuration (blue point in the plot), while \mathcal{S}_2 is prepared at temperature $T = 100$ K (orange point in the plot). After allowing the contact, \mathcal{S}_1 releases heat to \mathcal{S}_2 , splitting the total energy $E = E_1 + E_2$ equally between the two systems. At equilibrium \mathcal{S}_1 and \mathcal{S}_2 lie at the same temperature T_{eq} , with the same energy $E = \frac{E_1 + E_2}{2}$

Chapter 6

Conclusions

We saw that the existence of negative absolute temperatures is a direct result of the formal definition of temperature we use in thermodynamics.

In the first chapter I showed why this definition makes sense according to the general laws of thermodynamics which match our empirical observations. In particular I showed that there is a class of functions that assume the same value for systems that are in equilibrium with each other, and different values for systems that are not at reciprocal equilibrium. Any of these functions is a temperature.

By a combinatorial approach we saw that systems at equilibrium share the same value of the function $\frac{\partial S_i}{\partial E_i}$ which can then be considered a temperature.

The two levels system was then introduced showing some of the characteristic behaviours of the negative absolute temperatures as that they are hotter than any positive ones and that they describe systems with an inverted population state.

A general set of criteria to establish whether a system admits negative temperatures or not was then introduced, namely the Ramsey's criteria.

After that I discussed the first experimental observation of negative temperatures carried by Purcell and Pound, who observed a LiF spin lattice in an inverted population state by rapidly inverting a strong magnetic field.

Since the definition of temperature is a direct consequence of the definition of entropy, chapter 4 was devoted to the discussion of the historical issues and debates on the choice of the correct definition. Boltzmann's entropy was initially thought to be missing a thermodynamic condition, but was showed later to match it in the limit of a high number of degrees of freedom. Gibbs' entropy was instead showed to fail a basic condition in a specific example, that is to assume the same value for systems at equilibrium, and different values for systems out of equilibrium.

The last chapter was devoted to two simulations. The numerical analysis of the Ising model showed that the behaviour of a ferromagnetic system at negative temperatures is the same of an antiferromagnetic system at positive temperatures and viceversa. Instead the numerical experiment in which two system were put into contact, one of which at negative temperature, showed that the system at negative temperature released heat to the one at positive temperature.

One may wonder if negative temperatures have a direct application in everyday life: this is indeed the case. For example it was proved that Carnot machines working at negative temperature can obtain an efficiency equal to unity, giving an intuition on how more complex machines could produce more work when allowed to work at negative temperature.

Probably the most relevant application of negative temperatures is in the field of cosmology: negative absolute temperatures were found to be linked to negative pressures, fundamental to describe the accelerating expansion of the universe.

Bibliography

- [1] A. Abragam and W. G. Proctor. “Spin Temperature”. In: *Phys. Rev.* 109 (5 1958), pp. 1441–1458. DOI: [10.1103/PhysRev.109.1441](https://doi.org/10.1103/PhysRev.109.1441). URL: <https://link.aps.org/doi/10.1103/PhysRev.109.1441>.
- [2] Martin J. Klein. “Negative Absolute Temperatures”. In: *Phys. Rev.* 104 (3 1956), pp. 589–589. DOI: [10.1103/PhysRev.104.589](https://doi.org/10.1103/PhysRev.104.589). URL: <https://link.aps.org/doi/10.1103/PhysRev.104.589>.
- [3] Lars Onsager. “The Effects of Shape on the Interaction of Colloidal Particles”. In: *Annals of the New York Academy of Sciences* 51.4 (May 1949), pp. 627–659. DOI: [10.1111/j.1749-6632.1949.tb27296.x](https://doi.org/10.1111/j.1749-6632.1949.tb27296.x).
- [4] E. M. Purcell and R. V. Pound. “A Nuclear Spin System at Negative Temperature”. In: *Phys. Rev.* 81 (2 1951), pp. 279–280. DOI: [10.1103/PhysRev.81.279](https://doi.org/10.1103/PhysRev.81.279). URL: <https://link.aps.org/doi/10.1103/PhysRev.81.279>.
- [5] Eitan Abraham and Oliver Penrose. “Physics of negative absolute temperatures”. In: *Physical Review E* 95 (Jan. 2017). DOI: [10.1103/PhysRevE.95.012125](https://doi.org/10.1103/PhysRevE.95.012125).
- [6] Josiah Willard Gibbs. *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*. Cambridge Library Collection - Mathematics. Cambridge University Press, 2010. DOI: [10.1017/CB09780511686948](https://doi.org/10.1017/CB09780511686948).
- [7] V. Berdichevsky, I. Kunit, and F. Hussain. “Negative temperature of vortex motion”. In: *Phys. Rev. A* 43 (4 1991), pp. 2050–2051. DOI: [10.1103/PhysRevA.43.2050](https://doi.org/10.1103/PhysRevA.43.2050). URL: <https://link.aps.org/doi/10.1103/PhysRevA.43.2050>.
- [8] Jörn Dunkel and Stefan Hilbert. “Consistent thermostats forbids negative absolute temperatures”. In: *Nature Physics* 10.1 (Jan. 2014), pp. 67–72. DOI: [10.1038/nphys2815](https://doi.org/10.1038/nphys2815). arXiv: [1304.2066 \[cond-mat.stat-mech\]](https://arxiv.org/abs/1304.2066).
- [9] Daan Frenkel and Patrick B. Warren. “Gibbs, Boltzmann, and negative temperatures”. In: *American Journal of Physics* 83.2 (2015), 163–170. ISSN: 1943-2909. DOI: [10.1119/1.4895828](https://doi.org/10.1119/1.4895828). URL: <http://dx.doi.org/10.1119/1.4895828>.
- [10] Ernst Ising. “Beitrag zur theorie des ferromagnetismus”. In: *Zeitschrift für Physik* 31.1 (1925), pp. 253–258.

- [11] W. K. Hastings. “Monte Carlo sampling methods using Markov chains and their applications”. In: *Biometrika* 57.1 (Apr. 1970), pp. 97–109. ISSN: 0006-3444. DOI: [10.1093/biomet/57.1.97](https://doi.org/10.1093/biomet/57.1.97). eprint: <https://academic.oup.com/biomet/article-pdf/57/1/97/23940249/57-1-97.pdf>. URL: <https://doi.org/10.1093/biomet/57.1.97>.
- [12] Jos Uffink. *Compendium of the foundations of classical statistical physics*. Chapter for “Handbook for Philosophy of Physics”, J. Butterfield and J. Earman (eds) to appear. 2006. URL: <http://philsci-archive.pitt.edu/2691/>.
- [13] Kerson Huang. *Statistical Mechanics*. 2nd ed. John Wiley & Sons, 1987.
- [14] Mehran Kardar. *Statistical Physics of Particles*. Cambridge University Press, 2007. DOI: [10.1017/CB09780511815898](https://doi.org/10.1017/CB09780511815898).
- [15] Loris Ferrari. *Elementi di Termodinamica Statistica*.