Advanced Statistical Physics: Introduction

Leticia F. Cugliandolo
`leticia@lpthe.jussieu.fr`
Sorbonne Université
Laboratoire de Physique Théorique et Hautes Energies

September 9, 2018

Contents

1 Basic notions

1.1 Background

Equilibrium Statistical Mechanics is a very well-established branch of theoretical physics. Together with *Quantum Mechanics*, they form the basis of *Modern Physics*. It is based on *probability theory and mathematical statistics* on the one hand and Hamiltonian mechanics on the other. Its scope is very wide since it is, in principle, applicable to all states of matter. A nice introduction to the historical development of this theory can be found in the classical book of Pathria [1].

The goal of equilibrium statistical mechanics is to derive the *thermodynamic functions of state* of a *macroscopic* system from the *microscopic laws* that determine the behaviour of its constituents. In particular, it explains the origin of thermodynamic – and intuitive – concepts like pressure, temperature, heat, *etc.*

In Table 1 we recall the typical length, time and energy scales appearing in the microscopic (say, atomistic) and macroscopic World.

	Micro		Macro	
	Solid	Gas	Solid	Gas
dist (ℓ)	10^{-10}m	10^{-8}m	10^{-3}m	
# part (N)	1		$\left(\frac{10^{-3}}{10^{-10}}\right)^{d=3} = 10^{21}$	$\left(\frac{10^{-3}}{10^{-8}}\right)^{d=3} = 10^{15}$
energy (E)	$1\, eV$		$1J \approx 6\, 10^{18} eV$	
time (t)	$\hbar/1eV \approx 6\ 10^{-14}\ s$	$10^{-9}\ s$	$1\, s$	

Table 1: Typical length, energy and time scales in the microscopic and macroscopic World.

A reference number is the number of Avogadro, $N_A = 6.02\ 10^{23}$; it counts the number of atoms in a mol, *i.e.* 12gr of ^{12}C, and it yields the order of magnitude of the number of molecules at a macroscopic level. The ionization energy of the Hydrogen atom is 13.6 eV and sets the microscopic energy scale in Table 1.

It is clear from the figures in Table 1 that, from a practical point of view, it would be impossible to solve the equations of motion for each one of the $N \approx N_A$ particles – we keep the discussion *classical*, including quantum mechanical effects would not change the main conclusions to be drawn henceforth – and derive from their solution the macroscopic behaviour of the system. Moreover, the deterministic equations of motion may present a very high sensitivity to the choice of the initial conditions – *deterministic chaos* – and thus the precise prediction of the evolution of the ensemble of microscopic constituents becomes unfeasible even from a more fundamental point of view.

The passage from the microscopic to the macroscopic is then done with the help of *Statistical Methods*, *Probability Theory* and, in particular, the *Law of Large Numbers*. It assumes – and it has been very well confirmed – that there are no big changes in the fundamental *Laws of Nature* when going through all these orders of magnitude. However, a number of new and interesting phenomena arise due to the unexpected *collective behaviour* of these many degrees of freedom. For example, *phase transitions* when varying an external parameter occur; these are not due to a change in the form of the microscopic interactions but, rather, to the locking of the full system in special configurations.

In the theoretical derivation of phase transitions the necessity to take the infinite size clearly appears; this is called the *thermodynamic limit* and corresponds to taking the number of degrees of freedom, say N, to infinity and the volume, say V, also to infinity, while keeping the ratio between the two finite, N/V finite. In this limit, *extensive* quantities scale with N or V while *intensive* ones remain finite.

The main features of statistical mechanics are quite independent of the mechanics that describes the motion of the individual agents (classical or quantum mechanics, relativistic or not). Its foundations do need though different reasonings in different cases. For the sake of concreteness in this set of lectures we shall focus on *classical non-relativistic systems*.

In Table I we mentioned energy scales and length scales typical of atomic physics. Particle physics involves even shorter length scales and higher energy scales. Particle physics is studied using *field theories* (see the Lectures by J-B Fournier). While standard textbooks in Statistical Mechanics do not use a field theoretical formulation, it is indeed pretty straightforward to apply Statistical Mechanics notions to field theories – a theory with 'infinite' degrees of freedom, one for each space-point.

Equilibrium statistical mechanics also makes another very important assumption that we shall explain in more detail below. Namely, that of the equilibration of the macroscopic system. Some very interesting systems do not match this hypothesis, either because the *time scales* over which they evolve are too short to reach equilibrium or because the microscopic dynamics is *not Hamiltonian*. Still, one would like to use Probabilistic arguments to characterise the macroscopic behaviour of out of equilibrium systems as well. This is possible in a number of cases although we will not discuss them in these Lectures (see the Lectures by D. Mouhanna & F. van Wijland). Indeed, deriving a theoretical framework to describe the behaviour of *macroscopic systems out of equilibrium* is one the present major challenges in theoretical physics.

Analytic solutions of macroscopic systems (be them in equilibrium or not) are limited to rather simple cases (typically in very low dimension or in the mean-field limit. *Numerical methods*, viz. computer simulations, are very useful to treat complex systems. These are discussed in the Lectures by F. Krzakala.

1.2 Elements in statistical mechanics

Let us recall here some important features of Statistical Mechanics [4, 5, 6, 7, 8]. A short account of classical mechanics is given in [9] with special emphasis on the integrability properties that we will investigate in Sec. 3. A specially careful discussion of the passage from classical mechanics to statistical physics is given in [10, 11].

The state of a classical system with $i = 1, \dots, N$ particles moving in d-dimensional *real space* is fully characterised by a point in the $2dN$ dimensional *phase space* Γ. The coordinates of phase space are the real space coordinates of the particles, q_i^a, where i is the particle label and $a = 1, \dots, d$ is the label of the real space coordinates, and the particles' momenta, p_i^a. It is convenient to represent a point in phase space with a $2dN$-dimensional vector, *e.g.* $\vec{Y} = (\vec{Q}, \vec{P}) = (q_1^1, q_1^2, q_1^3, q_2^1, q_2^2, q_2^3, \dots, q_N^1, q_N^2, q_N^3, p_1^1, p_1^2, p_1^3, p_2^1, p_2^2, p_2^3, \dots, p_N^1, p_N^2, p_N^3)$ in $d = 3$. This problem has dN *degrees of freedom* and the phase space has *dimension* $2dN$.

The *Hamiltonian* of the system, H, is a function of the particles' position and momenta (we restrict the discussion to Hamiltonian systems). It can be explicitly time-dependent but we shall not consider these cases here. The particles' time evolution, $(\vec{Q}, \vec{P})(t)$, starting from a given initial condition, $(\vec{Q}, \vec{P})(t = 0)$, is determined by Hamilton's equation of motion that are equivalent to *Newton dynamics*. As time passes the representative point in phase space, $(\vec{Q}, \vec{P})(t)$, traces a (one dimensional) path in Γ, the *phase-space trajectory*. Through each point in phase space there passes one and only one trajectory. Any function of the $(\vec{Q}, \vec{P})(t)$ is called a *phase function*. The most important one in the Hamiltonian itself, that determines the evolution of the system *via* the equations of motion. The *energy*, E, or the value the Hamiltonian takes on any point of the trajectory, is conserved if the Hamiltonian does not depend on time explicitly and thus all points in any trajectory lie on a constant energy surface, $H(\vec{Q}, \vec{P}) = E$.

But, can one really describe the evolution of such a system? In practice, one cannot determine the position and momenta of all particles in a macroscopic system with $N \gg 1$ with great precision – due to uncertainty in the initial conditions, deterministic chaos, *etc.* A probabilistic element enters into play. What one really does is to estimate the probability that the representative point of the system lies in a given region of Γ at time t given that it started in some other region of Γ at the initial time. Thus, one introduces a *time-dependent probability density* $\rho(\vec{Q}, \vec{P}; t)$ such that $\rho(\vec{Q}, \vec{P}; t) d\Gamma$ is the probability that the representative point is in a region of volume $d\Gamma$ around the point $(\vec{Q}, \vec{P})$ at time t knowing the probability density of the initial condition, $\rho(\vec{Q}, \vec{P}; t = 0)$. Conditions on ρ to make it a probability density, $\rho(\vec{Q}, \vec{P}; t) \geq 0$ for all $(\vec{Q}, \vec{P})$ and t, and $\int_\Gamma d\Gamma \, \rho(\vec{Q}, \vec{P}; t) = 1$ at all times, need to be satisfied.

Note that if initially one knows the state of the system with great precision, the initial ρ will be concentrated in some tiny region of phase space. At later times, ρ can still be *localised* – perhaps in a different region of phase – or it may *spread*. This depends on the system and the dynamics. If the dynamics is of Newton-Hamilton tyoe, the phase space volume can change form but not volume, while if the dynamics are dissipative or if energy is injected in some way, the scape phase region occupied by the system can also change volume.

We now need to find an equation for the evolution of the probability density ρ knowing the evolution of the phase space coordinates $(\vec{Q}, \vec{P})$. The key point is that ρ behaves as the *density of a fluid*, it can change in time, decreasing somewhere and increasing somewhere else but this can only do continuously, with the density *flowing* from one region to another. The consequence is that ρ must satisfy a *continuity equation*. Analogously to the global conservation of the mass in a fluid, ρ conserves the probability.

ρ can vary in time due to two mechanisms; an explicit time variation, and the time variation of the coordinates and momenta as the representative point wanders in phase space:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial q_i^a}\dot{q}_i^a + \frac{\partial \rho}{\partial p_i^a}\dot{p}_i^a \; , \tag{1.1}$$

with $\dot{q}_i^a = dq_i^a/dt$ and $\dot{p}_i^a = dp_i^a/dt$, the summation convention over repeated indices (i labels particles and a labels coordinates), and $\rho(\vec{Q}, \vec{P}, 0)$ known. For Hamiltonian dynamics these variations are captured or represented by the dynamic equation

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \{\rho, H\} \tag{1.2}$$

with $\{f, g\}$ the Poisson bracket,

$$\{f, g\} \equiv \frac{\partial f}{\partial q_i^a}\frac{\partial g}{\partial p_i^a} - \frac{\partial f}{\partial p_i^a}\frac{\partial g}{\partial q_i^a} \; . \tag{1.3}$$

(See the attached handwritten notes for the derivation of eq. (1.2).)

Probability behaves like an *incompressible fluid in phase space* (see the attached handwritten notes for the proof) and one can then use knowledge of fluid mechanics to analyze the equation above. The partial derivative of ρ is taken at fixed $(\vec{Q}, \vec{P})$ and represents the time-variation of ρ as the fluid passes by the chosen point in phase space. The total derivative in the left-hand-side is, instead, the time-variation as we follow the displacement of a 'piece' of fluid in phase space.

Liouville's theorem states that the ensemble of systems (as represented by a point in phase space) in the vicinity of a given system remains constant in time:

$$\frac{d\rho}{dt} = 0 \; . \tag{1.4}$$

A detailed description of Liouville's theorem is given in [10, 12] (see handwritten notes).

In statistical equilibrium one expects the systems to reach stationarity and then the explicit time-variation to vanish

$$\frac{\partial \rho}{\partial t} = 0 \; . \tag{1.5}$$

In this case, the distribution ρ is constant on the phase trajectories. One may wonder whether this solution is reached from generic initial conditions.

Liouville's equation (1.1) is *invariant under time-reversal*, $t \to -t$ and $\vec{p} \to -\vec{p}$. Indeed, the existence of a conserved current implies, via *Noether's theorem*, the existence of a symmetry. The symmetry is invariance under time translations, and the generator of the symmetry (or Noether charge) is the Hamiltonian.

The time-reversal symmetry implies that, for generic initial conditions the solutions to Liouville's equation oscillate in time and do not approach a single asymptotic stationary solution that could be identified with equilibrium ($\partial_t \rho = 0$). The problem of how to obtain irreversible decay from Liouville's equation is a fundamental one in Statistical Mechanics. We shall not deepen this discussion here; let us just mention that the main attempt to understand the origin of irreversibility is in terms of flows in phase space, and this approach is called *ergodic theory*, as founded by Boltzmann by the end of the XIXth century [6].

In the absence of a good way to determine the evolution of ρ and its approach to a stationary state, one simply looks for *stationary solutions* without worrying about how the ensemble reaches them. This can be restated as the following hypothesis:

> As $t \to \infty$ one expects that the statistical properties of the system be independent of time and hence $\rho(\vec{Q}, \vec{P}; t) \to \rho(\vec{Q}, \vec{P})$.

Setting now $\partial_t \rho = 0$ one realizes that the remaining equation admits, as a solution, any function of the coordinate and momenta only *via* the Hamiltonian H. The characteristics of the ensemble are then determined by the chosen function $\rho(H)$.

1.3 The ergodic hypothesis & Gibbs ensembles

Finally, let us discuss Boltzmann's and Gibb's interpretation of averages and the *ergodic hypothesis*. Boltzmann interpreted macroscopic observations as time averages of the form[1]

$$\overline{A} \equiv \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau dt \; A(\vec{Q}(t), \vec{P}(t)) \; . \tag{1.6}$$

The fact that this limit exists is the content of a Theorem in Classical Mechanics initially proven by Birkhoff and later by Kolmogorov [10]. Note that in classical mechanics the choice of the initial time is irrelevant.

With the introduction of the *concept of ensembles* Gibbs gave a different interpretation (and an actual way of computing) macroscopic observations. For Gibbs, these averages

---

[1] In practice, in and experiment or numerical simulation initiated at time $t = 0$, averages are computed over a symmetric time interval around a measuring time t, in the form $\overline{A} \equiv \lim_{t_0 \ll \tau \leq t} \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} dt' \; A(\vec{Q}(t'), \vec{P}(t'))$ with the lower bound in the limit representing a microscopic time-scale. The result should be independent of the measuring time t.

are statistical ones over all elements of the statistical ensemble,

$$\langle A \rangle = c \int \prod_{i=1}^{N} \prod_{a=1}^{d} dq_i^a dp_i^a \; \rho(\vec{Q},\vec{P}) A(\vec{Q},\vec{P}) \;, \tag{1.7}$$

with ρ the measure. In the microcanonical ensemble this is an average over micro-states on the constant energy surface taken with the microcanonical distribution (2.10):

$$\langle A \rangle = c \int \prod_{i=1}^{N} \prod_{a=1}^{d} dq_i^a dp_i^a \; \delta(H(\vec{Q},\vec{P}) - E) A(\vec{Q},\vec{P}) \;, \tag{1.8}$$

and the normalization constant $c^{-1} = \int \prod_{i=1}^{N} \prod_{a=1}^{d} \delta(H(\vec{Q},\vec{P}) - E)$. In the canonical ensemble the average is computed with the Gibbs-Boltzmann weight:

$$\langle A \rangle = Z^{-1} \int \prod_{i=1}^{N} \prod_{a=1}^{d} dq_i^a dp_i^a \; e^{-\beta H(\vec{Q},\vec{P})} A(\vec{Q},\vec{P}) \;. \tag{1.9}$$

Z is the partition function $Z = \int \prod_{i=1}^{N} \prod_{a=1}^{d} dq_i^a dp_i^a \; e^{-\beta H(\vec{Q},\vec{P})}$.

The *(weak) ergodic hypothesis* states that under the dynamic evolution the representative point in phase space of a classical system governed by Newton laws can get as close as desired to any point on the constant energy surface.

The *ergodic hypothesis* states that time and ensemble averages, (1.6) and (1.7) coincide in equilibrium for all reasonable observables. This hypothesis cannot be proven in general but it has been verified in a large number of cases. In general, the great success of Statistical Mechanics in predicting quantitative results has given enough evidence to accept this hypothesis.

An important activity in modern Statistical Mechanics is devoted to the study of macroscopic systems that do not satisfy the ergodic hypothesis. A well-understood case is the one of phase transitions and we shall discuss it in the next section. Other cases are related to the breakdown of equilibration. This can occur either because they are externally driven or because they start from an initial condition that is far from equilibrium and their interactions are such that they do not manage to equilibrate. One may wonder whether certain concepts of thermodynamics and equilibrium statistical mechanics can still be applied to the latter problems. At least for cases in which the macroscopic dynamics are slow one can hope to derive an extension of equilibrium statistical mechanics concepts to describe their behaviour.

Finally, let us remark that it is usually much easier to work in the canonical ensemble both experimentally and analytically. Thus, in all our future applications we assume that the system is in contact with a heat reservoir with which it can exchange energy and that keeps temperature fixed.

2 The statistical (Gibbs) ensembles

In this Section we discuss some aspects of the standard Statistical Physics approach to equilibrium systems that are often overlooked in the elementary courses on Statistical Mechanics. They concern the differences encountered in the description of systems with short and long range interactions, as we define below.

Gibbs introduced the notion of statistical ensembles and, usually, equivalence between the thermodynamic descriptions derived from all of them is assumed (in the thermodynamic limit). We discuss here cases in which there can be *inequivalence of ensembles* even when the number of degrees of freedom diverges [16, 17, 18].

To start with, we recall the three ensembles used. These are:

- The microcanonical ensemble describes a completely isolated system at a given energy E.
- The canonical ensemble describes a system that can exchange energy with a large thermal reservoir characterised by its temperature T.
- The grand canonical ensemble describes a system that can exchange energy and particles with a reservoir withs temperature T and chemical potential μ.

As preliminary knowledge, let us define a *microstate* as a precise microscopic state (for example, a given $(\vec{Q}, \vec{P})$). A *macrostate* is defined in terms of a small set of macroscopic or coarse-grained variables and it refers to the (large) set of microstates that yield the same value of the macroscopic variables (for example, all the configurations that share the same kinetic and potential energy values in a system of interacting particles).

2.1 The microcanonical ensemble

In the *microcanonical ensemble* one makes the following hypothesis:

In the long-time limit the system does not prefer any special region on the constant energy surface in Γ – there is *a priori* no reason why some region in Γ should be more probable than others! – and thus $\rho(\vec{Q}, \vec{P})$ is expected to be a constant on the energy surface and zero elsewhere:

$$\rho(\vec{Q}, \vec{P}) = \begin{cases} \rho_0 & \text{if } H(\vec{Q}, \vec{P}) \in (E, E + dE) \, , \\ 0 & \text{otherwise} \, , \end{cases} \tag{2.10}$$

The constant ρ_0 is the inverse of the volume of the constant energy surface ensuring normalisation of ρ. This is indeed the simplest stationary solution to eq. (1.1).

These hypotheses can be valid only if the long-time dynamics is reasonably independent of the initial conditions.

Even though it is very difficult to show, the solution proposed above is very smooth as a function of $(\vec{Q},\vec{P})$ and it is then the best candidate to describe the *equilibrium* state – understood as the one that corresponds to the intuitive knowledge of equilibrium in thermodynamics.

This description corresponds to the *microcanonical ensemble* of statistical mechanics, valid for closed systems with fixed energy E and volume V. Each configuration on the constant energy surface is called a *microstate*. In the microcanonical ensemble all *microstates* are equivalent. We can think about all these microstates as being (many) independent *copies* of the original system. This is Gibbs' point of view: he introduced the notion of *ensemble* as the collection of mental copies of a system in identical macroscopic conditions.

The average of any phase space function $A(\vec{Q},\vec{P})$ can now be computed as

$$\begin{aligned}\langle\, A\,\rangle &= \int \prod_{i=1}^{N}\prod_{a=1}^{d} dq_i^a dp_i^a\ \rho(\vec{Q},\vec{P})A(\vec{Q},\vec{P}) \\ &= \left(\frac{1}{N!g(E)}\right)\int \prod_{i=1}^{N}\prod_{a=1}^{d} dq_i^a dp_i^a\ \delta[E-H(\vec{Q},\vec{P})]\, A(\vec{Q},\vec{P})\ . \end{aligned} \tag{2.11}$$

The normalisation constant $c = \int \prod_{i=1}^{N}\prod_{a=1}^{d} dq_i^a dp_i^a\ \delta[E-H(\vec{Q},\vec{P})] = N!g(E)$ is the volume of phase space occupied by the constant energy surface itself. The quantity $g(E)$ is called the *density of states*:

$$g(E) \equiv \frac{1}{N!}\int \prod_{i=1}^{N}\prod_{a=1}^{d} dq_i^a dp_i^a\ \delta[E-H(\vec{Q},\vec{P})]\ . \tag{2.12}$$

The *microcanonical entropy* is

$$S(E) \equiv k_B \ln g(E)\ . \tag{2.13}$$

The maximisation of the entropy is thus equivalent to the maximisation of the phase volume available to the system.

The temperature is defined, in the microcanonical ensemble as

$$\frac{1}{T(E)} = \frac{\partial S(E)}{\partial E}\ . \tag{2.14}$$

A positive temperature needs, therefore, that $S(E)$ be an increasing function of E.

Note that the nature of the interactions between the constituents has not been mentioned in this discussion. There is no reason to believe that the microcanonical description would fail for some type of interaction as the gravitational one, a case that we shall discuss in detail below.

2.2 Extensivity and additivity: short and long range interactions

Once the microcanonical ensemble has been established one usually goes further and derives what is called the canonical ensemble describing the statistical properties of a system that can exchange energy with its surrounding. When trying to derive the canonical description from the microcanonical one, the distinction between systems with short and long interactions becomes important. We discuss here two properties that allow one to differentiate these two classes.

First of all, we have to determine how do the constituents of the system interact, that is to say, which are the forces acting between them. We consider systems where the interaction potential is given by the sum, over pairs of the elementary constituents, of a *two-body translationally invariant potential*.

A system of N particles confined inside a volume V is said to be *extensive* if, when the number of particles and the volume are scaled by λ, the internal energy $E(\lambda N, \lambda V)$ of the system scales as $\lambda E(N, V)$.

It is easy to see that systems with short-range forces are extensive[2].

Let us consider a macroscopic system with volume V and divide it in two pieces with volumes V_1 and V_2, with $V = V_1 + V_2$. The aim is to characterise the statistical properties of one subsystem (say 1) taking into account the effect of its interaction with the rest of the macroscopic system (subsystem 2). It is clear that the energy of the subsystems is not fixed since these are not closed: they interact with each other. The total energy, E, is then the sum of the energies of the two sub-ensembles plus the interaction energy between the two pieces, $E = E_1 + E_2 + E_I$.

If the interactions between the constituents of the system are *short-ranged*, in the sense that the interaction energy is proportional to the surface between the two pieces, $E_I \propto S$, while the energy of each subsystem will be extensive and proportional to their volumes, $E_1 \propto V_1$ and $E_2 \propto V_2$. Thus, for a macroscopic system, E_I is negligible with respect to $E_1 + E_2$.

If, instead, the interactions between the constituents are sufficiently *long-ranged* the separation into volume and surface contributions to the total energy does not apply any longer. This remark allows one to define the following property.

---

[2] If the interaction potential is short-range, each particle will interact only with the particles which are within the range γ of the interaction potential. Suppose that a system is homogeneous, the number of particles within the distance γ of a given particle will then be proportional to $N\gamma^d/V$ and the internal energy must have the form of $E(N, V) = Nf(N/V)$, where $f(x)$ is a function that depends on the microscopic interactions between the particles. Actually, it is not necessary for the interaction potential to be bounded by γ; algebraically decaying potentials will lead to extensive thermodynamics as long as they decay sufficiently rapidly, i.e. if $\alpha > d$.

The *additivity property*, states that

$$E = E_1 + E_2 \; , \tag{2.15}$$

for any two subsystems 1 and 2 of a macroscopic system. An alternative definition of systems with *short range interactions* applies to those for which the additivity property holds, and systems with *long range interactions* as those for which this property fails.

In some cases, the energy of a system with long-range interactions can be rendered extensive (instead of super-extensive) by properly scaling the interaction parameters with the number of degrees of freedom.

One such example is the case of an *Ising spin model with fully-connected ferromagnetic interactions* (the Curie-Weiss model): $H = -J/2 \sum_{i\neq j} s_i s_j$ with $s_i = \pm 1$ and $J > 0$. The sum runs over *all pairs* of spins in the ensemble with the constraint $i \neq j$. The factor 1/2 ensures that each pair of spins contributes only once to the sum. Clearly, the ferromagnetic configurations have a super extensive energy, since $H(s_i = 1) = -J/2\ N(N-1) \to -J/2\ N^2$ for $N \gg 1$. This problem can be corrected by scaling $J \mapsto J/N$, in which case $H(s_i = 1) = -J/2\,(N-1)$ and energy extensivity is restored. However, the energy remains non-additive since all spins still interact with all other ones in the sample and there is no notion of interface between two subsystems. Indeed, let us explore the additivity property of the energy for the perfectly magnetised state. The total energy is $E = -J(N-1)$. If we now divide the system in two subsystems with $N/2$ spins each the total energy of each subsystem is $E_1 = -J(N/2-1)/4$ and $E_2 = -J(N/2-1)/4$ and one notices that $E \neq E_1 + E_2$. More precisely, $E_I \equiv E - (E_1 + E_2) = -JN/4$, still a macroscopic quantity. One has

$$E \sim E_1 \sim E_I \sim -JN \tag{2.16}$$

all these energies are of the order of the number of spins in the sample. (In contrast, in the usual Ising model defined on a d dimensional lattice with nearest-neighbour interactions the additivity properties holds.)

Another example is the *self-gravitating gas in a low dimensional space*. As the interaction is attractive, one and two dimensional gravitational systems are self-confining and one can safely take the infinite volume limit $V \to \infty$ and focus on the scaling of the energy with the number of particles. In order to avoid the non-extensivity of the energy, the gravitational constant can be rescaled by a factor $1/N$, the so-called *Kac prescription*. However, the energy remains non-additive after this rescaling.

Note that that non-additivity also occurs in systems with short-range interactions in which surface and bulk energies are comparable; this is realised in *finite size* problems.

2.3 The canonical ensemble

Let us review the derivation of the canonical distribution. Consider a system with volume V, divide it in two pieces with volumes V_1 and V_2, with $V = V_1 + V_2$, and energies E_1 and E_2. If we assume that the two systems are *independent* with the constraint that $E_1 + E_2 = E$, *i.e.* the *additivity property*, the probability of subsystem 1 to get an energy E_1 is

$$
\begin{aligned}
P(E_1)\, dE_1 &\propto \int dE_2\; g(E_1, E_2)\; \delta(E - E_1 - E_2)\, dE_1 \\
&= \int dE_2\; g_1(E_1) g_2(E_2)\; \delta(E - E_1 - E_2)\, dE_1 \\
&= g_1(E_1) g_2(E - E_1)\, dE_1 \\
&= g_1(E_1) e^{k_B^{-1} S_2(E-E_1)}\; dE_1 \\
&\simeq g_1(E_1)\; e^{k_B^{-1} S_2(E) + k_B^{-1} \partial_E S_2(E)(-E_1)}\; dE_1 \\
&\propto g(E_1) e^{-\beta_2 E_1}\; dE_1 \qquad (2.17)
\end{aligned}
$$

where we used the definition of the microcanonic inverse temperature of the second component, $\beta_2 = 1/(k_B T_2)$. Note that in the next to last passage we dropped all higher order terms in the Taylor expansion assuming that $E_1 \ll E$. After fixing the normalisation, and calling $\beta_2 \to \beta$,

$$
P(E_1) = Z^{-1}(\beta)\; g(E_1) e^{-\beta E_1} \qquad \text{with} \qquad Z(\beta) = \int dE_1\; g(E_1) e^{-\beta E_1}\; . \qquad (2.18)
$$

Let us recap the assumptions made: (i) energy additivity $E_2 = E - E_1$, (ii) independence, $g(E_1, E_2) = g(E_1) g(E_2)$, (iii) small system 1 ($E_1 \ll E$), (iv) constant inverse 'temperature' $k_B \beta \equiv \partial_E S_2(E)$. Note that assumptions (i) and (ii) fail in systems with long-range interactions. In these cases the microcanonical ensemble is well-defined though difficult to use, and the canonical one is not even defined!

First example: *the power-law potential*

In the field of particle systems with two-body interactions falling-off with distance as a power law

$$V(r) \sim r^{-\alpha} \qquad (2.19)$$

one finds that the interactions are

- long-ranged if $\alpha < d$,
- quasi long-ranged if $\alpha = d$,

- short-ranged if $\alpha > d$,

with d the dimension of space. A simple way of showing this general result is the following. Take a point particle with unit mass and place it at the origin of coordinates. Consider its interaction with a homogeneous massive spherical shell with internal radius ϵ and external radius R. The energy, e, felt by the central particle is

$$e = -\int_V d^d x \, \frac{J\rho}{r^\alpha} = -\Omega_d \, J\rho \int_\epsilon^R dr \, \frac{r^{d-1}}{r^\alpha} = -\frac{\Omega_d \, J\rho}{d-\alpha} \left[R^{d-\alpha} - \epsilon^{d-\alpha} \right] \, , \tag{2.20}$$

where we adopted the potential $V(r) = Jr^{-\alpha}$ for all r, and the particle density is ρ. Ω_d is the angular volume, $\Omega_d = 2\pi$ in $d = 2$, $\Omega_d = 4\pi$ in $d = 3$, etc. For $\alpha > d$ the contribution from the external surface ($r = R$) is negligible while for $\alpha \leq d$ it grows with the volume $V \propto R^d$ as $V^{1-\alpha/d}$. In the latter case surface effects cannot be neglected and the total energy $E = Ve$, is super-linear with the volume:

$$E \simeq JVR^{d-\alpha} \simeq JR^{2d-\alpha} = JR^{d(2-\alpha/d)} = JV^{1+1-\alpha/d} \tag{2.21}$$

for $1 - \alpha/d > 0$, or is *super extensive*. This problem can be solved by scaling the interactions analogously to what was done in the fully-connected spin model (defined on the complete graph), *i.e.* one can redefine the coupling constant $J \mapsto J \; V^{\alpha/d-1}$ and get an extensive system, $E \propto V$. However, the lack of additivity will not be resolved, as there is no clear distinction between *bulk* and *surface* in models with long-range interactions. This type of power-law decaying potential is also sometimes referred to as being *non-integrable*.

This definition implies then that the *gravitational interaction*, $V_G(r) = -Gm^2r^{-1}$ is long-ranged in three spatial dimensions while the Van der Waals interaction, $V_{VW}(r) \propto r^{-6}$, is short-ranged. The long-ranged interactions are sometimes called *non-integrable* in the literature. *Plasma physics* also provides examples of non-additive systems through an effective description.

The failure of energy additivity is at the origin of the unusual equilibrium and dynamic behaviour of systems with long-range interactions. Surprisingly enough, one finds that many usual thermodynamic results are modified with, for example, systems having negative microcanonical specific heat; moreover, the statistical ensembles (microcanonic, canonic and macrocanonic) are no longer equivalent, as we saw above with the failure of the derivation of the canonical ensemble from the microcanonical.

The statistical physics of self-gravitating systems falls into this class of *bizarre* problems and there is much current research [13, 14, 15] to try to elucidate their properties.

The canonical ensemble does not describe the fluctuations of a small subsystem of a system with constituents interacting via long-range interactions. One can, however, argue that it will describe the fluctuations of a system with long-range internal interactions connected, *via* short-range interactions, with another larger systems that acts as a heat-bath and, itself, with only internal short-range interactions. This claim leads us to the issue of the *reduction* of a larger system into a smaller one by integrating away a large part of the combined ensemble, a problem that we will discuss in Sec. 4.2.

2.4 Gaussian energy fluctuations in the canonical ensemble

The probability distribution $P(E)$ in the canonical ensemble, $P(E) = g(E)e^{-\beta E}/Z$, has an extremum at $E = U(\beta)$, with $U(\beta)$ given by[3]

$$\partial_E S(E)|_{E=U(\beta)} = k_B \beta = T^{-1} , \tag{2.22}$$

where $S(E) \equiv k_B \ln g(E)$ is the (microcanonic) entropy defined from the density of states $g(E)$, see eq. (2.13). The stability of this extremum depends on the sign of $\partial^2_E S(E)|_{E=U(\beta)}$ and we examine it below.

The Taylor expansion of $\ln[ZP(E)]$ around $E = U$ (not writing the β dependence of U to lighten the notation) yields

$$\begin{aligned} \ln[ZP(E)] &\sim -\beta[U - TS(U)] + \frac{1}{2k_B} \left.\frac{\partial^2 S(E)}{\partial E^2}\right|_{E=U} (E-U)^2 + \dots \\ &\sim -\beta[U - TS(U)] - \frac{k_B}{2}\beta^2 \frac{1}{C_V^{can}}(E-U)^2 + \dots \end{aligned} \tag{2.23}$$

where we took the derivative of (2.22) with respect to β

$$k_B = \left.\frac{\partial^2 S(E)}{\partial E^2}\right|_{E=U(\beta)} \frac{\partial U(\beta)}{\partial \beta} \tag{2.24}$$

to replace the coefficient of the quadratic term as a function of the *canonical specific heat*[4]

$$C_V^{can} \equiv \frac{\partial U(T)}{\partial T} = -k_B \beta^2 \frac{\partial U(\beta)}{\partial \beta} . \tag{2.25}$$

Within the quadratic approximation of eq. (2.23) that neglects all higher order terms, $C_V^{can} > 0$ and the energy probability density $P(E)$ is a Gaussian centred at U (thus $U = \langle E \rangle$) with variance $\sigma^2 = C_V^{can}/(k_B\beta^2)$ (thus $\langle (E-U)^2 \rangle = C_V^{can}/(k_B\beta^2)$).

Is the assumption $C_V^{can} > 0$ true? Indeed, it is. In the canonical ensemble the constant volume specific heat (note that it is defined through $U(\beta)$ an entity that is canonical by construction)

$$\begin{aligned} C_V^{can} &= -k_B\beta^2 \frac{\partial U(\beta)}{\partial \beta} = -k_B\beta^2 \frac{\partial \langle E \rangle(\beta)}{\partial \beta} = k_B\beta^2 \frac{\partial^2 \ln Z(\beta)}{\partial \beta^2} \\ &= k_B\beta^2 \left[-\frac{1}{Z^2(\beta)} \left(\frac{\partial Z(\beta)}{\partial \beta}\right)^2 + \frac{1}{Z(\beta)} \frac{\partial^2 Z(\beta)}{\partial \beta^2} \right] = k_B\beta^2 \ \left[\langle E^2 \rangle(\beta) - \langle E \rangle^2(\beta)\right] \\ &= k_B\beta^2 \langle (E - \langle E \rangle)^2 \rangle(\beta) > 0 \end{aligned} \tag{2.26}$$

---

[3] One proves it by taking the ln (a monotonic function of its argument), replacing $g(E)$ by its relation to $S(E)$, and taking the derivative.

[4] In order to check the presence of the k_B factors one can do a dimensional analysis and see that $[C_V^{can}] = [k_B]$.

is positive definite. All averages in these expressions have to be computed with the canonical pdf.

In this process we have proven a form of *fluctuation-dissipation* theorem or, in other words, a relation between a *susceptibility* (to changes in temperature) and *connected fluctuations* (of the energy):

$$-\frac{\partial\langle E\rangle}{\partial\beta} = \langle (E-\langle E\rangle)^2\rangle \qquad \Rightarrow \qquad \frac{\partial\langle E\rangle}{\partial T} = k_B\beta^2\,\langle (E-\langle E\rangle)^2\rangle\,. \tag{2.27}$$

In an extensive macroscopic system with N particles, $U \propto N$ and $C_V^{can} \propto N$ in which case the ratio between dispersion and typical energy, or relative fluctuation, vanishes as $\sigma/U \propto N^{-1/2}$. In the large N limit, fluctuations are 'killed', the energy in the canonical ensemble does not fluctuate, it is locked to the value U, and it is related to the temperature through (2.22).

In the thermodynamic limit $N \to \infty$ the relative fluctuations $\sigma/U \propto N^{-1/2} \to 0$ and $E \to U(\beta)$. This implies that the microcanonical and canonical descriptions should coincide as, in practice, in the canonical formalism the energy of a macroscopic system is concentrated on a single value U determined by the temperature $T = 1/(k_B\beta)$.

One thus proves the equivalence between the microcanonical and canonical results.

2.5 Negative specific heat and convex entropy function in the microcanonic ensemble

What happens when the equivalence fails? What kind of peculiar effect can one expect to find? One of the simplest mismatches found is the possibility of having negative specific heat in some range of energies in the microcanonical description of systems with long-range interactions. This is impossible in a canonical formalism. Indeed, the *microcanonical constant volume specific heat* is defined as

$$C_V^{micro} \equiv -k_B\beta^2\,\frac{\partial E(\beta)}{\partial\beta} = \frac{\partial E(T)}{\partial T} \tag{2.28}$$

[one inverts $\beta(E) = k_B^{-1}\partial_E S(E)$ to write $E(\beta)$ or $1/T(E) = \partial_E S(E)$ to write $E(T)$] and this quantity is not positive definite if the *entropy is not concave*. Indeed, after the simple manipulation,

$$\frac{\partial^2 S}{\partial E^2} = \frac{\partial}{\partial E}\left(\frac{\partial S}{\partial E}\right) = \frac{\partial}{\partial E}\left(\frac{1}{T}\right) = -\frac{1}{T^2}\frac{\partial T}{\partial E} = -\frac{1}{C_V^{micro}}\frac{1}{T^2} \tag{2.29}$$

and

the convexity of the entropy, $\partial^2 S/\partial E^2 > 0$ for some values of the energy implies that $C_V^{micro} < 0$ for those same values.

When the two ensembles are equivalent the micronanonical C_V^{micro} should be identical to the canonical one, C_V^{can}, and thus positive. However, it is possible to construct models, for instance those with long-range interactions, such that C_V^{micro} is negative in some range of energies (these are usually related to first order phase transitions in the canonical ensemble).

It turns out that when $C_V^{micro} < 0$ the system undergoes a first-order phase transition in the canonical ensemble. In a nutshell, the energy interval in which there is a convex entropy in the microcanonical setting, corresponds to a single value of the temperature in the canonical ensemble at which the system undergoes a first order phase transition. We will revisit this problem in the Chapter in which we will review phase transitions.

2.6 The macrocanonical ensemble

Finally in the macrocanonical ensemble one characterises the macroscopic state with the volume V, the temperature T, and the chemical potential μ.

Summarizing, in the microcanonical ensemble the system is isolated and temperature is defined as $T^{-1} \equiv \partial S/\partial E|_E$. In the canonical ensemble the system is in contact with a reservoir – considered to be a much larger system – with which it can exchange energy to keep temperature fixed to be the one of the external environment. In the macrocanonical ensemble the system is in contact with a reservoir with which it can exchange energy *and* particles. The equivalence between them is ensured only for systems with short-range interactions.

2.7 Equivalence of ensembles

The practical consequence of ensemble equivalence is that one has the freedom to choose the ensemble in which calculations are easier. In spite of its fundamental importance in the construction of statistical mechanics, the microcanonical ensemble is practically never used to perform analytical calculations, since calculations are much harder in its setting than in the canonical one. The microcanonical ensemble is, however, commonly used in numerical simulations, since it is at the basis of molecular dynamic simulations.

Paraphrasing H. Touchette [19], the questions as to whether there is equivalence of ensembles can be posed in different ways:

- *Thermodynamic equivalence.* Are the microcanonical thermodynamic properties of a system determined from the entropy as a function of energy the same as the canonical thermodynamic properties determined from the free energy as a function of temperature? Are energy and temperature always one-to-one related?

- *Macrostate equivalence.* Is the set of equilibrium values of macrostates (e.g., magnetization, energy, velocity distribution, etc.) determined in the microcanonical ensemble the same as the set of equilibrium values determined in the canonical ensemble? What is the general relationship between these two sets?
- *Measure equivalence.* Does the Gibbs distribution defining the canonical ensemble at the microstate level converge (in some sense to be made precise) to the microcanonical distribution defined by Boltzmann's equiprobability postulate?

Equivalence of ensembles holds at the thermodynamic level whenever the entropy is concave, that is to say, $\partial^2 S(E)/\partial E^2 < 0$.

(see the proof below). This also implies, under additional conditions, the equivalence of the microcanonical and canonical ensembles at the macrostate level. Less is known about the measure equivalence. A survey of recent research on this field can be found in [19]. Systems with short-range interactions have concave entropies and for them equivalence of ensembles is ensured. Systems with long-range interactions can have non-concave[5] entropies and, therefore, inequivalence of ensembles can apply to them [16, 17, 18].

The proof goes as follows. The partition function Z and the free-energy density f are linked by

$$Z(\beta, V, N) = e^{-\beta N f(\beta, v)} \tag{2.30}$$

with $v = V/N$ and the large N limit assumed. For a system of N identical particles, the partition function can be expressed as

$$\begin{aligned} Z(\beta, V, N) &= \frac{1}{N!}\int dE \int \prod_{i=1}^{N}\prod_{a=1}^{d} dq_i^a dp_i^a \; \delta(E - H(\vec{Q}, \vec{P}))\, e^{-\beta E} \\ &= \int dE\; g(E, V, N)\, e^{-\beta E} \\ &= \int dE\; e^{-N[\beta\varepsilon - k_B^{-1} s(\varepsilon, v)]} \end{aligned} \tag{2.31}$$

where, for large N, we replaced $S(E, V, N) = N s(\varepsilon, v)$ with $\varepsilon = E/N$ and $v = V/N$. In the large N limit we can evaluate the last integral by saddle-point *iff* the entropy is concave, $\partial^2 S(E)/\partial E^2 < 0$, otherwise the Gaussian corrections would make the remaining integral explode. Assuming concavity,

$$\beta f(\beta, v) = \inf_{\varepsilon} \left[\beta\varepsilon - k_B^{-1} s(\varepsilon, v)\right] \tag{2.32}$$

---

[5] A concave entropy density, $\lim_{N\to\infty} S(E, V, N)/N$, function satisfies $s(ce_1 + (1-c)e_2, v) \geq c s(e_1, v) + (1-c)s(e_2, v)$ for any choice of $e_1 = E_1/V$, $e_2 = E_2/V$, $v = N/V$ and $0 \leq c \leq 1$.

obtaining that $\beta f(\beta, v)$ is the Legendre transform of $s(\varepsilon, v)$. One can proceed backwards and derive that s is the Legendre transform of βf

$$k_B^{-1} s(\varepsilon, v) = \inf_{\beta} \left[\beta\varepsilon - \beta f(\beta, v)\right] . \tag{2.33}$$

For each value of β there is a value of ε that satisfies (2.32) and viceversa for (2.33).

In cases in which the entropy has a convex region, the inverse Legendre transform leading from βf to s yields the concave envelope of s (and cannot access its convex piece).

Ensemble inequivalence is not merely a mathematical drawback, but it is the cause of physical properties that can be experimentally verified.

2.8 Non-equilibrium steady states

The dynamics of systems with long-range interactions can be extremely slow and the approach to equilibrium can take a very long time, that increases with the number N of elementary constituents. This feature is induced by the long-range nature of the interaction itself and it is not a consequence of the existence of a collective phenomenon. The state of the system during this long transient is quasi-stationary but it is not a thermodynamic metastable state, as it does not lie on local extrema of equilibrium thermodynamic potentials. The nature of quasi-stationary states can depend on the initial condition. In addition, a variety of macroscopic structures can form spontaneously in out-of-equilibrium conditions for isolated systems: a fact that should not be a surprise given that already the equilibrium states of long-range systems are usually inhomogeneous.

2.9 Many physical examples

Computing the microcanonical distribution function of macroscopic systems with realistic long-range interactions is usually prohibitly difficult. One then works with *toy models* that are much simpler but capture the essential features of the realistic problems. Some of the toy models that have been studied in detail are

- The *self-gravitating two-body problem*:

  $$H(\vec{P}, \vec{Q}, \vec{p}, \vec{r}) = \frac{P^2}{2M} + \frac{p^2}{2\mu} - \frac{Gm^2}{r} \tag{2.34}$$

  where $\vec{P}$ and $\vec{Q}$ are the momentum and coordinate of the center of mass and $\vec{p}$ and $\vec{q}$ are the relative moment and coordinate, $M = 2m$ is the total mass, $\mu = m/2$ is the reduced mass and m is the mass of the individual particles. The distance r is the modulus of the vector $\vec{q}$. One also restricts the range of the r coordinate to the interval (a, R). The short-distance cut-off mimics hard spherical particles of radius $a/2$. In the limits $a \to 0$ and $R \to \infty$ this is the standard Kepler problem.

The statistical mechanics of this system are described in detail in [13]. The system has two natural energy scales $E_1 = -Gm^2/a$ and $E_2 = -Gm^2/R$ with $E_1 < E_2$. For $E \gg E_2$ gravity is irrelevant, there is a long distance between the particles ($r > R$) and the system behaves like a gas, confined by a container. The microcanonic heat capacity is positive. As one lowers the energy the effects of gravity begin to be felt. For $E_1 < E < E_2$ neither the box nor the short-distance cut-off have an effect and there is a negative specific heat. As $E \sim E_1$ the hard core nature of the particles becomes important and gravity is again resisted, this is the low energy phase with positive microcanonic specific heat. The $T(E)$ dependence (and hence the microcanonical specific heat) is shown in Fig. 2.1. It must be noticed that astrophysical systems are in the intermediate energy scales with negative specific heat; moreover, this range is pretty wide since $E_1 \ll E_2$.

One can also analyse the canonical partition function – knowing already that it should predict a different behaviour from the above in the region $[E_1, E_2]$. In particular, one can compute the mean energy and its relation with temperature to compare with the microcanonical behaviour. One finds that at very low and very high energies the curves coincides. In the intermediate region the canonical $T(E)$ relation is almost flat and the canonical specific heat takes a very large value, almost a divergent one. This is similar to a phase transition in which the specific heat would diverge. The divergence is smoothened in this case due to the fact that there is a finite number of degrees of freedom in the two body problem.

- The *Lynden-Bell model* is a model of $(2N+1)$ coordinates evolving through the Hamiltonian
$$H = \frac{p^2}{2m} + \sum_{i=1}^{N} \frac{1}{2mr^2}\left({p_\theta}_i^2 + \frac{{p_\phi}_i^2}{\sin^2\theta_i}\right) - \frac{Gm^2}{2r} \qquad (2.35)$$
with r constrained to take values in (a, R). In this model one can take the large N limit and recover a true phase transition in the canonical formulation.

- The *Thirring model* is one with a set of N particles in a volume V. The particles interact with a constant potential if they come within an interaction volume V_o. In both the Lynden-Bell and Thirring models in the high energy regime the particles occupy space uniformly: it is a homogeneous phase. In the low energy regime instead the particles are close together in a collapsed phase.

- The *self-gravitating gas.* Consider a system of N particles interacting through Newtonian gravitational forces alone. The properties of this system depend on N. If $N = 2$ it is the exactly solvable Kepler problem, for $N = 3 - 50$, say, it cannot be solved exactly but it can be tackled with a computer. For larger N, $N = 10^5 - 10^{11}$, say, one is interested in averaged properties and statistical methods should be used.

  First, one must recall that a short-distance cut-off is necessary to render all phase space integrals convergent. This is justified by arguing that at very short distances not only the gravitational force acts on the particles and other forces regularise the $r \to 0$ behaviour of the total interaction potential.

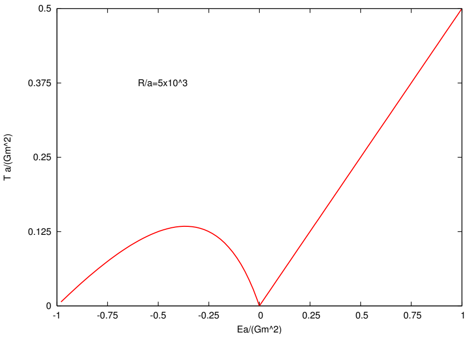

Figure 2.1: The temperature against the energy for the two-body problem with gravitational interaction with cut-off a in a box of size R. The microcanonical curve is non-monotonic with a negative specific heat in the region $E \sim [-0.4, 0]$. The canonical curve is monotonic and almost flat – as in a smoothened phase transition – in the region in which $C_V^{micro} < 0$.

It turns out that an interesting thermodynamic limit of the three ensembles (microcanonical, canonical and macrocanonical) is achieved in the very dilute limit [14, 15]

$$\eta \equiv \frac{Gm^2N}{V^{1/3}T} \qquad \text{finite} \; . \tag{2.36}$$

In this limit the thermodynamic quantities (free energy, energy, *etc.*) are functions of η and T and scale with N. Instead, the chemical potential and specific heat are just functions of T and η. The system undergoes collapse phase transitions in microcanonical and canonical ensembles though their location is different.

- The fully-connected *Blume-Capel* lattice spin-1 model,

$$H = \Delta \sum_{i=1}^{N} s_i^2 - \frac{J}{2N} \left(\sum_{i=1}^{N} s_i \right)^2 \;, \tag{2.37}$$

with s_i taking three values, $s_i = \pm 1, 0$, $J > 0$ the ferromagnetic coupling, $\Delta > 0$ the parameter that controls the energy difference between the ferromagnetic ± 1 states and the paramagnetic $s_i = 0$ state. The normalisation of the interaction term with N ensures that the total energy is extensive, *i.e.* $E = \mathcal{O}(N)$. This model has a phase diagram with two phases, ferromagnetic and paramagnetic, separated by a transition curve that is of second order for small Δ (the Curie-Weiss limit is attained for $\Delta = 0$) and of first order for large Δ. Details on the solution of this mean-field model are found in [16].

- The *Hamiltonian mean-field model* is defined by

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2} + \frac{J}{2N} \sum_{ij} [1 - \cos(\theta_i - \theta_j)] \tag{2.38}$$

  It represents unit mass particles moving on a unit circle with momenta p_i and position represented by an angular variable $\theta_i \in [0, 2\pi)$. The interactions are such that each particle interacts with all other ones in the same attractive ($J > 0$) or repulsive ($J < 0$) way (classical XY rotors).

3 Integrable systems and generalised Gibbs ensembles

In the past decade, atomic physics experiments have been able to test the global coherent dynamics of many-body interacting systems. As a result, interest in *far from equilibrium many-body Hamiltonian dynamics*, *coherent many-body dynamics* has emerged. Moreover, some of the quantum systems realised in the laboratory are modelled by *Integrable systems* and this fact re-boosted the interest in these models and their statistical physics description. These lectures treat only classical system and the rest of this Section will therefore focus on the Statistical Mechanics of classical integrable systems.

Integrable systems are nonlinear differential equations which, in principle, can be solved analytically, *i.e.* the solution can be reduced to a finite number of algebraic operations and integrations. The integrability of ordinary differential equations can be defined based on the existence of sufficiently many well-behaved *first integrals* or *constant of motions*. Integrability is related to a *sufficiently large symmetry group* that allows a construction of the most general solution.

In more concrete terms, an integrable system consists of a $2n$-dimensional phase space Γ together with n independent functions (in the sense that the gradients ∇O_k are linearly independent vectors on a tangent space to any point in Γ) $O_1, \ldots, O_n$: $\Gamma \to \mathbb{R}$, such that the mutual Poisson brackets vanish:

$$\{O_j, O_l\} = 0 \qquad \text{for all } \; j, l \tag{3.39}$$

(it is often said that the O_k are then *in involution*). We will assume henceforth that the O_k do not depend explicitly on time and that, threfore $dO_k/dt = 0$ is equivalent to $\{H, O_k\} = 0$. Conventionally, the first function O_1 is the Hamiltonian itself and therefore the first constant of motion is the energy. All other O_k with $k \neq 1$ are also constant of motion since their Poisson bracket with H vanishes. The dynamics of the system, can then be seen as the motion in a manifold with dimension $2n - n = n$. Under these conditions Hamilton's equations of motion are solvable. The way to solve them is to perform a *canonical transformation*[6] into *action, angle* variables (I_k, ϕ_k), with $k = 1, \ldots, n$ such that the Hamiltonian transforms into $\tilde{H}(I_k)$ and

$$I_k(t) = I_k(0) \; , \qquad \phi_k(t) = \phi_k(0) + t \, \frac{\partial \tilde{H}}{\partial I_k} = \phi_k(0) + t \, \omega_k(I) \; . \tag{3.40}$$

The action functions $I_k(t)$ are conserved quantities and we collected them all in I in the dependence of the frequencies $\omega_k(I)$. The remaining motion is given by n circular motions with constant angular velocities. Both deciding whether a system is integral and finding

---

[6] A canonical transformation is a change of variables that does not change the Hamiltonian structure of the system.

the canonical transformation that leads to the pairs (I_k, ϕ_k) are in practice very difficult questions. In case the system is integrable and one knows the action-angle pairs, the statement above is part of the *Liouville-Arnold theorem* [22].

If the systems of interest have more than one global conserved quantity, meaning that not just the total energy is constant but other quantities O_k are as well, the region of phase space on which the dynamics of the isolated system takes place is restricted to the configurations that share the initial values of all these quantities. As said above this is a space of $2n - n$ dimensions.

The *solar-planet* two body system is integrable. Therefore, the solar system, if one neglects the interplanetary interactions is an integrable system as well.

3.1 Generalized Microcanonical Measure

The fact that the microcanonical measure

$$\rho_{\rm GME}(\vec{X}) = c^{-1} \prod_{k=1}^{N} \delta(I_k(\vec{X}) - i_k) \ , \tag{3.41}$$

with

$$c = \int d\vec{X} \, \prod_{k=1}^{N} \delta(I_k(\vec{X}) - i_k) \equiv g(i_1, \dots, i_N) \tag{3.42}$$

(and a possible $N!$ factor, with N the number of indistinguishable particles, accompanying the definition of the *density of states* or *volume of the available phase space*, that we ignore) is sampled asymptotically (iff the *frequencies of the periodic motion on the torus are independent*, that is, $\vec{k} \cdot \vec{\omega} = 0$ for $\vec{k} = (k_1, \dots, k_N)$ with integer k_k has the unique solution $\vec{k} = 0$) is ensured by the Liouville-Arnold theorem [22]. One can call this ensemble the Generalized Microcanonical Ensemble (GME).

3.2 Generalized Canonical Measure

One can now construct the Generalized Canonical Ensemble (GCE), commonly called Generalized Gibbs Ensemble (GGE), from the GME following the usual steps. The idea is to look for the joint probability distribution of the n extensive (as for the Hamiltonian in the usual case) constants of motion, $P(i_1, \dots, i_n) di_1 \dots di_n$. The construction goes as follows.

Consider a system with volume V, divide it in two pieces with volumes V_1 and V_2, with $V = V_1 + V_2$, and values of the integrals of motion on the two partitions i_{k1} and i_{k2}. If we assume that the two systems are *independent* with the constraint that $i_{k1} + i_{k2} = i_k$, *i.e.* the *additivity property* for *each* integral of motion, the probability density of subsystem 1 to get the ensemble of values i_{k1} is

$$\begin{aligned} P(i_{11},\dots,i_{n1}) &\propto \int \prod_k di_{k2}\ \delta(i_k - i_{k1} - i_{k2})\ g(i_{11},i_{12},\dots,i_{n1},i_{n2}) \\ &= \int \prod_k di_{k2}\ \delta(i_k - i_{k1} - i_{k2})\ g_1(i_{11},\dots,i_{n1})\ g_2(i_{12},\dots,i_{n2}) \\ &= g_1(i_{11},\dots,i_{n1})\ g_2(i_1 - i_{11},\dots,i_n - i_{n1}) \end{aligned} \tag{3.43}$$

Introducing now the *entropy*

$$S_2(i_1 - i_{11},\dots,i_n - i_{n1}) = k_B \ln g_2(i_1 - i_{11},\dots,i_n - i_{n1}) \tag{3.44}$$

we have

$$\begin{aligned} P(i_{11},\dots,i_{n1}) &\propto g_1(i_{11},\dots,i_{n1})\ \exp\left[k_B^{-1} S_2(i_1 - i_{11},\dots,i_n - i_{n1})\right] \\ &\simeq g_1(i_{11},\dots,i_{n1})\ \exp\left[k_B^{-1} S_2(i_1,\dots,i_n) + k_B^{-1} \sum_{k=1}^{n} \partial_{i_k} S_2(i_k)(-i_{k1})\right] \\ &\propto g_1(i_{11},\dots,i_{n1})\ \exp\left[-\sum_{k=1}^{n} \beta_k i_{k1}\right] \end{aligned} \tag{3.45}$$

with

$$k_B \beta_k \equiv \frac{\partial S_2(i_1,\dots,i_k,\dots,i_n)}{\partial i_k}\ . \tag{3.46}$$

The Taylor expansion in the second line is justified by the assumption that system 1 is much smaller than system 2 and, hence, the full system. After fixing the normalisation, and erasing the subindex 1 for the selected subsystem,

$$P(i_1,\dots,i_n) = Z^{-1}(\beta_1,\dots,\beta_n)\ g(i_1,\dots,i_n) \exp\left[-\sum_{k=1}^{n} \beta_k i_k\right] \tag{3.47}$$

with

$$Z(\beta_1,\dots,\beta_n) = \int \prod_k di_k\ g(i_1,\dots,i_n) \exp\left[-\sum_{k=1}^{n} \beta_k i_k\right]\ . \tag{3.48}$$

Let us recap the assumptions made:
(i) independence, $g(i_{11},i_{12},\dots,i_{n1},i_{n2}) = g_1(i_{11},\dots,i_{n1})\ g_2(i_{12},\dots,i_{n2})$,
(ii) additivity of *all* constants of motion $i_k = i_{k1} + i_{k2}$,
(iii) small system 1 ($i_{k1} \ll i_k$ for all k),
(iv) constant inverse 'temperatures' $k_B \beta_k \equiv \partial_{i_k} S_2(i_1,\dots,i_n)$.
Note that assumptions (i) and (ii) fail in systems with long-range interactions [16, 17].

One way to ensure assumption (iii) is that i_ks are all extensive but, is it the only one? See the discussion below on the scaling with N in cases in which there are as many constants of motion as degrees of freedom in the system.

As in cases with just one conserved quantity, it is convenient to interpret P as a probability over position and momenta variables, and write

$$P_{\rm GGE}(q_1, p_1, \dots, q_N, p_N) = Z^{-1}_{\rm GGE}(\beta_1, \dots, \beta_n) \ \exp\left[-\sum_{k=1}^{n} \beta_k I_k(q_1, p_1, \dots, q_N, p_N) \right] \quad (3.49)$$

or in a more compact form

$$P_{\rm GGE}(\vec{X}) = Z^{-1}_{\rm GGE}(\beta_1, \dots, \beta_N) \ \exp\left[-\sum_{k=1}^{n} \beta_k I_k(\vec{X}) \right] \quad (3.50)$$

Concerning observables, the claim is then

$$\overline{A} \equiv \lim_{t_0 \ll \tau \ll t} \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} dt \ A(\vec{Q}(t), \vec{P}(t)) = \int \prod_{i=1}^{N} \prod_{a=1}^{d} dq_i^a dp_i^a \ A(\vec{Q}, \vec{P}) \ p_{\rm GGE}(\vec{Q}, \vec{P}) \ . \quad (3.51)$$

3.3 Generalised inverse temperatures

In the dynamic problem, the values of the parameters β_k should be fixed by the values of the constants of motion in the initial state

$$i_k = I_k(\vec{X}(t=0)) \quad (3.52)$$

and these should coincide with the statistical averages computed with the GGE

$$\langle I_k \rangle_{\rm GGE} = \int d\vec{X} \ P_{\rm GGE}(\vec{X}) \ I_k(\vec{X}) \qquad \text{for all } k \quad (3.53)$$

if the dynamics is characterised by this measure.

3.4 Scaling with system size?

Let us the scaling with system size in a case in which there are as many constants of motion as degrees of freedom, that is to say, $\{I_1, \dots, I_N\}$ or $n = N$ in the derivation above.

First of all the canonical weight

$$e^{-\sum_{k=1}^{N} \beta_k I_k} \quad (3.54)$$

would involve the exponential of a quantity $O(N^2)$ if all the I_ks scaled with N. This sounds weird. There are at least two ways out from this. One is to claim that all the I_ks are actually $O(1)$ but in this case we are in trouble with the derivation of the canonical weight, since we cannot claim $i_{k_1} \ll i_k$ for all k and hypothesis (iii) would fail. Another possibility would be that only a finite number of the I_ks are actually $O(N)$ while the rest are $O(1)$, but we would still have a problem with (iii).

Related to the scaling problems just mentioned, we would face severe problems in showing the equivalence between the microcanonical and canonical descriptions explained in Sec. 2.4. Following similar steps, that is to say, performing a Taylor expansion of $\ln[Z(\{\beta_k\})\, P(\{i_k\})]$ around the extrema $\{i_k = i_k^*\}$ one gets to

$$\begin{aligned}\ln[Z(\{\beta_k\})\, P(\{i_k\})] \;\simeq\;& -\sum_{k=1}^{N} \beta_k i_k + k_B^{-1} S(\{i_k^*\}) \\ & + \frac{1}{2} \sum_{k=1}^{N} \sum_{l=1}^{N} \left. \frac{\partial^2 \ln g(\{i_k\})}{\partial i_k \partial i_l} \right|_{\{i_k^*\}} (i_k - i_k^*)(i_l - i_l^*) \\ & + \dots \end{aligned} \qquad (3.55)$$

and we see, once again, that there might be scaling problems in the thermodynamic limit. In particular, the first term looks $O(N^2)$ if all i_k are proportional to N while the second would be $O(N)$ if this were the scaling of the entropy S. Moreover, the quadratic term also looks $O(N^2)$ if all the $\{i_k\}$ contributed equally and being $O(N)$ each of them. A further condition on the kl matrix in the term that controls the Gaussian fluctuations, namely that it should only have negative eigenvalues, should also be satisfied. As already mentioned, one possibility would be that only a finite number of $\{i_k\}$ are actually $O(N)$ and contribute to the measure.

This sort of questions are currently being posed in the context of quantum and classical problems.

3.5 Jaynes maximum entropy principle

The GGE probability distribution is sometimes justified advocating Jaynes maximum Shannon entropy principle. This principle is nowadays widely used in *inference problems*.

Inference is a brach of information theory with multidisciplinary applications and that shares concepts and techniques in common with statistical physics and, in particular, the theory of disordered systems. In its most general form one defines the inference problem as follows. Take a set of having a set of variables x_i, $i = 1, \dots, N$ and make M observations y_k, $k = 1, \dots, M$ of them, produced in some way by an unknown signal. We want to estimate the signal $\vec{x}$ of dimension N. A visual example is given by a pixelled photograph from which we only have a blurred image. Inference aims to uncover structure in data, in other words, to extract useful information from large amounts of data, and it is particularly useful nowadays with the *Big Data* revolution. Two main concerns in statistical inference are sufficient information and computational efficiency, explained by the questions (i) Under what conditions is the information contained in the observations sufficient for satisfactory recovery of the variables? (ii) Can the inference be done in an algorithmically efficient way? These concerns can be set in statistical physics terms and are related to phase transitions [23].

The inference problem we focus on here is defined as follows. Imagine that you have a random process, the probability of which you do not know. However, you do know the average value of a number, say m, of functions of this process. The idea is to use this (partial) information to infer the probability distribution of the process under the assumption that it maximises the *Shannon entropy* conditioned to satisfying the observed average values.

Let us explain how this construction works using an example, chosen for its notational simplicity. Consider a random variable X that can take n discrete values, $x_1, \dots, x_n$ with probabilities $p_1, \dots, p_n$ that we do not know. These probabilities have to be normalised so we know one constraint

$$\sum_{i=1}^{n} p_i = 1 \ . \tag{3.56}$$

Furthermore, we know the averages c_k, with $k = 1, \dots, m$ of m functions $f_k(\{x_i\})$:

$$\sum_{i=1}^{n} p_i f_k(\{x_i\}) = c_k \qquad \text{for} \quad k = 1, \dots, m \ . \tag{3.57}$$

The probabilities p_i are then estimated by requiring that the Shannon entropy

$$S_{\text{Shannon}}(\{p_i\}) = -\sum_{i=1}^{n} p_i \ln p_i \tag{3.58}$$

considered as a function of the n probabilities $\{p_i\}$, be maximised under the constraints (3.56) and (3.57). The construction is achieved by imposing the maximisation on the Lagrangian

$$\mathcal{L} = S_{\text{Shannon}}(\{p_i\}) - (\lambda_0 - 1)\left(\sum_{i=1}^{n} p_i - 1\right) - \sum_{k=1}^{m} \lambda_k \left(\sum_{i=1}^{n} p_i f_k(\{x_i\}) - c_k\right) \tag{3.59}$$

with $m+1$ Lagrange multipliers $\lambda_0, \lambda_1, \dots, \lambda_m$, and yields

$$p_i = e^{-\lambda_0 - \sum_{k=1}^{m} \lambda_k f_k(\{x_i\})} = \frac{e^{-\sum_{k=1}^{m} \lambda_k f_k(\{x_i\})}}{Z} \ . \tag{3.60}$$

The values of the Lagrange multipliers are fixed by the insertion of this expression in the conditions (3.56) and (3.57).

In the case of the GGE that we presented above, the functions $f_k(\{x_i\})$ are the action variables I_k, that themselves depend on the position and momenta of the particles, and the Lagrange multipliers are the generalised inverses temperatures β_k determined by the values of the conserved quantities. Note that in the common case in which the only conserved quantity is the energy one derives in this way the canonical Boltzmann measure. (Of course, the subtleties, and problems linked, for example to the long-range interacting potentials are not detected in this derivation.)

4 Reduced systems

In this section we will consider the reduction operation that consists in integrating away a portion of the system and investigate which are the statistical properties of what remains. We will distinguish the weak and strong coupling between the part retained and the part integrated away.

4.1 Canonical setting

In this lecture we always think of the system of interest being coupled to a mega environment with which it can exchange energy.

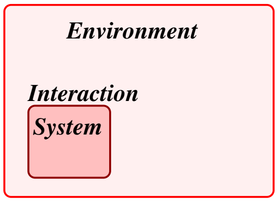

Figure 4.2: Sketch of the system and bath coupling.

4.2 The reduced partition function

We analyze the statistical static properties of a *classical canonical system* in equilibrium at inverse temperature β and itself formed by two sub-parts, one that will be treated as an environment (not necessarily of infinite size) and another one that will be the (sub-)system of interest. We study the *partition function* or Gibbs functional, $Z_{\rm tot}$:

$$Z_{\rm tot}[\eta] = \sum_{\substack{\text{conf env}\\ \text{conf syst}}} \exp(-\beta H_{\rm tot}) \tag{4.1}$$

where the sum represents an integration over the phase space of the full system, i.e. the system's and the environmental ones. We take

$$H_{\rm tot} = H_{\rm syst} + H_{\rm env} + H_{\rm int} \ . \tag{4.2}$$

For simplicity we use a single particle moving in $d = 1$: $H_{\rm syst}$ is the Hamiltonian of the isolated particle,

$$H_{\rm syst} = \frac{p^2}{2M} + V(x) \ , \tag{4.3}$$

with p and x its momentum and position. $H_{\rm env}$ is the Hamiltonian of a 'thermal bath' that, for simplicity, we take to be an ensemble of N independent harmonic oscillators with masses m_a and frequencies ω_a, $a = 1, \dots, N$

$$H_{\rm env} = \sum_{a=1}^{N} \frac{\pi_a^2}{2m_a} + \frac{m_a\omega_a^2}{2} q_a^2 \tag{4.4}$$

with π_a and q_a their momenta and positions. This is indeed a very usual choice since it may represent phonons. (These oscillators could be the normal modes of a generic Hamiltonian expanded to quadratic order around its absolute minimum, written in terms of other pairs of conjugate variables; the bath could be, for instance, a chain of harmonic oscillators with nearest-neighbor couplings.) $H_{\rm int}$ is the coupling between system and environment. We will restrict the following discussion to a linear interaction in the oscillator coordinates, q_a, and in the particle coordinate,

$$H_{\rm int} = x \sum_{a=1}^{N} c_a q_a \;, \tag{4.5}$$

with c_a the coupling constants.

The generalization to more complex systems and/or to more complicated baths and higher dimensions is straightforward. The calculations can also be easily generalized to an interaction of the oscillator coordinate with a more complicated dependence on the system's coordinate, $\mathcal{V}(x)$, that may be dictated by the symmetries of the system at the expense of modifying the counter-term. Non-linear functions of the oscillator coordinates cannot be used since they render the problem unsolvable analytically (although, of course, they can exist in Nature!).

Having chosen a quadratic bath and a linear coupling, the integration over the oscillators' coordinates and momenta can be easily performed. This yields the *reduced* partition function

$$Z_{\rm red} \propto \sum_{\rm conf\ syst} \exp\left[-\beta\left(H_{\rm syst} - \frac{1}{2}\sum_{a=1}^{N}\frac{c_a^2}{m_a\omega_a^2}\,x^2\right)\right] \;. \tag{4.6}$$

We notice that a quadratic term in x has been generated by the integration of the q_a variables. The reduced partition function can be written as

$$Z_{\rm red} \propto \sum_{\rm conf\ syst} \exp(-\beta H^*) \qquad \text{with} \qquad H^* = H_{\rm syst} - \frac{1}{2}\sum_{a=1}^{N}\frac{c_a^2}{m_a\omega_a^2}\,x^2 \;. \tag{4.7}$$

At this level different choices can be made.

- One can argue that the coupling constants are small enough, say $c_a^2 = \mathcal{O}(N^{-2})$, and drop the last term to recover the partition function of the selected system. This is a *weak coupling limit* between the selected system and the rest.

- One can claim that $c_a^2 = \mathcal{O}(N^{-1})$ and keep the modification of the Hamiltonian of the reduced system as a relevant one (and start in this way a *renormalisation* procedure by reproducing these steps many times).
- One can argue that, although the couplings are not small, say $c_a^2 = \mathcal{O}(N^{-1})$, the original Hamiltonian of the coupled system needs to be modified by adding a counter-term in order to recover only the Hamiltonian of the selected system after integrating away the oscillators. We explain this procedure in detail below.

In the last item above we propose to work with

$$H_{\rm tot} = H_{\rm syst} + H_{\rm env} + H_{\rm int} + H_{\rm counter} = H_{\rm syst} + \tilde{H}_{\rm env} \tag{4.8}$$

with

$$H_{\rm counter} = \frac{1}{2} \sum_{a=1}^{N} \frac{c_a^2}{m_a \omega_a^2} x^2 \tag{4.9}$$

so that the combination of the environmental, interaction and *counter-term* Hamiltonians take a rather simple and natural form

$$\tilde{H}_{\rm env} = H_{\rm env} + H_{\rm int} + H_{\rm counter} = \sum_a \frac{m_a \omega_a^2}{2} \left(q_a + \frac{c_a}{m_a \omega_a^2} x \right)^2 \; . \tag{4.10}$$

In this way,

$$H^* = H_{\rm syst} + H_{\rm counter} - \frac{1}{2} \sum_{a=1}^{N} \frac{c_a^2}{m_a \omega_a^2} \, x^2 = H_{\rm syst} \tag{4.11}$$

and

$$Z_{\rm red}[\eta] \propto \sum_{\rm conf\ syst} \exp\left[-\beta \left(H_{\rm syst}\right)\right] = Z_{\rm syst} \; . \tag{4.12}$$

The counter-term $H_{\rm counter}$ is chosen to cancel the term generated by the integration over the oscillators and it avoids the renormalization of the coefficient of the quadratic term in the potential that could have even destabilized the potential by taking negative values. For a non-linear coupling $H_{\rm int} = \sum_{a=1}^{N} c_a q_a \mathcal{V}(x)$ the counter-term is

$$H_{\rm counter} = \frac{1}{2} \sum_{a=1}^{N} \frac{c_a^2}{m_a \omega_a^2} [\mathcal{V}(x)]^2 \; . \tag{4.13}$$

The interaction with the reservoir does not modify the statistical properties of the particle since $Z_{\rm red} \propto Z_{\rm syst}$, independently of the choices of c_a, m_a, ω_a and N.

If one is interested in the *dynamics* of a coupled problem, the characteristics of the sub-system that will be considered to be the bath have an influence on the reduced dynamic equations found for the system, that are of generic Langevin kind. To take a weak coupling limit $c_a^2 \to 0$ to get rid of the generated force might be problematic since it implies also a very slow relaxation. One usually adds a counter-term in the dynamic formalism to get a good Langevin for the system.

5 Asymptotic states after quenches

Another field of research that is very active at present, motivated by analytic solutions of (simple) quantum systems and experiments in cold atomic systems, is the evolution following a *quench*. A quench is a rapid change in a parameter of the Hamiltonian, or the full potential in the Hamiltonian, that can be taken to be instantaneous. By this procedure one abruptly sets the system out of equilibrium and the question is what are the statistical properties of an ensemble of systems prepared with respect to the initial Hamiltonian and let evolve with the latter.

Take, as an example, a system constituted of a single point-like particle with mass m, and *initialise* it in canonical equilibrium within a harmonic potential $V_0(x)$. The probability distribution of the initial conditions is

$$P_0(x_0, p_0) = Z_0^{-1} \, e^{-\beta_i [\frac{p_0^2}{2m} + V_0(x_0)]} \; . \tag{5.14}$$

The averaged kinetic energy of the ensemble of initial states sampled with this pdf is

$$\frac{1}{2m} \langle p_0^2 \rangle = \frac{k_B T_i}{2} \; , \tag{5.15}$$

the *equipartition* of the kinetic energy. The averaged total energy is

$$\langle H_0 \rangle = \frac{\partial}{\partial \beta} \ln Z_0(\beta) \tag{5.16}$$

Make now a quench in the potential that corresponds to $V_0 \mapsto V$ and do it so quickly that the phase space variables do not change and remain p_0, x_0. By performing this abrupt change one injects or extracts a finite amount of energy,

$$\Delta E = H(x_0, p_0) - H_0(x_0, p_0) = V(x_0) - V_0(x_0) \; . \tag{5.17}$$

The energy surface on which the dynamics will take place is the one of the post-quench energy $E(0^+) = p_0^2/(2m) + V(x_0)$.

As a solvable problem, take V to be a simple harmonic potential. The Newton evolution of each initial configuration is

$$x(t) = x_0 \cos \omega t + \frac{p_0}{m\omega} \sin \omega t \tag{5.18}$$

$$p(t) = -m\omega x_0 \sin \omega t + p_0 \cos \omega t \tag{5.19}$$

The first question is,

- can one obtain constant values from the long-time averages of a single trajectory?

The answer is yes. For example, long-time average of $x(t)$ and $p(t)$ vanish, and the ones of the squares are

$$\overline{x^2(t)} = \frac{1}{2}\left(x_0^2 + \frac{p_0^2}{m^2\omega^2}\right) \tag{5.20}$$

$$\overline{p^2(t)} = \frac{1}{2}\left(m\omega^2 x_0^2 + p_0^2\right) \tag{5.21}$$

$$\overline{p(t)x(t)} = 0 \tag{5.22}$$

Another question is

- what is the probability distribution of the time-dependent phase space points $\{x(t), p(t)\}$?

Let us call $y = x(t)$ and $z = p(t)$ the position and momentum at a time t. The probability density of y, z at time t is

$$P(y,z,t) = \int dx_0 \int dp_0 \; P_0(x_0,p_0)\; \delta(y - x_0\cos\omega t - \frac{p_0}{m\omega}\sin\omega t) \\ \times\; \delta(z + m\omega x_0 \sin\omega t - p_0\cos\omega t)$$

We use the second δ function to integrate over p_0,

$$P(y,z,t) = \int dx_0 \; P_0\left(x_0, \frac{z}{\cos\omega t} + m\omega x_0 \tan\omega t\right)\frac{1}{\cos\omega t} \\ \times\; \delta(y - x_0\cos\omega t - \frac{z + m x_0 \omega \sin\omega t}{m\omega\cos\omega t}\sin\omega t)$$

The remaining δ function implies

$$y - \frac{z}{m\omega}\tan\omega t - x_0\left(\cos\omega t + \tan\omega t \sin\omega t\right) = y - \frac{z}{m\omega}\tan\omega t - x_0\frac{1}{\cos\omega t} = 0 \tag{5.23}$$

and we use it to integrate over x_0. Indeed, replacing $x_0 = y\cos\omega t - \frac{z}{m\omega}\sin\omega t$ and taking care of the Jacobian

$$P(y,z,t) = P_0\left(y\cos\omega t - \frac{z}{m\omega}\sin\omega t, \frac{z}{\cos\omega t} + m\omega y \sin\omega t - z\frac{\sin^2\omega t}{\cos\omega t}\right) \\ = P_0\left(y\cos\omega t - \frac{z}{m\omega}\sin\omega t, z\cos\omega t + m\omega y\sin\omega t\right) \tag{5.24}$$

No quench

Let us check this formula in the case in which no quench is performed. The initial potential has to be, then, harmonic $V_0(u) = m\omega^2 u^2/2$ and the equation above implies

$$\ln Z_0 + \ln P(y,z,t) = -\frac{\beta_i}{2m}\left(z\cos\omega t + m\omega y \sin\omega t\right)^2 - \frac{\beta_i}{2}m\omega^2\left(y\cos\omega t - \frac{z}{m\omega}\sin\omega t\right)^2 \\ = -\frac{\beta_i}{2m}z^2 - \frac{\beta_i}{2}m\omega^2 y^2$$

The equilibrium distribution is conserved by the dynamics, as it should.

Harmonic to harmonic quench

Imagine now that the quench corresponds to a change in the spring parameter of the quadratic potential $\omega_0 \mapsto \omega$. The equation for $P(y,z)$ implies

$$\ln Z_0 + \ln P(y,z,t) = -\frac{\beta_i}{2m}\left(z\cos\omega t + m\omega y \sin\omega t\right)^2 - \frac{\beta_i}{2} m\omega_0^2 \left(y\cos\omega t - \frac{z}{m\omega}\sin\omega t\right)^2$$

and expanding the squares

$$\begin{aligned}\ln Z_0 + \ln P(y,z,t) = &-\frac{\beta_i}{2m}\left(z^2\cos^2\omega t + m^2\omega^2 y^2 \sin^2\omega t + 2zm\omega y\cos\omega t\sin\omega t\right)\\ &-\frac{\beta_i}{2} m\omega_0^2\left(y^2\cos^2\omega t + \frac{z^2}{m^2\omega^2}\sin^2\omega t - 2y\frac{z}{m\omega}\cos\omega t\sin\omega t\right)\end{aligned}$$

and one can now collect terms

$$\begin{aligned}\ln Z_0 + \ln P(y,z,t) = &-\frac{\beta_i}{2}\left((\cos^2\omega t + \frac{\omega_0^2}{\omega^2}\sin^2\omega t)\frac{z^2}{m} + m\omega^2(\sin^2\omega t + \frac{\omega_0^2}{\omega^2}\cos^2\omega t)y^2\right.\\ &\left. +2zym\omega(1+\frac{\omega_0^2}{\omega^2})\cos\omega t\sin\omega t\right)\end{aligned}$$

Although the measure P is still Gaussian it does not have the same covariance as the initial P_0. It is simpler to see the time evolution of the variances of the position and momentum directly from the solutions to the equations of motion. One finds

$$\sigma_x^2(t) = \langle x^2(t)\rangle = \langle x_0^2\rangle\cos^2\omega t + \langle p_0^2\rangle\,\frac{1}{m^2\omega^2}\sin^2\omega t \tag{5.25}$$

$$\sigma_p^2(t) = \langle p^2(t)\rangle = \langle x_0^2\rangle\; m^2\omega^2\sin^2\omega t + \langle p_0^2\rangle\cos^2\omega t \tag{5.26}$$

Replacing now the averages of the initial values $\langle p_0^2\rangle/m = m\omega_0^2\langle x_0^2\rangle = T_i$

$$m\omega^2\sigma_x^2(t) \;=\; m\omega^2\langle x^2(t)\rangle = T_i\left(\frac{\omega^2}{\omega_0^2}\cos^2\omega t + \sin^2\omega t\right) \tag{5.27}$$

$$\frac{1}{m}\sigma_p^2(t) \;=\; \frac{1}{m}\langle p^2(t)\rangle = T_i\left(\cos^2\omega t + \frac{\omega^2}{\omega_0^2}\sin^2\omega t\right) \tag{5.28}$$

One readily verifies that, as expected, the averaged total energy is conserved

$$\langle E(t)\rangle = m\omega^2\sigma_x^2(t) + \frac{1}{m}\sigma_p^2(t) = T_i\left(1+\frac{\omega^2}{\omega_0^2}\right) \qquad \text{for} \quad t>0 \tag{5.29}$$

since each trajectory does conserve its initial energy. The averaged total energy is, however, different from the one right before the quench, $T_i = \langle E(t=0^-)\rangle \neq \langle E(t=0^+)\rangle = T_i\,(1+\omega^2/\omega_0^2)$.

Vanishing values of the position and momentum averages over time are readily obtained, $\overline{x(t)} = \overline{p(t)} = 0$. Time-independent values of the variances are found from the average over a long time window:

$$\begin{aligned} m\omega^2\overline{\sigma_x^2(t)} &= \frac{T_i}{2}\left(\frac{\omega^2}{\omega_0^2}+1\right) && (5.30) \\ \frac{1}{m}\overline{\sigma_p^2(t)} &= \frac{T_i}{2}\left(\frac{\omega^2}{\omega_0^2}+1\right) && (5.31) \\ \overline{x(t)p(t)} &= 0 && (5.32) \end{aligned}$$

- and these could be identified with equilibrium ones at a new temperature

$$T_f = T_i\left(\frac{\omega^2}{\omega_0^2}+1\right) \tag{5.33}$$

A Some useful formulæ

A.1 Stirling

Stirling formula for the factorial of a large number reads:

$$\ln N! \sim N \ln N - \ln N \;, \qquad \text{for} \qquad N \gg 1 \;. \tag{A.1}$$

A.2 Moments

Introducing a source h that couples linearly to a random variable x one easily computes all moments of its distribution $p(x)$. Indeed,

$$\langle\, x^k \,\rangle = \frac{\partial^k}{\partial h^k} \int dx \; p(x) e^{hx} \bigg|_{h=0} \;. \tag{A.2}$$

A.3 Gaussian integrals

The Gaussian integral is

$$I_1 \equiv \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi\sigma^2}} \; e^{-\frac{(x-\mu)^2}{2\sigma^2}} = 1 \;. \tag{A.3}$$

It is the normalization condition of the Gaussian probability density written in the *normal form*. One has

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi\sigma^2}} \; e^{-\frac{(x-\mu)^2}{2\sigma^2}} \; x = \mu \;, \\ \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi\sigma^2}} \; e^{-\frac{(x-\mu)^2}{2\sigma^2}} \; x^2 = \sigma^2 \;. \end{aligned} \tag{A.4}$$

From (A.3) one has

$$\int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi\sigma^2}} \; e^{-\frac{x^2}{2\sigma^2} + \frac{\mu x}{\sigma^2}} = e^{\frac{\sigma^2\mu^2}{2}} \;. \tag{A.5}$$

The generalization to N variables

$$I_N \equiv \int_{-\infty}^{\infty} \prod_{i=1}^{N} dx_i e^{-\frac{1}{2}\vec{x}^t A \vec{x} + \vec{x}^t \vec{\mu}} \tag{A.6}$$

with

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_N \end{pmatrix} , \qquad \vec{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \dots \\ \mu_N \end{pmatrix} , \qquad A = \begin{pmatrix} A_{11} & \dots & A_{1N} \\ A_{21} & \dots & A_{2N} \\ & \dots & \\ A_{N1} & \dots & A_{NN} \end{pmatrix} ,$$

and

$$-\frac{1}{2}\vec{x}^t A\vec{x} + \vec{x}^t\vec{\mu} \tag{A.7}$$

is the most generic quadratic form. Note that A plays here the role σ^{-2} in the single variable case. One can keep the symmetric part $(A + A^t)/2$ of the matrix A only since the antisymmetric part $(A - A^t)/2$ yields a vanishing contribution once multiplied by the vectors $\vec{x}$ and its transposed. Focusing now on a symmetric matrix, $A^t = A$, that we still call A we can ensure that it is diagonalizable and all its eigenvalues are positive definite, $\lambda_i > 0$. One can then define $A^{1/2}$ as the matrix such that $A^{1/2}A^{1/2} = A$ and its eigenvalues are the square root of the ones of A. Writing $\vec{x}^t A\vec{x} = (\vec{x}^t A^{1/2})(A^{1/2}\vec{x}) = \vec{y}\vec{y}$, the integral I_N in (A.6) becomes

$$I_N = \int_{-\infty}^{\infty} \prod_{i=1}^{N} dy_i J e^{-\frac{1}{2}\vec{y}^t\vec{y} + \vec{y}^t(A^{-1/2}\mu)} \tag{A.8}$$

where $J = \det(A^{1/2})^{-1} = (\det A)^{-1/2}$ is the Jacobian of the change of variables. Calling $\vec{\mu}'$ the last factor one has the product of N integrals of the type I_1; thus

$$I_N = (2\pi)^{N/2}(\det A)^{-1/2} e^{\frac{1}{2}\vec{\mu}^t A^{-1}\vec{\mu}} \tag{A.9}$$

Finally, the functional Gaussian integral is the continuum limit of the N-dimensional Gaussian integral

$$\vec{x} \equiv (x_1, \dots, x_N) \to \phi(\vec{x}) \tag{A.10}$$

and

$$I = \int \mathcal{D}\phi \; e^{-\frac{1}{2}\int d^dx d^dy \; \phi(\vec{x})A(\vec{x},\vec{y})\phi(\vec{y}) + \int d^dx \; \phi(\vec{x})\mu(\vec{x})} \; . \tag{A.11}$$

The sum runs over all functions $\phi(\vec{x})$ with the spatial point $\vec{x}$ living in d dimensions. The first and the second term in the exponential are quadratic and linear in the field, respectively. In analogy with the I_N case the result of the *path integral* is

$$I \propto e^{\frac{1}{2}\int d^dx d^dy \; \mu(\vec{x}) \, A^{-1}(\vec{x},\vec{y}) \, \mu(\vec{y})} \tag{A.12}$$

where we ignore the proportionality constant. Indeed, this one depends on the definition of the path-integral measure $\mathcal{D}\phi$. Usually, the actual value of this constant is not important since it does not depend on the relevant parameters of the theory. The inverse A^{-1} is defined by

$$\int d^dy \; A^{-1}(\vec{x},\vec{y})A(\vec{y},\vec{z}) = \delta(\vec{x} - \vec{z}) \; . \tag{A.13}$$

A.4 Wick's theorem

Take a Gaussian variable x with mean $\langle\, x \,\rangle = \mu$ and variance $\sigma^2 = \langle\, x^2 \,\rangle - \langle\, x \,\rangle^2$. Its pdf is

$$p(x) = (2\pi\sigma^2)^{-1/2}\, e^{-(x-\mu)^2/(2\sigma^2)} \; . \tag{A.14}$$

All moments $\langle\, x^k \,\rangle$ can be computed with (A.2). One finds

$$\langle\, e^{hx} \,\rangle = e^{\frac{h^2\sigma^2}{2}+h\mu} \tag{A.15}$$

and then

$$\langle\, x^k \,\rangle = \frac{\partial^k}{\partial h^k}\, e^{\frac{h^2\sigma^2}{2}+\mu h}\Big|_{h=0} \tag{A.16}$$

from where

$$\begin{array}{ll} \langle\, x \,\rangle = \mu \; , & \langle\, x^2 \,\rangle = \sigma^2 + \mu^2 \; , \\ \langle\, x^3 \,\rangle = 3\sigma^2\mu + \mu^3 \; , & \langle\, x^4 \,\rangle = 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4 \end{array}$$

etc. One recognizes the structure of Wick's theorem: given k factors x one organises them in pairs leaving the averages μ aside. The simplest way of seeing Wick's theorem in action is by drawing examples.

The generalization to N Gaussian variables is immediate. Equation (A.15) becomes

$$\langle\, e^{\vec{h}\,\vec{x}} \,\rangle = e^{\frac{1}{2}\vec{h}A^{-1}\vec{h}+\vec{h}\vec{\mu}} \tag{A.17}$$

and the generalization of (A.16) leads to

$$\langle\, x_i \,\rangle = \mu_i \; , \qquad\qquad \langle\, x_i x_j \,\rangle = {A^{-1}}_{ij} + \mu_i\mu_j \; , \tag{A.18}$$

etc. In other words, whereever there is σ^2 in the single variable case we replace it by ${A^{-1}}_{ij}$ with the corresponding indices.

The generalization to a field theory necessitates the introduction of functional derivatives that we describe below. For completeness we present the result for a scalar field in d dimensions here

$$\langle\, \phi(\vec{x}) \,\rangle = \mu(\vec{x}) \; , \qquad\qquad \langle\, \phi(\vec{x})\phi(\vec{y}) \,\rangle = A^{-1}(\vec{x},\vec{y}) + \mu(\vec{x})\mu(\vec{y}) \; , \tag{A.19}$$

etc.

A.5 The Hubbard-Stratonovich or Gaussian decoupling

$$e^{bm^2} = \sqrt{\frac{b}{\pi}} \int_{-\infty}^{\infty} dx\; e^{-bx^2+2bmx} \tag{A.20}$$

for all b and m.

A.6 Functional analysis

A functional $F[h]$ is a function of a function $h: \vec{x} \to h(\vec{x})$. The variation of a functional F when one changes the function h by an infinitesimal amount allows one to define the functional derivative. More precisely, one defines $\delta F \equiv F[h+\delta h] - F[h]$ and one tries to write this as $\delta F = \int d^dx\, \alpha(\vec{x})\delta h(\vec{x}) + \frac{1}{2}\int d^dx d^dy\, \beta(\vec{x},\vec{y})\,\delta h(\vec{x})\delta h(\vec{y}) + \dots$ and one defines the functional derivative of F with respect to h evaluated at the spatial point $\vec{x}$ as

$$\frac{\delta F}{\delta h(\vec{x})} = \alpha(\vec{x})\ , \qquad \frac{\delta^2 F}{\delta h(\vec{x})\delta h(\vec{y})} = \beta(\vec{x},\vec{y}) \tag{A.21}$$

etc. All usual properties of partial derivatives apply.

A.7 Fourier transform

We define the Fourier transform (FT) of a function $f(\vec{x}$ defined in a volume V as

$$\tilde{f}(\vec{k}) = \int_V d^dx\ f(\vec{x})\, e^{-i\vec{k}\vec{x}} \tag{A.22}$$

This implies

$$f(\vec{x}) = \frac{1}{V}\sum_{\vec{k}} \tilde{f}(\vec{k})\, e^{i\vec{k}\vec{x}} \tag{A.23}$$

where the sum runs over all $\vec{k}$ with components k_i satisfying $k_i = 2m\pi/L$ with m an integer and L the linear size of the volume V.

In the large V limit these equations become

$$\tilde{f}(\vec{k}) \ = \ \int_V d^dx\ f(\vec{x})\, e^{-i\vec{k}\vec{x}} \tag{A.24}$$

$$\tilde{f}(\vec{x}) \ = \ \int_V \frac{d^dk}{(2\pi)^d}\ f(\vec{k})\, e^{i\vec{k}\vec{x}} \tag{A.25}$$

The Fourier transform of a real function $f(\vec{x})$ satisfies $\tilde{f}^*(\vec{k}) = \tilde{f}(-\vec{k})$.

B The saddle-point method

Imagine one has to compute the following integral

$$I \equiv \int_a^b dx\ e^{-Nf(x)}\ , \tag{B.1}$$

with $f(x)$ a positive definite function in the interval $[a, b]$, in the limit $N \to \infty$. It is clear that due to the rapid exponential decay of the integrand, the integral will be dominated by the minimum of the function f in the interval. Assuming there is only one absolute minimum, x_0, one then Taylor expands $f(x)$ upto second order

$$f(x) \sim f(x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 \tag{B.2}$$

and obtains

$$I \sim e^{-Nf(x_0)} \int_a^b dx \; e^{-N\frac{1}{2}f''(x_0)(x-x_0)^2} = e^{-Nf(x_0)} [N f''(x_0)]^{-1/2} \int_{y_a}^{y_b} dy \; e^{-\frac{1}{2}(y-y_0)^2} \; , \tag{B.3}$$

with $y_0 \equiv \sqrt{N f''(x_0)} x_0$ and similarly for y_a and y_b. The Gaussian integral is just an error function that one can find in Tables.

This argument can be extended to multidimensional integrals, cases in which there is no absolute minimum within the integration interval, cases in which the function f is not positive definite, etc.

References

[1] R. K. Pathria, *Statistical Mechanics*, 2nd ed. (Butterworth-Heinemann, Oxford, 1996).

[2] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill (New York and London, 1965).

[3] W. Feller, *An introduction to probability theory and its applications*, (J. Wiley & sons, 1971).

[4] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill Series in Fundamentals of Physics).

[5] K. Huang, *Statistical mechanics*, (J. Wiley & sons, 2nd ed., 1987).

[6] L. E. Reichl, *A modern course in statistical physics* (E. Arnold publ., 1980).

[7] W. Pauli, *Pauli lectures on physics*, Volume 4 Statistical Mechanics (MIT Press, 1973).

[8] L. D. Landau and E. M. Lifshitz, *Statistical physics I*, (Pergamon Press, 1969).

[9] M. Dunajski, *Integrable systems*, Cambridge University Lectures (2012).

[10] A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics*, (Dover, New York, 1949).

[11] P. Castiglione, M. Falcioni, A. Lesne and A. Vulpiani, *Chaos and coarse-graning in statistical physics*, (Cambridge University Press, Cambridge, 2008).

[12] H. Goldstein, C. Pool and J. Safko, *Classical mechanics*, 3rd ed. (Pearson Education, Singapore, 2002).

[13] T. Padmanabhan, *Statistical mechanics od gravitating systems*, Phys. Rep. **188**, 285 (1990).

[14] H. de Vega and N. G. Sánchez, *Statistical mechanics of the self-gravitating gas: thermodynamic limit, phase diagrams and fractal structures*, 7th PAris Cosmology Colloquium, Observatoire de Paris, 2002 & 9th Course of the International School of Astrophysics ‘Daniel Chalonge’, Palermo, Italy, 2002, N. G. Sámchez and Y. Parijskij eds. NATO ASI, Kluwer Series II vol 130, 2003. astro-ph/0505561.

[15] P-H Chavanis, *Statistical mechanics of two-dimensional vortices and stellar systems* in ‘Dynamics and Thermodynamics of Systems with Long Range Interactions’, T. Dauxois, S. Ruffo, E. Arimondo, M. Wilkens Eds., Lecture Notes in Physics Vol. 602, Springer (2002).

[16] A. Campa, T. Dauxois, and S. Ruffo, *Statistical mechanics and dynamics of solvable models with long-range interactions*, Phys. Rep. **480**, 57 (2009).

[17] *Long-Range Interacting Systems*, vol. 90. Lecture Notes of the Les Houches Summer School 2008, Oxford University Press, Oxford (2010) T. Dauxois, S. Ruffo, and L. F. Cugliandolo, L.F. (eds.).

[18] Y. Levin, R. Pakter, F. B. Rizzato, T. N. Teles, and F. P. C. Benetti, *Nonequilibrium statistical mechanics of systems with long-range interactions*, Phys. Rep. **535**, 1 (2014).

[19] H. Touchette, *Equivalence and Nonequivalence of Ensembles: Thermodynamic, Macrostate, and Measure Levels*, J. Stat. Phys. **159**, 987 (2015).

[20] J. T. Lewis and R. Russell, *An Introduction to Large Deviations for Teletraffic Engineers* (1997).

[21] H. Touchette, *The large deviation approach to statistical mechanics*, Phys. Rep. **478**, 169 (2009).

[22] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Graduate Texts in Mathematics, 60) (Springer-Verlag, Berlin, 1978).

[23] L. Zdeborová and F. Krzakala, Adv. Phys. **65**, 453 (2016).

[24] E. T. Jaynes, *Information Theory and Statistical Mechanics*, Phys. Rev. Series II **106**, 620-630 (1957).

[25] E. T. Jaynes, *Information Theory and Statistical Mechanics II*, Phys. Rev. Series II **108**, 171-190 (1957).

[26] E. T. Jaynes, *Prior Probabilities*, IEEE Transactions on Systems Science and Cybernetics **4**, 227-241 (1968).

[27] H. E. Stanley, *Introduction to phase transitions and critical phenomena* (Oxford University Press, New York, 1971).

[28] N. Goldenfeld, *Lectures on phase transitions and the renormalization group* (Addison-Wesley, 1992).

[29] J. Cardy, *Scaling and renormalization in Statistical Physics*, Cambridge Lecture notes in physics (Cambridge University Press, 1996).

[30] D. J. Amit, *Field theory, the renormalization group and critical phenomena*(World Scientific, Singapore, 1984).

[31] G. Parisi, *Statistical field theory* (Addison-Wesley, 1988).

[32] J. M. Kosterlitz and D. J. Thouless, J. Phys. C: Solid State Phys. **6**, 1181 (1973).

[33] H. J. Maris, L. P. Kadanoff, *Teaching the renormalization group*, Am. J. Phys. **46**, 652 (1978).

[34] A. Das, *Finite temperature field theory*, (World Scientific, 1997).

[35] J. I. Kapusta and C. Gale, *Finite temperature field theory theory and applications*, 2nd Ed. (Cambridge University Press, 2006).

[36] J. W. Negele and H. Orland, *Quantum Many-Particle Systems* (Addison Wesley, New York, 1988).

[37] G. Fletcher, *A mechanical analog of first- and second-order phase transitions*, Am. J. Phys. 65, 74 (1997). E. Guyon, *Second-order phase transitions: Models and analogies*, Am. J. Phys. 43, 877 (1975). R. V. Mancuso, *A working mechanical model for first- and second-order phase transitions and the cusp catastrophe*, Am. J. Phys. 68, 271 (2000).

Papers

1. Inequivalence of ensembles in systems with long range interactions

Julien Barré, David Mukamel, Stefano Ruffo, *Inequivalence of ensembles in a system with long range interactions*, cond-mat/0102036, Phys. Rev. Lett. **87**, 030601 (2001).

We study the global phase diagram of the infinite range Blume-Emery-Griffiths model both in the canonical and in the microcanonical ensembles. The canonical phase diagram is known to exhibit first order and continuous transition lines separated by a tricritical point. We find that below the tricritical point, when the canonical transition is first order, the phase diagrams of the two ensembles disagree. In this region the microcanonical ensemble exhibits energy ranges with negative specific heat and temperature jumps at transition energies. These results can be extended to weakly decaying nonintegrable interactions.

D. Mukamel, S. Ruffo, N. Schreiber, *Breaking of ergodicity and long relaxation times in systems with long-range interactions*, cond-mat/0508604.

The thermodynamic and dynamical properties of an Ising model with both short range and long range, mean field like, interactions are studied within the microcanonical ensemble. It is found that the relaxation time of thermodynamically unstable states diverges logarithmically with system size. This is in contrast with the case of short range interactions where this time is finite. Moreover, at sufficiently low energies, gaps in the magnetization interval may develop to which no microscopic configuration corresponds.

As a result, in local microcanonical dynamics the system cannot move across the gap, leading to breaking of ergodicity even in finite systems. These are general features of systems with long range interactions and are expected to be valid even when the interaction is slowly decaying with distance.

Alessandro Campa, Andrea Giansanti, David Mukamel, Stefano Ruffo *Dynamics and thermodynamics of rotators interacting with both long and short range couplings*, cond-mat/0510508.

The effect of nearest-neighbor coupling on the thermodynamic and dynamical properties of the ferromagnetic Hamiltonian Mean Field model (HMF) is studied. For a range of antiferromagnetic nearest-neighbor coupling, a canonical first order transition is observed, and the canonical and microcanonical ensembles are non-equivalent. In studying the relaxation time of non-equilibrium states it is found that as in the HMF model, a class of non-magnetic states is quasi-stationary, with an algebraic divergence of their lifetime with the number of degrees of freedom N. The lifetime of metastable states is found to increase exponentially with N as expected.

Takayuki Tatekawa, Freddy Bouchet, Thierry Dauxois, Stefano Ruffo, *Thermodynamics of the self-gravitating ring model*, Phys. Rev. E **71** (2005) 056111, cond-mat/0501583.

We present the phase diagram, in both the microcanonical and the canonical ensemble, of the Self-Gravitating-Ring (SGR) model, which describes the motion of equal point masses constrained on a ring and subject to 3D gravitational attraction. If the interaction is regularized at short distances by the introduction of a softening parameter, a global entropy maximum always exists, and thermodynamics is well defined in the mean-field limit. However, ensembles are not equivalent and a phase of negative specific heat in the microcanonical ensemble appears in a wide intermediate energy region, if the softening parameter is small enough. The phase transition changes from second to first order at a tricritical point, whose location is not the same in the two ensembles. All these features make of the SGR model the best prototype of a self-gravitating system in one dimension. In order to obtain the stable stationary mass distribution, we apply a new iterative method, inspired by a previous one used in 2D turbulence, which ensures entropy increase and, hence, convergence towards an equilibrium state.

2. Equilibrium phase transitions

G. Fletcher, *A mechanical analog of first- and second-order phase transitions*, Am. J. Phys. **65**, 74 (1997).

A mechanical model that exhibits first- and second-order phase transitions is analyzed. The possible configurations are found first by using Newtonian mechanics and second by determining the minimum of the effective potential energy taken from the Lagrangian. A comparison is made between the effective potential energy method and the Landau theory of phase transitions. Phase diagrams are obtained for the mechanical system and are compared with those of a ferromagnet.

D. Boyanovsky, H. J. de Vega, D. J. Schwarz, *Phase transitions in the early and the present Universe*, to appear in Ann. Rev. Nucl. Part. Sci 2006, hep-ph/0602002.

The evolution of the Universe is the ultimate laboratory to study fundamental physics across energy scales that span about 25 orders of magnitude: from the grand unification scale through particle and nuclear physics scales down to the scale of atomic physics. The standard models of cosmology and particle physics provide the basic understanding of the early and present Universe and predict a series of phase transitions that occurred in succession during the expansion and cooling history of the Universe. We survey these phase transitions, highlighting the equilibrium and non-equilibrium effects as well as their observational and cosmological consequences. We discuss the current theoretical and experimental programs to study phase transitions in QCD and nuclear matter in accelerators along with the new results on novel states of matter as well as on multi- fragmentation in nuclear matter. A critical assessment of similarities and differences between the conditions in the early universe and those in ultra- relativistic heavy ion collisions is presented. Cosmological observations and accelerator experiments are converging towards an unprecedented understanding of the early and present Universe.

An improved apparatus for demonstrating first- and second-order phase transitions: Ball bearings on a rotating hoop, Richard V. Mancuso and Guy A. Schreiber, Am. J. Phys. 73, 366 (2005)

A working mechanical model for first- and second-order phase transitions and the cusp catastrophe, Richard V. Mancuso Am. J. Phys. 68, 271 (2000).

We have modified a toy to demonstrate first- and second-order phase transitions. The toy consists of a ball constrained to move on a rotating hoop. Analysis of the equilibrium positions of the ball as a function of the angular velocity and location of the axis of rotation shows that this system contains a cusp catastrophe.

On water, steam, and string theory, Christof Schmidhuber Am. J. Phys. 65, 1042 (1997)

At a pressure of 220 atm and a temperature of 374 °C there is a second-order phase transition between water and steam. Understanding it requires a key concept of both condensed matter and elementary particle physics: the renormalization group. Its basic ideas are explained with images from computer simulations of the lattice gas model. Then I briefly review how the renormalization group is used to compute critical coefficients for the water?steam phase transition. The results of this calculation are in good agreement with experiment. Finally, some applications in particle physics and string theory are mentioned.

A hand-held demonstration of cosmological phase transitions, David Lange, Marc Sher, Joel Sivillo, and Robert Welsh, Am. J. Phys. 61, 1049 (1993).

Illustrating phase transitions with soap films, David R. Lovett and John Tilley, Am. J. Phys. 59, 415 (1991).

First-order and second-order phase transitions are demonstrated using soap-film models. The models consist of two-dimensional parallel plates or three-dimensional frameworks in which film patterns are maintained. By making the sizes of the frameworks variable, it is possible to induce switching between film patterns analogous to transitions between phases. These phase changes are discussed thermodynamically and using a catastrophe theory model.

A simple geometrical model of spontaneous symmetry breaking, P. K. Aravind, Am. J. Phys. 55, 437 (1987).

A common, everyday phenomenon has a geometrical aspect that, in some respects, is analogous to a second-order phase transition. Concepts such as spontaneous symmetry breaking, order parameter, critical point, and critical exponent can thus be exhibited in a purely geometrical context. The purpose of the model, which is entirely pedagogical, is to illustrate in an elementary and amusing way some of the concepts associated with phase transition theory.

The self-gravitating gas

H. J. de Vega, N. G. Sanchez, *Statistical Mechanics of the Self-Gravitating Gas: Thermodynamic Limit, Unstabilities and Phase Diagrams*, Comptes Rendus Physique 7 (2006) 391-397, astro-ph/0601600.

We show that the self-gravitating gas at thermal equilibrium has an infinite volume limit in the three ensembles (GCE, CE, MCE) when $(N, V) \to \infty$, keeping $N/V^{1/3}$ fixed, that is, with $\eta = Gm^2 N/[V^{1/3}T]$ fixed. We develop MonteCarlo simulations, analytic mean field methods (MF) and low density expansions. We compute the equation of state and find it to be locally $p(r) = Trho_V(r)$, that is a local ideal gas equation of state. The system is in a gaseous phase for $\eta < \eta_T = 1.51024...$ and collapses into a very dense object for $\eta > \eta_T$ in the CE with the pressure becoming large and negative. The isothermal compressibility diverges at $\eta = \eta_T$. We compute the fluctuations around mean field for the three ensembles. We show that the particle distribution can be described by a Haussdorf dimension $1 < D < 3$.

H. J. de Vega, N. Sanchez, *Statistical Mechanics of the self-gravitating gas: thermody-*

namic limit, phase diagrams and fractal structures, Lecture given at the 7th. Paris Cosmology Colloquium, Observatoire de Paris, June 11-15, 2002 and at the 9 th Course of the International School of Astrophysics 'Daniel Chalonge', Palermo, Italy, 7-18 September 2002, NATO ASI, astro-ph/0505561.

We provide a complete picture to the self-gravitating non-relativistic gas at thermal equilibrium using Monte Carlo simulations, analytic mean field methods (MF) and low density expansions. The system is shown to possess an infinite volume limit in the grand canonical (GCE), canonical (CE) and microcanonical (MCE) ensembles when $(N, V)- > \infty$, keeping $N/V^{1/3}$ fixed. We compute the equation of state (we do not assume it as is customary in hydrodynamics), as well as the energy, free energy, entropy, chemical potential, specific heats, compressibilities and speed of sound; we analyze their properties, signs and singularities. All physical quantities turn out to depend on a single variable $\eta = Gm^2N/[V^{1/3}T]$ that is kept fixed in the $N- > \infty$ and $V- > \infty$ limit. The system is in a gaseous phase for $\eta < \eta_T$ and collapses into a dense object for $\eta > \eta_T$ in the CE with the pressure becoming large and negative. At $\eta \simeq \eta_T$ the isothermal compressibility diverges and the gas collapses. Our Monte Carlo simulations yield $\eta_T \simeq 1.515$. We find that $PV/[NT] = f(\eta)$. The function $f(\eta)$ has a second Riemann sheet which is only physically realized in the MCE. In the MCE, the collapse phase transition takes place in this second sheet near $\eta_{MC} = 1.26$ and the pressure and temperature are larger in the collapsed phase than in the gaseous phase. Both collapse phase transitions (in the CE and in the MCE) are of zeroth order since the Gibbs free energy has a jump at the transitions.

H. J. de Vega, N. Sanchez, *The Cluster Expansion for the Self-Gravitating gas and the Thermodynamic Limit*, Nucl. Phys. B 711 (2005) 604-620, astro-ph/0307318.

We develop the cluster expansion and the Mayer expansion for the self-gravitating thermal gas and prove the existence and stability of the thermodynamic limit N, V to infty with $N/V^{1/3}$ fixed. The essential (dimensionless) variable is here $\eta = [Gm^2N]/[V^{1/3}T]$ (which is kept fixed in the thermodynamic limit). We succeed in this way to obtain the expansion of the grand canonical partition function in powers of the fugacity. The corresponding cluster coefficients behave in the thermodynamic limit as $[\eta/N]^{j-1}c_j$ where c_j are pure numbers. They are expressed as integrals associated to tree cluster diagrams. A bilinear recurrence relation for the coefficients c_j is obtained from the mean field equations in the Abel form. In this way the large j behaviour of the c_j is calculated. This large j behaviour provides the position of the nearest singularity which corresponds to the critical point (collapse) of the self-gravitating gas in the grand canonical ensemble. Finally, we discuss why other attempts to define a thermodynamic limit for the self-gravitating gas fail.

H. J. de Vega, J. A. Siebert, *The Self-Gravitating Gas in the Presence of Dark Energy: Monte-Carlo Simulations and Stability Analysis*, Nucl. Phys. B 726 (2005) 464-480, astro-ph/0410147.

The self-gravitating gas in the presence of a positive cosmological constant Lambda is studied in thermal equilibrium by Monte Carlo simulations and by the mean field approach. We find excellent agreement between both approaches already for $N = 1000$ particles on a volume V [The mean field is exact in the infinite N limit]. The domain of stability of the gas is found to increase when the cosmological constant increases. The particle density is shown to be an increasing (decreasing) function of the distance when the dark energy dominates over self-gravity (and vice-versa).We confirm the validity of the thermodynamic limit: N, $V- > \infty$ with $N/V^{1/3}$ and $\Lambda V^{2/3}$ fixed. In such dilute limit extensive thermodynamic quantities like energy, free energy, entropy turn to be proportional to N. We find that the gas is stable till the isothermal compressibility diverges. Beyond this point the gas becomes a extremely dense object whose properties are studied by Monte Carlo.

4. Monte Carlo methods

A new approach to Monte Carlo simulations in statistical physics: Wang-Landau sampling, D. P. Landau, Shan-Ho Tsai, and M. Exler, Am. J. Phys. 72, 1294 (2004).

We describe a Monte Carlo algorithm for doing simulations in classical statistical physics in a different way. Instead of sampling the probability distribution at a fixed temperature, a random walk is performed in energy space to extract an estimate for the density of states. The probability can be computed at any temperature by weighting the density of states by the appropriate Boltzmann factor. Thermodynamic properties can be determined from suitable derivatives of the partition function and, unlike "standard" methods, the free energy and entropy can also be computed directly. To demonstrate the simplicity and power of the algorithm, we apply it to models exhibiting first-order or second-order phase transitions.

Monte Carlo Calculations as an Aid in Teaching Statistical Mechanics, D. P. Landau, R. Alben

A simple Monte Carlo sampling method is used to illustrate the principles of statistical mechanics as applied to a simple magnetic system. The concepts of ensembles, statistical averages, and responses are clarified particularly with respect to the role of statistical fluctuations. The basic properties of magnetic phase transitions are also demonstrated using small systems of interacting moments.

A simple algorithm for the transport of gamma rays in a medium, F. Arqueros and G. D. Montesinos, Am. J. of Phys. 71 38-45 (2003).

A simple Monte Carlo algorithm for the simulation of the passage of gamma rays of about 1 MeV in a medium is presented. In this energy range the only relevant processes are

Compton scattering and photoelectric absorption. The algorithm allows the visualization of the photon tracks as well as the calculation of many quantities of interest. Several problems for a layer and a cylinder are proposed. In particular, the energy transferred to electrons in a water cylinder as a function of depth and the line shape of a NaI scintillator is calculated.

Study of radiation-matter interaction processes below 1 MeV from simulated data, Fernando Arqueros and Sergio Martinez, American Journal of Physics 60, 232-238 (1992).

The passage of 1-MeV photons through thin and thick foils of Al and Pb has been simulated. The results have been used to show the basic properties of the dominant effects in radiation-matter interaction processes at energies just below 1 MeV. For the interpretation of the data it is necessary to handle total and differential cross sections of the involved processes, as well as the conservation laws, mainly of the Compton effect, whose phenomenology is studied in depth. In particular, the effect of multiple interactions in thick foils is studied and some results of the simulation are compared with suggested theoretical approaches.

Interdisciplinary applications of computational statistical physics, Dietrich Stauffer, Am. J. of Phys. 67, 1207-1211 (1999).

Biological and financial applications of computational methods in statistical physics are discussed. Examples are given of evolutionary models of sexual reproduction and stock markets.

Monte Carlo estimations of e, P. Mohazzabi, Am. J. of Physics 66, 138-14 (1998).

Three physical processes and the corresponding Monte Carlo algorithms are outlined, in which the number e, the base of the natural logarithm, can be obtained. The value of e is estimated in each case, and the three algorithms are compared.

Updating Monte Carlo algorithms, J. R. Drugowich de Felício, Valter L. Líbero, Am. J. Phys, 64, 1281-1285 (1996).

Using the long-range Ising model, we present modern Monte Carlo techniques2̌014single and multiple histogram and entropic sampling2̌014which permit increasing the amount of information obtained from a simulation. Numerical results for the density of states, mean energy and specific heat are compared with exact calculations, easily handled in this case. As a consequence of the simplicity of the model, the ability of those methods to generate continuous plots of thermodynamical quantities can be appreciated even by students taking basic courses of statistical physics.

Compton scattering, the electron mass, and relativity: A laboratory experiment, P. L. Jolivette and N. Rouze, American Journal of Physics 62, 266-271 (1994).

Compton scattering in a semiconductor detector is used to "discover" the relativistic relation between energy and momentum and to demonstrate the dependence of p, E and gamma on beta. The motivation is to measure the (rest) mass of the electron, and this can be done to within 1 keV with a commonly available set of gamma ray sources. To determine precisely where the Compton edge occurs in a spectrum, a Monte Carlo calculation of detector response is described which also helps the student to understand the physics of the detection process.