DYNAMICS AND STATISTICAL MECHANICS

CARLANGELO LIVERANI

1. Some questions

If we consider a portion of matter, we know that it consists of a large number, say N, of atoms. Such atoms interact with each other and move according to dynamical laws. For simplicity let us assume that they are classical, that is their motion obeys Newtown's law. This means that the equations of motion are the Hamiltonian equations connected to some Hamiltonian of the type

$$H(q,p) = \frac{1}{2}\sum_{i=1}^{N} \|p_i\|^2 + \sum_{i,j=1}^{N} V(q_i - q_j), \tag{1.1}$$

where V is the interaction potential among the atoms.

The Hamiltonian (1.1) yields the $2dN$ differential equations (d is the dimension of the space, typically $d = 3$)

$$\begin{aligned} \dot{q}_i &= p_i \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}. \end{aligned} \tag{1.2}$$

Clearly there is no hopes to solve them explicitly (typical N are of the order of 10^{24}), yet they show that the motion of the atoms should be expected to be very complex. Nevertheless, our experience is that the behaviour of the matter is rather regular: for example the pressure of a gas or the temperature of a solid do not change widely from one moment to the next although the typical value of the velocity of the atoms (say at room temperature) can be of almost one kilometer per second.

Question 1: *How comes that all these atoms performing, at great speed, very complex motions appear to us as essentially motionless ?*

Although one cannot hope to solve the equations of motion, the simple fact that the motion satisfies (1.1) does have some implications. For example, if $(q(t), p(t))$ is a solution of (1.1), with initial conditions $(q(0), p(0)$, then it is easy to check that $(q(-t), -p(-t))$ is also a solution with initial conditions $(q(0), -p(0))$. This is called *reversibility* and says that if something happens and we record it on a camera, then showing it backward we still have a motion governed by the (1.1). That is: looking at the movie we cannot be sure if it is projected forward or backward. But in fact we are!

Question 2: *How comes that the motions of the atoms do not distinguish the direction of time while to us such a direction is very clear?*

Date: Rome, October 11, 2003.

These questions have motivated more than hundred years of researches giving rise to the field commonly known as *Statistical Mechanics*. In fact, we still do not have a completely satisfactory answer, yet it is slowly emerging a setting in which such questions can, at least in principle, be answered.

2. Points of view: Equilibrium Statistical Mechanics

In Equilibrium Statistical Mechanics one assumes that a systems in a box Λ and with Hamiltonian (1.1) is in fact described by the Gibbs measures

$$\mu_\beta(f) := Z_\beta \int dq\, dp\, e^{-\beta H(q,p)} f(q,p); \quad \mu(1) = 1, \tag{2.1}$$

where β plays the rôle of the inverse of the temperature and the condition that μ_β be a probability measure determine Z_β.

Lemma 2.1. *Calling Φ_t the Hamiltonian flow, $\Phi_t^* \mu_\beta = \mu_\beta$, for all $t \in \mathbb{R}$.*

Proof. It is well known that $H \circ \Phi_t = H$ and, by Liouville theorem, $\Phi_t^* m = m$, where m is the Lebesgue measure. Hence

$$\Phi_t^* \mu_\beta(f) = \mu_\beta(f \circ \Phi_t) = Z_\beta m(e^{-\beta H} f \circ \Phi_t) = Z_\beta m([e^{-\beta H} f] \circ \Phi_t) = \mu_\beta(f).$$

□

That is, one assumes that the system is described by a particular class of invariant measures for the dynamics. But, of course, the system, at least in the present framework of classical mechanics, must be in some precise state with definite values for all the positions and velocities of the atoms.

Question 3: *In which sense we can think that a system is described by a measure rather than the list of position and velocities?*

The are at least to possible way to answer to the above queston. On the one hand one can take the statistical point of view: we many prepare many systems in, as far as we can tell, a similar way (in this case all at inverse temperature β, note that β can be determinded by measuring the expectation of p^2) and then (2.1) would describe their statistic. On the other hand we could simply measure p^2 and then consider (2.1) a prescription (for example it could be obtained by requiring that the measure be the invariant measure with me smaller relative entropy with respect to Lebesgue and the with the expectation of p^2 in agreement with the measurement) to determine the probabilities of the possible (unknown) configuration. It does not matter so much which point of view one takes, in both cases one has a *prescription* that allows to compute probabilites of given events and their validity must be checked against the experiments.

At first site the above setting does not seem very interesting: we know that the temperature in a given solid does not changes wildly from place to place and this appears very different from the statement that if we have a bunch of solids then their average temperature does not changes widely. Yet, let us take the above point of view and explore more closely its consequences.

Let us consider a simplified situation: $V \equiv 0$ (no interactions) and $d = 1$: a non interacting gas in an interval. Moreover, let us fix $\beta = 1$. Let us consider the pressure against the right wall. If the measurement is done very fast then the pressure will simply be given by

$$P(q,p) := \sum_{i=1}^{N} \chi_\Delta(q_i) \max\{0, p_i\},$$

where Δ is some very small neighborhood of the wall containing the particles that interact with the wall at the given time. Clearly the pressure will depend on the configuration (p, q) yet we can compute the average pressure:

$$\mu(P) = z^N |\Lambda|^{-N} \sum_{i=1}^{N} \int e^{-H} \chi_\Delta(q_i) \max\{0, p_i\} = N \frac{|\Delta|}{|\Lambda|} \bar{p},$$

where $\bar{p} := z \int_{\mathbb{R}^+} e^{-\frac{1}{2}p^2} p \, dp$. At this point we ask ourselves: *what is the probability to see a different pressure?*

Let us define the set

$$A_\delta = \{(q,p) \in \Lambda^N \times \mathbb{R}^N \ : \ |P(q,p) - \mu(P)| \geq \delta\mu(P)\}$$

Set $x_i = \chi_\Delta(q_i) \max\{0, p_i\} =: \chi_\Delta(q_i) p_i^+$, then $P = \sum_{i=1}^N x_i$, the x_i are then independently distributed random variables and $\mathbb{E}(x_i) := \mu(x_i) = \frac{|\Delta|}{|\Lambda|}\bar{p} =: \varrho\bar{p}$.

Let $A_\delta^+ = \{(q,p) \in \Lambda^N \times \mathbb{R}^N \ : \ P(q,p) - \mu(P) \geq \delta\mu(P)\}$ and $A_\delta^- = \{(q,p) \in \Lambda^N \times \mathbb{R}^N \ : \ P(q,p) - \mu(P) \leq -\delta\mu(P)\}$, clearly $A_\delta = A_\delta^+ \cup A_\delta^-$. Then, for each $\lambda \geq 0$,

$$\begin{aligned}
\mu(A_\delta^+) &= \mu\left(\{e^{\lambda(P-(1+\delta)\mu(P))} \geq 1\}\right) \leq \mu\left(e^{\lambda(P-(1+\delta)\mu(P))} \chi_{\{e^{\lambda(P-(1+\delta)\mu(P))} \geq 1\}}\right) \\
&\leq \mu\left(e^{\lambda(P-(1+\delta)\mu(P))}\right) \leq \left[\mathbb{E}(e^{\lambda(x-(1+\delta)\mathbb{E}(x))})\right]^N
\end{aligned}$$

It is then natural to define the random variable $y := x - (1+\delta)\mathbb{E}(x)$.

Lemma 2.2. *For each $\lambda \in \mathbb{R}$, setting $\varphi(\lambda) := \mathbb{E}(e^{\lambda y})$, holds $\varphi \in \mathcal{C}^\infty(\mathbb{R})$. Moreover φ is strictly convex and $\varphi'(0) < 0$.*

Proof. Just notice that $\varphi^{(n)}(\lambda) = \mathbb{E}(y^n e^{\lambda y})$, but, for $n > 1$,

$$\mathbb{E}(y^n e^{\lambda y}) \leq \frac{z}{|\Lambda|} \int_\Lambda \int_{\mathbb{R}} e^{-\frac{1}{2}p^2 + \lambda(p^+\chi_\Delta(q) - (1+\delta)\varrho\bar{p})} (p^+\chi_\Delta(q) - (1+\delta)\varrho\bar{p})^n \, dp \, dq < \infty.$$

The convexity follows from

$$\varphi''(\lambda) = \mathbb{E}(y^2 e^{\lambda y}) > 0,$$

while $\varphi'(0) = \mathbb{E}(y) = -\delta\mathbb{E}(x) < 0$. □

From Lemma 2.2 it follows that the minimum of φ is in $\mathbb{R}^+$ and it is less than one (since $\varphi(0) = 1$). A simple computation, expanding in powers of λ, show that the minimum is close to $\lambda_* = \frac{\delta\mathbb{E}(x)}{\mathbb{E}(x^2) - (1+\delta)\mathbb{E}(x)^2} =: \delta c$, provided that δ is so small that $\mathbb{E}(x^2) - (1+\delta)\mathbb{E}(x)^2 > 0$. Such a choice yields

$$\mathbb{E}(e^{\lambda_*(x-(1+\delta)\mathbb{E}(x))}) = 1 + \delta c\mathbb{E}(y) + \frac{c^2}{2}\delta^2\mathbb{E}(y^2) + \mathcal{O}(\delta^3) \leq e^{-\rho|\Delta|^{-2}\delta^2}$$

for some $\rho > 0$ independent of Δ. Thus,

$$\mu(A_\delta) \leq 2e^{-\rho N|\Delta|^{-2}\delta^2}.$$

This means that, unless Δ is incredibly small (microscopic), the probability to see even a very small fluctuation is completely negligible. It is then possible, at least in principle, to obtain predictions in complete agreement with our experience notwithstanding our naive feeling that this was a inadequate setting. Nevertheless, one must consider that if the measure is thought to be instantaneous, then Δ must be microscopic, say of the order of an atomic radii (about $10^{-1)}$ meters). In such a case one could expect big fluctuations, contrary to our experience. Yet, the real measures do last a finite time, actually normally a quite a long time from the perspective of atomic motion given the high speed at which such a motion takes place (one atom in the air at room temperature has about 10^9 collisions per second). It is then clear that the length of the measurement must play a rôle, hence the need to consider a bit more in depth the dynamics.