The chaotic route from classical to statistical mechanics: the meaning of entropy

Antonio Costa

1 Phase space and Liouville's theorem

During this course, you have seen how we can generally use the statistics of large ensembles to make definitive statements about thermodynamic properties. But if you think about it, fundamentally, the air in this room is composed of an Avogadro's number of particles, each of them with its own dynamics, so you could imagine trying to describe the system by studying these trajectories. But we don't that though do we? In fact, we can somehow forget about the details of each molecule's dynamics and still learn relevant emerging properties of the system: like pressure, temperature, volume, etc. What is it that allows us to do this? In order to understand this, let's first look at the dynamics of a single particle.

In order to study the dynamics of a particle, we essentially look at its position and momentum: these two variables compose the phase space of the particle. Let's generalize the concept of phase space. A dynamical system essentially consists of a phase space, $\vec{x}$ acted upon by some law of evolution ϕ,

$$\dot{\vec{x}} = \phi(\vec{x}, t). \tag{1}$$

In discrete time, this differential equation can be written as a map, F, that takes the system from the state $\vec{x}_t$ to $\vec{x}_{t+1}$,

$$\vec{x}_{t+1} = F(\vec{x}_t, t). \tag{2}$$

Therefore, the full state of the system $\vec{x}$ is composed of the set of variables that allow us to predict the future $\vec{x}_{t+1}$ from the present $\vec{x}_t$.

To make this notion more concrete, let's look at a simple one dimensional classical harmonic oscillator for which, according to Hooke's law, $F = -kx$. Newton's laws of motion allows us to write down the acceleration as,

$$\ddot{x} = -\frac{k}{m}x \tag{3}$$

where k is the spring constant, and m is the mass of the oscillator.

Is x alone enough the predict the future of the oscillator? In other words, is x the state of the system? Well, clearly, in order to predict the next state of the oscillator we not only need to know its position, but also its momentum, $p = m\dot{x}$. If I just give you x you won't be able to know whether the oscillator is moving towards or away from the resting state. Including p in the state of the system allows us to write down Eq. (3) as a system of first order differential equations,

$$\begin{aligned} \dot{x} =& \frac{p}{m} \\ \dot{p} =& -kx, \end{aligned} \tag{4}$$

and therefore the state of the system is composed of $\{x, p\}$. In state space, the full trajectory of the harmonic oscillator looks like Fig. (1). Say we start by pushing the oscillator to a point x_M to the right of the origin, and we release it. This corresponds to siting at the point $(x, p) = (x_M, 0)$. As the system evolves, the momentum is negative as the oscillator moves towards the origin, and it moves pass the origin with maximum speed, at the point $(x, p) = (0, -p_M)$. The momentum then starts decreasing as the oscillator reaches the point $(x, p) = (-x_M, 0)$, at which the momentum is 0. Finally, the oscillator reverses its movement and performs a reflected version of what we just described.

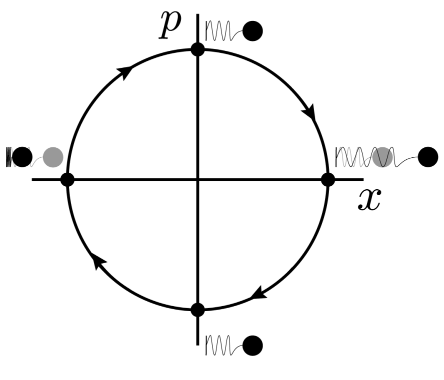


Figure 1: Phase space of the classical harmonic oscillator. The inset figures represent different positions of the oscillator: gray represents the ground state of the oscillator and black is the current position. Arrows indicate the direction of the flow.

If we think about a Hamiltonian picture of the dynamics, the notion of phase space falls naturally from Hamilton's equations. For each coordinate, there's a conjugate momentum, such that,

$$\dot{x} = \frac{\partial H}{\partial p}, \;\; \dot{p} = -\frac{\partial H}{\partial x}. \tag{5}$$

For the harmonic oscillator,

$$H = \frac{p^2}{2m} + \frac{kx^2}{2},$$

where the first term is the kinetic energy and the second is the harmonic potential energy. Solving Eq. (5) with this Hamiltonian, we easily get back Eq. (4), which was obtained simply by thinking principally about the definition of the state of a dynamical system.

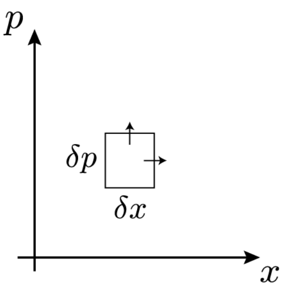
The very notion of a Hamiltonian is fundamental to the understanding of statistical mechanics. We know that in equilibrium statistical mechanics, the energy is conserved. In Hamiltonian mechanics, that's equivalent to saying that the Hamiltonian has no time dependent term,

$$\begin{aligned}
\dot{H} &= \frac{\partial H}{\partial p}\dot{p} + \frac{\partial H}{\partial x}\dot{x} + \cancel{\frac{\partial H}{\partial t}} \\
\dot{H} &= \dot{x}\dot{p} - \dot{p}\dot{x} \\
\dot{H} &= 0
\end{aligned}$$

And thus energy conservation comes from the fact that the Hamiltonian is time translation invariant. Consider now a system of N particles, each with their own positions and momenta. The Hamiltonian of the system is thus $H(x_i, p_i)$, where i is the index of each of the particles. Therefore, thinking about the phase space as the collection of all positions and momenta, we have a $2N$ dimensional phase space. The high dimensionality of phase space, combined with the properties of the dynamics, give rise to a probabilistic description of the system with which we can do statistical mechanics. The main goal of this lecture is to show how microscopic deterministic laws of motion can give rise to randomness, and therefore how we can forget about the fine scale dynamics of the system and learn its emergent thermodynamic properties using statistical arguments.

Now let's move on the idea of a flow in phase space. Instead of taking a certain initial condition and following it in time, let's imagine all possible starting points. You can think of it as having points uniformly distributed in phase space. The movement of the collection of points is analogous to a fluid. For the harmonic oscillator, this movement is really simple: all points move around the origin in the direction of the flow. More generally, for an arbitrary phase space, we will be interested in seeing how a volume of points evolves in time for a Hamiltonian flow.

To study volumes changes in phase space, let's think about a single particle, with phase space variables (x, p). Consider an infinitesimal phase space volume, which is essentially an infinitesimal square of area $dA = \delta x \delta p$. Moving the top side of the square moves p by $p \to \dot{p}dt$. If we move the top and the bottom side together, then we stretch δp by $\frac{\partial \dot{p}}{\partial p} dt \delta_p$: it's a change in p of a change in t. Therefore, the total change in volume in the p direction is



$$dA_p = \frac{\partial \dot{p}}{\partial p} \delta p dt \delta x,$$

the base times the height of the infinitesimal square. In much the same way, we can describe the change in the x direction as

$$dA_x = \frac{\partial \dot{x}}{\partial x} \delta x dt \delta p$$

Therefore, the total change in volume is,

$$dA = \left(\frac{\partial \dot{p}}{\partial p} + \frac{\partial \dot{x}}{\partial x} \right) \delta p \delta x dt \tag{6}$$

Filling in Hamilton's equations, Eq. (5), we get,

$$\begin{aligned} dA &= \left(-\frac{\partial H}{\partial p \partial x} + \frac{\partial H}{\partial x \partial p} \right) \delta p \delta x dt \\ dA &= 0 \end{aligned} \tag{7}$$

This is Liouville's theorem: in a Hamiltonian dynamical system, phase space volumes are conserved by the flow.

2 Second law of thermodynamics and conservation of phase space volumes

As we have seen, according to Liouville's theorem, if we take some initial volume in phase space V and act upon it with the dynamics, this volumes stays constant. In other words, if we take a ball of points and evolve it in time, it might get distorted, but the volume stays preserved.

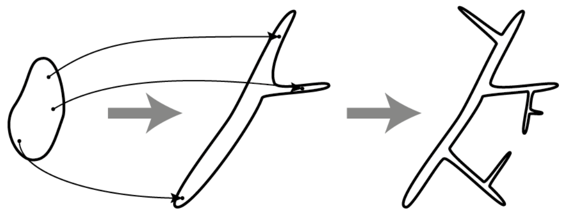


Figure 2: Phase space volume conservation. If we evolve a ball of points in time according to the dynamics, the phase space volume will generally stretch and contract along different directions, while preserving the total volume.

What does this mean for entropy? Let's assume that we have a uniform distribution of states in a phase space region, the leftmost in Fig. (2). In the context of the microcanonical ensemble, for which all states are equally probable, we define the entropy as

$$S = k_B \ln\Omega, \tag{8}$$

where Ω represents the number of microstates, k_B is the Boltzmann constant and ln represents the natural logarithm. In a continuous phase space, you can think of Ω as tiling the phase space volume into a collection of small patches, and counting how many times the systems visits each of these coarse-grained states. Therefore, the generalization of the number of microstates in a continuous state space is the phase space volume, such that the entropy can be written as

$$S \propto \ln V. \tag{9}$$

According to Liouville's theorem, as the system evolves, the phase space volume is preserved, implying that the entropy of the system stays preserved. Well, this is clearly inconsistent with the second law of thermodynamics, that states that any irreversible process is one that increases the system's entropy. If I spray perfume on a corner of the room, the perfume molecules will eventually distribute themselves around the room in an irreversible way. However, the dynamics of each molecule is perfectly deterministic and Newtonian mechanics tell us that we should be able to precisely reverse the trajectory of each individual particle and push it back into the perfume container. But we never observe that do we? Classical mechanics seems to be at odds with thermodynamics and statistical mechanics. It is both inconsistent with the irreversibility of thermodynamic processes and with the second law of thermodynamics. Let's try to understand this apparent contradiction, through the understanding of chaos.

3 A game of billiards

To illustrate chaos, let's play a game of *dissipationless* billiards. Imagine a billiards table that is absolutely frictionless, and in which the balls are perfectly reflected when they collide with each other or with the walls of the table, which have no holes, Fig. (3).

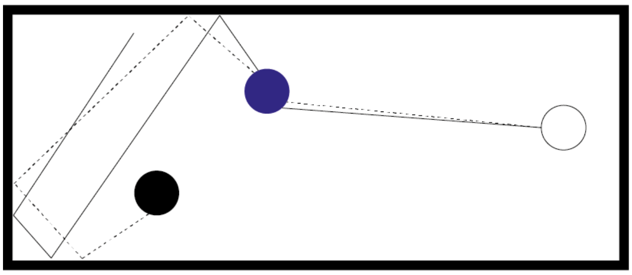


Figure 3: The game of *dissipationless* billiards. In this game, the table has no friction, there are no holes, and the balls are perfectly reflected upon collision. A tiny imprecision in the angle at which the cue ball is shot (white ball) can make a dramatic difference in the overall trajectory after just a few collisions: errors grow exponentially.

If we have infinite precision in the speed and orientation at which we shoot the cue ball (in white), we know precisely how the game will unravel. If we make a tiny mistake, however, the trajectory of the system will depart from the original one, and at each collision the errors will grow exponentially. Therefore, the system is extremely sensitive to the initial conditions, which is one of the properties of a chaotic system. Most systems, given enough complexity, exhibit this kind of behavior. To make the notion of chaos more precise, let's look at a canonical example: the logistic map.

4 Chaos: the logistic map

The billiards game can be described in terms of a discrete dynamical evolution, a map, in which the position of the ball changes in discrete jumps from collision to collision. In order to simplify, we will study instead a one dimensional map: the logistic map,

$$x_{t+1} = \mu x_t(1 - x_t). \tag{10}$$

The logistic map is a simple model of the evolution of a population level, which varies between 0 and 1, where the parameter μ governs the growth rate: the first term μx_t is analogous to the probability of growth, and $(1-x_t)$ is the probability of decay.

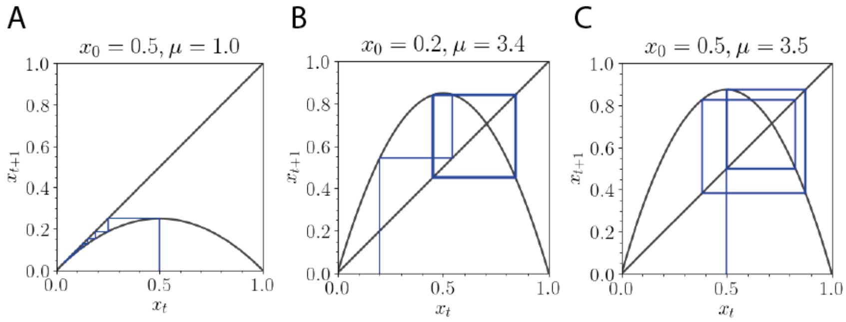


Figure 4: Trajectories of the logistic map for varying growth rates: $\mu = 1.0$ (A), $\mu = 3.4$ (B), $\mu = 3.5$ (C).

A simple way of visualizing the trajectory of a map is to make a cobweb plot: at each iteration, we take x_t to $F(x_t)$, and then draw a line from $F(x_t)$ to the diagonal to get the new location of x and iterate this process. Fig. (4) shows sample cobweb plots for different growth rates of the logistic map. As you can see, for $\mu = 1$ the logistic map flows towards to origin, which is a fixed point of the map. Fixed points are points in phase space that are mapped to themselves (unchanged by the dynamics): $x^*_{t+1} = x^*_t$. Solving

$$x^*_t = \mu x^*_t(1 - x^*_t),$$

it is easy to show that the fixed points of the logistic map are

$$x^* = 0$$
$$x^* = 1 - \mu^{-1}.$$

Notice that in Fig. (4B,C) the trajectory doesn't flow towards a trivial fixed point: it oscillates between two values. To understand the behavior of the trajectories, we need to discuss the stability of the fixed points. If a fixed point is stable, small perturbation will quickly decay back to the fixed point. On the other hand, for an unstable fixed point, small perturbations grow and the systems moves away from the fixed point. If we push the system away from a fixed point by $\delta x \ll 1$, the Taylor expansion around the fixed point yields $x^* + F'(x^*)\delta x + O((\delta x)^2)$. This means that for the perturbation to decay we need $|F'(x^*)| \leq 1$, while $|F'(x^*)| > 1$ yields unstable trajectories. Therefore in order to know whether the fixed point is stable or unstable, we need to look at $F'(x^*)$,

$$F'(x) = \mu(1 - 2x).$$

At the origin, $x^* = 0$, we have $F'(0) = \mu$: the origin is stable for $\mu \leq 1$ and unstable for $\mu > 1$. For $x^* = 1 - \mu^{-1}$, we have $F'(1 - \mu^{-1}) = 2 - \mu$, which is stable when $1 \leq \mu \leq 3$, and unstable otherwise. That is why in Fig. (4B,C) the trajectory never settles onto $x^* = 1 - \mu^{-1}$, but instead oscillates around it.

5 Coarse-graining, entropy and predictability

Let's now think about the dynamics of the logistic map in a different way. Instead of assuming we have infinite precision, let's be more realistic and assume that we have can only know the values of the logistic map up to some finite precision (say a few decimal points). This is equivalent to dicing the phase space into small little bins, and therefore obtaining a coarse-grained description of the dynamics. In fact, let's simply split the phase space into two bins: if $x < 0.5$ we label it L, if $x \geq 0.5$ we label in R (the left and right sides of phase space).

The question now is: how well can we predict the next coarse-grained state of the system, given the current state? In other words, if I am now in the L state, can you give me a estimate of the likelihood of the next state being L or R? Well, if we look at the examples of Fig. (4), we see that the sequence of states is

Example	Sequence of states
A	R L L L L L L L L L L L L L L ...
B	L R L R L R L R L R L R L R ...
C	R L R R R L R R R L R R R L ...

Table 1: Sequences of states on the example in Fig. (4).

In example A it is trivial to predict the next state of the system. Given the first $n - 1$ states, we would predict that the n-th state is L with almost

100% probability: the system is falling into a fixed point, which is also known as a period-1 orbit. In much the same way, in example B, we would predict that the next state is L, given the fact that the system has been oscillating between left and right and the sequence RL appears repeatedly: period-2 cycle. In example C, we can also predict that the next state will be R, given that the system repeats the sequence LRRR multiple times: period 4 cycle. In these regimes the logistic map is thus predictable.

5.1 Entropy and predictability

Let's now make a connection between the concepts of predictability and entropy. Recall the definition of entropy in the microcanonical ensemble, Eq. (8), for which all states are equally likely and, therefore, the probability of each state equals $p = 1/\Omega$. More generally though, as we know for the canonical ensemble, the probability of each state may depend on other properties of the system. Therefore, a more general definition of entropy is,

$$S = \sum_i -p_i \ln p_i. \tag{11}$$

The units of entropy are *nat* if we use the natural logarithm, or *bit* if we use a $\log_2$ basis. This definition of entropy was introduced by Shannon, in a seminal paper entitled "A Mathematical Theory of Communication" [1]. The entropy of a probability distribution essentially measures how much spread there is in the distribution. If among all states there is only one with probability $p(s) = 1$, meaning that the probability distribution is a delta function $p(x) = \delta(x - s)$, then the entropy of the system is $S = 0$ and we can precisely define the state of the system. On the other hand, if all states are uniformly distributed, such that $p(s) = 1/\Omega$, we recover the expression for the entropy of the microcanonical ensemble, Eq. (8) (up to the Boltzmann constant). In exactly the same fashion, you can also show that if p is a Boltzmann factor, then Eq. (11) resolves to the entropy in the canonical ensemble.

Let's now go back to our logistic map. We can think of predictability as the spread in the likelihood of finding the system in the state $s_{t+1} = j$, given that the current state of the system is $s_t = i$.

$$h = S(p(s_{t+1} = j | s_t = i))$$

In other words, if I am in state $s_t = \mathrm{L}$, and the probability of the next state is $p(s_{t+1} = \mathrm{R}) = 1$, then the system is totally predictable and we get an entropy of 0. If, on the other hand, there is some degree of unpredictability of the next state of the system, such that the probability of falling into states L and R is distributed across both states according to some probability distribution, then we get non zero entropy and the system becomes unpredictable. You can also think of unpredictability as the rate at which entropy grows as we increase the size of the sequence,

$$h = \lim_{N\to\infty} \frac{1}{N} S(p(s_N)) \tag{12}$$

where s_N is the sequence of length N. The units of entropy rate h are nat/symbol if we use a natural log, or bit/symbol is a $\log_2$ basis is used. In a predictable system, for example that of Fig. (4A), we see that as we increase the sequence size, the probability of finding that sequence (which is R, RR, RRR, etc.) anywhere in the full string of states is exactly one, and therefore the entropy rate is 0: the system is entirely predictable. In example Fig. (4B), the system oscillates between the left and right state, and therefore the probability of finding RL and LR is equal, $p(\mathrm{RL}) = 0.5$, $p(\mathrm{LR}) = 0.5$. Increasing the size of the symbol sequence to 3, we also find two sequences, LRL and RLR, with equal probability and therefore the $S(p(s_3)) = S(p(s_2))$. As we keep on increasing the size of the sequence, the entropy stays unchanged and, therefore, the entropy rate, or the unpredictability, as we take the limit $N \to \infty$ goes to $h = 0\, nat$/symbol. Likewise, in the example of Fig. (4C), since we have a repeated sequence (a period-4 cycle), the entropy rate is also going to tend to 0. If the dynamics exhibits a cyclic behavior, for which a sequence of states is repeated in a predictable fashion, the entropy rate vanishes.

5.2 Entropy and chaos

Let's now look at what happens when we push the growth rate to $\mu \to 4$. The obtained sequence, in this example, is

RRRLLLRRRRLRRLRLRLLL.

Can you know predict what the next letter is going to be? It turns out you can't, the trajectory is unpredictable. If the next state is completely unpredictable, then at any given sequence size, N, all sequences are possible. For sequences of length 1, we have $s = R$ or $s = L$, for $N = 2$ we have four possibilities: $s = RR$, $s = LL$, $s = LR$, $s = RL$, for $N = 3$ we have eight possibilities: $s = RRR$, $s = RRL$, $s = RLL$, $s = LLL$, $s = LLR$, $s = LRR$, $s = LRL$, $s = RLR$, so the number of allowed sequences grows with 2^N (for each element in the sequence of length N we have two possibilities, thus 2^N). Therefore, the entropy rate is

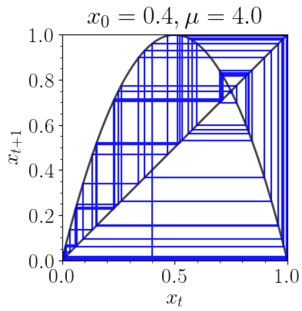


Figure 5: Cobweb plot of the logistic map for $\mu \to 4$.

$$
\begin{aligned}
h &= \lim_{N\to\infty} \frac{1}{N} \sum_{2^N} -\frac{1}{2^N} \ln\left(2^{-N}\right) \\
h &= \lim_{N\to\infty} \frac{1}{N} N \ln 2 \\
h &= \ln 2 \, nat/\text{symbol}
\end{aligned}
$$

As we have showed in the lecture, it is simple to simulate the logistic map: an example simulation can be found in *LogisticMap_Simulator.ipynb*. Using the simulation, and partitioning the phase space into L and R, it is easy to get an estimate of the entropy rate using Eq. (12). Doing so, we find $h \approx \ln 2 \, nat/\text{symbol} = 1 \, bit/\text{symbol}$ for $\mu \to 4$. Therefore, as we increase the growth rate μ, the system undergoes a cascade of period doubling bifurcations, and reaches chaotic behavior when essentially the period of the cycle diverges to infinity. In this scenario, we turn a deterministic law of evolution into a random number generator that is essentially indistinguishable from a coin toss experiment.

5.3 Chaos and thermodynamics

What does this mean for the system? Well, on the one hand, when the dynamics is chaotic, its behavior becomes random and irreversible. On the other hand, as you can see in Fig. (5), the systems visits all of phase space. As we have stated before, most system are chaotic. If we think back about the example of the perfume molecules spreading across the room in terms of a game of billiards with an Avogadro's number balls, it is easy to understand that the dynamics of the molecules is extremely chaotic. Being chaotic, as in the logistic map, the dynamics will push the system to populate the entire phase space.

Let's bring back the example of Fig. (2). In a Hamiltonian system, Liouville's theorem enforces that the phase space volume stays fixed as the system evolves. This is true for chaotic Hamiltonian systems, but the phase space volume becomes fractal, and branches out in all directions, filling phase space like a large piece of cotton. In a real physical system, we never have absolute precision over the exact position of a point in phase space. Instead, we should think about the phase space as a collection of coarse-grained little boxes. If that is the precision at which we can observe the system, then as the dynamics turns the phase space volume into a fractal, the effective volume spanned by the coarse-grained phase space grows as the system evolves, in an irreversible fashion, and therefore the entropy does increase. So chaotic dynamics plays a dual role in the relationship between classical mechanics and statistical mechanics/thermodynamics. On the one hand, it generates randomness from deterministic laws of evolution, such that we can achieve an accurate description of the system from a probabilistic/statistical perspective, without having to worry about the details of the underlying dynamics. On the other hand, chaotic dynamics makes the phase space volume evolve into a fractal, such that the effective coarse-grained entropy of the system increases, in accordance with the second law of thermodynamics.

6 Ising spins and entropy

The Hamiltonian (energy) of the Ising model in one dimension, can be written as,

$$H = -\epsilon \sum_{\substack{\text{neighboring} \\ \text{pairs } i,j}} s_i s_j, \tag{13}$$

where ϵ is the mutual interaction energy, and s_i represent the state of a spin, which is either up ($s_i = 1$) or down ($s_i = -1$). The partition function of the 1D Ising model, Eq.(8.43) in Schröder, is

$$Z \approx (2\cosh \beta\epsilon)^N \,, \tag{14}$$

where $\beta = (k_B T)^{-1}$, and N is the number of spins.

What is the entropy of this one dimensional chain of spins? Using the thermodynamic identity for the Helmotz free energy, we have

$$F = \langle E \rangle - TS, \tag{15}$$

and from this we can easily get the entropy. The average energy is simply

$$\langle E \rangle = -\frac{\partial}{\partial \beta} \ln Z = -N\epsilon \tanh{(\beta\epsilon)}, \tag{16}$$

and the free enery, F, is

$$F = -\frac{\ln Z}{\beta} = -\frac{N}{\beta} \ln(2\cosh \beta\epsilon). \tag{17}$$

Therefore we can write an expression for the entropy per spin, S/N, as,

$$\frac{1}{k_B}\frac{S}{N} = \ln(2\text{cosh}\beta\epsilon) - \beta\epsilon \tanh(\beta\epsilon). \tag{18}$$

If we focus on the high T limit, $\beta \to 0$, Eq.(18) reduces to,

$$\frac{1}{k_B}\frac{S}{N} = \ln 2 \, nat/\text{spin} \tag{19}$$

Given a sequence of spins, can we use Eq. (12) to estimate the entropy per spin? Let's simulate a 1D Ising chain using the Metropolis-Hastings algorithm, Fig. (6), *IsingModel_Simulator.ipynb* (we take $\epsilon = 1$ and $k_B = 1$). For more details on the implementation of the algorithm, see [2].

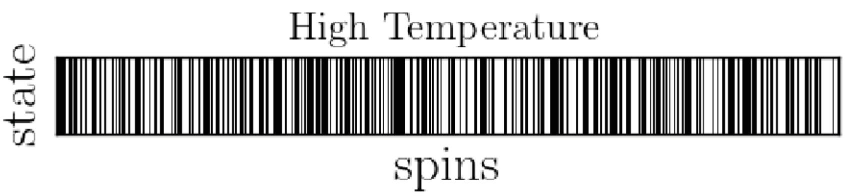


Figure 6: Final configuration, after 10^6 iterations, of a 1D spin chain using the Metropolis-Hastings algorithm with the Ising Hamiltonian, Eq. (13). Black represents spins in the "up" configuration, while white represents spins in the "down" configuration. At T=100, the spins assume random orientations.

Given the sequence of spin configuration, we can, as in the logistic map, estimate the entropy rate, or the entropy per state, h, which in this case will be the entropy per spin, S/N. Using the exact same method we have used to estimate the entropy rate of the logistic map, we obtain,

$$\frac{S}{N} \approx 0.692781 \approx \ln 2\, nat/\text{spin, with } T = 100$$

Therefore, using a simulation of a one dimensional chain of Ising spins, and estimating the entropy rate through counting the sequences of length N while taking $N \to \infty$, we are able to recapitulate the analytical result that $S/N \to \ln 2\, nat/\text{symbol}$ as $T \to \infty$. Also, you can see that the entropy rate of the high temperature 1D Ising model is equivalent to that of the logistic map for $\mu \to 4$, due to the fact that, in both cases, the system assumes any random configuration and therefore all sequences are equally probable and we obtain $h = \ln 2\, nat/\text{symbol} = 1\, bit/\text{symbol}$

Conclusion

In today's lecture, we have tried to highlight the intricate connections between classical mechanics and statistical mechanics, using the concepts of chaos and entropy as a guiding principle. We have seen how randomness can emerge from deterministic laws of motion, and how chaos gives rise to fractal phase space volumes and therefore to an increase in the coarse-grained entropy of the system, in accordance to the second law of thermodynamics. In addition, we have generalized the concept of entropy in the context of dynamical systems, specifically with the logistic map, and we have seen how the notions of entropy and predictability are related. Finally, we have used the 1D Ising model to show how we can estimate thermodynamic entropies using the same notion of entropy that allowed us to estimate the degree of unpredictability of a dynamical system.

References

[1] - Shannon, Claude E. (1948). “A Mathematical Theory of Communication”. Bell System Technical Journal. 27 (3): 379–423

[2] - Newman MEJ, Barkema GT, (1999) “Monte Carlo Methods in Statistical Physics”, Clarendon Press