On the Statistical Viewpoint of Entropy Increase

— A Reminder on the Ehrenfests' Urn Model —

Domenico Giulini
Physikalisches Institut
Universität Freiburg

Konstanz, 10.06.2005

Basic Statements

- **Assumption:** At time t_i the system is in a state $z(t_i)$ of *non-maximal* entropy.

- **Statement 1:** The probability that the state $z(t_i)$ will develop *in the future* to a state $z(t_{i+1})$ of larger entropy is larger than the probability for a development into a state of smaller entropy.

- **Statement 2:** The probability that the state $z(t_i)$ developped *in the past* out of a state $z(t_{i-1})$ of larger entropy is larger than the probability that it developped out of a state of smaller entropy.

Consequences and Remarks

- **Consequence 1:** The likely increase of entropy in the future state development $z(t_i) \mapsto z(t_{i+1})$ does not imply a likely decrease for the (fictitious) past development $z(t_i) \mapsto z(t_{i-1})$. Rather, the latter is also connected with a likely increase of entropy.

- **Remark:** To properly understand the last consequence, recall that our **condition** is placed on $z(t_i)$, that is, at time t_i. For $z(t_i) \mapsto z(t_{i+1})$ this means a *retarded* or *initial* condition, for $z(t_{i-1}) \mapsto z(t_i)$, however, an *advanced* or *final* condition. It is this change of condition which makes this behaviour of entropy possible.

- **Consequence 2** The likely increase of entropy in any direction away from a low-entropy state does not provide an orientation ("thermodynamic arrow") of time. Rather, an orientation is usually given by considering a finite time-interval and imposing a low-entropy condition at one of its two *ends*. However, without further structural elements, which would independently allow to distinguish the two ends, the apparently existing *two* possibilities are, in fact, identical. An apparent distinction is sometimes introduced by stating that the condition at one end is to be understood as *initial*. But, at this level, this merely defines 'initial' to be used for that end where the condition is placed.

Ehrenfests' Urn Model

- Let U_0 and U_1 be two urns with N (even) numbered (i.e. distinguishable) balls distributed amongst them.

- **Microstate:** Individual numbers of balls contained in U_1. Space of microstates ('phase space') is $\Gamma := \{0,1\}^N$ with elements $(x_1, \cdots x_N)$ (tells urn for each ball). Have $\mathrm{card}(\Gamma) = 2^N$.

- **Macrostate:** Cardinality of set of balls in U_1. Space of macrostates ('coarse grained phase space') is $\Omega := \{0, \cdots N\}$ with elements z. Have $\mathrm{card}(\Omega) = N + 1$.

- **Coarse Graining:** Projection map ('forget the individual numbers')

$$\pi : \Gamma \to \Omega \,, \qquad (x_1, \cdots x_N) \mapsto \sum_{i=1}^{N} x_i$$

- Have $$\mathrm{card}(\pi^{-1}(z)) = \binom{N}{z}$$

Probability Measures, Dynamics, and Observables

- A priori probability distribution on Ω:

$$W_{\mathrm{ap}}(z) := 2^{-N} \cdot \mathrm{card}(\pi^{-1}(z)) = 2^{-1}\binom{N}{z}$$

  To make this a **physical** probability measure, one has to prove from the **dynamical laws** that each microstate is equally probable in the sense of being reached equally often on time average.

- **Markoffian Dynamics:** At equidistant points in time, t_i, choose a random number $\in \{1, \cdots, N\}$ and 'instantaneously' let the corresponding ball change urns.

- **Observables:** Microscopically all functions $\Gamma \to \mathbb{R}$, i.e. $f(x_1, \cdots, x_N)$. Macroscopically the only 'coarse grained' observables ('relevant observables') are functions of $z = \sum_i x_i$.

- Let random variable $X : \Omega \to \mathbb{R}$ be $X(z) := z$; then

$$E(X, \mathrm{ap}) = \frac{N}{2} \qquad S(X, \mathrm{ap}) = \frac{\sqrt{N}}{2}.$$

Dynamics

- Let the macrostate at time t_i be z. For evolution $z(t_i) \mapsto z(t_{i+1})$ have two possibilities:
  - Picked number corresponds to ball in U_0; then $z(t_{i+1}) = z(t_i) + 1$.
  - Picked number corresponds to ball in U_1; then $z(t_{i+1}) = z(t_i) - 1$.
- **Conditional Probabilities** for $z \pm 1$ to occur at t_{i+1}, given macrostate at t_i is z, are:

$$
\begin{aligned}
W(z+1, t_{i+1}|z, t_i) &= \frac{N-z}{N} =: W_{\text{ret}}(z+1|z) \\
W(z-1, t_{i+1}|z, t_i) &= \frac{z}{N} =: W_{\text{ret}}(z-1|z)
\end{aligned}
$$

  Where 'ret' indicates that probabilities are **past-conditioned** or **retarded**.

Dynamics on Distributions and Stationary States

- Induced **forward** (in time) dynamics on probability distributions $W(z, t)$ is:

$$
\begin{aligned}
W(z; t_{i+1}) &= W(z, t_{i+1}|z+1, t_i)\, W(z+1, t_i) \\
&+ W(z, t_{i+1}|z-1, t_i)\, W(z-1, t_i) \\
&= \frac{z+1}{N} W(z+1, t_i) + \frac{N-z+1}{N} W(z-1, t_i)
\end{aligned}
$$

- **Proposition:** W_{ap} is the unique stationary distribution for this evolution law.

Bayes' Rule and Backward (in time) Dynamics I

- Given a probability space and set of events, $\{A_1, \ldots, A_n\}$, which is 1) complete (cover) and 2) exclusive (disjoint). Let B be some event. **Bayes' Rule** states:

$$W(A_k|B) = \frac{W(B|A_k)W(A_k)}{\sum_{i=1}^{n} W(B|A_i)W(A_i)} = \frac{W(B|A_k)W(A_k)}{W(B)}$$

- We identify the A_i with the $N+1$ events $(z'; t_i)$, where for $z' \in \{0, \cdots, N\}$, and A_k with the special event $(z \pm 1; t_i)$. The event B we identify with $(z; t_{i+1})$, i.e. the occurrence of z at the **later** time t_{i+1}. Then we can calculate the backward (in time) propagator:

$$\begin{aligned} W(z \pm 1, t_i \mid z, t_{i+1}) &= \frac{W(z, t_{i+1}|z \pm 1, t_i)W(z \pm 1, t_i)}{\sum_{z'=0}^{N} W(z, t_{i+1}|z', t_i)W(z', t_i)} \\ &= \frac{W(z, t_{i+1}|z \pm 1, t_i)W(z \pm 1, t_i)}{W(z, t_{i+1})} \end{aligned}$$

Bayes' Rule and Backward (in time) Dynamics II

- Using the known values for **past-conditioned** probabilities (forward propagators), which were

$$W(z+1, t_{i+1}|z, t_i) = \frac{N-z}{N} =: W_{ret}(z+1|z) \,,$$
$$W(z-1, t_{i+1}|z, t_i) = \frac{z}{N} =: W_{ret}(z-1|z) \,,$$

we can now calculate the **future-conditioned** probabilities (backward propagators):

$$W(z+1, t_i|z, t_{i+1}) = \frac{W(z+1, t_i)}{W(z+1, t_i) + \frac{N-z+1}{z+1} W(z-1, t_i)}$$
$$W(z-1, t_i|z, t_{i+1}) = \frac{W(z-1, t_i)}{W(z-1, t_i) + \frac{z+1}{N-z+1} W(z+1, t_i)}$$

Flow Equilibrium

- The condition for having flow equilibrium between times t_i and t_{i+1} reads

  $$W(z \pm 1; t_{i+1} | z; t_i) W(z; t_i) = W(z; t_{i+1} | z \pm 1; t_i) W(z \pm 1; t_i)$$

- **Proposition:** The above condition implies $W(z, t_i) = W_{ap}(z)$, i.e. the stationary distribution. Hence we also have stationarity and flow equilibrium for all $t_j > t_i$.

Time-Reversal Invariance I

- To be distinguished from flow equilibrium is time-reversal invariance. The latter is given by the following equality of past- and future-conditioned probabilities:

$$\begin{aligned} W(z \pm 1, t_{i+1} | z, t_i) &= W(z \pm 1, t_i | z, t_{i+1}) \\ &= W(z, t_{i+1} | z \pm 1, t_i) \frac{W(z \pm 1, t_i)}{W(z, t_{i+1})}, \\ \Longleftrightarrow \quad W(z, t_{i+1}) &= \frac{z+1}{N-z} W(z+1, t_i) \\ &= \frac{N-z+1}{z} W(z-1, t_i). \end{aligned}$$

- The condition of time-reversal invariance is strictly weaker that that of flow equilibrium. The former is implied, but does not itself imply the equilibrium distribution.

Time-Reversal Invariance II

- **Proposition:** Time-reversal invariance is stable under time-evolution. It is equivalent to the following 'constraint' on initial distribution:

$$W(z+1;t_i) = \frac{N-z}{z+1}\,\frac{N-z+1}{z}\,W(z-1;t_i).$$

  which has a one-parameter family of solutions. On those future and past time-evolutions coincide.

- Past and future evolution are **not** mutually inverse operations. The reason being that such a change in the direction of development is linked with a change from retarded to advanced conditionings.

Back to Statements 1 and 2

- Restrict to W_{ap}, then future-conditioned probabilities, too, are time-independent. Have $W(z \pm 1; t_i | z; t_{i+1}) =: W_{av}(z \pm 1|z)$, hence

$$
\begin{aligned}
W_{ret}(z+1|z) &= W_{av}(z+1|z) = \frac{N-z}{N} \\
W_{ret}(z-1|z) &= W_{av}(z-1|z) = \frac{z}{N}
\end{aligned}
$$

- Can now give a qualitative expression of **Statement 1** and **Statement 2**. Let $W_{max}(z)$, $W_{min}(z)$, $W_{up}(z)$, and $W_{down}(z)$ denote the probabilities for z to be a local maximum, minimum, to be on an ascending or descending branch respectively. Then

$$
\begin{aligned}
W_{max}(z) &= W_{av}(z-1|z)W_{ret}(z-1|z) = (z/N)^2 \\
W_{min}(z) &= W_{av}(z+1|z)W_{ret}(z+1|z) = (1-z/N)^2 \\
W_{up}(z) &= W_{av}(z-1|z)W_{ret}(z+1|z) = (z/N)(1-z/N) \\
W_{down}(z) &= W_{av}(z+1|z)W_{ret}(z-1|z) = (z/N)(1-z/N)
\end{aligned}
$$

Back to Statements 1 and 2

- Let's use instead of $z \in \{1, \cdots, N\}$ the bounded (in limit $N \to \infty$) variable σ, where $z = \frac{N}{2}(1+\sigma)$. Then

$$W_{\mathrm{max}}(\sigma) : W_{\mathrm{min}}(\sigma) : W_{\mathrm{up}}(\sigma) : W_{\mathrm{down}}(\sigma)$$
$$= \frac{1+\sigma}{1-\sigma} : \frac{1-\sigma}{1+\sigma} : 1 : 1$$

- Boltzmann entropy is

$$\begin{aligned} S_B(|\sigma|) &:= \ln(\mathrm{card}(\pi^{-1}(z))) \\ &\approx N\ln(N) - z\ln(z) - (N-z)\ln(N-z) \\ &= -\frac{N}{2}\left[\ln\frac{1-\sigma^2}{4} + \sigma\ln\frac{1+\sigma}{1-\sigma}\right] \end{aligned} \quad (1)$$

so that $S_B(|\sigma|) : [0,1] \to [\ln 2^N, 0]$ is strictly monotonically decreasing.

Thermodynamic Limit and Deterministic Dynamics

- Deterministic evolution for random variables results in the limit $N \to \infty$. Take $\Sigma = \sigma = \frac{2z}{N} - 1$; have

$$\begin{aligned} E(\Sigma, t_{i+1}) &= (1 - 2/N)\, E(\Sigma, t_i) \\ V(\Sigma, t_{i+1}) &= (1 - 4/N)\, V(\Sigma, t_i) + \frac{4}{N^2}\big(1 - E^2(\Sigma, t_i)\big) \end{aligned}$$

- In order to have a seizable fraction of balls moved within a macroscopic time span τ, we have to appropriately decrease the time steps $\Delta t := t_{i+1} - t_i$ with growing N, e.g. like $\Delta t = \frac{2}{N}\tau$, where τ is some positive real constant (the time span in which $N/2$ balls change urns). Now we can take the limit $N \to \infty$:

$$\begin{aligned} \frac{d}{dt}E(\Sigma, t) &= -\frac{1}{\tau}E(\Sigma, t) \;\Rightarrow\; E(\Sigma, t) &= E_0 \exp\left(\tfrac{-(t-t_1)}{\tau}\right) \\ \frac{d}{dt}V(\Sigma, t) &= -\frac{2}{\tau}V(\Sigma, t) \;\Rightarrow\; V(\Sigma, t) &= V_0 \exp\left(\tfrac{-2(t-t_2)}{\tau}\right) \end{aligned}$$

The Moral Once More

- According to the previous discussions it is clear that identical formulae would have emerged if W_{av} instead of W_{ret} had been used. Most importantly, the backward evolution is **not** obtained by taking the forward evolution and replacing in it $t \mapsto -t$. The origin of this difference is the fact already emphasized before, that $W_{av}(z; z')$ is not the inverse matrix to $W_{ret}(z; z')$, but rather the matrix computed according to Bayes' rule !!

<u>THE END</u>