The statistical mechanical arrow of time is often defined in terms of the tendency of "isolated mechanical system consisting of a very large numbers of molecules, to approach thermal equilibrium in which all ’macroscopic’ variables have reached steady values" (Uhlenbeck and Ford). An explanation of such tendency, in case it exists, is a step twards a definition the SM arrow of time. However, we should be clear first on what is the meaning of equilibrium, or more precisely a state of equilibrium (a state is a probability distribution on phase space).

My intention is NOT to eliminate the need for an independent notion of probability, by reducing it to long term relative frequency; nevertheless, the sense in which probability is objective and its connection to relative frequency will become clear.

The need for a realistic notion of equilibrium stems from various difficulties, some conceptual some technical. In the equilibrium state "all ’macroscopic’ variables reach steady values", meaning that their observed values remain most of the time near their phase-space average (weighted by the state). The standard accounts, however, suffer from considerable problems:

1. Following Boltzmann and the Ehrenfests the microcanonical distribution is taken by many to be the conceptually superior definition of equilibrium state for an isolated system. However, it often does not behave as expected: the ergodicity of the motion relative to it can rarely be established for realistic systems; worse, it can often be proved to fail.

2. The common interpretation of "macroscopic observable" is too liberal and includes many functions on phase-space which cannot possibly be measured (this is a point made by Khinchin).

3. The microcanonical distribution corresponds to completely and ideally closed systems. Here we encounter a schizophrenic attitude: The so called foundations of the theory rely on a theoretical notion of equilibrium state (that hardly and rarely delivers the goods), but all the calculations are done with Gibbs' canonical (or macro-canonical) distributions for systems that are not isolated

4. The excuse given for using Gibbs' (as opposed to Boltzmann's) formulation in practice is that all ensembles (alledgedly) become equivalent in the thermodynamic limit. But if, in the end, the large number of particles plays such a role in the explanation of thermal equilibrium why not put it from the start in the definition of equilibrium (this remark is also in the spirit of Khinchin).

The aim is therefore to define equilibrium in a more realistic way, that will not only apply to closed systems but include among others the canonical ensembles; a definition which imply that in equilibrium all reasonable macroscopic observables "have steady value", and reflect the fact that the system is made of a great number of particles.

I will begin with the mathematical definition and subsequently explain the physical motivation

Consider classical gas of n particles with positions $x_1, x_2, \ldots, x_n$ and momenta $p_1, p_2, \ldots, p_n$. Let $\rho(x_1, x_2, \ldots, x_n, p_1, p_2, \ldots, p_n)$ be its state, that is, a probability distribution defined over the phase space associated with it. We shall make four preliminary simplifying assumptions:

1.The probability distribution $\rho(x_1, x_2, \ldots, x_n, p_1, p_2, \ldots, p_n)$ is integrable with respect to the Lebesgue measure on R^{6n}.

2. The distribution ρ is stationary.

3. ρ is symmetric with respect to the interchange of particles. In other words:

$$\rho(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}, p_{\pi(1)}, p_{\pi(2)}, \ldots, p_{\pi(n)}) = \rho(x_1, x_2, \ldots, x_n, p_1, p_2, \ldots, p_n) \qquad 1$$

for every permutation π of $1, 2, \ldots, n$.

4. To avoid the mathematical complications associated with the thermodynamic limit, and make the discussion as general as possible, we assume that for each n the corresponding distribution ρ exists and is obtained from the higher-dimensional distributions as marginal. In other words, Kolmogorov's consistency conditions obtain (see eg Chow and Teicher)

Given a distribution ρ over n particles we can define its marginals:

$$\rho^{(1)}(x_1,p_1) = \int\int\ldots\int \rho(x_1,x_2,\ldots,x_n,p_1,p_2,\ldots,p_n)d^3x_2d^3x_3\ldots d^3x_nd^3p_2d^3p_3\ldots d^3p_n$$

$$\rho^{(2)}(x_1,x_2,p_1,p_2) = \int\int\ldots\int \rho(x_1,x_2,\ldots,x_n,p_1,p_2,\ldots,p_n)d^3x_3d^3x_4\ldots d^3x_nd^3p_3d^3p_4\ldots d^3p_n$$

And s o on for all values $\rho^{(k)}$.

Definition ρ *is an equilibrium state if the local density, and the correlations between local momentum fluctuations are stable.*

Explanation: let A be a region in physical space and B a region in the (3D) momentum space. We assume that these are small regions, but still of macroscopic size. The average number of particles in region A whose momenta lie in the region B at time t is

$$D(A,B,t) = \frac{1}{n}\sum_{i=1}^{n} \chi_B(p_i(t))\chi_A(x_i(t))$$

with χ_A being the indicator function of the set A etc. The local density is *stable* when its long term average converge as expected

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int_0^{\tau} D(A,B,t)dt = \int_B\int_A \rho^{(1)}(x,p)d^3xd^3p$$

for all regions A, B, as above.

Likewise for correlations. Take the case of local pair correlations Let $C(A_1,B_1,A_2,B_2,t)$ be the correlation between the (average) number of particles in A_1 with momenta in B_1 and the number of particles in A_2 with momenta in B_2 at time t

$$C(A_1,B_1,A_2,B_2,t)$$
$$= n^{-2}\sum_{ij}\chi_{B_1}(p_i)\chi_{B_2}(p_j)\chi_{A_1}(x_i)\chi_{A_2}(x_j)dt$$
$$- n^{-1}\sum_{i}\chi_{A_1}(x_i)\chi_{B_1}(p_i)dt \times n^{-1}\sum_{j}\chi_{A_2}(x_j)\chi_{B_2}(p_j)dt$$

these correlations are stable if

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau C(A_1,B_1,A_2,B_2,t)dt$$
$$= \int_{B_1}\int_{B_2}\int_{A_1}\int_{A_2}\rho^{(2)}(x_1,x_2,p_1,p_2)d^3x_1d^3x_2d^3p_1d^3p_2$$
$$- \int_{B_1}\int_{A_1}\rho^{(1)}(x_1,p_1)d^3x_1d^3p_1 \times \int_{B_2}\int_{A_2}\rho^{(1)}(x_2,p_2)d^3x_2d^3p_2$$

Denote by $c^{(2)}(x_1,x_2,p_1,p_2) = \rho^{(2)}(x_1,x_2,p_1,p_2) - \rho^{(1)}(x_1,p_1)\rho^{(1)}(x_2,p_2)$ the correlation density, and so on for higher correlation $c^{(3)}$, $c^{(4)}$,...we can represent the state as

$$\rho(x_1,x_2,\ldots,x_n,p_1,p_2,\ldots,p_n) =$$ 3
$$= \prod_{j=1}^{n}\rho^{(1)}(x_j,p_j)\Big(1 + \sum_{0\le i<j\le n}\frac{c^{(2)}(x_i,x_j,p_i,p_j)}{\rho^{(1)}(x_i,p_i)\rho^{(1)}(x_j,p_j)} +$$
$$+ \sum_{0\le i<j<k\le n}\frac{c^{(3)}(x_i,x_j,x_k,p_i,p_j,p_k)}{\rho^{(1)}(x_i,p_i)\rho^{(1)}(x_j,p_j)\rho^{(1)}(x_k,p_k)} +\ldots\Big)$$

and so our definition is that ρ is an equilibrium state if $\rho^{(1)}$ and all the relevant correlation densities $c^{(2)}$, $c^{(3)}$,... are locally stable. This leads to the following:

An equivalent "operational" definition ρ is a state of equilibrium if it is internally determined; meaning that idealized local observers, who are immersed in the gas (or liquid) and form a part of it, can determine its state “from within” .

One notorious local “observer” whose business is to measure local momentum fluctuations is Maxwell’s demon. However, there is a crucial difference between our observers and the demon in that our observers do not change the entropy of the system. All they are capable of is observing the velocity of the particles flying around them, store in memory the data concerning past observations, and perform simple calculations. Here is how it goes for pair correlations:

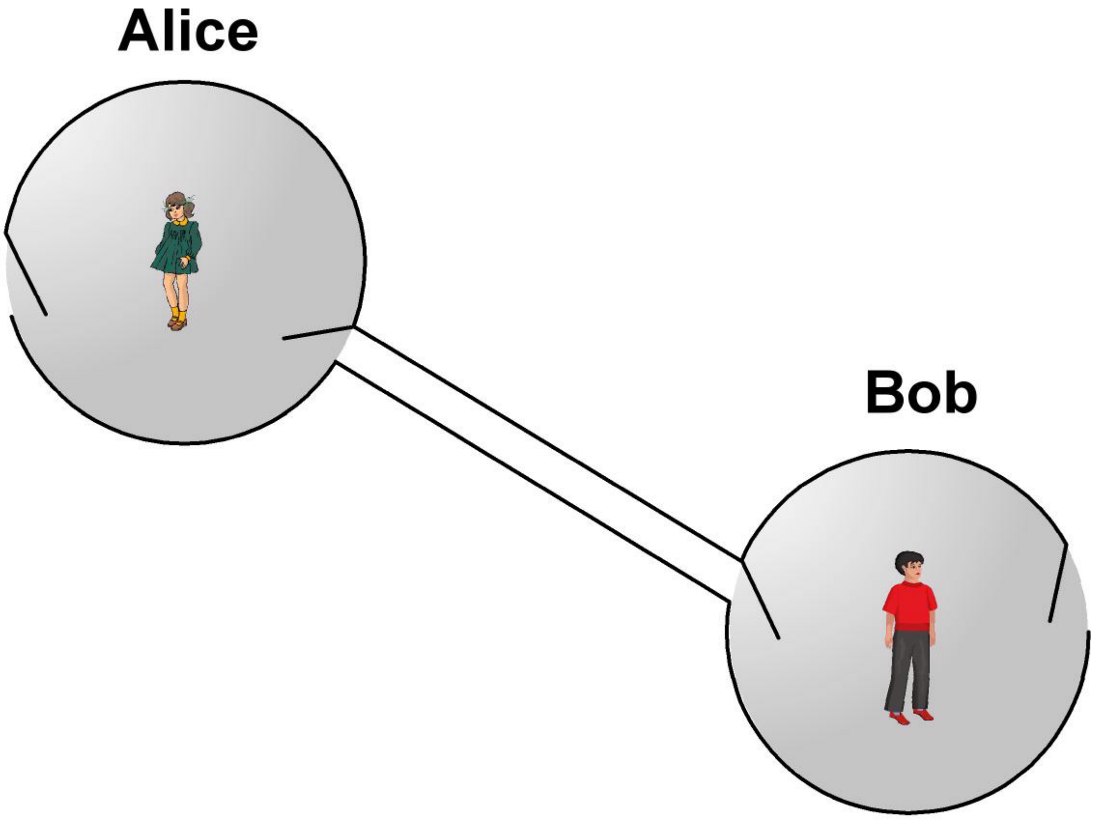

Figure 1

To illustrate the approach consider the magnitude $C^{(2)}(A_1,A_2,B_1,B_2)$, the correlation between the fluctuations in the (average) number of particles in A_1 with momentum values in B_1, and the fluctuations in the number of particles in A_2 having momentum in B_2. Assume that two local observers, call them Alice and Bob, are located in regions A_1 and A_2 respectively. Each of these regions is enclosed with solid walls and a gate with a door through which particles can come and go. The regions have a small but macroscopic dimensions, and are connected to one another by a pipe with doors at its ends.

To prepare their system Alice and Bob leave the doors connecting their cells to the outside world open, and the doors to the pipe connecting their cells closed. Now, suppose that at a certain moment Alice senses that the number of particles with the momentum range B_1 increases above average or decreases below average. She immediately closes the door to the outside world. Bob does the same if he senses a change above or below in the average number of particles with momenta in B_2. Suppose that $C^{(2)}(A_1,A_2,B_1,B_2) < 0$. In this case it is probable that an above average count in Alice's cell will correspond to a simultaneous below average count in Bob's cell and vice versa. Suppose, for example, that Alice's count is above average and Bob's count is below average. In this case, upon opening the doors to the pipe, the densities of particles of the kind B_1 and the kind B_2 will become more uniform. Therefore, Alice will observe particles of both kinds leaving her cell, and Bob will see particles of both kinds arrive at his cell. The number of such particles, which is known to both Alice and Bob, indicates the strength of correlation. Since we do not want our observers to change the distribution (and entropy), we assume that the entire affair is very swift, and takes no longer than it takes the fluctuation to recede normally. Similarly, if $C^{(2)}(A_1,A_2,B_1,B_2) > 0$, and Alice's count is above average and so is Bob's, particles of one kind B_1 will seem to flow from Alice to Bob, and of the other kind B_2 from Bob to Alice. Note that in order to “compare notes” the observers need no external means of communication. If the fluctuation correlations are non zero it provides for the information flow. If it is zero no information exchange is needed.