If more quantities are considered, process analogously.

Why is microcanonical ensemble suitable to describe an isolated mechanical system?

Def (ergodic): Let \((r, E, \mu)\) be a prob space, \(g^t: X \rightarrow X\) a measure-preserving flow.

\(g^t\) is ergodic over \(E\) if \(\forall E \in \Sigma\) with \(\mu(g^t(E))=\mu(E)=0\) for all \(t\) and \(\rho(E)\neq 0\). \(g^t\) is not decomposable and \(X\) reaches all of \(E\).

Thm (Birkhoff 1931): Let \((r, E, \mu)\) be prob space, \(g^t: X \rightarrow X\) a measure-preserving flow, \(f \in L^1(E)\).

Then \(\frac{1}{t} \int_0^t f(g^t(x)) dt \xrightarrow{a.e.} \hat{f}(x) \) for some \(\hat{f} \in L^1(E)\) with \(\int_X \hat{f} \, d\mu = \int_X f \, d\mu\).

\(\hat{f}\) is ergodic, \(\frac{1}{t} \int_0^t f(g^t(x)) dt \xrightarrow{a.e.} \hat{f}(x) \) for some \(\hat{f} \in L^1(E)\) with \(\int_X \hat{f} \, d\mu = \int_X f \, d\mu\).

Proof: (a) implies (b) directly.

(b): \(g^t\) measure-preserving \(\Rightarrow \forall E \in \Sigma\): \(\int_E f g^t \, d\mu = \int_E f \, d\mu\) for all \(t\).

proof: Take \(f = 1_E\) then \(\int_E f g^t = \mu(\sigma_g(E)) = \mu(E) = \int_E f \, d\mu\). Next show for simple fns. Next for \(L^1\) fns.

Maximal ergodic thm.: Let \(M(f) = \sup_{T \geq 0} \int_0^T f g^t(x) dt\), then \(\frac{M(f)}{T} \rightarrow 0\) as \(T \rightarrow \infty\).

proof: Let \(\mathcal{M}_n = \{x \in X\} \mid \sup_{0 \leq t \leq n} \int_0^T f g^t(x) dt > 0\) as \(n \rightarrow \infty\).

By dominated convergence, \(\int_X f g^t \, d\mu \rightarrow \int_X f \, d\mu\). Will prove \(\int_X f \, d\mu = 0\) a.e.

Fix \(n\) from now on.

\(\int_X f g^t \, d\mu = \int_X \mathcal{M}_n g^t \, d\mu = \int_X (f g^t) g^t \, d\mu \rightarrow \int_X f \, d\mu\) as \(T \rightarrow \infty\).

Will show: \(\exists C = C(n) > 0\) s.t. for all large enough \(T\), \(\int_0^T (f g^t) g^t \, d\mu > -C\).

Claim: For every \(X \subset E\), \(\exists C > 0\) s.t. \(\int_0^T (f g^t) g^t \, d\mu > 0\)

proof: Fix \(X \subset E\), take \(T > 0\) for which \(\mathcal{M}_n \subset X\), ok since \(T \rightarrow \infty\) as \(n \rightarrow \infty\).

Decompose \(T \int_0^T (f g^t) g^t \, d\mu = \int_0^T (f g^t) g^t \, d\mu + \int_0^T (f g^t) g^t \, d\mu\)

\(\rightarrow 0\)

\(\rightarrow 0\)

\(\rightarrow 0\)

\(\rightarrow 0\)

\(\rightarrow 0\)

(1) Show a.e. convergence of \(\frac{1}{t} \int_0^T f g^t \, d\mu\).

Let \(f^* = \lim sup \frac{1}{t} \int_0^T f g^t \, d\mu\) and \(f^* = \lim inf \frac{1}{t} \int_0^T f g^t \, d\mu\).

To derive a contradiction, suppose \(f^* \neq f^*\) a.e. A \(\subset \mathbb{R}\) is a \(f^* < f < f^* \subset f \subset C\) has pos. meas.

Note: \((f^*) = f \sup\) and \(f \inf\) consider \((f A) \subset A \subset \mathbb{R}\) as measure-preserving flow.

Apply mean erg. thm. to \((f A) \subset A\) with \(f^* = f < f\) \(\Rightarrow 0 = \int f \, d\mu \rightarrow \int f^* \, d\mu \geq 0\).

Analogously show \(\int f^* \, d\mu \geq 0\).

(2) Let \(f^* = f^* - f^*\) a.e.; show \(f^* \in L^1\) and \(\int f^* \, d\mu = \int f \, d\mu\).
Divide \( \mathcal{R} = (V_S, x_V) \), \( x_V \), disjoint intervals, \( \mu (f^\tau ((v_S), t)) = 0 \) \( 10^6 \). Let \( \mathcal{V} = (a, b, c) \) and \( f_1 = \{ a < b \} \). Let \( f \) be \( \gamma \)-morphism.

Then \( \gamma : \mu (f_1) \in \{ f : \mu (f_1) \} \).

Applying theorem to \( \gamma^\tau : f \rightarrow \gamma^\tau \) \( \Rightarrow \gamma : \mu (f_1) \in \{ f : \mu (f_1) \} \).

By Liouville's theorem, for an isolated mechanical system we know the phase flow \( \gamma^\tau \) is measure-preserving w.r.t. to the microcanonical distribution.

By Birkhoff's theorem, the time average \( \bar{f} \) exists a.e.

Furthermore, assuming the system to be ergodic (i.e., to reach every state), the time average equals the microcanonical phase space average by Birkhoff's theorem,

\[ \bar{f} = \langle f \rangle. \]

This is the major unsolved problem of statistical mechanics; it can only be proven for very few systems (most complex: hard spheres in a box). An important prerequisite is that the phase space is "fair" (otherwise, a uniform distribution of states does not make sense).

In view of these problems, one simply defines the microcanonical ensemble as that system which has exactly the prescribed measure \( \mu \) (and ignores whether that corresponds to the actual dynamics).

Note: The ensemble phase space \( \mathcal{D}(\mathcal{V}) \) is weighted in \( \mu \) according to the inverse velocity of trajectories \( \mathcal{V} \), i.e., if the velocities are high, trajectories don't stay in that region for long = less contribution to time (and thus also phase space) average.

**Coupled Systems**

Given: two systems \( A, B \) with states \( X_A, X_B \), \( X_A, X_B \), Hamiltionians \( H_A, H_B \),

- e.g., each with conserved quantities: energy \( E \), volume \( V \), molecular number \( N \) (consider only \( E \)).

The size of their accessible phase space is \( S_A(E, V, N) \), \( S_B(E, V, N) \).

Now bring the system into contact and let them exchange a quantity, e.g., energy (moves heat from a microcanonical ensemble).

**Assumption:** the systems are independent, i.e.,

\[ S_1(E_1, E_2, V_1, V_2, N_1, N_2) = S_A(E_1, V_1, N_1) S_B(E_2, V_2, N_2). \]

For the time being ignore fixed quantities \( V_1, V_2, N_1, N_2 \),

**probability of state** \( X_A \) **given total energy** \( E \): \( \text{prob}(X_A | H = E) = \frac{S_\mathcal{D}(E - H_A(X_A))}{Z(E)} \)

**probability of energy** \( E_A \) **given total energy** \( E \): \( \text{prob}(H_A = E_A | H = E) = \frac{S_A(E_A)S_B(E - E_A)}{Z(E)} \)

**normalization constant** \( Z(E) = Z(E) = \sum_{X_A, X_B, X_1, X_2} S_\mathcal{D}(E - E_A) \text{ for discrete phase space} \)

indeed: \( \# \) of states with \( X_A \) fixed \( = \# \) of states of \( B \) with energy \( E - H_A(X_A) = S_\mathcal{D}(E - H_A(X_A)) \)

**total # of states**: \( = \sum_{X_A} \# \) of states with \( X_A \) fixed \( = Z(E) \)

**# of states with** \( H_A = E_A \) \( = \sum_{X_A \text{ with } H_A = E_A} \# \) of states with \( X_A \) fixed \( = \sum_{X_A \text{ with } H_A = E_A} S_\mathcal{D}(E - E_A) = S_A(E_A)S_B(E - E_A) \).
\begin{itemize}
    \item cont. phase space, fixed $E$
    \item microcanonical measure on accessible phase space: $\text{d}\mu(X_A, X_B) = \frac{\text{d}\lambda(X_A, X_B)}{10H(X_A, X_B)}$
    \item total measure of accessible phase space $= \int_{H(E)} \frac{\text{d}\lambda(X_A, X_B)}{10H}$
    \item marginal measure for fixed $X_A$
    \begin{align*}
        \text{prob}(X_A, X_B, \delta X_A \mid H=E) &= \int \frac{\text{d}\lambda(X_A, X_B)}{10H} \\
        &= \int \frac{\text{d}\lambda(X_B)}{10H(X_A, X_B)} \cdot \delta X_A
    \end{align*}
    \item $\text{prob}(H_A = E_A \mid H=E) = \int \frac{\text{d}\lambda(X_A)}{10H} \cdot \delta E_A$

\end{itemize}

Ex: microscopic switches with $N_A = 2$, $N_B = 3$

\begin{center}
\begin{tabular}{|c|c|c|c|c|}
\hline
$X_A$ & $E_A$ & $E_B$ & $P(E_A)$ & $\text{Prob}(X_A \mid E)$ \\
\hline
(1,1) & 1 & 1 & 3 & 0.3 \\
(1,0) & 1 & 2 & 3 & 0.3 \\
(0,1) & 1 & 2 & 3 & 0.2 \\
(0,0) & 2 & 0 & 4 & 0.4 \\
\hline
\end{tabular}
\end{center}

Boltzmann entropy of a given system (e.g. with conserved quantities $E, V, N$):

\[
S(E,V,N) = k_B \sum_i P_i \ln P_i
\]

entropy is extensive: $S(E_A, E_B) = (V_A + V_B)(N_A + N_B)$

\[
S(E_A, E_B) = \ln [\Omega_A(E_A, N_A) \Omega_B(E_B, V_B, N_B)]
\]

If phase space continuous or systems large enough so entropy can be approximated by a smooth fn,

\[
S(E, V, N) = \frac{E}{k_B \beta} + \frac{V}{\beta} + \frac{N}{\beta M}
\]

Most likely $E_A$ in a coupled system $A, B$ while exchange energy:

\[
\text{argmax } \text{prob}(H_A = E_A \mid H=E) = \text{argmax } \left[ \Omega_A(E_A) \Omega_B(E-E_A) \right] = \text{argmax } \left[ S_A(E_A) + S_B(E-E_A) \right]
\]

\Rightarrow most likely macrostate found by maximizing entropy

"equilibrium" = observed macrostate \approx most likely state at $\frac{\partial S_A}{\partial E_A} = \frac{\partial S_B}{\partial E_B} \Rightarrow E_A = E_B$

\Rightarrow equilibrium temperatures are same!

Likewise, if coupled system additionally exchange volume or particles, pressure or chemical potential are also same as $A, B$!

Why we identify equilibrium with most likely state?

\[
\text{prob}(H_A = E_A \mid H=E) = \exp \left( \frac{S_A(E_A) + S_B(E-E_A) - S^*}{T} \right) \quad \text{for most likely } E_A^* \text{ and } S^*
\]

For large system sizes, $\exp^{-}$ becomes very negative for $E_A$ away from $E_A^*$

\Rightarrow $E_A^*$ becomes overwhelmingly likely