How does μ_E look like?

Let $D(E,\delta E) = H^{-1}([E, E+\delta E]) \subset \Gamma$.

$$\mu_E(A) \sim \lim_{\delta E \to 0} \frac{\int_{D(E,\delta E)} 1_A(X)\,dX}{|D(E,\delta E)|} \sim \lim_{\delta E\to 0} \frac{\int_{D(E)} \frac{\delta E}{|\nabla H(X)|} 1_A(X)\,d\alpha(X)}{\int_{D(E)} \frac{\delta E}{\nabla H(X)}\,d\alpha(X)} = \frac{1}{\Omega(E)} \int_{D(E)} \frac{1_A(X)}{|\nabla H(X)|}\,d\alpha(X)$$

$$\Omega(E) = \int_{D(E)} \frac{d\alpha(X)}{|\nabla H(X)|}$$

$$\Rightarrow d\mu = d\mu_E = \frac{1}{\Omega(E)} \frac{d\alpha(X)}{|\nabla H(X)|}$$ ($d\alpha$ stands for the $(2N-1)$-dimensional volume element)

(c) If more quantities are conserved, proceed analogously.

<u>Why is microcanonical ensemble suitable to describe an isolated mechanical system?</u>

<u>Def. (ergodic)</u>: Let (Γ, Σ, μ) be a prob. space, $g^t: \Gamma \to \Gamma$ a measure-preserving flow.
g^t is ergodic wrt μ if $\forall E \in \Sigma$ with $\mu((g^t)^{-1}(E) \Delta E) = 0$ for all t: $\mu(E)=0$ or $\mu(E)=1$.

ergodicity $\Rightarrow$ Γ is not decomposable; flow reaches all of Γ

<u>Thm (Birkhoff 1931)</u>: Let (Γ, Σ, μ) be prob. space, $g^t: \Gamma \to \Gamma$ a measure-preserving flow, $f \in L^1(\mu)$.
Then (a) $\frac{1}{T}\int_0^T f(g^t(X))\,dt \xrightarrow[T\to\infty]{a.e.} f^*(X)$ for some $f^* \in L^1(\mu)$ with $f^* \circ g^t = f^*$ & $\int_\Gamma f^* d\mu = \int_\Gamma f\,d\mu$

(b) if g^t is ergodic, $\frac{1}{T}\int_0^T f(g^t(X))\,dt \xrightarrow[T\to\infty]{a.e.} \int_\Gamma f(X)\,d\mu(X)$

proof: • (a) implies (b) directly

• (a): <u>Prop:</u> g^t measure-preserving $\Rightarrow$ $\forall f \in L^1(\mu)$: $\int_\Gamma f\circ g^t d\mu = \int_\Gamma f d\mu$ for all t.
proof: Take $f = \chi_A$, then $\int_\Gamma f\circ g^t d\mu = \mu((g^t)^{-1}(A)) = \mu(A) = \int_\Gamma f\,d\mu$; next show for simple fcns, next for L^1-fcns. □

<u>Maximal ergodic thm.</u>: Let $M(f)(X) = \sup_{T>0} \int_0^T f(g^t(X))\,dt$, then $\int_{\{M(f)\geq 0\}} f\,d\mu \geq 0$.

proof: Let $M_n = \{X \in \Gamma \mid \sup_{T\leq n} \int_0^T f\circ g^t(X)\,dt \geq 0\}$ $\Rightarrow$ $M_n \uparrow M_\infty = \{X\in\Gamma \mid M(f)(X) \geq 0\}$.

By dominated convergence, $\int_{M_n} f\,d\mu \longrightarrow \int_{M_\infty} f\,d\mu$. Will prove $\int_{M_n} f\,d\mu \geq 0$ $\forall n$.

Fix n from now on. $\int_{M_n} f\,d\mu = \int_\Gamma f\chi_{M_n}\,d\mu = \frac{1}{T}\int_\Gamma \int_0^T (f\chi_{M_n})\circ g^t\,dt\,d\mu$ $\forall T>0$.

Will show: $\exists C = C(n) > 0$ s.t. for all large enough T, $\int_\Gamma \int_0^T (f\chi_{M_n})\circ g^t\,dt\,d\mu \geq -C$.

Claim: For every X $\exists \tau(X)$, $1 \leq \tau(X) \leq n$ s.t. $\int_0^{\tau(X)} (f\chi_{M_n})(g^t(X))\,dt \geq 0$

proof: For $X \notin M_n$ take $\tau(X) = 1$, for $X \in M_n$ OK since $f\chi_{M_n} \geq f$ ($f \leq 0$ on $\Gamma \setminus M_n$) □

Decompose $\int_0^T (f\chi_{M_n})(g^t(X))\,dt = \underbrace{\int_0^{\tau(X)} (f\chi_{M_n})(g^t(X))\,dt}_{\geq 0} + \underbrace{\int_{\tau(X)}^{\tau(X)+\tau(g^{\tau(X)}(X))} (f\chi_{M_n})(g^t(X))\,dt}_{\geq 0}$

$+ \dots + \underbrace{\int \dots dt}_{\geq 0}$ + remainder $R(X)$

$$\Rightarrow \int_\Gamma \int_0^T (f\chi_{M_n})(g^t(X))\,dt\,d\mu(X) \geq \int_\Gamma R(X)\,d\mu(X) \geq -\int_0^n \int_\Gamma |f\chi_{M_n}|(g^{T-n+t}(X))\,d\mu(X)\,dt$$
$$\geq -n \int_\Gamma |f\chi_{M_n}|\,d\mu = -C. \quad \square$$

(1) Show a.e. convergence of $\frac{1}{T}\int_0^T f(g^t(X))\,dt$:

Let $f_+^* = \limsup_{T\to\infty} \frac{1}{T}\int_0^T f(g^t(X))\,dt$, $f_-^* = \liminf_{T\to\infty} \frac{1}{T}\int_0^T f(g^t(X))\,dt$.

To derive a contradiction, suppose $\exists a < b$ s.t. $A = \{f_+^* > b, f_-^* < a\} \subset \Gamma$ has pos. measure.

Note: $(g^t)^{-1}(A) = A$ $\forall t$, so consider $g_A^t: A \to A$ as measure-preserving flow.

Apply max. erg. thm. to g_A^t with $\tilde f = f - b$ $\Rightarrow \int_A \tilde f\,d\mu \geq 0 \Rightarrow \int_A f\,d\mu \geq b\,\mu(A)$

Analogously show $\int_A f\,d\mu < a\,\mu(A)$ ↯

(2) Let $f^* = f_+^* = f_-^*$ a.e.; show $f^* \in L^1(\mu)$ & $\int_\Gamma f^* d\mu = \int_\Gamma f\,d\mu$

Divide $\mathbb{R} = \bigcup_i I_i$, I_i disjoint intervals, $\mu(f^{-1}(\bigcup_i \partial I_i)) = 0$, $|I_i| < \varepsilon$.
Let $I_i = (a_i, b_i)$ and $F_i = \{a_i < f^* < b_i\} \subset \Gamma$ (g^t-invariant).
Then $a_i\, \mu(F_i) \le \int_{F_i} f^* \, d\mu \le b_i\, \mu(F_i)$. (1)
Apply max. erg. thm. to $g^t|_{F_i} : F_i \to F_i \Rightarrow a_i\, \mu(F_i) \le \int_{F_i} f \, d\mu \le b_i\, \mu(F_i)$ (2)
(1) & (2) $\Rightarrow$ $\left| \int_\Gamma f \, d\mu - \int_\Gamma f^* d\mu \right| < \varepsilon$ □
$|a_i - b_i| < \varepsilon$, $\sum \mu(F_i) = 1$

By Liouville's theorem, for an isolated mechanical system we know the phase flow g^t is measure-preserving wrt to the microcanonical distribution!
$\Rightarrow$ 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 thm,

$$\bar{f} = \langle f \rangle.$$

This assumption of ergodicity is the major unsolved problem of statistical mechanics; it could only be proven for very few systems (most complex: hard spheres in a box). Another important prerequisite is that the phase space is bounded (otherwise a uniform distribution on Γ does not make sense).
In view of these problems one simply defines the microcanonical ensemble as that system which has exactly the prescribed measure μ (and ignores whether that corresponds to the correct dynamics).

Note: The accessible phase space $D(E)$ is weighted in μ according to the inverse velocity of trajectories, $\frac{1}{|\nabla H|}$, i.e. if the velocities are high, trajectories don't stay in that region for long $\Rightarrow$ less contribution to time (and thus also phase space) average.

Coupled Systems

Given: two systems A & B with states $X_A \in \Gamma_A$, $X_B \in \Gamma_B$, Hamiltonians H_A, H_B,
e.g. each with conserved quantities: energy E, volume V, molecule number N (thinking of gas in isolated box)
the size of their accessible phase space is $\Omega_A(E_A, V_A, N_A)$, $\Omega_B(E_B, V_B, N_B)$
Now bring the systems into contact and let them exchange a quantity, e.g. energy (imagine heat flux through wall)
Together they form a microcanonical ensemble.
Assumption: • conserved quantities are "extensive" or "additive", i.e. for coupled system we have

$$E = E_A + E_B \quad (\text{or } H = H_A + H_B), \quad V = V_A + V_B, \quad N = N_A + N_B$$

• the systems are independent, i.e.

$$\underset{\uparrow \text{ \# of states with these macroscopic quantities}}{\Omega((E_A, E_B), (V_A, V_B), (N_A, N_B))} = \Omega_A(E_A, V_A, N_A)\, \Omega_B(E_B, V_B, N_B)$$

For the time being ignore fixed quantities V_A, V_B, N_A, N_B.
probability of state X_A given total energy E: $\text{prob}(X_A | H = E) = \frac{\Omega_B(E - H_A(X_A))}{Z(E)}$

probability of energy E_A given total energy E: $\text{prob}(H_A = E_A | H = E) = \frac{\Omega_A(E_A)\, \Omega_B(E - E_A)}{Z(E)}$

normalization constant $Z(E) = \Omega(E) = \begin{cases} \sum_{X_A \in \Gamma_A} \Omega_B(E - E_A) & \text{for discrete phase space} \\ \int_{H^{-1}(E)} \frac{d\alpha(X_A, X_B)}{|\nabla H|} & \text{for cont. phase space} \end{cases}$

$$\nabla H = \nabla_{(X_A, X_B)} (H_A + H_B) = \begin{pmatrix} \nabla_{X_A} H_A \\ \nabla_{X_B} H_B \end{pmatrix}$$

Indeed: → discrete phase space, fixed total energy E:
• # of states with X_A fixed = # of states of B with energy $E - H_A(X_A) = \Omega_B(E - H_A(X_A))$
• 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 $= \sum_{X_A \text{ with } H_A = E_A} \Omega_B(E - E_A) = \Omega_A(E_A)\, \Omega_B(E - E_A)$

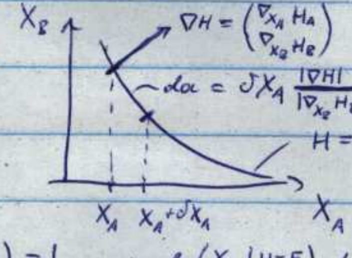
→ cont. phase space, fixed E:

- microcanonical measure on accessible phase space: $d\mu(X_A, X_B) = \frac{da(X_A, X_B)}{|\nabla H(X_A, X_B)|}$
- total measure of accessible phase space $= \int_{H^{-1}(E)} \frac{da(X_A, X_B)}{|\nabla H|}$
- marginal measure for fixed X_A

$$\text{prob}([X_A, X_A+\delta X_A] \mid H=E) = \int_{H^{-1}(E)\cap[X_A, X_A+\delta X_A]\times\Gamma_B} \frac{da(X_A, X_B)}{|\nabla H|}$$

$$\doteq \int_{H^{-1}(E)\cap\{X_A\}\times\Gamma_B} \frac{da(X_B)}{|\nabla H(X_A, X_B)|} \cdot \frac{\delta X_A \, |\nabla H(X_A, X_B)|}{|\nabla_{X_B} H_B|}$$

$$= \int_{H^{-1}(E)\cap\{X_A\}\times\Gamma_B} \frac{da(X_B)\,\delta X_A}{|\nabla_{X_B} H_B(X_B)|} = \Omega_B(E - H_A(X_A))\,\delta X_A$$



- $\text{prob}(H_A = E_A \mid H=E) = \int_{H_A^{-1}(E_A)} \underbrace{\text{prob}(X_A \mid H=E)}_{=\Omega_B(E-H_A(X_A))} \underbrace{d\mu_A(X_A)}_{\frac{da(X_A)}{|\nabla_{X_A} H_A|}} = \Omega_B(E-E_A)\int_{H_A^{-1}(E_A)} \frac{da(X_A)}{|\nabla_{X_A} H_A|} = \Omega_B(E-E_A)\,\Omega_A(E_A)$

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

for $E = 2$:

X_A	E_A	E_B	$\Omega_B(E_B)$	Prob $(X_A \mid E)$
(0,0)	0	2	3	0.3
(0,1)	1	1	3	0.3
(1,0)	1	1	3	0.3
(1,1)	2	0	1	0.1

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

$$S(E,V,N) = \ln \Omega(E,V,N) \quad \text{(measure of possibilities / undeterminedness / disorder)}$$

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

$$= \ln\left[\Omega_A(E_A, V_A, N_A)\,\Omega_B(E_B, V_B, N_B)\right] = S_A(E_A, V_A, N_A) + S_B(E_B, V_B, N_B)$$

If phase space continuous or systems large enough s. t. entropy can be approximated by a smooth fcn, define

$\frac{\partial S}{\partial E} = \beta = \frac{1}{\theta}$; β = inverse temperature (temp. in Kelvin is given by $k_B / \frac{\partial S}{\partial E}$ for Boltzmann's constant k_B)

$\frac{\partial S}{\partial V} = \beta p$; p = pressure

$\frac{\partial S}{\partial N} = \beta M$; M = chemical potential

Most likely E_A in a coupled system A, B which exchanges energy:

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

⇒ most likely macrostate found by maximizing entropy

"equilibrium" = observed macrostate ≈ most likely state at $\left.\frac{\partial S_A}{\partial E}\right|_{E_A} = \left.\frac{\partial S_B}{\partial E}\right|_{E_B = E - E_B}$, i.e. $\beta_A = \beta_B$ or $\theta_A = \theta_B$

⇒ in equilibrium temperatures are same!

Likewise, if coupled system additionally exchanges volume or particles, pressure or chemical potential are also same in A & B!

Why may we identify equilibrium with most likely state?

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

For large system sizes, exponent becomes very negative for E_A away from E_A^*
⇒ E_A^* becomes overwhelmingly likely