Some Basic Concepts in the Foundations of Quantum Statistical Mechanics

Roderich Tumulka

Department of Mathematics

RUTGERS UNIVERSITY

Genova, 23 May 2013
Quantum mechanics

For system of $N$ particles, wave function $\psi_t : \mathbb{R}^{3N} \to \mathbb{C}$

Schrödinger equation $i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi$

with $\hat{H} = \text{Hamilton operator}$

$$\hat{H}\psi = -\sum_{k=1}^{N} \frac{\hbar^2}{2m_k} \nabla_k^2 \psi + V\psi$$

$V(q_1, \ldots, q_N) = \text{potential energy, } q \in \mathbb{R}^3$

- e.g., $V(q_1, \ldots, q_N) = \sum_{j=1}^{N} U(q_j) + \sum_{1 \leq j < k \leq N} W(q_j - q_k)$

- $U = \text{external potential, e.g., } U(q) = \begin{cases} -mgq_3 & \text{if } q \in \Lambda \\ \infty & \text{if } q \notin \Lambda \end{cases}$
  (i.e., confined to a bounded volume $\Lambda \subset \mathbb{R}^3$)

- $W = \text{interaction potential, e.g., Coulomb potential } W(q) = \frac{1}{|q|}$
Quantum mechanics, applied to a macro-system

- Consider macroscopic quantum system $S$ (say, $N > 10^{20}$ particles)
- isolated, evolves unitarily according to the Schrödinger equation
- confined to a bounded volume $\Lambda \subset \mathbb{R}^3$
- as a consequence, $\hat{H}$ has pure point spectrum (i.e., no continuous spectrum)
- $\hat{H} = \sum_\alpha E_\alpha |\phi_\alpha\rangle \langle \phi_\alpha|$
A chapter of math: the trace

- The trace of an \( n \times n \) matrix \( A \) is \( \text{tr} A = \sum_{i=1}^{n} A_{ii} \).
- The trace of an \( \infty \times \infty \) matrix \( A \) is \( \text{tr} A = \sum_{i=1}^{\infty} A_{ii} \).
- The series may or may not converge.
- For an operator \( \hat{A} \) on Hilbert space \( \mathcal{H} \), choose any orthonormal basis (ONB) \( B = \{ b_i : i = 1, 2, \ldots \} \) and define \( \text{tr} \hat{A} = \sum_i \langle b_i | \hat{A} | b_i \rangle \).
- If \( \dim \mathcal{H} = \infty \), then the trace is ill-defined (e.g., infinite) for some operators. It is well defined for all operators in the trace class \( \{ \hat{A} : \text{tr}(\sqrt{\hat{A}^* \hat{A}}) < \infty \} \).
- For \( \hat{A} \in \text{trace class} \), \( \text{tr} \hat{A} \) does not depend on the choice of ONB.
- Note that
  - If \( \hat{A} \) is diagonalizable (e.g., self-adjoint with pure point spectrum), then \( \text{tr}(\hat{A}) \) is the sum of the eigenvalues.
  - \( \text{tr}(\hat{A} \hat{B} \cdots \hat{Y} \hat{Z}) = \text{tr}(\hat{Z} \hat{A} \hat{B} \cdots \hat{Y}) \) (invariant under cyclic permutation)
  - \( \text{tr} \) is linear
  - \( \text{tr}(\hat{A} | \psi \rangle \langle \psi |) = \langle \psi | \hat{A} | \psi \rangle \)
  - \( \text{tr}(\hat{A}^*) = \text{tr}(\hat{A})^* \)
Density operator vs. distribution of the wave function

- Hilbert space $\mathcal{H}$, e.g., $L^2(\mathbb{R}^3, \mathbb{C})$
- unit sphere $S(\mathcal{H}) = \{ \psi \in \mathcal{H} : \|\psi\| = 1 \}$
- probability distribution $\mu$ on $S(\mathcal{H})$
- equivalence relation $\mu \cong \nu : \Leftrightarrow$ empirically indistinguishable

$\Leftrightarrow \hat{\rho}_\mu = \hat{\rho}_\nu$, where 
\[
\hat{\rho}_\mu = \int_{S(\mathcal{H})} \mu(d\psi) \langle \psi | \phi \rangle \langle \phi | \psi \rangle
\]
is the pertaining density operator (DO, also known as density matrix) (= covariance operator of $\mu$, provided $E_\mu |\psi\rangle = 0$.)

Proof: For every conceivable yes-no experiment there is a positive operator $\hat{E}$ (i.e., self-adjoint with spectrum in $[0, \infty)$) such that $P(\text{yes}) = \langle \psi | \hat{E} | \phi \rangle$. Thus, if $\psi$ has distribution $\mu$, $P(\text{yes}) = \int_{S(\mathcal{H})} \mu(d\psi) \langle \psi | \hat{E} | \phi \rangle = \int_{S(\mathcal{H})} \mu(d\psi) \operatorname{tr}(\hat{E} |\psi\rangle \langle \psi|) = \operatorname{tr}(\hat{E} \hat{\rho}_\mu)$. $\square$
Density operators

\[ \hat{\rho}_\mu = \int_{\mathbb{S}(\mathcal{H})} \mu(d\psi) |\psi\rangle \langle \psi| \]

- \( \hat{\rho} \) is a DO \( \iff \hat{\rho} \) is positive with \( \text{tr}(\hat{\rho}) = 1 \).
- Proof: Check that \( \langle \phi|\hat{\rho}_\mu|\phi\rangle \geq 0 \) and \( \text{tr}(\hat{\rho}_\mu) = 1 \).

Conversely, for a positive trace-one operator \( \hat{\rho} \), diagonalize it, \( \hat{\rho} = \sum_n p_n |\chi_n\rangle \langle \chi_n| \). The eigenvalues \( p_n \geq 0 \) add up to 1. Define distribution \( \mu \) so that \( \mu\{\chi_n\} = p_n \) and otherwise 0. Then \( \hat{\rho}_\mu = \hat{\rho} \). □

- If \( \mu \) is concentrated on a single vector, \( \mu\{\psi\} = 1 \), then \( \hat{\rho}_\mu = |\psi\rangle \langle \psi| \) ("pure state").
- \( \mu \mapsto \hat{\rho}_\mu \) is not injective, but many-to-one: \( \hat{\rho}_\mu = \hat{\rho}_\nu \not\Rightarrow \mu = \nu \). Ex:
  1. If \( D := \text{dim} \mathcal{H} < \infty \) then there exists a uniform probability distribution \( u = u_{\mathbb{S}(\mathcal{H})} \) proportional to surface area on \( \mathbb{S}(\mathcal{H}) \). Then \( \rho_u = D^{-1} \hat{1} \) with \( \hat{1} = \text{identity operator} \).
  2. Let \( B = \{b_n : n = 1, \ldots, D\} \) be an orthonormal basis of \( \mathcal{H} \). Let \( \mu \) give probability \( D^{-1} \) to each \( b_n \). Then \( \rho_\mu = D^{-1} \hat{1} \) (independently of the choice of \( B \)).
Use density operator

for calculating probability distribution of the result \( Z \) of an experiment:

\[
\mathbb{P}(Z = z) = \text{tr}(\hat{\rho} \, \hat{E}(z)),
\]

with \( \hat{E}(z) = \) positive operators such that

\[
\sum_z \hat{E}(z) = \hat{I}
\]

Ex: \( \hat{E}(z) = \) projection to the eigenspace of observable \( \hat{A} \) with eigenvalue \( z \)

Use distribution

- if every system of an ensemble has a wave function
- for typicality statements that hold for most wave functions (relative to \( \mu \)):
  \[
  \mu\{\omega : p(\omega)\} > 1 - \varepsilon
  \]
Some more math: Gaussian distribution in 2 dimensions

\[ \rho(x, y) = \frac{1}{Z} e^{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}} \]

E.g., 2 independent Gaussian-distributed random variables \( X, Y \)

If \( \sigma_x = \sigma_y \), then rotationally symmetric (otherwise elliptical)
Gaussian distribution in 3 dimensions

\[ \rho(x, y, z) = \frac{1}{Z} e^{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} - \frac{z^2}{2\sigma_z^2}} \]

or, if the orthonormal basis (ONB) gets rotated by \( \hat{R} \in SO(3) \):

\[ \rho(x) = \rho(x_1, x_2, x_3) = \frac{1}{Z} \exp\left(-\frac{1}{2}(x_1, x_2, x_3) \hat{P} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right) = \frac{1}{Z} e^{-\frac{1}{2}\langle x|\hat{P}x\rangle} \]

mit \( \langle x|y \rangle = \sum_{i=1}^{3} x_i y_i \) und \( \hat{P} = \hat{R}^{-1} \begin{pmatrix} \sigma_x^{-2} & 0 & 0 \\ 0 & \sigma_y^{-2} & 0 \\ 0 & 0 & \sigma_z^{-2} \end{pmatrix} \hat{R} \)

Thus, \( \hat{P} \) is symmetric and positive definite, otherwise arbitrary. Covariance matrix \( \hat{C} = (C_{ij})_{ij} \), \( C_{ij} = \mathbb{E}(X_iX_j) \). We have that \( \hat{P} = \hat{C}^{-1} \).

**Example:** Maxwell’s distribution of velocities in the ideal gas

\[ \rho(v) = \frac{1}{Z} \exp\left(-\frac{m|v|^2}{2kT} \right) \]

\( m = \text{mass/molec.} \), \( T = \text{abs. temp.} \), \( k = \text{Boltzmann const.} \approx 10^{-23} \text{ J/K} \)
Macro-states and thermal equilibrium
State: point \( X = (q_1, \ldots, q_N, p_1, \ldots, p_N) \) in phase space

energy shell
\( \Gamma = \{ X : E \leq H(X) \leq E + \delta E \} \)

depending on a choice of macro-variables, partition \( \Gamma \) into macro-states \( \Gamma_\nu \)
corresponding to different (small ranges of) values of the macro-variables,

\[
\Gamma = \bigcup_\nu \Gamma_\nu.
\]

Often, one cell \( \Gamma_{eq} \) has the overwhelming majority of volume,

\[
\frac{\text{vol } \Gamma_{eq}}{\text{vol } \Gamma} \approx 1.
\]

**Def:** A system is in equilibrium \( \iff \) its phase point lies in the set \( \Gamma_{eq} \).
Example: gas in a box (1)

$N \approx 10^{23}$ tiny billiard balls in a box, moving according to Newtonian mechanics (uniform motion between collisions, perfect reflection at collision between balls or ball–wall).

Macro-state

Subdivide box into cells $C_i$ (say, of size $(1 \text{ mm})^3$).
Subdivide velocity space into cells $V_j$.
For each $C_i \times V_j$, specify number of balls (rounded to billions) with position in $C_i$ and velocity in $V_j$.  

Ludwig Boltzmann (1844–1906)
Every macro-state corresponds to a subset $\Gamma_\nu$ of phase space $\mathbb{R}^{6N}$.

Basic facts:

- Macro-states have very different volumes.
- Most micro-states at a given energy lie in the macrostate $\Gamma_{eq}$ of thermal equilibrium, characterized by an empirical distribution of particles that is uniform in position and Gaussian in velocity.
- Most $X \in \Gamma_{eq}$ stay in $\Gamma_{eq}$ for a VERY long time (of order $10^{10^{10}}$ years).
- The choice of macro-variables and ranges is somewhat subjective: Different physicists may draw the cells $\Gamma_\nu$ somewhat differently. But this is no practical difficulty when $N$ is large.
An even simpler example: gas without interaction in a box $\Lambda \subset \mathbb{R}^3$, phase space $= \Lambda^N \times \mathbb{R}^{3N} \subseteq \mathbb{R}^{6N}$

Hamiltonian function $H(q, p) = \sum_{k=1}^{N} \frac{1}{2m} |p_k|^2$

Energy surface $H^{-1}(E) = \Lambda^N \times S_R(\mathbb{R}^{3N})$ for $R = \sqrt{2mE}$, where $S_R(\mathbb{R}^{3N}) = $ sphere of radius $R$

**Typicality theorem**

For most phase points on $S_R(\mathbb{R}^{3N})$ with large $N$, the empirical distribution of velocities is close to the Maxwellian (Gaussian) distribution.
energy shell:
Let \( [E, E + \delta E] \) be an energy interval that is small on the microscopic scale but contains many eigenvalues \( E_\alpha \) of \( \hat{H} \). Usually, the eigenvalues are VERY dense; typical separation of order \( 10^{-10^10} \) J. So, \( [E, E + \delta E] \) usually contains of order \( 10^{10^10} \) eigenvalues.

An energy shell is \( \mathcal{H}_{[E, E+\delta E]} = \text{span}\{ \phi_\alpha : E \leq E_\alpha \leq E + \delta E \} \), the spectral subspace of \( \hat{H} \) for the interval \( [E, \delta E] \).

microcanonical density operator:

\[
\hat{\rho}_{mc} = \frac{1}{Z} 1_{[E, E+\delta E]}(\hat{H}) = \frac{1}{Z} \hat{P}_{[E, E+\delta E]}
\]

with \( Z = \text{normalization constant} = \text{tr} \ \hat{P}_{[E, E+\delta E]} = \text{dim} \ \mathcal{H}_{[E, E+\delta E]} \)

microcanonical ensemble of wave functions:
uniform distribution \( u \) on \( \mathcal{S}(\mathcal{H}_{[E, E+\delta E]}) \)
Thermal equilibrium in quantum mechanics

- State: wave fct $\psi = \psi(q_1, \ldots, q_N)$, $||\psi|| = 1$.
- Hamiltonian $\hat{H} = \sum_\alpha E_\alpha |\phi_\alpha\rangle\langle\phi_\alpha|$.

- Energy shell $\mathcal{H} = \text{span}\left\{ \phi_\alpha : E \leq E_\alpha \leq E + \delta E \right\}$. $\dim \mathcal{H} \approx 10^{10}$
- Macro-states correspond to subspaces $\mathcal{H}_\nu$, mutually orthogonal, 

$$\mathcal{H} = \bigoplus_\nu \mathcal{H}_\nu$$

- Thermal equilibrium subspace $\mathcal{H}_{eq} \subset \mathcal{H}$ with 

$$\frac{\dim \mathcal{H}_{eq}}{\dim \mathcal{H}} \approx 1$$

- **Def:** A system is in thermal equilibrium $\iff$ $\psi$ is close to $\mathcal{H}_{eq} \iff$ 

$$\langle \psi | \hat{P}_{eq} | \psi \rangle \approx 1$$

(Where $\hat{P}_{eq}$ = projection to $\mathcal{H}_{eq}$)

- Equilibrium is typical: $\langle \psi | \hat{P}_{eq} | \psi \rangle \approx 1$ for most $\psi$.
  Proof: $\mathbb{E}_\psi \langle \psi | \hat{P}_{eq} | \psi \rangle = \dim \mathcal{H}_{eq}/ \dim \mathcal{H} \approx 1$. □
How are the macro-states defined?

There is no precise, simple, universally valid prescription. Maybe in the limit $N \to \infty$ in special cases (e.g., dilute gas with short-range interaction) there is a precise prescription.

Use judgement. Different physicists will choose the $\Gamma_\nu$, $\mathcal{H}_\nu$ differently. But we expect that this makes no significant difference for the inferences we draw.
A proposal for the macro-spaces $\mathcal{H}_\nu$

- $\mathcal{H} = \bigoplus \nu \mathcal{H}_\nu$
- $\mathcal{H}_\nu$ correspond to different macro-states
- proposed construction: [von Neumann, book (1932)]

- choose macro-observables $\hat{M}_i$
- choose ranges of values and obtain coarse-grained macro-observables $\hat{M}_i'$
- they don’t commute but commutator is small
- find nearby operators $\hat{M}_i''$ that do commute mathematically non-trivial, see
  - [Choi: Proc. AMS 102 (1988)],
  - [Ogata: arXiv 1111.5933 (2011)]
- the joint spectral subspaces are the $\mathcal{H}_\nu$

- examples of macro-observables: particle number and/or total energy/ momentum/magnetization of a region $C_i$ of space that is not too small
Consider, for a gas consisting of \( N > 10^{20} \) atoms enclosed in a box \( \Lambda \subset \mathbb{R}^3 \), the following 51 macro-spaces \( \mathcal{H}_0, \mathcal{H}_2, \mathcal{H}_4, \ldots, \mathcal{H}_{100} \): \( \mathcal{H}_\nu \) contains the quantum states for which the number of atoms in the left half of \( \Lambda \) lies between \( \nu - 1 \) percent of \( N \) and \( \nu + 1 \) percent of \( n \). Note that, in this example, \( \mathcal{H}_{50} \) has the overwhelming majority of dimensions.

(Actually, these subspaces form an orthogonal decomposition of \( \mathcal{H}_{\text{total}} \) rather than of the energy shell \( \mathcal{H} \), since the operator of particle number in the left half of \( \Lambda \) fails to map \( \mathcal{H} \) to itself. Thus, certain approximations are necessary in order to obtain an orthogonal decomposition of \( \mathcal{H} \).)
Example of $\mathcal{H}_\nu$ (2)

- $N = 10^{23}$ particles, weakly interacting,
- $\mathcal{H}_1 = 1$-particle Hilbert space,
- $\mathcal{H} = \text{Sym} \mathcal{H}_1 \otimes N$ or Anti $\mathcal{H}_1 \otimes N$,
- $\hat{H}_1 = 1$-particle Hamiltonian $= \sum_\alpha E_{1,\alpha} |\phi_{1,\alpha}\rangle \langle \phi_{1,\alpha}|$,
- $\hat{H} = \sum_{i=1}^N \hat{I} \otimes \cdots \otimes \hat{H}_1 \otimes \cdots \otimes \hat{I} + \hat{V}$

Partition of the 1-particle energy axis $I_k = [E_k, E_{k+1})$,

- $\mathcal{H}_{1,k} = \text{range } 1_{I_k}(\hat{H}_1) = \text{span}\{\phi_{1,\alpha} : E_k \leq E_{1,\alpha} < E_{k+1}\}$
- $\hat{N}_k = \text{occupation number operator of } \mathcal{H}_{1,k}$ (an operator on $\mathcal{H}$!)
- the $\hat{N}_k$ commute
- coarse-graining fct $f(n) = 10^9[n/10^9]$, $[\cdot] = \text{nearest integer}$
- macro observables $\hat{M}_k = f(\hat{N}_k)$; the $\hat{M}_k$ commute
- $\nu = (\ldots, m_k, \ldots)$, $\mathcal{H}_\nu = \text{joint eigenspace of the } \hat{M}_k$
Approach to thermal equilibrium
Most $X \in \Gamma_\nu$, $\nu \neq \text{eq}$, move (in both time directions) to bigger macro-states.
The Poincaré recurrence theorem

Poincaré recurrence theorem, in a version for QM

Suppose that $\hat{H}$ has pure point spectrum. For every $\varepsilon > 0$ and $T > 0$ there exist (infinitely many) times $t > T$ such that

$$\| e^{-i\hat{H}t} - \hat{I} \| < \varepsilon.$$

As a consequence, for all $\psi_0 \in \mathcal{H}$ with $\|\psi_0\| = 1$,

$$\|\psi_t - \psi_0\| < \varepsilon.$$

If dim $\mathcal{H} < \infty$, then automatically $\hat{H}$ has pure point spectrum. Also, if an isolated system is confined to a finite volume $\Lambda \subset \mathbb{R}^3$ then its $\hat{H}$ has pure point spectrum.
It is impossible that \( \langle \psi_t | \hat{P}_{eq} | \psi_t \rangle \approx 1 \) for all \( t \to \infty \), by Poincaré recurrence. Instead, *approach to equilibrium* means that \( \langle \psi_t | \hat{P}_{eq} | \psi_t \rangle \approx 1 \) for most \( t \).

### Claim

For most \( \hat{H} \) it is the case that for every initial state \( \psi_0 \), the system will spend *most of the time* (as \( t \to \infty \)) in thermal equilibrium.
Let \( 1 \ll \text{dim } \mathcal{H} < \infty \) (think of \( \mathcal{H} = \) an energy shell of a closed system), \( \mathcal{H}_{\text{eq}} \subset \mathcal{H} \) a subspace with \( \frac{\text{dim } \mathcal{H}_{\text{eq}}}{\text{dim } \mathcal{H}} > 1 - \varepsilon, \ \varepsilon \ll 1 \).

**Theorem 1** [Goldstein, Lebowitz, Mastrodonato, Tumulka, Zanghì: *Phys.Rev.E* 81 (2010)]

If \( \hat{H} = \sum_{\alpha} E_\alpha |\phi_\alpha\rangle \langle \phi_\alpha| \) is non-degenerate (i.e., \( E_\alpha \neq E_\beta \)) and all eigenvectors \( \phi_\alpha \) are in thermal equilibrium

then any \( \psi_0 \in \mathcal{H} \) with \( \|\psi_0\| = 1 \) will, under the Schrödinger evolution, spend most of the time as \( t \to \infty \) in thermal equilibrium, i.e.,

\[
\liminf_{T \to \infty} \frac{1}{T} \left\{ 0 < t < T : \langle \psi_t | \hat{P}_{\text{eq}} | \psi_t \rangle > 1 - 2\varepsilon \right\} > 1 - \delta'.
\]

\(|M| = \) Lebesgue measure (length) of \( M \)

Recall: \( \psi \) in thermal equilibrium \( \iff \psi \in U_{\sqrt{2\varepsilon}}(\mathcal{H}_{\text{eq}}) \iff \langle \psi | \hat{P}_{\text{eq}} | \psi \rangle \approx 1. \)
Proof

time average \( \bar{f}(t) = \lim_{T \to \infty} \frac{1}{T} \int_0^T f(t) \, dt \)

\[
\langle \psi_t | \hat{P}_{eq} | \psi_t \rangle = ?
\]

\[
\psi_0 = \sum_{\alpha=1}^{\dim \mathcal{H}} c_\alpha | \phi_\alpha \rangle,
\psi_t = \sum_{\alpha=1}^{\dim \mathcal{H}} e^{-iE_\alpha t} c_\alpha | \phi_\alpha \rangle
\]

\[
\langle \psi_t | \hat{P}_{eq} | \psi_t \rangle = \sum_{\alpha, \beta} \underbrace{e^{i(E_\alpha - E_\beta) t}}_{\delta_{\alpha \beta}} c_\alpha^* c_\beta \langle \phi_\alpha | \hat{P}_{eq} | \phi_\beta \rangle
\]

\[
= \sum_{\alpha} |c_\alpha|^2 \underbrace{\langle \phi_\alpha | \hat{P}_{eq} | \phi_\alpha \rangle}_{>1-\eta \delta'} > 1 - \eta \delta'
\]

If error(\(t\)) > \(\eta\) for more than \(\delta'\) of the time then error(\(t\)) > \(\eta \delta'\).

Thus, \(\langle \psi_t | \hat{P}_{eq} | \psi_t \rangle > 1 - \eta\) for \((1 - \delta')\)-most of the time. \(\square\)
“Most” Hamiltonians satisfy non-degeneracy.

**Theorem 2**  [Goldstein, Lebowitz, Mastrodonato, Tumulka, Zanghì: *Phys.Rev.E* 81 (2010)]

“Most” non-degenerate Hamiltonians satisfy the “eigenstate thermalization hypothesis” (1).

- This is the general reason for approach to equilibrium.
- However, exceptional Hamiltonians exist (e.g., Anderson localization).

**Meaning of “most” in Theorem 2**

Fix the (non-degenerate) eigenvalues $E_\alpha$. Consider the uniform measure over all ONBs for $\phi_\alpha$. 
Probability, typicality, universality
Talking about “probability” is just a way of saying what’s true of most Hamiltonians; doesn’t mean that $\hat{H}$ is so distributed in nature.
What does “typical” mean? (1)

The statement $p(\omega)$ holds for typical $\omega$

$\iff$

$\mu\{\omega : p(\omega)\} > 1 - \varepsilon$

“overwhelming majority”

Ex: weak law of large numbers

For at least $1 - 10^4/n$ of all real numbers between 0 and 1, the relative frequency of the digit 7 among the first $n$ decimal digits lies between 9 and 11%. ($\mu = \text{Lebesgue measure}$)

slight abuse of language: “typical number”
What does “typical” mean? (2)

\[ f(\omega) \approx y \] is typical for \( \omega \in \Omega \)
\[ \iff y \text{ is the typical value of } f \text{ on } \Omega \]
\[ \iff f(\omega) \approx y \text{ for most } \omega \in \Omega \]
\[ \iff \mu\{\omega \in \Omega : |f(\omega) - y| > \delta\} < \varepsilon \]
\[ \iff f \text{ is nearly constant on } \Omega \text{ (in fact, } y \approx \mathbb{E}f\text{)} \]
\[ \iff f \text{ has small variance } \mathbb{E}(f - \mathbb{E}f)^2 \]
\[ \iff \text{more precisely, consider} \]
\[ \text{sequence } (\Omega_n, \mu_n)_{n \in \mathbb{N}} \text{ of probability spaces, } f_n : \Omega_n \to Y \]
\[ \forall \delta > 0 : \mu_n\{\omega \in \Omega_n : |f_n(\omega) - y| < \delta\} \to 1 \]

Slightly different: convergence in probability
(here, \( \Omega_n = \Omega, \mu_n = \mu \), only \( f_n \) varies,
convergence to \( f_\infty \) rather than constant \( y \); otherwise the same)

**Ex:** In classical statistical mechanics, thermodynamic functions are often nearly constant on the energy surface.

\( \omega \) doesn’t have to be random.

It suffices that \( \omega \) is not special.
Canonical typicality
A chapter of math: tensor products

- There is an operation $\otimes$ with Hilbert spaces such that, for configuration spaces $X$, $Y$,

$$L^2(X \times Y) = L^2(X) \otimes L^2(Y).$$

- $\psi(x, y)$

- If system 1 has $\mathcal{H}_1$ and system 2 has $\mathcal{H}_2$ then $\mathcal{H}_{1\cup 2} = \mathcal{H}_1 \otimes \mathcal{H}_2$.

- There is also an operation $\otimes$ on functions:

$$\psi = f \otimes g \text{ means } \psi(x, y) = f(x) g(y).$$

Note that most $\psi \in L^2(X \times Y)$ are not product functions. But each is a sum of product functions, $\psi = \sum_i f_i \otimes g_i$.

- There is also an operation $\otimes$ on operators:

$$(\hat{U} \otimes \hat{V}) \left( \sum_i f_i \otimes g_i \right) = \sum_i (\hat{U} f_i) \otimes (\hat{V} g_i).$$

- For non-interacting systems, $\hat{U}_{1\cup 2}(t) = \hat{U}_1(t) \otimes \hat{U}_2(t)$ and

$$\hat{H}_{1\cup 2} = \hat{H}_1 \otimes \hat{I}_2 + \hat{I}_1 \otimes \hat{H}_2.$$
Let $\hat{A}$ be an operator on $\mathcal{H}_1 \otimes \mathcal{H}_2$. Then $\text{tr}_2(\hat{A})$ is an operator on $\mathcal{H}_1$.

Let $\{\phi_n^1\}$ be an ONB of $\mathcal{H}_1$ and $\{\phi_m^2\}$ an ONB of $\mathcal{H}_2$. Then

$$\langle \phi_n^1 | \text{tr}_2 \hat{A} | \phi_k^1 \rangle = \sum_{m=1}^{\dim \mathcal{H}_2} \langle \phi_n^1 \otimes \phi_m^2 | \hat{A} | \phi_k^1 \otimes \phi_m^2 \rangle,$$

$\text{tr}(\text{tr}_2(\hat{A})) = \text{tr}(\hat{A})$; $\text{tr}_2$ is linear; $\text{tr}_2(\hat{A}^*) = (\text{tr} \hat{A})^*$

If $\hat{A}$ is a positive operator then $\text{tr}_2(\hat{A})$ is a positive operator.

$\text{tr}_2(\hat{U} \otimes \hat{V}) = (\text{tr} \hat{V}) \hat{U}$.

For a yes-no experiment on system 1,

$$\mathbb{P}(\text{yes}) = \langle \psi | \hat{E} \otimes \hat{I}_2 | \psi \rangle = \text{tr} (\hat{\rho}_1 \hat{E})$$

with $\hat{\rho}_1 = \text{tr}_2(\langle \psi \rangle \langle \psi |)$; $\hat{\rho}_1$ is called the reduced density operator of system 1.

All positive operators with trace 1 (and only those) can occur as reduced density operators.
Canonical density operator

classical canonical distribution:
density \( \rho(q, p) = \frac{1}{Z} \exp\left(-\beta H(q, p)\right) \)
A system in thermal equilibrium with a heat bath at (absolute) temperature \( T \) has \( \rho \)-typical phase point \((q, p)\).

inverse temperature \( \beta = \frac{1}{kT} \)
\( k = \) Boltzmann's constant \( \approx 10^{-23} \) J/K

canonical density operator:
\[
\hat{\rho}_\beta = \frac{1}{Z} \exp\left(-\beta \hat{H}\right)
\]
Justification of the canonical density operator

Canonical typicality

Let system 1 and system 2 consist of $N_1 \ll N_2$ particles. Let the interaction between systems 1 and 2 be weak, $\hat{H} \approx \hat{H}_1 \otimes \hat{I}_2 + \hat{I}_1 \otimes \hat{H}_2$. Let $\mathcal{H}_{mc} \subset \mathcal{H}_1 \otimes \mathcal{H}_2$ be an energy shell.

For most $\psi \in \mathcal{H}_{mc}$ with $\|\psi\| = 1$, the reduced density operator of system 1 is (approximately) canonical,

$$\text{tr}_2 |\psi\rangle\langle\psi| \approx \frac{1}{Z} e^{-\beta \hat{H}_1}$$

for suitable $\beta = \beta(E)$.

[Schrödinger: book (1952)]
System $s$, heat bath $b$, coupling negligible:

$$\hat{H} = \hat{H}_s \otimes \hat{I}_b + \hat{I}_s \otimes \hat{H}_b.$$  

“average” statement:

$$\text{tr}_b \hat{\rho}_{[E,E+\delta E]} \approx \hat{\rho}_\beta$$

in the thermodynamic limit $N_b \to \infty$, $E/N_b \to e < \infty$.

“almost always” statement: canonical typicality

$$u_{[E,E+\delta E]} \left\{ \psi : \text{tr}_b |\psi\rangle \langle \psi| \approx \hat{\rho}_\beta \right\} \to 1$$

↑ uniform measure on $\mathcal{S}(\mathcal{H}_{[E,E+\delta E]})$

in the thermodynamic limit

For most $\psi$ of $s + b$ from the microcanonical ensemble, the reduced density matrix of $s$ is canonical.
Theorem on canonical typicality [Popescu, Short, Winter 2005]

Let $\mathcal{H}_R \subseteq \mathcal{H}_s \otimes \mathcal{H}_b$ arbitrary subspace (e.g., $\mathcal{H}_R = \mathcal{H}_{E, \delta E}$ microcanonical), $u_R$ the uniform distribution on $S(\mathcal{H}_R)$,

$$\hat{\rho}_R = \frac{1}{\dim \mathcal{H}_R} \hat{P}_{\mathcal{H}_R}$$ and $\varepsilon > 0$.

Then

$$u_R\left\{ \psi : \|\text{tr}_b |\psi\rangle\langle\psi| - \text{tr}_b \hat{\rho}_R \|_1 \geq \eta \right\} \leq \eta'$$

where $\eta, \eta'$ are small when $\dim \mathcal{H}_s \ll 1 / \text{tr}(\text{tr}_s \hat{\rho}_R)^2$ (system $\ll$ bath) and $\varepsilon \ll 1 \ll \varepsilon^2 \dim \mathcal{H}_R$ (many states allowed).

$$\|\hat{M}\|_1 = \text{tr} |\hat{M}| = \text{tr} \sqrt{\hat{M}^* \hat{M}}$$
Thank you for your attention