A metric for quantum states (matrices) issued from von Neumann's entropy

Ref: R. Balian, Y. Alhassid and H. Reinhardt

Physics Reports 131 (1986) 1-146

§3 and appendix B
I. Quantum formalism

QM = The most fundamental theory, but probabilistic

• "Physical variables" = "observables" $\hat{O}$: elements of a C*-algebra. Behave as random variables, ... but do not commute (≠ classical physics)

• "State" = encoding of information: linear, real and non-negative mapping $\hat{O} \mapsto \langle \hat{O} \rangle$ onto real numbers. $\langle \cdots \rangle$ analogous to expectation values... but violate ordinary probability theory (Bell)

$\Rightarrow$ Correspondence implemented as scalar product $\langle \hat{O} \rangle = (\hat{O}; \hat{D})$ with an element $\hat{D}$ of the space of states, dual of the space of observables (regarded as vector space)
Two vector spaces, dual but not symmetric:

- **Observeable space**: supplemented with C*-algebraic structure, physical observables restricted to $\hat{O} = \hat{O}^\dagger$
- **State space**: physical states restricted by $(\hat{O}; \hat{D})$ real, $(\hat{I}; \hat{D}) = 1$ and $(\hat{O}^2; \hat{D}) \geq 0$

"Representation" = choice of basis in the two spaces

In a change of representation, only scalar products $\langle \hat{O} \rangle = (\hat{O}; \hat{D})$ are physically meaningful.

**Finite systems**: $\hat{O}$ = hermitean matrix; $\hat{D}$ = "density matrix" $\hat{D} = \hat{D}^\dagger$, $\text{Tr} \hat{D} = 1$, $\| \hat{D} \| \geq 0$

$\langle \hat{O} \rangle = (\hat{O}; \hat{D}) = \text{Tr} \hat{O} \hat{D}$

How can one define physically meaningful distance between states?
II. Von Neumann entropy (1932)

\[ S(\hat{D}) = -\text{Tr} \hat{D} \ln \hat{D} \]

thermodynamic entropy
if \( \hat{D} \) is Boltzmann–Gibbs equilibrium state

analysis of quantum measurement
if \( \hat{D} \) is the state of system + apparatus

\( \Rightarrow \) Shannon (1948) \( S(p) = -\sum_i p_i \ln p_i \) Information

\( \Rightarrow \) Brillouin (1957) \( S(\hat{D}) \) interpreted as missing information
when \( \hat{D} \) is known. Entropy = Information \( \Uparrow \)

\( \Rightarrow \) Jaynes (1958) Maximum entropy criterion for assigning
a density operator without bias when
partial information is available on a quantum system. Experimental confirmation

\( \Rightarrow \) quantum information

4
Further justifications

- Axiomatic construction (additivity, unitary invariance, concavity, continuity)
- Derivation of maximum von Neumann entropy from Laplace’s principle of equiprobability ("insufficient reason")

\[ S(\hat{D}) = -\text{Tr} \hat{D} \ln \hat{D} \text{ is the sole physically meaningful quantity, besides the expectation values } <\hat{\mathcal{O}}> = (\hat{\mathcal{O}}; \hat{D}) = \text{Tr} \hat{\mathcal{O}} \hat{D} \]

⇒ enrich the mathematical structure
III. Mathematical structures issued from $S(\hat{\mathcal{D}})$

- Concavity $d^2 S < 0$ induces natural metric
  \[ ds^2 ≡ -d^2 S(\hat{\mathcal{D}}) = \text{Tr} \ d\hat{\mathcal{D}} \ d\ln\hat{\mathcal{D}} \]
  $ds$ = distance between neighbouring density matrices

- Metric tensor = Hessian of $S(\hat{\mathcal{D}})$

- Legendre transform $\hat{X} =$ observable, $F(\hat{X}) = \ln \text{Tr} \ \exp \hat{X}$
  \[ dF = \text{Tr} \ d\hat{X} \ \hat{\mathcal{D}} \quad \text{with} \quad \hat{\mathcal{D}} = \frac{\exp \hat{X}}{\text{Tr} \ \exp \hat{X}} \]
  \[ F - \text{Tr} \ \hat{X} \ \hat{\mathcal{D}} = S(\hat{\mathcal{D}}) \]

  $F(\hat{X}) \leftrightarrow S(\hat{\mathcal{D}})$: Legendre
  (von Neumann entropy)

Generalization of Legendre transform of thermodynamics
Massieu potential $F \leftrightarrow$ Entropy of thermodynamic $S$

- Canonical mapping between the two dual vector spaces
  observable $\hat{X} \leftrightarrow$ state $\hat{\mathcal{D}}$
IV. Projection method of statistical mechanics

- Quantum dynamics: \[ i\hbar \frac{d\hat{D}(t)}{dt} = [\hat{H}, \hat{D}(t)] \]

- Follow only "relevant observables" \( \{\hat{A}_k\} \) \[ \langle A_k \rangle = \text{Tr} \hat{A}_k \hat{D} \]

- Reduced state \( \hat{\mathcal{D}}_R \)
  \[ \begin{cases} \text{Tr} \hat{A}_k \hat{\mathcal{D}}_R = \langle A_k \rangle & \text{all relevant information} \\ S(\hat{\mathcal{D}}_R) \text{ max} & \text{but no extra information} \end{cases} \]
  \[ \Rightarrow \hat{\mathcal{D}}_R = \exp \sum_k \lambda_k \hat{A}_k \quad \text{(Lagrange)} \]

- Relevant manifold \( \mathcal{R} \) (parametrized by \( \{A_k\} \) or \( \{A_k^*\} \))

  The correspondence \( \hat{\mathcal{D}} \mapsto \hat{\mathcal{D}}_R \) is a projection on \( \mathcal{R} \)

  With the metric \( ds^2 = -d^2S \) it is an orthogonal projection

  \( \hat{\mathcal{D}}_R(t) = "\text{best}\) approximation of \( \hat{\mathcal{D}}(t) \) within \( \mathcal{R} \) \)
V. Mathematical properties of the metric

\[ ds^2 = -d^2S = \text{Tr} \int_0^\infty d\xi \left( d\hat{\mathcal{D}} \frac{1}{\xi + \hat{\mathcal{D}}} \right)^2 \]

\[ \rightarrow = \sum_{i,j} \frac{\ln D_i - \ln D_j}{D_i - D_j} dD_{ij} dD_{ji} \]

- Natural generalization for density matrices of the Fisher information metric \( \sum_i (d\rho_i)^2 \)

- More physical than the current Bures metric

\[ \sum_{i,j} \frac{2}{D_i + D_j} dD_{ij} dD_{ji} \]
• **Riemannian structure**

Basis $\hat{\Sigma}_\mu$ in the matrix space of states.

Coordinates of $\hat{D}$: $\hat{D} = \sum D^\nu \hat{\Sigma}_\nu$

→ Metric tensor $g_{\mu\nu} = -\frac{\partial^2 S(\hat{D})}{\partial D^\mu \partial D^\nu}$

→ Christoffel connection $\Gamma_{\mu\nu\rho} = -\frac{1}{2} \frac{\partial^3 S(\hat{D})}{\partial D^\mu \partial D^\nu \partial D^\rho}$

→ Riemann curvature $R_{\mu\nu,\rho\sigma} = g^{\tau\nu'}(\Gamma_{\mu\rho\sigma} \Gamma_{\tau\nu'} - \Gamma_{\mu\nu\rho} \Gamma_{\tau,\nu'})$

Curvature is issued from the non-commutativity of quantum observables.

The space of classical states is flat.