Radboud Universiteit

Demystifying quantum mechanics

Mikhail Katsnelson

Collaborators

Hans De Raedt,
RUG

Kristel Michielsen,
Julich

Dennis Willsch,
Julich

Vitaly Vanchurin,
Bethesda

LI approach - References

Quantum theory as the most robust description of reproducible experiments

Hans De Raedt[a], Mikhail I. Katsnelson[b], Kristel Michielsen[c,d,*]

Annals of Physics 347 (2014) 45–73

Quantum theory as a description of robust experiments: Derivation of the Pauli equation

Hans De Raedt[a], Mikhail I. Katsnelson[b], Hylke C. Donker[b], Kristel Michielsen[c,d,*]

Annals of Physics 359 (2015) 166–186

Logical inference approach to relativistic quantum mechanics: Derivation of the Klein–Gordon equation

H.C. Donker[a,*], M.I. Katsnelson[a], H. De Raedt[b], K. Michielsen[c]

Annals of Physics 372 (2016) 74–82

Logical inference derivation of the quantum theoretical description of Stern–Gerlach and Einstein–Podolsky–Rosen–Bohm experiments

Hans De Raedt[a], Mikhail I. Katsnelson[b], Kristel Michielsen[c,d,*]

Annals of Physics 396 (2018) 96–118

Quantum theory as plausible reasoning applied to data obtained by robust experiments

H. De Raedt[1], M. I. Katsnelson[2] and K. Michielsen[3,4]

PHILOSOPHICAL TRANSACTIONS A

Cite this article: De Raedt H, Katsnelson MI, Michielsen K. 2016 Quantum theory as plausible reasoning applied to data obtained by robust experiments. *Phil. Trans. R. Soc. A* **374**: 20150233.

Other relevant references

Separation of conditions as a prerequisite for quantum theory

Annals of Physics 403 (2019) 112–135

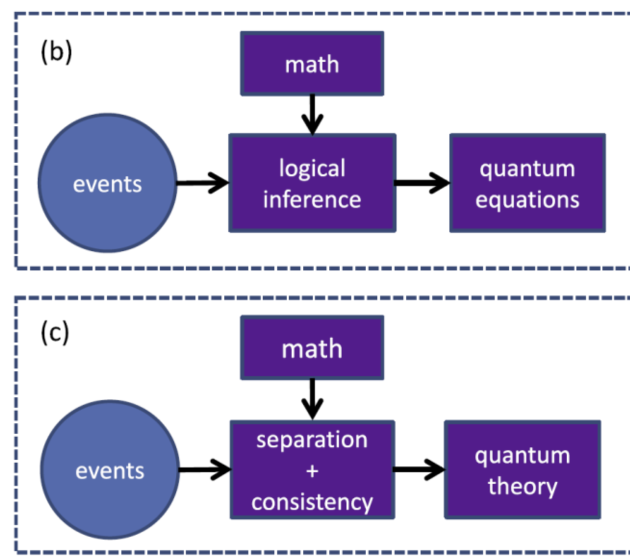
Hans De Raedt [a], Mikhail I. Katsnelson [b], Dennis Willsch [c], Kristel Michielsen [c,d,*]

Foundations of Physics (2021) 51:94
https://doi.org/10.1007/s10701-021-00503-3

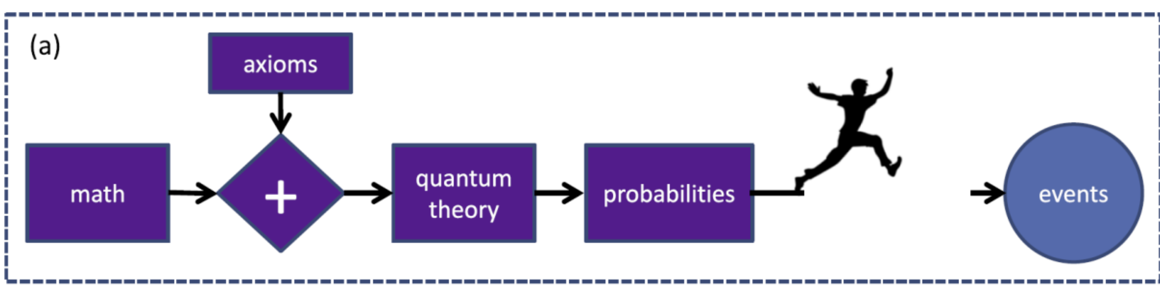
Emergent Quantumness in Neural Networks

Mikhail I. Katsnelson[1] · Vitaly Vanchurin[2,3]

Starting from ***events***

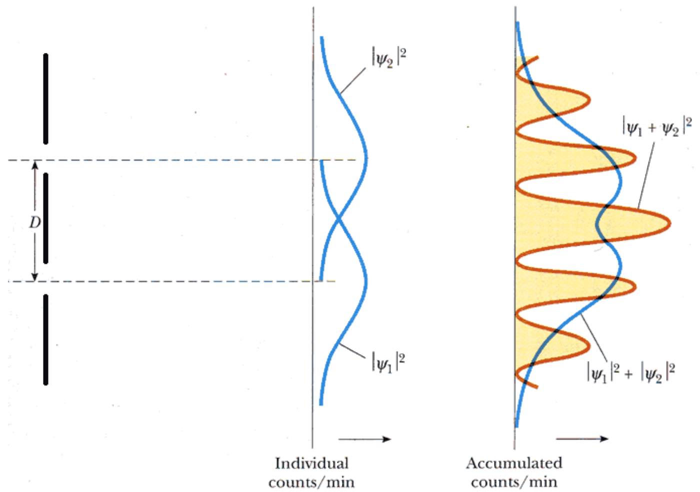


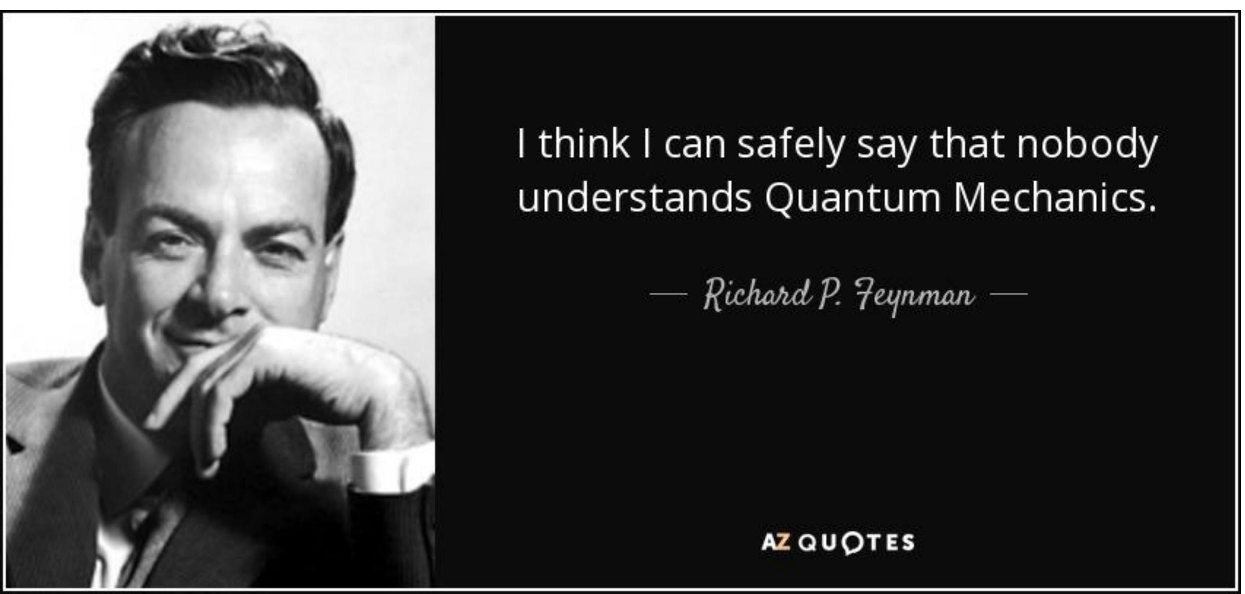
Conventional presentation



Microworld: waves are corpuscles, corpuscles are waves

Einstein, 1905 – for light (photons)
L. de Broglie, 1924 – electrons and other microparticles





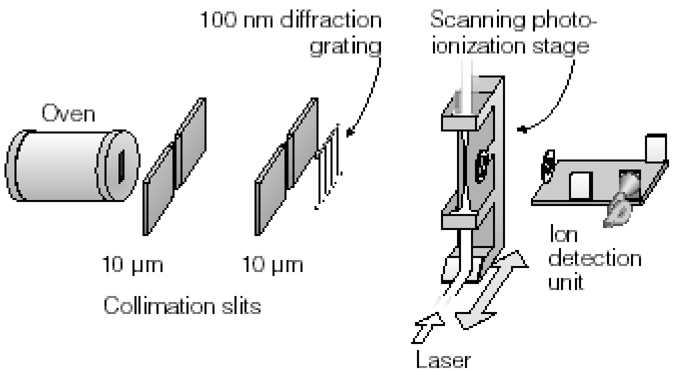
THINK!!!

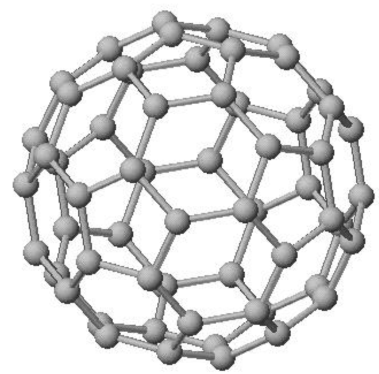
Universal property of matter

Wave–particle duality of C_{60} molecules

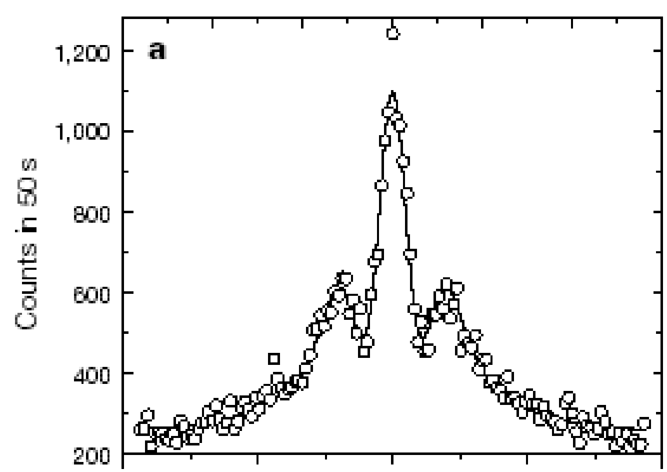
Markus Arndt, Olaf Nairz, Julian Vos-Andreae, Claudia Keller, Gerbrand van der Zouw & Anton Zeilinger

NATURE | VOL 401 | 14 OCTOBER 1999 |



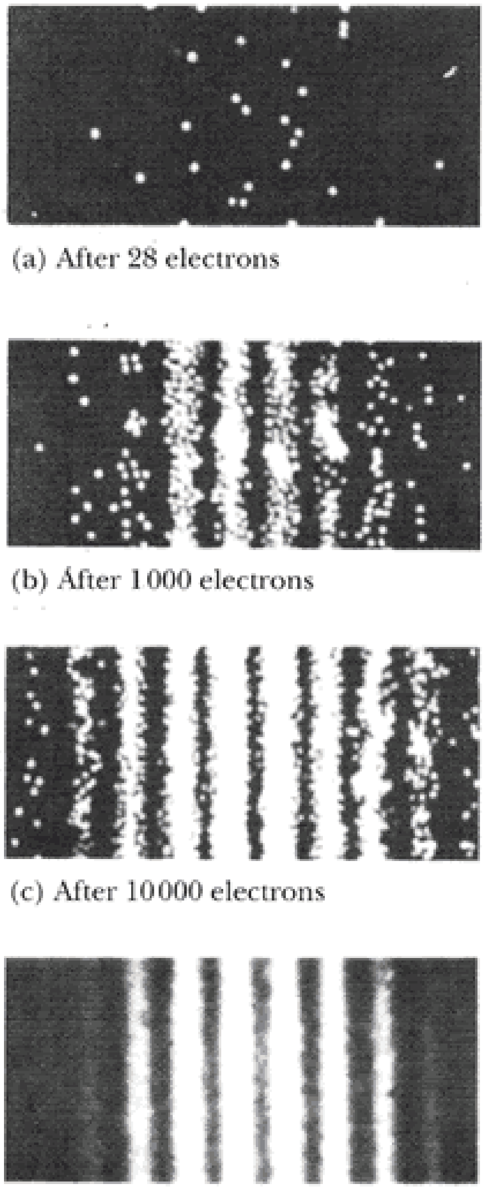


C_{60}



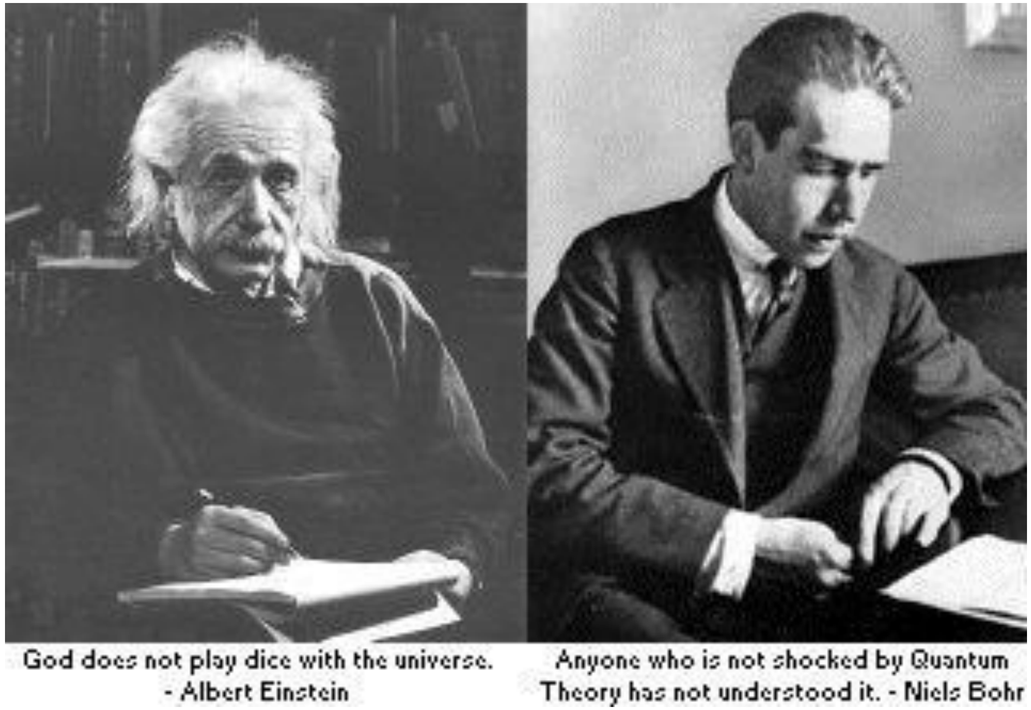
Matter waves for C_{60} molecules

Electrons are particles (you cannot see half of electron) but moves along ***all*** possible directions (interference)



We cannot describe individual events, individual spots seem to be completely random, but ensemble of the spots forms regular interference fridges

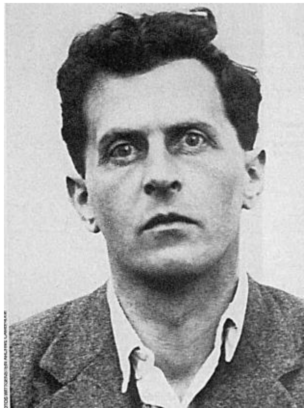
Randomness in the foundations of physics?!



A. Einstein: Quantum mechanics is incomplete; superposition principle does not work in the macroworld

N. Bohr: Classical measurement devices is an important part of quantum reality; we have to describe quantum world in terms of a language created for macroworld

The limits of my language mean the limits of my world
(Ludwig Wittgenstein)

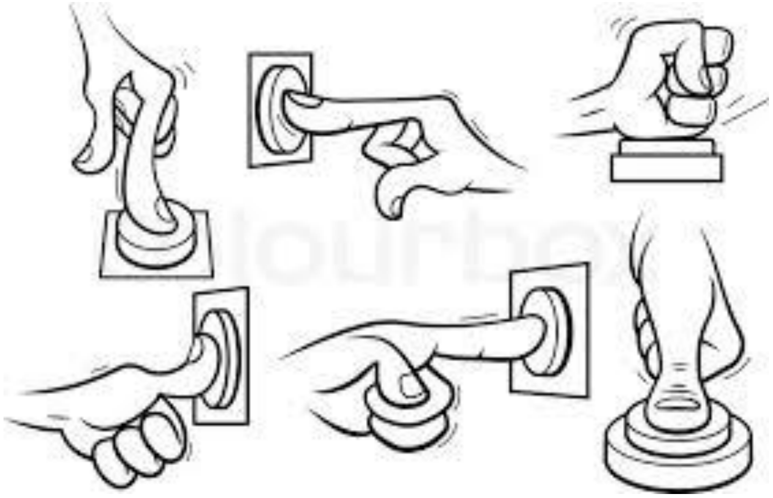


Two ways of thinking

I. Reductionism ("microscopic" approach)

Everything is from water/fire/earth/gauge fields/quantum space-time foam/strings... and the rest is your problem

II. Phenomenology: operating with "black boxes"



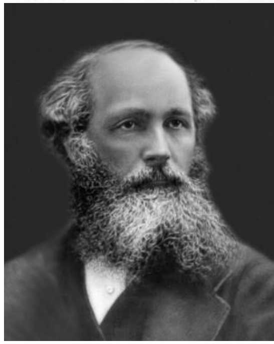
Two ways of thinking II

Knowledge begins, so to speak, in the middle, and leads into the unknown - both when moving upward, and when there is a downward movement. Our goal is to gradually dissipate the darkness in both directions, and the absolute foundation - this huge elephant carrying on his mighty back the tower of truth - it exists only in a fairy tales (Hermann Weyl)

We never know the foundations! How can we have a reliable knowledge without the base?

Mathematics & Physics

Newton: It is useful to solve (ordinary) differential equations



Maxwell: It is useful to solve *partial* differential equations

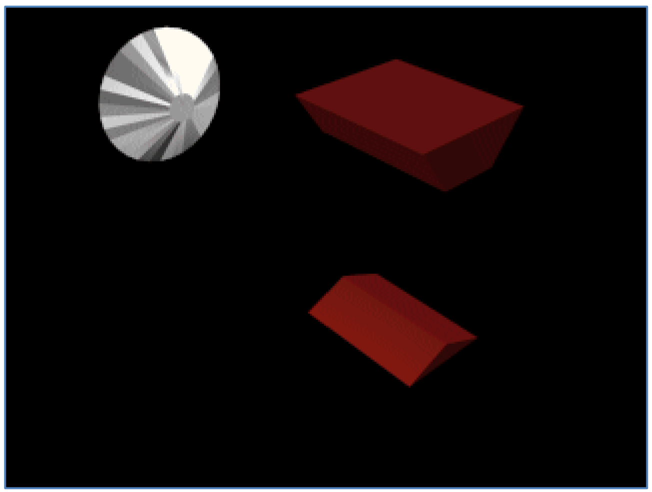
Heisenberg, Dirac, von Neumann et al: It is useful to consider state vectors and operators in Hilbert space

But this is much farther from usual human intuition – may be, too far?!

Can we demistify it?!

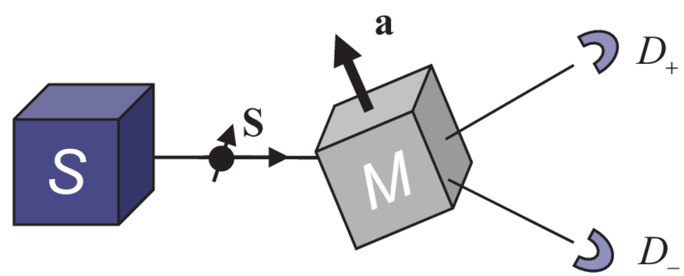
Stern-Gerlach experiment

- Neutral atoms (or neutrons) pass through an inhomogeneous magnetic field



- Inference from the data: directional quantization

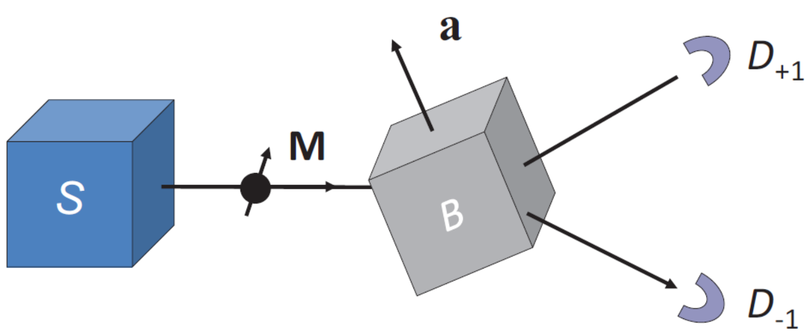
- Idealization



- Source *S* emits particles with magnetic moment
- Magnet *M* sends particle to one of two detectors
- Detectors count every particle

Logical inference

- Shorthand for propositions
  - x=+1 ⇔ D_+ clicks
  - x=-1 ⇔ D_- clicks
  - **M** ⇔the value of **M** is **M**
  - **a** ⇔the value of **a** is **a**
  - **Z** ⇔everything else which is known to be relevant to the experiment but is considered to fixed



- We assign a real number $P(x|\mathbf{M},\mathbf{a},Z)$ between 0 and 1 to express our expectation that detector D_+ or (exclusive) D_- will click and want to derive, not postulate, $P(x|\mathbf{M},\mathbf{a},Z)$ from general principles of rational reasoning
- What are these general principles ?

Plausible, rational reasoning → inductive logic, logical inference

- G. Pólya, R.T. Cox, E.T. Jaynes, …
  - From general considerations about rational reasoning it follows that the plausibility that a proposition *A* (*B*) is true given that proposition *Z* is true may be encoded in real numbers which satisfy

$$0 \le P(A|Z) \le 1$$

$$P(A|Z) + P(\bar{A}|Z) = 1 \quad ; \quad \bar{A} = \text{NOT } A$$

$$P(AB|Z) = P(A|BZ)P(B|Z) \quad ; \quad AB = A \text{ AND } B$$

  - Extension of Boolean logic, applicable to situations in which there is uncertainty about some but not all aspects
    - Kolmogorov's probability theory is an example which complies with the rules of rational reasoning
    - Is quantum theory another example?

Plausible, rational reasoning ➔ logical inference II

- Plausibility
  - Is an intermediate mental construct to carry out inductive logic, rational reasoning, logical inference
  - May express a degree of believe (subjective)
  - May be used to describe phenomena independent of individual subjective judgment

    plausibility ➔ i-prob (inference-probability)

Application to the Stern-Gerlach experiment

We repeat the experiment *N* times. The number of times that D_+ (D_-) clicks is n_+ (n_-)

i-prob for the individual event is

$$P(x|\mathbf{a}\cdot\mathbf{M},Z)=P(x|\theta,Z)=\frac{1+xE(\theta)}{2} \quad , \quad E(\theta)=E(\mathbf{a}\cdot\mathbf{M},Z)=\sum_{x=\pm1}xP(x|\theta,Z)$$

Dependent on $\cos\theta=\mathbf{a}\cdot\mathbf{M}$ Rotational invariance

Different events are logically independent:

$$P(x_1,\ldots,x_N|\mathbf{a}\cdot\mathbf{M},Z) \quad = \quad \prod_{i=1}^{N}P(x_i|\theta,Z)$$

The i-prob to observe n_+ and n_- events is

$$P(n_{+1},n_{-1}|\theta,N,Z)=N!\prod_{x=\pm1}\frac{P(x|\theta,Z)^{n_x}}{n_x!}$$

How to express robustness?

- Hypothesis H_0: given θ we observe n_+ and n_-
- Hypothesis H_1: given $\theta+\varepsilon$ we observe n_+ and n_-
- The evidence Ev(H_1/H_0) is given by

$$Ev(H_1 \mid H_0) = \ln \frac{P(n_+, n_- \mid \theta+\varepsilon, N, Z)}{P(n_+, n_- \mid \theta, N, Z)} = \sum_{x=\pm1} n_x \ln \frac{P(x \mid \theta+\varepsilon, Z)}{P(x \mid \theta, Z)} =$$

$$= \sum_{x=\pm1} n_x \left\{ \varepsilon \frac{P'(x \mid \theta, Z)}{P(x \mid \theta, Z)} - \frac{\varepsilon^2}{2} \left[\frac{P'(x \mid \theta, Z)}{P(x \mid \theta, Z)} \right]^2 + \frac{\varepsilon^2}{2} \frac{P''(x \mid \theta, Z)}{P(x \mid \theta, Z)} \right\} + O(\varepsilon^3)$$

- Frequencies should be robust with respect to small changes in θ ➔ we should minimize, in absolute value, the coefficients of ε, ε^2,...

Remove dependence on ϵ (1)

$$Ev(H_1 \mid H_0) = \sum_{x=\pm 1} n_x \left\{ \varepsilon \frac{P'(x \mid \theta, Z)}{P(x \mid \theta, Z)} - \frac{\varepsilon^2}{2} \left[\frac{P'(x \mid \theta, Z)}{P(x \mid \theta, Z)} \right]^2 + \frac{\varepsilon^2}{2} \frac{P''(x \mid \theta, Z)}{P(x \mid \theta, Z)} \right\} + O(\varepsilon^3)$$

- Choose

$$P(x \mid \theta, Z) = \frac{n_x}{N}$$

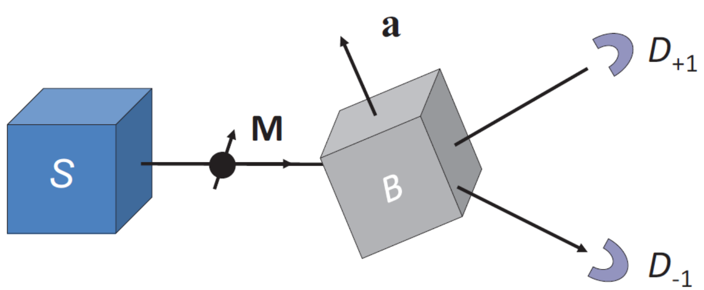
  - Removes the 1st and 3rd term
  - Recover the intuitive procedure of assigning to the i-prob of the individual event, the frequency which maximizes the i-prob to observe the whole data set

Remove dependence on ϵ (2)

$$Ev(H_1 | H_0) = \sum_{x=\pm 1} n_x \left\{ \varepsilon \frac{P'(x|\theta,Z)}{P(x|\theta,Z)} - \frac{\varepsilon^2}{2}\left[\frac{P'(x|\theta,Z)}{P(x|\theta,Z)}\right]^2 + \frac{\varepsilon^2}{2}\frac{P''(x|\theta,Z)}{P(x|\theta,Z)} \right\} + O(\varepsilon^3)$$

- Minimizing the 2nd term (Fisher information) for all possible (small) ε and θ

$$I_F = \sum_{x=\pm 1} \frac{1}{P(x|\theta,Z)} \left(\frac{\partial P(x|\theta,Z)}{\partial \theta} \right)^2$$



$$P(x|\mathbf{a}\cdot\mathbf{M},Z) = P(x|\theta,Z) = \frac{1 \pm x\mathbf{a}\cdot\mathbf{M}}{2}$$

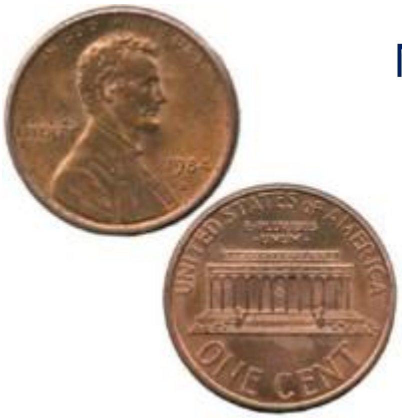
- In agreement with quantum theory of the idealized Stern-Gerlach experiment

Bernoulli trial

Two outcomes (head and tails in coin flypping)

Results are dependent on a single parameter θ which runs a circle (periodicity); what is special in quantum trials?

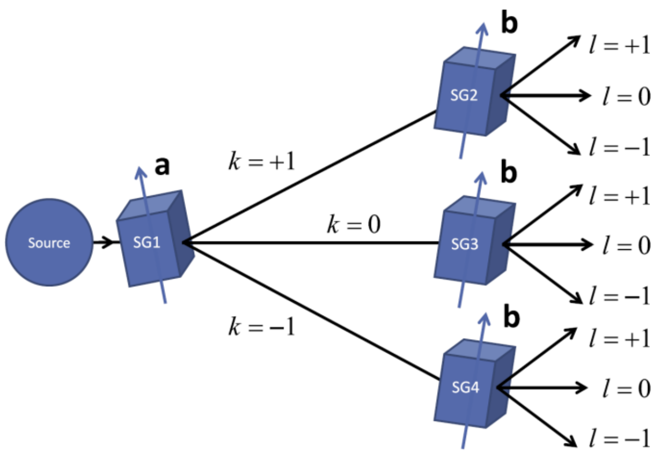
The results of SG experiment are the most robust, that is, correspond to minimum Fisher information

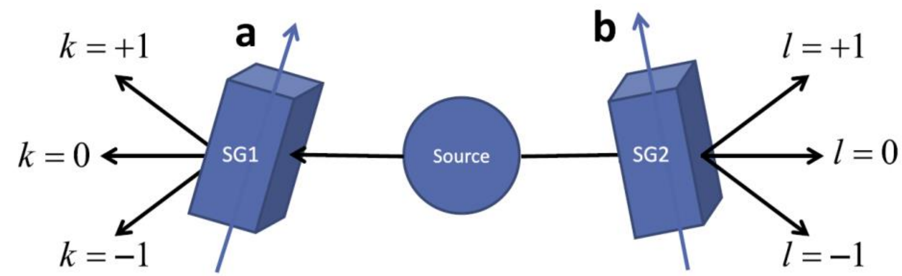


No assumptions on wave functions, Born rules and other machinery Of quantum physics, just looking for the most robust description of the results of repeating "black box" experiments

Double SG experiment or EPRB experiment for S >1/2

Annals of Physics 396 (2018) 96–118





$$P(k,l|\mathbf{a}\cdot\mathbf{b},Z)=\frac{1}{12}\begin{pmatrix}(1+\mathbf{a}\cdot\mathbf{b})^2 & 2(1-(\mathbf{a}\cdot\mathbf{b})^2) & (1-\mathbf{a}\cdot\mathbf{b})^2\\ 2(1-(\mathbf{a}\cdot\mathbf{b})^2) & 4(\mathbf{a}\cdot\mathbf{b})^2 & 2(1-(\mathbf{a}\cdot\mathbf{b})^2)\\ (1-\mathbf{a}\cdot\mathbf{b})^2 & 2(1-(\mathbf{a}\cdot\mathbf{b})^2) & (1+\mathbf{a}\cdot\mathbf{b})^2\end{pmatrix}_{2-k,2-l}$$

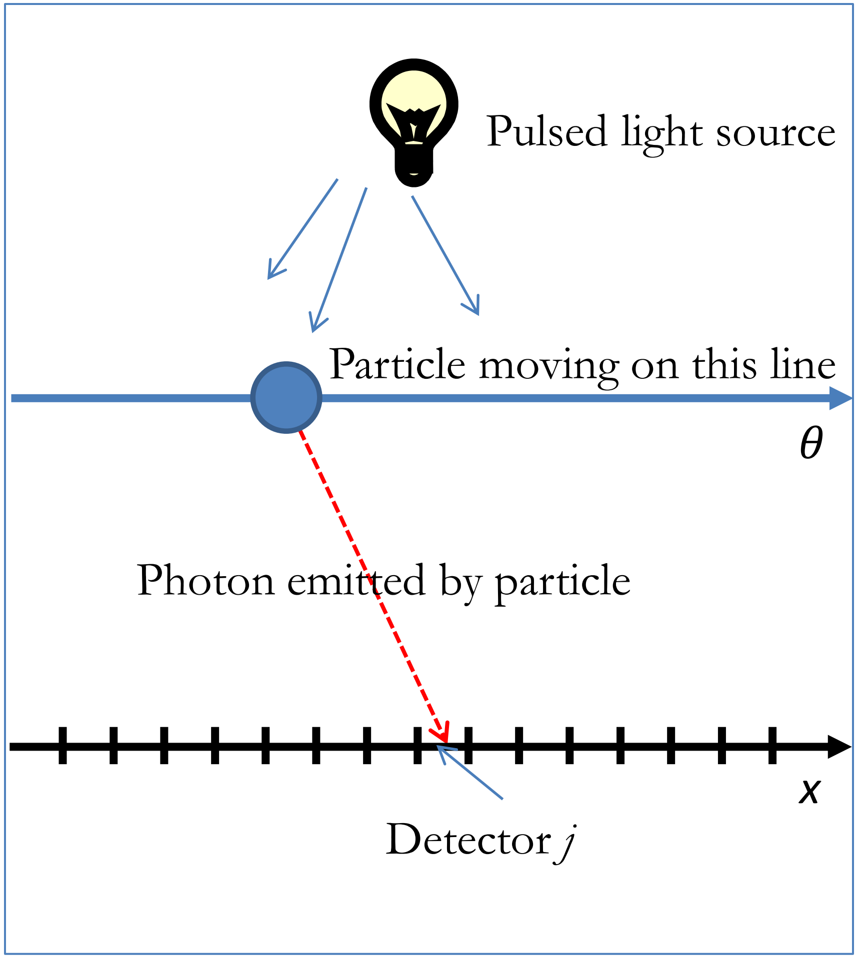
In agreement with the predictions of QM ***but*** there is a second solution with the same Fisher info

$$\mathbf{a}\cdot\mathbf{b}\ \rightarrow\ -\mathbf{a}\cdot\mathbf{b}$$

The LI framework includes quantum theory as a special case.

Logical inference ➔ Schrödinger equation

- Generic procedure:
- Experiment ➔
- The "true" position θ of the particle is uncertain and remains unknown
- i-prob that the particle at unknown position θ activates the detector at position x : $P(x \mid \theta, Z)$



Robustness

- Assume that it does not matter if we repeat the experiment somewhere else ➔

$$P(x\,|\,\theta,Z)=P(x+\zeta\,|\,\theta+\zeta,Z)\quad;\quad \zeta \text{ arbitrary}$$

- Condition for robust frequency distribution ⬄ minimize the functional (Fisher information)

$$I_F(\theta)=\int_{-\infty}^{\infty}dx\,\frac{1}{P(x\,|\,\theta,Z)}\left(\frac{\partial P(x\,|\,\theta,Z)}{\partial x}\right)^2$$

with respect to $P(x|\theta,Z)$

We need to add some “dynamical” information on the system

Impose classical mechanics (á la Schrödinger)

- If there is no uncertainty at all ➔classical mechanics ➔ Hamilton-Jacobi equation

$$\frac{1}{2m}\left(\frac{\partial S(\theta)}{\partial \theta}\right)^2 + V(\theta) - E = 0 \qquad \text{(X)}$$

- If there is "known" uncertainty

$$\int_{-\infty}^{\infty} dx \left[\left(\frac{\partial S(x)}{\partial x}\right)^2 + 2m[V(x) - E]\right] P(x \mid \theta, Z) = 0 \qquad \text{(XX)}$$

  – Reduces to (X) if $P(x|\theta, Z) \rightarrow \delta(x - \theta)$

Robustness + classical mechanics

- $P(x|\theta, Z)$ can be found by minimizing $I_F(\theta)$ with the constraint that (XX) should hold

➔ We should minimize the functional

$$F(\theta) = \int_{-\infty}^{\infty} dx \left\{ \frac{1}{P(x\,|\,\theta, Z)} \left(\frac{\partial P(x\,|\,\theta, Z)}{\partial x} \right)^2 + \lambda \left[\left(\frac{\partial S(x)}{\partial x} \right)^2 + 2m[V(x) - E] \right] P(x\,|\,\theta, Z) \right\}$$

  - λ = Lagrange multiplier
  - Nonlinear equations for $P(x|\theta, Z)$ and $S(x)$

Robustness + classical mechanics

- Nonlinear equations for $P(x|\theta, Z)$ and $S(x)$ can be turned into linear equations by substituting*

$$\psi(x|\theta,Z) = \sqrt{P(x|\theta,Z)}e^{iS(x)\sqrt{\lambda}/2} \quad \Rightarrow$$

$$F(\theta) = \int_{-\infty}^{\infty} dx \left\{ 4\frac{\partial \psi^*(x|\theta,Z)}{\partial x}\frac{\partial \psi(x|\theta,Z)}{\partial x} + 2m\lambda[V(x)-E]\psi^*(x|\theta,Z)\psi(x|\theta,Z) \right\}$$

- Minimizing with respect to $\psi(x|\theta,Z)$ yields

$$-\frac{\partial^2 \psi(x|\theta,Z)}{\partial x^2} + \frac{m\lambda}{2}\left[V(x)-E\right]\psi(x|\theta,Z) = 0$$

➔ Schrödinger equation $\lambda = 4K^{-2} = 4\hbar^{-2}$

*E. Madelung, "Quantentheorie in hydrodynamischer Form," Z. Phys. 40, 322 – 326 (1927)

Time-dependent, multidimensional case

The space is filled by detectors which are fired (or not fired) at some discrete (integer) time $\tau = 1, \ldots, M$

At the very end we have a set of data presented as 0 (no particle in a given box at a given instant or 1

$$\Upsilon = \{\boldsymbol{j}_{n,\tau} | \boldsymbol{j}_{n,\tau} \in [-L^d, L^d]; \; n = 1, \ldots, N; \; \tau = 1, \ldots, M\}$$

or, denoting the total counts of voxels $\boldsymbol{j}$ at time τ by $0 \leq k_{\boldsymbol{j},\tau} \leq N$, the experiment produces the data set

$$\mathcal{D} = \left\{ k_{\boldsymbol{j},\tau} \middle| \tau = 1, \ldots, M; N = \sum_{\boldsymbol{j} \in [-L^d, L^d]} k_{\boldsymbol{j},\tau} \right\}. \tag{55}$$

Logical independence of events:

$$P(\mathcal{D} | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M, N, Z) = N! \prod_{\tau=1}^{M} \prod_{\boldsymbol{j} \in [-L^d, L^d]} \frac{P(\boldsymbol{j} | \boldsymbol{\theta}_\tau, \tau, Z)^{k_{\boldsymbol{j},\tau}}}{k_{\boldsymbol{j},\tau}!}$$

Time-dependent case II

Homogeneity of the space: $P(\boldsymbol{j}|\boldsymbol{\theta}, Z) = P(\boldsymbol{j}+\boldsymbol{\zeta}|\boldsymbol{\theta}+\boldsymbol{\zeta}, Z)$

Evidence: $$\mathrm{Ev} = \sum_{\boldsymbol{j},\tau} \sum_{i,i'=1}^{d} \frac{\epsilon_{i,\tau}\epsilon_{i',\tau}}{P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)} \frac{\partial P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)}{\partial \theta_i} \frac{\partial P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)}{\partial \theta_{i'}}$$

$$\mathrm{Ev} = \sum_{\boldsymbol{j},\tau} \left(\sum_{i=1}^{d} \frac{\epsilon_i, \tau}{\sqrt{P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)}} \frac{\partial P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)}{\partial \theta_i} \right)^2 \geq 0,$$

and, by using the Cauchy–Schwarz inequality, that

$$\mathrm{Ev} \leq \sum_{\boldsymbol{j},\tau} \left(\sum_{i=1}^{d} \epsilon_{i,\tau}^2 \right) \left(\sum_{i=1}^{d} \frac{1}{P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)} \left(\frac{\partial P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)}{\partial \theta_i} \right)^2 \right)$$

$$\widehat{\epsilon}^2 = \max_{i,\tau} \epsilon_{i,\tau}^2$$

$$\leq d\widehat{\epsilon}^2 \sum_{\boldsymbol{j},\tau} \sum_{i=1}^{d} \frac{1}{P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)} \left(\frac{\partial P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)}{\partial \theta_i} \right)^2,$$

Time-dependent case III

Minimizing Fisher information: $$I_F = \sum_{j,\tau} \sum_{i=1}^{d} \frac{1}{P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)} \left(\frac{\partial P(\boldsymbol{j}|\boldsymbol{\theta}_\tau, \tau, Z)}{\partial \theta_i} \right)^2$$

Taking into account homogeneity of space; continuum limit:

$$I_F = \int d\boldsymbol{x} \int dt \sum_{i=1}^{d} \frac{1}{P(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z)} \left(\frac{\partial P(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z)}{\partial x_i} \right)^2$$

Hamilton – Jacobi equations:

$$\frac{\partial S(\boldsymbol{\theta}, t)}{\partial t} + \frac{1}{2m} \left(\nabla S(\boldsymbol{\theta}, t) - \frac{q}{c} \boldsymbol{A}(\boldsymbol{\theta}, t) \right)^2 + V(\boldsymbol{\theta}, t) = 0$$

Time-dependent case IV

Minimizing functional:

$$F = \int d\boldsymbol{x} \int dt \sum_{i=1}^{d} \left\{ \frac{1}{P(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z)} \left(\frac{\partial P(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z)}{\partial x_i} \right)^2 + \lambda \left[\frac{\partial S(\boldsymbol{x}, t)}{\partial t} + \frac{1}{2m} \left(\frac{\partial S(\boldsymbol{x}, t)}{\partial x_i} - \frac{q}{c} \boldsymbol{A}(\boldsymbol{x}, t) \right)^2 + V(\boldsymbol{x}, t) \right] P(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z) \right\}$$

Substitution $\psi(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z) = \sqrt{P(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z)} e^{iS(\boldsymbol{x},t)\sqrt{\lambda}/2}$

Equivalent functional for minimization:

$$\begin{aligned} Q = 2 \int d\boldsymbol{x} \int dt \, \Bigg\{ & mi\sqrt{\lambda} \Bigg[\psi(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z) \frac{\partial \psi^*(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z)}{\partial t} \\ & - \psi^*(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z) \frac{\partial \psi(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z)}{\partial t} \Bigg] \\ & + 2 \sum_{j=1}^{d} \left(\frac{\partial \psi^*(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z)}{\partial x_j} + \frac{iq\sqrt{\lambda}}{2c} A_j(\boldsymbol{x}, t) \psi^*(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z) \right) \\ & \times \left(\frac{\partial \psi(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z)}{\partial x_j} - \frac{iq\sqrt{\lambda}}{2c} A_j(\boldsymbol{x}, t) \psi(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z) \right) \\ & + m\lambda V(\boldsymbol{x}, t) \psi^*(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z) \psi(\boldsymbol{x}|\boldsymbol{\theta}(t), t, Z) \Bigg\}, \end{aligned}$$

$$\lambda = 4/\hbar^2$$

Time-dependent case V

Time-dependent Schrödinger equation

$$i\hbar \frac{\partial \psi(\mathbf{x}|\boldsymbol{\theta}(t), t, Z)}{\partial t} = \left[-\frac{\hbar^2}{2m} \sum_{j=1}^{d} \left(\frac{\partial}{\partial x_j} - \frac{iq}{\hbar c} \boldsymbol{A}(\mathbf{x}, t) \right)^2 + V(x, t) \right] \psi(\mathbf{x}|\boldsymbol{\theta}(t), t, Z)$$

It is linear (superposition principle) which follows from classical Hamiltonian (kinetic energy is $mv^2/2$) and, inportantly, from building one complex function from two real (S and S +2πħ are equivalent).

A very nontrivial operation dictated just by desire to simplify the problem as much as possible (to pass from nonlinear to linear equation).

Requires further careful thinking!

The model of neural network

Vanchurin, V.: The world as a neural network. Entropy **22**, 1210 (2020)

The dynamics of machine learning close to learning equilibrium leads to Madelung equation; non surprisingly, machine learning provides a model of system satisfying axioms of "rational thinking"

Dynamics of trainable variables: diffusion equation

$$\frac{\partial p(t,\mathbf{q})}{\partial t} = \sum_k \frac{\partial}{\partial q_k}\left(D\frac{\partial p(t,\mathbf{q})}{\partial q_k} - \frac{dq_k}{dt}p(t,\mathbf{q})\right)$$
$$= \sum_k \frac{\partial}{\partial q_k}\left(D\frac{\partial p(t,\mathbf{q})}{\partial q_k} - \gamma\frac{\partial F(t,\mathbf{q})}{\partial q_k}p(t,\mathbf{q})\right)$$

$$\frac{dq_k}{dt} = \gamma\frac{\partial F(t,\mathbf{q})}{\partial q_k}$$

F is the free energy of the network

D and γ are parameters of the network (as well as its step ε)

The model of neural network II

Foundations of Physics (2021) 51:94

The key step: passing to grand canonical ensemble
(variable number of "neurons")

$$F \cong F + \mu n \qquad \forall n \in \mathbb{Z}$$

$$\Psi \equiv \sqrt{p}\exp\left(\frac{iF\epsilon}{\hbar}\right) \qquad -i\hbar\frac{\partial}{\partial t}\Psi = \left(\frac{\hbar^2}{2m}\sum_k \frac{\partial^2}{\partial q_k^2} - V\right)\Psi$$

$$m \equiv \frac{\epsilon}{2\gamma} \qquad \hbar \equiv \epsilon\sqrt{\frac{4D}{\gamma\lambda}}$$

V is related to loss function

Parameters should be chosen such as $\hbar = \pm\frac{\mu\epsilon}{2\pi}$

and we can use advantages of quantum learning in non-quantum system

Separation of conditions principle

Separation of conditions as a prerequisite for quantum theory

Annals of Physics 403 (2019) 112–135

Hans De Raedt [a], Mikhail I. Katsnelson [b], Dennis Willsch [c], Kristel Michielsen [c,d,*]

LI allows to derive also Pauli equation, Klein-Gordon equation (Dirac is in progress) but... Superposition principle arises as a trick. Why linear equation? Why wave function? Last not least – what about *open* quantum systtems?

Slightly different view but also based on data analysis

Standard logic: Shrödinger equation → von Neumann prescription → description of meaurements. We invert this logic!

Starting point: the way how we deal with the data (reproduced as binary sequences)

Von Neumann theory of measurement (1932)

Density matrix for subsystem A of a total system A + B

$$\rho(\alpha,\alpha')=Tr_{\beta}\Psi^{*}(\alpha',\beta)\Psi(\alpha,\beta)$$

$$\rho=\sum_{a}W_{a}|a\rangle\langle a|$$

Pure state $\rho=|a\rangle\langle a|$

$$\rho^{2}=\rho$$

Mixed state $Tr\rho^{2}<Tr\rho$

Two ways of evolution

1. Unitary evolution

$$i\hbar\frac{\partial\rho}{\partial t}=[H,\rho]$$

$$\rho(t)=\exp(iHt/\hbar)\rho(0)\exp(-iHt/\hbar)$$

Entropy is conserved

$$S=-Tr\rho\ln\rho$$

2. Nonequilibrium evolution by the measurement

$$\rho_{after}=\sum_{n}P_{n}\rho_{before}P_{n}$$

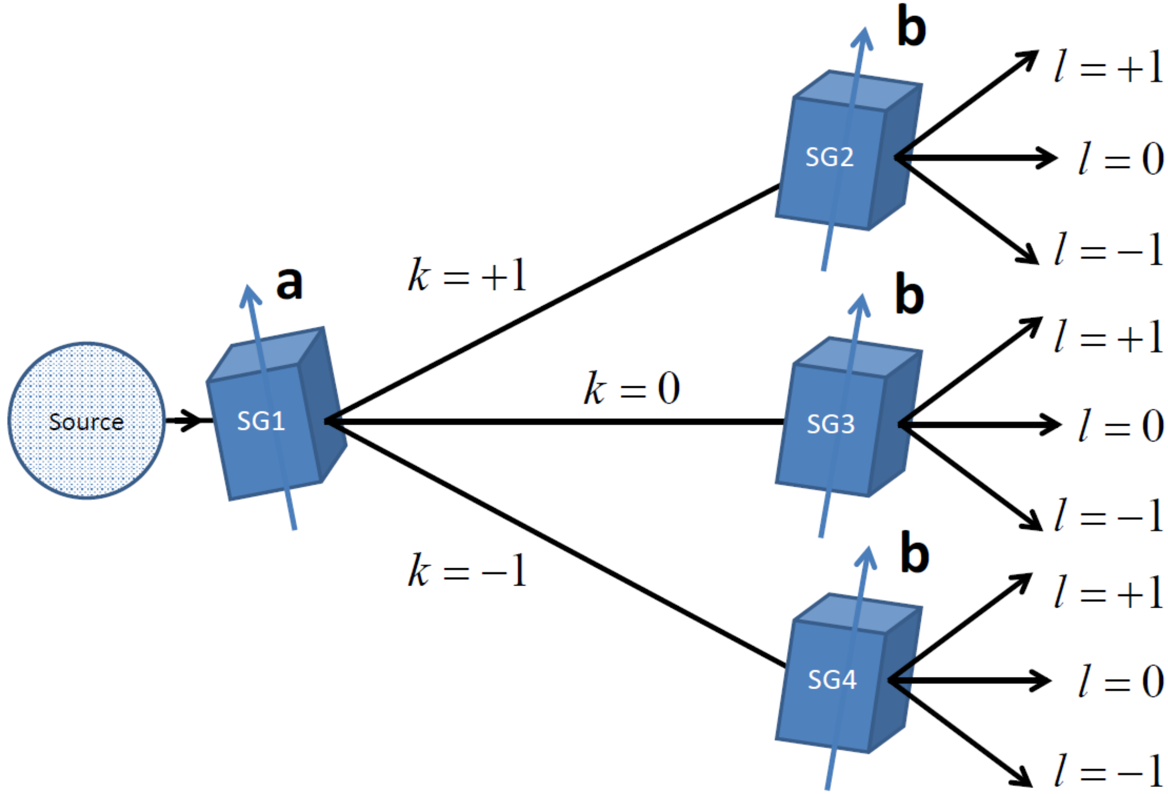
$$P_{n}=|n\rangle\langle n|$$

$$S_{after}>S_{before}$$

Density matrix after the measurement is diagonal in *n*-representation

Separation procedure

Double SG experiment with three possible outcomes (“spin 1”) is generic enough



The first SG device prepares the initial state for the second device

Separation procedure II

The data set for the first device

$$\mathscr{K} = \{k_n \mid k_n \in \{+1,0,-1\} \; ; \; n = 1,\ldots,N\}$$

$$f(k|\mathbf{a},P,N) = \frac{1}{N}\sum_{n=1}^{N} \delta_{k,k_n}$$

P properties of the particles emitted by source

Representation in terms of momenta

$$f(k|\mathbf{a},P,N) = 1 - m_2(\mathbf{a},P,N) + \frac{m_1(\mathbf{a},P,N)}{2}k + \frac{3m_2(\mathbf{a},P,N)-2}{2}k^2$$

$$m_p(\mathbf{a},P,N) = \langle k^p \rangle_{\mathbf{a}} = \frac{1}{N}\sum_{n=1}^{N} k_n^p = \sum_{k=+1,0,-1} k^p f(k|\mathbf{a},P,N) \quad , \quad p = 0,1,2$$

Separation procedure III

Let us try to represent the data as strings (sequences)

$$\mathbf{k} = (+1, 0, -1)^T \qquad \mathbf{f} = (f(+1|\mathbf{a}, P, N), f(0|\mathbf{a}, P, N), f(-1|\mathbf{a}, P, N))^T$$

$$\langle 1 \rangle_{\mathbf{a}} = (1,1,1) \cdot \mathbf{f} = \mathbf{Tr}\,(1,1,1) \cdot \mathbf{f} = \mathbf{Tr}\,\mathbf{f} \cdot (1,1,1) = \mathbf{Tr} \begin{pmatrix} f(+1|\mathbf{a}, P, N) & 0 & 0 \\ 0 & f(0|\mathbf{a}, P, N) & 0 \\ 0 & 0 & f(-1|\mathbf{a}, P, N) \end{pmatrix}$$

$$\langle k \rangle_{\mathbf{a}} = \mathbf{k}^T \cdot \mathbf{f} = \mathbf{Tr}\,\mathbf{k}^T \cdot \mathbf{f} = \mathbf{Tr}\,\mathbf{f} \cdot \mathbf{k}^T = \mathbf{Tr} \begin{pmatrix} f(+1|\mathbf{a}, P, N) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -f(-1|\mathbf{a}, P, N) \end{pmatrix}$$

$\langle k^2 \rangle_{\mathbf{a}} = \mathbf{Tr}\,\mathbf{f} \cdot (\mathbf{k}^{(2)})^T \qquad \mathbf{k}^{(2)} = (+1, 0, +1)^T$ is the other vector

Separation procedure IV

But with matrice multiplication rule we need only two matrices

$$\widetilde{\mathbf{K}} = \begin{pmatrix} +1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad \widetilde{\mathbf{F}}(\mathbf{a},P,N) = \begin{pmatrix} f(+1|\mathbf{a},P,N) & 0 & 0 \\ 0 & f(0|\mathbf{a},P,N) & 0 \\ 0 & 0 & f(-1|\mathbf{a},P,N) \end{pmatrix}$$

$$\langle k^p \rangle_{\mathbf{a}} = \mathbf{Tr}\, \widetilde{\mathbf{F}}(\mathbf{a},P,N)\widetilde{\mathbf{K}}^p \quad , \quad p = 0,1,2$$

When we rotate the axis of the first SG device and assume rotational invariance (+1 means along the device axis, -1 means opposite, 0 means perpendicular to the axis, for any direction of the axis)

$$\mathbf{K}(\mathbf{a}) = \mathbf{a} \cdot \mathbf{S} \qquad S^x = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad , \quad S^y = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ +i & 0 & -i \\ 0 & +i & 0 \end{pmatrix} \quad , \quad S^z = \begin{pmatrix} +1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Nothing is quantum yet, except the assumption of *three* outcomes!

Separation procedure V

$$\mathbf{M}_k(\mathbf{e}_z) = \mathbb{1} - (S^z)^2 + \frac{k}{2}S^z + \frac{k^2}{2}\left[3(S^z)^2 - 2\mathbb{1}\right]$$

Introduce projector operator:

$$\mathbf{M}_k(\mathbf{e}_z)\mathbf{M}_l(\mathbf{e}_z) = \delta_{k,l}\mathbf{M}_k(\mathbf{e}_z)$$

$$= \begin{pmatrix} \frac{k^2+k}{2} & 0 & 0 \\ 0 & 1-k^2 & 0 \\ 0 & 0 & \frac{k^2-k}{2} \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , & k=+1 \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} , & k=0 \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} , & k=-1 \end{cases}$$

From rotational invariance:

$$\mathbf{M}_k(\mathbf{a}) = 1 - (\mathbf{a}\cdot\mathbf{S})^2 + \frac{k}{2}\mathbf{a}\cdot\mathbf{S} + \frac{k^2}{2}\left[3(\mathbf{a}\cdot\mathbf{S})^2 - 2\mathbb{1}\right]$$

$$f(k|\mathbf{a},P,N) = \mathbf{Tr}\,\mathbf{F}(P,N)\mathbf{M}_k(\mathbf{a}) = \mathbf{Tr}\,\mathbf{M}_k(\mathbf{a})\mathbf{F}(P,N) = \mathbf{Tr}\,\mathbf{M}_k(\mathbf{a})\mathbf{F}(P,N)\mathbf{M}_k(\mathbf{a})$$

Only the last form gives Hermitian density matrix for the next use!

Separation procedure VI

As all SG magnets are assumed to be identical, consistency demands that their description should be the same, that is the filtering property of SG2, SG3 and SG4 should be described by $\mathbf{M}_l(\mathbf{b})$.

The first SG device plays the role of the source for the second device etc. – this is the separaction of conditions requirement!

$$\mathscr{D} = \left\{ (k_n, l_n) \mid k_n, l_n \in \{+1, 0, -1\} \,;\, n = 1, \ldots, N \right\}$$

$$f(k|\mathbf{a}, P, N) = \sum_{l=+1,0,-1} f(k, l|\mathbf{a}, \mathbf{b}, P, N)$$

$$f(k, l|\mathbf{a}, \mathbf{b}, P, N) = \mathbf{Tr}\, \mathbf{M}_l(\mathbf{b}) \mathbf{M}_k(\mathbf{a}) \mathbf{F}(P, N) \mathbf{M}_k(\mathbf{a}) \mathbf{M}_l(\mathbf{b})$$

Consequence: $$f(k|\mathbf{a}, P, N) = \sum_{l=+1,0,-1} f(k, l|\mathbf{a}, \mathbf{b}, P, N)$$

Separation procedure VII

Until now *P* (the properties of source) is arbitrary. Illustration:

$$\mathbf{F}(P,N) = \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad f(k|\mathbf{a},P,N) = \mathbf{Tr}\,\mathbf{M}_k(\mathbf{a})\mathbf{F}(P,N)\mathbf{M}_k(\mathbf{a}) = \frac{1}{3}$$

(sourse of unpolarized particles, full isotropy in single SG)

$$f(k,l|\mathbf{a},\mathbf{b},P,N) = \mathbf{Tr}\,\mathbf{M}_l(\mathbf{b})\mathbf{M}_k(\mathbf{a})\mathbf{F}(P,N)\mathbf{M}_k(\mathbf{a})\mathbf{M}_l(\mathbf{b})$$

$$= \begin{cases} \frac{1}{12}(1+\mathbf{a}\cdot\mathbf{b})^2 & , \quad k=l=+1,-1 \\ \frac{1}{3}(\mathbf{a}\cdot\mathbf{b})^2 & , \quad k=l=0 \\ \frac{1}{12}(1-\mathbf{a}\cdot\mathbf{b})^2 & , \quad (k,l)=(+1,-1),(-1,+1) \\ \frac{1}{6}(1-(\mathbf{a}\cdot\mathbf{b})^2) & , \quad (k,l)=(+1,0),(-1,0),(0,+1),(0,-1) \end{cases}$$

This is the result of QM – but strictly speaking not the *derivation*

$$\mathrm{SOC} \models \mathrm{QT}$$

Separation procedure VIII

Dependence on parameters (e.g., time) $\mathscr{D}(\lambda)$ $f(k,l|\mathbf{a},\mathbf{b},P,N,\lambda)$

$$\langle k^p \rangle_\lambda = \mathbf{Tr}\,\mathbf{F}(P,N,\lambda)\mathbf{K}^p(\mathbf{a}) \quad , \quad p = 0,1,2$$

$$\mathbf{Tr}\,\mathbf{F}(P,N,\lambda) = 1 \qquad \mathbf{Tr}\,\frac{\partial^n \mathbf{F}(P,N,\lambda)}{\partial \lambda^n} = 0 \quad , \quad n > 0$$

Traceless matrix is a commutator

K. Shoda, "Einige Sätze über Matrizen," Jap. J. Math. **13**, 361–365 (1936).
A. A. Albert and B. Muckenhoupt, "On matrices of trace zeros," Michigan Math. J. , 1–3 (1957).

$$\frac{\partial \mathbf{F}(P,N,\lambda)}{\partial \lambda} = [Y(\lambda), Z(\lambda)]$$

$\mathbf{F}(P,N,\lambda)$ is a Hermitian (non-negative definite) matrix

$\mathbf{F}(P,N,\lambda) = U^\dagger(\lambda)D(\lambda)U(\lambda)$ $\quad D(\lambda)$ is diagonal

Separation procedure IX

$$\frac{\partial \mathbf{F}(P,N,\lambda)}{\partial \lambda} = \left[\mathbf{F}(P,N,\lambda), U^{\dagger}(\lambda)\frac{\partial U(\lambda)}{\partial \lambda}\right] + U^{\dagger}(\lambda)\frac{\partial D(\lambda)}{\partial \lambda}U(\lambda)$$

If we assume $\partial D(\lambda)/\partial \lambda = 0$ $\quad \dfrac{\partial \mathbf{F}(P,N,\lambda)}{\partial \lambda} = i[\mathbf{F}(P,N,\lambda), H(\lambda)]$

$iH(\lambda) = U^{\dagger}(\lambda)(\partial U(\lambda)/\partial \lambda)$

H is Hermitian and cannot dependent on *F* due to separation requirement

Von Neumann equation:

$$i\hbar\frac{\partial \rho(t)}{\partial t} = [H(t), \rho(t)]$$

If $\rho(t) = |\Psi(t)\rangle\langle\Psi(t)|$ (its eigenvalues are not dependent on time in this case!)

we have Schrödinger equation $i\hbar\dfrac{\partial}{\partial t}|\Psi(t)\rangle = H(t)|\Psi(t)\rangle$

but to find the “Hamiltonian” one needs other considerations (e.g. like in logical inference part)

To conclude

The way how we deal organize the “data” adds a lot of restrictions on mathematical apparatus which deals with predictions of outcomes of *uncertain* measurements (QT does not predict individual outcomes): (1) Robustness and (2) Separation of conditions

It is not enough to derive QM as a unique theory, some physics should be added but in restricts enormously a class of possible theories

Unexpected consequence: emergent quantumness in systems which are not quantum per se

A lot of thing to do but, at least, one can replace (some) (quasi)philosophical declarations by calculations – as we like

Thank you