The unreasonable effectiveness of quantum theory: Logical inference approach

Hans De Raedt[1], Mikhail Katsnelson[2],

Hylke Donker[2] and Kristel Michielsen[3]

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

Annals of Physics 359 (2015) 166–186

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

1 University of Groningen
Zernike Institute
for Advanced Materials

2 Radboud Universiteit

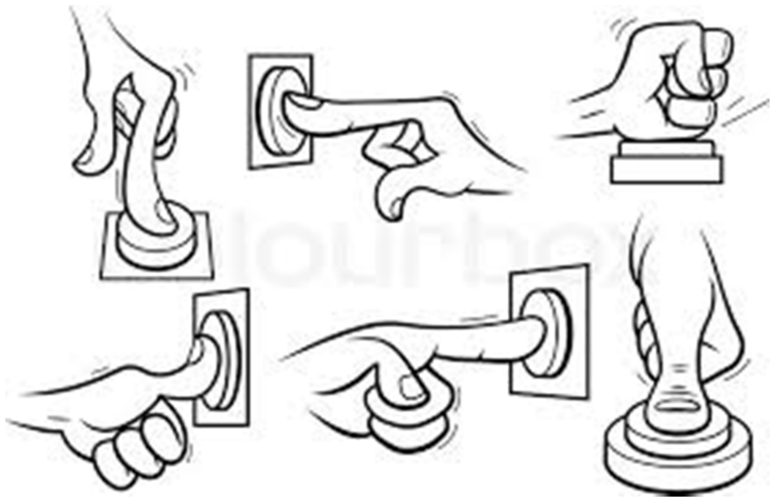
3 JÜLICH
FORSCHUNGSZENTRUM

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"



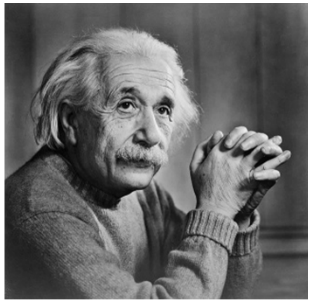
Two ways of thinking II

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?

Is fundamental physics fundamental?

Classical thermodynamics is the only physical theory of universal content which I am convinced will never be overthrown, within the framework of applicability of its basic concepts (A. Einstein)



The laws describing our level of reality are essentially independent on the background laws. I wish our colleagues from *true* theory (strings, quantum gravity, etc....) all kind of success but either they will modify electrodynamics and quantum mechanics at atomic scale (and then they will be wrong) or they will not (and then I do not care). Our way is *down*

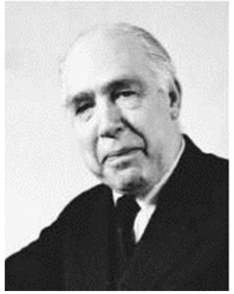
But how can we be sure that we are right?!

Unreasonable effectiveness

- Quantum theory describes a vast number of different experiments very well

- WHY ?

- Niels Bohr*:
  It is wrong to think that the task of physics is to find out how nature *is*.
  Physics concerns what we can *say* about nature.



*A. Petersen, "The philosophy of Niels Bohr," Bulletin of the Atomic Scientists 19, 8 – 14 (1963).

Main message of this talk

- Logical inference applied to experiments for which
  1. There is uncertainty about each individual event
  2. The frequencies of observed events are robust with respect to small changes in the conditions

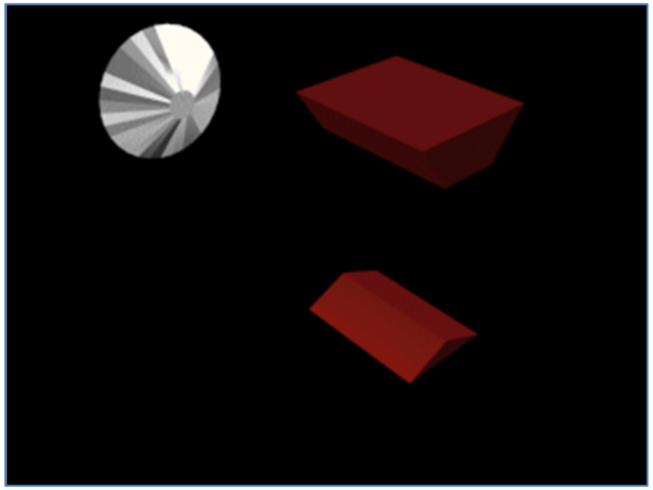
➔ Basic equations of quantum theory

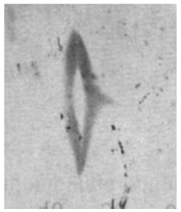
- Not an interpretation of quantum theory

- Derivation based on elementary principles of human reasoning and perception

Stern-Gerlach experiment

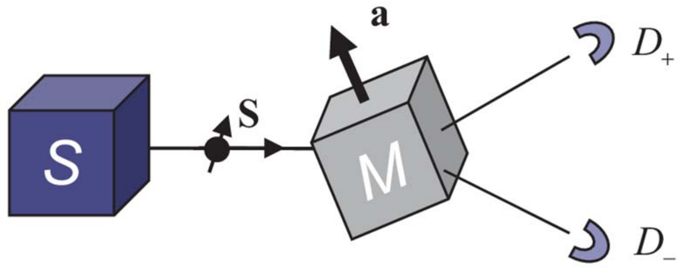
- Neutral atoms (or neutrons) pass through an inhomogeneous magnetic field





- Inference from the data: directional quantization

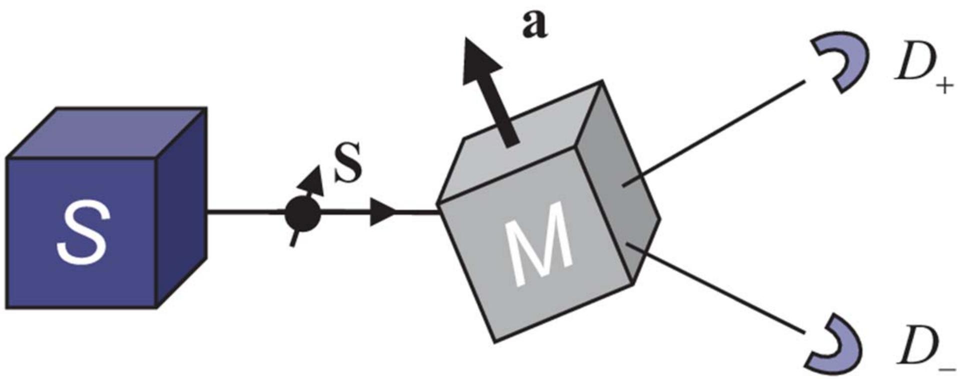
- Idealization



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

Idealized Stern-Gerlach experiment

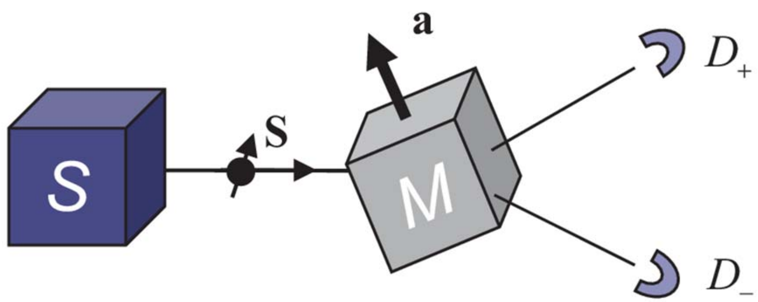
- Event = click of detector D_+ or (exclusive) D_-



- There is uncertainty about each event
  - We do not know how to predict an event with certainty

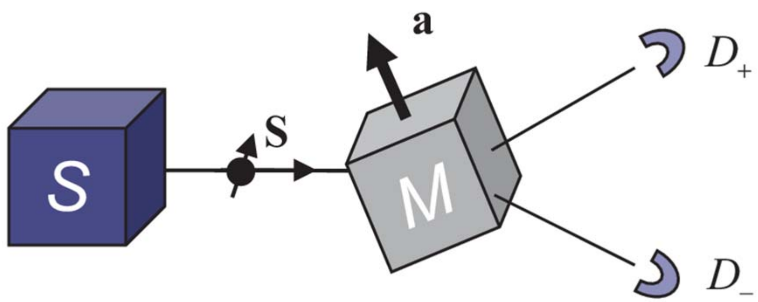
Some reasonable assumptions (1)

- For fixed **a** and fixed source *S*, the frequencies of + and – events are reproducible
- If we rotate the source *S* and the magnet *M* by the same amount, these frequencies do not change



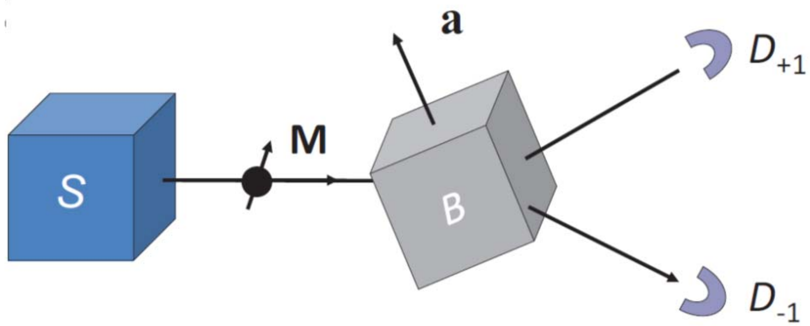
Some reasonable assumptions (2)

- These frequencies are robust with respect to small changes in **a**
- Based on all other events, it is impossible to say with some certainty what the particular event will be (logical independence)



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

$$
\boxed{
\begin{aligned}
&0 \leq P(A|Z) \leq 1 \\
&P(A|Z) + P(\bar{A}|Z) = 1 \quad ; \quad \bar{A} = \text{NOT } A \\
&P(AB|Z) = P(A|BZ)P(B|Z) \quad ; \quad AB = A \text{ AND } B
\end{aligned}
}
$$

  - 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

- 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

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

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

Different events are logically independent:

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

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

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

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

$$Ev(H_1 \mid H_0) = \ln \frac{P(n_+, n_- \mid \theta+\varepsilon, N, Z)}{P(n_+, n_- \mid \theta, N, Z)} = \sum_{x=\pm 1} n_x \ln \frac{P(x \mid \theta+\varepsilon, Z)}{P(x \mid \theta, Z)} =$$

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

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

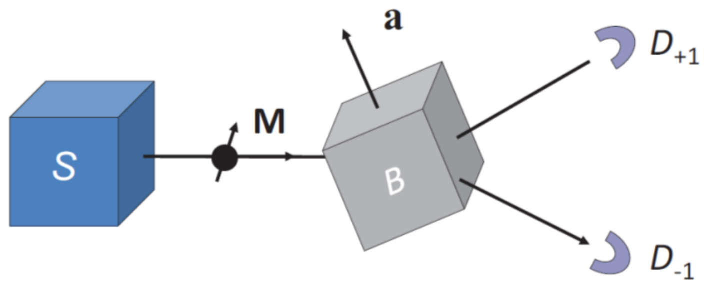
  - Removes the 1st and 3rd term
  - Recover the intuitive procedure of assigning to the i-prob of the individual event, the frequency which maximizes the i-prob to observe the whole data set

Remove dependence on ϵ (2)

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

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

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



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

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

Derivation of basic results of quantum theory by logical inference

- Generic approach
  1. List the features of the experiment that are deemed to be relevant
  2. Introduce the i-prob of individual events
  3. Impose condition of robustness
  4. Minimize functional ➔ equation of quantum theory when applied to experiments in which
     1. There is uncertainty about each event
     2. The conditions are uncertain
     3. Frequencies with which events are observed are reproducible and robust against small changes in the conditions

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

Schrödinger's first "derivation" of his equation*

- Start from time-independent Hamilton-Jacobi equation

$$\boxed{\frac{1}{2m}\left(\frac{\partial S(x)}{\partial x}\right)^2 + V(x) - E = 0}$$

- Write for the action $S(x) = K \ln \psi(x)$ to obtain

$$\boxed{\frac{K^2}{2m}\left(\frac{\partial \psi(x)}{\partial x}\right)^2 + [V(x) - E]\psi^2(x) = 0} \qquad \text{(X)}$$

- "Incomprehensible" step: instead of solving (X), minimize

$$\boxed{Q = \int_{-\infty}^{+\infty} dx \left[\frac{K^2}{2m}\left(\frac{\partial \psi(x)}{\partial x}\right)^2 + [V(x) - E]\psi^2(x)\right]}$$

- to find $\boxed{-K^2 \dfrac{\partial^2 \psi(x)}{\partial x^2} + [V(x) - E]\psi(x) = 0}$

*Ann. Physik 384, 361 (1926) (Band 79 Folge 4 Seite 361)

Schrödinger's second "justification" of his equation*

1. *Quantisierung als Eigenwertproblem;*
von E. Schrödinger.

(Zweite Mitteilung.)[1]

§ 1. **Die Hamiltonsche Analogie zwischen Mechanik und Optik.**

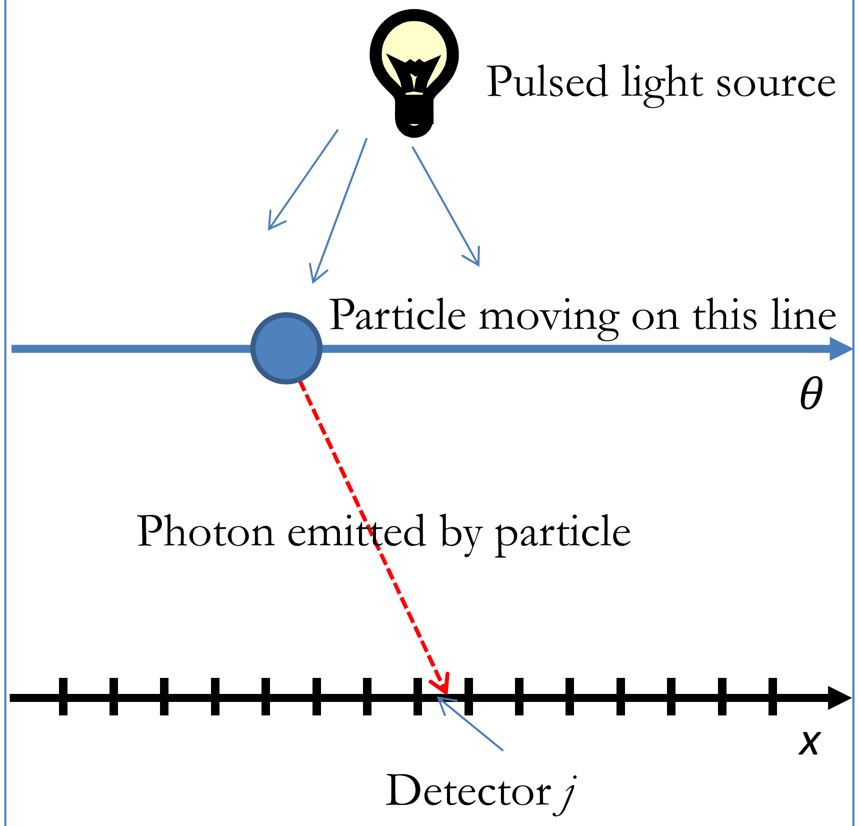
Bevor wir daran gehen, das Eigenwertproblem der Quantentheorie für weitere spezielle Systeme zu behandeln, wollen wir den *allgemeinen* Zusammenhang näher beleuchten, welcher zwischen der Hamiltonschen partiellen Differentialgleichung (H. P.) eines mechanischen Problems und der „zugehörigen" *Wellengleichung*, d. i. im Falle des Keplerproblems der Gleichung (5) der ersten Mitteilung, besteht. Wir hatten diesen Zusammenhang vorläufig nur kurz seiner äußeren analytischen Struktur nach beschrieben durch die an sich unverständliche Transformation (2) und den ebenso unverständlichen Übergang von der *Nullsetzung* eines Ausdrucks zu der Forderung, daß das *Raumintegral* des nämlichen Ausdruckes *stationär* sein soll.[2]

- Calls his first derivation "incomprehensible" and then gives a "physical" justification of his equation by analogy with optics
- From the viewpoint of rational reasoning, the first derivation is all but incomprehensible

*Ann.der.Physik 384, 489 (1926) (Band 79 Folge 4 Seite 489)

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 : $\boxed{P(x\,|\,\theta, Z)}$



Robustness

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

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

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

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

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

Impose classical mechanics (á la Schrödinger)

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

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

- If there is "known" uncertainty

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

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

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*

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

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

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

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

$\rightarrow$ 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 $0 \le k_{\boldsymbol{j},\tau} \le N$, 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:

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

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:

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

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

Time-dependent case V

Time-dependent Schrödinger equation

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

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).

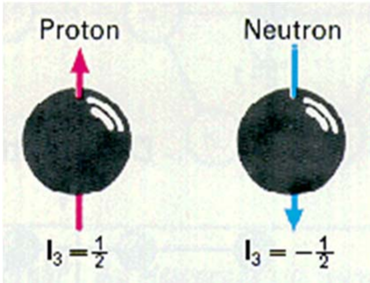
A very nontrivial operation dictated just by desire to simplify the problem as much as possible (to pass from nonlinear to linear equation).

Requires further careful thinking!

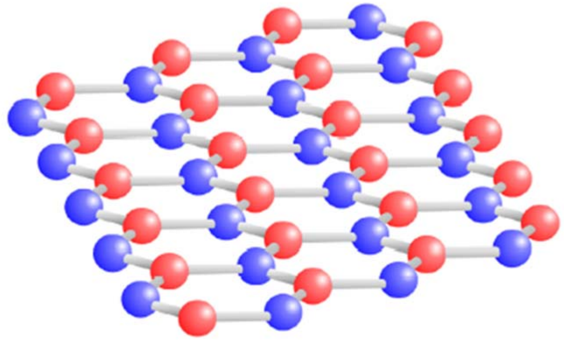
Pauli equation

What is spin? Just duality (e.g., color – blue or red). Nothing is rotating (yet!)

Isospin in nuclear physics



Sublattice index in graphene



Just add color (k=1,2)

The result of N repetitions of the experiment yields the data set

$$\Upsilon = \{(\mathbf{j}, k)_{n,\tau} | \mathbf{j}_{n,\tau} \in \mathcal{V};\ k = \pm 1;\ n = 1, \ldots, N;\ \tau = 1, \ldots, M\}, \tag{1}$$

or, denoting the total counts of voxels $\mathbf{j}$ and color k at time τ by $0 \leq c_{\mathbf{j},k,\tau} \leq N$, the data set can be represented as

$$\mathcal{D} = \left\{ c_{\mathbf{j},k,\tau} \middle| \tau = 1, \ldots, M\ ;\ \sum_{k=\pm 1} \sum_{\mathbf{j} \in [-L^d, L^d]} c_{\mathbf{j},k,\tau} = N \right\}. \tag{2}$$

Pauli equation II

Fisher information part just copies the previous derivation

$$\mathrm{Ev} = \ln \frac{P(\mathcal{D}|\mathbf{X}_\tau + \boldsymbol{\varepsilon}_\tau, \tau, N, Z)}{P(\mathcal{D}|\mathbf{X}_\tau, \tau, N, Z)}$$
$$= \sum_{\mathbf{j},k,\tau} c_{\mathbf{j},k,\tau} \ln \frac{P(\mathbf{j}, k|\mathbf{X}_\tau + \boldsymbol{\varepsilon}_\tau, \tau, Z)}{P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)}$$

for the evidence

Expansion

$$\mathrm{Ev} = \sum_{\mathbf{j},k,\tau} c_{\mathbf{j},k,\tau} \ln \left[1 + \frac{\boldsymbol{\varepsilon}_\tau \cdot \boldsymbol{\nabla}_\tau P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)}{P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)} + \frac{1}{2} \frac{(\boldsymbol{\varepsilon}_\tau \cdot \boldsymbol{\nabla}_\tau)^2 P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)}{P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)} + \mathcal{O}(\varepsilon_\tau^3) \right]$$
$$= \sum_{\mathbf{j},k,\tau} c_{\mathbf{j},k,\tau} \left[\frac{\boldsymbol{\varepsilon}_\tau \cdot \boldsymbol{\nabla}_\tau P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)}{P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)} - \frac{1}{2} \left[\frac{\boldsymbol{\varepsilon}_\tau \cdot \boldsymbol{\nabla}_\tau P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)}{P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)} \right]^2 \right.$$
$$\left. + \ \frac{1}{2} \frac{(\boldsymbol{\varepsilon}_\tau \cdot \boldsymbol{\nabla}_\tau)^2 P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)}{P(\mathbf{j}, k|\mathbf{X}_\tau, \tau, Z)} \right] + \mathcal{O}(\boldsymbol{\varepsilon}_\tau^3),$$

Pauli equation III

$$I_F = \sum_{\mathbf{j},k,\tau} \frac{1}{P(\mathbf{j},k|\mathbf{X}_\tau,\tau,Z)}\left[\nabla_\tau P(\mathbf{j},k|\mathbf{X}_\tau,\tau,Z)\right]^2 = \int d\mathbf{x}\,dt \sum_{k=\pm 1} \frac{1}{P(\mathbf{x},k|\mathbf{X},t,Z)}\left[\nabla P(\mathbf{x},k|\mathbf{X},t,Z)\right]^2$$

$$P(\mathbf{x},k=+1|\mathbf{X},t,Z) = P(\mathbf{x}|\mathbf{X},t,Z)\cos^2\frac{\theta(\mathbf{x},\mathbf{X},t,Z)}{2}$$

$$P(\mathbf{x},k=-1|\mathbf{X},t,Z) = P(\mathbf{x}|\mathbf{X},t,Z)\sin^2\frac{\theta(\mathbf{x},\mathbf{X},t,Z)}{2}$$

$$I_F = \int d\mathbf{x}\,dt \left\{ \frac{1}{P(\mathbf{x}|\mathbf{X},t,Z)}\left[\nabla P(\mathbf{x}|\mathbf{X},t,Z)\right]^2 + \left[\nabla\theta(\mathbf{x},\mathbf{X},t,Z)\right]^2 P(\mathbf{x}|\mathbf{X},t,Z) \right\}$$

$\theta(\mathbf{x},\mathbf{X},t,Z)$ has no dynamical or geometric meaning (yet)

Pauli equation IV

Dynamical part is less trivial; we restrict ourselves only by d=3 (spin is introduced in 3D space, it is important!)

Alternative representation of the Newton's laws (or HJE)

$$\frac{d\mathbf{x}}{dt} = \mathbf{U}(\mathbf{x}, t)$$ The velocity field is derived by (numerical) differentiation of position data

Decomposition for any vector field in 3D:

$$\mathbf{U}(\mathbf{x}, t) = \nabla S(\mathbf{x}, t) - \mathbf{A}(\mathbf{x}, t)$$

$$\mathbf{A}(\mathbf{x}, t) = \nabla \times \mathbf{W}(\mathbf{x}, t)$$

$$\nabla \cdot \mathbf{A} = 0$$

Direct differentiation:

$$\begin{aligned}
\frac{d^2 x_i}{dt^2} &= \frac{\partial U_i}{\partial t} + \sum_{j=1}^{3} \frac{\partial U_i}{\partial x_j} U_j \\
&= \frac{\partial^2 S}{\partial x_i \partial t} - \frac{\partial A_i}{\partial t} + \sum_{j=1}^{3} \left(\frac{\partial^2 S}{\partial x_i \partial x_j} - \frac{\partial A_i}{\partial x_j} \right) \left(\frac{\partial S}{\partial x_j} - A_j \right) \\
&= \frac{\partial}{\partial x_i} \left[\frac{\partial S}{\partial t} + \frac{1}{2} \sum_{j=1}^{3} \left(\frac{\partial S}{\partial x_j} - A_j \right)^2 \right] + \sum_{j=1}^{3} \left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} \right) \left(\frac{\partial S}{\partial x_j} - A_j \right) - \frac{\partial A_i}{\partial t}
\end{aligned}$$

Pauli equation V

Introducing the vector field $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$

$$\frac{d^2\mathbf{x}}{dt^2} = \boldsymbol{\nabla}\left[\frac{\partial S}{\partial t} + \frac{1}{2}(\boldsymbol{\nabla} S - \mathbf{A})^2\right] + \frac{d\mathbf{x}}{dt} \times \mathbf{B} - \frac{\partial \mathbf{A}}{\partial t}$$

Hypothesis (alternative form of HJE): Existence of scalar field $\phi(\mathbf{x}, t)$

such that $$\frac{\partial S}{\partial t} + \frac{1}{2}(\boldsymbol{\nabla} S - \mathbf{A})^2 = -\phi$$

Then, upon introducing the vector field $\mathbf{E} = -\boldsymbol{\nabla}\phi - \partial \mathbf{A}/\partial t$,

$$\frac{d^2\mathbf{x}}{dt^2} = \mathbf{E} + \frac{d\mathbf{x}}{dt} \times \mathbf{B}.$$

Nothing but equation of motion of particle in electromagnetic field (in proper units)

Pauli equation VI

Dynamical information on the system (constrain):

we will require that there exists two scalar fields $V_k(\mathbf{x}, t)$ for $k = \pm 1$ such that

$$\int d\mathbf{x}\, dt \sum_{k=\pm 1} \left[\frac{\partial S_k(\mathbf{x}, t)}{\partial t} + \frac{1}{2m} \left(\nabla S_k(\mathbf{x}, t) - q\mathbf{A}(\mathbf{x}, t) \right)^2 + V_k(\mathbf{x}, t) \right] P(\mathbf{x}, k|\mathbf{X}, t, Z) = 0,$$

$$S_k(\mathbf{x}, t) = S(\mathbf{x}, t) - kR(\mathbf{x}, t) \text{ for } k = \pm 1$$

$$V_0(\mathbf{x}, t) = [V_{+1}(\mathbf{x}, t) + V_{-1}(\mathbf{x}, t)]/2,\; V_1(\mathbf{x}, t) = [V_{+1}(\mathbf{x}, t) - V_{-1}(\mathbf{x}, t)]/2$$

Constrain functional:

$$\begin{aligned}
\Lambda &= \int d\mathbf{x}\, dt \sum_{k=\pm 1} \left[\frac{\partial S_k(\mathbf{x}, t)}{\partial t} + \frac{1}{2m} \left(\nabla S_k(\mathbf{x}, t) - q\mathbf{A}(\mathbf{x}, t) \right)^2 + V_k(\mathbf{x}, t) \right] P(\mathbf{x}, k|\mathbf{X}, t, Z) \\
&= \int d\mathbf{x}\, dt \left\{ \frac{1}{2m} \left[\left(\nabla S(\mathbf{x}, t) - q\mathbf{A}(\mathbf{x}, t) \right)^2 + \left(\nabla R(\mathbf{x}, t) \right)^2 \right.\right. \\
&\quad \left. - 2\cos\theta(\mathbf{x}, \mathbf{X}, t, Z) \nabla R(\mathbf{x}, t) \left(\nabla S(\mathbf{x}, t) - q\mathbf{A}(\mathbf{x}, t) \right) \right] \\
&\quad + \left[\frac{\partial S(\mathbf{x}, t)}{\partial t} - \cos\theta(\mathbf{x}, \mathbf{X}, t, Z) \frac{\partial R(\mathbf{x}, t)}{\partial t} \right] + V_0(\mathbf{x}, t) \\
&\quad \left. + V_1(\mathbf{x}, t) \cos\theta(\mathbf{x}, \mathbf{X}, t, Z) \right\} P(\mathbf{x}|\mathbf{X}, t, Z),
\end{aligned}$$

Pauli equation VII

Up to know we did not assume that "color" is related to any rotation or any magnetic moment. But we know experimentally (anomalous Zeeman effect) that electron has magnetic moment, with its energy in external magnetic field $\widehat{V}_1(\mathbf{x},t) = -\gamma\mathbf{m}(\mathbf{x},t)\cdot\mathbf{B}(\mathbf{x},t)$. We have correct classical equations of precession if we identify $\theta(\mathbf{x},\mathbf{X},t,Z)$ and $\varphi(\mathbf{x},t) = R(\mathbf{x},t)/a$ with the polar angles of the unit vector $\mathbf{m}(\mathbf{x},t)$

$$
\begin{aligned}
\Lambda = \int d\mathbf{x}\,dt \Bigg\{ & \frac{1}{2m}\Big[(\nabla S(\mathbf{x},t) - q\mathbf{A}(\mathbf{x},t))^2 + a^2\big(\nabla\varphi(\mathbf{x},t)\big)^2 \\
& - 2a\cos\theta(\mathbf{x},\mathbf{X},t,Z)\nabla\varphi(\mathbf{x},t)\,(\nabla S(\mathbf{x},t) - q\mathbf{A}(\mathbf{x},t))\Big] \\
& + \left[\frac{\partial S(\mathbf{x},t)}{\partial t} - a\cos\theta(\mathbf{x},\mathbf{X},t,Z)\frac{\partial\varphi(\mathbf{x},t)}{\partial t}\right] + V_0(\mathbf{x},t) \\
& - a\gamma\mathbf{m}(\mathbf{x},t)\cdot\mathbf{B}(\mathbf{x},t)\Bigg\}P(\mathbf{x}|\mathbf{X},t,Z).
\end{aligned}
$$

Pauli equation VIII

$$V_0(\mathbf{x},t) = q\phi(\mathbf{x},t),\ a = \hbar/2,\ \gamma = q/m \qquad \lambda = \hbar^2/8m$$

$$\Phi(\mathbf{x},t) = \begin{pmatrix} P^{1/2}(\mathbf{x}, k=+1|\mathbf{X}, t, Z)e^{iS_1(\mathbf{x},t)/\hbar} \\ P^{1/2}(\mathbf{x}, k=-1|\mathbf{X}, t, Z)e^{iS_2(\mathbf{x},t)/\hbar} \end{pmatrix}$$

Extremum of the functional $F = \lambda I_F + \Lambda$

gives the Pauli equation: $i\hbar\frac{\partial}{\partial t}\Phi = H\Phi$

$$\begin{aligned} H &= \frac{1}{2m}\{\boldsymbol{\sigma}\cdot[-i\hbar\nabla - q\mathbf{A}(\mathbf{x},t)]\}^2 + q\phi(\mathbf{x},t) \\ &= \frac{1}{2m}[-i\hbar\nabla - q\mathbf{A}(\mathbf{x},t)]^2 + q\phi(\mathbf{x},t) - \frac{q\hbar}{2m}\boldsymbol{\sigma}\cdot\mathbf{B}(\mathbf{x},t) \end{aligned}$$

$$\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)^T$$

What to do next?

1. Relativistic wave equations. Currently: Klein-Gordon is basically done, for Dirac – new ideas are required.
2. Understanding the origin of wave function and linear operators: why matrices and state vectors? The first step (Stern-Gerlach and Einstein-Podolsy-Rosen-Bohm experiments): derived from "separability" condition $\langle x \rangle = \mathbf{Tr}\,\widehat{\rho}\widehat{X}$

where *X* and *ρ* characterize only setup and only system, e.g., for SG

$$\widehat{\rho} = (\mathbb{1} + \mathbf{M} \cdot \boldsymbol{\sigma})/2 \quad , \quad \widehat{X} = \mathbf{a} \cdot \boldsymbol{\sigma},$$

arXiv:1506.03373 [pdf, ps, other] Quantum theory as a description of robust experiments: Application to Stern-Gerlach and Einstein-Podolsky-Rosen-Bohm experiments
H. De Raedt, M.I. Katsnelson, H.C. Donker, K. Michielsen
Comments: Paper to be presented at SPIE15

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

Thank you