Four-dimensional understanding of quantum mechanics and Bell violation

Jarek Duda
Jagiellonian University, Golebia 24, 31-007 Krakow, Poland, Email: *dudajar@gmail.com*

***Abstract*—While our natural intuition suggests us that we live in 3D space evolving in time, modern physics presents fundamentally different picture: 4D spacetime, Einstein's block universe, in which we travel in thermodynamically emphasized direction: arrow of time. Suggestions for such nonintuitive and nonlocal living in kind of "4D jello" come among others from: Lagrangian mechanics we use from QFT to GR saying that history between fixed past and future situation is the one optimizing action, special relativity saying that different velocity observers have different present 3D hypersurface and time direction, general relativity deforming shape of the entire spacetime up to switching time and space below the black hole event horizon, or the CPT theorem concluding fundamental symmetry between past and future for example in the Feynman-Stueckelberg interpretation of antiparticles as propagating back in time.**

Accepting this nonintuitive living in 4D spacetime: with present moment being in equilibrium between past and future - minimizing tension as action of Lagrangian, leads to crucial surprising differences from intuitive "evolving 3D" picture, in which we for example conclude satisfaction of Bell inequalities - violated by the real physics. Specifically, particle in spacetime becomes own trajectory: 1D submanifold of 4D, making that statistical physics should consider ensembles like Boltzmann distribution among entire paths, what leads to quantum behavior as we know from Feynman's Euclidean path integrals or similar Maximal Entropy Random Walk (MERW). It results for example in Anderson localization, or the Born rule with squares - allowing for violation of Bell inequalities. As e.g. for S-matrix, quantum amplitude turns out to describe probability at the end of half-spacetime from a given moment toward past or future, to randomly get some value of measurement we need to "draw it" from both time directions, getting the squares of Born rules. Tension in both time directions is also suggested in quantum experiments like Wheeler's delayed choice experiment, it will be argued that it is also crucial in quantum algorithms like Shor's.

Keywords: quantum mechanics, nature of time, spacetime, Einstein's block universe, Born rule, Bell inequalities, Shor algorithm, Euclidean path integrals, statistical physics, Maximal Entropy Random Walk

I. Introduction

Starting with special relativity (SR) a century ago, modern physics uses 4D spacetime view of our world - Einstein's block universe in which we travel in time direction. Also a century ago quantum mechanics (QM) was born, bringing many nonintuitive consequences, like violation of Bell inequalities. As briefly presented in Fig. 1, 2, 3 and 4, this article argues that these two revolutions of our understanding - violating our natural intuitions, are in fact

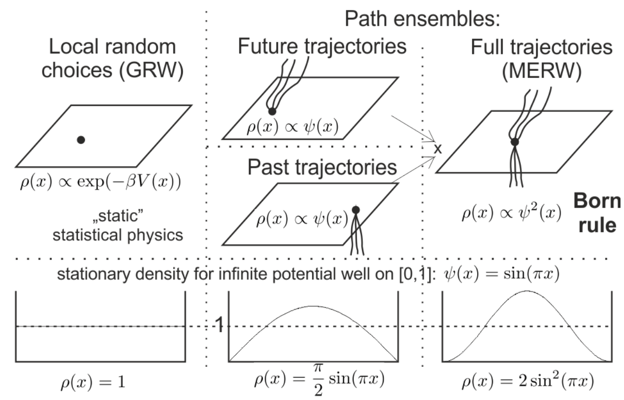

Figure 1. Three philosophies for finding probability distribution of a walker, for example electron. Left: "static 3D" situation. Center: example of "evolving 3D" philosophy of time, assuming Boltzmann distribution among trajectories from the past to the current moment. As it will be discussed, it leads to probability distribution proportional to the first power of quantum ground state amplitude. Right: "4D spacetime" situation where particle becomes its full trajectory. Like in Feynman's Euclidean path integrals, Boltzmann distribution among such full trajectories leads to Born rule: focusing on a fixed-time cut, to randomly get a given position, we need "to draw it" from both past and future half-trajectories. Hence, probability is proportional to the product of two (identical here) "evolving 3D" probabilities, getting the Born rule. Bottom: stationary probability distribution predicted by them for $[0, 1]$ infinite potential well, only the "spacetime" consideration agrees with the QM ground state.

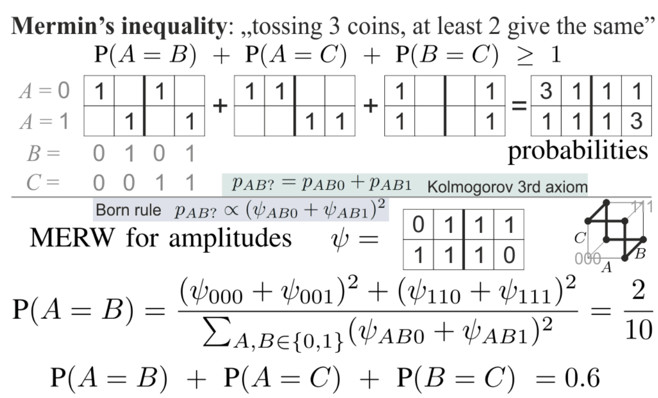

Figure 2. Top: schematic proof of simple Bell-like inequalities ([1], [2]): assuming any hidden probability distribution among 8 possibilities $\{0,1\}^3$ for 3 binary variables, we get inequality written at the top, which can be violated in QM. Bottom: example of its violation using Born rules like in MERW: as in Fig. 1 amplitude corresponds to probability toward one of two time directions, to randomly get some value we need to "draw it" twice, getting probability as normalized square of sum of amplitudes. Amplitude for this violation example can be obtained by MERW using any regular graph on the 6 vertices. Details in Section III.

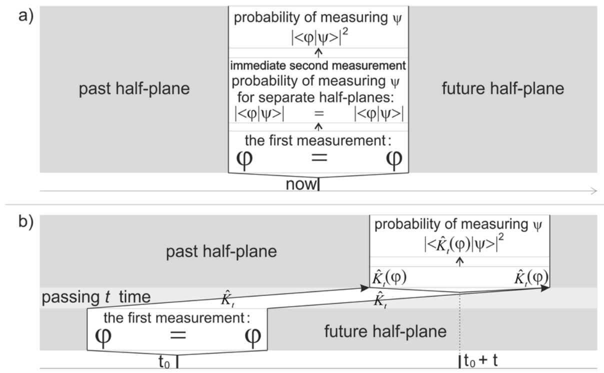

Figure 3. The lesson from Fig. 1 taken to quantum evolution and measurement: in the "spacetime" philosophy, quantum amplitude describes not only probability distribution at the end of the past, but simultaneously also at the beginning of the future. To calculate probability of some value of measurement in a fixed time cut, we need to multiply the (identical) predicted probabilities for the past and future directions, getting Born rule: squares known in QM formalism, which lead e.g. to correlations violating Bell's inequalities.

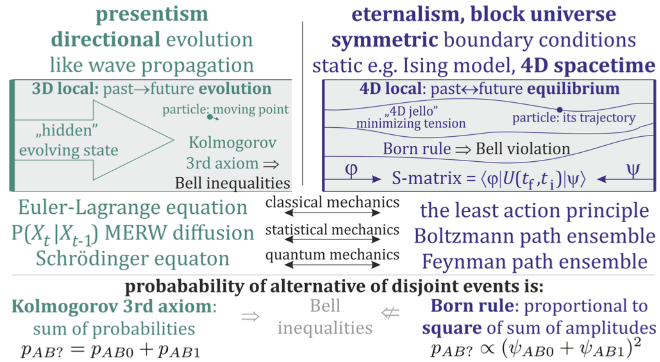

Figure 4. Two philosophies of time: presentism evolving current 3D situation, and eternalism, block universe imagining travel through some already chosen 4D solution as past-future equilibrium - like in "4D jello", shape of 4D spacetime in the general relativity (trying "to develop" spacetime with Euler-Lagrange equations seems highly problematic). We can translate between such two types of solutions e.g. in classical, quantum, statistical mechanics. However, solutions originally found with one of them have subtle differences, 3D locality suggested by our intuition is different from 4D locality. For example while directional solutions should satisfy Bell inequalities, symmetric ones allow to replace Kolmogorov 3rd axiom with Born rule, allowing for violation of Bell-like inequalities. We will focus here on MERW as diffusion chosen accordingly to the maximal entropy principle - as required by statistical mechanics models.

deeply connected: that living in spacetime has surprising consequences present in QM formalism. Let us start with reminding well known arguments and consequences of the spacetime view of modern physics.

Our natural human intuition has evolved to handle past-future reason-result chains of consequences: initiated in our Big Bang, leading to us through creation of our planet, evolution, our development. However, since the special relativity we know that "situation in a given moment", more formally called the hypersurface of the present, in fact depends on the observer's frame of reference: it changes with his velocity accordingly to Lorentz boost, which also modifies the direction of time. The general relativity takes it even further, modifying the entire spacetime accordingly to local mass/energy concentrations, up to extreme situation below the black hole event horizon, where time and space directions literally switch places, making "situation in a given moment" far from our biological intuition.

This ambiguity of time direction is also seen in the Lagrangian formalism we successfully use to describe reality in all scales: from quantum field theories to the general relativity. It has multiple equivalent formulations, starting with our intuitive "evolving 3D" picture: Euler-Lagrange equation allows to evolve the situation forward in time, from a situation (as values and the first derivatives) in a given moment. However, mathematically it also allow to use these equations to evolve situation backward in time, as Lagrangian mechanics is usually time (or CPT) symmetric. Quite different formulation is through action optimization: fixing situation (only values, without the first derivatives) in two moments in time, the history between them is the one optimizing action. While we can translate between such solutions, those originally found with each of them have subtle differences, visualized in Fig. 4.

Lagrangian mechanics for field theories can additionally be Lorentz invariant: compatible with changing the direction of time by Lorentz boosts. To get some intuition, let us briefly remind the simplest scalar field theory: with Hamiltonian (energy density) and Lagrangian:

$$\mathcal{H} = \frac{1}{2}\sum_{\nu=0,1,2,3} (\partial_\nu \phi)^2 + V(\phi)$$

$$\mathcal{L} = \frac{1}{2}\left((\partial_0 \phi)^2 - \sum_{i=1,2,3} (\partial_i \phi)^2\right) - V(\phi)$$

Surprisingly, energy density (Hamiltonian) is often completely 4D symmetric like here: does not emphasize any time direction in 4D. Choosing a frame of reference, it determines time direction '0', for which we can find the Lagrangian which Legendre transform is the given Hamiltonian ($\pi = \partial_0\phi,\ \mathcal{L} = \pi\,\partial_0\phi - \mathcal{H}$). This Lagrangian emphasizes the chosen time direction. To summarize, energy density (Hamiltonian) of a Lorentz invariant field theory often allows to imagine the spacetime as completely symmetric "4D jello": minimizing tension as Hamiltonian. Choosing some time direction and situation in its hypersurface of the present, we can transform Hamiltonian to Lagrangan, find Euler-Lagrange equation for it, and use it to evolve situation from this arbitrary hyperplane in a chosen direction of time.

Fundamental similarity of past and future is also the base of quantum field theories requiring CPT symmetry due to the Schwinger's CPT theorem [3]. Feynman diagrams represent antiparticles as propagating backward in time (Feynman-Stueckelberg interpretation) [4].

In contrast, against e.g. SR and fundamental CPT symmetry, 2nd law of thermodynamics wants to emphasize some "arrow of time" direction in spacetime

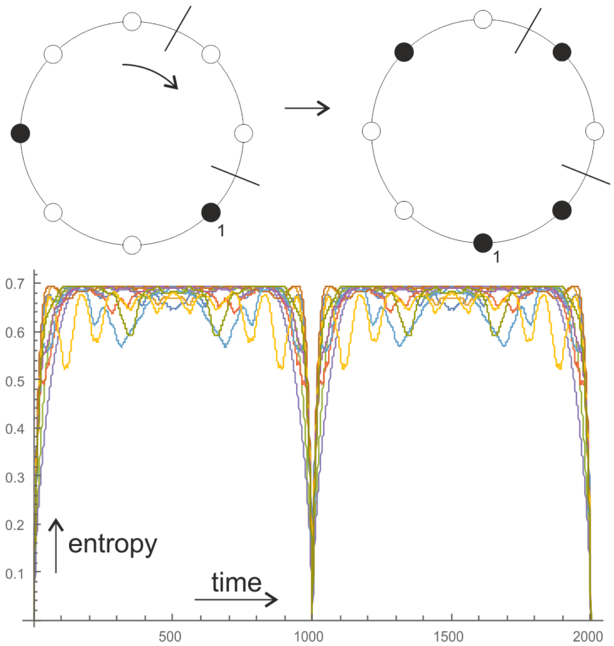

Figure 5. Kac ring [5]: valuable lesson regarding entropy growth in fundamentally time-symmetric models. We have a ring with n white and black balls, and there are chosen some marked positions (two drawn segments). In each step all balls rotate by one position, each ball going through a marked position flips color. Denote $p \in [0,1]$ as the current number of white balls divided by n. A natural looking assumption (Stosszahlansatz) is that this p also describes proportion of colors of balls before the marked positions - this assumption allows to conclude that $p \to 1/2$, tending to maximize the entropy as $h(p) = -p\lg(p) - (1-p)\lg(1-p)$. However, if starting from all white balls and performing n steps, all balls go through the same number of flips, returning to the zero entropy situation. The bottom plots contain entropy for 10 simulations with $n = 500$ balls and randomly chosen marked positions with 0.01 probability. Analogously, for also time-symmetric classical mechanics, the Stosszahlansatz used in Boltzmann H-theorem [6] to conclude entropy growth is assumption that two particles participating in a collision have independently randomly chosen energies, directions and starting points.

(and its "arrow"). However, thermodynamics is not fundamental, only effective modelling: describes the most probable statistical behavior, averaging over unknowns. While physics fundamentally suggests quite symmetric "4D jello", this symmetry is clearly broken on thermodynamical level in the actual solution we live in. Like a fundamentally symmetric water surface can obtain a state (solution) with this symmetry broken e.g. by throwing a rock. As proven for example in Boltzmann H-theorem [6], entropy growth is a natural statistical tendency even for time-symmetric models, what seems self-contradictory as after applying such symmetry, the same proof should conclude opposite entropy growth. Figure 5 presents a simple Kac model which gives a valuable lesson about this apparent paradox - such proofs of entropy growth have to rely on looking natural assumptions of some uniformity (called Stosszahlansatz), allowing to break time-symmetry of the model. However, hidden structure of such time-symmetric system can also lead to entropy decrease and bouncing from zero entropy

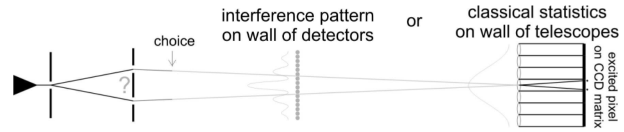

Figure 6. Wheeler experiment as one of QM examples suggesting future-past causality: imagine there is some double-slit type of experiment in a distance, for example light from a star can pass near one of sides of a planet on the way. Later (by travel time) we choose how to observe it: if resolution of our telescope allows to distinguish the two slits we get corpuscular behavior, otherwise we get wave behavior: quantum interference. Hence it seems like in the future we can choose between wave or particle past nature of photons. In experimental confirmation [7] there was used Mach-Zehnder setting: choosing between classical and quantum photon behavior (lifting or not the last half-silvered mirror) after it went through the first half-silvered mirror. There are also more sophisticated QM experiments suggesting future-past causality, like the delayed choice quantum erasure [8] or experiment from "Asking photons where they have been" article [9].

as in the plot in Fig. 5. The direct reason for entropy growth with our time arrow might be our Big Bang: having all matter localized and so low entropy, starting the cascade of reason-result chains of consequences leading to our current situation. Assuming our Universe will finally gravitationally collapse and bounce starting new Universe, the fundamental CPT symmetry of our physics suggests that such Big Bounce situation might be also nearly symmetric from the point of view of entropy, like in bounces for Kac ring in Fig. 5.

While it is difficult for us to really accept, we see that especially special relativity and Lagrangian mechanics for fields provide a picture that, against our natural intuitions, we live in spacetime as kind of "4D jello" minimizing tension defined by energy density (Hamiltonian). Hence, the present moment is kind of equilibrium between past and future situation (like in time/CPT symmetry of fundamental theories we use), what makes physics nonlocal in "evolving 3D" sense (but local in 4D view). In contrast, our intuition considers only consequences from the past time direction. Nature provides many suggestions that the resulting non-intuitiveness is connected with the strangeness of quantum mechanics, like the Wheeler experiment briefly presented in Fig. 6, or the Delayed Choice Quantum Erasure. For example there is Cramer's transactional interpretation of QM [10] based on this inherent time symmetry of quantum unitary evolution, suggesting propagation of information in both time directions.

Time symmetric formulation of QM is also advocated by Aharonov [11] in so called two-state vector formalism: seeing the present moment as a result of two propagators: from minus and plus infinity, what is analogous to view presented here. Article "Asking photons where they have been" [9] presents its very nice experimental conformation: vibrating mirrors allows to conclude from the final light beam which mirrors have been visited. However, in properly chosen setting they obtain also signal from mirrors which naively should not be visited, unless we

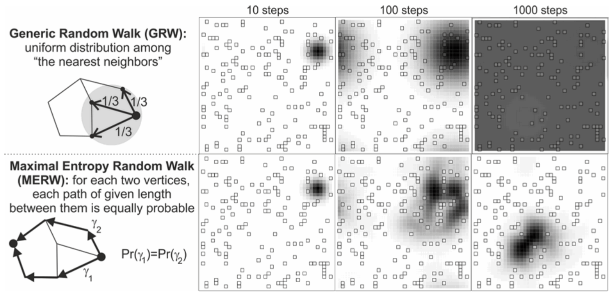

Figure 7. Left: GRW and one of formulations of MERW for a simple graph. Right: example of comparison of their evolution for a defected 40×40 lattice with cyclic boundary conditions - all but the marked vertices have self-loop: edge to itself. Stationary probability density of GRW is proportional to degree of vertex, getting very weak localization: defects have $4/5$ probability of the remaining nodes. In contrast, while MERW seems to have similar local behavior (density after 10 steps), it leads to a completely different stationary density, exactly like for the quantum ground state: square of coordinates of the dominant eigenvector of minus Hamiltonian (simplified Bose-Hubbard here: $\mathcal{H} = -\sum_{\text{edge }(ij)} a_j^+ a_i$). Such defected lattice can be seen as a simple model of semiconductor: regular lattice with rare dopants. Nearly uniform electron density predicted by standard diffusion (GRW) means that electrons should flow if attached electric field - making it a conductor against experiment. In contrast, QM and MERW predict electron prisoned e.g. in local defect-free regions (called Lifhitz spheres). Simple GRW/MERW conductance simulator is available in [14]. MERW transition probabilities use eigenvector and so it depends on the entire system, making this model nonlocal. However, in such thermodynamical view the walker does not directly use these nonlocal probabilities, only we use them to predict the most probable evolution of its probability distribution accordingly to our knowledge.

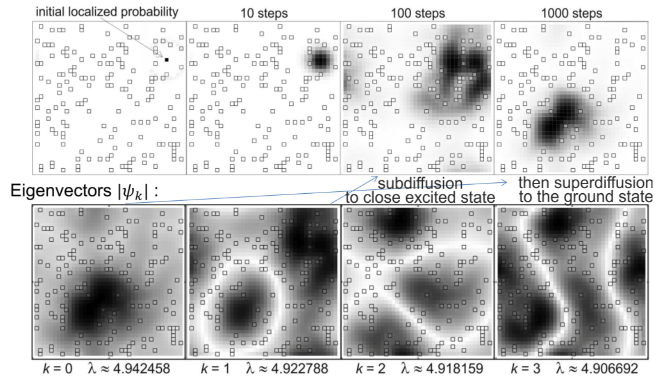

Figure 8. While MERW evolution leads to stationary probability distribution exactly as in the quantum ground state, the excited quantum states are also seen in this evolution. Specifically, the coefficient corresponding to $M\psi_i = \lambda_i \psi_i$ eigenfunction/eigenvector, drops $\approx (\lambda_i/\lambda_0)^t$ times during t steps, where λ_0 is the dominant (largest) eigenvalue. As energy turns out to correspond to $-\lambda$, we get exponential weakening of contributions of excited states - in contrast to standard QM they are not stable, however, excited states in nature are also unstable. The top row of above diagram shows example of such evolution: as initial contribution of the ground state is very small, the evolution first gets close to the first excited state (subdiffusion), then it kind of tunnels to the ground state (superdiffusion). This simple model neglects conservation laws e.g. of energy, which might prevent the walker from going to the ground state - its perturbed averaged trajectory should instead smoothen toward some close excited probability distribution. Also for multiple repelling walkers (like electrons), they should choose densities of successive eigenstates here (like in Pauli exclusion principle).

focus exactly on mirrors visited by photons propagating in both time directions.

Here we will focus on statistical consequences of given moment (hypersurface of the present) being in equilibrium between past and future in spacetime as kind of "4D jello". In this picture particles are no longer just "moving points", but rather their trajectories: one-dimensional submanifolds of the spacetime. From statistical physics perspective, it brings the question of what objects should we use in the considered ensembles, e.g. while assuming Boltzmann distribution like presented in Fig. 1 - only the last one: considering Boltzmann distribution among full trajectories, like in Feynman's Euclidean path integrals or related Maximal Entropy Random Walk ([12], [13]), has thermodynamical agreement with predictions of QM: leading to stationary probability distribution of the quantum ground state, with crucial differences like avoiding the boundaries for $[0, 1]$ infinite potential well.

In the next Section we will focus on MERW philosophy, briefly presented in Fig. 7, as a reparation of standard diffusion models - which for example wrongly predict that semiconductor should have nearly uniform probability distribution of electrons, making it a conductor against experiment (tiny electric field would cause electron flow). This fundamental disagreement was repaired by QM showing strong localization property for these electrons (e.g. Anderson), preventing the conductance. This crucial problem of stochastic modelling has caused that it is currently seen as a completely different realm than QM. However, electrons are indivisible quants of electric charge, what should prevent them from being objectively blurred. Heisenberg uncertainty principle limits abilities of measurement, which are sophisticated destructive processes idealized for example by the Stern-Gerlach experiment. In contrast, this principle is commonly seen as limitation for example for objective position of electron in atom. However, modern techniques like field-emission electron microscopy already allow to get resolution below the size of atomic orbital: strip electrons from single atom, use EM field shaped to act as a lens for magnification, and measure positions of theses single electron in detector matrix [15]. This way they literally obtained photography of orbitals: densities by averaging over positions of single electrons. Anyway, using Heisenber principle as an excuse for ignoring questions about objective dynamics is no longer valid when increasing the scale, for example while asking question about local currents in a lattice: what is the probability distribution for electron jumping to neighboring parts of this lattice - getting a stochastic model for its dynamics (conductance). There are also examples of larger objects for which we should expect objective positions and so stochastic models, but in some situations their quantum description works surprisingly good, for example the nuclear shell model for baryons - MERW shows that such success of QM formalism does not disqualify stochastic description, in contrast it is also supported from this perspective, there is universality of quantum predictions.

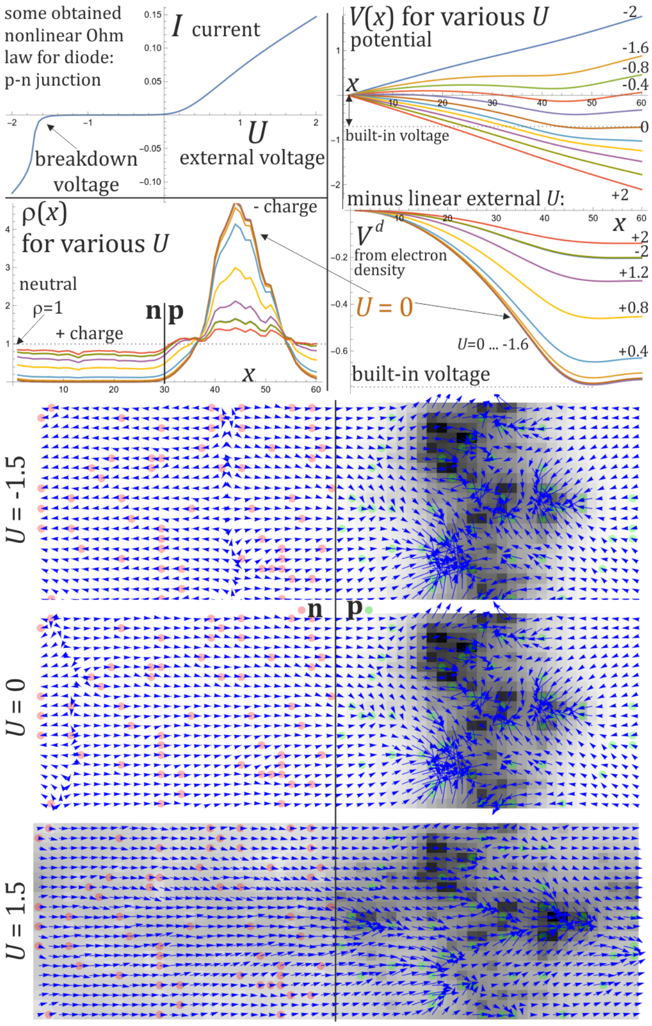

Figure 9. MERW model of conductance in p-n junction (diode) ([16], simple simulator: [14]) for 60×30 lattice with dopants: defects as nodes with lowered (p) or increased (n) potential. In GRW we would get nearly uniform electron density, leading to conductance for any external voltage as in (linear) Ohm law, incorrectly making it a conductor. In contrast, MERW has the same stationary probability distribution as QM - with Anderson localization preventing conductance (experimental using scanning tunneling microscope can be found in [17]), requiring some minimal breakdown voltage to start conductance in reverse bias, getting current-voltage dependence as in diode.

MERW allows to understand and repair this problem for example of seeing electron conductance as some statistical flow of charges - also where standard diffusion models had to give up, like defected lattice of semiconductors. The reason for this disagreement of standard stochastic models is that what looked as a natural choice for transition probabilities (like GRW) or stochastic propagator, often turns out only approximation of what is expected by statistical physics: entropy maximization. By repairing this approximation, MERW turns as close QM as we could expect from a diffusion model, like recreating equilibrium probability distribution exactly as the quantum ground state density. Also probability densities of excited states appear there as preferred, but can diffuse further (unless adding some constraint) like in Fig. 8.

Hence MERW can be seen as quantum correction to diffusion models. However, this is still only diffusion, not a complete QM - it ignores interference, which requires e.g. some internal clock (de Broglie's $E = mc^2 = \hbar\omega$, zittebewegung) of particle. However, beside providing clear intuitions for looking problematic properties of QM (like squares in formalism leading to violation of Bell's inequalities), like in Fig. 9 such quantum corrections to diffusion can be also useful especially as practical approximations of extremely demanding complete quantum modeling of conductance: in semiconductor, microscopic scale, or single molecule electronics.

In the third Section there will explained MERW's analogue of measurement, especially for violation of Bell's inequalities. Fourth Section focuses on quantum computation - we will argue that Shor's algorithm also exploits the fact that we live in a spacetime, suggest a general approach for designing quantum algorithms. Finally the last Section briefly discuses a possible ways to expand this simple but surprisingly successful effective model: just Boltzmann distribution among possible paths, into a more complete picture of physics, effectively described by quantum field theories - in perturbative approximation using ensemble of scenarios with varying number of particle: Feynman diagrams.

II. Maximal Entropy Random Walk as quantum corrections to diffusion

Let us start with the common problem of choosing a random walk (as Markov process) on a graph defining the space of interest - which later will be chosen for example as a lattice, where we can introduce inhomogeneity (defects) like in Fig. 7, or perform infinitesimal limit to get diffusion as continuous random walk. This section contains a condensed informal introduction, more complete description can be found as PhD Thesis of the author [13].

For simplicity assume here that the graph is indirected and defined by its (symmetric) adjacency matrix: $M_{ij} = M_{ji} = 1$ if there is edge between vertex i and j, 0 otherwise. From the perspective of physics, this adjacency matrix can be seen as simplified (zero potential) Bose-Hubbard Hamiltonian for a particle travelling between a set of sites connected as in this graph, jumping for i to j is first annihilation a_i then creation a_j^+:

$$\mathcal{H} = - \sum_{\text{edge } (ij)} a_j^+ a_i \equiv -M \tag{1}$$

A. Standard random walk (GRW) and its suboptimality

We would like to choose a stochastic matrix S for this graph: $S_{ij} = \Pr(\gamma_{t+1} = j | \gamma_t = i)$, which is nonzero only for graph edges, outgoing probabilities for each vertex have to sum to 1:

$$0 \le S_{ij} \le M_{ij} \le 1, \qquad \forall_i \sum_j S_{ij} = 1. \tag{2}$$

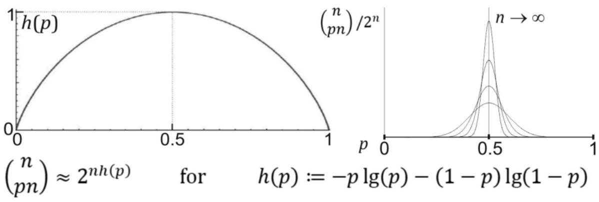

Figure 10. If among all 0/1 sequences of length n we focus on subset of sequences with p probability of '1', the size of this subset grows exponentially with $n \cdot h(p)$, where h is Shannon entropy. In the $n \to \infty$ thermodynamical limit, the $p = 1/2$ subset completely dominates all sequences - uniform probability distribution among all sequences has asymptotically $p = 1/2$, what is a special case of Asymptotic Equipartition Property. Generally, entropy being such exponent leads to the Principle of maximum entropy: the safest choice of statistical parameters is the one maximizing entropy, in a random configuration we asymptotically should almost certainly get these statistical parameters.

The standard way to choose random walk, referred as Generic Random Walk (GRW), is assigning equal probability to each outgoing edge, what for indirected graph leads to stationary probability distribution ($\sum_i \rho_i S_{ij} = \rho_j$) with probability of vertex being proportional to its degree d:

$$S^G_{ij} = \frac{M_{ij}}{d_i} \qquad \rho^G_i = \frac{d_i}{\sum_j d_j} \qquad \text{for} \quad d_i = \sum_j M_{ij} \quad (3)$$

Before commenting the above choice, let us remind the Principle of maximum entropy of Jaynes [18]. Imagine a length n sequence of '0' and '1', the number of such sequences is 2^n. Now focus on subspace of possibilities with density $p \in [0,1]$ of value '1': with approximately pn of '1'. Using Stirling formula $\left(k! \approx \sqrt{2\pi k}\,(k/e)^k\right)$ we can find asymptotic number of such combinations, plotted in Fig. 10:

$$\binom{n}{pn} \approx 2^{nh(p)} \quad \text{for} \quad h(p) = -p\lg(p) - (1-p)\lg(1-p)$$

being the Shannon entropy ($\lg \equiv \log_2$), which has single maximum $h(1/2) = 1$. Hence splitting the set of all 0/1 sequence into disjoint subsets with p density, asymptotically ($n \to \infty$) the $p = 1/2$ uniform probability case will combinatorially completely dominate all the other subsets.

Generally, like in the famous Boltzmann's formula, entropy is just (normalized) logarithm of the number of possibilities. Hence focusing on subset described by parameters (like density), maximizing entropy means focusing asymptotically on nearly all possibilities - contribution of subsets corresponding to suboptimal parameters asymptotically vanishes in exponential way with the size of the system. It can be summarized in the **Principle of maximum entropy: probability distribution which best represents the current knowledge is the one with largest entropy**. Without additional knowledge, entropy is maximized for uniform probability distribution on a given set. Assigning energy to objects/possibilities and fixing total energy, we get Boltzmann distribution instead. These two distributions are the base of statistical physics.

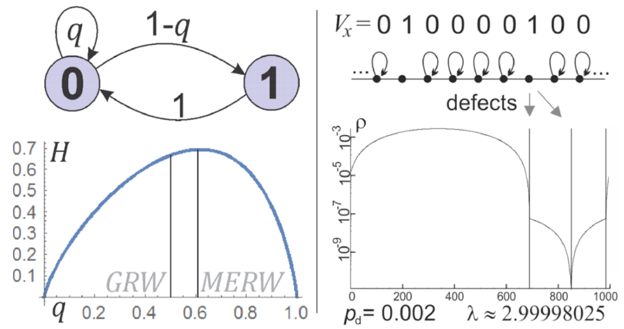

Figure 11. Left: Fibonacci coding as simple nontrivial example of suboptimality of GRW repaired by MERW. We have 0/1 sequences for which it is forbidden to use '11'. So after '0' we can use '0' or '1', but after '1' we have to use '0'. While seeing such sequence as random walk, the remaining question is to choose probability of '0' after '0' (parameter q). GRW suggests to choose $q = 1/2$, but its entropy rate $H(S)$ is suboptimal: subset of sequences described by such parameter is asymptotically completely dominated by other sequences without '11'. In contrast, MERW chooses golden ratio $q = (\sqrt{5} - 1)/2 \approx 0.618$, maximizing $H(S)$ and properly describing statistics in the set of all sequences without '11', or in such single "typical" infinite sequence. Right: MERW on 1D defected lattice with cyclic boundary conditions: all vertices are connected with 2 neighbors, all but the marked defects have additional self-loop (can remain in the vertex). As stationary probability distribution of GRW is proportional to degree of vertex, we get 2/3 times lower density in defects. In contrast to such very weak localization property, the drawn density of MERW has very strong localization in the largest defect-free region (so called Lifshitz sphere), exactly as for the quantum ground state for this lattice.

Returning to random walk on a graph, GRW clearly maximizes entropy for every vertex - is kind of local maximization. The question is if it maximizes average entropy per step: averaged over probability distribution of being in a given vertex. This measure is also called entropy rate:

$$H(S) = \sum_i \rho_i \sum_j S_{ij} \lg(1/S_{ij}) \qquad (4)$$

It turns out to be equal to normalized entropy in the space of sequences generated by such Markov process S:

$$H(S) = \lim_{n\to\infty} \frac{1}{n} \sum_{\gamma=\gamma_0\ldots\gamma_n} \Pr(\gamma)\lg(1/\Pr(\gamma)) \qquad (5)$$

$$\text{where} \qquad \Pr(\gamma_0 \ldots \gamma_n) = \rho_{\gamma_0} S_{\gamma_0\gamma_1} \ldots S_{\gamma_{n-1}\gamma_n}$$

is probability of obtaining sequence γ.

By Maximal Entropy Random Walk (MERW) we will refer to the choice of S matrix which maximizes $H(S)$ over all random walks on a given graph: fulfilling conditions (2). As this maximization involves corresponding stationary probability distribution: dominant eigenvector of S matrix to eigenvalue 1: $\rho S = \rho$, for maximization it is more convenient to use formula (5), which reaches maximum for uniform probability distribution among (infinite) paths generated by a given Markov process.

In many cases GRW already maximizes $H(S)$ making it equal with MERW, for example for regular graphs (all vertices have the same degree), like regular lattice and so standard diffusion in empty homogeneous space obtained as continuous limit of the lattice. The simplest

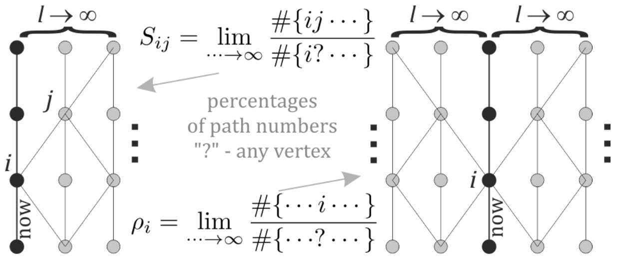

Figure 12. Deriving MERW formulas for simple four-vertex graph: term in its 4×4 adjacency matrix M_{ij} is 1 iff $|i-j| = 1$. Left: terms of stochastic matrix S_{ij} can be found by considering all length t paths starting in i and calculating what percentage of them makes the first step to j. For $t = 1$ we GRW this way, generally we can also consider GRW_t family for $t \in \mathbb{N}^+$, finally in $t \to \infty$ limit we obtain MERW stochastic matrix. Right: MERW stationary density can be found by considering all length $2t+1$ trajectories: ρ_i is percentage of those having x in the center in $t \to \infty$ limit is ρ_x. These formulas come from assumption of unique dominant eigenvalue of adjacency matrix: $M^t \to \lambda^t \psi \psi^T$.

example of nonoptimality of GRW is Fibonacci coding case, presented in Fig. 11. More physical examples are defected lattices, for example representing a semiconductor, or its continuous limit: diffusion in inhomogeneous space. In contrast to standard diffusion which leads to nearly uniform stationary probability distribution, MERW leads to very strong localization properties - exactly as the quantum ground state (for e.g. Bose-Hubbard or Schrödinger Hamiltonian), what we would from QM consideration and for example prevents semiconductor from being a conductor by prisoning electrons in entopic wells as mentioned in Fig. 7.

The GRW nearly uniform stationary probability distribution can be seen as maximizing entropy in spatial "static 3D" picture like in Fig. 1: for a fixed time cut of spacetime, leading e.g. to unform density in $[0, 1]$ potential well there. In contrast, MERW philosophy maximizes entropy in 4D spacetime picture: where particles become their trajectories, leading to $\rho(x) \propto \sin^2(\pi x)$ there, exactly as predicted also by QM.

B. MERW formulas and Born rule

While GRW assumes uniform probability distribution among outgoing edges: paths of length one, let as analogously define GRW_k to assume uniform probability distribution among length t paths ($\mathrm{GRW} \equiv \mathrm{GRW}_1$). The number of length t paths from vertex i to k for which the first step is to j is $M_{ij}(M^{t-1})_{jk}$, hence the stochastic matrix of GRW_t is $S_{ij} \propto M_{ij} \sum_k (M^{t-1})_{jk}$.

MERW assumes uniform probability distribution among possible infinite paths, what allows to see it as $t \to \infty$ limit of GRW_t like in Fig. 6. To derive formula for $t \to \infty$ limit, for simplicity let us assume that our indirected graph ($M = M^T$) is connected and aperiodic, where the Frobenius-Perron theorem says that M has non-degenerated (single) dominant eigenvalue $\lambda > 0$ and its corresponding eigenvector (left and right are equal for symmetric M) has real nonnegative coordinates:

$$\psi M = M\psi = \lambda\psi \quad \text{maximizing } \lambda \in \mathbb{R}^+, \quad \psi_i \in \mathbb{R}^+$$

Non-degenerated dominant eigenvalue makes that in the $t \to \infty$ limit we have $M^t \to \lambda^t \psi \psi^T$ (or $\lambda^t |\psi\rangle\langle\psi|$ in bra-ket formalism), getting MERW $S_{ij} \propto M_{ij}\psi_j$ as limit of GRW_t. For normalization, as $\sum_j M_{ij}\psi_j = \lambda\psi_j$, we get the final **formula for MERW** stochastic matrix:

$$S_{ij} = \frac{M_{ij}}{\lambda}\frac{\psi_j}{\psi_i} \qquad\qquad \rho_i \propto \psi_i^2 \tag{6}$$

where the above formula for stationary probability distribution, $\rho_i = \psi_i^2 / \sum_j \psi_j^2$ after probabilistic normalization, can be easily verified:

$$\sum_i \psi_i^2 \cdot \frac{M_{ij}}{\lambda}\frac{\psi_j}{\psi_i} = \sum_i \psi_i M_{ij} \frac{\psi_j}{\lambda} = \lambda\psi_j \frac{\psi_j}{\lambda} = \psi_j^2$$

The above derivation of MERW stochastic formula has used ensemble of infinite half-paths going forward in time, with (normalized) ψ describing probability distribution at the beginning of such ensemble. To analogously derive the formula for its stationary probability distribution, we can fix a position i and use ensemble of infinite half-paths toward both past in future. Using bra-ket formalism, both derivation can be informally written as:

$$|i\rangle\langle i||i\rangle\langle j|M^{t-1} \to \propto \lambda^{t-1}\psi_j \quad \text{imples} \quad S_{ij} \propto M_{ij}\psi_j$$
$$M^t|i\rangle\langle i|M^t \to \propto \lambda^{2t}\psi_i^2 \quad \text{imples} \quad \rho_i \propto \psi_i^2 \tag{7}$$

This way we get a natural intuition for $\rho_i \propto \psi_i^2$ containing **Born rule** as in Fig. 1, 3: the quantum amplitude ψ describes situation at the beginning of past or future half-spacetime (usually equal). If we want to measure a position or some observable in a given moment, we need to "draw" this random value from both past and future directions, getting final probability being product of both original (identical) probabilities, getting squares known from the quantum formalism, which as we know for example lead to violation of Bell inequalities wrongly expected by our natural "evolving 3D" intuition.

C. Boltzmann path distribution

Observe that taking a power of MERW stochastic matrix (6), or calculating probability distribution of a path γ, the intermediate ψ_j/ψ_i terms cancel, getting:

$$(S^t)_{ij} = \frac{(M^t)_{ij}}{\lambda^t}\frac{\psi_j}{\psi_i}$$

$$\Pr(\gamma_0 \ldots \gamma_t) = \frac{\psi_{\gamma_0}\ M_{\gamma_0\gamma_1} \ldots M_{\gamma_{t-1}\gamma_t}\ \psi_{\gamma_t}}{\lambda^t} \tag{8}$$

This way we get another "local" equivalent condition for MERW (written in Fig. 7): for all two vertices, each path of given length between them is equally probable.

In statistical physics we emphasize some possibilities by introducing energy, going from uniform to Boltzmann distribution $p_i \propto e^{-\beta E_i}$, which is obtained from the Jaynes

Principle of Maximum entropy by maximizing entropy under constraint of fixed average energy.

To take Boltzmann distribution to the MERW philosophy we need first to define energy of paths. A simple way is through choosing energy (potential) corresponding to each edge: V_{ij}, then define energy of a path as sum over all its edges:

$$\text{energy of path } (\gamma_0 \ldots \gamma_t) \quad \text{is} \quad V_{\gamma_0\gamma_i} + \ldots + V_{\gamma_{t-1}\gamma_t}$$

In equation (8) we can change from uniform to such Boltzmann distribution among paths by just using more general M matrix: still real nonnegative, but carrying weights corresponding to related potential:

$$M_{ij} = A_{ij}\, e^{-\beta V_{ij}} \qquad \text{for} \quad A_{ij} \in \{0,1\} \tag{9}$$

being the adjacency matrix.

To use vertex potential V_i instead, we can take e.g. $V_{ij} = \frac{1}{2}(V_i + V_j)$.

D. Lattice and continuous limit to Schrödinger equiation

As discussed, adjacency matrix can be seen as minus simplified (without potential) Bose-Hubbard Hamiltonian (1). Lattices are basic graphs used in physics, continuous situation can be realized as infinitesimal limit of a lattice. Let us start with 1D lattice from Fig. 11: MERW with potential barriers realized using self-loops (1 at diagonal of adjacency matrix, possibility to remain in the vertex), here removed in defects: $V_x = 0$ if vertex x contains self-loop, 1 otherwise.

While for GRW stationary probability distribution is proportional to degree of a vertex, getting very weak localization, for MERW we first need to find the dominant eigenvector of adjacency matrix ($\lambda\psi = M\psi$):

$$(\lambda\psi)_x = \psi_{x-1} + (1 - V_x)\psi_x + \psi_{x+1} \qquad / - 3\psi_x, \ \cdot - 1$$

$$E\psi_x = -(\psi_{x-1} - 2\psi_x + \psi_{x+1}) + V_x\psi_x \tag{10}$$

where maximization of λ has became minimization of energy $E := 3 - \lambda$ due to change of sign.

The $(\psi_{x-1} - 2\psi_x + \psi_{x+1})$ term is discrete Laplacian, making formula (10) discrete analogue of stationary Schrödinger equation. Hence MERW predicts going to exactly the same stationary probability distribution $\rho_x \propto \psi_x^2$ as predicted by quantum mechanics here - with very strong localization properties, for example preventing semiconductor from being a conductor.

To get the standard 1D continuous Schrödinger equation, let us take a regular lattice with ϵ time step and δ lattice constant. To consider real potential V, we can assume Boltzmann distribution among paths as in (9). The $\lambda_\epsilon\psi_x = (M_\epsilon\psi)_x$ eigenequation becomes for example:

$$\lambda_\epsilon\psi_x = e^{-\beta\epsilon V_{x-1}}\psi_{x-1} + e^{-\beta\epsilon V_x}\psi_x + e^{-\beta\epsilon V_{x+1}}\psi_{x+1}$$

As we are interested in $\epsilon, \delta \to 0$ limit, let us use approximations: $e^{-\beta\epsilon V} \approx 1 - \beta\epsilon V$ and $V_{x-1} \approx V_x \approx V_{x+1}$. After simple transformations and multiplying by -1 as previously (to change from maximization of λ to minimization of E), we get:

$$\frac{3 - \lambda_\epsilon}{3\beta\epsilon}\psi_x \approx -\frac{1}{3\beta}\frac{\psi_{x-1} + \psi_x + \psi_{x+1}}{\epsilon} + V_x\psi_x$$

Defining energy and choosing relation between time and space steps: with square required for getting from discrete to continuous random walk (as standard deviation grows with square root of time):

$$E_\epsilon = \frac{3 - \lambda_\epsilon}{3\beta\epsilon} \qquad\qquad \epsilon = \frac{\delta^2}{3\alpha}$$

for some α parameter, in the $\delta \to 0$ limit we get standard 1D stationary Schrödinger equation:

$$E\Psi = \left(-\frac{\alpha}{\beta}\Delta + V\right)\Psi = \left(-\frac{\hbar^2}{2m}\Delta + V\right)\Psi \tag{11}$$

assuming $\alpha/\beta = \hbar^2/2m$. Considering time dependant situation and comparing the continuity equations [13], suggests to choose: $\alpha = \hbar/2m, \ \beta = 1/\hbar$. For such generalized MERW there are also appearing other QM properties like Heisenberg principle.

The nature and values of such constants describing "fundamental noise" are crucial but not well understood. A natural source might be intrinsic periodic process of particles, so called Zitterbewegung or de Broglie's clock $E = \hbar\omega$, which has been directly observed for electrons [19], [20]. The MERW behavior also sees excited states, as presented in Fig. 8, suggesting that random thermodynamical distortion of a classical trajectory should make it smoothen toward probability cloud of close (overlapping) potentially excited quantum state.

What is surprising in the above lattice derivations is that Laplacian, which in standard QM is related with momentum, here appears only from spatial structure of the lattice as these corrected diffusion models do not consider velocity of particle. To add kinetic energy into considerations, we could perform MERW in phase space: (space, velocity), like in the Langevin equation, however, it becomes much more complicated.

The fact that Boltzmann distribution among paths leads to quantum thermodynamical predictions is not surprising as this MERW philosophy seems close to Feynman's Euclidean path integrals (EPI), however, there are some essential differences:

- Philosophy: EPI starts with assuming the axioms of QM, then performs "Wick rotation" to imaginary time - both having questionable clarity. In contrast, MERW just repairs diffusion: accordingly to the principle of maximum entropy, repairing known disagreements with reality of standard diffusion.
- Formula: standard EPI propagator lacks stochastic normalization, especially the crucial ψ_j/ψ_i term, which modifies the behavior with position in a very nonlocal way (dependent on the entire system).

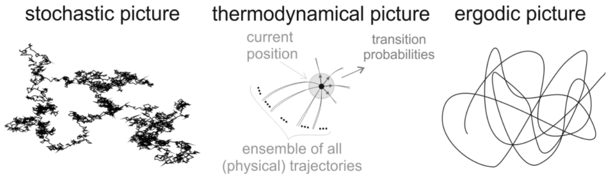

Figure 13. Three philosophies for probabilistic modelling of dynamics. Left: stochastic picture where evolution is imagined as succeeding random choices, with arbitrarily chosen probabilities e.g. locally maximizing entropy (GRW), for which it has nearly no localization property. Right: ergodic picture usually assuming a complex deterministic evolution (neglecting e.g. thermodynamical fluctuations), density is obtained by averaging over e.g. a single trajectory. Center: the discussed MERW-like philosophy based the principle of maximum entropy required by thermodynamical model, using Boltzmann distribution among possible trajectories, leading to quantum-like densities. Local transition probabilities (stochastic propagator) are calculated from ensemble of trajectories. The object does not directly use these probabilities (nonlocal: depending on the entire system), just performs some complex dynamics - only we use probabilities of this thermodynamical model to predict the safest evolution of our knowledge.

- Statistics: standard EPI assumes Boltzmann distribution among paths in a given time period, like in the GRW_t philosophy - conditioning the behavior on this arbitrarily chosen time period. In contrast, MERW uses ensemble of paths infinite in both time directions.
- Complexity: EPI starts with the continuous case, which path integration is mathematically problematic. In contrast, MERW philosophy starts with well understood discrete case.

Mathematically closer to MERW is Zambrini's Euclidean quantum mechanics [21], but like for Nelson's stochastic quantum mechanics [22], the motivation is fitting the expected behavior of QM, instead of MERW's just concluding from required fundamental mathematical principle: of entropy maximization.

III. Measurement and Bell inequalities

While our intuition of living in 3D space evolving in time requires satisfaction of Bell-like inequalities, they can be violated in real world or QM - let us understand it from perspective of living in 4D spacetime instead: where the basic objects are trajectories and we should consider their ensembles like in Euclidean path integrals or MERW.

For simplicity let us focus on Mermin's [2] Bell-like inequality for three binary variables $A, B, C \in \{0,1\}$:

$$\Pr(A=B)+\Pr(A=C)+\Pr(B=C)\geq 1 \qquad (12)$$

It can be intuitively explained that **tossing three coins, at least two of them give the same outcome**. More formally, choosing any probability distribution among their $2^3=8$ possibilities $\left(\sum_{ABC=0}^{1} p_{ABC}=1\right)$, each of three equalities correspond to 4 out of 8 possibilities - as shown in Fig. 2, summing $\Pr(A=B)+\Pr(A=C)+\Pr(B=C)$ we get $\left(\sum_{ABC=0}^{1} p_{ABC}\right)+2p_{000}+2p_{111}\geq 1$.

While it seems impossible for this inequality to be false, it is somehow violated in QM and real world. For this purpose it is crucial that we measure only 2 out of 3 variables, otherwise we would operate on $\{p_{ABC}\}$ probability distribution - which satisfies inequality (12).

So we measure 2 out of 3 variables - each outcome represents two possibilities for the unmeasured variable, e.g. $AB?$ outcome represents $\{AB0, AB1\}$ set. To violate the inequality we need something nonintuitive, like Born rules characteristic for QM and MERW:

- **Kolmogorov 3rd axiom**: probability of alternative of disjoint events is sum of individual probabilities: $p_{AB?}=p_{AB0}+p_{AB1}$, leading to the inequality (12).
- **Born rule**: probability of alternative of disjoint events is proportional to **square** of sum of their amplitudes: $p_{AB?}\propto(\psi_{AB0}+\psi_{AB1})^2$.

As in Fig. 2, this Born rule assumption allows to violate inequality (12): for example taking $\psi_{000}=\psi_{111}=0$, $\psi_{001}=\psi_{010}=\psi_{011}=\psi_{100}=\psi_{101}=\psi_{110}>0$ we get $\Pr(A=B)=\Pr(A=C)=\Pr(B=C)=1/5$ and so violation of the inequality to $3/5<1$.

Assuming as in Euclidean path integrals or MERW: uniform probability distribution among paths, we got the squares like in Born rules by multiplying amplitudes from both time directions. To formalize it we need to define MERW measurement, which in QM is destructive process: transforms usually continuous initial state into a discrete set of possibilities: eigenvectors of measurement operator.

To understand destructiveness of measurements, adapt it to a simple model like MERW, let us look at **Stern-Gerlach experiment** which is used as idealization of measurement: it applies strong magnetic field to transform initially continuous space of spin directions into one of two possibilities: parallel or anti-parallel alignment (only these two do not have Larmor precession hence minimize kinetic energy), which can be later separated using field gradient.

We can imagine that after aligning in strong magnetic field, spin can no longer flip during this flight in strong magnetic field - it leads to idealized condition which can be adapted for MERW and turns out sufficient: **during measurement its outcome cannot change**.

A. *Realization of Bell violation example*

Let us take "during measurement its outcome cannot change" rule to MERW like in Figure. 15. In all but time 0-1 step there are allowed steps accordingly to the assumed graph: blue edges in cube on the left (can be also used simpler, e.g. width 3 Ising lattice with interaction forbidding $\downarrow\downarrow\downarrow$ and $\uparrow\uparrow\uparrow$). In the remaining time 0-1 step we measure AB: first two out of three variables. This step is governed by the measurement rule: it cannot change the measured coordinates. However, it can change the third (unmeasured) coordinate - which is volatile in this measurement, otherwise inequality (12) would be satisfied.

For this spacetime diagram presented in the right part of Fig. 15, let us assume MERW rule that all possible paths (using blue edges) are equally probable - asking what percentage of them goes through the four boxes corresponding

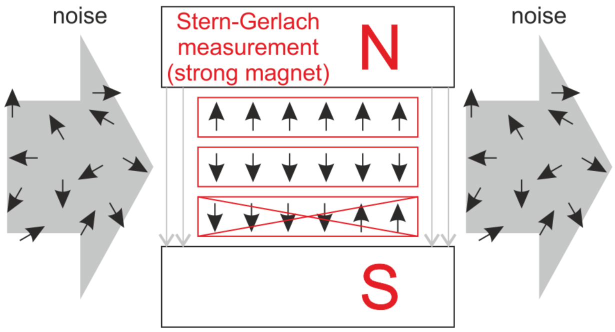

Figure 14. Measurement is a destructive process. We need to extract idealization of its influence on the system to take it to ensemble of paths considered here. Stern-Gerlach experiment is seen as idealization of measurement: with strong magnetic field enforcing initially random spin direction to choose between aligned or anti-aligned direction. This choice cannot be changed during the measurement: strong magnetic field does not allow to flip spin after it was already chosen. Such rule can be taken to ensemble of paths and turns out sufficient to get Born rules: we will **restrict ensemble to paths not changing outcome during measurement**.

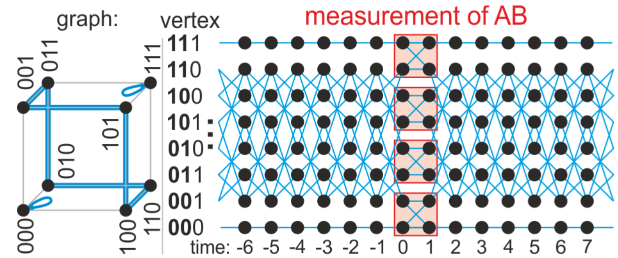

Figure 15. Left: graph we are considering, allowing for steps using the blue edges. Right: considered trajectories among which we assume uniform probability distribution (MERW) as analogue of Stern-Gerlach situation from Fig. 14, noise before and after measurement corresponds to random walk. Each column represents 8 possible vertices from assumed graph (cube), column's number represents time moment. All but the measurement moment we assume allowed paths accordingly to the marked (blue) edges of the cube (can be also used simpler, e.g. width 3 Ising lattice with interaction forbidding $\downarrow\downarrow\downarrow$ and $\uparrow\uparrow\uparrow$). This behavior is interrupted by measurement of AB coordinates in 0-1 time step (marked red) - single transition enforcing to remain outcome value: using only edges inside marked squares. Probability of given measurement outcome (square) is the number of paths going through its square, divided by the total number of paths.

to measurement outcomes, we will correspondingly get $1/10,\ 4/10,\ 4/10,\ 1/10$ probabilities, which violate the inequality.

Specifically, let us calculate the number of past paths from time $t=-l$ to $t=0$ in Fig. 15. For the 000 and 111 final vertices there is only a single such path. For the remaining 6 vertices there are 2 possibilities for each time step, hence there are 2^l paths ending in each of these vertices. Analogously for future paths: from time $t=1$ to $t=l+1$, their number is 1 for the 000 and 111, and 2^l for the remaining 6 vertices.

Now let us count bidirectional paths: from $t=-l$ to $t=l+1$ going through each of 4 measurement boxes. For the top and bottom box, corresponding to measuring 00 or 11 of AB coordinates, the number of paths is $(2^l+1)^2$. For the remaining two boxes, the number of paths is $(2^l+2^l)^2$. Asymptotically ($l\to\infty$) we get:

$$\Pr(A=0,B=0)=\lim_{l\to\infty}\frac{(2^l+1)^2}{2(2^l+1)^2+2(2^l+2^l)^2}=\frac{1}{10}$$

and analogously for the remaining three measurement outcomes, getting probability of going through each of the four boxes being correspondingly: $1/10,\ 4/10,\ 4/10,\ 1/10$.

Hence, in this scenario $\Pr(A=B)=2/10$ when measuring the first two coordinates. Analogously measuring the remaining two pairs (different grouping into pairs of vertices), we get violation of the inequality:

$$\Pr(A=B)+\Pr(A=C)+\Pr(B=C)=6/10<1$$

B. Born rules in general MERW measurement

To generalize this Born rule to adjacency matrix M, assume that measurement chooses between disjoint subsets of possibilities (4 red squares in Fig. 15), splitting the set of vertices X into disjoint subsets (components) distinguished by the measurement:

$$X=\bigsqcup_i X_i \qquad \text{and we need} \quad P(X_i)\propto\left(\sum_{j\in X_i}\psi_j\right)^2$$

Using the rule that **two neighboring steps are in the same component** during measurement as before:

Definition: *MERW measurement* in time $t=0$ for split $X=\bigsqcup_i X_i$ modifies the original uniform ensemble among paths $(\gamma_t)_{t\in\mathbb{Z}}$, to all paths satisfying: $\exists_i\ \gamma_0,\gamma_1\in X_i$ (beside usual $M_{\gamma_t\gamma_{t+1}}=1$ for $t\neq 0$).

The number of paths from $t=-l$ to $t=l+1$ going through X_i in time 0 and 1 is:

$$\sum_{ab}\sum_{jk\in X_i}(M^l)_{aj}(M^l)_{kb}\to\lambda^{2l}\sum_{jk\in X_i}\psi_j\psi_k=\lambda^{2l}\left(\sum_{j\in X_i}\psi_j\right)^2$$

where ψ, λ are dominant eigengenvector/eigenvalue ($M\psi=\lambda\psi$, assume it is unique), asymptotically ($l\to\infty$) getting general Born rule: $\Pr(X_i)\propto\left(\sum_{j\in X_i}\psi_j\right)^2$.

IV. Four dimensional understanding of quantum computation

While violation of Bell inequalities is rather only an interesting fact regarding consequences of QM, much deeper and applied exploitation of quantum strangeness is proposed for quantum computers, especially the Shor's algorithm [23] shifting the factorization problem from exponential to polynomial complexity. This believed exponential classical cost is crucial for safeness of widely used asymmetric cryptography like RSA, which could be endangered by quantum computers if overcoming technical (or deeper?) difficulties of their implementation.

Such possibility of shifting from classical exponential to quantum polynomial complexity suggests some computational superiority coming with the nonintuitive properties of quantum mechanics - understanding of which might help us designing new quantum algorithms, especially to understand the question of existence of polynomial quantum algorithms for NP-complete problems, for which positive answer could, among others, endanger all kind of used cryptography.

One characteristic property of quantum algorithmics is the requirement to use only reversible operations (gates) as quantum evolution is unitary. Observe that $(x, y, z) \to (x, y, z \text{ XOR } g(x, y))$ is its own reverse and allows to realize any boolean function like AND, OR and XOR if using prepared auxiliary bit $z = 0$. While we could classically reverse such gates and their sequences realizing some function, the requirement of a large number of prepared auxiliary bits prevents such use of reversible operations to actually reverse a difficult function, like the discrete logarithm - such reversing would require fixing on both ends of the process: of final values of the function and initial values of the auxiliary bits.

Hence the question is if we could influence some complex (reversible/time symmetric) computational process on its both ends (initialization and output/measurement) in order to obtain a somehow more superior computation capabilities, e.g. shifting some problem from exponential to polynomial complexity? We could fix a system of rubber bands on its both ends, like for anyons forming braids in Kitaev's hypothetical topological quantum computers [24]. However, it seems technically difficult to realize logic gates on such rubber bands in 3D. Even if we could realize basic gates for them, minimizing the tension of such rubber band system might be physically very difficult to stabilize for solving our computational problem, especially that for hard problems the number of local energy minima grows exponentially with problem size [25]. Another way to fix values in both boundaries is replacing Feynman with Boltzmann path ensemble: going to Ising model as in bottom of Fig. 16 - bringing question if such ensemble (also Feynman) of e.g. 2^{1000} paths is more than an idealization?

The situation seems more optimistic if this "rubber band setting" is in 4D spacetime: is a system of trajectories of some qubit carriers. One reason is that realizing logic gates is simpler in 4D than in 3D thanks to more freedom. More importantly, optimization of such system to solve our problem is no longer a continuous process, but from perspective of action optimization formulation of Lagrangian mechanics: we can imagine that nature has already solved the problems we are planning to ask.

Figure 16 contains such schematic picture for Shor's algorithm. Fixing situation in the past is easy: just prepare the qubits in some chosen states. Additionally, quantum measurement gives some possibility to affect the system also from the future direction: in case of Shor's algorithm it restricts the original ensemble to only possibilities having

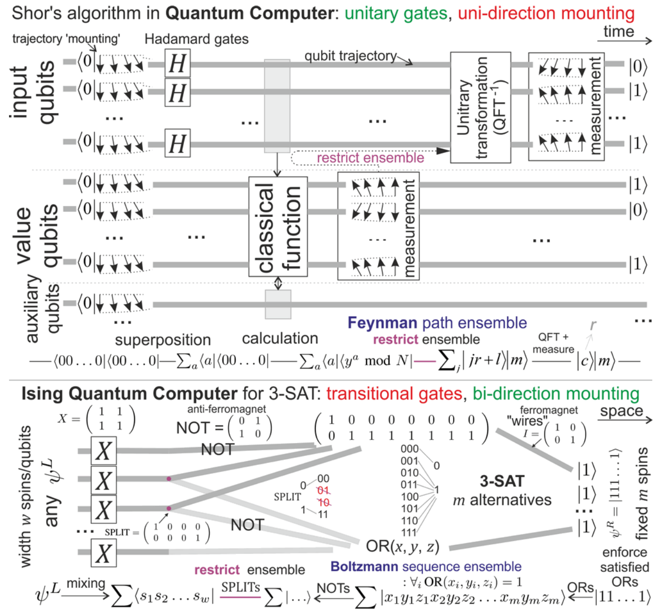

Figure 16. **Top**: Schematic diagram of quantum subroutine of Shor's algorithm [23] for finding prime factors of natural number N. For a random natural number $y < N$, it searches for period r of $f(a) = y^a \bmod N$. This period can be concluded from measurement of value c after Quantum Fourier Transform (QFT^{-1}) and with some large probability ($O(1)$) allows to find a nontrivial factor of N. The Hadamard gates produce state being superposition of all possible values of a. Then classical function $f(a)$ is applied, getting superposition of $|a\rangle|f(a)\rangle$. Due to necessary reversibility of applied operations, this calculation of $f(a)$ requires use of auxiliary qubits, initially prepared as $|0\rangle$. Now measuring the value of $f(a)$ returns some random value m, and restricts the original superposition to only a fulfilling $f(a) = m$. Mathematics ensures that $\{a : f(a) = m\}$ set has to be periodic here ($y^r \equiv 1 \bmod N$), this period r is concluded from the value of Fourier Transform (QFT^{-1}). Seeing the above process as a situation in 4D spacetime, qubits become trajectories, state preparation mounts their values (chosen) in the past (beginning), measurement mounts their values (random) in the future (end). Superiority of this quantum subroutine comes from future-past propagation of information (tension) by restricting the original ensemble in the first measurement. **Bottom**: NP complete problem (3SAT) approach with Ising model [26] - restring Boltzmann path ensemble to corresponding to solutions of given instance of problem. It brings fundamental question if Boltzmann (or Feynman in quantum computers) ensemble of e.g. 2^{1000} paths is more than an idealization?

the same (randomly chosen) value of calculated function. As emphasized in this diagram, the consequence of this restriction (tension) seems to propagate backward in time here, like in Wheeler's or delayed choice quantum erasure QM experiments, or in action optimizing formulation of Lagrangian mechanics.

Hence the suggested general approach to exploit the quantum superiority e.g. to search for polynomial algorithm for some NP-complete problem is:

1) Use Hadamard gates to get superposition of exponentially large set of possibilities, for example of all inputs to the problem among which we search for the satisfying one,
2) Perform some chosen classical function on these inputs, getting superposition like $\sum_a |a\rangle|f(a)\rangle$,
3) Measure value of this function, restricting the ensemble to $\sum_{a:f(a)=m} |a\rangle|m\rangle$,

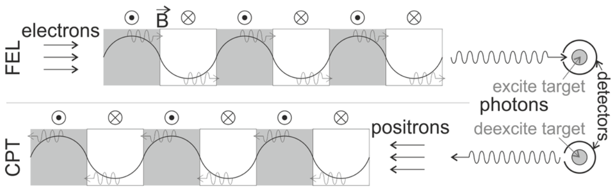

Figure 17. Schematic diagram of CPT transformation of free electron laser (CPTFEL). Standard laser stimulates emission of photons, causing their later absorbtion by the target, exciting it. As physics is believed to conserve CPT symmetry, in theory there should be possible to construct CPT analogue of a laser, what seems quite direct for the conceptually simplest: free electron laser (FEL). After CPT transformation of the above situation (bottom), initially excited target (e.g. a sodium lamp for narrow spectrum) is deexcited and photons escape through the hole in detector, travel to the CPTFEL, which absorbs them. Surprisingly, stimulated emission becomes stimulated absorbtion after CPT transformation: deexcitation of initially excited target seems stimulated by turning on the CPTFEL, earlier by the optical length - affecting energy balance between the target and detectors, which can be monitored.

4) Ask a question about this final ensemble, for example about its periodicity using QFT. Another basic question we can realize is if the size of resulting superposition is larger than one, what can be done by first producing multiple copies of bits of the inputs (e.g. using $(x, y) \rightarrow (x, y \text{ XOR } x)$ for auxiliary $y = |0\rangle$), then measuring them: values of their measurements will vary iff the superposition contains more than one possibility.

However, due to randomness of value of measurement restricting the ensemble, using this possibility of restriction to help with NP-complete problems does not seem to bring a direct solution. A hypothetical physical possibility which might help here is being able to affect the value of measurement to restrict the ensemble in a more controlled way. As physics is believed to fundamentally conserve the CPT symmetry, a CPT analogue of state preparation might help affecting the final qubit values. Figure 17 briefly presents such hypothetical CPT analogue of conceptually simple free electron laser.

V. Conclusions and further perspectives

While our natural intuition suggests us ”evolving 3D” picture of the world: which e.g. allows to conclude Bell inequalities for resulting correlations, quantum mechanics has Born rules: squares relating amplitudes and probabilities - leading to violation of such inequalities. Living in 4D spacetime, what is nonituitive but required by modern understanding of physics like special relativity or Lagrangian mechanics, requires to consider paths/scenarios (Feynman diagrams) as the basic objects - what leads to quantum-like behavior, starting with Anderson localization, Born rules and Bell violation. Specifically, its consequence is the present moment being in equilibrium between past and future, tension from both of these directions is described by the limit behavior for example of $M^t = (-\mathcal{H})^t$ as in (7), which is quantum amplitude of the ground state. Finally to randomly get some value of a measurement, we intuitively need to draw it from both time directions, getting probability being the square of amplitude. Consequences of living in 4D spacetime are seen in many quantum experiments like Wheeler's, delayed choice quantum erasure, as discussed also in Shor's algorithm.

Another conclusion from MERW is that the stochastic and quantum realms of physics, which have historically split their ways for example due to disagreement in predictions for semiconductor due to Anderson localization, can reunite if repairing the subtle approximation in entropy maximization, using ensembles of full paths. Especially interesting and important are situations in intersection of both worlds, like good understanding of electron flow in microscopic systems, what is crucial in modern electronics reaching level of single atoms and molecules, where standard Ohm's law is not fulfilled: a fixed potential difference for identical local situations can lead to different currents, behaving in a nonlocal way: depending on the entire system like in Fig. 9. MERW-based modelling can be used as a practical approximation of extremely costly complete quantum calculations e.g. of p-n junction - discussed in [16].

As discussed, Boltzmann distribution among paths is a simple model suggested by many perspectives, like the principle of maximum entropy of statistical physics, living in 4D spacetime, or agreement with expected quantum predictions and confirming experiments - making it a promising direction for understanding of physics governing our world. However, it obviously also contains essential simplifications, some of which are visualized in Fig. 18. One of them is varying number of particles in real physical scenarios, what is repaired in perturbative quantum field theories by considering ensemble among more sophisticated scenarios: Feynman diagrams. A real scenario represented by such simplified diagram contains

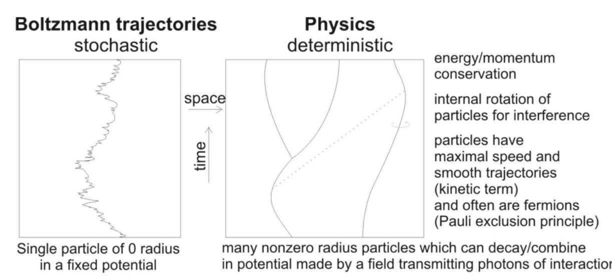

Figure 18. Boltzmann distribution among trajectories is a simple effective models recreating some basic properties of physics, quantum mechanics. One expansion toward fundamental physics is adding possibility to combine or decay particles, requiring to extend the ensemble of objects to consider from paths to more complex scenarios like Feynman diagrams. This way we would go to still universal picture of perturbative quantum field theories. Further steps toward fundamental physics will eventually require asking for field configurations behind such single diagrams, including gluing of EM fields in the center of charge, avoiding energy infinity obtained by assuming that charge is a perfect point. Such localized field configurations for particles are technically called solitons, for which quantization of charge has well known mathematical analogue: topological charge, presented in Fig. 19.

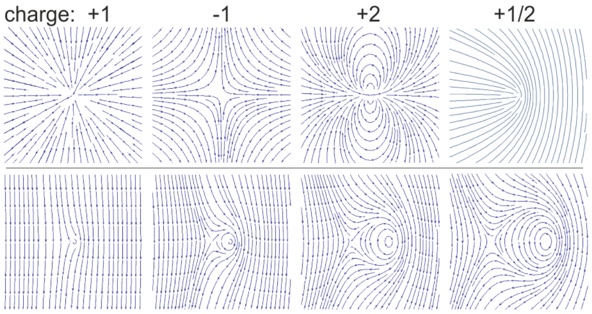

Figure 19. Top: examples of 2D topological charges of vector field - looking at a loop around such singularity, charge is the (integer) number of complete field rotations. If "removing arrows of vectors" we can also get charges being multiplicity of 1/2 this way (like spin). Well known example in nature of such field configuration stabilized by topology (topological soliton) is fluxon (Abrikosov vortex) in superconductor, carrying quant of magnetic field. Going to 3D analogs, e.g. in liquid crystals there are observed also Coulomb-like interactions [27] for such topological charges. Mathematically, defining EM field as curvature of vector field, Gauss-Bonnet theorem acts exactly as the Gauss law: integrating curvature over a closed surface, we get the total topological charge inside this surface - leading to electrodynamics with Maxwell's equations governing dynamics of such quantized charges [28].

additionally configuration of fields of interactions, for example electromagnetic, suggesting field theories for more fundamental description. Particles having a charge maintain nearly singular configuration of electric field - robust configurations of fields are technically called solitons, topology brings a natural mathematical tool to explain their charge quantization like in Fig. 19.

One of essential properties ignored by Boltzmann distribution among paths is the requirement for interference: some particle's internal periodic process, called de Broglie's clock ($E = mc^2 = \hbar\omega$) or Zitterbewegung, which has been directly observed for electron for example as increased absorbtion when synchronizing period of such clock with lattice constant of silicon crystal ([19], [20]). It causes coupled so called pilot waves in the surrounding field, confirmed as de Broglie-Bohm interpretation of QM for example by experiment measuring average trajectories of interfering photons in double-slit experiment [30]. While principle of complementarity forbids measuring both corpuscular and wave natures simultaneously, it does not exclude that particle has objectively both natures at a time, especially that no conditions for choosing one of them are specified (e.g. in which moment meeting electron and proton become hydrogen?), or mechanisms for such change of nature, and there are experiments successfully exploiting both natures at a time like Afshar's [31].

There are also lots of experimental hydrodynamical analogs of QM, especially started by Yves Couder, which show that such classical objects coupled with waves they create (droplet on a vibrating liquid surface) allow to recreate many quantum phenomena like: interference [32] in double-slit experiment (particle goes one trajectory, interacting with waves it created - going through all trajectories), tunneling [33] (depending on practically random

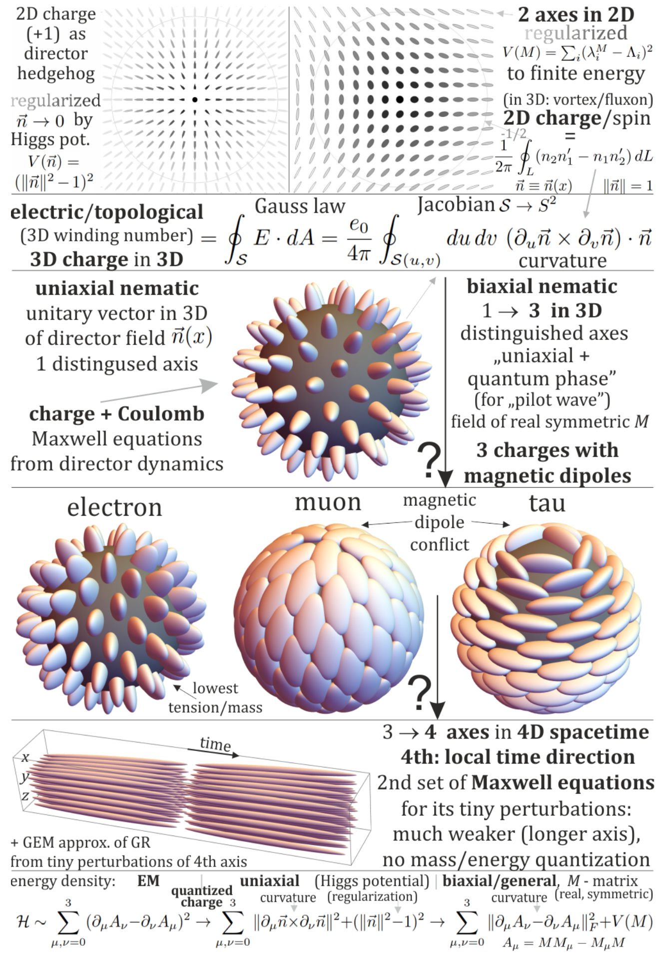

Figure 20. Topological soliton particle model approach [29] exploring resemblance of topological defects in liquid crystals, e.g. experimentally allowing to obtain Coulomb-like interaction [27]. Higgs-like potential enforces e.g. vector field to unitary vectors far from particles, also allowing to regularize singularity (charge) to finite energy. Interpreting field curvature as EM field, Gauss law returns topological charge - getting Maxwell equations with built-in charge quantization (Faber's apprach [28]). In biaxial nematic we have 3 distinguishable axes, leading to 3 types of hedgehog: the same charge, but different mass - resembling 3 leptons, also requiring magnetic dipole moment due to the hairy ball theorem. Extending further to field of 4 distinguishable axes in 4D, represented by real symmetric tensor field (M), leads to second set of Maxwell equations - as in gravitoelectromagnetism approximation of the general relativity theory, e.g. using the written Hamiltonian extending electromagnetism.

hidden parameters - highly complex state of the field), orbit quantization [34] including double quantization [35]: of both radius and angular momentum like in Bohr-Sommerfeld (particle has to "find a resonance" with the field - its internal phase has to return to initial state during full orbit), Zeeman splitting analogue [36] (using Coriolis as Zeeman force) and like in MERW: recreating quantum eigenstates with statistics of trajectory [37].

The universality of quantum formalism is also recreated in other hydrodynamical analogues, like Casimir effect: two plates in vibrating water tank also are experimentally shown to attract [38] as wave energy between them is lower due to restriction by the plates. There is also hydrodynamical analogue of Aharonov-Bohm effect suggested by Berry [39]: using vortex, vorticity and Casimir force analogues of solenoid, magnetic field and Lorentz force.

To summarize, while there is unsuccessful belief that we need to find a boundary between classical and quantum world, this boundary blurs e.g. with hydrodynamical analogues or MERW - it might turn out nonexistent: they can be just different perspectives/descriptions of the same system. For example coupled oscillators can be described by evolution of their positions ("classical"), or in the base of their normal modes - where this evolution becomes literally unitary ("quantum"). Lattices of such oscillators are used to model crystals: classically, or equivalently in Fourier basis: using phonons as normal modes - which are treated as (quasi)particles in Feynman diagrams. In continuous limit of such lattice we get field theories - which can be modelled with hydrodynamical analogues. Solitons are localized particle-like configurations of fields, effectively described by QFT. Using topological solitons we get charge/spin quantization, pair creation/annihilation, and electromagnetism-like interaction for them. Finally Couder's quantization suggests how to understand atoms: Schrödinger equations describes coupled wave, which to minimize energy needs to become a standing wave - this resonance between particle's clock and the field gives quantization conditions. Fig. 19, 20 presents some basics of such particle approach, discussed e.g. in [28], [29].

References

[1] J. S. Bell, "On the einstein podolsky rosen paradox," 1964.

[2] N. D. Mermin, "Bringing home the atomic world: Quantum mysteries for anybody," *American Journal of Physics*, vol. 49, no. 10, pp. 940–943, 1981.

[3] J. Schwinger, "The theory of quantized fields. i," *Physical Review*, vol. 82, no. 6, p. 914, 1951.

[4] R. P. Feynman, "Space-time approach to non-relativistic quantum mechanics," *Reviews of Modern Physics*, vol. 20, no. 2, p. 367, 1948.

[5] G. A. Gottwald and M. Oliver, "Boltzmann's dilemma: An introduction to statistical mechanics via the kac ring," *SIAM review*, vol. 51, no. 3, pp. 613–635, 2009.

[6] L. Boltzmann, "Weitere studien über das wärmegleichgewicht unter gasmolekülen." Sitzungsberichte Akademie der Wissenschaften, 1872.

[7] V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, and J.-F. Roch, "Experimental realization of wheeler's delayed-choice gedanken experiment," *Science*, vol. 315, no. 5814, pp. 966–968, 2007.

[8] S. Walborn, M. T. Cunha, S. Pádua, and C. Monken, "Double-slit quantum eraser," *Physical Review A*, vol. 65, no. 3, p. 033818, 2002.

[9] A. Danan, D. Farfurnik, S. Bar-Ad, and L. Vaidman, "Asking photons where they have been," *Physical review letters*, vol. 111, no. 24, p. 240402, 2013.

[10] J. G. Cramer, "The transactional interpretation of quantum mechanics," *Reviews of Modern Physics*, vol. 58, no. 3, p. 647, 1986.

[11] B. Reznik and Y. Aharonov, "Time-symmetric formulation of quantum mechanics," *Physical Review A*, vol. 52, no. 4, p. 2538, 1995.

[12] Z. Burda, J. Duda, J.-M. Luck, and B. Waclaw, "Localization of the maximal entropy random walk," *Physical review letters*, vol. 102, no. 16, p. 160602, 2009.

[13] J. Duda, "Extended maximal entropy random walk," Ph.D. dissertation, Jagiellonian University, 2012. [Online]. Available: http://www.fais.uj.edu.pl/documents/41628/d63bc0b7-cb71-4eba-8a5a-d974256fd065

[14] ——, "Electron conductance models using maximal entropy random walks." [Online]. Available: http://demonstrations.wolfram.com/author.html?author=Jarek+Duda/

[15] I. Mikhailovskij, E. Sadanov, T. Mazilova, V. Ksenofontov, and O. Velicodnaja, "Imaging the atomic orbitals of carbon atomic chains with field-emission electron microscopy," *Physical Review B*, vol. 80, no. 16, p. 165404, 2009.

[16] J. Duda, "Diffusion models for atomic scale electron currents in semiconductor, pn junction," *arXiv preprint arXiv:2112.12557*, 2021.

[17] A. Richardella, P. Roushan, S. Mack, B. Zhou, D. A. Huse, D. D. Awschalom, and A. Yazdani, "Visualizing critical correlations near the metal-insulator transition in Ga1-xMnxAs," *science*, vol. 327, no. 5966, pp. 665–669, 2010.

[18] E. T. Jaynes, "Information theory and statistical mechanics," *Physical review*, vol. 106, no. 4, p. 620, 1957.

[19] D. Hestenes, "Electron time, mass and zitter," *The Nature of Time Essay Contest. Foundational Questions Institute*, 2008.

[20] P. Catillon, N. Cue, M. Gaillard, R. Genre, M. Gouanère, R. Kirsch, J.-C. Poizat, J. Remillieux, L. Roussel, and M. Spighel, "A search for the de broglie particle internal clock by means of electron channeling," *Foundations of Physics*, vol. 38, no. 7, pp. 659–664, 2008.

[21] J. Zambrini, "Euclidean quantum mechanics," *Physical Review A*, vol. 35, no. 9, p. 3631, 1987.

[22] E. Nelson, "Derivation of the schrödinger equation from newtonian mechanics," *Physical review*, vol. 150, no. 4, p. 1079, 1966.

[23] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM review*, vol. 41, no. 2, pp. 303–332, 1999.

[24] M. Freedman, A. Kitaev, M. Larsen, and Z. Wang, "Topological quantum computation," *Bulletin of the American Mathematical Society*, vol. 40, no. 1, pp. 31–38, 2003.

[25] J. Duda, "P=? np as minimization of degree 4 polynomial, or grassmann number problem," *arXiv preprint arXiv:1703.04456*, 2017.

[26] ——, "Nearly accurate solutions for ising-like models using maximal entropy random walk," *arXiv preprint arXiv:1912.13300*, 2019.

[27] B.-K. Lee, S.-J. Kim, J.-H. Kim, and B. Lev, "Coulomb-like elastic interaction induced by symmetry breaking in nematic liquid crystal colloids," *Scientific reports*, vol. 7, no. 1, pp. 1–8, 2017.

[28] M. Faber, "Particles as stable topological solitons," in *Journal of Physics: Conference Series*, vol. 361, no. 1. IOP Publishing, 2012, p. 012022.

[29] J. Duda, "Framework for liquid crystal based particle models," *arXiv preprint arXiv:2108.07896*, 2021.

[30] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. K. Shalm, and A. M. Steinberg, "Observing the average trajectories of single photons in a two-slit interferometer," *Science*, vol. 332, no. 6034, pp. 1170–1173, 2011.

[31] S. S. Afshar, "Violation of the principle of complementarity, and its implications," in *The Nature of Light: What Is a Photon?*, vol. 5866. International Society for Optics and Photonics, 2005, pp. 229–245.

[32] Y. Couder and E. Fort, "Single-particle diffraction and interference at a macroscopic scale," *Physical review letters*, vol. 97, no. 15, p. 154101, 2006.

[33] A. Eddi, E. Fort, F. Moisy, and Y. Couder, "Unpredictable tunneling of a classical wave-particle association," *Physical review letters*, vol. 102, no. 24, p. 240401, 2009.

[34] E. Fort, A. Eddi, A. Boudaoud, J. Moukhtar, and Y. Couder, "Path-memory induced quantization of classical orbits," *Proceedings of the National Academy of Sciences*, vol. 107, no. 41, pp. 17 515–17 520, 2010.

[35] S. Perrard, M. Labousse, M. Miskin, E. Fort, and Y. Couder, "Self-organization into quantized eigenstates of a classical wave-driven particle," *Nature communications*, vol. 5, p. ncomms4219, 2014.

[36] A. Eddi, J. Moukhtar, S. Perrard, E. Fort, and Y. Couder, "Level splitting at macroscopic scale," *Physical review letters*, vol. 108, no. 26, p. 264503, 2012.

[37] D. M. Harris, J. Moukhtar, E. Fort, Y. Couder, and J. W. Bush, "Wavelike statistics from pilot-wave dynamics in a circular corral," *Physical Review E*, vol. 88, no. 1, p. 011001, 2013.

[38] B. C. Denardo, J. J. Puda, and A. Larraza, "A water wave analog of the casimir effect," *American Journal of Physics*, vol. 77, no. 12, pp. 1095–1101, 2009.

[39] M. Berry, R. Chambers, M. Large, C. Upstill, and J. Walmsley, "Wavefront dislocations in the aharonov-bohm effect and its water wave analogue," *European Journal of Physics*, vol. 1, no. 3, p. 154, 1980.