Approaches to Lorentz Invariant Bohmian Mechanics

Bachelorarbeit
von

Niklas Boers

an der Fakultät für Mathematik und Informatik
der
Ludwig-Maximilians-Universität München

München im April 2008

Betreuer: Prof. Dr. Detlef Dürr
Zweitgutachter: Prof. Dr. Dieter Lüst

Contents

1 Introduction

In this Bachelorthesis, I will present recent attempts to find a relativistic deterministic quantum theory. The most common argument against deterministic quantum theories, as for example Bohmian Mechanics, is that they cannot be generalized to the relativistic case. It is true that no complete relativistic Bohmian Theory that accounts for all relativistic quantum phenomena has been found so far. Nevertheless, there is no reason why this should be impossible.

I will restrict myself to Bohmian Mechanics, which is well-understood in the non-relativistic limit. An overview will be given in section 2. Bohmian Mechanics is a *deterministic quantum theory*. In addition, it is a *realistic quantum theory* in the sense that there really *is* some kind of reality on the microscopic level, in contrast to the orthodox (Copenhagen-) interpretation of quantum mechanics. In Bohmian Mechanics, the fundamental entities are particles. They could be fields or strings or something completely different, as long as there *is* some ontology on the microscopic level, the theory can be called *realistic* or *Bohmian*. The big advantage of this realism is that the theory does not depend on the role of an observer on a fundamental level - as it is the case in ordinary quantum mechanics. Hence, it might also be called *objective*: The exterior world, described by physical laws, does not depend on whether I look at it or not. Particles always have positions, whether they are "measured" (whatever that means) or not. The moon is there, whether I see it right now or not. Finally, the cat (and, indeed, the atom possibly causing its death) is whether dead or alive, I do not decide this unless I take a gun and kill it. But if I'm a nice guy and just take a look into the box to see if the cat is doing well, I won't collapse her wave function into "dead" or "alive". That falls under the regime of chance, with always one being true and the other one false, but never both at the same time. This is what logic tells us.

As it will turn out, any quantum theory which assumes some kind of reality on the microscopic level must be *nonlocal* if it wants to reproduce the results of measurements correctly. This was shown by John Bell, since his famous inequality is violated in nature. I will explain what is meant by this in section 3.

From the above, it seems natural to demand three properties of the theory to be found: realism or objectivity, non-locality and, of course, relativistic invariance.

In section 4, I will first discuss David Bohm's original work on relativistic invariance of what he calls his "ontological interpretation". As it will turn out, the case is rather trivial for one Dirac-particle. However, as soon as one considers a many-body-system, many problems arise that are still not solved. They will be addressed in the subsequent sections and attempts to solve them will be presented: The *Hypersurface Bohm-Dirac model*, that introduces a foliation of spacetime as an additional dynamical part of the theory to reconcile nonlocality and Lorentz invariance, and the *Opposite Arrows of Time Model*, that tries the same by introducing a second time-direction on the microscopic level.

2 Nonrelativistic Bohmian Mechanics

In nonrelativistic Bohmian Mechanics (BM) the equations of motion are the Schrödinger Equation

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi \tag{1}$$

and the guiding equation for the particle trajectories

$$\dot{\mathbf{Q}}_j = \frac{\hbar}{m_j}\Im\frac{\psi^*\nabla_j\psi}{\psi^*\psi}, \tag{2}$$

where $\mathbf{Q} = \mathbf{Q}(t)$ is the configuration of the particles at time t. Already at this point, one sees the explicit nonlocal character of BM: In addition to the fact that the wave function ψ is defined on configuration space, and hence influenced instantaneously by any changes of the configuration, the velocities of the Bohmian particles, calculated from the wave function, depend on the positions of all other particles in an instantaneous way.

The two fundamental dynamical variables of the theory are the complex- (or spinor-) valued wave function ψ on the space of possible configurations $\mathbf{q}$ and the actual configurations $\mathbf{Q} = \mathbf{Q}(t)$ of the particles at time t. The particle trajectories are given as integral curves to the velocity field

$$\mathbf{v}_j^{\psi} = \dot{\mathbf{Q}}_j. \tag{3}$$

Defining the current

$$\mathbf{j}^{\psi} = \Im\psi^*\psi, \tag{4}$$

the Schrödinger Equation implies a continuity equation (the quantum flux equation):

$$\frac{\partial|\psi|^2}{\partial t} + \nabla\cdot\mathbf{j}^{\psi}, \tag{5}$$

which makes it seem natural to interpret $|\psi|^2$ as a probability density ρ (it is obviously positive definite). ρ is not stationary, since $\nabla\cdot\mathbf{j}$ does not vanish, but *equivariant*: Due to (5), if at some time t, the configuration of the particles is $|\psi_t|^2$-distributed, then this holds for all other times as well.

$|\psi|^2$ is called the *Quantum Equilibrium Distribution* and plays the role of a *typicality measure*: In BM, a property $\mathcal{P}$ of a configuration $\mathbf{Q}(t)$ is defined to be *typical*, if for $0 < \epsilon \ll 1$ holds [8]

$$\int_{S_0(p)} |\psi_0(p)|^2 dq = 1 - \epsilon, \tag{6}$$

where $S_0(p)$ is the set of initial configurations $Q(0)$ leading to Q(t). Born's Statistical Law can thus be reformulated as: Typically, the position distribution of the particles is close to the $|\psi|^2$-distribution. According to Bell, all measurements can, in principle, be reduced to position measurements:

> ...in physics the only observations we must consider are position observations, if only positions of instrument pointers. It is a great merit of the de Broglie-Bohm picture to force us to consider this fact. ([1], p. 166)

Hence, all quantum mechanical predictions can be derived from this law. Within the framework of BM, it can actually be derived rather than just taken as an axiom (see [8]). Furthermore, the whole operator formalism of ordinary quantum mechanics arises quite natural from the two equations of motion (1) and (2), taking the fact as serious that in BM, every apparatus that is to "measure" something is also made of Bohmian particles ([7], chapter 12), which implies that, in order to describe an experiment, the combined system, including the observed system as well as the measuring system, has to be considered. The configuration of this bigger system then transforms over time, which is described by the guiding equation of the system[1]. The quantitative scale attached to the pointers of the measuring instrument will (thanks to quantum equilibrium) reflect the probability distribution obtained via the spectral measure of a self-adjoint operator related to the measurement [6]. Therefore, nonrelativistic BM is experimentally (but *not* conceptually) equivalent to orthodox quantum mechanics, with the advantage of being objective and deterministic.

[1]Note that, in particular, no *ad hoc* collapse of the wave function is needed to describe a measurement process.

3 Bell's Theorem and Nonlocality

Bohmian Mechanics is commonly said to be incompatible with special relativity because it is a nonlocal theory. John Bell has shown that nonlocality is not just a feature of Bohmian Mechanics (it is just very obvious there), but of all quantum theories or interpretations of the quantum mechanical formalism. For the derivation of his inequality, Bell only assumed locality, i.e. that no influences can travel faster than light, and that there is some kind of reality at the microscopic level. I shall work out in detail what is meant by this. Since the inequality and all of the equivalent formulations are violated by the predictions of quantum theory as well as by experiments, its assumptions have to be wrong. Bell himself concluded that our world is nonlocal and *not*, as in common misunderstanding, the impossibility of "hidden variables":

> It is the requirement of locality, or more precisely that the result of a measurement on one system be unaffected by operations on a distant system with which it has interacted in the past, that creates the essential difficulty. There have been attempts to show that even without such a separability or locality requirement no 'hidden variable' interpretation of quantum mechanics is possible. These attempts have been examined elsewhere and found wanting. Moreover, a hidden variable interpretation of elementary quantum theory[2] has been explicitly constructed.([1], page 14)

By this "explicitly constructed hidden variable interpretation", Bell means Bohmian Mechanics. From the foregoing it is clear that the question of compatibility of Special Relativity and nonlocality should arise in any attempt to construct a relativistic quantum theory. Bell's argument goes as follows ([1], pages 14 to 21):

At spatially separated points, two Stern-Gerlach-Magnets (SGMs) measure the components $\sigma_1 \cdot \mathbf{a}$ and $\sigma_2 \cdot \mathbf{b}$ of the spin of two particles ($\mathbf{a}$ and $\mathbf{b}$ are unit vectors). Since, by measuring the spin-component of particle 1, the spin-component of particle 2 is immediately exactly known for a setup where $\mathbf{a}$ is parallel to $\mathbf{b}$, although influences of the one measurement apparatus to the other are excluded by the locality-condition, Einstein, Podolsky and Rosen (EPR) concluded that quantum mechanics is not complete. Now comes Bell. The completion of quantum mechanics, which is necessary from the EPR-point of view, shall be achieved by the introduction of some parameters λ. For the result A of a measurement of $\sigma_1 \cdot \mathbf{a}$ and B of $\sigma_2 \cdot \mathbf{b}$ holds

$$A(\mathbf{a}, \lambda) = \pm 1 \quad \text{and} \quad B(\mathbf{b}, \lambda) = \pm 1. \tag{7}$$

Bell assumed that the two measurements cannot in any way influence each other. If $\rho(\lambda)$ denotes the probability density of the parameter λ, the expectation value of the product $\sigma_1 \cdot \mathbf{a} \sigma_2 \cdot \mathbf{b}$ is given by

$$P(a, b) = \int d\lambda \rho(\lambda) A(a, \lambda) B(b, \lambda) \tag{8}$$

[2]The remark here refers to the paper of Bohm from 1952

which should equal the quantum mechanical expectation

$$\langle \sigma_1 \cdot \mathbf{a} \sigma_2 \cdot \mathbf{b} \rangle = -\mathbf{a} \cdot \mathbf{b}. \tag{9}$$

From (7) follows that P in (8) cannot be less than -1, because ρ is normalized. But for $\mathbf{a} = \mathbf{b}$, $P = -1$ can only be reached if

$$A(\mathbf{a}, \lambda) = -B(\mathbf{a}, \lambda), \tag{10}$$

so that (8) can be rewritten as

$$P(\mathbf{a}, \mathbf{b}) = -\int d\lambda \rho(\lambda) A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda). \tag{11}$$

Introducing another unit vector $\mathbf{c}$, one finds that

$$\begin{aligned} P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c}) &= -\int d\lambda \rho(\lambda) \left[A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) - A(\mathbf{a}, \lambda) A(\mathbf{c}, \lambda)\right] & (12) \\ &= \int d\lambda \rho(\lambda) A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) \left[A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda) - 1\right]. & (13) \end{aligned}$$

for which (7) implies that

$$|P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})| \leq \int d\lambda \rho(\lambda) \left[1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)\right] \tag{14}$$

which is equivalent to

$$1 + P(\mathbf{b}, \mathbf{c}) \geq |P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})| \tag{15}$$

This is Bell's famous inequality. Because the right hand side is of order $|\mathbf{b} - \mathbf{c}|$ for small $|\mathbf{b} - \mathbf{c}|$, Bell concludes that $P(\mathbf{b}, \mathbf{c})$ cannot be stationary at the minimal (-1) at $\mathbf{b} = \mathbf{c}$. But the quantum mechanical expectation value *is* stationary at that point, hence the two values cannot be equal in general. Thus, quantum mechanics violates Bell's inequality. More generally, every formulation of quantum mechanics that reproduces (9) and is hence in accordance with all experiments that have been carried out so far is in contradiction to (15). Moreover, the quantum mechanical expectation value cannot even be approximated arbitrarily closely by (8). According to Bell, this shows that

> In a theory in which parameters are added to quantum mechanics to determine the result of individual measurement, without changing the statistical predictions, there must be a mechanism whereby the setting of one measuring device can influence the reading of another instrument, however remote. Moreover, the signal involved must propagate instantaneously, so that such a theory could not be Lorentz invariant.([1], page 20)

So far, one could, in order to restore locality, also deny the existence of any parameter λ. But the violation of Bell's inequality only excludes the possibility of *local* hidden variable theories and since *nonlocal* hidden variable theories which reproduce the whole operator formalism of quantum mechanics and the $|\psi|^2$-distribution (and hence violate Bell's inequality) do exist (e.g. Bohmian Mechanics), it seems rather senseless to deny their existence. We are left with the fact that, as long as we trust in quantum mechanical predictions (indeed verified by experiments, e.g. by Alain Aspect), nature *is* nonlocal. Any theory, especially relativistic extensions of quantum mechanics, has to cope with this property of our world. The question whether this nonlocality is somehow compatible with Lorentz invariance or not will be addressed in the next sections.

But first, I would like to present what Robert Griffiths answered to the question about how locality can be saved, taking into account the foregoing. At the ASC summer school 2007 in Munich, he presented the following picture:

Charles in Copenhagen prepares two letters, one filled with red, the other with blue paper, and sends them to Alice in Vienna and Bob in Munich, both knowing about the letters but *not* knowing which one they would receive. Then Alice immediately knows which color bob received after opening her own letter. But, there is no nonlocal influence between Alice and Bob; facts have already been made in Copenhagen before the letters were sent off. Physically, this is close to the idea that the information responsible for the correlations in a 2-photon-measurement is localized in the overlap of the two backward light cones. There would be no problem with local causality then. One can also reformulate this argument as follows: The EPR-setup shows that, whenever the two analyzers are parallel, the correlations force us to accept that if we measure on one side, we *immediately* know the outcome at the other side. This still holds if the experimental setup excludes signals that could travel from the one apparatus to the other with velocity less or equal than the speed of light. Hence, if we deny the possibility of immediate influences between different positions, the measurement outcomes must have been determined beforehand by variables quantum mechanics does not provide us with (e.g. "information" that lies in the backwards light cone of the spacetime point where the source emits the two photons). But this is exactly what Bell has shown to be impossible and he emphasizes this point in *"Bertlmann's socks and the nature of reality"* [1] using a very general EPR-Bohm-like setup: He neither assumes determinism nor any particular ontology, i.e. whether it is particles, or fields or whatever that constitutes the microcosmos. The setup consists of a "go"-signal as input, followed by two outputs in opposite directions, displaying "yes", when a signal (e.g. on the "upper" screen of a SGM) has reached them and displaying "no", when not. In addition, right before the signals reach the output, there will be two other signals a and b injected at the two ends such that $c\delta \ll L$, where c is the speed of light, L the length of the system and δ is the time between the injection of a resp. b and the signal arriving at the output. These signals a and b could e.g. rotate the SGMs. After a long run experiment, one will find out that the conditional probability distribution for results A and B at the two ends for given signals a and b does not separate:

$$P(A, B|a, b) \neq P_1(A|a)P_2(B|b). \tag{16}$$

If we wanted to explain certain quantum mechanically realizable correlations of A and B in a local way, i.e. excluding instantaneous effects from one end to the other, we would have to introduce additional variables which then allow for a separation of the probability distributions. This is the only way to *explain* the correlations if one does not allow for causal connections: look for "spurious correlations". But again, Bell shows in a very similar way to the derivation above that this leads to contradictions with quantum mechanics. Bell concludes:

> So the quantum correlations are locally inexplicable.([1], page 153)

4 Relativistic Bohmian Mechanics

As mentioned before, commonly the first argument against Bohmian Mechanics is that it cannot be brought to a relativistic extension. This is, according to the argument, probably due to its apparent nonlocal character. But, as we have seen, nonlocality is a feature of nature itself as long as we want physics to be about a *real* world where there really *is* something happening also on smaller scales.

4.1 The Bohm-Dirac Model for One Particle

For one particle, the situation is very simple because in this case, the configuration space on which the wave function is defined is the physical space itself. The following is referring to David Bohm [2].

From the one particle Dirac equation ($\hbar = c = 1$)

$$i\frac{\partial\psi}{\partial t} = (-i\alpha\cdot\nabla - e\alpha\cdot A(\mathbf{q},t) + e\phi(\mathbf{q},t) + \beta m)\,\psi \tag{17}$$

or

$$(i\gamma^\mu\partial_\mu - e\gamma^\mu A_\mu - m)\psi = 0$$

one can derive that

$$\partial_\mu\overline{\psi}\gamma^\mu\psi = 0, \tag{18}$$

where $\overline{\psi} \equiv \psi^\dagger\gamma^0$, which makes it seem natural to introduce the conserved current density

$$j^\mu = \overline{\psi}\gamma^\mu\psi, \tag{19}$$

whose time component $j^0 = \overline{\psi}\gamma\psi = \psi^\dagger\psi$ is positive definite. Identifying this current with a vector field of particle trajectories (it is always timelike and future oriented),

$$\frac{dX^\mu}{d\tau} = j^\mu \tag{20}$$

we can interpret j^0 as the density of crossings ρ of the particle trajectories through a $t = const.$ hypersurface[3].

The equations of motion (17) for the spinor and the spacetime configuration of the Bohm-Dirac-particle (20) give us, together with the 4-divergence free Dirac current (19) a well-defined Lorentz invariant deterministic quantum theory for one Dirac particle. Unfortunately, the generalization to a system of N particles is not at all straightforward and still not achieved.

[3] This crossing will for each particle path and each hypersurface only happen once, because j^μ is timelike and future oriented, i.e. the trajectories (integral curves of j^μ) can never go backwards in time. Note as well that the distribution of the points through which the trajectory crosses the hyperplane is the same as the probability distribution for the particles at that certain time.

4.2 The Bohm-Dirac Model for N Particles

Given the (not covariantly formulated) N particle Dirac equation ($\hbar = c = 1$)

$$i\frac{\partial\psi}{\partial t} = \sum_i \left(-i\alpha_i \cdot \nabla_i - e\alpha_i \cdot A(\mathbf{q}_i, t) + e\phi(\mathbf{q}_i, t) + \beta_i m\right)\psi \tag{21}$$

where $\psi = \psi(\mathbf{q}_1, ..., \mathbf{q}_N, t)$ takes values in $(\mathbb{C}^4)^N$ and α_k and β_k operate only on the four indices belonging to the k-th particle, we can again derive a continuity equation:

$$\frac{\partial\psi^\dagger\psi}{\partial t} + \sum \nabla_i \psi^\dagger \alpha_i \psi = 0 \tag{22}$$

where over all indices of all particles is summed. The guiding equation for the particle trajectories is now

$$\mathbf{v}_i = \dot{\mathbf{Q}}_i = \frac{\psi^\dagger \alpha_i \psi}{\psi^\dagger\psi}. \tag{23}$$

Introducing $\rho = \psi^\dagger\psi$, (22) can be rewritten as

$$\frac{\partial\rho}{\partial t} + \sum \nabla_i \cdot \rho\mathbf{v}_i = 0 \tag{24}$$

Hence, if the probability distribution is given by $\rho = \psi^\dagger\psi$ at one time in one distinguished Lorentz frame, this will be true for all times and we have an equivariant measure on the N-particle configuration space in that frame. Bohm pointed out that for the positions $\mathbf{q}_i$ and velocities $\mathbf{v}_i$ of the Bohm-Dirac particles as well as for the equivariant ensemble density ρ, a particular reference frame was chosen ([2], page 274), which of course conflicts with the idea of relativity. E.g., equation (21) considers all particles at the same time and, from a relativistic point of view, it is not clear what this means, since a Lorentz transformation would in general break this symmetry. Moreover, it is not clear how to extend the velocities to 4-vectors. I will later present a different approach, called *multitime formalism*, which seems to be more natural and direct for the question of a relativistic multi-particle quantum theory. But first, I will summarize Bohm's discussion of the Lorentz invariance of the velocities (23) and make some general remarks about quantum equilibrium in different Lorentz frames. For simplicity, Bohm considers a 2-particle-system ([2], page 280). If the wave function happens to have product structure, the velocities of the particles will be independent of one another and Lorentz invariance follows immediately. But if the state is given by a linear combination

$$\psi = A\phi_a(x_1, t)\phi_b(x_2, t) + B\phi_c(x_1, t)\phi_d(x_2, t) \tag{25}$$

the two particles will only move independently if the two summands do not overlap, because then the mixed terms in $\psi^\dagger\psi$ would vanish.

In the general case where they do overlap, there will be a nonlocal dependence between the two, which will also be displayed in the expressions for $\psi^\dagger\psi$ and $\mathbf{j}_1$ resp. $\mathbf{j}_2$. Therefore, the velocities $\mathbf{v}_i = \frac{j_i}{\rho}$ will be connected in a explicitly nonlocal way. Thus, the guiding equation is not Lorentz invariant. Nevertheless, one can say that the statistical predictions derived from quantum equilibrium are not anymore fixed to a certain frame although the problem of Lorentz invariance remains for the fundamental level of the theory. As far as the statistics of measurement results are concerned, i.e. on the observational level, the theory *is* indeed relativistically invariant in the sense that all predictions for measurement outcomes can be seen as predictions for pointer positions (or equivalent) and can thus be derived from the position-predictions of $\rho = \psi^\dagger\psi$ evaluated at a common time in a preferred frame, which cannot anymore be identified in the statistics[4]. Bohm argues that

> The theory of measurement processes that we have given, along with the demonstration that the probability distribution $P = \rho$ is covariant, then ensures that the statistical results of these measurement processes will have a Lorentz invariant content.([2], page 285)

The fundamental problem remains: How can a covariant theory be obtained that treats particles which are nonlocally connected? Bringing together covariance and nonlocality easily leads to causal loops which should really not arise in a reasonable physical theory. So, instantaneous transmission of *signals* must be excluded.

4.3 The $|\Psi|^2$-Distribution in Different Lorentz Frames

In this section, I will discuss the work of Berndl, Dürr, Goldstein and Zanghì in [3] on the question of quantum equilibrium in different reference frames. The quantum equilibrium hypothesis states that *at a given time t, the spatial distribution of particles is close to the $|\psi|^2$-distribution* $(\star)$ [7]. This is the observer-independent version of Born's Law. In nonrelativistic Bohmian Mechanics, it can be derived from the fundamental equations of motion via the quantum flux equation, which provides us with an equivariant measure of typicality[5]: Typically, the particles in the universe are $|\Psi|^2$-distributed, where Ψ is the wave function of the universe. This measure of typicality then allows for a proof of the $|\psi|^2$-distribution in subsystems (ψ is the "effective wave function"), i.e. Born's Law [7]. If this is possible, the model shall be called *statistically transparent.* It is this demand of statistical transparancy rather than its nonlocal structure on its own that makes it so difficult to formulate a relativistic Bohmian quantum theory.

"Relativistically speaking", $(\star)$ translates to ([3], page 6): For all $t = const.$-hypersurfaces in Minkowski space, the distribution of crossings by the spacetime-trajectories of N particles ("N-paths") is given by $\rho^\Sigma = |\psi^\Sigma|^2$. The distribution of crossings

$$\rho^\Sigma : \Sigma^N \to \mathbb{R} \tag{26}$$

is determined by a probability measure P on the set of N-paths. As we will see later, the $t = const.$-hypersurfaces may be replaced by any spacelike hypersurfaces. The problem is this: These *simultaneity surfaces* are always defined with respect to a particular Lorentz

[4]See also [3], page 4.

[5]See for example [7] for a detailed discussion.

frame Λ and the wave function determining the distribution of particles is also given with respect to this frame: $\psi^{\Sigma} = \psi^{\Lambda}$. Berndl, Dürr, Goldstein and Zanghì state that

> There does not in general exist a probability measure P on N-paths for which the distribution of crossings ρ^{Σ} agrees with the corresponding quantum mechanical distribution on all spacelike hyperplanes Σ. ([3], page 6)

This very general claim can be shown to be true using the setup of Hardy's nonlocality-theorem ([3], [10]). Consider the experimental setup in Fig. 1 with three different Lorentz frames.

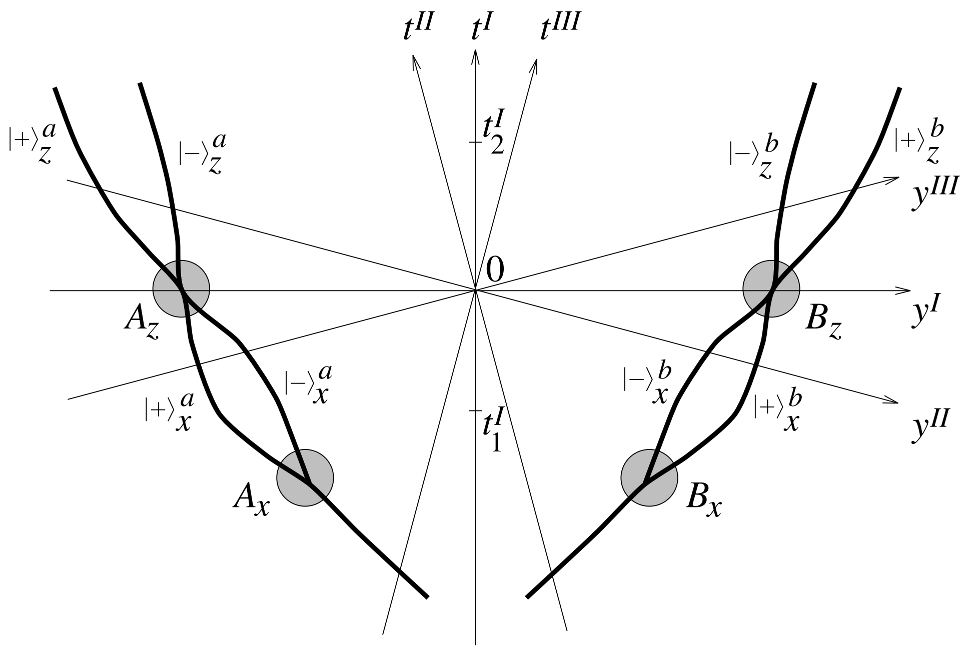

Figure 1: Spacetime diagram of the evolution of the wave function in Hardy's experiment, taken from [3]

Assume that the two particles are prepared in a state (in frame I)

$$\psi_{Hardy} = \frac{1}{\sqrt{3}} \left(|+\rangle_z^a |-\rangle_z^b - \sqrt{2} |-\rangle_x^a |+\rangle_z^b \right) \tag{27}$$

Then, for the different hyperplanes of the different frames, the wave function takes the form

$$\psi_1^I = \frac{1}{\sqrt{12}} \left(|+\rangle_x^a |+\rangle_x^b - |+\rangle_x^a |-\rangle_x^b - |-\rangle_x^a |+\rangle_x^b - 3|-\rangle_x^a |-\rangle_x^b \right) \tag{28}$$

$$\psi_0^{II} = \frac{1}{\sqrt{6}} \left(|-\rangle_z^a |+\rangle_x^b + |-\rangle_z^a |-\rangle_x^b - 2|+\rangle_z^a |-\rangle_x^b \right) \tag{29}$$

$$\psi_0^{III} = \frac{1}{\sqrt{6}} \left(|+\rangle_x^a |-\rangle_z^b + |-\rangle_x^a |-\rangle_z^b - 2|-\rangle_x^a |+\rangle_z^b \right) \tag{30}$$

$$\psi_2^I = \frac{1}{\sqrt{3}} \left(|+\rangle_z^a |-\rangle_z^b - |+\rangle_z^a |+\rangle_z^b + |-\rangle_z^a |+\rangle_z^b \right) \tag{31}$$

Now assume that quantum equilibrium holds on all spacelike hyperplanes, i.e. $\rho^\Sigma = |\psi^\Sigma|^2$ for all frames. This together with (28) implies that on $\Sigma = \Sigma^I(t_1^I)$

$$\int_{supp|+\rangle_x^a \times supp|+\rangle_x^b} \rho_1^I(q_a, q_b) dq_a dq_b = \int_{supp|+\rangle_x^a \times supp|+\rangle_x^b} \psi_1^{II}(q_a, q_b) dq_a dq_b = \frac{1}{12} \tag{32}$$

and on $\Sigma^{II}(0)$ and $\Sigma^{III}(0)$, (29) resp. (30) imply that

$$\rho_0^{II}(q_a, q_b) = 0 \quad \text{for} \quad (q_a, q_b) \in supp|+\rangle_z^a \times supp|+\rangle_x^b \tag{33}$$

and

$$\rho_0^{III}(q_a, q_b) = 0 \quad \text{for} \quad (q_a, q_b) \in supp|+\rangle_x^a \times supp|+\rangle_z^b \tag{34}$$

Finally, from (31) follows that on $\Sigma^I(t_2^I)$

$$\rho_2^I(q_a, q_b) = 0 \quad \text{for} \quad (q_a, q_b) \in supp|-\rangle_z^a \times supp|-\rangle_z^b \tag{35}$$

Now, look only at 2-paths which cross $supp|+\rangle_x^a \times supp|+\rangle_x^b$. According to (32), they have probability 1/12. (33) implies that particle a will be in $supp|-\rangle_z^a$ at t_2^I and, from (34), particle b will be in $supp|-\rangle_z^b$ at t_2^I. Hence

$$\int_{supp|-\rangle_z^a \times supp|-\rangle_z^b} \rho_2^I(q_a, q_b) dq_a dq_b \geq \frac{1}{12} \tag{36}$$

in contradiction with (35)([3], page 9).

Thus, the distribution of the actual particle positions in different Lorentz frames (on different spacelike hyperplanes) cannot be represented by the joint distributions given by quantum mechanics for a position measurement. This proofs the claim.

Nonetheless, this does not lead to *observable* contradictions : Any measurement process (e.g. by insertion of photo plates) will influence the future paths of the particle in such a way that the actual measurement can still be predicted correctly. Only the "unmeasured" distributions might conflict with quantum mechanical predictions.

4.4 The Multitime Formalism

According to Bohm ([2], page 278), with the usual N-particle Dirac equation

$$i\frac{\partial\psi}{\partial t} = \sum H_n\psi \tag{37}$$

where

$$H_n = \alpha\cdot\left[\frac{\nabla_n}{i} - e\mathbf{A}(\mathbf{q}_n,t)\right] + m\beta_n + \phi_n(q_n,t) \tag{38}$$

(the scalar potential Φ is actually missing in [2]), we can only derive probabilities for results of measurements carried out at a common time t. If, for instance, one would like to measure the positions x_n each at a different time τ_n (which is surely closer to the spirit of relativity), a *multitime formalism* has to be used: Consider the N wave equations

$$i\frac{\partial\chi}{\partial\tau_n} = H_n\chi \tag{39}$$

which are all integrable if all the Hamiltonians H_n commute. This is guaranteed if the spacetime points $(x_1\tau_1,...,x_N\tau_N)$ are on a spacelike plane, i.e. outside each other's light cones.

4.4.1 Multitime Translation Invariance

Using this formalism, Berndl, Dürr, Goldstein and Zanghì derived a multitime translation invariant Bohmian theory [3]. Multitime translation invariance can be seen as a certain limit of Lorentz invariance, having in common the basic idea of relativity that there is no absolute simultaneity. Berndl, Dürr, Goldstein and Zanghì consider a two-particle system described by two observers O and O'. Two widely separated events (t_α,x_α) and (t_β,x_β), $x_\beta\gg 1$, which are simultaneous for O $(t_\alpha=t_\beta)$ have the primed coordinates

$$\begin{aligned} t'_\alpha = t_\alpha = 0, &\qquad x'_\alpha = x_\alpha = 0 \\ t'_\beta = \gamma\,(t_\beta - vx_\beta) \approx -\vartheta, &\qquad x'_\beta = \gamma\,(x_\beta - vt_\beta) \approx x_\beta \end{aligned} \tag{40}$$

For v small, $\gamma = \frac{1}{\sqrt{1-v^2}} \approx 1$. The authors assume that x_β is that large that vx_β is of order unity. Obviously, $t'_\beta \neq t'_\alpha$, hence there is for this system no such thing as absolute simultaneity due to the relative time translation. This is a special case of a Lorentz transformation ($x_\beta\to\infty, v\to 0$ such that $vx_\beta\neq 0$), as mentioned above.

On configuration spacetime, Berndl, Dürr, Goldstein and Zanghì formulate the multitime translation introducing distinct coordinate systems for the two subsystems a and b:

$$\begin{aligned} L_\tau : \mathbb{R}\times\mathbb{R}^{3N_a}\times\mathbb{R}\times\mathbb{R}^{3N_b} &\longrightarrow \mathbb{R}\times\mathbb{R}^{3N_a}\times\mathbb{R}\times\mathbb{R}^{3N_b}, \tau=(\tau_a,\tau_b)\in\mathbb{R}^2 \\ z := (t_a,q_a,t_b,q_b) &\longmapsto (t_a-\tau_a,q_a,t_b-\tau_b,q_b) = z' = L_\tau z \end{aligned} \tag{41}$$

For independent subsystems (i.e. on a spacelike hyperplane), their corresponding Hamiltonians will commute and one can write down the Schrödinger evolution in a multitime translation invariant way:

The two-time wave function $\psi(t_a, t_b) \in L^2(\mathbb{R}^{3N_a}) \otimes L^2(\mathbb{R}^{3N_b})$,

$$\psi(t_a, t_b) = e^{-iH_a t_a} e^{-iH_b t_b} \psi_0 = U^a_{t_a} U^b_{t_b} \psi_0 \tag{42}$$

transforms according to

$$\psi(z) = \psi \circ L_\tau^{-1}(L_\tau z) =: \psi'(z') \tag{43}$$

and

$$\psi' = e^{-iH_a \tau_a} e^{-iH_b \tau_b} \psi = U^a_{\tau_a} U^b_{\tau_b} \psi = U_\tau \psi. \tag{44}$$

The two-time wave function satisfies the two Schrödinger equations

$$i\frac{\partial \psi}{\partial t_n} = H_n \psi \qquad \text{for} \quad n = a, b \tag{45}$$

From this setting, the authors derive a multitime translation invariant measurement formalism using the Heisenberg picture ([3], page 13): Given that the two subsystems are independent (no interaction potential), all observables and unitary evolutions mutually commute: For all j, j', t_a,t_b:

$$[M^a_j, M^b_{j'}] = [M^a_j, U^b_{t_b}] = [M^b_j, U^a_{t_a}] = [U^a_{t_a}, U^b_{t_b}] \tag{46}$$

Using the projection operators $\pi^a_{j,\alpha}$ and $\pi^b_{j,\beta}$ onto the eigenspaces of M^a_j and M^b_j corresponding to the eigenvalues α and β, Berndl, Dürr, Goldstein and Zanghì construct the Heisenberg operators ([3], page 13)

$$\pi^a_{j,\alpha}(t_a) := U^a_{-t_a} \pi^a_{j,\alpha} U^a_{t_a} \tag{47}$$

and

$$\pi^b_{j,\beta}(t_b) := U^b_{-t_b} \pi^b_{j,\beta} U^b_{t_b}. \tag{48}$$

Then, the probability for a certain measurement outcome is[6]

$$\begin{aligned} &P(M^a_1 = \alpha_1, ..., M^a_k = \alpha_k, M^b_1 = \beta_1, ..., M^b_l = \beta_l) \\ &= \|\pi^b_{l,\beta_l}(t^b_l) \cdots \pi^b_{1,\beta_1}(t^b_1), \pi^a_{k,\alpha_k}(t^a_k) \cdots \pi^a_{1,\alpha_1}(t^a_1)\psi\|^2, \end{aligned} \tag{50}$$

Arguing that the Heisenberg operators transform under a multitime translation as

$$\begin{aligned} \pi^a_{j,\alpha}(t'_a) &= U^a_{-t'_a} \pi^a_{j,\alpha} U^a_{t'_a} \\ &= U^a_{\tau_a - t_a} \pi^a_{j,\alpha} U^a_{-(\tau_a - t_a)} \\ &= U^a_{\tau_a} \pi^a_{j,\alpha}(t_a) U^a_{-\tau_a} \\ &= U_\tau \pi^a_{j,\alpha}(t_a) U_{-\tau} \end{aligned} \tag{51}$$

[6]This is (morally) clear from

$$\begin{aligned} |\langle \mu|\psi\rangle|^2 &= \langle \psi|\mu\rangle\langle \mu|\psi\rangle \\ &= \langle \psi|\mu\rangle\langle \mu|\mu\rangle\langle \mu|\psi\rangle \\ \text{"="} &\ \ \langle \pi\psi|\pi\psi\rangle = \|\pi\psi\|^2 \end{aligned} \tag{49}$$

and the state ψ itself according to (44), the authors conclude that (50) is multitime translation invariant, in particular independent of a certain frame of reference:

> Thus, the quantum mechanical measurement formalism for a system, which consists of independent widely separated subsystems, is multitime translation invariant.([3], page 13)

Experimental setups of the type suggested by EPR can be described by this formalism: The two subsystems (e.g. the two spin-$\frac{1}{2}$-particles) do not interact via an interaction potential, although they are entangled and thus nonlocally connected. The commutativity assumption (46) guarantees that there can still be no transmission of signals. This is what is meant by "not interacting": It refers to spatially separated systems so that causal influences in the sense of special relativity can be excluded.

4.4.2 A Multitime Translation Invariant Bohmian Theory

As already mentioned above, the formulation of a multitime translation invariant Bohmian theory might be seen as a first step towards a Lorentz invariant Bohmian theory, since it inherits and actually focuses on a basic feature of special relativity, the non-existence of absolute simultaneity. Berndl, Dürr, Goldstein and Zanghì consider a system of n ($n = 2$ for simplicity) widely separated subsystems with total particle number $N = N_a + N_b$, described by an n-time wave function ψ satisfying n Schrödinger equations ([3], page 14)

$$i\frac{\partial\psi}{\partial t_i} = H_i\psi \tag{52}$$

Let $Q_a(t)$ and $Q_b(t)$ denote the trajectories of the particle configurations in the two subsystems. To formulate the theory, a synchronization has to be introduced: an equivalence class of maps

$$(T_a, T_b) : \mathbb{R} \longrightarrow \mathbb{R}^2 \tag{53}$$

$$s \longmapsto (T_a(s), T_b(s)) \tag{54}$$

Together with $Q_a(t)$ and $Q_b(t)$, this yields an synchronized N-path in configuration spacetime:

$$(T_a(s), Q_a(s), T_b(s), Q_b(s)) =: Z(s) \tag{55}$$

where $Q_i(s) = Q_i(T(s))$. The guiding equation for the particle trajectories in spacetime will be

$$\frac{dT_a}{ds} = 1$$
$$\frac{dT_b}{ds} = 1 \tag{56}$$

together with

$$\frac{dQ_a}{ds} = v_a^{\psi}(Z) = \Im\frac{\nabla_{q_a}\psi}{\psi}$$
$$\frac{dQ_b}{ds} = v_b^{\psi}(Z) = \Im\frac{\nabla_{q_b}\psi}{\psi} \tag{57}$$

Hence,

$$\frac{dZ(s)}{ds} = \begin{pmatrix} 1 \\ v_a^\psi(Z) \\ 1 \\ v_b^\psi(Z) \end{pmatrix} \tag{58}$$

If a tuple (ψ, Z) solves (52) and (58), then so does $(\psi', Z') = (\psi \circ L_\tau^{-1}, L_\tau \circ Z)$, which makes the theory invariant under time translations L_τ. In addition, since the parameter s that labels what is simultaneous and hence "chooses" a certain frame of reference[7], is arbitrary, the theory does not have a preferred frame: Equation (58) is equivalent to all other equations of the form

$$\frac{dZ(s)}{ds} = A(Z) \begin{pmatrix} 1 \\ v_a^\psi(Z) \\ 1 \\ v_b^\psi(Z) \end{pmatrix}$$

In the same way as for the Bohmian Theories described so far, Berndl, Dürr, Goldstein an Zanghì look for a distinguished measure ([3], page 15), which is provided by the appropriate extension of the subsystem-continuity equations

$$div_{Z_i} J_i^\psi = 0, \tag{59}$$

which follows from (52) with

$$J_k^\psi = \begin{pmatrix} |\psi|^2 \\ j_k^\psi \end{pmatrix} \tag{60}$$

and

$$j_k^\psi = |\psi|^2 v_k^\psi = \Im(\psi^* \nabla_{q_k} \psi) \tag{61}$$

For Z, this yields

$$\frac{\partial |\psi|^2}{\partial s} + div_{Z_a} J_a^\psi + div_{Z_b} J_b^\psi = 0 \tag{62}$$

since $|\psi|^2$ is stationary with respect to the synchronization parameter s. Thus, the dynamics itself points on $|\psi|^2$ as a distinguished measure on $\mathbb{R}^{2+3N}$, the space of initial values for (58).

Unfortunately, $|\psi|^2$ is not normalizable. Indeed, the authors point out that for the dynamics defined by (58), there can be no normalizable density on $\mathbb{R}^{2+3N}$ that is stationary with respect to the parameter s: All stationary measures for Z give stationary measures for T_i, which are, due to (56), proportional to the Lebesgue-measure on $\mathbb{R}$ (which is not normalizable). Therefore, $|\psi|^2$ is in this case very bad as a probability measure.

The way out of this goes as follows: Eliminate the explicit T_i-dependence from the equations of motion by first fixing a synchronization via setting $T_a(0) =: s_0$ and $T_b(0) =:$

[7] The velocity v_a^ψ depends via the wave function on the configuration of the a-system as well as on the configuration of the b-system in an instantaneous way, i.e. the nonlocal interaction goes along a $s = const.$-surface.

$s_0 + h$. This yields $T_a(s) = s_0 + s$ and $T_b(s) = s_0 + h + s$. The wave function is then given by

$$\psi^h(s, q_a, q_b) := \psi(T_a(s), q_a, T_b(s), q_b) = \psi(s_0 + s, q_a, s_0 + h + s, q_b) \tag{63}$$

and the guiding equations do not anymore depend on the subsystem times

$$\begin{aligned} \frac{dQ_a}{ds} &= v_a^{\psi^h}(s, Q_a(s), Q_b(s)) \\ \frac{dQ_b}{ds} &= v_b^{\psi^h}(s, Q_a(s), Q_b(s)), \end{aligned} \tag{64}$$

nor does the continuity equation:

$$\frac{\partial \rho^h}{\partial s} + div_{q_a}(\rho^h v_a^{\psi^h}) + div_{q_b}(\rho^h v_b^{\psi^h}) = 0 \tag{65}$$

with an equivariant density $\rho^h = |\psi^h|^2$ that is normalizable for $\psi^h(s) \in L^2(\mathbb{R}^{3N})$. It represents the distribution of crossings for the set of N-paths given by ψ^h of any hyperplane corresponding to the synchronization defined by h.

This distribution of crossings ρ^h allows for a statistical analysis and, in particular, the derivation of the quantum mechanical measurement formalism in the way described in the last section. The authors conclude that

> We thus have, with regard to our multitime Bohmian model, three levels of description: the microscopic dynamical level (...) which is multitime translation invariant; the statistical mechanical level, given by the quantum equilibrium hypothesis, which is, in precisely the same way, also multitime translation invariant (...); and the observational level given by the quantum measurement formalism, which is also apparently multitime translation invariant. ([3], page 19)

Nevertheless, on the first two levels, invariance is only achieved via an additional spacetime structure, the synchronization. Berndl, Dürr, Goldstein and Zanghì remark that any model can be made invariant under any kind of spacetime transformation by introducing such additional structures ([3], page 20). Hence, if one follows this way to achieve Lorentz invariance, the question of the physical significance of that structure has to be answered.

In the next section, I will describe the somehow straightforward "generalization" of multitime translation invariance to Lorentz invariance of a Bohmian theory. This will still inherit the problem of a not (yet) physically explained additional structure. Finally, in the last chapter, an approach by Goldstein and Tumulka will be presented, that combines nonlocality and Lorentz invariance without using more structure than already given by physics.

4.5 The Hypersurface Bohm-Dirac Model

A synchronized N-path

$$\begin{aligned} (X_1,...,X_N):\mathbb{R} &\longrightarrow \mathbb{R}^{4N} \\ s &\longmapsto (X_1(s),...,X_N(s)) \end{aligned} \tag{66}$$

should in general satisfy the analogue of (58):

$$\frac{dX_k}{ds} = v_k(X_1(s),...,X_N(s)). \tag{67}$$

The naive covariant reformulation of the guiding equation (23) for the N-particle Bohm-Dirac model yields

$$v_k^\mu = \overline{\psi}\gamma_1^0\cdots\gamma_k^\mu\cdots\gamma_N^0\psi \qquad \left(\text{"} = \text{"} \begin{pmatrix} \psi^\dagger\psi \\ \psi^\dagger\alpha_k\psi \end{pmatrix} \right) \tag{68}$$

which is *not* a 4-vector. Alternatively, one might set

$$v_k^\mu = \overline{\psi}\gamma_k^\mu\psi, \tag{69}$$

which *is* a 4-vector, but lacks statistical transparency: Unlike in the case of (68), where v_k^i/v_k^0 is of the form $\mathbf{j}/\rho$, this is not true for the dynamics given by (70), where

$$\frac{v_k^i}{v_k^0} = \frac{\overline{\psi}\gamma_k^i\psi}{\overline{\psi}\gamma_k^\mu\psi} = \frac{\psi^\dagger\gamma_1^0\cdots(\gamma_k^0)^2\alpha_k\cdots\gamma_N^0\psi}{\psi^\dagger\gamma_1^0\cdots(\gamma_k^0)^2\cdots\gamma_N^0\psi} \neq \frac{\mathbf{j}}{\rho} \tag{70}$$

Hence, it is not clear why quantum equilibrium should hold in this case, because there is no equivariant measure at hand to define typicality. But without quantum equilibrium, there is nor a known way to compare the theory with ordinary quantum mechanics, neither can any predictions for measurement results be obtained from it.

Albeit (68) is not a 4-vector, due to its physical equivalence to

$$\frac{dX_k}{ds} = \begin{pmatrix} 1 \\ \frac{\psi^\dagger\alpha_k\psi}{\psi^\dagger\psi} \end{pmatrix}, \tag{71}$$

(they only differ by a factor $\psi^\dagger\psi$, a real-valued function on $\mathbb{R}^{4N}$), it is statistically transparent in the sense of section 4.4.2 and should therefore be worked out in more detail anyway.

If the v_k in (67) were 4-velocities ($v_{k\mu}v_k^\mu = 1$), then the synchronization (now in accordance with proper time parametrization) would, in the nonrelativistic limit, reduce to the one of section 4.4.2.

In [4], Berndl, Dürr, Goldstein and Zanghì investigate the dynamics implied by (68) and suggest to achieve synchronization (the needed additional structure mentioned above) in a more geometrical way introducing a foliation on Minkowski spacetime. It will still come ad-hoc, since no equation of motion will be presented, nor will the question of where the foliation comes from be touched.

A foliation $\mathcal{F}$ on Minkowski space can be defined via a smooth function $f : M \to \mathbb{R}$ which has no critical points ($df \neq 0$ everywhere). The smooth hypersurfaces $f^{-1}(s)$ are then called "leaves of the foliation". Since $f = const.$ on each leaf, df_x vanishes on the tangent space of the leaf to which $x \in M$ belongs. Using the Lorentz metric, one can associate with df_x a normal vector field $\partial f(x)$. If this is everywhere timelike, the hypersurfaces of the foliation are everywhere spacelike and one gets, by normalization of $\partial f(x)$, a unit normal vector field $n(x)$, characteristic (indeed uniquely determined) for a given foliation. ([4], page 5)

4.5.1 The Equations of Motion

For N Dirac particles, the wave function satisfies N multitime Dirac equations

$$(i\gamma_k \cdot \partial_k - e\gamma_k \cdot A(x_k) - m)\psi = 0 \tag{72}$$

where $\gamma_k = I \otimes \cdots \otimes I \otimes \gamma \otimes I \otimes \cdots \otimes I$ with γ at the k-th entry. The authors point out that for the multitime equation (72), it is impossible (as far as it is known) to add an interaction potential. Despite this fact, the N Dirac particles might of course depend on each other (nonlocally) due to an entangled wave function.

The current

$$j_k = \overline{\psi}\gamma_1^0 \cdots \gamma_k \cdots \gamma_N^0 \psi \tag{73}$$

already discussed can be written covariantly by replacing γ_1^0 by $\gamma \cdot n$, where n is the future oriented unit vector normal to the constant time hyperplanes[8]:

$$j_k = \overline{\psi}(\gamma_1 \cdot n) \cdots \gamma_k \cdots (\gamma_N \cdot n)\psi \tag{74}$$

This suggests to take for the guiding equation ([4], page 7)

$$\frac{dX_k}{ds} = j_k \tag{75}$$

or, more generally, by reparameterization:

$$\frac{dX_k}{d\tau} = a j_k \tag{76}$$

with a positive scalar field a.

To see that the theory is invariant under such reparameterization, denote by $X_k(\Sigma_t)$ the crossing point of a N-path X_k with the $t = const.$-hypersurface Σ_t and by $\dot{X}_k(\Sigma_t)$ the tangent line to X_k at the points $X_k(\Sigma_t)$. The equation of motion for the N-path can then be written geometrically as

$$\dot{X}_k(\Sigma_t) \| j_k(X_1(\Sigma_t), ..., X_N(\Sigma_t)) \tag{77}$$

[8] See how "simple" it is to achieve Lorentz invariance by the implemetion of additional structure (here the synchronization).

Here, the dependence on a reference frame is only due to the specification of simultaneity hypersurfaces Σ_t of the foliation. The generalization to arbitrary spacelike hypersurfaces is thus straightforward. In (77), the particle trajectories are via the wave function explicitly evaluated with respect to the foliation. This is not true for (75) or (76), where the trajectories need not in general agree with the given foliation; the derivative with respect to an arbitrary parameter does not "force" the particles to move in accordance with any spacelike hypersurface as in the case of (77). The reason for this is that in (75), (76), the j_k's are not necessarily evaluated on the leaves of the foliation ([4], page 7).

The extension of (75) to a foliation by arbitrary curved spacelike hypersurfaces is obvious by generalizing the current (74):

$$j_k = \overline{\psi}(\gamma_1 \cdot n_1) \cdots \gamma_k \cdots (\gamma_N \cdot n_N)\psi \tag{78}$$

with $n_i = n(x_i)$ the unit vector fields normal to the leaves of the foliation at the points x_i. The equations of motion for the particle trajectories with respect to an arbitrary spacelike foliation (with leaves Σ) is then

$$\dot{X}_k(\Sigma) \| j_k(X_1(\Sigma), ..., X_N(\Sigma)). \tag{79}$$

It is important to assure that j_k and hence the particle trajectories are everywhere time- or lightlike in order to exclude causal loops (each leaf Σ should only be crossed once by each particle trajectory). Writing the current as

$$j_k = \psi^\dagger(\gamma_1^0\gamma_1 \cdot n_1) \cdots \gamma_k^0\gamma_k \cdots (\gamma_N^0\gamma_N \cdot n_N)\psi \tag{80}$$

(note that $\gamma^0\gamma \cdot n$ is a positive operator on each $\mathbb{C}^4$) and scalar multiplication by n yields

$$j_k \cdot n \geq 0, \tag{81}$$

and, therefore, j_k is future oriented (the 0-component is positive definite) and nowhere spacelike.

Equation (79) can also be written in a parameterized form using the function $f : M \to \mathbb{R}$, which generates the foliation ([4], page 8). Using the parameter s that corresponds to the hypersurface $f^{-1}(s)$ for the parameterization, one finds that $X_k(s)$ is on the leaf $f^{-1}(s)$. Then, from (79):

$$\frac{dX_k}{ds} \| j_k(X_1(s), ...X_N(s)) \tag{82}$$

Furthermore, it must hold that $f(X_k(s)) = s$ for all k and s, which is the same as asking that[9]

$$\frac{dX_k}{ds} \cdot \partial f(X_k(s)) = 1. \tag{83}$$

The equation of motion is thus given by

$$\frac{dX_k}{ds} = \frac{j_k(X_1(s), ...X_N(s))}{\partial f(X_k(s)) \cdot j_k(X_1(s), ..., X_N(s))}. \tag{84}$$

[9]morally speaking: differentiate the equation $f(X_k(s)) = s$ with respect to s

For a flat foliation with $x^0 = const.$-hyperplanes as leaves, this reproduces the Bohm-Dirac law

$$\frac{d\mathbf{Q}_k}{ds} = \frac{\psi^\dagger \alpha_k \psi}{\psi^\dagger \psi}, \tag{85}$$

since, in that case, $\partial f = (1, 0, 0, 0)$.

4.5.2 Statistical Analysis

The roadmap is the usual one ([4], page 8ff.). Berndl, Dürr, Goldstein and Zanghì first look for a distinguished measure on the set of N-paths that fulfil (79). A natural candidate for this crossing probability density is suggested by the fact that the currents (74) are 4-divergence free,

$$\partial_k \cdot j_k = 0. \tag{86}$$

This can easily be derived from the Dirac equation (72) and its adjoint. In the covariant spirit of the model, instead of taking the 0-component of j_k, multiply it by n_k and take the resulting product as the crossing probability density:

$$\rho = j_k \cdot n_k = \overline{\psi}(\gamma_1 \cdot n_1) \cdots (\gamma_N \cdot n_N)\psi \tag{87}$$

This does not anymore depend on k[10]. From this fact, together with (86), follows the equivariance of ρ, which is thus the measure that yields quantum equilibrium ([4], page 9).

It should be stressed that this measure ρ only depends on the wave function, which solves (72). This wave function will be evaluated on a hypersurface Σ belonging to the foliation $\mathcal{F}$. For each $\Sigma \in \mathcal{F}$, there will be a distinct ψ^Σ and hence ρ^Σ, providing quantum equilibrium on each leaf due to the equivariance of ρ: This is what lies behind the sentence: "For each hypersurface $\Sigma \in \mathcal{F}$, the distribution of crossings by N-paths (i.e. the spatial probability density on that simultaneity surface) is given by ρ^Σ."

The foliation $\mathcal{F}$ and its hypersurfaces Σ can therefore (at least in this formulation) hardly be seen as only *mathematical* artifacts. On the contrary, the foliation has fundamental physical importance: First, it defines the dynamics in equation (79) and secondly, it defines where quantum equilibrium holds: for the intersection points of the N-paths with its leaves ([4], page 8).

To proof the equivariance of ρ, Berndl, Dürr, Goldstein and Zanghì first formulate the continuity equation for the dynamics given by (79) and then show that ρ solves it ([4], page 9f.):

Proof of the equivariance of ρ:

The probability distribution for the positions of the N particles on a foliation leaf Σ is given by a function $R_\Sigma : \Sigma^N \longrightarrow \mathbb{R}$ such that

$$\begin{aligned} &P(\text{the i-th particle crosses } \Sigma \text{ in the 3-volume } \delta x_i,\ i = 1, ..., N) \\ &\qquad = R_\Sigma(x_1, ..., x_N)\delta x_1 \cdots \delta x_N \end{aligned} \tag{88}$$

[10] Although the alternative current $j_k = \overline{\psi}\gamma_k\psi$ is also 4-divergence free, for that current $j_k \cdot n_k$ is *not* independent of k. Therefore, the corresponding model is not statistically transparent.

On an infinitesimally close surface Σ', the probability density is given by $R_{\Sigma'}$ calculated at $(x'_1, ..., x'_N) \in (\Sigma')^N$, with $x'_i \in \Sigma'$ obtained from $x_i \in \Sigma$ by transition normal to the hyperplanes.

The continuity equation to be established will be a law for the difference between R_Σ and $R_{\Sigma'}$: This must be equal to the probability flux through the *lateral* boundaries of the box between $\delta x_1 \times \cdots \times \delta x_N \subset \Sigma^N$ and $\delta x'_1 \times \cdots \times \delta x'_N \subset \Sigma'^N$ in configuration spacetime:

$$R_{\Sigma'}(x'_1, ..., x'_N)\delta x'_1 \cdots \delta x'_N - R_\Sigma(x_1, ..., x_N)\delta x_1 \cdots \delta x_N = \tag{89}$$

$$-\sum_{k=1}^{N} \delta x_1 \cdots \widehat{\delta x_k} \cdots \delta x_N \int_{\partial(\delta x_k)} (R_\Sigma v_k)(x_1, ..., x_{k-1}, y, x_{k+1}, ..., x_N) \cdot (u_k \delta\tau)(y) dS_k \tag{90}$$

where $\partial(\delta x_k)$ is the 2D-boundary of $\delta x_k \subset \Sigma$, dS_k its area-element, u_k is a unit normal vector field in the tangent space of Σ normal to $\partial(\delta x_k)$, $\delta\tau(y)$ is the distance in Minkowski space between $y \in \Sigma$ and $y' \in \Sigma'$ such that $y' = \delta\tau(y)n(y) + y$ and, finally,

$$v_k = \frac{j_k}{j_k \cdot n_k} \tag{91}$$

is the *covariant* velocity of particle k.

The authors point out that the continuity equation (89) is valid for any motion given by (79) ([4], page 10). In particular, the currents j_k do not have to be 4-divergence free.

ρ and ρ' are defined on configuration Minkowski space M^N and can therefore be decomposed as ([4], page 10)

$$\begin{aligned}
&\rho(x'_1, ..., x'_N)\delta x'_1 \cdots \delta x'_N - \rho(x_1, ..., x_N)\delta x_1 \cdots \delta x_N \\
&= \rho(x'_1, ..., x'_N)\delta x'_1 \cdots \delta x'_N \\
&-\rho(x_1, x'_2, ..., x'_N)\delta x_1 \delta x'_2 \cdots \delta x'_N \\
&+\rho(x_1, x'_2, ..., x'_N)\delta x_1 \delta x'_2 \cdots \delta x'_N \\
&- \cdots \\
&+ \cdots \\
&-\rho(x_1, ...x_{N-1}, x'_N)\delta x_1 \cdots \delta x_{N-1}\delta x'_N \\
&+\rho(x_1, ...x_{N-1}, x'_N)\delta x_1 \cdots \delta x_{N-1}\delta x'_N \\
&-\rho(x_1, ..., x_N)\delta x_1 \cdots \delta x_N
\end{aligned} \tag{92}$$

Since $\delta x' = \delta x + \mathcal{O}$, the "leading-order version" of (92) will only have one primed coordinate in each summand. Using that $j_k \cdot n_k$ is independent of k, one ends (to leading order) up with

$$\begin{aligned}
\sum_{k=1}^{N} \delta x_1 \cdots \widehat{\delta x_k} \cdots \delta x_N \big(& j_k(x_1, ..., x'_k, ..., x_N) \cdot n(x'_k)\delta x'_k \\
& - j_k(x_1, ..., x_k, ..., x_N) \cdot n(x_k)\delta x_k \big),
\end{aligned} \tag{93}$$

while the right hand side of (89) becomes

$$-\sum_{k=1}^{N}\delta x_1\cdots\widehat{\delta x_k}\cdots\delta x_N\int\limits_{\partial(\delta x_k)}(j_k\cdot u_k)\delta\tau dS_k \tag{94}$$

Subtracting (94) from (93), one obtains

$$\sum_{k=1}^{N}\delta x_1\cdots\widehat{\delta x_k}\cdots\delta x_N\big(j_k(x_1,...,x'_k,...,x_N)\cdot n(x'_k)\delta x'_k$$
$$-j_k(x_1,...,x_k,...,x_N)\cdot n(x_k)\delta x_k+\int\limits_{\partial(\delta x_k)}(j_k\cdot u_k)\delta\tau dS_k\big) \tag{95}$$

where the first two terms in the bracket are to leading order equal to the integral over the "time-boundary" and can thus be brought together with the integral over the "space-boundary" $\partial(\delta x_k)$ to yield an integral of j_k over the whole "spacetime-boundary" of the "box" over δx_k between Σ and Σ'. But since j_k is 4-divergence free, all these integrals for the different particles vanish and so does the sum. Hence, $\rho = j_k \cdot n_k$ solves (89) and is an *equivariant* probability density.□

For those who worry about the "leading-order-character" of the above proof, the authors also give a local version of it in ([4], page 10ff).

4.5.3 Statistical Value of the HBD-Model

The aim of the model presented above is to reproduce the quantum mechanical prediction for position measurement on each leaf of a foliation $\mathcal{F}$. This was achieved thanks to the equivariance of ρ on the hypersurfaces, which yields the probability distribution for the N-paths. For one particle as well as for N independent particles (the wave function will then have product structure), the motion will be independent of the foliation ([4], page 12).

But, in the case of an entangled wave function, this is not in general true. On hypersurfaces which do not belong to the foliation, the quantum mechanical distribution does not have to agree with the distribution of crossings of this surface. In fact, according to what was shown in section 4.3, quantum equilibrium *cannot* hold on all spacelike hypersurfaces. However, by the same argument as discussed at the end of section 4.3, this does not lead to *observable* conflicts with quantum mechanical predictions ([4], page 12f.).

The foliation used to derive the Hypersurface Bohm-Dirac (HBD) model clearly has *physical* significance, in particular; it cannot be seen as some kind of a gauge. If the model shall be called Lorentz invariant in a serious way, the foliation must be part of the objective physical world. Therefore, a law for its dynamics must be found in order to complete the theory ([4], page 13f.).

In the next section, I will present a different approach to combine quantum nonlocality and relativity, which has the advantage of not assuming any additional structure whose physical significance might be questionable.

4.6 Two Arrows of Time

In [9], Goldstein and Tumulka present a model that describes nonlocal quantum phenomena[11] while being completely relativistic. It can, like the models described so far, be called realistic or Bohmian in the sense that it assumes *some* kind of reality on the microscopic level, establishing a particle ontology. Therefore, it does, unlike ordinary quantum mechanics, not depend on the role of an observer on the fundamental level. Whether the model is favorable or not from the physical point of view, it proves the existence of objective, nonlocal *and* relativistic quantum theories without using more structure than implied by physics itself: Minkowski spacetime and the wave function.

Unfortunately, the theory lacks statistical transparency: according to the authors, the dynamics do not have a distinguished probability measure on the set of the particle trajectories. Thus, it cannot be justified why quantum equilibrium should hold and the model can therefore not be compared with ordinary quantum mechanical predictions.

4.6.1 The idea of two time-arrows pointing in opposite directions

The basic idea of the model is that there are two arrows of time, a *microcausal* (c) and a *macrocausal* (θ) one. The latter θ coincides with the usual thermodynamic time arrow, whose direction is given by the Second Law, whereas the microcausal arrow c will always point in the opposite direction of θ, i.e. "backwards in macrocausal time"([9], page 3).

The equation of motion of the model, an equation for particle velocities, will be such that the velocity of a certain particle will depend on the wave function of the system evaluated on the intersection points of the trajectories of all particles with its own *future* light cone as well as on their velocities at these points. Therefore, within this model, the past can be calculated from the future, but there is no obvious way how to compute the future from the past on the microscopic level: Causality "goes" from future to past, suggesting the microcausal arrow of time c. Hence, on the level of individual particles, past and future are reversed:

$$\begin{aligned} \text{future}_c &= \text{past}_\theta \\ \text{past}_c &= \text{future}_\theta \end{aligned} \tag{96}$$

Therefore, within this framework, there is no problem with causal paradoxes on the microscopic level: events $prior_c$ determine the $future_c$, but never vice versa. The arrow of time we feel in our every-day-life is of course the thermodynamic one pointing in the opposite direction.

Due to this structure, the equation of motion is local with respect to c and nonlocal with respect to θ. The nonlocality is explicitly macroscopic, since an experimenter might modify particle trajectories separated spacelike from a particle at spacetime point p, which $later_\theta$ cross the $future_\theta$ light cone of p. Since the velocity of the particle at p depends on the wave function evaluated at those crossing points, it can be changed by the intervention of the experimenter. Nevertheless, the theory is Lorentz invariant ([9], page 4).

[11] i.e., it includes the possibility of effects traveling faster than light

To get a better understanding of the maybe weird concept of two opposite arrows of time, it might be useful to compare it with classical statistical mechanics: Newton's equation of motion is time reversal invariant. There is thus no such thing as a microscopic arrow of time in classical physics. Despite this time symmetry at the individual particle level, time asymmetry arises in the statistical limit of many particles: the Second Law provides us with a macroscopic (thermodynamic) arrow of time. For the model under consideration here, this implies that we are in fact *free to choose* a direction at the microscopic level (it will simply not *affect* the macroscopic arrow θ). The advantage of choosing c to be opposite to θ is that this makes it possible to include nonlocality ([9], page 4).

4.6.2 The Equations of Motion

As in all models described above, the dynamics are completely defined by two fundamental equations:

First the (by now) well-known multitime Dirac equation (without an interaction potential):

$$\gamma_i^\mu(i\hbar\partial_{i,\mu} + eA_\mu(x_i))\psi = m\psi \tag{97}$$

where $\gamma_i^\mu = I \otimes \cdots \otimes I \otimes \gamma^\mu \otimes I \otimes \cdots \otimes I$ with γ^μ at the i-th entry. And secondly, the guiding equation for the particle trajectories (which have to be timelike):

$$\frac{dX_i^{\mu_i}}{ds_i} \| \overline{\psi}(\gamma^{\mu_1} \otimes \cdots \otimes \gamma^{\mu_N})\psi \prod_{j\neq i} \frac{dx_j^{\nu_j}}{ds_j}(s_{ret,j}(p_i))\eta_{\mu_j\nu_j} \tag{98}$$

Here, $s_{ret,j}(p_i)$ denotes the value of s such that $x_j^{\mu_j(s)}$ lies on the $past_c$ light cone of p_i, $\overline{\psi} = \psi^\dagger\gamma^0 \otimes \cdots \otimes \gamma^0$ is the adjoint N-particle spinor and ψ, $\overline{\psi}$ are evaluated at the intersection points of the N particle trajectories with the $past_c$ light cone of p_i: $(x_1(s_{ret,1}(p_i)), \ldots, x_N(s_{ret,N}(p_i)))$. Analogously, $u_j^{\nu_j} \equiv \frac{dx_j^{\nu_j}}{ds_j}(s_{ret,j}(p_i))$ denotes the velocities of all particles j at the intersection points with the $past_c$ light cone of the i-th particle at p_i.

In (98) the tensor $J \in T_{p_1}M \otimes \cdots \otimes T_{p_N}M$,

$$J^{\mu_1\cdots\mu_N} = \overline{\psi}(\gamma^{\mu_1} \otimes \cdots \otimes \gamma^{\mu_N})\psi \tag{99}$$

gets its indices contracted by

$$u_j^{\nu_j}\eta_{\mu_j\nu_j}$$

for all $j \neq i$, which yields $j_i^{\mu_i} \in T_{p_i}M$.

Hence, (98) may also be written as

$$\frac{dX_i^{\mu_i}}{ds_i} \| j_i^{\mu_i} = J^{\mu_1\cdots\mu_N} u_1^{\nu_1}\eta_{\mu_1\nu_1} \ldots \widehat{u_1^{\nu_1}\eta_{\mu_1\nu_1}} \cdots u_N^{\nu_N}\eta_{\mu_N\nu_N} \tag{100}$$

The result of this contraction, $j_i^{\mu_i}$, defines a 1D-subspace of $T_{p_i}M$ to which, according to the law, the i-th particle trajectory has to be tangent.

Note that for $N = 1$, this model reproduces the one particle Bohm-Dirac law. The many particle Bohm-Dirac law considered e.g. in the beginning of section 4.5 can now be reformulated to give

$$\frac{dX_i^{\mu_i}(s)}{ds} \| J^{0 \ldots \mu_i \ldots 0}(p_1, \ldots, p_N) \tag{101}$$

where a preferred Lorentz frame is needed. Indeed, this can be obtained by contracting all but the i-th index of J, using the unit normal vector fields of the simultaneity hypersurfaces belonging to a given foliation. In addition, J is then evaluated on the crossing points of the N-path with that surface.

In (98) resp. (100) however, such a foliation is not needed: Instead of spacelike hypersurfaces, light cones are used and contraction of the indices of J is performed using the 4-velocities of all other particles evaluated on those light cones.

The problem with (98) resp. (100) is, as already mentioned, a rather serious one: The dynamics implied by it do not yield a continuity equation for $\psi^\dagger\psi$. Hence, there is no equivariant measure available and no direct comparison with quantum mechanical predictions can be made. Only in the nonrelativistic limit, where the future light cones fall together with $t = const.$-planes, the model becomes the ordinary Bohm-Dirac model, which is consistent with the $\psi^\dagger\psi$-distribution for position measurement ([9], page 7).

For the model described above, it is important that $j_i^{\mu_i}$ is nowhere spacelike. This would be true if and only if its scalar product with any nonzero timelike vector $u_i^{\nu_i}$ was nonzero:

$$\lambda := j_i^{\mu_i} u_i^{\nu_i} \eta_{\mu_i \nu_i} = \overline{\psi}(\gamma^{\mu_1} \otimes \cdots \otimes \gamma^{\mu_N})\psi \prod_j u_j^{\nu_j} \eta_{\mu_j \nu_j} \tag{102}$$

By suitable Lorentz tranformations on $T_{p_1}, \ldots, T_{p_1}$, all u_j's can be brought to the form $(1, 0, 0, 0)$, changing also ψ into ψ'. But then $\lambda = \pm\overline{\psi'}(\gamma 0 \cdots \gamma^0)\psi' = \psi'^\dagger\psi'$, which is larger than zero unless $\psi' = 0$ and hence $\psi = 0$ ([9], page 8).

4.6.3 Micro- and Macrocausality

To explain the consequences of two different causalities, micro- and macrocausality, for their model, Goldstein and Tumulka give an example of two electromagnetic potentials A_μ and A'_μ which are everywhere equal except in a small spacetime region $U(p)$ around p (see Fig. 2), ([9], page 10).

The wave functions ψ and ψ' obtained from the multitime Dirac equation will then differ for those configuration spacetime points $(p_1, \ldots, p_N)$ for which at least on p_i lies insight or on the $future_\theta$ "light cone" of the area $U(p)$ ([9], page 10). Whereas any influences on the wave function will be such that the external cause was $prior_\theta$ to the effect, this is not true for the particle trajectories: If S is a solution of the equation of motion (98) for ψ and S' a solution for ψ', these two solutions won't agree in the past of $U(p)$ nor in the future of $U(p)$. Hence, causation goes for this model in both directions. Nonetheless, Goldstein and Tumulka point out that, despite causation in both directions, the equations of motion do not allow for any causal paradoxes: the model does not contain the possibility to construct causal loops, because the Dirac equation is solved from $initial_\theta$ conditions and, afterwards, the guiding equation is solved from

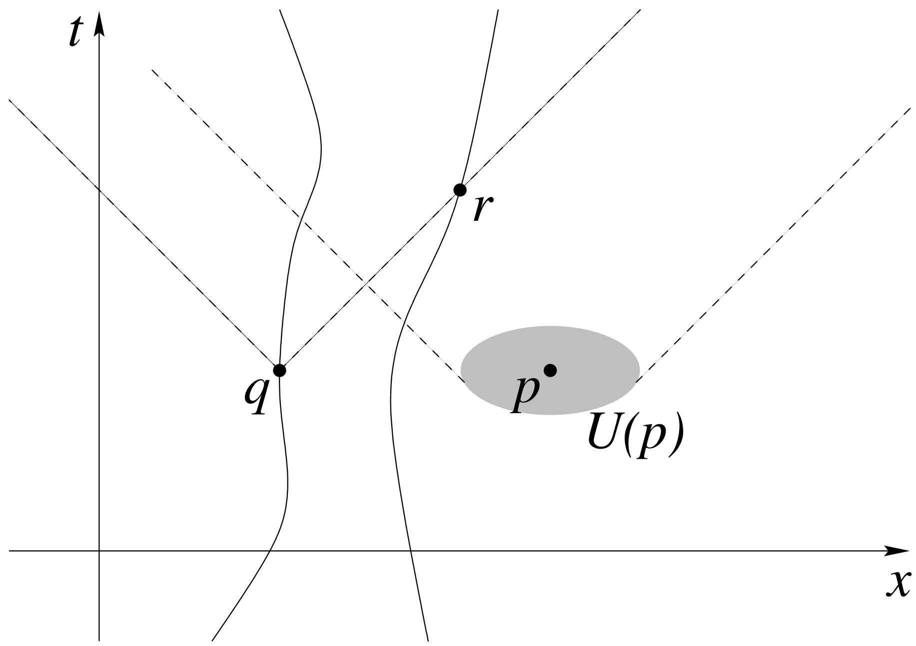

Figure 2: Spacetime diagram: Changing the electromagnetic potential in the region $U(p)$ will influence the wave function in its future "light cone". But this will change the wave function at r and hence influence the velocity of the particle at q (taken from [9]).

$initial_c$ conditions. Loosely speaking, the two kinds of causalities do not "interact" ([9], page 11).

The nonlocal backwards (from *our* macroscopic point of view) causation on the single particle level described above was suggested to account for nonlocal behavior of quantum systems in association with entanglement[12]. In connection with this entanglement, Bell has shown the nonlocal character of nature which, however, as far as we know, does not allow for the transmission of signals. The backwards microcausality was not designed for giving rise to the possibility of macroscopic backward causation ([9], page 11).

[12]Therefore, the two arrows were chosen in opposite directions. The wave function evaluated on the *future* light cone of a point p has traveled through spacetime regions spatially separated from p, carrying "information" from these regions finally to the point p to calculate the velocity of the particle there.

5 Conclusion

The aim of this thesis was to present recent attempts to achieve a Lorentz invariant deterministic quantum theory. For all models that have been presented, the dynamics on the microscopic level were given by the respective modifications of Bohmian Mechanics. As it turned out, the major obstacle on the way towards a relativistic Bohmian theory is the reconciliation of nonlocality and Lorentz invariance under the condition of statistical transparency. Within the Hypersurface Bohm-Dirac model this is achieved, albeit only by incorporating a foliation of spacetime, whose law of motion still has to be found. Nonetheless, the model shows that the three requirements, Lorentz invariance, nonlocality and statistical transparancy (leading to the reproduction of the predictions of ordinary quantum formalism), can all be met within one model. Moreover, the Two Arrows of Time model shows that Lorentz invariance and nonlocality are compatible even without the incorporation of additional structure. The fact that no complete relativistic objective quantum theory has been found so far should, in my opinion, not lead to the conclusion that looking for one is a waste of time. If one wants to avoid the vagueness of orthodox quantum theory and in addition prefers a deterministic description of nature, Bohmian Mechanics seems to be a vantage point for further research.

Selbstständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbstständig verfaßt und keine anderen als die genannten Quellen und Hilfsmittel verwendet zu haben.

München, den 3. April 2008

Niklas Boers

References

[1] J.S. Bell: *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, 2004

[2] D. Bohm and B.J.Hiley: *The Undivided Universe: An ontological interpretation of quantum theory*, Routledge, London and New York, 1993, p. 271 to p. 295

[3] K. Berndl, D. Dürr, S. Goldstein and N. Zanghì, Phys. Rev. A 53, 2062 (1996)

[4] K. Berndl, D. Dürr, S. Goldstein and N. Zanghì, Phys. Rev. A 60, 2729 (1999) and quant-ph/9801070

[5] D. Dürr, S. Goldstein, R. Tumulka and N. Zanghì, quant-ph/0407116

[6] D. Dürr, S. Goldstein, and N. Zanghì: *Quantum Equilibrium and the Role of Operators as Observables in Quantum Theory*, J. Stat. Phys. 116, 959-1055 (2004)

[7] D. Dürr: *Bohmsche Mechanik als Grundlage der Quantenmechanik*, Springer, Berlin, 2001

[8] D. Dürr, S. Goldstein, and N. Zanghì, J. Stat. Phys. 67, 843 (1992)

[9] S. Goldstein, R. Tumulka: *Opposite Arrows of Time Can Reconcile Relativity and Nonlocality*, submitted, arXiv: quant-ph/0105040

[10] L. Hardy, Phys. Rev. Lett. 68, 2981 (1992)

[11] T. Maudlin: *Quantum Non-locality and Relativity*, Blackwell, Cambridge, 1994