Fundamental Theories of Physics 215

Angelo Bassi
Sheldon Goldstein
Roderich Tumulka
Nino Zanghi *Editors*

Physics and the Nature of Reality

Essays in Memory of Detlef Dürr

Springer

Fundamental Theories of Physics

Volume 215

The international monograph series “Fundamental Theories of Physics” aims to stretch the boundaries of mainstream physics by clarifying and developing the theoretical and conceptual framework of physics and by applying it to a wide range of interdisciplinary scientific fields. Original contributions in well-established fields such as Quantum Physics, Relativity Theory, Cosmology, Quantum Field Theory, Statistical Mechanics and Nonlinear Dynamics are welcome. The series also provides a forum for non-conventional approaches to these fields. Publications should present new and promising ideas, with prospects for their further development, and carefully show how they connect to conventional views of the topic. Although the aim of this series is to go beyond established mainstream physics, a high profile and open-minded Editorial Board will evaluate all contributions carefully to ensure a high scientific standard.

Angelo Bassi · Sheldon Goldstein ·
Roderich Tumulka · Nino Zanghì
Editors

Physics and the Nature of Reality

Essays in Memory of Detlef Dürr

Springer

Editors
Angelo Bassi
Department of Physics
University of Trieste
Trieste, Italy

Sheldon Goldstein
Department of Mathematics
Rutgers University
Piscataway, NJ, USA

Roderich Tumulka
Fachbereich Mathematik
Eberhard Karls University of Tübingen
Tübingen, Baden-Württemberg, Germany

Nino Zanghì
Dipartimento di Fisica
Università di Genova
Genova, Italy

ISSN 0168-1222 ISSN 2365-6425 (electronic)
Fundamental Theories of Physics
ISBN 978-3-031-45433-2 ISBN 978-3-031-45434-9 (eBook)
https://doi.org/10.1007/978-3-031-45434-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Detlef Dürr (1951–2021)
(*Photo credit* Dr. Tilo Moser 2008, reproduced with kind permission)

Preface

This volume collects papers dedicated to the memory of Detlef Dürr (4 March 1951–3 January 2021), our dear friend and colleague. His unexpected and untimely passing came as a great shock to us. He was a great inspiration to our research, to our way of thinking about the laws of physics, and to our idea of honest and serious scientific debate. With this book, we would like to honor his legacy by collecting scientific works that present research influenced by Detlef, as well as some essays describing personal recollections.

Some of the chapters in this book are writeups of talks given at the *Laws of Nature Conference Remembering Detlef Dürr*, held on August 8–12, 2022, at Ludwig-Maximilians University in Munich (Germany), where Detlef worked for more than 30 years. The conference was organized by Angelo Bassi, Dirk-André Deckert, Dustin Lazarovici, Peter Pickl, Paula Reichert, and Ward Struyve.

Other chapters were written independently of the conference, upon our invitation to contribute to this collection of essays commemorating Detlef. In content and style, some chapters are physical, others philosophical or mathematical. Some present novel results in a continuation of lines of research inspired or initiated by Detlef, some review particularly relevant results, and some provide personal recollections. Some were written by former students, and some by colleagues.

The diversity of themes reflects Detlef's many interests. We have collected 31 chapters and grouped them by topic into five parts: recollections, foundations of quantum mechanics, relativistic quantum theory, mathematical physics, and philosophy of physics.

Detlef was fond of the ancient Greek philosophers. Reinhard Lang, who unfortunately could not contribute a chapter to this book, has sent us, as a memento of Detlef, an image of Pythagoras shown in Fig. 1 along with the following quotation from Whitehead's lecture *Science and the Modern World*[1]:

> The Platonic world of ideas is the refined, revised form of the Pythagorean doctrine that number lies at the base of the real world.

[1] A. N. Whitehead: Science and the Modern World. *Humana Mente* **1(3)**: 380–385 (1925).

Fig. 1 Sculpture of Pythagoras on the west façade of the cathedral of Chartres, France (*Photo credit* Dr. Nick Thompson, University of Auckland, 2009, reproduced with kind permission)

We the editors would like to thank the authors who contributed to this volume, the referees who helped bring the chapters to their optimal form, and the staff at Springer for their careful and efficient work.

Trieste, Italy — Angelo Bassi
Highland Park, NJ, USA — Sheldon Goldstein
Tübingen, Germany — Roderich Tumulka
Arquata Scrivia, AL, Italy — Nino Zanghì
July 2023

Contents

Contributors

Sergio Albeverio Institute for Applied Mathematics and Hausdorff Center for Mathematics, University of Bonn, Bonn, Germany

Valia Allori University of Bergamo, Bergamo, Italy

Jeffrey Barrett Logic and philosophy of science, University of California Irvine, Irvine, CA, USA

Christian Beck Department of Humanities and Arts, Technion-Israel Institute of Technology, Haifa, Israel

Lea Boßmann Mathematisches Institut, Ludwig-Maximilians-Universität München, München, Germany

Jean Bricmont Institut de recherche en mathématiques et physique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Craig Callender University of California San Diego, San Diego, CA, USA

Holly K. Carley Department of Mathematics, New York City College of Technology, City University of New York, New York, NY, USA

Eddy Keming Chen Department of Philosophy, University of California, San Diego, CA, USA

Dirk-André Deckert Mathematisches Institut, Ludwig-Maximilians-Universität München, München, Germany

Carlos F. Destefani Departament d'Enginyeria Electrónica, Universitat Autónoma de Barcelona, Bellaterra, Barcelona, Spain

Michael Esfeld Department of Philosophy, University of Lausanne, Lausanne, Switzerland

David K. Ferry School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ, USA

Rodolfo Figari Gran Sasso Science Institute, L'Aquila, Italy

Jürg Fröhlich Institut für theoretische Physik, Eidgenössische Technische Hochschule, Zurich, Switzerland

Zhou Gang Department of Mathematics and Statistics, Binghamton University, Binghamton, NY, USA

Michael J. W. Hall Centre for Quantum Dynamics, Griffith University, Brisbane, QLD, Australia

Hannes Herrmann Mathematisches Institut, Ludwig-Maximilians-Universität München, München, Germany

Günter Hinrichs Universität Augsburg, Augsburg, Germany

Mario Hubert Department of Philosophy, The American University in Cairo, Cairo, Egypt

Michael K.-H. Kiessling Department of Mathematics, Rutgers University, Piscataway, NJ, USA

Federico Laudisa Department of Humanities and Philosophy, University of Trento, Trento, Italy

Dustin Lazarovici Humanities and Arts Department, Technion–Israel Institute of Technology, Haifa, Israel

Nikolai Leopold Department of Mathematics and Computer Science, University of Basel, Basel, Switzerland

Matthias Lienert Marvel Fusion GmbH, Theresienhöhe 12, Munich, Germany

Barry Loewer Department of Philosophy, Rutgers University, New Brunswick, NJ, USA

Tim Maudlin Department of Philosophy, New York University, New York, NY, USA

Franz Merkl Mathematisches Institut, Ludwig-Maximilians-Universität München, München, Germany

David Mitrouskas Institute of Science and Technology Austria (ISTA), Klosterneuburg, Austria

Travis Norsen Smith College, Northampton, MA, USA

Markus Nöth Mathematisches Institut, Ludwig-Maximilians-Universität München, München, Germany

Xabier Oianguren-Asua Departament de Física, Universitat Autónoma de Barcelona, Bellaterra, Barcelona, Spain

Xavier Oriols Departament d'Enginyeria Electrónica, Universitat Autónoma de Barcelona, Bellaterra, Barcelona, Spain

Sören Petrat School of Science, Constructor University, Bremen, Germany

Peter Pickl Fachbereich Mathematik, Eberhard-Karls-Universität Tübingen, Tübingen, Germany

Alessandro Pizzo Dipartimento di Matematica, Università di Roma "TorVergata", Rome, Italy

Paula Reichert Mathematisches Institut, Ludwig-Maximilians-Universität München, München, Germany;
Department for Humanities & Arts, Technion-Israel Institute of Technology, Haifa, Israel

Aaron Schaal Fachbereich Mathematik, Eberhard-Karls-Universität Tübingen, Tübingen, Germany

Ward Struyve Instituut voor Theoretische Fysica, Katholieke Universiteit Leuven, Leuven, Belgium;
Centrum voor Logica en Filosofie van de Wetenschappen, Katholieke Universiteit Leuven, Leuven, Belgium

Daniel Sudarsky Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Mexico City, Mexico

A. Shadi Tahvildar-Zadeh Department of Mathematics, Rutgers University, Piscataway, NJ, USA

Alessandro Teta Dipartimento di Matematica, La Sapienza Università di Roma, Roma, Italy

Roderich Tumulka Fachbereich Mathematik, Eberhard-Karls-Universität Tübingen, Tübingen, Germany

Lev Vaidman Raymond and Beverly Sackler School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv, Israel

Matteo Villani Departament d'Enginyeria Electrónica, Universitat Autónoma de Barcelona, Bellaterra, Barcelona, Spain

Isaac Wilhelm Department of Philosophy, National University of Singapore, Singapore, Singapore

Howard M. Wiseman Centre for Quantum Dynamics, Griffith University, Brisbane, QLD, Australia

Stephanie Zhou Department of Mathematics, Rutgers University, Piscataway, NJ, USA

Recollections

Some Things I Have Learned From Detlef Dürr

Roderich Tumulka

Abstract Detlef Dürr (1951–2021) was a theoretical and mathematical physicist who worked particularly on the foundations of quantum mechanics, electromagnetism, and statistical mechanics. This piece is a rather personal look back at him and his science.

1 Introduction

Sadly, Detlef Dürr passed away after short and severe illness on January 3, 2021, at the age of 69. I had the privilege to be one of his Ph.D. students in 1998–2001, and we continued collaborating until his death. At first, he was my supervisor, then he became my colleague and dear friend. I still find his research achievements very impressive and inspiring, and I would like to try to convey here why.

2 An Example

Detlef's research intertwined mathematics, physics, and philosophy. To illustrate this, I pick as an example a mathematical result Detlef published in 2004 jointly with his long-time collaborators Shelly Goldstein and Nino Zanghì [12]. I call it the main theorem about POVMs. (A POVM, or positive-operator-valued measure, is for our purposes a family of positive operators on a Hilbert space that add up to the identity operator.) Here is a (somewhat informal) statement of the theorem.

Theorem *Let S be a quantum system with Hilbert space $\mathscr{H}$. For any conceivable experiment $\mathscr{E}$ that can be conducted on S when it has arbitrary wave function, there*

R. Tumulka (✉)
Fachbereich Mathematik, Eberhard-Karls-Universität Tübingen, Tübingen 72076, Germany
e-mail: roderich.tumulka@uni-tuebingen.de

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_1

is a POVM E on the set $\mathscr{Z}$ of possible outcomes of $\mathscr{E}$ acting on $\mathscr{H}$ such that for every $\psi \in \mathscr{H}$ with $\|\psi\| = 1$, the probability distribution of the outcome Z of $\mathscr{E}$ is

$$\mathbb{P}_\psi(Z = z) = \langle \psi | E(z) | \psi \rangle \quad \textit{for all } z \in \mathscr{Z}.$$

The statement is mathematics as it can be formulated rigorously and given a proof, see [12] or [20, Sect. 5.1]. It is physics as it concerns the physics question of which probability distributions occur in quantum experiments and how they can depend on the wave function ψ. It also has a flavor of philosophy (although it is an indisputable scientific fact) because of its foundational character, or even more so because it is based on an analysis of the measurement process that is usually treated as un-analyzable, but most of all because of the status it confers to the observables: In almost every textbook on quantum mechanics, observables are self-adjoint operators that enter the theory through a *postulate*; but here, observables are given by POVMs, and they come in through a mathematical *analysis* of the measurement process. That is a kind of radical break with the attitude prevailing in quantum mechanics: analysis instead of postulate!

This theorem opens up a new way of thinking about quantum observables. For me, it removes the mystery about quantum observables. Indeed, in textbook quantum mechanics the observables retained a mysterious air as they were thought of as physical quantities but do not actually have values, not to speak of the fact that they do not commute. Here, observables are something different: they are mathematical objects that encode how the probability distribution of the random outcome Z of an experiment depends on ψ. Since this dependence is quadratic, it is rather obvious that this mathematical object should be an operator $E(z)$. And operators do not commute in general. No mystery left.

The above theorem has various applications. For example, it provides a clear justification of why superselection rules hold under suitable conditions, of various versions of the no-signaling theorem, of why two probability distributions over wave functions with equal density matrices are empirically indistinguishable, and of why an experiment on one of two entangled systems has outcomes with distribution determined by the reduced density matrix.

3 Detlef's Questions

A key trait that makes Detlef's research findings come alive to me is that they are about unveiling how the world works: what its fundamental physical laws are, and how to explain macroscopic phenomena like randomness or the arrow of time. In other words, it impresses me that Detlef's research adds to our understanding of the world. That may almost seem impossible today. We are used to the idea that our understanding of the world is based on the works of Lavoisier and Copernicus and Einstein and other guys from previous centuries, but not on contributions from contemporary scientists. We are used to the idea that although present-day scientists

do computations much more difficult than those of Lavoisier and Copernicus and Einstein, the present-day results are of lesser significance and concern details that only few specialists have ever heard of. But Detlef didn't see it that way. He saw possibilities of progress in our days on rather fundamental questions of science.

I have learned from him that when we study scientific issues, we need to get to their bottom, that is, to reach full understanding. For example, he wanted to get to the bottom of the second law of thermodynamics (the statement that the entropy of a closed system can only increase) and understand the origin of the arrow of time. He thought it was part of a scientist's job to actually *understand* her or his field, and that the quality of a researcher's results will depend on how well she or he understands which of the available theories or approaches work and why. In particular, Detlef thought it was part of his job to understand whether and how statistical mechanics explains entropy increase and the arrow of time. Since a lot of different approaches to this question have been proposed in the literature, for example some using the Gibbs entropy and some the Boltzmann entropy [18], he felt it was important to think through which of these approaches were valid, and he inspired me to do the same. I arrived at the same conclusion as he had, not because I would repeat what he said but because he had arguments that made sense (and that was ultimately because he had honestly and seriously thought about it). The conclusion was, put very briefly, that Boltzmann's approach [6], based on typicality, provided the crucial elements. This understanding then formed the basis of Detlef's technical, mathematical work about the origin of randomness from chaos and typicality, in particular work on deriving that the velocity of a tracer particle in a classical many-body system (such as a hard sphere gas) follows approximately a Wiener process [8].

4 Quantum Mechanics

Physics today is really in a crisis; it is a quantum crisis. In future centuries, physicists will say they cannot figure out how mainstream physicists in the 20th and 21st centuries thought quantum mechanics works because what they wrote does not actually make sense. Detlef saw this crisis clearly, and instead of shrugging his shoulders, he tried to get to the bottom of quantum mechanics. He insisted that physical theories have to make sense. I have learned from him that it is important to openly criticize orthodox views where they deserve criticism.

In fact, Detlef was frank with criticism. He criticized my work and ideas a lot. But it was the kind of criticism that belongs in a fair debate. Debates have the purpose of finding the truth, not the purpose of defending a particular person. The debates with Detlef often led to agreement, or at least partial agreement; all people involved had learned something.

Detlef had known and collaborated with Shelly since the late 1970s and with Nino since the mid 1980s. The shared goal to get to the bottom of quantum mechanics led to their intensified joint collaboration on that subject in the late 1980s and to

a series of Dürr-Goldstein-Zanghì (DGZ) papers since about 1990, most of them collected later in Ref. [13]. They jointly arrived at the conclusion that, for quantum mechanics, Bohm's approach [5, 16] provided the crucial elements. They clarified Bohm's theory, defined it in a more coherent way than Bohm himself, and derived and motivated it in a more direct way; one could say they put it upside up. Detlef coined the name "Bohmian mechanics" in analogy to Newtonian mechanics. By combining the approaches of Bohm and Boltzmann, DGZ clarified the status of the Born rule in Bohmian mechanics as based on typicality [11], thereby removing the need for an approach to equilibrium through mixing.

Bohmian mechanics is not just a counter-example to some orthodox claims, it is the most serious theory of quantum mechanics that we have. By using a clear theory like this, we can clear up many confusing issues such as contextuality or the status of observables; we can see more clearly which general statements or proofs are relevant (such as the main theorem about POVMs); we can see more clearly what the problems are with various fields of quantum physics (such as quantum field theory) and take concrete steps to make progress on these problems. For example, quantum field theory involves the creation and annihilation of particles, which leads to the question of how to incorporate particle creation in Bohmian mechanics. I had the privilege to work with DGZ on the development of such extensions.

I also think it was important that DGZ took seriously the possibility that our relativistic space-time might have a preferred foliation (slicing) into spacelike hypersurfaces, even though this might at first seem against the spirit of relativity. Equations for an adaptation of Bohmian mechanics using a preferred foliation were worked out in 1999 [10] (actually, in a general-relativistic setting, although the published paper limits its discussion to the special-relativistic case). Since Bell's theorem [3] shows that nonlocality is inevitable, the nonlocality of Bohmian mechanics is good, not bad.

At the same time, Detlef was not at all dogmatic about Bohmian mechanics. For example, he supported (and contributed himself [1, 15] to) research about a competitor to Bohmian mechanics, the Ghirardi-Rimini-Weber theory of wave function collapse. It wasn't that he was dissatisfied with Bohmian mechanics and searched for alternatives. Rather, he felt that theories of quantum mechanics that actually make sense, that describe a coherent picture of reality, should be explored.

5 Mathematics

Some mathematicians tend to exaggerate the importance of proofs; they have a low opinion of the person who came up with a conjecture and a high opinion of the person who proved it, even if the proof did not introduce any particularly new ideas but arose mainly from persistence and diligence. Detlef was not like that. Although he was a mathematician and belonged to a math department, he had a deep appreciation for non-rigorous considerations. But that doesn't mean that he was imprecise or didn't

do proofs. On the contrary, he encouraged me and other students to provide rigorous studies of the physical concepts we considered, in part because that forces us to be careful and precise.

Correspondingly, he himself collaborated in a large number of such studies: for example, apart from the theorem mentioned initially [12], he contributed to mathematical proofs of global existence of Bohmian trajectories [4], that Bohmian trajectories become straight lines in the scattering limit [19], that the probability flux (and thus the Bohmian arrival places) agree with the scattering matrix [9], and to constructing the unitary time evolution of the quantized Dirac field in an external electromagnetic field [7]. He taught through his example how careful mathematical study should complement physical theory, and conversely how physics and other real-world applications give us orientation in a sea of mathematical possibilities. I believe that if you feed students with unmotivated mathematics, then it reduces their ability to think about mathematics.

Here is another thing that, although it is kind of obvious, I had never realized until Detlef pointed it out to me: classical electrodynamics is inconsistent. After all, the Lorentz force law requires that we evaluate the electromagnetic field $F_{\mu\nu}$ at the location of the charged particle, but exactly there it diverges as a consequence of the Maxwell equation. Textbooks on classical electrodynamics had never mentioned this basic fact. Going further, for solving the ultraviolet divergence problem of quantum electrodynamics, it might be a good start to try to solve it for classical electrodynamics. This has inspired Detlef's interest in Wheeler-Feynman electrodynamics [2] and in shape dynamics [14], two approaches that might offer ways out and that he published mathematical studies of.

6 Directions

Let me come back to Detlef's choice of research problems. It was motivated by the goal of understanding nature. It was not motivated by seeking applause, or trying to win (explicit or implicit) competitions against other scientists. People who see science as a competition between the smartest minds and wish to score highly have a tendency to choose problems that are popular (that many others have already worked on) and whose answer will be uncontroversial; they invest their time and energy in improving known results (providing, say, more accurate approximations and tighter bounds) and hope for recognition. Competition tends to remove the sense of intrinsic value: people who believe that the main reason for practicing the violin, or math, or swimming is to be better than others will also believe that nobody would play the violin, or think about math, or swim just for the joy of it. Competitions are usually competitions in doing something useless. Detlef's style was different. It was not that he didn't enjoy recognition, or that he didn't care what other people said. But on matters on which he could judge for himself, he did judge for himself, and didn't follow widespread opinion when he knew it was flawed. I think that research problems that will win competition and applause today but have little intrinsic value

will be of little interest to future generations of scientists. As I am writing this, I have to think of a quotation of Albert Einstein [17]:

> "Try not to become a man of success, but rather try to become a man of value."

This fits with what I described if the "man of success" is the scientist who primarily wants to gain applause for his work regardless of its actual value while the "man of value" primarily wants to do work of lasting scientific value regardless of how much applause it may win or not. I think that Detlef was a man of value.

References

1. A. Bassi, D. Dürr, G. Hinrichs, On the uniqueness of the equation for state-vector collapse. Phys. Rev. Lett. **111**, 210401 (2013). arXiv:1303.4284
2. G. Bauer, D.-A. Deckert, D. Dürr, On the existence of dynamics of Wheeler-Feynman electromagnetism. Zeitschrift für angewandte Mathematik und Physik **64**, 1087–1124 (2013). arXiv:1009.3103
3. J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, 1987)
4. K. Berndl, D. Dürr, S. Goldstein, G. Peruzzi, N. Zanghì, On the global existence of Bohmian mechanics. Commun. Math. Phys. **173**, 647–673 (1995). arXiv:quant-ph/9503013
5. D. Bohm, A suggested interpretation of the quantum theory in terms of "Hidden" variables I. Phys. Rev. **85**, 166–179 (1952)
6. L. Boltzmann, *Vorlesungen über Gastheorie*, Part I. (Barth, Leipzig, 1896) (Part II, 1898). English translation by S.G. Brush, *Lectures on Gas Theory* (University of California Press, Berkeley, 1964)
7. D.-A. Deckert, D. Dürr, F. Merkl, M. Schottenloher, Time evolution of the external field problem in QED. J. Math. Phys. **51**, 122301 (2010). arXiv:0906.0046
8. D. Dürr, S. Goldstein, J.L. Lebowitz, A mechanical model of Brownian motion. Commun. Math. Phys. **78**, 507–530 (1981)
9. D. Dürr, S. Goldstein, T. Moser, N. Zanghì, A microscopic derivation of the quantum mechanical formal scattering cross section. Commun. Math. Phys. **266**, 665–697 (2006)
10. D. Dürr, S. Goldstein, K. Münch-Berndl, N. Zanghì, Hypersurface Bohm–Dirac models. Phys. Rev. A **60**, 2729–2736 (1999) Reprinted in [13]. arXiv:quant-ph/9801070
11. D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the origin of absolute uncertainty. J. Stat. Phys. **67**, 843–907 (1992) Reprinted in [13]. arxiv:quant-ph/0308039
12. D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the role of operators as observables in quantum theory. J. Stat. Phys. **116**, 959–1055 (2004) Reprinted in [13]. arxiv:quant-ph/0308038
13. D. Dürr, S. Goldstein, N. Zanghì, *Quantum Physics Without Quantum Philosophy* (Springer, Heidelberg, 2013)
14. D. Dürr, S. Goldstein, N. Zanghì, Quantum motion on shape space and the gauge dependent emergence of dynamics and probability in absolute space and time. J. Stat. Phys. **180**, 92–134 (2020). arxiv:1808.06844
15. D. Dürr, G. Hinrichs, M. Kolb, On a stochastic Trotter formula with application to spontaneous localization models. J. Stat. Phys. **143**, 1096–1119 (2011). arxiv:1103.4018
16. D. Dürr, S. Teufel, *Bohmian Mechanics* (Springer, Heidelberg, 2009)
17. A. Einstein, in *An Interview to LIFE Magazine*. Published 2 May 1955. http://quoteinvestigator.com/2017/11/20/value/

18. S. Goldstein, J.L. Lebowitz, R. Tumulka, N. Zanghì, Gibbs and Boltzmann entropy in classical and quantum mechanics, in *Statistical Mechanics and Scientific Explanation*, ed. by V. Allori (World Scientific, Singapore, 2020), pp. 519–581. arxiv:1903.11870
19. S. Römer, D. Dürr, T. Moser, Asymptotic behavior of Bohmian trajectories in scattering situations. J. Phys. A: Math. Gen. **38**, 8421–8443 (2005). arxiv:math-ph/0505074
20. R. Tumulka, *Foundations of Quantum Mechanics* (Springer, Heidelberg, 2022)

Detlef Dürr's Path from Mechanics of the Brownian Motion to the Mechanics of the Quantum World: a Personal Point of View

Sergio Albeverio

Dedicated to the memory of a wonderful friend, an outstanding scientist and natural philosopher.

Abstract Some elements of the early work of Detlef Dürr are exposed. They involve a mathematical treatment of a mechanical model of the physical Brownian motion and other stochastic processes, as well as topics of statistical mechanics and stochastic analysis. Some comments are given on further developments, especially connected with Detlef's fundamental work on Bohmian mechanics and quantum mechanics.

1 Introduction

As a small tribute to Detlef Dürr I shall first present the early stage of his scientific development, perhaps as a complement to the presentation in this book by specialists and Detlef's coworkers, who were directly involved in the beautiful joint venture to develop Bohmian mechanics as a full alternative of standard quantum mechanics (in its "Copenhagen interpretation").

Detlef attended one of my earliest courses in Dirichlet forms and mathematical physics at the University of Bielefeld around 1977–1978. He would come by train from Münster, where he was working on his Ph.D thesis in physics. He impressed me right away for his strong motivations, asking sharp questions ranging from mathematics to physics. What stroke me particularly was his strong realistic view of the physical phenomena, his passion to understand how things are really made and work.

S. Albeverio (✉)
Institute for Applied Mathematics and Hausdorff Center for Mathematics, University of Bonn, Bonn, Germany
e-mail: albeverio@iam.uni-bonn.de

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_2

And he was ready to learn and use all kind of mathematical methods necessary for this understanding.
Soon a strong friendship between us developed and this also extended to our respective wives, Veronika and Solvejg. I remember well how we were sitting a summer day in the mini-garden adjoining the small apartment (full of different types of guitars) Veronika and Detlef were renting in Münster. We were discussing about almost everything, from Detlef's growing up at the sea side in the northern part of Germany, about his family, especially his brother (with whom he shared his passion for music)—about politics (in particular sharing our concerns how the "Berufsverbot" was spreading like a cold wind at that time in West Germany) and about philosophy (e.g. the presocratics, that Detlef liked particularly—I will come back to this theme later on). And of course we were talking about music (we both loved true folk singers like, e.g., Woody Guthrie)—Detlef was a very good guitar player and had a vision from insight. In fact, he had been actively playing with his life-long friend Peter Finger (who later became an excellent professional guitar player) and actively participating with him at that time in various folk song meetings (Detlef told me that for some time he was strongly split between physics and music—and as we know all life long he wholeheartedly and passionately maintained both activities). Detlef gave us as a present the beautiful instrumental record he had just produced with Peter. After our meeting in Münster I felt that something very special had happened, a kind of rare strong resonance between us that would last all life long.

Detlef's Ph.D thesis in Münster (in 1978) consisted in a study of the Onsager-Machlup function to determine the most probable paths of diffusion processes [BDS77] on which I will comment further in Sect. 2. After the thesis, Detlef started realizing his strong wish to go deeper in the comprehension of basic physical phenomena described by classical mechanics, kinetic theory and statistical mechanics, using all possible mathematical methods, that he was eager to learn and deepen.

The impression I had of Detlef already at that time was of a "natural philosopher", with a strong physical sense—this impression was confirmed over the years. I admired this attitude of him—in this he reminded me strongly of my own teachers at the ETHZ, Markus Fierz and Res Jost. I was happy that Detlef would find some stimulation for his research through the course he attended in Bielefeld and we remained in strong contact. And when I officially moved to the Ruhr-University in Bochum in 1979 he joined me as an assistant. For some years we would also commute together regularly between Bielefeld and Bochum, to attend BiBoS[1]-Seminars. He was very active in participating and often leading those seminars. Moreover he was interacting with several scientists visiting Bielefeld or Bochum, often initiating collaborations with them. Let me already mention some names in this connection like Paola Calderoni (Rome/Bremen), Rodolfo Figari (Naples), Shigeo Kusuoka (Kyoto/Tokyo), Danilo

[1] BiBoS stood first for Bielefeld-Bochum-Stochastics and was a Research Center founded by Philippe Blanchard, Ludwig Streit and myself, located in Bielefeld and Bochum with special funding through the Volkswagenstiftung. When I moved from Bochum to Bonn in 1997 "Bo" came to stand for Bonn.

Merlini (EPFL/Bochum/Locarno), Alessandro Teta (L'Aquila/Roma). I will come back to these collaborations in Sect. 2.

Starting already in 1979 Detlef was also staying at the Department of Physics of Rutgers University in the group of the coworkers of Joel Lebowitz. The collaboration of Detlef with Joel and his coworkers, in particular Sheldon Goldstein, also formed the basis for his Habilitation thesis in mathematics at the Ruhr-University in Bochum in 1983. I will come to say something on this in the next section.

After that, I will resume my personal account on the further evolution of Detlef's life and scientific production, before and after his nomination to a professorship in Mathematics at the Ludwig-Maximilian University of Munich in 1989.

2 Detlef's Work on Mechanical Models of Brownian Motion and Stochastic Processes

In the work [BDS77] done at the "Institute for Theoretical Physics I" of the University of Münster, the meaning of the Onsager-Machlup (OM) function associated with a diffusion process, introduced in physical terms by L. Onsager and S. Machlup in 1953, is investigated from a foundational point of view. In fact the paper provides a pioneering mathematical derivation of a probability functional related to the OM function. It also shows how the OM-functional permits to express interesting functionals of diffusion processes. The problem and method are deepened in [BD78a] where other original mathematical results are proven.

The OM method is namely characterized as a mean to find via a Lagrangian the most probable tube around a differentiable function of a diffusion process. Its relation with the Girsanov formula is stressed and in the case of a diffusion with constant diffusion coefficient, the OM-functional is made explicit. Moreover a variational principle is formulated and a critical assessment of previous discussions for the case of non-constant coefficients is presented. A further paper in this direction is [BD78a] and relates the potential part of the OM-functional to the entropy production density of a "most probable path", pointing out the relation between probability and entropy. A further paper [BD78b] discusses the equivalence between different methods of modelling diffusion processes that received a lot of attention at that time especially in the physical literature. The authors here relate the discussion to an important mathematical result by Wong and Zakai [WZ65]. The discussion was concluded and deepened in [BD78c]. In [DB79] an extension of the explicit formula for the OM-functional to non-linear drifts is proven. In the case of an anharmonic oscillator (quartic potential), the equation of motion for the "most probable path" is exhibited and the special case of a laser equation is analysed in detail. Let us stress that although much of this work has become nowadays a kind of implicit knowledge

in the probabilistic community, the influence of Detlef's work is acknowledged in most influential works, see for a sample e.g. [TW81,FK82,DZ91,Zei89,Aya+21].[2]

The strong will and determination of Detlef to understand more of the physical world, in particular the connection between classical and statistical mechanics methods, lead him in 1979 to go to the Rutgers University to take up a postdoctoral position in the group of Joel Lebowitz.[3] There indeed Detlef worked on the very important paper [DGL81], in collaboration with Shelly Goldstein and Joel Lebowitz. This opened a new line of research that also later lead to his Habilitationsschrift at Bochum University presented in 1983.

Let me describe first the content of this Habilitationsschrift and then briefly discuss the scientific work on which it was based. The "Habilitationsschrift" of Detlef is entitled "Dynamische Modelle für Brown'sche Bewegungen". It starts out with an introductory part (that I call here part I) consisting of the following sections: 0. Einführung, 1. Historisches, 2. Dynamische Modelle und die heutige Situation, 3. Ausblick. This part I consists of 25 densely written pages, that by themselves would deserve a separate publication.[4] It introduces the subject, its history and meaning, and discusses colloquially but very clearly the main results, and presents a beautiful outlook to further developments.

Section 0 recalls in compact but clear terms Brownian motion as a physical phenomenon and its mathematical correlate, namely Norbert Wiener (stochastic) process. Section 1 shortly presents the history of the observation of Brownian motion as a physical phenomenon, its relation with the heat equation established by A. Einstein (1905) and M. Smoluchowski (1906), and with the introduction in 1930 by Ornstein and Uhlenbeck of the physical stochastic process in space and velocity, in the style of P. Langevin stochastic equations proposal of 1911. Detlef refers to the "beautiful book" [Nel67] for this part, but also presents an original point of his on how to derive mathematically Einstein-Smoluchowski results starting from classical mechanical dynamical models. This is actually the program implemented by in the paper on which the Schrift is based. This is deepened in Sect. 2, where main results of the

[2] The methods discussed by Detlef and coworkers involve critical points of Lagrangians and have also influenced work in other areas, e.g. [Alb+09,AM11,AMB17,DFT04].

[3] I have found a moving and at the same time funny and optimistic letter from Detlef and Veronika sent to me on July 1st, 1979 from New Bunswick, containing in Detlef's typical terms a half-serious, half-critical but smiling tone a sort of report on their first month in the USA (of course accompanied by "Pauli", their small dog). From the letter is already visible that despite a certain difficulty to adopt to a new country and style of discussion, Detlef is already "taking foot" in the new surrounding, in which soon he would start a strong, intensive collaboration. Then he remarks that after having been already twice in New York, he found out that guitars are at least half price cheaper there than in Europe. "So kann unser USA-Aufenthalt gar nicht sinnlos sein." Then he also speaks about his wish of learning more, after having discussed with Joel, with whom he would hope very much to be able to work with, and for this he wants to get more familiarity with statistical mechanics and ergodic theory. And indeed he did become very familiar also with those areas, as documented by his successive brilliant work in these domains.

[4] The Habilitationsschrift—following the rules of the German university system—is deposited at the central library of Bochum University. As such is not easily accessible outside the German library system.

Schrift (and the papers in part II of the thesis) are presented. Beyond this Detlef also gives very useful insight about work by other authors e.g. [Spi69,Hol71]. In Sect. 3 he also gives important comments relating the various limit models obtained by the suitable scaling procedures, starting from mechanical models.

Let me now describe in some detail some of the papers in part II of the Habilitationsschrift. In the first paper [DGL81] Detlef and his coworkers consider a system consisting of one large heavy spherical particle of mass M and an infinite number of light particles of mass $0 < m \ll M$, all moving in $\mathbb{R}^3$. Initially the light particles form a Poisson point process (ideal gas) with density $m^{-1/2}$, with initial velocities that are independent and have a distribution with density $f_m(v) = m^{3/2} f_1(m^{1/2}v)$, f_1 being any probability density on $\mathbb{R}^3$ having finite moments up to order 4 and satisfying the symmetry condition $\int \|v\| v f_1(v) dv = 0$. The dynamics by assumption is provided by the elastic collisions between the point particles and the heavy particle. The authors obtained the beautiful result that in the limit $M/m \to \infty$ the joint distributions of position and velocity of the heavy particle converge weakly to those of the Ornstein-Uhlenbeck position and velocity process, with parameters determined by M, v and the first three moments of f_1. In this sense this constitutes a truly mechanical model for the physical (and mathematical) Brownian motion process. This is the first extension to 3 (or arbitrary many) dimensions of the pioneering work by Holley [Hol71] that was limited, by the very method it used, to the one-dimensional case. The difficulty of this extension is due to several factors, first of all the necessity of mastering the possibility of recollisions, depending on the geometric shape of the heavy particle.

The other papers in part II of Detlef's Habilitationsschrift are [DGL83a,DGL83b, Dür82]. In [DGL83b] a mechanical model for the Brownian motion of a convex body is discussed and the results of [DGL81] are extended in a natural (but by no means immediate) way to the case where the geometric shape of the heavy particle is that of a convex body. The extension is by no means immediate, because one has to take care of rotations of the body and complicated collisions with gas particles in absence of spherical symmetry of the body.

In [Dür82,DGL83a] the problem of recollisions is discussed in a general setting, together with related literature, including the one-dimensional equal masses case, e.g. [Har65,Spi69], and the case of a 2-dimensional periodic Lorentz gas [BS81]. References are also given to work in preparation by Detlef and coworkers, concerning relations with the central limit theorem in the case of dependent variables [Dür85,DG86]. Another topic discussed in [DGL85,DGL83a,DNZ87,DGL87], where the Landau model with Poisson distributed "soft-scatterer" with finite range isotropic smooth potentials is investigated and weak convergence under suitable scaling is obtained in the 2-dimensional case (the case of dimension $d \geq 3$ discussed by [KP80] is also mentioned). Finally the semi infinite case published in [CDK89] is also described. Many useful estimation infinite illustrations of methods and suggestions about treating other cases are provided in these very rich and clear papers and in the schrift.

The series of works collected in the beautiful Habilitation of Detlef in Bochum '83 stimulated much further work also by other authors. To just quote an example,

[DGL81] has 114 quotations on Semantic Scholar. For these let me just mention a few that in one way or the other I could follow personally, since they were connected with Detlef working in BiBoS and connecting there with other coworkers of mine e.g. Paola Calderoni[5] and Shigeo Kusuoka, who were visiting me from Rome and resp. Kyoto/Tokyo, worked with Detlef in [CDK89], extending the mechanical model of Brownian motion described in [DGL81] to a half-space. Shigeo Kusuoka had a PhD student, Song Liang, with whom he worked further on mechanical models of stochastic processes [KL08] and for models of different types of particles [KL10, KL12]. Song Liang also wrote on her own several important extensions of [DGL81], e.g. [Lia14a] (also related to the convex body case treated by Detlef in [DGL83b]). Even more recently we have the further important work by Song Liang [KL10,KL12, Lia13a,Lia13b,Lia14b,Lia18].[6]

3 Detlef's Work on the Mathematical and Physical Foundations of Bohmian Mechanics, and Some Closing Remarks

In '85, when Detlef was working in Bielefeld in the framework of the BiBoS Center, a young Italian postdoctoral student, Nino Zanghì, came along. In discussing with him I learned that he had been working on various problems of quantum mechanics and had developed a strong interest in looking for other approaches to the quantum world than the orthodox formulation of quantum mechanics (in particular he was reading work by authors like Norbert Wiener, Giacomo della Riccia and Takeyuki Hida, on the relation between classical phase space, Brownian motion and quantum theory). Myself, together with Raphael Høegh-Krohn, had been looking with interest into Ed Nelson's development of stochastic mechanics as an alternative vision to orthodox quantum mechanics and in fact at that time much activity, involving also Philippe Blanchard and Ludwig Streit in Bielefeld was developed inspired by this direction of investigations.

This was also enhanced by a group of enthusiastic friends from Ticino and more generally Insubria—the biologist Gabriele Losa, the physicists Uberto Cattaneo, Danilo Merlini and Giulio Casati (see also [ACM85,Alb+90,ACM95]) (who organized seminars and conferences in Ascona/Locarno with the participation of Ed, Detlef and Nino, see footnote 6 below). In this sense it was more than natural that I would strongly suggest to both Detlef and Nino that they should meet and discuss

[5] Paola Calderoni was an excellent young postdoctoral student visiting BiBoS from Rome. Detlef and Paola became close friends and collaborated in two papers with Detlef [CDK89,CD89]. When she died tragically in a car accident in Rome, Detlef was terribly shocked—he attended the funeral and stood for a while in Rome to be close to her family. This is one of his spontaneous manifestations of generosity that was characteristic of him as a wonderful person.

[6] Also roughly at the time of the Habilitationsschrift Detlef worked on several other problems of classical and statistical mechanics, let me mention [DP82], on the vortex flow, and [ADM83], on the independence of free energy from crystalline boundary conditions in plasma model.

together their points of view on quantum mechanics, since they had such common interests and orientations. In fact I was very happy to see how from their conversations (that sometimes seemed as debates—so energetic both of them were in defending their evolving points of view—)a strong life-long collaboration was developing.

The present book testifies how the cooperation between Detlef, Nino, Shelly Goldstein (with whom Detlef was already cooperating since his stay in Rutgers University) and coworkers of them growing over the years led to a full-fledged, coherent theory, both from the physical and mathematical points of view, of "Bohmian mechanics", as an alternative to orthodox, non-relativistic quantum mechanics.[7] I refer to other contributions in this book for systematic expositions and references to Detlef's fundamental contributions in the construction and elaboration of this theory. I will limit myself here to shortly explain my position. I consider that the theory initiated by Bohm (having inspired roots in work by De Broglie) in the beautiful mathematical and physical presentation and interpretation given by Detlef and coworkers, relating also to the physical inspiration from A. Einstein and E. Schrödinger, should be considered as a true alternative of other expositions of the essential context of quantum mechanics and its physical and epistemological interpretation (let me refer to Detlef's book and references therein).[8]

This said, the reader might ask the question about why I did not participate more directly in these developments. The main answer I can provide is that I was concentrating on mathematical problems arising from quantum mechanics itself—especially from its relativistic counterpart, namely quantum field theory, and I was postponing foundational and interpretational issues, taking a sort of agnostic position. But whenever we would meet with Detlef we would discuss intensively and plan future collaborations. Among other things, let me mention Detlef's interest in the work on mathematical Feynman path integrals, that we thought of applying e.g. around ideas of Feynman–Wheeler absorber theory, or of making direct use of them (in physical real time) as a constructive tool. Also I was following closely and with great interest the work of Detlef and coworkers towards relativistic quantum physics and quantum field theory. I also worked in these areas but whereas Detlef was concentrated primarily on a physical and foundational issue, I was mainly absorbed by developing mathematical methods that could in principle help to cope with the problem of divergence in the case of fields.

Let me also mention that when in 1989 Detlef was nominated to a permanent position at the Ludwig Maximilian University of Munich, he invited me to give a Doctorate Course on the theory of Dirichlet forms in infinite dimensions and its

[7] Detlef and Nino gave joint lectures in particular at the 3 Ascona/Locarno Conferences [ACM85, Alb+90, ACM95].

[8] In fact Bohmian mechanics also furnished motivations and momentum to raising and studying certain questions that had not been asked before in orthodox quantum mechanics. Let me mention e.g. "the scattering into cones" (see the work [Dau+97] and its Nelson's mechanical correlate in [Car85, PU04] and references therein); also let me quote developments connected with Schrödinger's analogy between quantum mechanics and Euclidean mechanics, in the sense of J.C. Zambrini (see, e.g [Zam87]); see also for further developments and connections with variational principles in quantum mechanics e.g. [AVU17, AVU23].

applications in quantum field theory. I accepted with great pleasure, we discussed a lot and at that time planned several collaborations in this direction, particularly on looking at canonical quantum mechanical objects both for particles and fields, and as such potentially fitting into the Bohmian view that Detlef and coworkers were actively and very successfully developing. Although our plans did not really materialize in direct collaborations, the mutual interest in each-other's progress of our insights was again and again renewed whenever we met.

The investigation in a similar spirit of the deeper structure lying below effective theories, that includes the problem of mastering the divergences in quantum field theory, has now to be left to the former students and scientists inspired by Detlef's work. Detlef's heritage is indeed very strong and we can expect for relativistic Bohmian mechanics and quantum field corresponding beautiful developments as we experienced with non-relativistic Bohmian mechanics. Elements of this are already in [Dür+04], see also e.g. [TT21] and references therein. I cannot close without mentioning another side of Detlef's strong personality and his search for truth, the strong wish to share with others this search. It is connected with his being in my eyes a "natural philosopher".[9]

The Weihnachtsvorlesungen [Dür05,Dür06,Dür07,Dür11,Dür12,Dür13,Dür15] that he was giving every year from 2006 to his retirement in 2016, which I unfortunately did not manage to attend, but read online with great interest, are for me impressive expressions of the reflections of a fully engaged true master in "philosophy of nature". In these conferences Detlef presents his reflections on the thoughts of the Presocratics-especially Parmenides and Heraclitus-confronting them with the evolution of our thinking about Nature. He also forcefully supports Schiller's view of what University should mean, criticizing then aspects of the way in which our present day's society is evolving. These lectures constitute nowadays rare examples of a genre with a noble tradition.

Let me close by stating that I always considered Detlef as a wise, close, dear, generous friend, a person of independent thinking, free and consequent in his expressions, with a critical view but also with a very empathic smile. We all miss him, but he will live further by his work and the strong inspiration he gave to all he would meet, in person or by the written word.

Acknowledgements I heartedly thank the organizers of the Conference in Memory of Detlef. Due to a sickness I caught at that time I was unfortunately unable to attend, but I am very grateful to the editors, and in particular to Roderich Tumulka, that they allowed me to get enough time to write this contribution. I also thank very warmly Elgi Orozi, who has been helping so much in setting up the text for publication. At the moment of handling in the final version of the text I was lying seriously ill at the Hospital, and it is only thanks to the generous help of Mielikki (my daughter) and my friends Rodolfo Figari and Alessandro Teta that my sketchy notes on parts of Sect. 2 and 3 could finally gathered and transformed, also with the help of Elgi, into a readable text. Let me express my deep gratitude and thank them all from the bottom of my heart!

[9] I hope some time to be able to write more on these aspects of Detlef's production, to give them visibility to a larger audience.

References

[BDS77] A. Bach, D. Dürr, B. Stawicki, Functionals of paths of a diffusion process and the Onsager-Machlup function. Zeitschrift für Physik. B. Condensed Matter **26**(2), 191–193 (1977). ISSN: 0722-3277. https://doi.org/10.1007/BF01325272

[BD78a] A. Bach, D. Dürr, Entropy density in function space and the Onsager-Machlup function. Phys. Lett. A **69**(4), 244–246 (1978). https://doi.org/10.1016/0375-9601(78)90174-3

[BD78b] A. Bach, D. Dürr, On the equivalence between different methods of modelling diffusion processes. Phys. Lett. A **66**(2), 79–82 (1978). ISSN: 0375-9601. https://doi.org/10.1016/0375-9601(78)90001-4

[WZ65] E. Wong, M. Zakai, On the convergence of ordinary integrals to stochastic integrals. Ann. Math. Stat. **36**(5), 1560–1564 (1965). https://doi.org/10.1214/aoms/1177699916

[BD78c] A. Bach, D. Dürr, Structural modeling of diffusion processes. Zeitschrift für Physik B Condensed Matter **29**, 265–272 (1978). https://doi.org/10.1007/BF01321191

[DB79] D. Dürr, A. Bach, Application of the Onsager-Machlup function to nonlinear diffusion processes. Zeitschrift für Physik B Condensed Matter and Quanta **32**(4), 413–417 (1979). https://doi.org/10.1007/bf02109471

[TW81] Y. Takahashi, S.Watanabe, The probability functionals (Onsager-Machlup functions) of diffusion processes. *Lecture Notes in Mathematics* (Springer, Berlin, Heidelberg, 1981), pp. 433–463. https://doi.org/10.1007/bfb0088735

[FK82] T. Fujita, S. Kotani, The Onsager-Machlup function for diffusion processes. Kyoto J. Math. **22**(1) (1982). https://doi.org/10.1215/kjm/1250521863

[DZ91] A. Dembo, O. Zeitouni, Onsager-Machlup functionals and maximum a posteriori estimation for a class of non-gaussian random fields. J. Multivar. Anal. **36**(2), 243–262 (1991). https://doi.org/10.1016/0047-259x(91)90060-f

[Zei89] O. Zeitouni, On the Onsager-Machlup functional of diffusion processes around non-C2-curves. Ann. Probab. **17**(3), 1037–1054 (1989). https://doi.org/10.1214/aop/1176991255

[Aya+21] B. Ayanbayev, I. Klebanov, H.C. Lie, T.J. Sullivan, Γ- convergence of Onsager-Machlup functionals. Part I: with applications to maximum a posteriori estimation in Bayesian inverse problems (2021). https://doi.org/10.48550/ARXIV.2108.04597

[Alb+09] S. Albeverio, Y. Kondratiev, Y. Kozitsky, M. Röckner, *The Statistical Mechanics of Quantum Lattice Systems* (EMS Press, 2009). https://doi.org/10.4171/070

[AM11] S. Albeverio, S. Mazzucchi, Path integral: mathematical aspects. Scholarpedia **6**(1), 8832 (2011). https://doi.org/10.4249/scholarpedia.8832

[AMB17] S. Albeverio, S. Mazzucchi, Z. Brzezniak, Probabilistic integrals: mathematical aspects. Scholarpedia **12**(5), 10429 (2017). https://doi.org/10.4249/scholarpedia.10429

[DFT04] D. Dürr, R. Figari and A. Teta, Decoherence in a two particle model. J. Math. Phys. **45**, 129–1309 (2004)

[DGL81] D. Dürr, S. Goldstein, J.L. Lebowitz, A mechanical model of Brownian motion. Commun. Math. Phys. **78**(4), 507–530 (1981). ISSN: 0010-3616

[Nel67] E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton University Press, 1967). https://doi.org/10.2307/j.ctv15r57jg

[Spi69] F. Spitzer, Uniform motion with elastic collision of an infinite particle system. J. Math. Mech. **18**(10), 973–989 (1969). ISSN: 00959057, 19435274. http://www.jstor.org/stable/24902097

[Hol71] R. Holley, The motion of a heavy particle in an infinite one dimensional gas of hard spheres. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **17**(3), 181–219 (1971). https://doi.org/10.1007/bf00536757

[DGL83a] D. Dürr, S. Goldstein, J.L. Lebowitz, Stochastic processes originating in deterministic microscopic dynamics. J. Stat. Phys. **30**(2), 519–526 (1983). https://doi.org/10.1007/bf01012325

[DNZ87] D. Dürr, V. Naroditsky and N. Zanghi, On the hydrodynamic motion of a test particle in a random harmonic chain, Ann. Phys. **178**, 74-88 (1987)

[DGL83b] D. Dürr, S. Goldstein, J.L. Lebowitz, A mechanical model for the Brownian motion of a convex body. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **62**(4), 427–448 (1983). ISSN: 0044-3719. https://doi.org/10.1007/BF00534196

[Dür82] D. Dürr, All that Brownian motion, in *Stochastic Processes in Quantum Theory and Statistical Physics*, ed. by S. Albeverio, Ph. Combe, M. Sirugue-Collin (Springer, Berlin, Heidelberg, 1982), pp. 119–135. ISBN: 978-3-540-39546-1. https://doi.org/10.1007/3-540-11956-6_114

[Har65] T.E. Harris, Diffusion with "collisions" between particles. J. Appl. Probab. **2**(2), 323–338 (1965). https://doi.org/10.2307/3212197

[BS81] L.A. Bunimovich, Ya..G.. Sinai, Statistical properties of Lorentz gas with periodic configuration of scatterers. Commun. Math. Phys. **78**(4), 479–497 (1981). https://doi.org/10.1007/bf02046760

[Dür85] D. Dürr, Diffusions and central limit theorems, in *Stochastic Aspects of Classical and Quantum Systems (Marseille, 1983)*, vol. 1109. Lecture Notes in Math (Springer, Berlin, 1985), pp. 55–67. https://doi.org/10.1007/BFb0101537

[DG86] D. Dürr, S. Goldstein, Remarks on the central limit theorem for weakly dependent random variables, in *Stochastic Processes – Mathematics and Physics*, ed. by S. Albeverio, P. Blanchard, L. Streit (Springer, Berlin, Heidelberg, 1986), pp. 104–118. https://doi.org/10.1007/bfb0080211

[DGL85] D. Dürr, S. Goldstein, J.L. Lebowitz, Asymptotics of particle trajectories in infinite one-dimensional systems with collisions. Commun. Pure Appl. Math. **38**(5), 573–597 (1985). ISSN: 0010-3640. https://doi.org/10.1002/cpa.3160380508

[DGL87] D. Dürr, S. Goldstein, J.L. Lebowitz, Asymptotic motion of a classical particle in a random potential in two dimensions: Landau model. Commun. Math. Phys. **113**(2), 209–230 (1987). https://doi.org/10.1007/bf01223512

[KP80] H. Kesten, G.C. Papanicolaou, A limit theorem for stochastic acceleration. Commun. Math. Phys. **78**(1), 19–63 (1980). https://doi.org/10.1007/bf01941968

[CDK89] P. Calderoni, D. Dürr, S. Kusuoka, A mechanical model of Brownian motion in half-space. J. Stat. Phys. **55**(3–4), 649–693 (1989). https://doi.org/10.1007/bf01041603

[CD89] P. Calderoni, D. Dürr, The Smoluchowski limit for a simple mechanical model. J. Stat. Phys. **55**(3–4), 695–738 (1989). https://doi.org/10.1007/bf01041604

[KL08] S. Kusuoka, S. Liang, A mechanical model of Markov processes, in *Proceedings of RIMS Workshop on Stochastic Analysis and Applications. RIMS K.oky.uroku Bessatsu, B6* (Research Institute for Mathematical Sciences (RIMS), Kyoto, 2008), pp. 167–176

[KL10] S. Kusuoka, S. Liang, A classical mechanical model of Brownian motion with plural particles. Rev. Math. Phys. J. Both Rev. Orig. Res. Papers Field Math. Phys. **22**(7), 733–838 (2010). ISSN: 0129-055X. https://doi.org/10.1142/S0129055X10004077

[KL12] S. Kusuoka, S. Liang, A classical mechanical model of Brownian motion with one particle coupled to a random wave field. Stoch. Anal. Appl. **30**(3), 493–528 (2012). https://doi.org/10.1080/07362994.2012.668444

[Lia14a] S. Liang, A mechanical model of Brownian motion with uniform motion area. The University of Tokyo. J. Math. Sci. **21**(2), 235–334 (2014). ISSN: 1340-5705

[Lia13a] S. Liang, Stochastic Hamiltonian equation with uniform motion area. Dyn. Syst. Appl. **22**, 557–590 (2013)

[Lia13b] S. Liang, A stochastic process with a Stochastic Hamiltonian equation phase and a Uniform Motion phase (Probability Symposium). RIMS Kokyuroku **1955**, 28–35 (2013)

[Lia14b] S. Liang, A mechanical model of Brownian motion with uniform motion area. J. Math. Sci.-Univer. Tokyo **21**, 235–334 (2014)

[Lia18] S. Liang, A mechanical model of Brownian motion for one massive particle including slow light particles. J. Stat. Phys. **170**(2), 286–350 (2018). ISSN: 0022-4715. https://doi.org/10.1007/s10955-017-1934-4

[ADM83] S. Albeverio, D. Merlini and D. Dürr, Remarks on the independence of the free energy from crystalline boundary conditions in the two-dimenisonal one-component plasma. J. Stat. Phys. **31**, 389–407 (1983)

[DP82] D. Dürr, and M. Pulvirenti, On the vortex flow in bounded domains. Commun. Math. Phys. **85**(2), 265–273 (1982)
[TT21] S. Teufel and R. Tumulka, Hamiltonian without ultraviolet divergence for quantum field theories, in Quantum Studies: Mathematics and Foundations **8**, 17–35 (2021)
[ACM85] S. Albeverio, G. Casati, D. Merlini, Stochastic processes in classical and quantum systems, in *Proceedings of the 1st Ascona-Como International Conference, Ascona, Switzerland*, ed. by S. Albeverio, G. Casati, D. Merlini. (Springer, Berlin, Heidelberg, 1985). https://doi.org/10.1007/3-540-17166-5
[Alb+90] S. Albeverio, G. Casati, U. Cattaneo, D. Merlini, R. Moresi, Stochastic processes, physics and geometry, in *Stochastic Processes, Physics and Geometry* (World Scientific, Ascona/Locarno, Switzerland, 1990). https://doi.org/10.1142/9789814541107
[ACM95] S. Albeverio, U. Cattaneo, D. Merlini, Stochastic processes, physics and geometry II, in *Stochastic Processes, Physics and Geometry II*. (World Scientific, Ascona/ Locarno, Switzerland, 1995). https://doi.org/10.1142/9789814533195
[Dau+97] M. Daumer, D. Dürr, S. Goldstein, N. Zanghì, On the quantum probability flux through surfaces. J. Stat. Phys. **88**(3/4), 967–977 (1997). https://doi.org/10.1023/b:joss.0000015181.86864.fb
[Car85] E.A. Carlen, Potential scattering in stochastic mechanics. Annales de l'Institut Henri Poincaré. Physique Th'eorique **42**(4), 407–428 (1985). ISSN: 0246-0211
[PU04] A. Posilicano and S. Ugolini, Asymptotic flux across hypersurfaces for diffusion processes, pp. 185–197 in Proc. Int. Conf. Stochastic Analysis and Appk., Hammamet 2001, Eds. S. Albeverio, A. Boutet de Monvel and H. Ouerdiane, Birkhäuser, (2004)
[Zam87] J.C. Zambrini, Euclidean quantum mechanics. Phys. Rev. A **35**(9), 3631–3649 (1987). https://doi.org/10.1103/physreva.35.3631
[AVU17] S. Albeverio, F.C. De Vecchi, S. Ugolini, Entropy chaos and Bose-Einstein condensation. J. Stat. Phys. **168**(3), 483–507 (2017). https://doi.org/10.1007/s10955-017-1820-0
[AVU23] S. Albeverio, F.C. De Vecchi and S. Ugolini, Some connections between stochastic mechanics, optimal control, and nonlinear Schrodinger equations, in *Mathematics going forward Collected mathematical brushstrokes*, eds. by J.M. Morel, B. Teissier, *Lecture Notes in Mathematics*, vol. 2313 (Springer, Cham, 2023), pp. 505-534
[Dür+04] D. Dürr, S. Goldstein, R. Tumulka, N. Zanghì, Bohmian mechanics and quantum field theory. Phys. Rev. Lett. **93**(9), 090402 (2004). https://doi.org/10.1103/physrevlett.93.090402
[Dür05] D. Dürr, Gedanken über das Höhlengleichnis von Plato. Weihnachtsvorlesung WS05/06 (2005)
[Dür06] D. Dürr, Wie kommt der Irrsinn zustande? Weihnachtsvorlesung WS06/07 (2006)
[Dür07] D. Dürr, Was heißt und zu welchem Ende studiert man Physik? Weihnachtsvorlesung WS07/08 (2007)
[Dür11] D. Dürr, Der Dialog. Weihnachtsvorlesung WS11/12 (2011)
[Dür12] D. Dür, Über die Genesis und über einige Bedeutungen des Höhlengleichnisses. Weihnachtsvorlesung WS12/13 (2012)
[Dür13] D. Dürr, Dialog zwischem dem Verstand und den Sinnen. Weihnachtsvorlesung WS13/14 (2013)
[Dür15] D. Dürr, In welcher Kulturepoche lebt Schrödingers Katze? Weihnachtsvorlesung WS15/16 (2015)

Detlef the Adventurer

Tim Maudlin

Abstract Detlef Dürr was a remarkable figure in many different ways. I recall some adventures we had with him in Abu Dhabi.

A volume of remembrance like this is largely devoted to scientific pieces on Detlef's work or on the topics he devoted himself to. It is altogether fitting and proper that we should do this. But it is also essential to memorialize Detlef's unique spirit. Those who knew him personally will never forget being with him, and for those who never had the fortune to spend time with him it is important to try to convey who he was as a human being.

Perhaps some of the stories—such as the rescue of Bertha the chicken, at substantial investment of care and expense (for a chicken), resulting in a devoted friendship—are fairly well known. Certainly, Detlef's role as Doktorvater to generations of students at the LMU will have a permanent impact on foundations of physics that goes beyond the lasting effect of his publications. His ability to guide and encourage and inspire dozens of students in an otherwise hostile academic environment is, to my knowledge, unparalleled in the field. He is irreplaceable.

Those stories—of his unstinting devotion to his students—are for them to tell. I never had the privilege of studying under him. Even the tale of Bertha I could only pass on second-hand, as told to me by Detlef himself. But there is an episode I can relate first-hand, which is not widely known and may perhaps give a small sense of Detlef the person to those who later pick up this volume.

In the fall of 2012, I spent half a semester teaching at NYU's campus in Abu Dhabi. This was not long after Abu Dhabi had opened as a gateway campus—to which students would directly apply for admission—in NYU's Global Network University scheme. The dedicated Abu Dhabi campus was just being designed, and so we were housed in a large apartment building downtown, with temporary classrooms set up a little way off. Relatively few of the Abu Dhabi permanent faculty had been hired,

T. Maudlin (✉)
Department of Philosophy, New York University, New York, NY, USA
e-mail: timmaudlin@johnbellinstitute.org

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_3

and teaching needs were met by visiting faculty from the Washington Square campus in New York.

Part of the arrangement for visiting faculty, in order to promote academic activity at the new institution, was the possibility of inviting a colleague to spend a week in Abu Dhabi. So I was presented with the question of who would be best to have around to spend time with, talking to and arguing with, and who would also create connections with other departments and generally light up both NYU Abu Dhabi and our own lives. And also, who would just be really excited and enthusiastic about the chance to come to such an exotic location. The obvious answer was: Detlef.

The visit was far more successful than I ever could have anticipated. At a purely academic level, Detlef was invited to talk to a mathematical physics class and was such a hit that long after I left—and with no impetus from me—he was invited back to speak and interact with the faculty there. Detlef just immediately won over, in the course of a few days, people in mathematical physics who previously didn't know him at all. There is not much I could say to convey how that happened that would add to the clarity and humor and passion that one finds in his published work and recorded lectures. But there is another aspect to Detlef's visit that would never be generally known except that I recount some of it here. That is the purely personal part, and I record bits of it for those who knew him but not this particular episode, and for those who never knew him but want to understand who we was, and mostly for myself just for the sheer pleasure of recalling it.

Detlef came with his daughter Anna, and as I mentioned they stayed for about a week. Detlef's attitude to the visit was already evident the day after they arrived. I had, myself, been somewhat concerned about going to Abu Dhabi: when I signed up for it I had never been anywhere in the region and really had no clear idea what to expect. It was *terra incognita* for me and my wife Vishnya and our son Maxwell, who had been taken from school in Princeton and was trying to keep up with classes remotely. From our NYU encampment in downtown Abu Dhabi we had ventured out to the Corniche and the huge malls and a few other places. But already just 12 h after having arrived, Detlef and Anna had been to places we hadn't seen. He loved the beach, and sought out the older areas to explore. Whereas I approached being in Abu Dhabi with a certain amount of trepidation, Detlef just threw himself into his visit from the first moment as an adventure.

As far as I can reconstruct, Detlef and Anna arrived on November 12. On the 14th he was presenting Bohmian mechanics to the undergraduates, opening their minds to physical possibilities that they would almost certainly never have been introduced to in any other place. Here he is in pedagogical mode:

Detlef Explaining

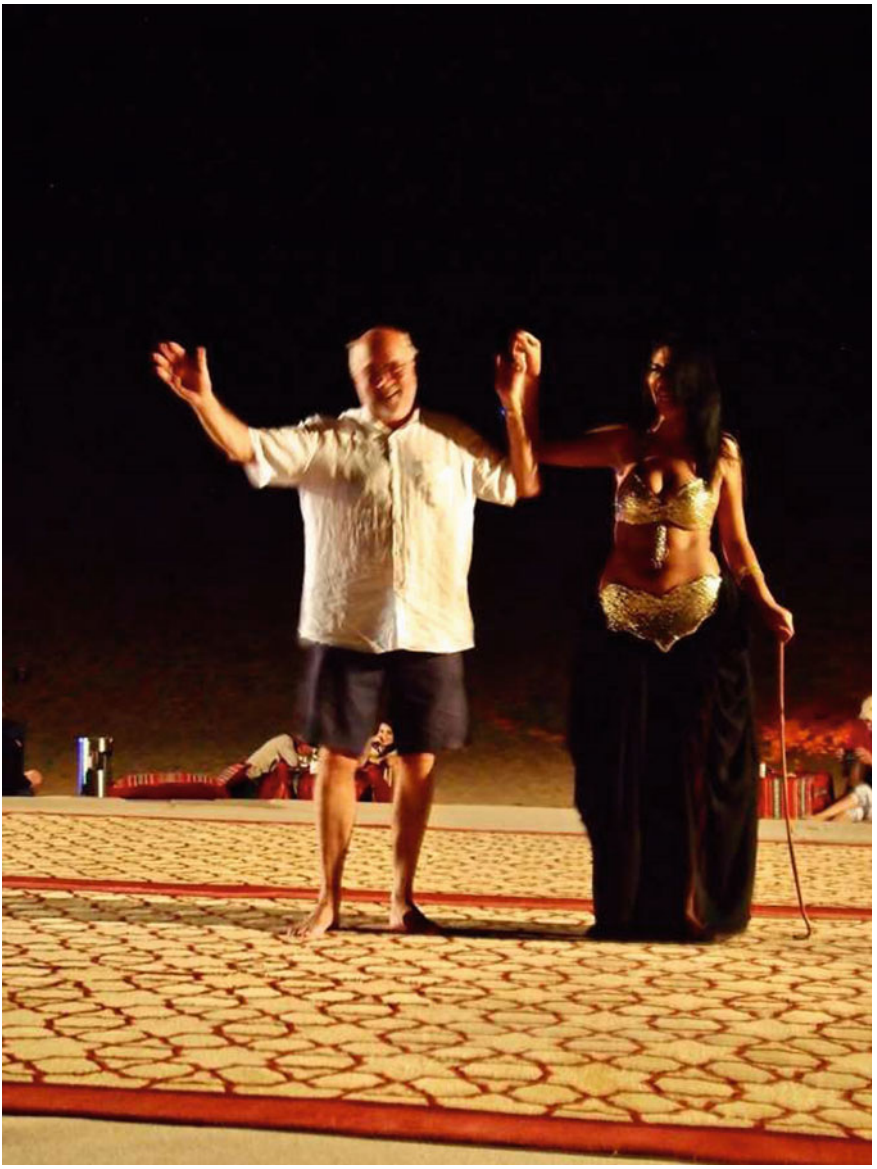
That, of course, is a picture similar to hundreds of others at thousands of other similar occasions.

But a couple of days later we decided to go to the desert to get a taste of the local customs and history. And from that experience there are some pictures which I hope will convey more of who Detlef was.

First, of course, he loved the sheer physical beauty of the place:

Detlef, Anna, Maxwell and me

Later, there was a belly dancer who would pick people out of the audience to dance with. Can it be entirely a mere co-incidence that of all the people in audience her eye settled on….Detlef? She just knew how he would rise to the occasion:



Detlef the Terpsichorist

Nor would Detlef (or Anna) miss a chance to try the hookah in the desert:

Detlef the Experimentalist

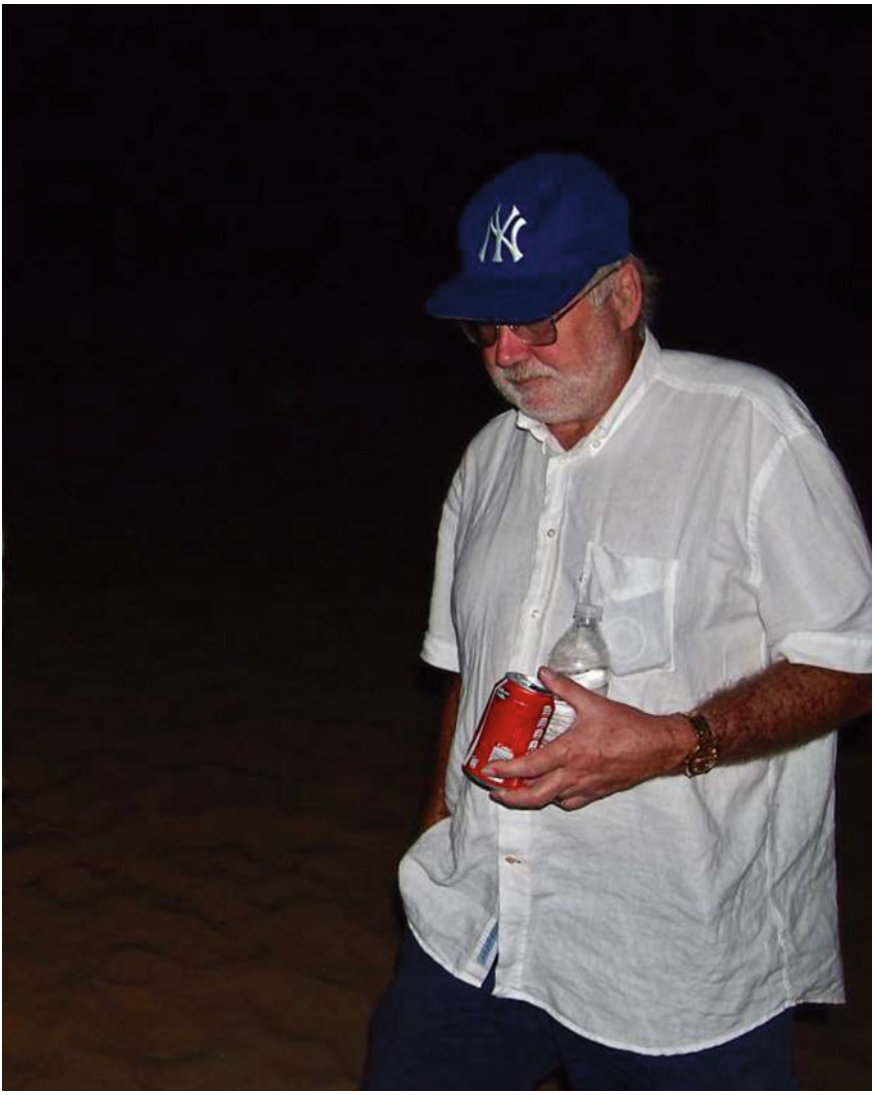
But most memorably, Detlef positively relished the chance to really immerse himself in the experience of being in the Empty Quarter. The hallmark of his approach to foundational questions in physics is asking "What it would be like if…..?". And there was no end of other contexts in which he would immerse himself in such questions…..

Vishnya, Detlef and Anna

The sense of playfulness—of joy in the variety and richness and perhaps absurdity of life—lies at the center of who Detlef was. That, and his boundless love and affection for his family, his friends, his colleagues and his students.

A picture, they say, is worth a thousand words. I hope that this little account, and moreso the pictures, are able to get across at least a small sense of who Detlef was for those who did not know him, and recall him more vividly to mind for those who did. A few more:

In the Desert



Detlef with his Thoughts

The Detlef Dürr Room, John Bell Institute, Hvar, Croatia

Foundations of Quantum Mechanics

Why Bohm and Only Bohm?

Jean Bricmont

> Particles move (Detlef's answer to a famous mathematical physicist with no interest in foundations of QM who was asking him what Bohm's theory was all about "in two words".)

Abstract It is often claimed that there are three "realist" versions of quantum mechanics: the de Broglie–Bohm theory or Bohmian mechanics, the spontaneous collapse theories and the many worlds interpretation. We will explain why the two latter proposals suffer from serious defects coming from their ontology (or lack thereof) and that the many worlds interpretation is unable to account for the statistics encoded in the Born rule. The de Broglie–Bohm theory, on the other hand, has no problem of ontology and accounts naturally for the Born rule.

1 A Misleading Problem: The Measurement One

The measurement problem is well known: at the end of an experiment where one measures the property of a particle that can take two values, the wave function (or quantum state) of the measuring device, or of the cat if we couple the device to the cat through a poison capsule, is (leaving aside normalization factors):

$$\Psi_{\text{cat alive}} + \Psi_{\text{cat dead}}.$$

And that cannot be a complete description of the cat, which is obviously either alive or dead but not both!

The way out of this problem from the point of view of ordinary quantum mechanics is to introduce the collapse postulate: when one looks at the cat, one sees whether

Dedicated to the memory of Detlef Dürr.

J. Bricmont (✉)
Institut de recherche en mathématiques et physique, Université Catholique de Louvain, chemin du Cyclotron 2, 1348 Louvain-la-Neuve, Belgium
e-mail: jean.bricmont@uclouvain.be

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_4

she is alive or dead and, depending on what one sees, one reduces the wave function of the cat (and of the particle that was measured and is thus coupled to the state of the cat) to either $\Psi_{\text{cat alive}}$ *or* $\Psi_{\text{cat dead}}$.

Since this is a *deus ex machina* from the point of view of the linear Schrödinger evolution, justifying it is often viewed as the main problem in foundations of quantum mechanics.

However there is a deeper problem: neither $\Psi_{\text{cat alive}}$ nor $\Psi_{\text{cat dead}}$ are cats: they are functions defined on a high dimensional space $\mathbb{R}^{3N}$ while cats are located in $\mathbb{R}^3$. And it is not clear what it means to say that $\Psi_{\text{cat alive}}$ or $\Psi_{\text{cat dead}}$ are descriptions of cats, let alone "complete descriptions" of them.[1]

What most people do is to mentally identify cats and wave functions of cats, which is illegitimate. So, the real problem is the meaning or the ontology one: what does the wave function mean outside of laboratories, and what does it say about what the world is made of?

2 An Intuitive Solution That Does Not Work

A simple and a priori attractive solution to this problem is the naïve statistical interpretation of quantum mechanics (and that is probably what is in the back of the minds of most of the "no worry about quantum mechanics" physicists): particles do have properties such as position, velocity, spin etc., but we cannot know or control them-we have only access to their wave function.

That object gives the statistical distribution of the values of those quantities (through the Born rule) over sets of particles having the same wave function. And, when we perform a measurement of a property of a given particle, we learn what that value is for that particle.

In that interpretation, the reduction of the wave function is no problem; we simply adjust our probabilities when we learn something new about the system.

And, if that worked, there would indeed be no reason to worry about the meaning of the wave function and we would have a decent meaning of that function outside of measurements. However, it cannot work, because of well-established theorems due to Bell [5] and to Kochen and Specker [25], but that unfortunately are not widely known among physicists.

Those theorems show that, if we assume that there exists a map v that assigns a value to each observable A corresponding to various properties that are simultaneously measurable according to ordinary quantum mechanics and that agrees with minimal quantum mechanical predictions concerning such observables, then one can deduce a contradiction.[2]

These theorems are called the "no hidden variables theorems".

[1] This idea is emphasized by Tumulka [34, Sect. 5.1].

[2] See Mermin [28] for pedagogical proofs and [11, 12] for more details.

Obviously, there cannot be a statistical distribution of maps that do not exist. In other words, what we call the statistical interpretation (individual quantum systems do have definite properties but we are only able to know their statistical distributions) does not work.

Note in passing that the ideas behind these no hidden variables theorems can be used to rule out the "decoherent histories" approach of Gell-Mann, Hartle, Griffiths and Omnès [17, 24, 31].

Indeed, as noticed by Goldstein [21] and Bassi and Ghirardi [4], this approach amounts to assigning simultaneous values to pairs of commuting observables; but if a series of such pairs is suitably chosen, a contradiction follows.

So, let us look for non obvious solutions.

3 Spontaneous Collapse Theories

These theories are modified versions of quantum mechanics, the first of which was introduced by Ghirardi, Rimini, and Weber [18] in which wave functions spontaneously collapse.[3]

To be precise, in that model, the wave function evolves according to the Schrödinger equation most of the time, but there is a set of spacetime points $(\mathbf{y}_i, t_i)$ chosen at random, such that the wave function $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N, t)$ for a system of N particles is multiplied at the chosen times t_i by a Gaussian function in the variable $\mathbf{x}_k$ (k chosen uniformly among $1, \ldots, N$), centered in space at the chosen space points $\mathbf{y}_i$.

The probability distribution of these random points is determined by the wave function of the system under consideration at the times when they occur, and is given by the familiar $|\Psi|^2$ distribution. This ensures that the predictions of the GRW theory will (almost) coincide with the usual ones.

The above-mentioned multiplication factors localize the wave function in space, and, for a system of many particles in a superposed state, effectively collapse the wave function onto one of the terms. Now, the trick is to choose the parameters of the theory so that spontaneous collapses are rare enough for a single or for a few particles in order to ensure that they do not lead to detectable deviations from the quantum predictions, but are frequent enough to ensure that a system composed of a large number of particles, say $N = 10^{23}$, will not stay in a superposed wave function for more than a split second.

Spontaneous collapse theories are not the same as ordinary quantum mechanics, since they lead to predictions that differ from the usual ones, even for systems made of a small number of particles. But the parameters of the theory are simply adjusted so as to avoid being refuted by present experiments, which is not exactly an appealing move.

[3] For reviews and further discussions of those theories, see Ghirardi [20], Ghirardi et al. [19], Allori et al. [2], Goldstein et al. [22].

Moreover there is the problem of making sense of a pure wave function ontology (even when the latter collapses, since the collapsed wave function is still just a function defined on a high-dimensional space).

Two solutions have been proposed to give a meaning to the GRW theory beyond the pure wave function ontology: the matter density ontology often denoted GRWm [2, 19], and the flash ontology denoted GRWf [6].

The matter density ontology associates a continuous matter density to the wave function of a system of N particles. For each $\mathbf{x} \in \mathbb{R}^3$, and $t \in \mathbb{R}$, one defines:

$$m(\mathbf{x}, t) = \sum_{i=1}^{N} m_i \int_{\mathbb{R}^{3N}} \delta(\mathbf{x} - \mathbf{x}_i) |\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N, t)|^2 d\mathbf{x}_1 \ldots d\mathbf{x}_N \ , \tag{1}$$

where $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N, t)$ is the usual wave function of the system at time t. This equation makes a connection between the wave function defined on the high-dimensional configuration space and an object, the matter density, existing in our familiar space $\mathbb{R}^3$.

In our three-dimensional world, there is just a continuous density of mass: no structure, no atoms, no molecules, etc., just an amount of "stuff", with high density in some places and low density elsewhere.

In the flash ontology, one has a world made only of spacetime points at the center of the Gaussian multipliers of the wave function that collapse it. No particles, no fields, nothing at all, except a "galaxy" of spacetime points, called "flashes".

Let us note that, if the God of the physicists was trying not to be malicious and if either the matter density or the flash ontologies are true, then He failed badly: indeed, it means that we were wrong all along when we "discovered" atoms, nuclei, electrons, etc., and that we are lying to schoolchildren when we tell them that matter is mostly void with a few pieces of matter (the atoms) here and there. Indeed, in the matter density ontology, matter is continuous after all, with higher and lower density in some places, and we have simply been fooled by this modified version of quantum mechanics into thinking that it is not.

On the other hand, if the flash ontology is true, then we have been fooled into thinking that there exists something most of the time (like atoms): if we take the visible universe since the Big Bang, it has contained only finitely many flashes. Since the flashes are all there is in that ontology, this means that, most of the time, the universe is just empty.

Of course, even if those ontologies are weird, we might be forced to accept one of them (after all the existence of atoms in empty space is also counterintuitive), if there were independent reasons for doing so, like a greater explanatory power or greater empirical adequacy.

But, and this is the most important point, empirical adequacy of any spontaneous collapse theory would mean that ordinary quantum mechanics is empirically wrong, since the predictions of both theories differ, at least in principle. So, if one found that a prediction of a spontaneous collapse theory is right, when it differs from ordinary quantum mechanics, it would be a major revolution in physics and we might be forced

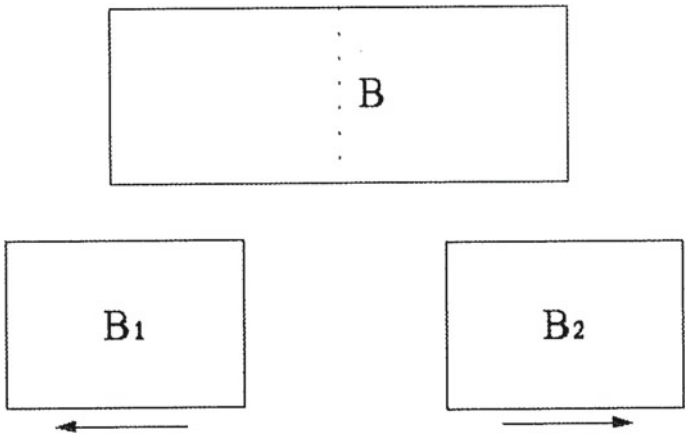


Fig. 1 Einstein's boxes. Reproduced with permission from Norsen [29]. Copyright 2005 American Association of Physics Teachers

to worry about those weird ontologies. But this hasn't happened yet and I would suggest not to worry about what to do after the revolution before that revolution has occurred.

To illustrate how odd the continuous matter and the flash ontologies are, compared to the particle ontology of the de Broglie–Bohm theory (to be discussed in Sect. 5), consider what happens, in those ontologies, with the thought experiment of Einstein's boxes: imagine a box containing just one particle that is cut in two parts; one part is sent to New York, the other to Tokyo [29], [26, Chap. 10].

The wave function of the particle is a superposition of a wave function located in box B_1 + a wave function located in box B_2. In ordinary quantum mechanics, when Alice in New York opens her half-box and sees the particle, the wave function collapses on the part of the wave function located in New York and, if she doesn't see it, the wave function collapses on the part located in Tokyo (Fig. 1).

In the de Broglie–Bohm theory, nothing surprising happens: the particle is in one of the half-boxes all along and is found where it is.

But in the GRWm theory, there is one-half of the matter density of a single particle in each half-box. When Alice couples her half-box with a detector of particles, the evolution of the wave function of the particle is coupled with a macroscopic object and many collapses occur quite rapidly, so that the matter density suddenly jumps from being one-half of the matter density of a single particle in each half-box to being the full matter density of a particle in that half-box and nothing in the other.

There is a nonlocal transfer of matter in the GRWm theory, while there is no such thing in the de Broglie–Bohm theory, and not even anything nonlocal when one deals with only one particle.

In the GRWf theory, there is simply nothing in either half-box, just a wave function traveling so to speak with the half-boxes. When Alice couples her half-box with

a detector of particles, the wave function of the particle becomes coupled with a macroscopic object, and there is suddenly a "galaxy of flashes" appearing (randomly) in that detector, either detecting the particle or not; we interpret the first possibility as meaning that the particle is in that half-box and the second one as meaning that the particle is in the other half-box.

Putting aside the weirdness of the ontologies, the spontaneous collapse theories are more nonlocal than the EPR-Bell result implies, since the latter concerns nonlocality for systems with at least two particles.

Here we have nonlocal effects or actions at a distance (Alice affects the physical situation in Tokyo by acting in New York) **even for one particle**!

In summary:

- Spontaneous collapse theories have unnatural ontologies.
- They are very ad hoc: parameters are chosen so as to avoid refutation and not on the basis of any evidence.[4]
- They can only be true if quantum mechanics itself is false.
- They are more nonlocal than they have to be.

4 The Many-Worlds Interpretation

This interpretation postulates that, when the proverbial cat (or any other macroscopic device) finds itself in a superposed state, then, instead of undergoing a collapse by fiat as in ordinary quantum mechanics, both terms simply continue to exist. But how can that be possible? We always see the cat alive *or* dead but not both! The short answer is that they both exist, but in different "worlds".

Hence, whenever an experiment leads to a macroscopic superposition, the world splits into two or more worlds, depending on the number of distinct macroscopic states produced by that experiment, one for each possible result.

Why do I always perceive only one of the results? It is simple: I, meaning my body, my brain (and thus also my consciousness) becomes entangled with the states of the cat, so there are two or more copies of me also, one seeing the dead cat in one world, another seeing the live cat in another world. And that, of course, is also true for everything else: every molecule in the entire world becomes copied twice (maybe not instantaneously, but that is another question).

In his original paper [16], Everett stressed that "*all* elements of a superposition are 'actual', none any more 'real' than the rest." Everett felt obliged to write this because "some correspondents" had written to him saying that, since we experience only one element of a superposition, we have only to assume the existence of that

[4] Moreover, recent results seem to provide evidence refuting spontaneous collapse theories: https://www.quantamagazine.org/physics-experiments-spell-doom-for-quantum-collapse-theory-20221020/.

unique element. This shows that some early readers of Everett were already baffled by the radical nature of the "many-worlds" proposal.

Putting aside the weirdness of this multiplication of "worlds", one should ask whether the many worlds scheme is coherent. Consider the Born rule. Suppose that the probabilities of having the cat alive or dead, as a result of an experiment, are $(\frac{1}{2}, \frac{1}{2})$. And suppose that I decide to repeat the same experiment successively many times, with different particles (and cats) but all having the same initial wave function.

After one experiment, there are two worlds, one with a dead cat and a copy of me seeing a dead cat and one with a live cat and a copy of me seeing a live cat. Since both copies of me are in the same state of mind as I was before the first experiment (after all, both copies are just copies of me!), each of them repeats that experiment.

Then, we have four worlds, one with two consecutive dead cats, one with two consecutive live cats and two with one dead cat and one live cat. "I" (by that I mean each copy of me in each of those four worlds) repeat the experiment again: we have now eight worlds, with one "history" of worlds with three dead cats, one history of worlds with three live cats, three histories of worlds with one live cats and two dead ones, three histories of worlds with one dead cat and two live ones.

Now, continue repeating that experiment: for every possible sequence of outcomes, there will be some of my "descendants" (i.e. copies of me, that exist in all the future worlds) that will see it. There will be a sequence of worlds in which the cats are always alive and another sequence where they are always dead. There are also many sequences of worlds where the cats are alive one quarter of the time and dead three quarters of the time, and that is true for any other statistics different from $(\frac{1}{2}, \frac{1}{2})$. So that we can be certain that many of our descendants will *not* observe Born's rule in their worlds.

But one could argue, on the basis of the law of large numbers, that, at least in the vast majority of worlds, the Born rule will be obeyed, since, in the vast majority of worlds, the frequencies of dead and live cats will be close to $(\frac{1}{2}, \frac{1}{2})$.

But what happens if, instead of being $(\frac{1}{2}, \frac{1}{2})$, the probabilities predicted by quantum mechanics are, say, $(\frac{3}{4}, \frac{1}{4})$? We will still have two worlds coming out of each experiment, because these experiments have two possible outcomes. So, the structure of the multiplication of worlds is exactly the same as when the predicted probabilities were $(\frac{1}{2}, \frac{1}{2})$.

But now, if one applies the law of large numbers as above, one arrives at the conclusion that, in the vast majority of worlds, the quantum predictions will *not* be observed, since our descendants will still see the cats alive in approximately $\frac{1}{2}$ of the worlds, and the cats dead also in approximately $\frac{1}{2}$ of the worlds, instead of the $(\frac{3}{4}, \frac{1}{4})$ frequencies predicted by the Born rule.

This is a serious problem for the many-worlds interpretation. There have been many proposals to solve this problem and it would be too long and too technical to discuss all of them here.

Some authors have argued that one should count the worlds differently, by weighting them with the coefficients that appear in the Born rule [13, 16].

However, this does not answer the objection above.

Another "solution" is to give to low probability worlds (according to Born's rule) a lower degree of existence or of reality, but it is unclear what it means to live in a low reality world since we cannot compare that life with one in a world with a high degree of existence: indeed, different worlds don't interact with each other.

And this "solution" is quite contrary to Everett's original idea that "*all* elements of a superposition are 'actual', none any more 'real' than the rest".

Moreover, one also has the problem of ontology: what proliferates are wave functions, but the latter are mathematical objects not "worlds" in space-time.

One solution is to associate to the wave function the continuous matter density (1), as in the spontaneous collapse theories, see [3].

However, this does not solve the problem of the probabilities discussed above. Coming back to our example with two possible outcomes, one having probability $\frac{3}{4}$ and the other $\frac{1}{4}$, the density of matter will be different in the world where one sees the outcome having probability $\frac{3}{4}$ from the one where one sees the outcome having probability $\frac{1}{4}$.

But what difference does it make? In which way does having a smaller or larger matter density affect my states of mind ? And if it does not, we are back to the problem that, if one repeats many times the experiment whose outcomes have probabilities $(\frac{3}{4}, \frac{1}{4})$, most of my descendants (some of course having a small matter density) will see massive violations of the Born rule.

5 The de Broglie–Bohm Theory

Nature and Nature's Laws lay hid in Copenhagen: God said, "Let de Broglie–Bohm be!" and all was light.[5]

In the de Broglie–Bohm theory, the complete state of a system with N variables at time t is specified by $\big(\Psi(t), \mathbf{X}(t)\big)$, where $\Psi(t)$ is the usual wave function, $\Psi(t) = \Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N, t)$ and $\mathbf{X}(t) = \big(\mathbf{X}_1(t), \ldots, \mathbf{X}_N(t)\big) \in \mathbb{R}^{3N}$ are the actual positions of the particles.[6]

The theory assumes that the particles have positions at all times, and therefore trajectories, independently of whether one measures them or not.

The evolution of the state $(\Psi, \mathbf{X})$ is given by two laws:

1. Ψ obeys the usual Schrödinger equation at all times. The wave function of an isolated system never collapses.
2. The evolution of the positions of the particles is guided by the wave function at time t.

[5] Adapted from Alexander Pope's epitaph about Newton.

[6] Our presentation of the de Broglie–Bohm theory follows the one of Bell [7] and of Dürr, Goldstein and Zanghì [14] rather than the one of Bohm [8]. Many expositions of the de Broglie–Bohm theory are available, see, e.g., Albert [1] or Tumulka [33] for elementary introductions and, e.g. Bohm and Hiley [9], Bricmont ([10], Dürr and Teufel [15], Goldstein [23], Maudlin [27], Norsen [30] and Towler [32] for more advanced ones.

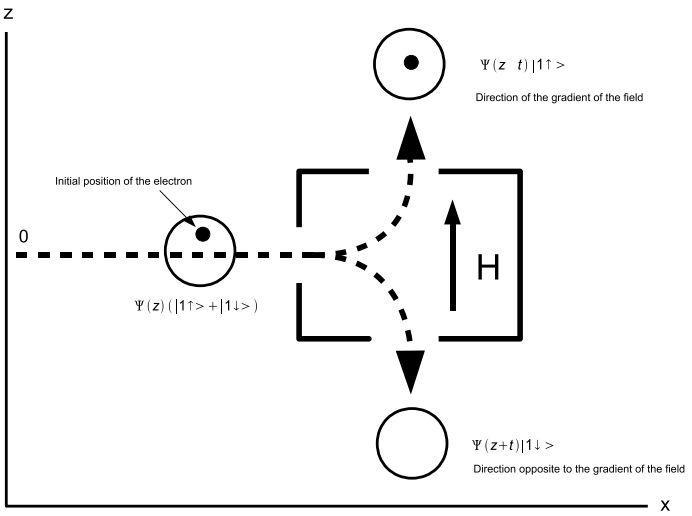


Fig. 2 An idealized spin measurement

Moreover, since the system is deterministic, one has to make statistical assumptions on the initial conditions of the system (as one does in chaos theory) in order to obtain statistical predictions.

These are the fairly natural "quantum equilibrium" ones: $\rho = |\psi|^2$.

Let me list the main qualities of this theory:

- Under the quantum equilibrium assumptions on initial conditions, one recovers the usual predictions of quantum mechanics.
- The ontology of the theory is the same as in classical physics and is thus unproblematic, unlike the ontologies of the spontaneous collapse and many worlds theories.[7] It can be summarized by Detlef's quote mentioned at the beginning of this article: "particles move". And while in classical physics, that motion is guided by gravitational or electromagnetic fields, here it is guided by the wave function, a more abstract notion but a perfectly well-defined one.
- The de Broglie–Bohm theory gives a clear physical meaning to the wave function, which is no longer simply a "probability wave" (whatever that means exactly) but a physical quantity determining the motion of particles, similar in some ways to classical Hamiltonians.
- This theory explains the "contextuality" of measurements.
  Consider an idealized spin measurement: if the wave function has a symmetry along the z axis and if the particle starts above the line of symmetry $z = 0$, it will be deflected upwards, meaning that its spin is "up", see Fig. 2.

[7] One might replace particle's positions by field configurations in quantum field theories, but that goes beyond the scope of this article.

But if we reverse the orientation of the gradient of the magnetic field and do the same "measurement" with the same initial wave function and the same initial position of the particle, the particle will still be deflected upwards, but that means now that its spin is "down", see Fig. 3.
So, "measurements" don't measure intrinsic properties of the system (except for position measurements).
Measurements are **interactions** between an apparatus and a system.
This was stated intuitively by Bohr, but here it follows from the equations of the theory.

- The de Broglie–Bohm theory is a statistical theory, but, unlike the naïve one mentioned in the Sect. 2, it is consistent and is not refuted by the no hidden variables theorems of Bell and Kochen and Specker, because it does not introduce the "hidden variables" that are forbidden by these theorems, like spin values preexisting to their "measurement".
  A subtle but crucial point: the de Broglie–Bohm theory is a hidden variable theory that is not refuted by the no hidden variables theorems.
  Often missed (to put it mildly)!
- One can use this contextuality of measurements to illustrate how nonlocality works in the de Broglie–Bohm theory.
  If the result of a "measurement" on one side of an EPR-Bell experiment with an entangled pair of particles depends on how the orientation of the gradient of the field is oriented on that side, then changing that orientation will affect the behavior of the particle on the other side: since the spins of both particles have to be anti-correlated, if changing that orientation on one side changes the "value" of the spin

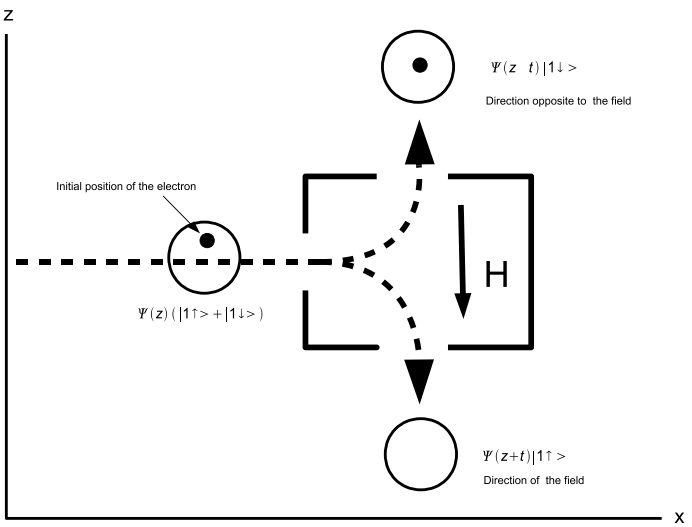


Fig. 3 An idealized spin measurement with the orientation of the gradient of the magnetic field reversed with respect to the one in Fig. 2

on that side from up to down, then the spin on the other side must go from down to up. But that means that the trajectory followed by the particle on the other side must also change.

This makes nonlocality explicit in the de Broglie–Bohm theory, which is a (big) quality rather than a defect since Bell has shown that nonlocality is a property of the world and not only of the quantum theory.

6 Conclusions

While we have left out certain questions such as quantum field theory and relativity (that can be dealt with within de Broglie–Bohm theory but it would be too long to discuss that), we want to stress certain aspects of that theory:

The de Broglie–Bohm theory is **not** a different theory from ordinary quantum mechanics. It is the rational completion of ordinary quantum mechanics. The latter is just the algorithm allowing us to predict "results of measurements" and that algorithm can be derived from the de Broglie–Bohm theory.

Most physicists either don't care about the meaning of their most fundamental theory ("shut up and calculate") or adhere (in the back of their mind) to the naïve statistical interpretation.

Most physicists won't be persuaded by spontaneous collapse theories (too ad hoc) unless some future experiments contradict ordinary quantum predictions.

If that happens, they are likely to look for an entirely different theory (nonlinear?)

An important minority of physicists "like" the Many-Worlds interpretation.

But I believe that this is because they haven't thought it through.

Apart from its fantastic nature, one has to provide it with an ontology and solve the problem of the statistical predictions (unsolved since 1957).

So, that leaves us with the de Broglie–Bohm theory as the only option. I believe it is gaining popularity, due in part to the work of Detlef Dürr, although it is still very marginal.

For it to become more popular, one needs:

- That physicists start to worry about the meaning of their most fundamental theory.
- That they be better aware of the nature of the problem: not the measurement one, but the meaning or ontology. And that they also become aware of the no hidden variable theorems that refute their naïve statistical interpretation.

But that is a long way to go!

Acknowledgements I thank Sheldon Goldstein and Tim Maudlin for many discussions on Bohmian mechanics and the many-worlds theories.

References

1. D. Albert, Bohm's alternative to quantum mechanics. Sci. Amer. **270**(1994), 32–39 (1994)
2. V. Allori, S. Goldstein, R. Tumulka, N. Zanghì, On the common structure of Bohmian mechanics and the Ghirardi-Rimini-Weber theory. Br. J. Philos. Sci. **59**, 353–389 (2008)
3. V. Allori, S. Goldstein, R. Tumulka, N. Zanghì, Many-worlds and Schrödinger's first quantum theory. Br. J. Philos. Sci. **62**, 1–27 (2011)
4. A. Bassi, G.C. Ghirardi, Can the decoherent histories description of reality be considered satisfactory? Phys. Lett. A **257**, 247–263 (1999)
5. J.S. Bell, On the problem of hidden variables in quantum mechanics. Rev. Modern Phys. **38**, 447–452 (1966). Reprinted as Chap. 1 in [7]
6. J.S. Bell, Are there quantum jumps?, in *Schrödinger. Centenary Celebration of a Polymath*. ed. by C.W. Kilmister (Cambridge University Press, Cambridge, 1987), pp.41–52. Reprinted as Chap. 22 in [7]
7. J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics. Collected Papers on Quantum Philosophy*, 2nd edn. (2004). With an introduction by Alain Aspect, 1st edn. (Cambridge University Press, Cambridge, 1987)
8. D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden" variables. I. Phys. Rev. **85**(2), 166–179 (1952). II. Phy. Rev. **85**(2), 180–193
9. D. Bohm, B.J. Hiley, *The Undivided Universe* (Routledge, London, 1993)
10. J. Bricmont, *Making Sense of Quantum Mechanics* (Springer, Switzerland, 2016)
11. J. Bricmont, S. Goldstein, D. Hemmick, Schrödinger's paradox and proofs of nonlocality using only perfect correlation. J. Stat. Phys. **180**, 74–91 (2019)
12. M. Daumer, D. Dürr, S. Goldstein, N. Zanghì, Naive realism about operators. Erkenntnis **45**, 379–397 (1996)
13. B. DeWitt, R.N. Graham (eds.), *The Many-Worlds Interpretation of Quantum Mechanics* (Princeton University Press, Princeton, 1973)
14. D. Dürr, S. Goldstein, N. Zanghì, *Quantum Physics Without Quantum Philosophy* (Springer, Berlin, 2013)
15. D. Dürr, S. Teufel, *Bohmian Mechanics. The Physics and Mathematics of Quantum Theory* (Springer, Berlin-Heidelberg, 2009)
16. H. Everett, 'Relative state' formulation of quantum mechanics. Rev. Modern Phys. **29**, 454–462 (1957). Reprinted in [13] (pp. 141–149)
17. M. Gell-Mann, J.B. Hartle, Alternative decohering histories in quantum mechanics, in *Proceedings of the 25th International Conference on High Energy Physics, Singapore, 1990*, ed. by K.K. Phua, Y. Yamaguchi (World Scientific, Singapore, 1991)
18. G.C. Ghirardi, A. Rimini, T. Weber, Unified dynamics for microscopic and macroscopic systems. Phys. Rev. D **34**, 470–491 (1986)
19. G.C. Ghirardi, R. Grassi, F. Benatti, Describing the macroscopic world: closing the circle within the dynamical reduction program. Found. Phys. **25**, 5–38 (1995)
20. G.C. Ghirardi, Collapse theories, in *The Stanford Encyclopedia of Philosophy*, E.N. Zalta, Winter 2011 Edition (2011). http://plato.stanford.edu/entries/qm-collapse/
21. S. Goldstein, Quantum theory without observers. Physics Today, Part I, March 1998, pp. 42–46; Part II, April 1998, pp. 38–42; S. Goldstein et al.: Letters and replies in Physics Today, February 1999, p. 11
22. S. Goldstein, R. Tumulka, N. Zanghì, The quantum formalism and the GRW formalism. J. Stat. Phys. **149**, 142–201 (2012)
23. S. Goldstein, Bohmian mechanics, in *The Stanford Encyclopedia of Philosophy*, ed. by E.N. Zalta, Spring 2013 Edition (2013). https://plato.stanford.edu/archives/spr2013/entries/qm-bohm/
24. R.B. Griffiths, *Consistent Quantum Theory* (Cambridge University Press, Cambridge, 2002)
25. S. Kochen, E.P. Specker, The problem of hidden variables in quantum mechanics. J. Math. Mech. **17**, 59–87 (1967)

26. T. Maudlin, *Quantum Non-locality and Relativity: Metaphysical Intimations of Modern Physics*, 3rd edn. (Wiley-Blackwell, Oxford, Malden, MA, USA, 2011)
27. T. Maudlin, *Philosophy of Physics: Quantum theory* (Princeton University Press, Princeton, 2019)
28. N.D. Mermin, Hidden variables and the two theorems of John Bell. Rev. Modern Phys. **65**, 803–815 (1993)
29. T. Norsen, Einstein's Boxes. Amer. J. Phys. **73**, 164–176 (2005)
30. T. Norsen, *Foundations of Quantum Mechanics: An Exploration of the Physical Meaning of Quantum Theory* (Springer International Publishing, Switzerland, 2017)
31. R. Omnès, *Understanding Quantum Mechanics* (Princeton University Press, Princeton, 1999)
32. M. Towler, De Broglie–Bohm pilot-wave theory and the foundations of quantum mechanics, lectures (2009). https://casinoqmc.net/pilot_waves.html
33. R. Tumulka, Understanding Bohmian mechanics - a dialogue. Amer. J. Phys. **72**, 1220–1226 (2004)
34. R. Tumulka, Paradoxes and primitive ontology in collapse theories of quantum mechanics, in *Collapse of the Wave Function*, ed. by S. Gao (Cambridge University Press, Cambridge, 2018), pp. 134–153

The Prodigy That Time Forgot: The Incredible and Untold Story of John von Newton

Craig Callender

Abstract By developing an absurd counterfactual history, I show that many objections launched against Bohmian mechanics could also have been made against Newtonian mechanics. This paper introduces readers to Koopman–von Neumann dynamics, an operator-based Hilbert space representation of classical statistical mechanics. Lessons for quantum foundations are drawn by replaying the battles between advocates of standard quantum theory and Bohmian mechanics in a fictional classical history.

Born in the year 1603 in a small hamlet in the Kingdom of Hungary, John von Newton was an extraordinary polymath. It was said that when he was only six years old that he could divide two nine digit numbers in his head while conversing fluently in Ancient Greek. Widely acclaimed as the last mathematician who was equally at home in pure algebra and applied alchemy, his contributions in the Wallachian Project of the Thirty Years War led to the development of the cannon known as the Orban II. While some may know him for his development of mechanical automata, "it's-not-a-game" theory, and numerical astrology, his unparalleled advances in physics were what made him famous amongst contemporaries. However, these advances were controversial and quickly forgotten. This essay is a recounting of the astonishing breakthroughs made by John von Newton and their equally extraordinary reception.

Due to the plague in 1620, von Newton (Fig. 1) was sent home and had to study remotely. Because lessons were wrapped in straw, it was called learning by Broom. While many students suffered greatly from Broom courses and the resulting social and intellectual isolation, the circumstances had the opposite effect on a prodigy like von Newton. Finally separated from teachers and students of inferior talent, he embarked on what can only be described as the most remarkable set of intellectual leaps to ever occur in world history. In short, in six months von Newton discovered an empirically adequate (then) new physics, a theory equivalent to classical statistical mechanics, and all of the mathematical innovations necessary to express this theory

C. Callender (✉)
University of California San Diego, San Diego, CA, USA
e-mail: ccallender@ucsd.edu

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_5

Fig. 1 John von Newton

(e.g., calculus, analysis). A month later he represented this theory with an operator formalism in a state space we now call Hilbert space.

This achievement was completed in 1620, yet by the time he died in 1699 this massive feat was forgotten. (It is speculated that the tumor that killed him may have been due to his work with toxic alchemical materials while developing Orban II.) It took until the late 19th and early 20th centuries for science to rediscover what von Newton already learned. In what follows I will summarize what he accomplished and his fate.

1 The Classical Schrödinger Equation

Contemporary writers said that von Newton would often go to bed troubled by a problem and wake up with the solution. That is why he kept a quill pen and parchment by his bedside. We don't know what problem he had in mind on the night of Feb 3, 1620, but the sepia-colored notepad from Feb 4 survives (Fig. 2). On it one can make out the faded remains of an equation that takes the following form

$$i\frac{\partial \Psi}{\partial t} = L\hat{\Psi} \tag{1}$$

when put in modern terminology. Here $\Psi(\varphi)$ is a classical wavefunction that is a ray in a complex Hilbert space. Unlike in quantum mechanics, its domain is phase space $\varphi = (q, p)$, not configuration space. The generator $\hat{L}$ is the Hermitian Louiville operator and it evolves the ray through Hilbert space with time.

Equation (1) wasn't (re)discovered until the 1930s by Koopman (1931) and a year later by von Neumann (1932). Working in the context of ergodic theory, Koopman showed that unitary transformations are central to classicla physics. In so doing

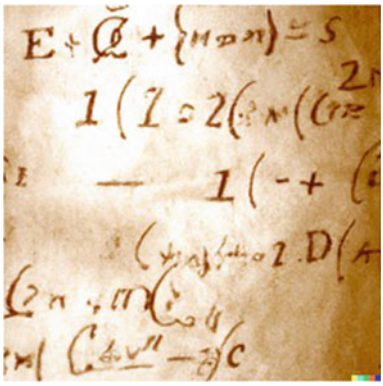


Fig. 2 Some of von Newton's calculation

he proved that if a wavefunction $\Psi(\varphi)$ satisfies (1) then the probability density $\rho(\varphi) = |\Psi(\varphi)|^2$ satisfies the classical Liouville equation $\partial_t \rho = \{H, \rho\}$, where H is the classical Hamiltonian of the system. Although this result should be widely known, apparently it is not as it has been rediscovered many times, often by very prominent physicists, e.g., Berry, Wiener, 't Hooft.[1]

von Newton's equation obviously bears a great similarity to its more famous cousin, the Schrödinger equation of quantum mechanics: $i\frac{\partial \Psi}{\partial t} = \hat{H}\Psi$. Working through von Newton's notebooks, we were astonished to see how "quantum mechanical" his formalism was. He began with four postulates:

1. The state of the system is represented by a vector $|\Psi\rangle$ in a complex Hilbert space.
2. The state space of a composite system is the tensor product of the subsystems' state spaces.
3. For any observable A, there is an associated Hermitian operator $\hat{A}$ and eigenvalue problem $\hat{A}\,|A\rangle = a\,|A\rangle$. The eigenvalue a is understood as representing a particular outcome measured in a lab.
4. The probability of measuring a is given by $P(a) = |\langle A|\ \Psi(t)\rangle|^2$. von Newton called this "Bodor's Law" .

von Newton called postulate 4 "Bodor's Law" in honor of a friend who sold the best goat milk in the hamlet. However, the name probably stuck because Bodor was renowned for his gambling prowess. von Newton interpreted Bodor's Law as arising due to an instantaneous collapse of the state $|\Psi\rangle$ into the eigenstate $|A\rangle$ associated with the measured eigenvalue a.

[1] See Berry (1992), Chirikov, Izrailev and Shepelyanskii (1988), Della Riccia and Wiene (1966), and 't Hooft (1997). The "classical Schrödinger equation" (1) should not be confused with another "classical Schrödinger equation" derived in the 1960's by Schiller (1962) and Rosen (1964). This later equation defines the wavefunction on configuration space $\Psi(q)$ whereas (1) applies to a wavefunction over phase space.

Because there is no uncertainty relation in classical physics, position and momentum have a common set of eigenstates in von Newton's Hilbert space. As mentioned, the eigenkets therefore live in phase space, not configuration space, i.e., $|A\rangle = |q, p\rangle = |q\rangle \otimes |p\rangle$. The vectors $|q, p\rangle$ form a basis of the space. By assuming what we would call a "classical commutator" $[\hat{q}, \hat{p}] = 0$ rule, von Newton was able to derive Eq. (1) from these four postulates (Bondar et al. 2021). Without modern mathematical physics at his disposal, unfortunately it took von Newton 250 pages of calculation to get this result. Helping ourselves to modern results such as the Ehrenfest Theorems and Stone's Theorem, today we can derive (1) very quickly (Wilczek 2023). Interestingly, recently (Bondar et al. 2012) show that replacing the classical commutator with the quantum commutation relations but otherwise retaining the same postulates $1 - 4$ as above leads to standard quantum mechanics. In 1620 von Newton was only one tiny adjustment from discovering quantum theory!

In any case, the resulting theory is an operator-based probabilistic theory that makes predictions about the values of measurements. The generator of motion evolves the state in a complex Hilbert space between measurements via (1) just as the Hamiltonian does in the Schrödinger equation. The norm $\langle \Psi(t)|\Psi(t)\rangle$ is conserved by the time evolution, which helps justify Bodor's Rule. And one can calculate expectation values of observables and even easily switch vector bases as one does in quantum mechanics. See Gozzi and Mauro (2004), Jordan and Sudarshan (1961), Mauro (2002), and Bondar et al. (2012) for the state of the art on Koopman–von Neumann dynamics.

As much as it looks like quantum mechanics, however, von Newton's theory was purely classical. The wavefunction lives in phase space, not configuration space. And the probabilities are the ones predicted by classical statistical mechanics, not quantum mechanics. The probabilities predicted by Bodor's Rule correspond precisely to solutions of classical statistical mechanics, i.e., the probability densities given by the classical Liouville equation. In a two slit experiment (see Mauro 2002 for a clear analysis) the phases of the classical waves cancel out and the total probability distribution on the screen is the sum of the probability distributions for each slit, reproducing what we expect classically. The theory was empirically adequate to then known empirical phenomena, which at this time consisted mostly of cannon ball trajectories.

2 Reception

When the plague ended, von Newton promoted his theory at various august academic bodies throughout Europe. With such a breathtaking set of advances, he expected to be lauded as having produced a great triumph of reason. "If I have seen further," he said, "it is because I stand as a giant." Instead the response was somewhat chilly. Scientists were impressed, but they felt uneasy about von Newton's product. His peers wanted to understand the nature of physical reality. Rene Descartes had posited a world consisting of corpuscles organized in complicated vortexes, but what was von

Newton offering? A kind of operationalist "black box" quality pervaded his theory, as his operator formalism provided only predictions for various observables.

At the University of Zurich he met a physicist named Albert Mechanstein, who would prove to be a real thorn in von Newton's side. Mechanstein was a disciple of the philosophy of Descartes. He said to von Newton that it's all well and good that you've accurately predicted the probability distribution of a bunch of cannon balls hitting a castle wall and of arrows entering sniper windows, but you don't say anything about what constitutes these balls, arrows, and walls, nor the reason *why* they behave the way they do. von Newton replied,

> Since a good theory must be based on directly observable magnitudes, I thought it more fitting to restrict myself to these.

As von Newton later recounted, Mechanstein was stunned:

> But you don't seriously believe that none but observable magnitudes must go into a physical theory?...It is the theory which decides what we can observe.

von Newton was equally upset, reporting that he was "completely taken aback by [Mechanstein]'s attitude."[2] He felt that it is "wrong to think that the task of physics is to find out how nature is"; rather, he thought, "Physics concerns what we can say about nature."[3]

As he travelled von Newton heard more objections. Pressure was put on the relationship between Bodor's Law and Eq. (1). von Newton held that we have "two fundamentally different types of interventions which can occur in a system; when an object is undisturbed, Eq. (1) "describes how the system changes continuously and causally in the course of time" but once measurement happens something "discontinuous, non-causal, and instantaneous" occurs, i.e., the collapse via Bodor's Law to an eigenstate.[4] The dynamics is deterministic when no measurement is happening, but indeterministic when it is.

This response, however, only focused attention on the role of measurement in von Newton's theory. Like standard quantum theory, von Newton's theory has a measurement problem.[5] Equation (1) is linear and allows superpositions of macroscopic outcomes; measurement collapses these superpositions to an eigenstate of

---

[2] Heisenberg recounting his discussions with Einstein, quoted in Becker (2018), 29.

[3] Bohr on physics after the Solvay conference, quoted in Becker (2018), 49.

[4] von Neumann describing his two dynamics, quoted in Becker (2018), 67.

[5] The two measurement problems are slightly different and interesting to consider. As Mauro (2002) emphasizes, the fundamental difference to consider. As Mauro (2002) emphasizes, the fundamental difference between Koopman–von Neumann theory and ordinary quantum theory is that in the former but not the latter the phase interacts with the modulus. Contrast a Madelung decomposition of Eq. (1) with the Schrödinger equation. Write the quantum wavefunction as $\psi(x) = A(x)exp[i/\hbar S(x)]$and substitute it into the Schrödinger equation and then separate real and imaginary parts. Then as is well known one obtains

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + V = \frac{\hbar^2}{2mA}\frac{\partial^2 A}{\partial x^2}$$

$$m\frac{\partial A}{\partial t} + \frac{\partial A}{\partial x}\frac{\partial S}{\partial x} + \frac{A}{2}\frac{\partial^{2S}}{\partial x^2} = 0$$

the relevant observable. What interaction qualifies as a measurement, Mechanstein asked?

von Newton's acolytes differed on this question. Some said that a measurement only occurs when an outcome has been registered by a divine soul. That raised the question of who had souls. The Synod of Mâcon had long ago clarified that women had souls and Descartes compellingly argued that animals do not. But what about pagans and serfs? They certainly looked and acted like they could apply von Newton's theory as well as anyone, but did they have divine souls? And what types of souls were necessary? Was having only a vegetative soul sufficient to collapse a wavefunction?

Other acolytes did not use souls but understood measurement as an interaction between systems described in different ways. Bohr taught us that measuring devices are inherently classical, that the interaction between the classical and the quantum is central to explaining measurement. It's amazing to learn that there was a counterpart to this Bohrian position back in von Newton's day. One of his followers held that *measuring devices are essentially medieval*. What triggers a measurement is the interaction of a classical system with a medieval one, e.g., catapult, plough, water mill. Opponents felt that "medieval" was too vague to be a fundamental category in a physical theory.

A common theme emerged: scientists of the day didn't like the fundamental split between measurer and measured. Shouldn't the measurer—be they a stone mason, a nobleman, or a scythe—be itself describable in the language of physics? Why must there always be this shifty subject/object split in physics? von Newton's protestations that "for all practical purposes" it didn't matter found few sympathetic ears.

---

where one can see that the phase S is coupled to the modulus A. Do the same for the classical wavefunction $\psi(x) = F(q, p)exp[i/\hbar G(q, p)]$ when inserted into (1). Then we get

$$i\frac{\partial F}{\partial t} = \hat{\mathcal{H}}\mathcal{F}$$

$$i\frac{\partial G}{\partial t} = \hat{\mathcal{H}}\mathcal{G}$$

and no coupling between F and G. (Why then introduce phases at all? They become necessary if one wants the freedom of basis one gets in Hilbert space; see Mauro 2002.) As a result of this decoupling, wavefunctions without phases cannot generate them in their time evolution. Hence the measurement problem is a bit different than quantum mechanically. In the language of foundations of physics, the classical measurement problem associated with Koopman–von Neumann is like the quantum one if decoherence worked perfectly, driving the off-diagnol terms to exactly zero. That still leaves a measurement problem, the so-called "and" to "or" problem of Bell (1990) (see also Maudlin 1995). On the classical measurement problem, see Chen (2022) (section 5.4), Katagiri (2020), and McCoy (2020).

3 Classical Mechanics Without Obs'rv'rs

What really put pressure on von Newton's theory was the remarkable development of a mechanistic theory by someone with essentially the same name, Isaac Newton. In 1687 Newton published the *Philosophiæ Naturalis Principia Mathematica*. The *Principia* posited an ontology of corpuscles who always evolve according to the same dynamical equation. Cannon balls, cannon ball parts, cannon ball operators, and cannon ball victims could all be described at once by Newton's famous second law. There was no subject/object split, no fine discussions of what types of souls or medieval devices collapse wavefunctions, or any of that. Positing one basic law rather than two, Newton offered what he called a "mechanics without obs'rv'rs."

In our age, Newton is famous for offering a physics that unified celestial and terrestrial spheres, the heavens and the earth. Back then he was also known for having provided a deeper unification of von Newton's process 1 dynamics (the deterministic Eq. 1) with von Newton's process 2 dynamics (Bodor's law). He unified the spheres *and* the two types of dynamics.

More than that, Newton offered the physical "nut-and-bolt" explanations that people didn't find in von Newton's physics. In a siege of a castle, one might shoot a cannon aimed at a wall many times. Cannon operators noticed a kind of normal statistical pattern developing on the wall. Again and again, attack after attack, similar probability distributions appeared on castle walls. Why? von Newton's physics would predict these distributions, but they couldn't answer *why* they might appear like this. It would be a very hard calculation to do, but Newton's physics at least offered one understanding of what must be going on. Small changes in the initial positions and velocities of the cannon balls, plus tiny fluctuations in their mass, are to be expected. Patterns in these differences are then responsible for why the cannon balls form these distributions.

More generally, going back to Mechanstein's complaint, the theory "decides" what is observable. That is, we can explain what is observable in terms of the posited ontology—corpuscles—and laws. We do not begin with observations as primitive, but offer explanations for why we observe what we do. These explanations were possible because Newton offered an ontology and clear laws, something that von Newton rejected.

When Newtonians ultimately derived von Newton's theory from their own, that was the death knell of the latter's influence. Suppose we have a swarm of Newtonian corpuscles sweeping out continuous trajectories through time. We can think of this as a kind of fluid described by a density $\rho(x, p, t)$. If we insist that its value is non-negative and real, it can be interpreted as the probability of a particle being at point x at time t with momentum p (using measure $\int dxdp$). It follows from Newtonian mechanics that the flow of this fluid is incompressible, which implies that

$$\frac{\partial p}{\partial t} = -\dot{x}\frac{\partial \rho}{\partial x} - \dot{p}\frac{\partial \rho}{\partial p} \tag{2}$$

holds, which provides a dynamics for ρ. Equation (1) can be derived from (2) by defining a wavefunction $\rho \equiv |\Psi(x, p)|^2$ and multiplying both sides by i (Wilczek 2023). So with Newton one could get all of von Newton's predictions but also explain *why* we were observing what we do. We could open the black box and see what's going on.

4 Criticism of Newton

von Newton and his advocates did not take these provocations lightly. They viciously attacked Newton and his physics. One giant defender of von Newton did not deign to comment on Newton's physics directly, but through intermediaries said it was "very foolish."[6] Another very distinguished physicist called it "artificial metaphysics."[7] Some took an extremely bold position (bold because manifestly false) and held that that there was no alternative to von Newton and his interpretation, that von Newton's physics "eminently possesses this character of uniqueness" in it.[8] Mostly inspired by an extreme empircist or even positivistic philosophy, these objections fell on deaf ears among the Cartesians and Newtonians of the day.

von Newton even made a political case against Newton. Like Leibniz, he wrote to Princess Caroline of Ansbach complaining about Newton's theory. Leibniz accused Newton of positing occult qualities through his non-local gravitational force and of requiring God to act as a clockmaker, fixing his product from time to time. von Newton picked up on this and also complained that Newton's clockwork universe deprived us of free will whereas his indeterministic theory made room for it. Newton was summoned before Parliament's House for Unpious Activities Committee as a result, but he answered the charges so well that no stain was left on his reputation and he was ultimately made Master of the Royal Mint.[9]

Finding fewer and fewer supporters, von Newton could only find an employment with a few of his followers at the University of Copenhagen. There he toiled in obscurity until the minstrels only sang of Newton and never the great von Newton. In some sense he had the last laugh, however, as his papers left in the gorgeous library at the University of Copenhagen were found by a young physicist named Niels Bohr.

[6] Bohr on Bohm, cited in Becker (2018), 107.

[7] Pauli on Bohm, cited in Becker (2018), 107.

[8] Rosenfeld (1957), 4–42.

[9] See Cushing (1994) for many objections to Bohm along these lines, especially by Pauli. Cushing also details the political attacks on Bohm.

5 Lessons from the Rise and Fall of von Newton

It's an honor for me to write in a volume dedicated to Detlef Dürr. He filled a room with both his warmth and knowledge. It would be impossible for me to quantify how much I learned from him and his group. One article that made a special impression is "Naive Realism about Operators", which inspired this paper. "Naive Realism..." shows in detail how the entire Hilbert space operator formalism mechanics can be derived from natural assumptions and moves from Bohmian mechanics. It argues that one should not confuse mathematical operators with physical properties of systems. Doing so leads to a fetishization of the quantum operator algebra that becomes an even bigger problem than the measurement problem.

In my absurd counterfactual history, I mimic this situation classically. I imagine that a measurement operator formalism arose first and then Newton came along with a dynamics for classical "beables" (an always determinate ontology). From this dynamics and ontology, one can then derive in detail how the entire Koopman–von Neumann Hilbert space operator formalism might arise. In the actual world, we had Newton first and Koopman–von Neumann second; and later, standard quantum theory first, Bohmian mechanics second (by only two years in the form of de Broglie). Should the temporal order of these appearances matter? I don't think so. Yet it seems almost unconscionable to launch the counterparts of the objections directed at Bohm in the actual world to Newton. Newtonian mechanics is rightly celebrated as one of the great achievements of science. While there are of course differences between the cases of Bohm and Newton, many common objections do not rely on these differences.

Since we can deduce the operator formalism of Koopman–von Neumann from Newtonian dynamics and had the latter first, we were never tempted to be "naive realists" about classical operators. But had things worked out differently, we might have been. We're often better at seeing mistakes in the past than the present, so I invented a counterfactual past and transported mistakes across times and worlds.[10]

Another lesson of the Koopman–von Neumann theory is that it is important to tease apart features of a particular mathematical representation of a theory from the theory itself. Features of a representation have a pernicious way of sneaking into our interpretation of the theory and how we evaluate it and alternatives. Jennings and Leifer (2016) ask "what phenomena of quantum theory are intrinsically non-classical?" To answer this question they apply a criterion:

> If a phenomenon of quantum physics also occurs within a classical statistical physics setting, perhaps with minor additional assumptions that don't violently clash with our everyday conceptions, then it should not be viewed as an intrinsically quantum mechanical phenomenon.

They conclude that many "commonly touted phenomena" such as randomness, complementarity, collapse of the wavepacket, the use of wavefunctions and Hilbert space, and more, cannot be marks of intrinsically quantum phenomena. I wholeheartedly agree. By placing classical statistical mechanics in an operator-based formalism in

[10] See Nikolić (2008) for a less incredible counterfactual history toward the same point.

Hilbert space, Koopman and von Neumann demonstrably show that many of these representational features are not inherently quantum mechanical. Not only is classical physics expressible in a similar formalism, but it can also employ collapses of the wavefunction, Born's Rule, a fundamental subject/object split, and two types of dynamics. The operator measurement formalism seems to almost invite an interpretation with an instrumentalist flavor.

If one is a naive realist about classical observables, Koopman–von Neumann even has a measurement problem. But that is the result of a choice, a bad choice. Classical statistical mechanics does not have a measurement problem. Neither does quantum mechanics if one adopts a decent interpretation, e.g., the Bohmian mechanics that Detlef prized. The measurement problem in Koopman–von Neumann makes this point plain. There it results not from the peculiarities of the classical world but from the peculiarities of "quantum philosophy" applied to the classical world. As Detlef saw much better than most, the same is true in quantum physics. That is the ultimate lesson of the tragedy of the great and forgotten John von Newton, the naive realist about classical observables.

Acknowledgements Thanks to Jacob Barandes, Eddy Chen, Casey McCoy, the audience at the University of Lisbon's Open Problems in Philosophy of Physics conference, and the UC San Diego philosophy of physics reading group for helpful comments. Details of von Newton's life were drawn from von Neumann's biography. Figures 1 and 2 were generated with the assistance of DALL·E 2. Figure 1 is a blend of the faces of John von Neumann and Isaac Newton.

References

A. Becker, *What is Real?* (Basic Books, NY, 2018)
J.S. Bell, Against "Measurement," in *Sixty-Two Years of Uncertainty*. ed. by A.I. Miller (Plenum Publishing, 1990), pp.17–31
M.V. Berry, True quantum chaos? An instructive example, in *New Trends in Nuclear Collective Dynamics*. ed. by Y. Abe, H. Horiuchi, K. Matsuyanagi (Springer, 1992)
D.I. Bondar, R. Cabrera, R.R. Lompay, M.Y. Ivanov, H.A. Rabitz, Operational dynamic modeling transcending quantum and classical mechanics. Phys. Rev. Lett. **109**, 190403 (2012)
E. Chen, The Wentaculus: density matrix realism meets the arrow of time, in *Physics and the Nature of Reality: Essays in Memory of Detlef Dürr*. ed. by A. Bassi, S. Goldstein, R. Tumulka, N. Zanghì (2022)
J. Cushing, *Quantum Mechanics: Historical Contingency and the Copenhagen Hegemony* (University of Chicago Press, Chicago, 1994)
M. Daumer, D. Dürr, S. Goldstein et al., Naive realism about operators. Erkenntnis **45**, 379–397 (1996)
G. Della Riccia, N. Wiener, Wave mechanics in classical phase space, brownian motion, and quantum theory. J. Math. Phys. **6**, 1372–1383 (1966)
D. Dürr, S. Goldstein, N. Zanghì, Quantum physics without quantum philosophy. Stud. Hist. Philos. Phys. **26**, 137–149 (1995)
D. Dürr, D. Lazarovici, *Understanding Quantum Mechanics: The World According to Modern Quantum Foundations* (Springer, 2020)
E. Gozzi, D. Mauro, On Koopman-Von Neumann waves II. Int. J. Mod. Phys. A **19**, 1475 (2004)

D. Jennings, M. Leifer, No return to classical reality. Contemp. Phys. **57**(1), 60–82 (2016)
T. Jordan, E.C.G. Sudarshan, Lie group dynamical formalism and the relation between quantum mechanics and classical mechanics. Rev. Mod. Phys. **33**, 515–524 (1961)
S. Katagiri, Measurement theory in classical mechanics. Progress Theor. Exp. Phys. (6), 063A02 (2020)
B.O. Koopman, Hamiltonian systems and transformations in Hilbert space. Proc. Natl. Acad. Sci. USA **17**, 315–318 (1931)
T. Maudlin, Three measurement problems. Topoi **14**, 7–15 (1995)
D. Mauro, On Koopman-von Neumann waves. Intl J. Mod. Phys. A **17**, 1301 (2002)
C. McCoy, An alternative interpretation of statistical mechanics. Erkenntnis **85**(1), 1–21 (2020)
H. Nikolić, Would Bohr be born if Bohm were born before Born? Amer. J. Phys. **76**, 143–146 (2008)
N. Rosen, The relation between classical and quantum mechanics. Amer. J. Phys. 597–600 (1964)
L. Rosenfeld, Misunderstandings about the foundations of quantum theory, in *Observation and Interpretation*. ed. by S. Korner (Butterworths Scientific Publications, London, 1957)
R. Schiller, Quasi-classical theory of the Non-spinning electron. Phys. Rev. **125**, 1100–1108 (1962)
G. 'T Hooft, Quantum mechanical behaviour in a deterministic model. Found. Phys. Lett. **10**, 105–111 (1997)
J. Von Neumann, Zur Operatorenmethode in der klassischen Mechanik. Ann. Maths **33**, 587–642 (1932)
F. Wilczek, Notes on Koopman von neumann mechanics, and a step beyond (2015). Accessed 5 April 2023. http://frankwilczek.com/2015/koopmanVonNeumann02.pdf

Bohmian Collapse

Isaac Wilhelm

Abstract I present and explain the Bohmian account of collapse in quantum mechanics.

1 Introduction

This paper is an explication of an idea due to Detlef Dürr and collaborators. The idea is striking: Bohmian mechanics has the formal and physical resources to fully account for the phenomenon of quantum collapse. This account is perhaps not as well understood, in philosophical circles at least, as it deserves to be. So my goal, in what follows, is to present a simple yet rigorous version of this account; hopefully, thereby, increasing its audience of appreciators.

But before doing so, it is worth making a brief remark about the passing of Detlef Dürr. There is joy, and also pain, in reading through this volume. The joy comes from seeing how many lives Detlef touched, and the community which formed around him. The pain comes from the reminder of his loss.

There is some comfort to be had, however, in Detlef's own research. For Bohmian mechanics suggests that when someone dies, their particles disperse according to a determined, coordinated dance, one in which—by virtue of nonlocality—we all participate. Just as clouds, gently floating across the sky, are carried towards the distant horizon by their particulate motions, so we are carried along by our composite particles, making tracks towards a horizon where everything familiar vanishes; a vanishing point through which Detlef has already passed, and through which all else passes too, and lucky for us that on the way there, our trajectories briefly crossed his.

I. Wilhelm (✉)
Department of Philosophy, National University of Singapore, Singapore, Singapore
e-mail: iwilhelm@nus.edu.sg

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_6

2 Basics

In this section, I present the basics of Bohmian mechanics.[1] Roughly put, according to the version of Bohmian mechanics on which I focus here, the universe consists of some particles and a universal wave function. The resources used to describe all this, it turns out, can also be used to describe subsystems of the universe: in particular, those resources can be used to define the wave functions that subsystems have.

To start, let N be the number of particles in the universe. For each i from 1 to N, let $\mathbf{q}_i$ be a variable which ranges over the candidate positions of the ith particle. Let $q = (\mathbf{q}_1, \ldots, \mathbf{q}_N)$ be a variable which ranges over the candidate configurations of all particles in the universe. For each time t, let $\Psi_t(q)$ be the universal wave function at that time. In addition, for each i, let $\mathbf{Q}_i(t)$ be the actual position of particle i at time t. Let $Q(t) = \big(\mathbf{Q}_1(t), \ldots, \mathbf{Q}_N(t)\big)$ denote the actual configuration, at time t, of the particles in the universe.

There is an important difference between the symbols 'q' and '$Q(t)$'. The former is a generic variable which ranges over all candidate configurations of the universe's particles. The latter is, for any given time t, a constant which denotes a single configuration of the particles in the universe: the configuration which the particles, at t, actually have. So in Bohmian mechanics, q acts as a generic symbol which can be used to specify the domain of all possible particle configurations to which the universal wave function Ψ_t assigns a complex number. $Q(t)$, however, specifies a specific particle configuration: the actual one at time t.

Bohmian mechanics posits two equations: one describes the evolution of the universal wave function $\Psi(t, q) = \Psi_t(q)$, while the other describes the evolution of particle configurations. The evolution of the universal wave function is given by the Schrödinger equation.

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

The evolution of particle configurations is given by the guidance equation.

$$\frac{d\mathbf{Q}_i}{dt} = \frac{\hbar}{m_i} Im \frac{\Psi^* \nabla_i \Psi}{\Psi^* \Psi}(\mathbf{Q}_1, \ldots, \mathbf{Q}_n) \qquad i = 1, \ldots, N \tag{2}$$

In (2), each m_i represents the mass of particle i. Together, (1) and (2) describe how the entire universe evolves.

In addition to describing the behavior of the universe as a whole, Bohmian mechanics also provides the resources needed to describe subsystems. A 'subsystem' is simply a collection of particles. The 'environment' of a subsystem consists of all particles in the universe which are not in that subsystem.

Subsystems and their environments can be represented by variables and constants, in the following way. Take any subsystem of M particles, where $M < N$. Index all

[1] For an early formulation of Bohmian mechanics, see (Bohm 1952a, b).

the particles in the universe so that particle 1, particle 2, ... , and particle M, are all and only the particles in this subsystem. For each i from 1 to M, let $\mathbf{x}_i = \mathbf{q}_i$: so each $\mathbf{x}_i$ is a variable which ranges over the candidate positions of particle i in the subsystem. Let $x = (\mathbf{x}_1, \ldots, \mathbf{x}_M)$ be a variable which ranges over the candidate configurations of the subsystem's particles. Similarly, for each i from $M + 1$ to N, let $\mathbf{y}_i = \mathbf{q}_i$. So each $\mathbf{y}_i$ is a variable which ranges over the candidate positions of particle i in the environment. Let $y = (\mathbf{y}_{M+1}, \ldots, \mathbf{y}_N)$ be a variable which ranges over the candidate configurations of the environment's particles. In addition, for each time t and each i from 1 to M, let $\mathbf{X}_i(t) = \mathbf{Q}_i(t)$: so for each time t, each $\mathbf{X}_i(t)$ is a constant which denotes the actual position of particle i in the subsystem. For each time t, let $X(t) = \big(\mathbf{X}_1(t), \ldots, \mathbf{X}_M(t)\big)$ be a constant which denotes the actual configuration, at t, of the subsystem as a whole. Finally, for each time t and each i from $M + 1$ to N, let $\mathbf{Y}_i(t) = \mathbf{Q}_i(t)$: so for each time t, each $\mathbf{Y}_i(t)$ is a constant which denotes the actual position of particle i in the environment. And for each time t, let $Y(t) = \big(\mathbf{Y}_{M+1}(t), \ldots, \mathbf{Y}_N(t)\big)$ be a constant which denotes the actual configuration, at t, of the environment as a whole.

Some more notation will be helpful. For any subsystem of M particles as described above, the variable q—which ranges over candidate configurations of the universe—may be rewritten as $q = (x, y)$. This equation conveniently represents the split between (i) the candidate configurations of the subsystem, and (ii) the candidate configurations of the environment. Similarly, for each time t, the constant $Q(t)$—which represents the actual configuration of the universe—may be rewritten as $Q(t) = \big(X(t), Y(t)\big)$. This equation conveniently represents the split between (i) the actual configuration of the subsystem, and (ii) the actual configuration of the environment.

These resources can be used to define a particular sort of wave function—called the 'conditional wave function'—for any given subsystem (Dürr et al. 1992, p. 864).[2] To see how, take the subsystem of M particles described above. Let x, Y, t, and Ψ_t be as defined earlier. Then for any given time t, the conditional wave function of this subsystem is the function $\psi_t(x)$ defined as follows.

$$\psi_t(x) = \Psi_t\big(x, Y(t)\big) \tag{3}$$

In other words, the wave function[3] of a given subsystem at a fixed time is obtained by (i) taking the actual positions of the particles in the subsystem's environment, and (ii) plugging those positions into the universal wave function.[4]

This feature of Bohmian mechanics—that it contains the resources required to formulate equation (3)—is striking. In more orthodox interpretations of the quantum

[2] For a more accessible account of wave function collapse, see (Goldstein 2010). For more discussion of different ways to interpret wave functions like these, and different ways to understand the physical significance of universal wave functions too, see (Goldstein and Zanghì 2013).

[3] Note that (3) is not normalized. This does not matter, however. All wave functions related by a constant non-vanishing multiple may be regarded as physically equivalent.

[4] For more on conditional wave functions, see (Dürr et al. 2004, pp. 966–968).

mechanical formalism, certain functions are simply stipulated to be the wave functions of subsystems. Subsystems' wave functions are not defined in terms of anything else. Similarly, in fact, for other interpretations of quantum mechanics, such as versions of the Everett interpretation. Bohmian mechanics, in contrast, can be used to define the wave functions of subsystems in terms of a few basic posits: the existence of a universal wave function, and the actual positions of the physically real particles which comprise the environment. So altogether, whereas Bohmian mechanics has the formal and physical resources to account for how certain wave functions are associated with certain subsystems, many other interpretations of quantum mechanics do not. And that is a significant point in favor of Bohmian mechanics.

3 Collapse

In this section, I discuss the Bohmian account of how the wave functions of subsystems—that is, conditional wave functions—collapse.[5] Then I briefly present the conditions under which conditional wave functions conform to a version of Schrödinger's equation. Finally, I explain why this version of Schrödinger's equation does not always describe how conditional wave functions evolve.

To start, here is the account of how conditional wave functions evolve in accord with the collapse postulate of quantum mechanics. Let t_1 be a time shortly before a measurement occurs. Suppose that at time t_1, the subsystem's conditional wave function $\psi_{t_1}(x)$ is in a superposition of the eigenstates $\psi_{t_1,\alpha_1}(x), \psi_{t_1,\alpha_2}(x), \ldots, \psi_{t_1,\alpha_n}(x)$ of the observable being measured. So for some constants $c_{\alpha_1}, c_{\alpha_2}, \ldots, c_{\alpha_n}$, the following holds.

$$\psi_{t_1}(x) = \sum_{\alpha=\alpha_0}^{\alpha_n} c_\alpha \psi_{t_1,\alpha}(x) \tag{4}$$

In addition, suppose that before measurement of the observable, the subsystem and the environment do not interact with one another. Moreover, let us assume that at time t_1, there is a function $\phi_{t_1}(y)$ such that the universal wave function is $\psi_{t_1}(x)\phi_{t_1}(y)$.[6] So (4) implies that at time t_1, the universal wave function is as follows.

$$\Psi_{t_1}(x, y) = \psi_{t_1}(x)\phi_{t_1}(y) = \sum_{\alpha=\alpha_0}^{\alpha_n} c_\alpha \psi_{t_1,\alpha}(x)\phi_{t_1}(y) \tag{5}$$

[5] This discussion is based on the theory developed in Dürr et al. (1992), Dürr and Teufel (2009), and Goldstein (2010).

[6] This assumption is unrealistic: the universal wave function generally does not factorize into a product state of functions $\psi_{t_1}(x)$ and $\phi_{t_1}(y)$. But as it turns out, this assumption is not really necessary, even approximately. It is sufficient that the universal wave function satisfies $\Psi_t(x, y) = \psi_{t_1}(x)\phi_{t_1}(y) + \Psi^{\perp}_{t_1}(x, y)$, where $\phi_{t_1}(y)$ and $\Psi^{\perp}_{t_1}(x, y)$ have macroscopically disjoint y-supports (Dürr et al. 1992 pp. 861–864).

In other words, before measurement, the universal wave function is in a superposition of the wave functions $\psi_{t_1,\alpha}(x)\phi_{t_1}(y)$, where each $\psi_{t_1,\alpha}(x)\phi_{t_1}(y)$ represents a universal wave function in which the subsystem's state is $\psi_{t_1,\alpha}(x)$ and the environment's state is $\phi_{t_1}(y)$.

Now to describe some post-measurement wave functions which will be relevant in what follows. Let t_2 be a time right after the measurement occurs. Take any α from $\alpha_0, \ldots, \alpha_n$. Then from t_1 to t_2, the wave function $\psi_{t_1,\alpha}(x)\phi_{t_1}(y)$ evolves to a new wave function $\psi_{t_2,\alpha}(x)\phi_{t_2,\alpha}(y)$, where $\psi_{t_2,\alpha}(x)$ and $\phi_{t_2,\alpha}(y)$ have several important properties. First, $\psi_{t_1,\alpha}(x) = \psi_{t_2,\alpha}(x)$: this corresponds to the fact that measurements of a given observables' eigenstates do not alter those eigenstates. Second, $\phi_{t_2,\alpha}(y)$ is the wave function associated with the environment recording the fact that the subsystem is in state $\psi_{t_2,\alpha}(x)$: that is just part of what it is for the event in question to count as a measurement of the observable in question.[7] Third, for all α' such that $\alpha \neq \alpha'$, the support of $\phi_{t_2,\alpha}(y)$ is macroscopically disjoint from the support of $\phi_{t_2,\alpha'}(y)$:[8] basically, this too is just part of what it is for the event in question to count as a measurement.[9] So each term $\psi_{t_1,\alpha}(x)\phi_{t_1}(y)$, in (5), evolves to a wave function $\psi_{t_2,\alpha}(x)\phi_{t_2,\alpha}(y)$ such that (i) $\psi_{t_1,\alpha}(x)$ is $\psi_{t_2,\alpha}(x)$, (ii) $\phi_{t_2,\alpha}(y)$ says that the system is in state $\psi_{t_2,\alpha}$, and (iii) the $\phi_{t_2,\alpha}(y)$ have macroscopically disjoint supports.

With all that as background, here is the Bohmian account of collapse. Since each $\psi_{t_1,\alpha}(x)\phi_{t_1}(y)$ evolves to $\psi_{t_2,\alpha}(x)\phi_{t_2,\alpha}(y)$, the linearity of the Schrödinger equation implies that the universal wave function in (5) evolves to the universal wave function below.

$$\Psi_{t_2}(x, y) = \sum_{\alpha=\alpha_0}^{\alpha_n} c_\alpha \psi_{t_2,\alpha}(x)\phi_{t_2,\alpha}(y) \tag{6}$$

By the definition of conditional wave functions from (3), the conditional wave function of the subsystem at time t_2 is obtained by substituting $Y(t_2)$—the actual configuration of the environment particles at t_2—for y in (6). The following results.

$$\begin{aligned}\psi_{t_2}(x) &= \Psi_{t_2}\big(x, Y(t_2)\big) \\ &= \sum_{\alpha=\alpha_0}^{\alpha_n} c_\alpha \psi_{t_2,\alpha}(x)\phi_{t_2,\alpha}\big(Y(t_2)\big)\end{aligned}$$

[7] For more discussion of why the state $\psi_{t_2,\alpha}$ must record the state of the subsystem—which is based on considerations of what it is to conduct a measurement—see (Albert 1992, pp. 74–79).

[8] In other words, if $\phi_{t_2,\alpha}(y)$ is non-zero for some configuration y, then $\phi_{t_2,\alpha'}(y)$ is zero for all configurations y from which that former configuration is macroscopically indistinguishable. And if $\phi_{t_2,\alpha'}(y)$ is non-zero for some configuration y, then $\phi_{t_2,\alpha}(y)$ is zero for all configurations y from which that former configuration is macroscopically indistinguishable. Put in intuitive terms, this all amounts to the following: the configurations of the environment, which record the outcome of the measurement, are macroscopically distinct from one another. In other words, the measurement device never enters a state in which it is somehow recording two distinct experimental outcomes.

[9] For more details, see (Dürr et al. 1992, pp. 863–866).

Since the functions $\phi_{t_2,\alpha}(y)$ all have macroscopically disjoint supports, at most one of the quantities $\phi_{t_2,\alpha}\big(Y(t_2)\big)$ is non-zero. Since each of the wave functions $\psi_{t_1,\alpha}(x)$ are eigenstates of the original conditional wave function $\psi_{t_1}(x)$, at least one of the quantities $\phi_{t_2,\alpha}\big(Y(t_2)\big)$ is non-zero. Therefore, for some α_j, the above sum reduces to $c_{\alpha_j}\psi_{t_2,\alpha_j}(x)\phi_{t_2,\alpha_j}\big(Y(t_2)\big)$ where $\phi_{t_2,\alpha_j}\big(Y(t_2)\big) \neq 0$. Dropping the unnecessary constant $\phi_{t_2,\alpha_j}\big(Y(t_2)\big)$[10]—and using the fact, mentioned earlier, that $\psi_{t_1,\alpha_j} = \psi_{t_2,\alpha_j}$—it follows that at time t_2, the conditional wave function of the subsystem is as follows.

$$\psi_{t_2}(x) = \psi_{t_1,\alpha_j}(x) \tag{7}$$

In other words, the conditional wave function of the subsystem after measurement is one of the eigenstates of the conditional wave function of the subsystem before measurement. The subsystem's conditional wave function has collapsed.

Basically, according to the Bohmian account, collapse results from two different features of subsystems, environments, and the universe as a whole. First, the wave function associated with any given subsystem is determined by (i) the wave function of the universe, and (ii) the actual positions of the environment particles. In other words, the wave function of any given subsystem is the conditional wave function given by (3). Second, after measurement, the universal wave function has the following property: when the post-measurement positions of the environment particles are plugged into the universal wave function, the resulting function is an eigenstate of the conditional wave function of the subsystem just before measurement. In slogan form: what it is to be a subsystem's wave function is to, among other things, exhibit collapse-like behavior.

Note that according to this account, collapse is real: the wave functions of subsystems really do, that is, undergo collapse. For when measurement occurs, a subsystem's wave function really does become an eigenstate of the wave function which the subsystem had before the measurement event. The subsystem starts out with one conditional wave function before measurement; after measurement, the subsystem's conditional wave function is an eigenstate of the conditional wave function from earlier. So collapse is a real, actual part of the physical world, according to the Bohmian account.

It is worth briefly explaining why conditional wave functions sometimes conform to a version of Schrödinger's equation. For conditional wave functions do not always collapse: they often exhibit Schrödinger evolution. Basically, that happens whenever the subsystem—corresponding to the conditional wave function in question—is suitably isolated from its environment.

For example, take the subsystem of M particles once more. Suppose that the universal wave function factorizes such that for all times t, there is a function Φ_t such that $\Psi_t(x, y) = \psi_t(x)\Phi_t(y)$; or at least, suppose that the universal wave function

[10] Recall that as mentioned in Footnote 3, wave functions related by a constant non-vanishing multiple are physically equivalent.

approximately obeys an equation of this form.[11] In addition, suppose that there is negligible interaction between the subsystem and the environment; so the universal Hamitonian H may be written as $H = H^x + H^y$.[12] Finally, let ψ be defined by $\psi(t, x) = \psi_t(x)$. Then it can be shown that the following holds.[13]

$$i\hbar\frac{\partial\psi}{\partial t} = H^x\psi \tag{8}$$

The conditional wave functions of subsystems for which the corresponding universal wave functions are approximately in product states, in other words, conform to a version of Schrödinger's equation.

Before concluding, it is worth discussing two reasons why the Bohmian account of collapse is preferable to the account of collapse that orthodox quantum mechanics endorses. First, that other account—call it the 'orthodox account'—simply stipulates that collapse occurs. The phenomenon of collapse, in other words, is a primitive posit of the orthodox account. The Bohmian account, however, does not merely posit a collapse principle. Instead, the Bohmian account shows how collapse derives from other, more basic posits: namely, posits about actual configurations and universal wave functions.

Second, and relatedly, the orthodox account does not offer a clear method for associating wave functions with subsystems in the first place.[14] To illustrate, consider the following question: for any given subsystem of the universe, at any given time t, what wave function should be associated with that subsystem? The answer to this question, that Bohmian mechanics supports, is clear: given that (i) the wave function of the universe at t is Ψ_t, (ii) the subsystem in question is defined as the collection of particles with actual positions $\mathbf{X}_1(t), \ldots, \mathbf{X}_M(t)$, so that the environment particles have actual positions $\mathbf{Y}_{M+1}(t), \ldots, \mathbf{Y}_N(t)$, and (iii) a condition about macroscopically disjoint supports of the sort mentioned earlier obtains, it follows that (iv) the wave function which should be associated with this subsystem is the conditional wave function $\psi_t(x) = \Psi_t\big(x, Y(t)\big) = \Psi_t\big(\mathbf{x}_1, \ldots, \mathbf{x}_M, \mathbf{Y}_{M+1}(t), \ldots, \mathbf{Y}_N(t)\big)$. Orthodox quantum mechanics does not support an analogously clear answer to this question. For orthodox quantum mechanics does not provide clear, precise principles which, for any given subsystem, define the wave function associated with that subsystem in terms of anything as well-defined as actual configurations and universal wave functions. And so whereas Bohmian mechanics can be used to provide a satisfying answer to this question, orthodox quantum mechanics cannot.

---

[11] For the reasons mentioned in Footnote 6, this assumption is unrealistic, but not necessary. It is made here merely in order to simplify the discussion.

[12] H^x is the contribution to H arising from terms involving only degrees of freedom from particles in the subsystem, while H^y is the contribution to H arising from terms involving only degrees of freedom from particles in the environment.

[13] See (Dürr et al. 1992, pp. 861–862).

[14] As mentioned earlier, this is arguably true for other accounts too, like accounts of collapse suggested by some versions of the Everett interpretation.

This is, in my view, one of the most attractive features of Bohmian mechanics. It provides precisely the resources needed to clearly define the wave functions which should be associated with subsystems: those resources consist of a few simple posits about particles and a universal wave function. And in so doing, it supports an account of how collapse occurs.

In short, by helping itself to physically real particles, Bohmian mechanics clarifies decades of confusion surrounding quantum collapse. The orthodox account exacerbates that confusion, since it resists positing an actual configuration for any given subsystem's environment: so given the orthodox account, there is nothing to plug into a universal wave function, to obtain the wave functions associated with subsystems—that is, according to the Bohmian account, the conditional wave functions—which undergo collapse. Bohmian mechanics does posit an actual configuration for each subsystem's environment, however. And as a result, Bohmian mechanics supports an illuminating account of how, and why, collapse occurs.

4 Conclusion

Bohmian mechanics can be used to provide an attractive, elegant, and simple account of collapse. The account says, basically, that collapse is a consequence of how conditional wave functions evolve over time. Their evolution generates the phenomenon of collapse because of how the environment particles, and the universal wave function, evolve.

Acknowledgements Thanks to Laura Ruetsche, and especially Shelly Goldstein, for much helpful feedback and discussion.

References

D. Albert, *Quantum Mechanics and Experience* (Harvard University Press, Cambridge, MA, 1992)
D. Bohm, A suggested interpretation of the quantum theory in terms of "Hidden" variables. I. Physical Review **85**(2), 166–179 (1952)
D. Bohm, A suggested interpretation of the quantum theory in terms of "Hidden" variables. II. Physical Review **85**(2), 180–193 (1952)
D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the origin of absolute uncertainty. J. Stat. Phys. **67**(5/6), 843–907 (1992)
D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the role of operators as observables in quantum theory. J. Stat. Phys. **116**(1/4), 959–1055 (2004)
D. Dürr, S. Teufel, *Bohmian Mechanics* (Springer, New York, NY, 2009)
S. Goldstein, Bohmian mechanics and quantum information. Found. Phys. **40**, 335–355 (2010)
S. Goldstein, N. Zanghì, Reality and the role of the wave function in quantum theory, in *The Wave Function*. ed. by D. Albert, A. Ney (Oxford University Press, New York, NY, 2013), pp. 91–109

Generic Contextuality

Travis Norsen

Abstract This paper reviews and develops the concept of "contextuality" whose profound significance—or lack thereof—was clarified especially by Detlef Dürr and his collaborators. In particular, we explore, in the context of a simple toy model of a measurement procedure described using Bohmian mechanics, the dependence of measurement outcomes on the (continuously variable) strength of the coupling between the system and measuring apparatus. This provides a revealing illustration of the fact that the outcomes of experiments may, and in general do, depend on details of the experiment other than simply the Hermitian operator which (as it is often misleadingly said) is "measured" in the experiment.

1 Context

Unlike many of the contributors to this volume, I was never (formally) a student of Detlef's, I did not collaborate with him on many projects across the years and decades, and I never even managed to visit him in Munich. Nevertheless, I learned a tremendous amount from his books and papers over the years, I was lucky enough to work with him on one brief project a few years ago, and I had the enormous pleasure of getting to know him via email discussions and, especially, in person, at a number of conferences and summer schools over the years. I considered him a friend and a mentor and I will never forget his kindness, his clarity, and especially his charming smile and glorious sense of humor.

I have included, at random points in the following paper and without any further explanation, what I think of as "Detlefisms"—short quotes (taken from his book on Bohmian Mechanics [1]) whose simultaneously sarcastic and reassuring rhetorical style is unmistakably Detlef's. When I read these, I cannot help but hear them in his voice while picturing his smiling face and twinkling eyes. I hope their inclusion here will trigger the same experience in others and thus help us all remember and appreciate Detlef's unique spirit.

T. Norsen (✉)
Smith College, Northampton, MA 01060, USA
e-mail: tnorsen@smith.edu

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_7

2 Contextuality

In the orthodox interpretation of quantum mechanics, quantum systems are claimed to be completely described in terms of wave functions. In ordinary circumstances, these wave functions evolve in time in accordance with Schrödinger's equation. But when a quantum system is *measured*, different rules apply: the system's wave function is postulated to "collapse" (a non-unitary change that is incompatible with the usual Schrödinger evolution) to one of the eigenfunctions of the Hermitian operator corresponding to the property of the system which is being measured, and the outcome of the measurement (corresponding to the particular eigenstate that is collapsed to) is said to be irreducibly random. This randomness is of course the locus in the orthodox theory of the idea, which many commentators regard as one of the central metaphysical innovations of the theory, that quantum physics refutes determinism.

> But that is not a very intelligent thing to say.

Critics of the orthodox interpretation [2, 3]—motivated both by the implausibility of having incompatible fundamental laws for "ordinary" and "measurement" situations, and also by the apparent non-locality that accompanied the measurement-induced wave function collapse for systems of spatially-separated but entangled particles—hypothesized the existence of so-called (and misleadingly named) "hidden variables" (HVs). HVs are simply properties that quantum systems are postulated to possess in addition to the wave function. Typically (though not necessarily) the HVs were assumed to *determine* the outcomes of measurements. From this point of view, the stochasticity of the orthodox interpretation was simply a result of its utilization of *incomplete* descriptions of the states of systems.

Various defenders of the orthodox viewpoint (most notably von Neumann [4] and then Kochen and Specker [5]) attempted to prove mathematically that deterministic HV theories could not reproduce the (apparently correct) empirical predictions of quantum mechanics. But as pointed out especially by Bell [6], the no-hidden-variables proofs invariably imposed arbitrary and unwarranted—and often unacknowledged—assumptions. The most interesting and important such assumption was that the hidden variables be "non-contextual". A non-contextual HV theory is one in which a definite value, to be revealed upon measurement, is assigned for the Hermitian operators which correspond to properties of the system which might be measured.

The basis for the terminology here is the idea that a given measurement, say of the property corresponding to Hermitian operator $\hat{A}$, may happen in several distinct ways. For example, if $\hat{A}$ and $\hat{B}$ commute, $[\hat{A}, \hat{B}] = 0$, (the properties corresponding to) $\hat{A}$ and $\hat{B}$ can in principle be measured simultaneously; doing so would constitute one possible way of measuring (the property corresponding to) $\hat{A}$. Alternatively, (the property corresponding to) $\hat{A}$ could be measured by measuring $\hat{A}$ and $\hat{C}$ simultaneously (again assuming $[\hat{A}, \hat{C}] = 0$). A non-contextual hidden variable theory is then one in which the value assigned to $\hat{A}$ – the value that will be the outcome of an experiment measuring (the property corresponding to) $\hat{A}$ – is the same, regardless of

the so-called "context" of the measurement, i.e., whether $\hat{B}$ or $\hat{C}$ (or neither or some other thing) is measured in conjunction with $\hat{A}$. (Note, by the way, that here $\hat{B}$ and $\hat{C}$ need not commute, so that the experiments of measuring $\hat{A}$ in conjunction with $\hat{B}$, and of measuring $\hat{A}$ in conjunction with $\hat{C}$, are incompatible.)

In a "contextual" hidden-variable theory, by contrast, definite values—to be revealed by the appropriate experiment—are not assigned to Hermitian operators *per se*, but instead to the more detailed and varied experiments by which (the properties corresponding to) a given Hermitian operator might be measured. For example, a contextual hidden-variable theory might have "the value of $\hat{A}$" (meaning, of course, the outcome of an experiment to measure the property corresponding to $\hat{A}$) when $\hat{A}$ is measured in conjunction with $\hat{B}$, being different from "the value of $\hat{A}$" when $\hat{A}$ is measured in conjunction with $\hat{C}$. "The value of $\hat{A}$", in short, can depend on the full context of the specific way in which $\hat{A}$ is measured.

Is there anything deep in all this? Well, no, there is not.

The assumption that HV theories should be non-contextual may seem innocuous; evidently it appeared so innocent to the authors of the no-hidden-variables theorems that many of them barely even acknowledged the assumption, and the idea of contextual hidden variables is still looked upon as somehow suspicious by many commentators [7].

But the supposed weirdness and implausibility of contextuality is belied by the naturalness and simplicity with which this feature arises in the context of the one extant hidden variable theory, the pilot-wave theory of de Broglie and Bohm. The most well-known example is probably the way in which the pilot-wave theory accounts for the familiar correlations from the EPR-Bell setup. To connect with the above discussion, if $\hat{A}$ here is (the operator corresponding to) the component of the spin of particle 1 along some direction, and $\hat{B}$ and $\hat{C}$ are respectively two different components of the spin of (the entangled but spatially separated) particle 2, the pilot-wave theory just straightforwardly predicts (on the basis of its postulated dynamics) that the outcome of measuring $\hat{A}$ in conjunction with $\hat{B}$ can, in appropriate circumstances, be different from the outcome of measuring $\hat{A}$ in conjunction with $\hat{C}$ (even when the initial state of the measured system is exactly the same). This example was discussed already by Bell in his analysis of what had gone wrong with the no-hidden-variables proofs [6]; for a more recent pedagogical discussion see Ref. [8].

Another concrete example of the utter straightforwardness with which the pilot-wave theory exhibits contextuality was pointed out by Albert [9]: the outcome of a measurement of a generic component of the spin of a particle can depend on which of two equally valid Stern-Gerlach devices for measuring that spin component are used. Concretely, for certain appropriate initial conditions of the incoming particle, a Stern-Gerlach device with its magnetic field oriented in a certain direction will deflect the particle upward and hence reveal the particle to be "spin up" along that direction. But for precisely the same initial conditions, a Stern-Gerlach device with its magnetic field oriented in the opposite direction will *also* deflect the particle upward and hence (due to the opposite calibration of the device associated with the

oppositely-directed field) reveal the particle to be instead "spin down" along that same axis. (See again [8] for a more detailed discussion.)

Albert's example is particularly nice for a number of reasons. First, unlike the EPR example mentioned above (in which the relevant "context" is the setting of a distant apparatus such that the contextuality is also an example of the theory's well-known non-locality), here we have a concrete example of pure contextuality without the interesting, but also somewhat distracting, feature of non-locality. Second, and relatedly, Albert's example involves only a single measurement. It thus shows clearly that the context-dependence of experimental outcomes need not involve additional measurements that are made in conjunction with the one for the property in question. Instead we simply have distinct measurement outcomes for distinct, but fully and equally valid, realizations of experiments to measure (the property associated with) the same one operator. So the example, in this sense as well, illustrates the heart of what contextuality means (as opposed to merely illustrating the particular sort of contextuality that had been formally excluded in the no-HV theorems).

Some other types of examples—not involving spin—of the pilot-wave theory's contextuality have also been given (see, e.g., Sect. 7.5 of Ref. [10]), but these are slightly less dramatic than the above and tend to less clearly convey the appropriate lesson, which has been put very clearly and forcefully by Detlef et al. as follows:

> In foundations of quantum mechanics circles this situation is referred to as *contextuality*, but we believe that this terminology, while quite appropriate, somehow fails to convey with sufficient force the rather definitive character of what it entails: Properties that are merely contextual are not properties at all; they do not exist, and their failure to do so is in the strongest sense possible! We thus believe that contextuality reflects little more than the rather obvious observation that the result of an experiment should depend upon how it is performed! [11]

The experiments we have previously spoken of as "measuring" properties that turn out to be contextual, therefore, should really instead not be thought of as "measurements" at all, if that word means an experimental intervention which reveals the pre-existing value of some property. They don't. But as revealed so clearly when we examine such experiments as described by the pilot-wave theory, there is nothing the least bit mysterious about this.

> The price to pay is that we need to be careful
> not to treat things which are not the same
> as being the same.

Our goal in the remainder of this paper is to explore a novel manifestation of so-called contextuality, for the pilot-wave theory, which we think helps reinforce and broaden this important lesson. Compared to Albert's example involving spin, the manifestation of contextuality that we will discuss has the virtues of being continuous as opposed to discrete (that is, instead of two distinct outcomes for two different experimental realizations, we see a range of distinct measurement outcomes as a continuously variable parameter is adjusted) and also of being straightforwardly generalizable to virtually any experiment of the sort that is traditionally, but we now know misleadingly, described as a measurement of some property of the system.

All of this will be discussed in more depth in Sect. 5, after first, in Sect. 3 laying out a concrete toy model and, in Sect. 4, displaying the results of some numerical simulations of it.

3 Toy Model of an Energy "Measurement"

One of the fundamentally important features of the pilot-wave theory is that it solves the measurement problem—that is, it allows us to view the processes we describe as "measurements" as examples of (rather than exceptions to) the standard dynamical principles of the theory. We can, in the pilot-wave theory, *analyze* measurements.

So if contextuality means that the outcome of a "measurement" (of, say, the property corresponding to some operator $\hat{A}$) can depend on details of the specific experimental realization of that "measurement" (rather than just on the operator $\hat{A}$ which is "measured"), it stands to reason that we could find a wealth of examples of contextuality by simply analyzing generic measurements from the point of view of the pilot-wave theory.

Here is a way to see that we can indeed do this.

Let us take, as our quantum system, a one-dimensional, length-L "particle in a box" (PIB) with Hamiltonian $\hat{H}_x = \frac{\hat{p}_x^2}{2m} + V(x)$ with $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$ and

$$V(x) = \begin{cases} 0 & \text{for } 0 < x < L \\ \infty & \text{otherwise} \end{cases} . \tag{1}$$

The normalized energy eigenstates are given by $\psi_n(x) = \sqrt{\frac{2}{L}}\sin(n\pi x/L)$ with corresponding energy eigenvalues $E_n = \frac{\hbar^2\pi^2 n^2}{2mL^2}$.

Now consider a measurement of the PIB's energy, as depicted schematically in Fig. 1. Following von Neumann, we treat the apparatus pointer (whose final position indicates the outcome of the energy measurement) as an explicit degree of freedom which couples to the PIB via the interaction Hamiltonian

$$H_{int} = \lambda \hat{H}_x \hat{p}_y \tag{2}$$

with $\hat{p}_y = -i\hbar\dfrac{\partial}{\partial y}$. The λ here is simply a (real) constant which parameterizes the strength of the PIB-apparatus coupling.

For simplicity we take the apparatus pointer to be infinitely heavy so that the Hamiltonian for the entire setup, the PIB and the measuring apparatus, is

$$\hat{H} = \hat{H}_x + \hat{H}_{int}. \tag{3}$$

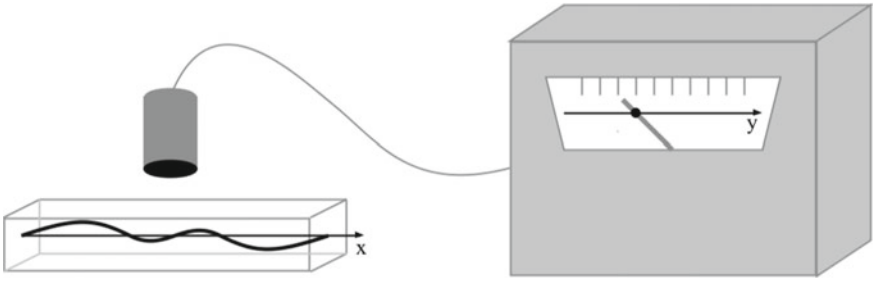


Fig. 1 Schematic depiction of the measurement setup we use here to illustrate the idea of generic contextuality. On the left is a one-dimensional "particle in a box" with degree of freedom x, whose energy is measured by the apparatus, shown on the right, whose pointer has coordinate y

The Schrödinger equation, $i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi$ for the joint system-apparatus wave function $\Psi = \Psi(x, y, t)$ thus reads

$$\frac{\partial\Psi}{\partial t} = \frac{i\hbar}{2m}\frac{\partial^2}{\partial x^2}\Psi + \lambda\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\frac{\partial}{\partial y}\Psi. \tag{4}$$

If the initial wave function $\Psi(x, y, 0)$ is a product of the energy eigenstate $\psi_n(x)$ for the PIB and a (say, width-σ Gaussian) wave packet $\phi(y)$ centered at the "ready" position of the apparatus pointer, it is easy to see that, at time t, the wave function will be

$$\Psi(x, y, t) = \psi_n(x)\, e^{-iE_n t/\hbar}\, \phi(y - \lambda E_n t). \tag{5}$$

Upon completion of the measurement at $t = T$, the support of the wave function, in the apparatus pointer degree of freedom, will be centered around the value $y_n = \lambda T E_n$. During the measurement, the pointer thus moves, to the right, a distance proportional to E_n and is thereby registering, in its final position, the expected outcome of the energy measurement.

And since the Schrödinger equation is linear and homogeneous, a general initial state $\Psi(x, y, 0) = \left(\sum_j c_j\psi_j(x)\right)\phi(y)$ (in which the PIB is in some arbitrary superposition of energy eigenstates) will time-evolve into

$$\Psi(x, y, t) = \sum_j c_j\psi_j(x)\, e^{-iE_j t/\hbar}\, \phi(y - \lambda E_j t) \tag{6}$$

which is an entangled superposition.

Of course, in the pilot-wave theory, the wave function alone does not provide a complete description of the state of the system. There are, in addition, the actual particles (corresponding to the PIB and apparatus pointer) with positions $X(t)$ and $Y(t)$ which evolve in time according to

$$\frac{dX}{dt} = \left.\frac{j_x}{|\Psi|^2}\right|_{x=X(t),y=Y(t)} \tag{7}$$

and

$$\frac{dY}{dt} = \left.\frac{j_y}{|\Psi|^2}\right|_{x=X(t),y=Y(t)} \tag{8}$$

where j_x and j_y are the components of the standard quantum probability current satisfying the continuity equation,

$$\frac{\partial}{\partial t}|\Psi|^2 = -\left(\frac{\partial}{\partial x}j_x + \frac{\partial}{\partial y}j_y\right). \tag{9}$$

The somewhat unfamiliar structure of the interaction Hamiltonian $\hat{H}_{int}$ gives the probability current a somewhat unusual form, but it is easy enough to verify that

$$j_x = \frac{\hbar}{2mi}\left(\Psi^*\frac{\partial\Psi}{\partial x} - \Psi\frac{\partial\Psi^*}{\partial x}\right) - \frac{\lambda\hbar^2}{2m}\left(\Psi^*\frac{\partial}{\partial x}\frac{\partial}{\partial y}\Psi + \Psi\frac{\partial}{\partial x}\frac{\partial}{\partial y}\Psi^*\right) \tag{10}$$

and

$$j_y = \frac{\lambda\hbar^2}{2m}\frac{\partial\Psi^*}{\partial x}\frac{\partial\Psi}{\partial x} \tag{11}$$

indeed satisfy the above continuity equation when Ψ evolves according to Equation (4).[1]

During the measurement interaction, the wave function Ψ branches into a number of (for an appropriately large T, approximately) non-overlapping packets. The particle configuration (X, Y) will end up in the support of one of these packets, and the outcome of the measurement will be indicated in particular by the final position of the apparatus pointer: if $Y(T)$ is near y_n, we will declare the outcome of the experiment (to "measure" the energy of the PIB) to be E_n. See Fig. 2 for a cartoon representation.

The dynamics here is completely deterministic, so the final state—comprising $X(T)$, $Y(T)$, and $\Psi(x, y, T)$, and so in particular the outcome of the measurement—depends, for a given value of λ, on the initial state and in particular (since the initial wave function is presumed fixed) on the initial particle configuration $(X(0), Y(0)) = (X_0, Y_0)$. The theory reproduces the quantum statistics—for example, that the outcome of the energy measurement should be E_n with relative frequency $|c_n|^2$—by treating the initial particle configuration as random with distribution $|\Psi|^2$.

[1] As usual, the probability current is not unique. Any divergenceless term could be added and the result would still satisfy the continuity equation. Here and in subsequent discussions we simply take the simplest possibility. There are contexts in which it is quite reasonable to consider other possibilities—see, e.g., Ref. [12]—but here we do not expect our central qualitative conclusions to be affected by this issue.

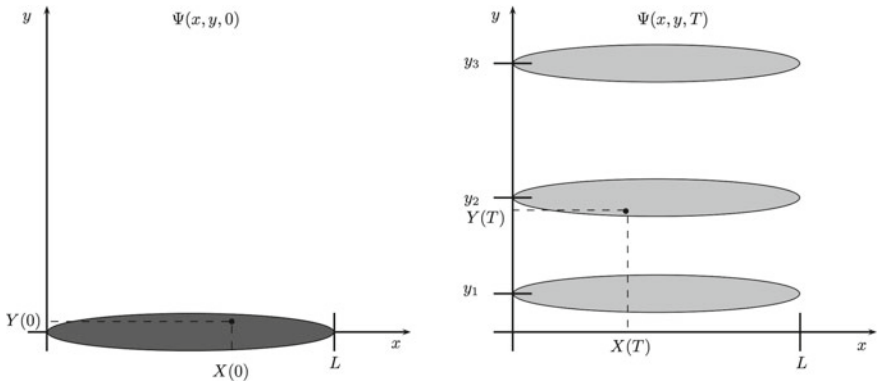


Fig. 2 The dark grey blob on the left shows the region of the two-dimensional configuration space where the initial wave function $\Psi(x, y, 0)$ has support, with the black dot representing a possible initial particle configuration, $(X(0), Y(0))$. The time-evolution splits the wave function into several approximately non-overlapping islands or branches, as shown on the right, with the particle configuration ending up in one of the branches. In the case shown, with the final pointer position $Y(T)$ being approximately equal to y_2, we would say that the outcome of the measurement of the PIB's energy was E_2

The interesting feature that we wish to explore here pertains to how the mapping from (X_0, Y_0) onto the measurement outcome, indicated by $Y(T)$, varies with λ. It should be clear that, for any nonzero value of λ, the system described constitutes (what is ordinarily called) a measurement of $\hat{H}_x$, the energy of the PIB. Each different value of the continuously-variable parameter λ, however, corresponds to a distinct way for the measurement to be instantiated.

In the following section we will show, via numerical solution of the equations of motion, that the outcome—the measured energy of the PIB—can vary with the PIB-apparatus interaction strength λ even as the detailed initial state of the PIB-apparatus system remains fixed. Before turning to that, however, it is worth thinking about the large-λ limit. For large λ, the terms not involving λ on the right hand sides of Equations (4) and (10) can be neglected. In the resulting equations of motion, only the product λt appears, which means that varying λ will not change the character of the solutions and, in particular, will not change the measurement outcome that results from a given detailed initial condition—it will only change the speed with which the experiment proceeds toward completion.[2]

[2] Note that this contradicts the implication of Project 7.9 of Ref. [10], which turns out to have been based on a coding error. The author regrets the earlier mistake and thanks Tim Maudlin, private communication, for expressing skepticism about the earlier claim. It is hoped that the present paper clarifies precisely the situations under which this λ-dependence type of contextuality does and does not arise.

The novel generic sort of contextuality we will demonstrate, then—that is, the dependence of the measurement outcome on the strength of the system-apparatus coupling—arises only for "non-impulsive" measurements with finite λ, i.e., measurements which have the realistic feature of taking a finite time to proceed to completion.

4 Numerical Illustration

Working in units where $\hbar = m = L = \sigma = 1$, we solve Equations (7) and (8) numerically with the wave function given by Equation (6) with $c_1 = 1.0$, $c_2 = -0.8$, $c_3 = 0.7i$, and $c_j = 0$ for $j > 3$. (These particular values were chosen solely on the grounds that they led to pretty pictures; the phenomenon we explore is obviously general.) For a given value of λ, we solve from $t = 0$ to $t = T$ where $T = 0.5/\lambda$ so that the separation of the wave packets in configuration space at the final time (and hence the correspondence between $Y(T)$ and measurement outcomes) is identical even as λ is varied.

There are thus three possible outcomes of the energy measurement: E_1 (indicated by $Y(T) \approx 2.5$), E_2 (indicated by $Y(T) \approx 9.9$), and E_3 (indicated by $Y(T) \approx 22.2$).

For a particular value of λ, we solve the equations of motion to find $Y(T)$ on a rectangular grid of initial particle configuration points (X_0, Y_0) within the support of $\Psi(x, y, 0)$. We then display the results by coloring the points (X_0, Y_0) which lead to measurement outcome E_1 (i.e., $Y(T) \approx 2.5$) blue, E_2 yellow, and E_3 red. An example plot, for $\lambda = 1.0$, is shown in Fig. 3.

Figure 4 shows similar maps for several values of $\lambda > 1$. Note that while the images are not all identical, they are clearly converging in the large-λ limit. This illustrates the point, mentioned at the end of the previous section, that the

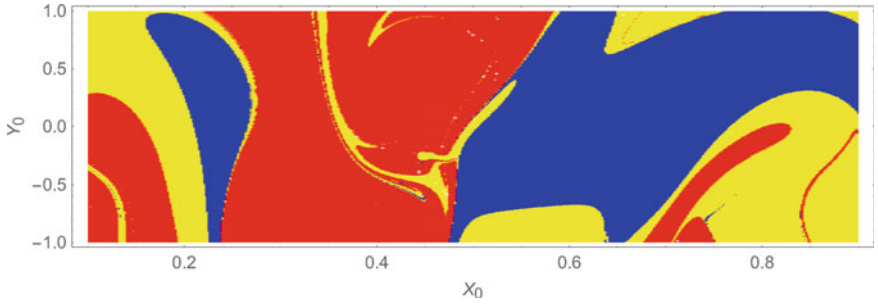
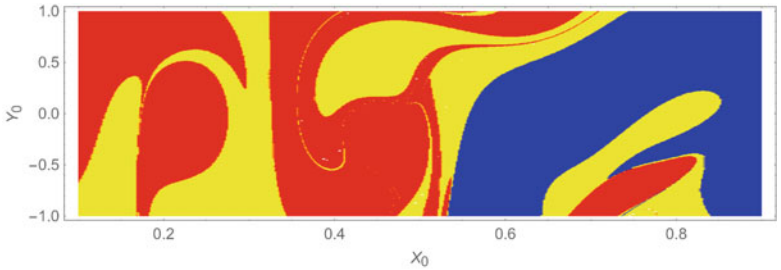
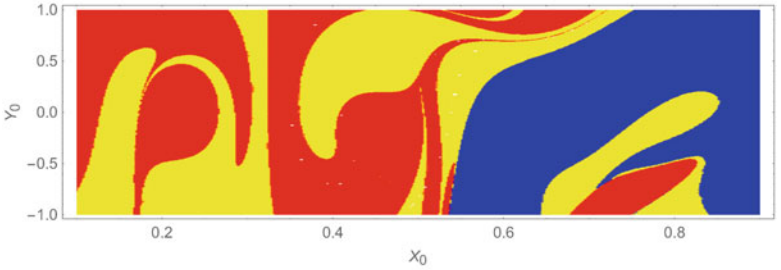


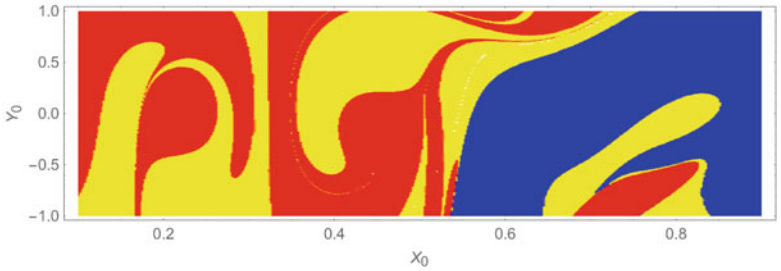
Fig. 3 In this image, each point in the configuration space of PIB-apparatus initial conditions, (X_0, Y_0), is color-coded to represent the outcome of the energy measurement that will result, with $\lambda = 1.0$, from those particlar initial positions for the PIB and the apparatus pointer. Initial particle positions that yield the outcome E_1 (meaning that $Y(T) \approx y_1$) are colored blue, yellow is E_2, and red is E_3



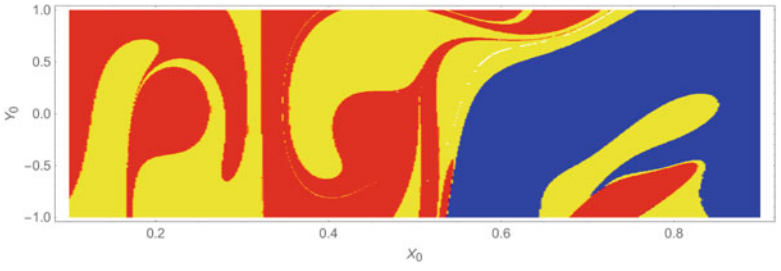
(a) $\lambda = 10.0$



(b) $\lambda = 30.0$



(c) $\lambda = 100.0$



(d) $\lambda = 300.0$

Fig. 4 Maps of the measurement outcomes for several larger values of λ. Note that, as expected, in the large-λ limit the results are visually identical—i.e., the λ-dependence, the contextuality, disappears

measurement outcome for a given initial configuration is not sensitive to λ if λ is so large that $\hat{H}_{int}$ swamps all other terms in the overall Hamiltonian.

The novel type of contextuality that it is our main goal here to highlight is thus best seen for smaller values of λ.

The reader is encouraged to examine Fig. 5, which shows outcome maps for (a) $\lambda = 0.3$, (b) $\lambda = 0.1$, (c) $\lambda = 0.03$, and (d) $\lambda = 0.01$.

The contextuality we wish to highlight is manifest simply in the fact that the images in Fig. 5 are different from one another. This means that, for generic initial conditions of the PIB-apparatus system, the outcome of the measurement of the PIB's energy is different depending on details of the specific experiment (in particular, the strength of the PIB-apparatus coupling, λ) even though all of these experiments are equally good, equally valid measurements of the property corresponding to the same operator, here $\hat{H}_x$.

So it is a nice concrete illustration of the phenomenon of contextuality, i.e., of the fact that "the result of an experiment should depend upon how it is performed!" [11].

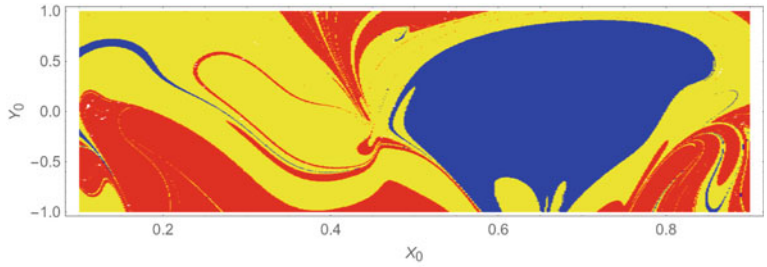
5 Discussion

We have shown that, in the pilot-wave theory, the outcome of a (realistic, non-impulsive) "measurement" of the energy of a simple system can vary depending on the precise value of the coupling constant that controls the strength of the system-apparatus interaction.

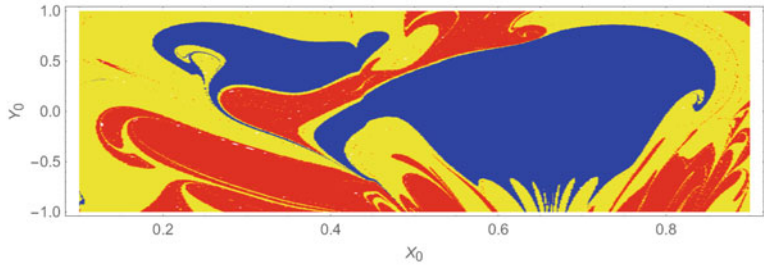
This already has the interesting implication that, according to the pilot-wave theory, energy—just like the various components of particles' spins—"do[es] not exist" [11].

This will not be surprising to those familiar with the pilot-wave theory and in particular the idea that, for this theory, every property other than position is "contextual"—i.e., not really a property at all, not (therefore) something that can be "measured" (if we accept the usual connotations of that word). And even from the point of view of orthodox quantum theory, it is not terribly surprising that a system which is in a superposition of energy eigenstates with several different energies, should fail to possess any particular amount of energy.

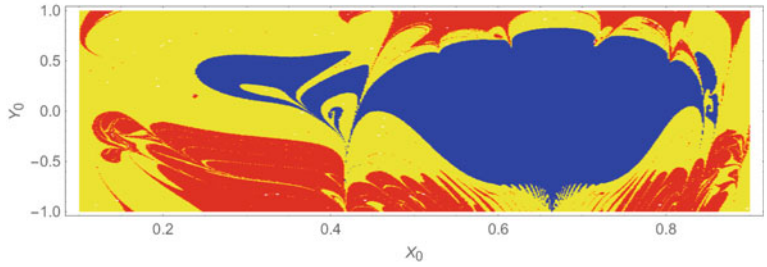
Nevertheless, the commonly-used language—that an experiment of the sort we have been analyzing can be thought of as a "measurement of the energy"—tends to almost irresistably suggest that, although we might not have *known* and/or couldn't formally *define* the amount of energy possessed by the system prior to our experimental intervention, when we "measure the energy" of the system we are, surely, finding out how much energy it in fact had. And beyond arguably-misleading terminology, the idea that conservation of energy is a foundational bedrock principle of physics also contributes to the feeling that, while maybe it is tolerable for spin components to be contextual and hence unreal, we must, surely, take energy much more seriously.



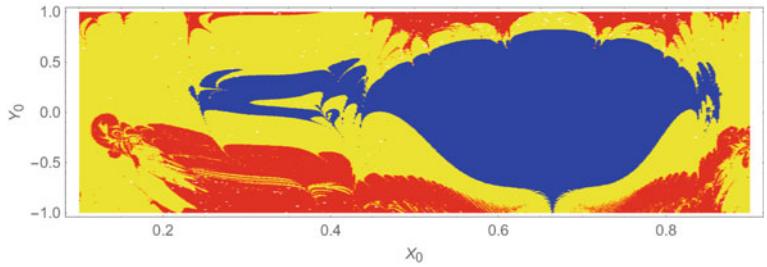
(a) $\lambda = 0.3$



(b) $\lambda = 0.1$



(c) $\lambda = 0.03$



(d) $\lambda = 0.01$

Fig. 5 Maps of the measurement outcomes for several smaller values of λ. The essential point is simply that the pictures are different: for many possible initial conditions of the PIB-apparatus system, distinct (non-large) values of λ produce distinct measurement outcomes

Witness, for example, the struggle to make sense of the conservation of energy (in a not-too-dissimilar setup) exhibited recently in Ref. [13].

Trivial it is, but food for mysticism nevertheless.

But we can and should resist the misleading connotations of the word "measurement" [14]. If several different but equally valid ways of "measuring the energy" of a system yield different outcomes, then none of them are actually measuring the energy of that system and indeed there is no reason to believe there is any such property as the energy of that system.

Indeed, note that the worry about exactly how to define and understand the conservation of energy in QM [13] is simply dissolved by facing the fact that, like spin, energy is contextual, i.e., unreal, at least according to the pilot-wave theory.

But this is no world-shattering discovery,
for there is absolutely no reason why it should be the case.

And it is not just energy. It should be clear not only that our analysis could have used virtually any example quantum system in virtually any state, but that we could have replaced $\hat{H}_{int}$ with $\lambda\hat{A}\hat{p}_y$—i.e., we could have analyzed a "measurement" of many other "properties", with similar results. There is nothing particularly special here about energy. We could seemingly demonstrate this same sort of contextuality, in an appropriate setup, for other operators.

And, interestingly, this is true even for position!

Consider, for example, a system with degree of freedom x and Hamiltonian $\hat{H}_x$ which couples to a measuring apparatus whose pointer has coordinate y via the interaction Hamiltonian $H_{int} = \lambda\hat{x}\hat{p}_y$. (As before, for simplicity, we assume the pointer is infinitely massive so $\hat{H}_y = 0$.) According to the pilot-wave theory, the x-system particle will evolve in the usual way, e.g.,

$$\frac{dX(t)}{dt} = \frac{\hbar}{m}\mathrm{Im}\left[\frac{\partial\Psi}{\partial x}/\Psi\right]_{x=X(t),\, y=Y(t)} \tag{12}$$

if $\hat{H}_x = \frac{\hat{p}_x^2}{2m} + V(x)$. That is, for $\hat{H}_{int} = \lambda\hat{x}\hat{p}_y$ there is no term in j_x proportional to λ.

What is interesting is that the position of the pointer evolves according to

$$\frac{dY(t)}{dt} = \lambda X(t). \tag{13}$$

For an impulsive measurement in which the coupling strength λ is large and the required duration of the interaction $T \sim 1/\lambda$ is therefore small, the pointer will simply move a distance ΔY proportional to the actual position X of the system particle at the "moment" of the interaction. That is, in the large-λ limit, the final pointer position does indeed register the actual position of the particle whose position is being measured.

But in general—for realistic "position measurements" with finite λ—$X(t)$ will vary during the course of the experiment and the displacement of the pointer (i.e., the outcome of the measurement) will be proportional to the system particle's *average* position during the course of the interaction:

$$\Delta Y = \lambda\, T \langle X \rangle \tag{14}$$

where $\langle X \rangle = \frac{1}{T}\int_0^T X(t)dt$. To be sure, this is as sensible an outcome as one might hope for. But, strictly speaking, it does show that, in the pilot-wave theory, even (realistic, non-impulsive) position measurements are contextual—i.e., they do not simply reveal the pre-measurement value of the position of the particle; different experiments to "measure the position" (here, the different experiments parameterized by different values of λ) will yield different outcomes even for identical initial conditions.

And in this way, many irrelevant questions arise.

Does this mean that (like we have said about spin components and energy and by implication generic other properties) particle positions don't really exist according to the pilot-wave theory?

Of course not. The theory postulates that they exist, so according to the theory they exist!

It's just that, in general, realistic "position measurements" (just like all other experiments) involve a definite, finite interaction between the system and the measuring apparatus. They stretch across a finite time period during which, so to speak, the system sloshes around a bit. And the outcome of the experiment depends, in general, at least a bit on the details and extent of this sloshing. That the outcome fails to somehow magically reveal the exact position of the system particle, say, just prior to the initiation of the system-apparatus interaction, is as problematic as it is shocking... namely: not at all!

Still, it is interesting that when we analyze the kinds of experiments we too often call "measurements", in detail, using the pilot-wave theory, we find this novel interaction-strength-dependent sort of contextuality virtually everywhere, even in so-called position measurements. Contextuality, in this sense, is a truly generic feature of measurements in the pilot-wave picture.

Disaster? Of course not.

References

1. D. Dürr, S. Teufel, *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory* (Springer, 2009)
2. A. Einstein, B. Podolsky, N. Rosen, Can quantum mechanical description of reality be considered complete? Phys. Rev. **47**, 777–780 (1935)

3. E. Schrödinger, The present situation in quantum mechanics, in *Naturwissenschaften*, vol. 23 (1935), translated by J. Trimmer, in *Proceedings of the American Philosohpical Society*, vol. 124, 10 (Oct. 1980), pp. 323–338
4. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Julius Springer - Verlag, Berlin, 1932), (English translation Princeton University Press, 1955)
5. S. Kochen, E. Specker, The problem of hidden variables in quantum mechanics. J. Math. Mech. **17**, 59–87 (1967)
6. J.S. Bell, On the problem of hidden variables in quantum mechanics. Rev. Mod. Phys. **38**, 447–452 (1966)
7. C. Held, *"The Kochen-Specker Theorem", The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), ed. by E.N. Zalta. https://plato.stanford.edu/archives/spr2018/entries/kochen-specker/
8. T. Norsen, The pilot-wave perspective on spin. Am. J. Phys. **82**, 337 (2014)
9. D.Z. Albert, *Quantum Mechanics and Experience* (Harvard University Press, Cambridge, MA, 1992), p. 153
10. T. Norsen, *Foundations of Quantum Mechanics: An Exploration of the Physical Meaning of Quantum Theory* (Springer, 2017)
11. M. Daumer, D. Dürr, S. Goldstein, N. Zanghi, Naive Realism about operators. Erkenntnis **45**, 379–397 (1996). (quant-ph/9601013)
12. S. Das, D. Dürr, *"Arrival Time Distributions of Spin-1/2 Particles", Nature Scientific Reports*, vol. 9, p. 2242 (2019). https://doi.org/10.1038/s41598-018-38261-4
13. Y. Aharonov, S. Popescu, D. Rohrlich, On conservation laws in quantum mechanics. *PNAS* (2020). https://doi.org/10.1073/pnas.1921529118
14. J.S. Bell, "Against 'Measurement'," in *62 Years of Uncertainty: Erice, 5-14 Aug. 1989* (Plenum Publishers, New York); reprinted in *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, 2004)

The Wentaculus: Density Matrix Realism Meets the Arrow of Time

Eddy Keming Chen

Dedicated to the memory of Detlef Dürr.

Abstract In this paper, I characterize and elaborate on the "Wentaculus" theory, a new approach to time's arrow in a quantum universe that offers a unified solution to the problems of what gives rise to the arrow of time and what the ontology of quantum mechanics is.

1 Introduction

Two of the most difficult problems in the foundations of physics are (1) what gives rise to the arrow of time and (2) what the ontology of quantum mechanics is. They are difficult because the fundamental dynamical laws of physics do not privilege an arrow of time, and the quantum-mechanical wave function describes a high-dimensional reality that is radically different from our ordinary experiences.

In this paper, I characterize and elaborate on the "Wentaculus" theory, a new approach to time's arrow in a quantum universe that offers a unified solution to both problems. Central to the Wentaculus are (i) Density Matrix Realism, the idea that the quantum state of the universe is objective but can be impure, and (ii) the Initial Projection Hypothesis, a new law of nature that selects a unique initial quantum state. On the Wentaculus, the quantum state of the universe is sufficiently simple to be a law, and the arrow of time can be traced back to an exact boundary condition. It removes the intrinsic vagueness of the Past Hypothesis, eliminates the Statistical Postulate,

E. K. Chen (✉)
Department of Philosophy, University of California, San Diego 9500, CA, USA
e-mail: eddykemingchen@ucsd.edu
URL: http://www.eddykemingchen.net

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_8

provides a higher degree of theoretical unity, and contains a natural realization of "strong determinism." I end by responding to four recent objections. In a companion paper, I elaborate on Density Matrix Realism.

2 Conceptual Foundations

I start by reviewing the conceptual foundations of Density Matrix Realism and the arrow of time that are necessary to formulate the Wentaculus in Sect. 3.

2.1 Density Matrix Realism

To understand Density Matrix Realism, let us start with a more familiar thesis—*Wave Function Realism*. When we consider realist solutions to the quantum measurement problem, such as Bohmian mechanics, objective collapse theories, and Everettian quantum mechanics, it is natural to consider the quantum state of the universe as an objective feature of reality. Moreover, it is widely believed that it has to be a pure state, represented by a wave function. Let us formulate the thesis as follows:

Wave Function Realism The quantum state of the universe is objective; it has to be pure.

This characterization of Wave Function Realism is broader than that of Albert (1996), Ney (2021). For them, Wave Function Realism carries a specific commitment to understanding the universal wave function as a physical field that lives on a vastly high dimensional "configuration" space, from which the ordinary 3-dimensional space is emergent. This is, however, not the only way to be a realist about the wave function. For example, the multi-field interpretation, spacetime state realism, and the nomological interpretation also count as versions of Wave Function Realism (Chen 2019b).

We often assume that the quantum state of the universe, if objective, must be pure. In quantum mechanics, mixed states are often used to represent reduced or statistical states, as expressions of entanglement with other systems or of our ignorance of the actual pure state. However, there is no compelling argument why the universe cannot be in a fundamental mixed state, one that does not arise from entanglement or lack of knowledge. In fact, it is easy to formulate Bohmian mechanics, collapse theories, and Everettian quantum mechanics with a fundamental density matrix (Allori et al. 2013; Dürr et al. 2005; Maroney 2005; Wallace 2012). Let us consider an alternative to Wave Function Realism, called *Density Matrix Realism*:

Density Matrix Realism The quantum state of the universe is objective; it can be pure or impure.

This thesis allows more choices of the quantum state of the universe; it can be pure or impure (a "mixed state"). The freedom to use impure states is crucial to the Wentaculus, which I discuss in Sect. 3.

Let us clarify some key terms in both theses.

(i) "The": it implies uniqueness. Both theses differ from the account considered by Wallace (2016) that the universe has two physical states at the same time: a fundamental pure state and a fundamental mixed state.
(ii) "Quantum state of the universe": both theses are about the quantum state of the universe. It does not logically entail that subsystem quantum states must be objective or that they must be pure (or impure).
(iii) "Objective": it means that the universal quantum state corresponds to an objective feature of reality, is not merely epistemic (encoding lack of knowledge), or pragmatic (merely a useful instrument for calculations). The meaning of objectivity is left open-ended, making room for different ways to be a realist (Chen 2019b).
(iv) "Must be pure" versus "Can be pure or impure": this is the only difference between the two theses. "Must" and "can" are modal concepts. Wave Function Realism restricts universal quantum states to only pure ones, while Density Matrix Realism allows both pure and impure universal quantum states. However, the latter is compatible with additional laws of physics (such as the Initial Projection Hypothesis) that make it physically impossible for the universe to be in a pure state.

Are there modifications of Bohmian mechanics, collapse theories, and Everettian quantum mechanics for which Density Matrix Realism is a natural framework? Yes, there are. They have been discussed, but not necessarily endorsed, by several authors in the foundational literature. For example, Dürr et al. (2005), Allori et al. (2013) have considered density-matrix realist versions of Bohmian mechanics, GRW theory, and Everettian quantum mechanics. See also Maroney (2005), Wallace (2012). For example, in the Bohmian framework, we can evolve the fundamental universal density matrix W by the Von Neumann Equation:

$$i\hbar\frac{\partial \hat{W}}{\partial t} = [\hat{H}, \hat{W}] \tag{1}$$

particle configuration by a new guidance equation (Dürr et al. 2005):

$$\frac{dQ_i}{dt} = \frac{\hbar}{m_i}\mathrm{Im}\frac{\nabla_{q_i} W(q, q', t)}{W(q, q', t)}(q = q' = Q) \tag{2}$$

and distribute the initial particle configuration by:

$$P(Q(t_0) \in dq) = W(q, q, t_0)dq \tag{3}$$

This version of Bohmian mechanics satisfies equivariance just as the wave-function version does (Dürr et al. 2005; Dürr, et al. 1992). In the Everettian framework, we can unitarily evolve the fundamental universal density matrix W by the same von Neumann equation (1), understand the emergent branching structure via decoherence, and apply decision theory or self-locating probability to recover the Born rule (Chua and Chen 2023). We also have the option to add a separable fundamental ontology in spacetime by defining a mass-density field (Allori et al. 2013):

$$m(x,t) = \mathrm{tr}(M(x)W(t)), \tag{4}$$

with $M(x) = \sum_i m_i \delta(Q_i - x)$, and $Q_i \psi(q_1, q_2, ...q_n) = q_i \psi(q_1, q_2, ...q_n)$. Finally, in the GRW framework, we can interrupt the unitary evolution of the fundamental universal density matrix W with spontaneous collapses that occur at rate $N\lambda$ (where N is the number of "particles" in the universe):

$$W_{T^+} = \frac{\Lambda_{I_k}(X)^{1/2} W_{T^-} \Lambda_{I_k}(X)^{1/2}}{\mathrm{tr}(W_{T^-} \Lambda_{I_k}(X))} \tag{5}$$

with W_{T^+} the post-collapse density matrix, W_{T^-} the pre-collapse density matrix, and X distributed by the probability density $\rho(x) = \mathrm{tr}(W_{T^-} \Lambda_{I_k}(x))$, where the collapse rate operator is defined as $\Lambda_{I_k}(x) = \frac{1}{(2\pi\sigma^2)^{3/2}} e^{-\frac{(Q_k - x)^2}{2\sigma^2}}$. Similar to the wave-function versions of GRW, we can define density-matrix versions of GRW with local beables such as a mass-density field $m(x,t)$ or flashes F (Allori et al. 2013).

Each of the preceding theories posits a fundamental universal density matrix with precise laws of nature. The three theories are called W-BM, W-EQM, and W-GRW. Given appropriate choices of the universal quantum state, each density-matrix theory is empirically equivalent to its wave-function counterpart, so that they cannot be distinguished even in principle by experiment or observation (Chen 2019a). I return to this issue in Sect. 5.1.

2.2 *The Arrow of Time*

To appreciate the Initial Projection Hypothesis in the Wentaculus, we need to review a standard account about the arrow of time. Given the reversibility of the fundamental dynamical laws, we must locate the origin of macroscopic irreversibility somewhere else. A proposal that has been influential in foundational literature posits a low-entropy boundary condition called the Past Hypothesis.[1] Roughly speaking, if our universe started in the Past-Hypothesis region of the global state space, it will (with

[1] See Albert (2000), Goldstein (2001), Callender (2004), Callender (2011), Lebowitz (2008), North (2011), Wallace (2023), Loewer (2020), Goldstein et al. (2020), and Chen (2023). For some criticisms, see Earman (2006); Winsberg (2004).

overwhelming probability) wander into states of higher entropy and eventually arrive at the thermodynamic equilibrium.

For concreteness, I focus on the Mentaculus theory of (Albert, 2000, 2015) and (Loewer, 2007, 2012), a particular version of the neo-Boltzmannian approach to the foundation of statistical mechanics. It contains two assumptions in addition to the fundamental dynamical laws. Following Albert and Loewer, let us call this package **the Mentaculus**:

1. **Fundamental Dynamical Laws (FDL):** the classical microstate of the universe is represented by a point in phase space (encoding the positions and momenta of all particles in the universe) that obeys $F = ma$, where F encodes the classical interactions.
2. **The Past Hypothesis (PH)**: at a temporal boundary of the universe, the microstate of the universe lies inside M_0, a low-entropy macrostate that corresponds to a small-volume set of points on phase space that are macroscopically similar.
3. **The Statistical Postulate (SP)**: given the macrostate M_0, we postulate a uniform probability distribution with respect to the natural measure over the microstates compatible with M_0.

Some comments:

(i) Loewer borrowed the name "Mentaculus" from the Coen Brothers movie *A Serious Man* (2009). It means the "probability map of the universe." The Classical Mentaculus provides a probability assignment for every proposition formulable in the language of the classical phase space. If correct, it may account for all the temporally asymmetric phenomena and underly the objective probability in deterministic physics and the special sciences (Loewer 2020).

(ii) There is a certain degree of vagueness in the partition of state space into macrostates, and hence in PH and SP. A macroscopic description of the initial state does not correspond to any exact region in phase space. Any choice of an exact region risks a certain kind of objectionable arbitrariness (Chen 2022b).

(iii) The probability distribution in SP can be regarded as an objective notion of "most" with which we can ignore the anti-thermodynamic initial microstates in the PH region of the state space. This can be interpreted as a kind of deterministic objective probability (Loewer 2001) or typicality measure (Goldstein 2012). In the quantum case, it is distinct from and in addition to the Born rule postulate.

How should we implement the Mentaculus in quantum theory? On Wave Function Realism, the natural strategy is to replace the classical state with a quantum pure state. This is a standard picture of Boltzmannian quantum statistical mechanics (Goldstein et al. 2010). We posit the quantum state of the universe as represented by a unitarily evolving wave function (obeying the Schrödinger equation) that started out in a low-entropy region in the Hilbert space, represented by a low-dimensional Past-Hypothesis subspace that corresponds to low quantum Boltzmann entropy. We further postulate a uniform probability distribution over all wave functions compatible with the PH subspace, with respect to the natural surface area measure on the unit sphere of the subspace. Call this the **Wave-Function Mentaculus**.

What about on Density Matrix Realism? A similar strategy is to replace the universal wave function with a universal density matrix that can be pure or impure. We postulate a unitarily evolving density matrix (obeying the von Neumann equation) that started out in the PH subspace. We further postulate a uniform probability distribution over all density matrices compatible with the PH subspace, with respect to the natural measure on the space of all such density matrices (Chen and Tumulka 2022). Call this the **Density-Matrix Mentaculus**.

An interesting feature of Density Matrix Realism is that there is another and perhaps more compelling way to implement the key idea. It satisfactorily removes the inherent vagueness of PH and eliminates the SP. That is the **Wentaculus**.

3 The Wentaculus

3.1 The Initial Projection Hypothesis

Can we do better than the preceding quantum versions of Mentaculus, in the sense of obtaining a unique initial quantum state? I suggest that we can. Recall that for any finite-dimensional Hilbert space $\mathscr{H}$ there is a natural density matrix in that Hilbert space—its normalized projection operator $\frac{I}{dim\mathscr{H}}$, where I is the identity / projection operator on $\mathscr{H}$ and $dim\mathscr{H}$ is the dimension of $\mathscr{H}$. Moreover, in general $\frac{I}{dim\mathscr{H}}$ is the simplest object one can associate with $\mathscr{H}$, containing no more information than is contained by $\mathscr{H}$ itself. Hence, if $\mathscr{H}$ is simple to characterize, then $\frac{I}{dim\mathscr{H}}$ is also simple to characterize.

As a special case, consider the particular Past-Hypothesis subspace $\mathscr{H}_{PH}$ (which, among other things, has very low dimension and thus very low quantum Boltzmann entropy $S_B(\mathscr{H}) = k_B dim\mathscr{H}$). There is a natural density matrix in $\mathscr{H}_{PH}$, namely $\frac{I_{PH}}{dim\mathscr{H}_{PH}}$, with I_{PH} the identity / projection operator on $\mathscr{H}_{PH}$ and $dim\mathscr{H}_{PH}$ the dimension of $\mathscr{H}_{PH}$. It is as simple to characterize as the Past-Hypothesis subspace itself (modulo a normalization constant). Therefore, if the PH is sufficiently simple to be considered a law, then the natural density matrix $\frac{I_{PH}}{dim\mathscr{H}_{PH}}$ is too.

I propose (Chen 2018) the following posit about the initial density matrix of the universe, called the *Initial Projection Hypothesis* (IPH):

$$\hat{W}_{IPH}(t_0) = \frac{\mathbb{I}_{PH}}{dim\mathscr{H}_{PH}}, \tag{6}$$

All the arguments that the PH should be nomological apply to IPH. I think the best understanding of this posit is a fundamental law of nature. After all, it is no more complicated and no less informative than usual versions of the PH (Chen 2023).

The posit can be generalized to other types of initial constraints. Here is a recipe: starting from the full Hilbert space (energy shell) $\mathscr{H}$, we can use simple principles (if there are any) to determine an initial subspace $\mathscr{H}_0 \subset \mathscr{H}$, choose the natural

quantum state in that subspace—the normalized projection $\hat{W}_0(t_0) = \frac{I_0}{dim\mathscr{H}_0}$. The natural choice will be simple and unique.

3.2 *Three Versions of the Wentaculus*

When we add the IPH to Density Matrix Realism, we arrive at the Wentaculus[2]:

1. **Fundamental Dynamical Laws (FDL):** the quantum state of the universe is represented by a density matrix $\hat{W}(t)$ that obeys the von Neumann equation (1).[3]
2. **The Initial Projection Hypothesis (IPH)**: at a temporal boundary of the universe, the density matrix is the normalized projection onto $\mathscr{H}_{PH}$, a low-dimensional subspace of the total Hilbert space. (That is, the initial quantum state of the universe is $\hat{W}_{IPH}(t_0)$ as described in equation (6).)

The Wentaculus implements the key idea of the Mentaculus for quantum theory. However, unlike the the Wave-Function Mentaculus, it requires Density Matrix Realism; moreover, unlike the Density-Matrix Mentaculus, its boundary condition law narrows down the choices of the initial density matrix to a unique one. It contains one fewer fundamental postulate than each of the preceding versions of the Mentaculus. The Statistical Postulate becomes redundant, because there is exactly one nomologically possible initial density matrix. In earlier versions of the Mentaculus, there are infinitely many nomologically possible initial states compatible with the PH. When we replace PH with IPH, we have only one initial state left.

The Wentaculus is compatible with realist solutions to the quantum measurement problem. For the Bohmian Wentaculus, we postulate that the state of the universe is described by particle configuration and the universal density matrix, and we add the IPH to the list of fundamental laws, described by Eqs. (1), (2), and (3). Given the IPH, the initial quantum state of the universe is nomologically necessary. Hence, there is only one nomologically possible velocity field for the particle configuration. This differs from the Bohmian Mentaculus or standard versions of Bohmian mechanics, where there is nomological contingency about the initial quantum state.

For the Everettian Wentaculus, we postulate that the state of the universe is described by the universal density matrix, and we add the IPH to the deterministic dynamical law, described by Eq. (1). Given the IPH, the initial quantum state of the universe is nomologically necessary, rendering the theory strongly deterministic (Chen 2022c). Given the fundamental laws, there is only one possible history of the universal density matrix, and hence only one possible history of the Everettian multiverse. I return to this issue in Sect. 4.3.

---

[2] The Wentaculus is so named because (1) it is inspired by the Mentaculus, and (2) "W" is sometimes used to denote the fundamental density matrix.

[3] For GRW-type theories, the density matrix obeys the stochastic modification of the von Neumann equation described in footnote #22.

For the GRW Wentaculus, we add the IPH to the stochastic dynamical law, described by Eq. (5). Given the IPH, the initial quantum state of the universe is nomologically necessary, but because of the stochastic dynamics, the theory is not strongly deterministic. There are many nomologically possible histories of the universal quantum state, corresponding to different collapse histories permitted by the theory.

Each Wentaculus theory is empirically equivalent to its Mentaculus counterpart, but they have physically inequivalent sets of models. For example, in the Bohmian case, the two theories will yield different particle trajectories. In the Everettian case, the two theories will yield different multiverses with different branches and different local descriptions.

3.3 Realist Interpretations of the Density Matrix

The Wentaculus is compatible with realist interpretations of quantum mechanics and of the quantum state.

We have four ontological interpretations of the universal density matrix. First, we can understand $W(q, q', t)$ as representing a physical field evolving in a $6N$-dimensional fundamental space represented by $\mathbb{R}^{6N}$. The field assigns properties to every point on that space. Second, we can understand it as representing a low-dimensional multi-field. The fundamental space is a 3-dimensional space represented by $\mathbb{R}^3$, and $W(q, q', t)$ assigns properties (represented by complex numbers or vectors) to every $2N$-tuple of points on that space. Third, we can understand it as representing properties of spacetime regions. We can obtain, from the universal density matrix, reduced density matrices that correspond to physical properties of regions in a 4-dimensional manifold. Such properties are in general non-separable due to quantum entanglement. Finally, we can understand it as representing a geometric object in Hilbert space. The Wentaculus is also compatible with the nomological interpretations of the quantum state. Moreover, as I explain in Sect. 4.1, it solves the problem of complexity. While a generic universal quantum state of both quantum versions of Mentaculus is enormously complicated, the initial density matrix postulated by the Wentaculus is sufficiently simple to be a law. For more details on these realist interpretative options, see Chen (2019b).

Of course, we need not be realists to accept the Wentaculus. There are non-realist quantum interpretations according to which the quantum state represents our knowledge (or the lack thereof) or practical guidance for how we should act. QBists and quantum pragmatists can regard the Wentaculus as giving them the best epistemic or practical guidance for what to believe and what to act. In fact, they may be more comfortable with mixed states than some realists are.

4 Implications

The Wentaculus has implications for several debates in foundations of physics.

4.1 The Nature of the Quantum State

On the Wentaculus, we have the option to regard the quantum state as ontological. However, we also have an improved option for the nomological interpretation of the quantum state. As already mentioned, if PH is sufficiently simple to be a law, then the normalized projection onto the PH subspace is sufficiently simple.

What does it mean to say the quantum state is nomological? As an analogy, consider the standard suggestion that the Hamiltonian function in classical mechanics is nomological. In Hamilton's equations:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \tag{7}$$

the q_i and p_i obviously represent something in the ontology. They have the "marks of the ontic." They take on complicated values, and their values are not completely fixed by the theory and thus nomologically contingent. In contrast, the Hamiltonian function H is very different: H generates motion; H is simple; and H is fixed by the theory (and nomologically necessary). According to the standard interpretation, H is not ontological but nomological. It does not represent things like particles or fields but a law that tells particles and fields how to move.

Consider the guidance equation in Bohmian Wentaculus, with the right hand side expanded with a fixed initial density matrix W_0:

$$\frac{dQ_i}{dt} = \frac{\hbar}{m_i}\mathrm{Im}\frac{\nabla_{q_i}\langle q|e^{-i\hat{H}t/\hbar}\hat{W}_{IPH}(t_0)e^{i\hat{H}t/\hbar}|q'\rangle}{\langle q|e^{-i\hat{H}t/\hbar}\hat{W}_{IPH}(t_0)e^{i\hat{H}t/\hbar}|q'\rangle}(q = q' = Q) \tag{8}$$

W_0 has a similar character as the Hamiltonian function in Hamilton's equations: W_0 generates motion; W_0 is simple; and W_0 is fixed by the theory (nomologically necessary). W_0 has the marks of the nomic and *can* be given a nomological interpretation. This is to be contrasted with the Bohmian Mentaculus: its initial quantum state (either pure or mixed) is not guaranteed to be simple by the PH. Implementing the nomological interpretation would require a different argument, perhaps by appealing to considerations about quantum gravity (Dürr, et al. 1996; Goldstein and Teufel 2001).

The nomological interpretation can also apply to certain versions of the Everettian Wentaculus with a mass-density ontology. We can understand that the mass density is constrained by a law:

$$m(x,t) = \mathrm{tr}(M(x)e^{-i\hat{H}t/\hbar}\hat{W}_{IPH}(t_0)e^{i\hat{H}t/\hbar}) \quad (9)$$

Here W_0 does not generate motion in the sense of giving a velocity field, but it still generates the exact shape of the mass-density field in ordinary spacetime. Since W_0 is simple, we can take this equation as the fundamental law in the Everettian theory with a mass-density ontology, and W_0 is again nomological. What exists in the material ontology is just a separable field on spacetime constrained by a simple law.

What about the GRW Wentaculus? This is a more delicate issue, because of interpretational questions about the nature of stochastic laws. However, if we think of GRW as giving us guidance about which histories are typical, then the initial density matrix together with stochastic dynamics will fix a class of the histories that are typical according to the laws, and any local probabilities can be obtained by conditionalizing the universal history on available records.

The nomological interpretation is much more attractive on the Wentaculus than on the Mentaculus, because the universal density matrix is guaranteed to be as simple as the PH. The nomological interpretation of the initial density matrix is compatible with both Humeanism and non-Humeanism. In particular, it demonstrates that quantum entanglement need not be a threat to Humean supervenience (Chen 2022).

4.2 Statistical Mechanical Probabilities

Another advantage of the Wentaculus is that there is only one kind of probability left—quantum mechanical probability. The Statistical Postulate (understood either as a probability or typicality measure) in the Mentaculus becomes redundant, because the Wentaculus allows only one nomologically possible initial quantum state.

For example, on the Bohmian Wentaculus, the only probability law we need is the distribution postulate of the initial particle configuration. On the Everettian Wentaculus, the only probability corresponds to the weights of the actual branches, which are interpreted decision-theoretically or using self-locating uncertainty. On the GRW Wentaculus, the only probability corresponds to the collapse chances. Hence, the probability map of the universe is entirely based on quantum probabilities (Chen 2020). This way of reducing the sources of probability is more conservative than the proposals of Albert (1994) and of Wallace (2023). For the Everettians, this has an additional bonus, to which I turn now.

4.3 Strong Determinism

The elimination of statistical mechanical probability and the postulate of a maximal constraint on the initial density matrix leads to an interesting consequence for Everettians. The Everettian Wentaculus becomes strongly deterministic, in the sense

that there is only one nomologically possible history given the fundamental laws. In a sense, we have eliminated all sources of arbitrariness in the theory, including that of the initial microstate. The theory would have explained everything at the fundamental level, leaving nothing nomologically contingent (Chen 2022c).

The Everettian Wentaculus, I believe, is a realistic and simple example of strong determinism. This has implications for thinking about the relevance of strong determinism to our world. Even if one regards Everettian quantum mechanics as the wrong solution to the measurement problem, it would be dogmatic to regard it as impossible, because it may be empirically equivalent to the other quantum theories. Hence, strong determinism may be closer to the actual world than we have imagined.

For non-Everettian versions of the Wentaculus where the quantum state obeys unitary dynamics, such as the Bohmian Wentaculus, we have strong determinism with respect to the quantum state history. Given the fundamental laws, the history of the quantum state could not have been different. The only nomological contingency resides in the initial particle configuration.

4.4 Nomic Vagueness

Standard versions of the PH, such as those found in all three versions of the Mentaculus, are best understood as fundamental yet vague laws. Removing their vagueness by picking an exact set of microstates leads to an objectionable kind of arbitrariness, that I call untraceability (Chen 2022b). The reason is that the PH is a macrostate law that does not directly enter into the micro-dynamical laws.

We can regard the Initial Projection Hypothesis as an exact law without committing to such arbitrariness in nature. The initial density matrix simultaneously plays the role of the macrostate and the role of the microstate. It spans the entire macrospace—$\mathscr{H}_{PH}$ and yet it also appears in the micro-dynamical laws. Its exact values will make a difference to what there is in spacetime or how it evolves. For example, in versions where the initial density matrix is ontic, it is automatically traceable. In versions where it is nomic, it is still traceable just like a constant of nature. Any slight change in its values will in general affect how things evolve, for example, by making a difference in the Bohmian velocity field or in the mass-density field.

This has implications for the mathematical expressibility of fundamental laws. The untraceable arbitrariness is a cost for eliminating fundamental nomic vagueness by fiat, but the Wentaculus takes that cost away. Hence, the Wentaculus saves the exactness of fundamental laws without any cost.

4.5 Theoretical Unity

With Density-Matrix Mentaculus, the Wentaculus offers more theoretical unity than Wave-Function Mentaculus. In Everettian and GRW cases, there is more kinematical unity. Even if the universe is in a pure state, most quasi-isolated subsystems do not have pure states.

The Wentaculus (as well as Density-Matrix Mentaculus) provides additional dynamical unity in the Bohmian case. Suppose the universe is partitioned into a system S_1 and its environment S_2. Since there are no actual spins to plug into the spin indices of the wave function, we cannot always define conditional wave functions in an analogous way. Still, we can follow Dürr et al. (2005) to define a *conditional density matrix* for S_1, by plugging in the actual configuration of S_2 and tracing over the spin components in the wave function associated with S_2. The conditional density matrix for S_1 is defined as:

$$W_{cond}{}^{s_1}_{s_1'}(q_1, q_1') = \frac{1}{N} \sum_{s_2} \Psi^{s_1 s_2}(q_1, Q_2) \Psi^*_{s_1' s_2}(q_1', Q_2), \tag{10}$$

with the normalizing factor: $N = \int_{\mathcal{Q}_1} dq_1 \sum_{s_1 s_2} \Psi^{s_1 s_2}(q_1, Q_2) \Psi^*_{s_1 s_2}(q_1', Q_2)$. Even if the universe is in a pure state, the configurations of most subsystems are guided by mixed states according to W-BM.

5 Objections and Replies

5.1 Reliability of Records

Albert (2022) raises two worries about the Wentaculus. Here I address his worry about the reliability of records on the Bohmian Wentaculus. In Sect. 5.2, I address his worry about time-translation invariance.

Objection: Suppose we have two momentum eigenstates, one with momentum $+1$ (moving uniformly to the right) and the other -1 (moving uniformly to the left). On the Bohmian Mentaculus, one of them is the actual quantum state guiding the particle. The particle will be either moving to the left or to the right. However, on the Bohmian Wentaculus, if we regard the equal mixture of the two as the fundamental density matrix, the particle guided by this density matrix will be "entirely, and permanently, and with certainty, at rest."[4] Nevertheless, when we measure the particle, the record may indicate that the particle is moving. Hence, that is a trou-

[4] Albert's thought experiment relies on certain idealizations about momentum eigenstates. We usually require the particle configuration to be guided by square-integrable wave functions or density matrices with finite traces, which do not include momentum eigenstates. However, as Sheldon Goldstein points out to me, the example can be fixed by considering momentum eigenstates defined on a closed circle instead of on an infinite line.

bling *mismatch between the experimental outcome and the state of the particle* on the Bohmian Wentaculus: the particle is not moving but the record says it is. The mismatch undermines our confidence in the reliability of records in a way analogous to the worry about Bell's Everett (?) theory (1981).

Reply: In a single-particle universe, there are no macroscopic records or human experiences. In a many-particle universe, things become more realistic and interestingly different. To analyze measurements and records, we need to consider the full physical setup of the subsystem and the environment.

Does the particle move when we include the environment? The answer is yes. Before measurement, the conditional density matrix of the particle, obtained by plugging the environmental configuration (where the recording device is in a ready state) into the universal density matrix, is an equal mixture of the "+1" momentum eigenstate and the "–1" momentum eigenstate. After measurement, the particle is measured and suppose the record indicates that it is moving to the left. Plugging the new environmental configuration into the universal density matrix, the conditional density matrix of the particle is approximately the "–1" momentum eigenstate, because the two parts of the universal density matrix have become macroscopically disjoint and decohered, with the configuration sitting in the "–1" part. Hence, the particle is uniformly moving to the left. There is no mismatch between the record and the state of the particle. The reply here is similar to Bohm's reply (1953) to Einstein (1953).[5] We should have no less confidence in the reliability of records on the Bohmian Wentaculus than on the Bohmian Mentaculus.

5.2 Time-Translation Invariance

Albert's second worry targets the nomological interpretation of the quantum state on the Wentaculus. The reason for postulating an initial universal quantum state, on the nomological interpretation, is to constrain how material objects move in spacetime. Hence, the universal quantum state is not some material object with its own independent dynamics. The only things that move and change should be particles (or fields) in physical spacetime, with (8) being the only fundamental dynamical law on the Bohmian Wentaculus and (9) on the Everettian Wentaculus. Since the right hand sides of (8) and (9) change their functional forms over time, they are not time-translation invariant.

Objection: The Wentaculus on the nomological interpretation of the quantum state is fundamentally non-time-translation-invariant, but the world described by such a theory is phenomenologically time-translation-invariant. The theory is "divided against itself" Albert (2022, p. 28).

[5] In response to Einstein's worry about the particle in a box of length L with a real-valued wave function $\psi(x) = Asin\frac{2\pi nx}{L}$, Bohm points out that ordinary Bohmian mechanics does not "contradict any known *experimental* facts," because when we carry out a "momentum measurement," the wave function (of a stationary state) is transformed and the particle starts to move, even though its original momentum is exactly zero.

Reply: First, it is hard to see why this is a bug from Albert's own perspective. For example, Bohmian mechanics is not Lorentz invariant, but the phenomenological world described by such a theory is. Albert's high-dimensional ontological interpretation of the wave function tells us the world is fundamentally $3N$-dimensional (with $N \approx 10^{80}$), while the phenomenological world described by such a theory is 3-dimensional. Fundamental reality may not be an exact image of our phenomenological world. Sometimes theoretical reasons take us to surprising conclusions about fundamental reality. As long as there is a natural and simple explanation of the phenomena from the fundamental, the theory is not self-undermined.

Second, there is a formal and technical sense in which we can recover a set of non-fundamental laws of the motion that are time-translation invariant. One can define a universal quantum state W_t from the fundamental laws (the Initial Projection Hypothesis and the von Neumann equation). With respect to this derivative object W_t, the particle motion will be time-translation invariant. This explains why there exists a predictive recipe that is time-translation invariant; the violation at the fundamental level makes no practical difference.

Third, I regard invariances and symmetries as only defeasible indicators for simplicity, and the lack thereof as defeasible indicators for complexity. Overall simplicity is something we should strive for. Does the violation of time-translation invariance render the Wentaculus more complicated than the Mentaculus? No; in fact the theory becomes simpler because of it. The defeasible indicators can be ignored in this case, because we already know we have a simple theory.

Finally, the non-invariance may be regarded not as a cost but an advantage of the theory, as manifestation of a deeper unity. In the Mentaculus, the theory is not fundamentally time-translation-invariant, because PH applies only at a particular time. However, we can still understand a sense in which the Mentaculus is time-translation-invariant. We can separate the dynamics from the boundary condition constraint; the dynamics is invariant even though the boundary condition is not. But in the Wentaculus, on the nomological interpretation, there is no such clean separation. The two are genuinely intertwined and unified into a single law.

5.3 Ontological Redundancy

The objection from ontological redundancy has come up in conversations. (Wallace (2012, p. 399) mentions but does not necessarily endorse an objection like this.)

Objection: The Wentaculus requires us to accept density matrix realism, which leads to "a major expansion of our ontology, from admitting only pure states, to admitting also mixed states." And this seems problematic and unjustified.

Reply: Density Matrix Realism does not have a larger ontology (about what actually exists) than Wave Function Realism. In fact, both frameworks postulate that there is exactly one actual quantum state of the universe. Their difference is a modal one, having to do with which states are possible. However, the possibility here is stronger than metaphysical possibility but potentially weaker than nomological

possibility. They impose substantive constraints on what kind of states the universe can have. There can be additional fundamental laws of nature, such as the Initial Projection Hypothesis, that further limit such possibilities.

Perhaps what is behind this worry is the following parsimony principle:

Parsimony of Nomological Possibilities All else being equal, we should prefer theories with smaller sets of nomological possibilities.

Ironically, this principle works against the objection, because it supports the Wentaculus over the Mentaculus. The Wentaculus is compatible with exactly one nomologically possible initial quantum state, while the Mentaculus is compatible with infinitely many. Hence, by Parsimony of Nomological Possibilities, all else being equal, we should prefer the Wentaculus to the Mentaculus. From the perspective of the Wentaculus, it is the Mentaculus that leads to a major expansion of our nomological possibilities.

5.4 The Classical Analogue

The final worry is hard to articulate but important to address, because many philosophers of physics have raised this objection in conversations.

Objection: The same "trick" can be played in the classical context. This means that all the advantages of the Wentaculus are too easy to achieve and therefore trivial. On first glance, the suggested maneuver is to take the "probability distribution" (ρ) as "ontic" or "nomic." The same thing can presumably be done in the classical context (see McCoy (2020) for an example), where the probability distribution on phase space can be given a similarly ontic or nomic interpretation, thus avoiding the problems in the classical domain as well. If that is possible, it seems to show that either we have proven too much, or that it does not depend on the details of quantum theory.

Reply: It is much less natural to give an ontic or nomic interpretation of the probability distribution in classical statistical mechanics. If we use the same idea in the classical domain, we will get a many-worlds version of classical mechanics or lose determinism. The classical probability distribution ρ plays no dynamical role (unlike the density matrix in the W-quantum theories). Since ρ follows the Hamiltonian dynamics, it will in general be supported on many macroscopically distinct regions on phase space. If we reify ρ as ontic and do not modify the dynamics, we arrive at a many-worlds theory for classical mechanics. If we modify the dynamics to introduce objective "collapses" of ρ that take it to some "branch" of ρ, it will look much more artificial and complex than the original deterministic classical theory. In contrast, on each of the three interpretations of QM, the artificial effects do not arise on the Wentaculus. The Bohmian version remains deterministic (and single-world), the GRW version remains stochastic (and single-world), while the Everettian/many-worlds version is still deterministic. On the other hand, even if a classical extension

of our maneuver is possible, it is unclear how it makes the quantum case trivial, since presumably both require different choices of the ontology and the dynamics.

6 Conclusion

The Wentaculus is an attractive picture of quantum mechanics in a time-asymmetric universe. It is a coherent theory, with arguably a better package of theoretical virtues than the standard picture. It illuminates the differences between Density Matrix Realism and Wave Function Realism, and displays the advantages of permitting fundamental mixed states. It has implications for our discussions about laws, chance, randomness, symmetries, vagueness, determinism, and the quantum reality. If the Wentaculus is correct, then solutions to the puzzles of time's arrow and quantum ontology are deeply related. Nature is so unified that we can solve both problems with one key.

Acknowledgements For helpful discussions, I thank David Albert, Valia Allori, Jeffrey Barrett, Craig Callender, Eugene Chua, Barry Loewer, Katie Robertson, Simon Saunders, Shelly Yiran Shi, David Wallace, audiences in the Oxford Philosophy of Physics Seminar, the Southern California Philosophy of Physics Group, Density Matrix Realism Workshop (FraMEPhys), and editors of this volume. I have learnt much from Detlef Dürr over the years, and I am grateful for his insights, suggestions, and encouragements.

References

D.Z. Albert, The foundations of quantum mechanics and the approach to thermodynamic equilibrium. Br. J. Philos. Sci. **45**(2), 669–677 (1994)

D.Z. Albert, Elementary quantum metaphysics, in *Bohmian Mechanics and Quantum Theory: An Appraisal*. ed. by J.T. Cushing, A. Fine, S. Goldstein (Kluwer Academic Publishers, Dordrecht, 1996), pp.277–84

D.Z. Albert, *Time and Chance* (Harvard University Press, Cambridge, 2000)

D.Z. Albert, *After Physics* (Harvard University Press, Cambridge, 2015)

D. Z. Albert, Physical laws and physical things (manuscript) (2022)

V. Allori, S. Goldstein, R. Tumulka, N. Zanghì, Predictions and primitive ontology in quantum foundations: a study of examples. Br. J. Philos. Sci. **65**(2), 323–352 (2013)

J.S. Bell, Quantum mechanics for cosmologists, in *Quantum Gravity 2*. ed. by C. Isham, R. Penrose, D. Sciama (Clarendon Press, 1981), pp. 611–637

D. Bohm, A discussion of certain remarks by Einstein on Born's probability interpretation of the ψ-function, in *Scientific Papers Presented to Max Born* (Hafner, 1953), pp. 13–19

C. Callender, Measures, explanations and the past: should 'special' initial conditions be explained? Br. J. Philos. Sci. **55**(2), 195–217 (2004)

C. Callender, Thermodynamic asymmetry in time, in *The Stanford Encyclopedia of Philosophy*, ed. by E.N. Zalta (Metaphysics Research Lab, Stanford University, 2011) fall 2011 edition. https://urldefense.com/v3/–https://plato.stanford.edu/entries/timethermo/;!!NLFGqXoFfo8MMQ!vprs9m5l-WnJD262YutLOEjrctYKHQxmVxyqMmK3j9jx9VGPMCweSj3h5e1WNlCQJYochYf0E1PwDFXTbQsT4iw7TmlMeLmhMDzemjYoR2fFYoG1

E.K. Chen, Quantum mechanics in a time-asymmetric universe: on the nature of the initial quantum state. Br. J. Philos. Sci. (2018). https://doi.org/10.1093/bjps/axy068
E.K. Chen, Quantum states of a time-asymmetric universe: wave function, density matrix, and empirical equivalence. Master's Thesis, Department of Mathematics, Rutgers University, New Brunswick (2019a). arXiv:1901.08053
E.K. Chen, Realism about the wave function. Philos. Compass **14**(7) (2019b)
E.K. Chen, Time's arrow in a quantum universe: on the status of statistical mechanical probabilities, in *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, ed. by V. Allori (World Scientific, Singapore, 2020)
E.K. Chen, From time asymmetry to quantum entanglement: the humean unification. Noûs **56**, 227–255 (2022)
E.K. Chen, Fundamental nomic vagueness. Philos. Rev. **131**(1) (2022b)
E.K. Chen, Strong determinism. Philosophers' Imprint, forthcoming (2022c). arXiv:2203.02886
E.K. Chen, The past hypothesis and the nature of physical laws, in *The Probability Map of the Universe: Essays on David Albert's Time and Chance*, ed. by B. Loewer, E. Winsberg, B. Weslake (Harvard University Press, 2023)
E.K. Chen, R. Tumulka, Uniform probability distribution over all density matrices. Quantum Stud.: Math. Found. **9**(2), 225–233 (2022)
E.Y.S. Chua, E.K. Chen, Decoherence, branching, and the Born rule in a mixed-state Everettian multiverse (2023). arXiv:2307.13218
D. Dürr, S. Goldstein, R. Tumulka, N. Zanghì, On the role of density matrices in Bohmian mechanics. Found. Phys. **35**(3), 449–467 (2005)
D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the origin of absolute uncertainty. J. Stat. Phys. **67**(5–6), 843–907 (1992)
D. Dürr, S. Goldstein, N. Zanghì, Bohmian mechanics and the meaning of the wave function, in *Experimental Metaphysics: Quantum Mechanical Studies in honor of Abner Shimony* (1996)
J. Earman, The "past hypothesis": not even false. Stud. Hist. Philos. Sci. Part B: Stud. Hist. Philos. Mod. Phys. **37**(3), 399–430 (2006)
S. Goldstein, Boltzmann's approach to statistical mechanics, in *Chance in Physics*. ed. by J. Bricmont, D. Dürr, M.C. Galavotti, G. Ghirardi, F. Petruccione, N. Zanghì (Springer, Berlin, 2001), pp.39–54
S. Goldstein, Typicality and notions of probability in physics, in *Probability in Physics* (Springer, Berlin, 2012), pp. 59–71
S. Goldstein, J.L. Lebowitz, C. Mastrodonato, R. Tumulka, N. Zanghì, Approach to thermal equilibrium of macroscopic quantum systems. Phys. Rev. E **81**(1), 011109 (2010)
S. Goldstein, J.L. Lebowitz, R. Tumulka, N. Zanghì, Gibbs and Boltzmann entropy in classical and quantum mechanics, in *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*. ed. by V. Allori (World Scientific, Singapore, 2020)
S. Goldstein, S. Teufel, Quantum spacetime without observers: ontological clarity and the conceptual foundations of quantum gravity, in *Physics meets Philosophy at the Planck Scale* (2001), pp. 275–289
J.L. Lebowitz, Time's arrow and Boltzmann's entropy. Scholarpedia **3**(4), 3448 (2008)
B. Loewer, Determinism and chance. Stud. Hist. Philos. Sci. Part B: Stud. Hist. Philos. Mod. Phys. **32**(4), 609–620 (2001)
B. Loewer, Counterfactuals and the second law, in *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, ed. by H. Price, R. Corry (Oxford University Press, 2007)
B. Loewer, Two accounts of laws and time. Philos. Stud. **160**(1), 115–137 (2012)
B. Loewer, The Mentaculus vision, in *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, ed. by V. Allori (World Scientific, Singapore, 2020)
O. Maroney, The density matrix in the de Broglie-Bohm approach. Found. Phys. **35**(3), 493–510 (2005)
C. McCoy, An alternative interpretation of statistical mechanics. Erkenntnis **85**(1), 1–21 (2020)

A. Ney, *The World in the Wave Function: A Metaphysics for Quantum Physics* (Oxford University Press, 2021)
J. North, Time in thermodynamics, in *The Oxford Handbook of Philosophy of Time* (2011), pp. 312–350
D. Wallace, *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation* (Oxford University Press, Oxford, 2012)
D. Wallace, Probability and irreversibility in modern statistical mechanics: classical and quantum, in *Quantum Foundations of Statistical Mechanics* (Oxford University Press, forthcoming, 2016)
D. Wallace, The logic of the past hypothesis, in *The Probability Map of the Universe: Essays on David Albert's Time and Chance*, ed. by B. Loewer, E. Winsberg, B. Weslake, (Harvard University Press, 2023)
E. Winsberg, Can conditioning on the "past hypothesis" militate against the reversibility objections? Philos. Sci. **71**(4), 489–504 (2004)

Bohmian Mechanics as a Practical Tool

Xabier Oianguren-Asua, Carlos F. Destefani, Matteo Villani, David K. Ferry, and Xavier Oriols

Abstract In this chapter, we will take a trip around several hot-spots where Bohmian mechanics and its capacity to describe the microscopic reality, even in the absence of measurements, can be harnessed as computational tools, in order to help in the prediction of phenomenologically accessible information (also useful for the followers of the Copenhagen theory). As a first example, we will see how a Stochastic Schrödinger Equation, when used to compute the reduced density matrix of a non-Markovian open quantum system, necessarily seems to employ the Bohmian concept of a conditional wavefunction. We will see that by dressing these conditional wavefunctions with an interpretation, the Bohmian theory can prove to be a useful tool to build general quantum simulation frameworks, such as a high-frequency electron transport model. As a second example, we will explain how a Copenhagen "observable operator" can be related to numerical properties of the Bohmian trajectories, which within Bohmian mechanics, are well-defined even for an "unmeasured" system. Most importantly in practice, even if these numbers are given no ontological meaning, not only we will be able to simulate (thus, predict and talk about) them, but we will see that they can be operationally determined in a weak value experiment. Therefore, they will be practical numbers to characterize a quantum system irrespective of the followed quantum theory.

X. Oianguren-Asua
Departament de Física, Universitat Autónoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

C. F. Destefani · M. Villani · X. Oriols (✉)
Departament d'Enginyeria Electrónica, Universitat Autónoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
e-mail: Xavier.Oriols@uab.cat

D. K. Ferry
School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ 85287, USA

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_9

1 Introduction

Questioning whether "there are" electrons inside our mobile phones sounds like an absurd reflection, and yet the standard (also called Copenhagen or orthodox) quantum theory is not able to affirm it [1, 2]. Under this theory, a quantum object has a well-defined property (like the position) only when its wavefunction is an eigenstate of the associated operator. We know that this happens when the property is "strongly measured". But in general, the wavefunction is in a superposition of eigenstates for the operator of that observable, meaning nothing can be said about it: the property becomes "unspeakable" until measured. Consequently, the Copenhagen theory affirms that it is meaningless to talk about, say, the positions of the electrons inside the active regions of nanoscale devices, because while in operation, their position is never (strongly) measured. Thus, there is no chance for an affirmative answer to our initial question. And yet, consciously or not, no engineer or applied physicist can seriously accept there is no electron in an operating nano-device like a transistor [1, 2]. Fortunately, alternatives to the Copenhagen interpretation of quantum mechanics exist, by which electrons have a defined position irrespective of their measurement and the state of superposition of their wavefunction, e.g., the well-known Bohmian interpretation [3–6].

What might be more relevant from a practical point of view, however, is that even if one turns a blind eye to these "picky unspeakabilities" of the Copenhagen theory, their implications also limit the employable modelling tool-set, making some scenarios look (unnecessarily) pathological. For example, the explained undefined position of electrons comes into conflict with a well-defined dwell time for the electrons in the active region of a nano-scale transistor, which is an essential parameter to predict the performance of next generation computers. Similar practical issues can be found in the search of measurement operators (like the multi-electron displacement current [7, 8]) in scenarios where their mathematical shape is far from obvious (e.g. in nano-scale devices operating at THz frequencies [9]), or when looking for pure state "unravellings" in non-Markovian open quantum systems. As we will see in this chapter, such problems happen to be unambiguously solvable under the Bohmian quantum theory. Interestingly, it turns out that the ones who came up with the mathematical tools that allow the prediction of the phenomenological manifestations of these "pathological" scenarios, were many times non-Bohmian physicists, who, maybe accidentally, reached concepts natural to Bohmian mechanics, like position post-selected weak values [10] or the conditional wavefunction [11, 12]. We will see in this chapter that, giving to these frameworks their (natural) Bohmian narrative, makes them even more capable computational tools (even useful for the followers of the Copenhagen theory).

1.1 A Suggestive Review

We can arrive at these conclusions through the inherently Bohmian concepts of a conditional wavefunction (CWF) and an effective wavefunction (EWF), introduced by Dürr et al. [13], together with the understanding of the measurement dilemma they illuminate [14]. But before going into the details, we note that only **non-relativistic quantum phenomena** will be discussed in this chapter. The spirit is to show that, for this kind of phenomena and their formulation, the Bohmian theory provides a most convenient narrative.

Given an isolated quantum system of N degrees of freedom described by the real coordinate vector $\vec{q} = (q_1, \ldots, q_N) \in \mathbb{R}^N$, its evolution in time t is given by two entities: a complex wavefunction $\Psi(\vec{q}, t)$, which in polar form $\Psi(\vec{q}, t) = \rho^{1/2}(\vec{q}, t)e^{i\mathcal{S}(\vec{q},t)/\hbar}$ encodes the real fields $\mathcal{S}$ and ρ, and a trajectory $\vec{q}^{\,\xi}(t) \equiv \vec{q}\,(\vec{\xi}, t)$ for the degrees of freedom, the initial condition of which, $\vec{q}^{\,\xi}(t = 0) = \vec{\xi} \in \mathbb{R}^N$, labels the actual trajectory among the possible ones. This trajectory is piloted by the wavefunction, which provides the velocity field v_k for the k-th degree of freedom as $v_k(\vec{q}, t) := \frac{1}{m_k}\frac{\partial \mathcal{S}(\vec{q},t)}{\partial x_k}$ [3–6]. Meanwhile, the wavefunction itself is guided by the Schrödinger Equation

$$i\hbar\frac{\partial \Psi(\vec{q}, t)}{\partial t} = \Big[\sum_{k=1}^{N}\frac{-\hbar^2}{2m_k}\frac{\partial^2}{\partial q_k^2} + U(\vec{q})\Big]\Psi(\vec{q}, t), \tag{1}$$

where m_k is the mass associated with the k-th degree of freedom and U denotes the potential energy field describing the interaction between the degrees of freedom.

The most general isolated system we can consider is the entire Universe, where $\vec{q}$ would reflect its possible **configurations**. Then, if we partition it into a subsystem of interest S, of $n < N$ degrees of freedom $\vec{x} = (x_1, \ldots, x_n)$, and its (big) environment, of degrees of freedom $\vec{y} = (y_{n+1}, \ldots, y_N)$, such that $\vec{q} \equiv (\vec{x}, \vec{y})$, Bohmian mechanics allows us to associate to the system and the environment their own wavefunctions, labelled by the initial joint configuration $\vec{\xi}$, as $\psi^{\xi}(\vec{x}, t) := \Psi(\vec{x}, \vec{y}^{\,\xi}(t), t)$ and $\varphi^{\xi}(\vec{y}, t) := \Psi(\vec{x}^{\,\xi}(t), \vec{y}, t)$. These are particular cases of the so called **conditional wavefunctions**. In general, a CWF is a "slice" of a wavefunction, obtained by evaluating some of its degrees of freedom along a (Bohmian) trajectory, while leaving the rest of them un-evaluated [6, 13]. Now, a priori, the actual trajectories $\vec{x}^{\,\xi}(t)$ and $\vec{y}^{\,\xi}(t)$ are unknown, but by the Quantum Equilibrium principle [13], if the trajectory of the whole Universe had a "typical" initial condition $\vec{\xi}$, the probability density of the position $\vec{x}$ at time t (resp. $\vec{y}$), will be given by the CWF as $|\psi^{\xi}(\vec{x}, t)|^2$ (resp. $|\varphi^{\xi}(\vec{y}, t)|^2$).

As proved in [15], the full Schrödinger Equation (1) can be rewritten exactly into a coupled pair of dynamical equations ruling the motion of the two CWFs. Assuming we can write $U(\vec{x}, \vec{y}\,) = U_x(\vec{x}\,) + U_{xy}(\vec{x}, \vec{y}\,)$, for the system S we have[1]

[1] For the environment the equation will be the same but changing the CWF and the index ranges in (2) and (3).

$$i\hbar\frac{\partial\psi^{\xi}(\vec{x},t)}{\partial t} = \left[\sum_{k=1}^{n}\frac{-\hbar^2}{2m_k}\frac{\partial^2}{\partial x_k^2} + U_x(\vec{x}) + U_{xy}(\vec{x},\vec{y}^{\,\xi}(t)) + \mathfrak{W}(\vec{x},\vec{y}^{\,\xi}(t),t)\right]\psi^{\xi}(\vec{x},t), \tag{2}$$

where $\mathfrak{W}$ is the so-called **quantum correlation potential**

$$\mathfrak{W}(\vec{x},\vec{y}^{\,\xi}(t),t) := \sum_{j=n+1}^{N}\left[-\frac{\hbar^2}{2m_j\rho^{1/2}}\left(\frac{\partial^2\rho^{1/2}(\vec{x},\vec{y},t)}{\partial y_j^2}\right) - \frac{1}{2}m_j v_j^2(\vec{x},\vec{y},t) - i\frac{\hbar}{2}\frac{\partial v_j(\vec{x},\vec{y},t)}{\partial y_j}\right]\Big|\vec{y}=\vec{y}^{\xi}(t) \tag{3}$$

where we recognize as its real part $\mathbb{R}e\{\mathfrak{W}\}$ the difference between the Bohmian quantum potential [5, 6] and the kinetic energy of the environment degrees of freedom y_j; and as the imaginary part $\mathbb{I}m\{\mathfrak{W}\}$, the spatial variation in the environment axes y_j of their associated Bohmian velocity. The evaluation of both parts involves, at each $\vec{x}$, a derivative of the phase $\mathcal{S}$ or of the magnitude ρ of the full wavefunction Ψ along the environment coordinates $\vec{y}$, centered at the trajectory position $\vec{y}^{\,\xi}(t)$. This means $\mathfrak{W}$ requires information about the wavefunction over nearby possible trajectories $\vec{y}^{\,\xi'}(t) = \vec{y}^{\,\xi}(t) + \Delta\vec{y}$, with $|\Delta\vec{y}|$ small. That is, the evolution of the CWF $\psi^{\xi}(\vec{x},t)$ depends on other adjacent CWFs or slices of the full wavefunction (with different $\vec{\xi}$). This feature is known as "quantum wholeness" [6].

Now, we might ask when the subsystem CWF $\psi^{\xi}(\vec{x},t)$ behaves as if it was an independent closed quantum system wavefunction, ruled by a unitary Schrödinger Equation like Eq. (1). We see in Eq. (2) that this happens for instance when $\mathfrak{W} = f(t)$ (adding only a global phase) and $U_{xy}(\vec{x},\vec{y}^{\,\xi}(t)) \simeq V(\vec{x},t)$ with a same shape irrespective of the trajectory $\vec{\xi}$.[2] Whenever this is the case, we can say that the CWF of the system is its **effective wavefunction**. The question is then: when do these two conditions happen? One of the most important cases is just after a "strong measurement" of the subsystem.

This is well-known in the Bohmian literature about measurement [5, 13, 14], but let us review it qualitatively, because it will be key to understand Markovianity. Given an initially closed quantum system S with EWF $|\psi(0)\rangle_S = \int\psi^{\xi}(\vec{x},t=0)d^nx$, following the standard von Neumann protocol [16], as part of the environment of S, let us consider the degree of freedom of the pointer of a macroscopic measuring apparatus M, $z \equiv y_{n+1}$. Initially this pointer will be around its repose position, independently of the rest of the environment, meaning it should have a localized EWF $|\varphi(0)\rangle_M$. Then, S and M are made to interact until $t = T$, through the von Neumann coupling Hamiltonian $\hat{H}_{MS} := \bar{\mu}(t)\,\hat{p}_M \otimes \hat{B}_S$, where $\hat{p}_M$ is the pointer's momentum

[2] If only $\mathbb{I}m\{\mathfrak{W}\}$ vanished, the CWF would already seem to be ruled by a unitary Schrödinger Equation of a closed system, with a real potential energy field defined as $V(\vec{x},t) := U(\vec{x},\vec{y}^{\,\xi}(t)) + \mathbb{R}e\{\mathfrak{W}(\vec{x},\vec{y}^{\,\xi}(t),t)\}$. Computationally though, in order to evaluate $\mathbb{R}e\{\mathfrak{W}\}$ and the trajectory $\vec{y}^{\xi}(t)$, a quantum description of the environment would still be required, making the CWF of S not independent of the environment's evolution and thus, not an EWF.

operator, $\hat{B}_S = \sum_k b_k |b_k\rangle \langle b_k|$ with $b_k \in \mathbb{R}$ is the diagonalized self-adjoint operator related to the property B of the system we wish to "measure" and $\mu := \int_0^T \bar{\mu}(t)dt$ is the interaction strength. This Hamiltonian entangles the position of the pointer z with the eigenstates $|b_k\rangle_S$ of $\hat{B}_S$ such that, the system-pointer wavefunction will separate the eigenstates $|b_k\rangle_S$ along the configuration space axis z, by enveloping each with a differently displaced version of the localized $|\varphi(0)\rangle_M$, scaled by $P_k := |\langle b_k|_S |\psi\rangle_S|^2$. For a reproducible measurement [14] and the pointer to show us macroscopically distinguishable results b_k, the interaction μ, proportional to the separation of the envelopes, must be strong enough to leave them macroscopically disjoint in z. Then, the pointer will show a position $z^{\xi}(T)$ around one of these envelopes, which will "slice" a CWF for S equal to the eigenstate $|b_j\rangle_S$, modulated by that envelope. This will happen with probability P_j (area of the envelope) given by the Quantum Equilibrium principle [13]. The CWFs linked to the rest of possible envelopes, are called "empty waves". At this point, the Copenhagen theory postulates a so-called "wavefunction collapse", that transforms the entangled wavefunction into a product of a single eigenstate $|b_j\rangle_S$ and its corresponding envelope [16]. In Bohmian mechanics, there is no need to postulate any physical "collapse", instead the orthodox collpase is seen just as an apparent process. Because the different CWF "groupings" in z have a macroscopically disjoint support, and because for $t > T$, the coupling potential vanishes, $\bar{\mu}(t) = 0$, the correlation potential $\mathfrak{W}$ for S "vanishes". In consequence, the CWF for S selected by the Bohmian position of the pointer, will evolve for $t > T$ as if it were again an independent closed quantum system wavefunction: it will be an EWF. Since the EWF is enough for the complete future description of S, we can consider an "effective collapse" $|\psi\rangle_S \to |b_j\rangle_S$.

Notice that either the assumption that, for time $t > T$, M does not interact anymore with S, or that its entanglement with S is lost by some sort of thermalisation (by which the empty waves get macroscopically dispersed in configuration space [13]), mean that the information of S "leaked" to the environment M, the "empty waves" do not interact back with the EWF of S. These assumptions thus imply that the environment effectively "forgets" the entanglement achieved with S. This is an environment behaviour we could call memory-less or **Markovian**.

Using this effective collapse idea, we can extract more general information about the subsystem. If part of the environment, let us call it the "ancilla" A, gets entangled with S and this ancilla then suffers an effective collapse as in the measurement we just described, S will also suffer an effective "collapse", but now into non-necessarily orthogonal, nor linearly independent states. If say, $|\theta_0\rangle_A$ and $|\psi\rangle_S$ are the EWFs of A and S before their interaction, then a general unitary evolution coupling them will yield $\hat{U}_{AS}|\theta_0\rangle_A \otimes |\psi\rangle_S = \sum_m |\theta_m\rangle_A \otimes \hat{M}_m|\psi\rangle_S$, with $\{|\theta_m\rangle_A\}_m$ an orthonormal basis of A's Hilbert space and $\{\hat{M}_m\}_m$ a family of bounded linear operators on S, such that $\sum_m \hat{M}_m^{\dagger}\hat{M}_m$ equals the identity. By measuring the observable of A with eigenstates $\{|\theta_m\rangle_A\}_m$, the composite will collapse into the (unnormalized) EWFs $|\theta_m\rangle \otimes \hat{M}_m|\psi\rangle_S$, with probabilities $P_m := \langle\psi|_S \hat{M}_m^{\dagger}\hat{M}_m|\psi\rangle_S$, where the corresponding CWF of S would be $|\phi_m\rangle_S := \hat{M}_m|\psi\rangle_S$. If A and S stop interacting, $|\phi_m\rangle_S$ will be EWFs of S, called

"conditional states" (which are a particular types of CWFs),[3] and this process is called a **generalized measurement** of S [14, 17].

Within the described scenario, consider the interpretation of density matrices in Bohmian mechanics, as useful tools for statistical predictions about stochastic ensembles of wavefunctions (even if they provide an incomplete microscopic description) [14, 18]. Then, the partial trace of A in the state $\hat{U}_{AS}|\theta_0\rangle_A \otimes |\psi\rangle_S$ will yield the unconditional (meaning we keep track of all possible measurement outcomes) post-measurement density matrix $\hat{\rho}_S = \sum_m |\phi_m\rangle_S \langle\phi_m|_S$. In general, this suggests, and turns out to be the case, that the partial trace of an ancilla partition A of a composite AS state, called, the reduced density matrix of S, can always be interpreted as how an ensemble of identical subsystems S, each coupled to an identical ancilla A, would be left if an strong measurement was performed on each ancilla A [17]. By the uniqueness of the partial trace, this "fictitious measurement" of A could be for an arbitrary observable. In consequence, since we could choose the position operator of A, the reduced density matrix can always be computed by the ensemble average over possible CWF-s of S. But importantly, if the traced partition A is not really measured at t and the entanglement between A and S is not "thermalised" and their interaction does not cease indeterminately, then the reduced density matrix of S will just be a "fiction" if interpreted as describing independent possible quantum states. Each conditional state of S, each $|\phi_m\rangle_S$ will still interact through the environment's degrees of freedom with each other, namely, they will not be (unnormalized) EWFs. Therefore, even if the reduced density matrix is enough to predict measurement statistics on S in typical (non-Markovian) scenarios it will not convey enough information to predict its time evolution. The traced out environment's entanglement with the subsystem will need to be tracked for that. This could then be called an environment with memory, or non-Markovian environment. For example, the microscopic information about the spatial distribution of the CWFs along the environment's axes would be required for that, which is encapsulated in the correlation potential (3) (or in the so-called "memory time superoperator" [19] of the standard quantum theory). To know this without explicit simulation of the environment is the challenge of open quantum systems.

2 How Markovian and Non-markovian Stochastic Schrödinger Equations Tacitly Employ Conditional Wavefunctions

There are scenarios where the "fiction" we indicated above does provide a reasonable description of an open quantum system. Let us consider a scenario where every Δt time units, a different portion of the environment (a different ancilla) got coupled with S and was then (the ancilla) strongly measured. If these ancillas never again inter-

[3] If the measurement was for the position operator of A, $|\phi_m\rangle_S$ would be the CWFs of the system for the state of $\hat{U}_{AS}|\theta_0\rangle_A \otimes |\psi\rangle_S$ as it was **before the strong measurement of A**, otherwise, they will only be CWFs of the collapsed $|\theta_m\rangle_A \otimes \hat{M}_m|\psi\rangle_S$.

acted with S (or their entanglement was somehow "thermalised" before their next interaction with S), the result would be equivalent to a generalized measurement of S every Δt. Then, following our previous comments, we could call such a system S, a Markovian open quantum system [20], satisfying, among others, the "Past-Future Independence" characterization of Ref. [21]. It turns out, as shown by Ref. [22], that if $\Delta t \to 0$, such a continuous measurement of ancillas that sequentially get coupled to the subsystem S, can be used to derive the dynamical equation for the reduced density matrix of S (also known as the master equation) of several Markovian environments. Moreover, it is proven that for any Markovian master equation a (perhaps fictitious) environment and a set of observables for it exist, such that the equation is interpretable as due to their simultaneous and continuous measurements [21, 22].

As a consequence, for a Markovian environment, instead of directly solving the master equation for S, we could do the following. First, find (fictitious or not) environmental ancillas and observables W, such that if the ancillas got entangled with S one after the other, and their properties W were sequentially measured, they would cause the same (unconditional) evolution of the reduced density matrix of S, as the one described by the master equation (which in principle is possible for all Markovian master equations). Then, if a pure state-vector of S was evolved, by choosing for each projective measurement of the bath ancillas, one of the possible post-measurement conditional states, this would generate pure states $|\psi_{w(t)}(t)\rangle_S$, associated with a certain chain of measurement results (an unravelling) for the bath ancillas: $w(t)$.[4] This pure state $|\psi_{w(t)}(t)\rangle_S$ is called the **quantum trajectory**, linked to the "noise realization" $w(t)$ for its environment [17, 20, 21]. As we saw previously, the reduced density matrix of S defines how S would be left if an unconditional measurement was performed on its environment. Since in Markovian systems, this can actually be assumed to be happening, the reduced density matrix for S is obtainable by averaging the quantum trajectories for the ensemble of possible bath measurement chains $w(t)$ [20, 21]

$$\hat{\rho}_S(t) := tr_E[\hat{\rho}_{ES}(t)] = \mathbb{E}_{w(t)}\left[|\psi_{w(t)}(t)\rangle_S \langle\psi_{w(t)}(t)|_S\right]. \tag{4}$$

Computationally, this means that if for a given master equation, we obtain the stochastic equation ruling the time evolution of such state-vectors $|\psi_{w(t)}\rangle_S$, we would be able to parallelize the computation of the reduced density matrix by solving several independent "vector equations", instead of a single big "matrix equation" [20, 21]. Equations of this kind are the so-called **Stochastic Schrödinger Equations** (SSEs) [17, 22]. Note that such a quantum trajectory $|\psi_{w(t)}\rangle_S$ for a Markovian environment, can always be physically interpreted in the Copenhagen explanation as a so-called pure unravelling [21] (where one would invoke the collapse of the subsystem wavefunction at each Δt). In the Bohmian view on the other hand, such a quantum trajec-

[4] At each time a different generalized measurement is performed on S, meaning the stochastic trajectory $w(t)$ reflects the Bohmian positions of different measurement pointers at each Δt. Its non-differentiability is thus unproblematic.

tory is just a normalized CWF of the subsystem S which is converted into an EWF (thus the normalization), after every significant Δt.

However, what if we had an environment that got entangled with S, but which never really allowed us to consider an effective collapse? What if the different CWFs of the subsystem S were allowed to interact in any future time, instead of being converted into EWF every Δt? That is, what if the "quantum trajectories" $|\psi_{w(t)}\rangle_S$ for different $w(t)$ could interact between them in future times, making their time evolutions not independent (and not parallelizable unless approximations are made)? This would mean that "the information leaked" into the environment from S, the Bohmian "empty waves", would be able to affect the evolution of the system at any time. Such an environment with "memory" of the entanglement achieved with S could be called a **non-Markovian** environment [21].

It turns out that Bohmian mechanics still allows a "pure state" description for S, since, given the position of the environment ancillas interacting with S,[5] S has always a CWF, wether the conditioning variables are measured or not [11, 12]. In the Copenhagen view, a CWF, does not have a physical interpretation, unless it is an EWF, e.g., unless the conditioning variable is strongly measured. As a consequence, under the Copenhagen view, if a SSE is found for a non-Markovian master equation, the pure state $|\psi_{w(t)}\rangle_S$ at time t, can only be understood as the state in which S would be left in if the environment ancillas were strongly measured to give $w(t)$. But, since this would produce a very different subsequent evolution of $|\psi_{w(t)}\rangle_S$, such a measurement can only be seen as "a fiction". Of course, non-Markovian SSEs under the Copenhagen view are still useful as pragmatic computational tools to obtain the reduced density matrix of S. However, dynamical information inherent to each pure state (each CWF), like two-time correlations, should be avoided, unless one accepts some sort of ontological reality (independent of measurement) for the conditioning property of the environment, such as the one provided by the Bohmian theory [11, 12].

This narrative in terms of Bohmian CWFs for non-Markovian open quantum systems is not only theoretically insightful, but also a **practical** tool to look for reasonable SSEs as we will exemplify now. In the first section, we arrived at an exact system of equations, Eq. (2), that described the general time evolution of CWFs in arbitrary settings. In principle, in those equations the CWF of the subsystem S and its environment E are coupled at all times, not only between them, but also with the rest of possible CWFs (signature of the non-Markovianity). However, for specific scenarios, we can make educated guesses for the correlation term $\mathfrak{W}$ of Eq. (3), and the classical potential U, to generate a SSE for CWFs of the subsystem S (which **need** to be independently evolvable to be a valid SSE). Thus, Eq.(2) is a general framework to look for position SSEs. In fact, this equation system is also a detector of non-Markovian behaviour. As long as the CWFs of the subsystem S,

[5] To allow non-Markovian SSEs "unravelled" through non-position variables, consider the positions of environment "pointers" coupled with non-position observables of the ancillas around the system. Else, consider the associated unmeasured system information $\mathfrak{B}^{\psi}$ presented in Sect. 3, or the modal theory corresponding to the unravelled observable.

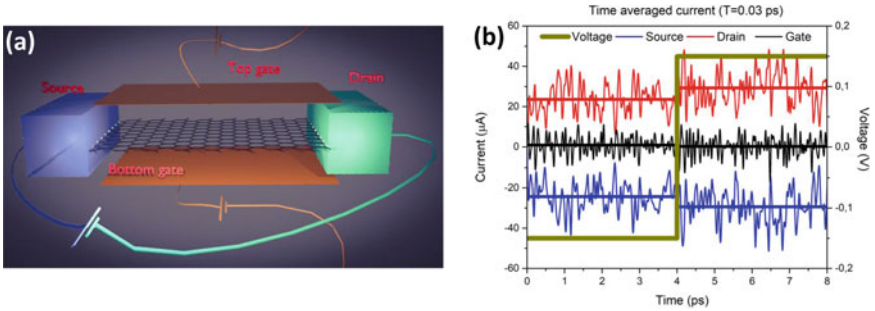


Fig. 1 (a) Schematic representation of the graphene-based FET, with a channel composed of a single-crystal mono-layer graphene. (b) The fluctuating lines are the instantaneous currents (time-averaged at a window of 0.03 ps) as a function of time. The straight lines are due to a wider averaging window of 4 ps, where we can clearly assert the binary response. We can conclude that, 4 ps is a reasonable operating time for the transistor

$\psi^{\xi}(\vec{x}, t) := \Psi(\vec{x}, \vec{y}^{\,\xi}(t), t)$, are described by a U or $\mathfrak{W}$ that depend explicitly on $\vec{y}^{\,\xi}(t)$, the system will be notably non-Markovian. As an example application of this method, we developed the BITLLES simulator [9, 23, 24]. In this simulator, we consider a two-terminal nano-scale electronic device operating at high frequencies (in the order of THz), where both the relevant dynamics of the active region electrons and the current measurement times on the reservoirs are in the sub-picosecond range. We consider the active region to be a non-Markovian open quantum system within the language of Eq. (2) [9]. The simulator computes the potential U as a solution of the Poisson equation [24], while $\mathfrak{W}$ is modelled by proper boundary conditions [24, 25] including the correlations between the active region of the device and the reservoirs. Even electron-phonon and electron-photon decoherence effects can be included [26, 27].

In Fig. 1 the ability of the present method is demonstrated [9, 28] by predicting for a field-effect transistor (FET) with a graphene channel, the time needed to acknowledge a stable reaction of the drain and source currents when the gate voltage is changed. The Klein tunneling suffered by the electrons while traversing the channel (partition noise) and the random energies of the electrons when injected into the system (thermal noise), cause a fluctuation in the instantaneous current that can be diminished in the laboratory by window averaging. The required window size for providing error-free current (averaged) values for digital applications (binary messages) defines the operating frequency of the transistor.

3 Speakable and Operational Information About an Unmeasured System?

Returning to the discussion at the beginning of the chapter, under the Copenhagen eigenstate-eigenvalue link, we can only say that a quantum system **has** a defined property when its wavefunction is an eigenstate of the operator related to that property. Since a strong measurement, as we have seen, always forces the system to adopt an eigenstate, while the unitary evolution in the meantime, will typically cause a superposition, it seems we are only allowed to speak about properties of **measured** quantum systems. This makes the predictions about what measurement apparatus pointers show, privileged in front of the rest of the information computable using the state of the pre-measurement quantum system. It is true that a quantum theory that correctly predicts what the measurement apparatus pointer will show, is by construction enough for phenomenological predictions. This is why it is argued (even by some Bohmian physicists) that if these predictions are obtainable with empirical agreement, dealing with the rest of the information concealed in the system's state (before its interaction with the measurement apparatus), is just adding unnecessary controversy. However, there are scenarios where the characterization of a quantum system, without the effect of a "collapse backaction" by some measurement pointer, would solve serious **practical** difficulties.

As a paradigmatic example, in order to obtain the maximum working frequency of nano-scale transistors (to test the performance of modern computers) [29], the time spent by electrons in the active region of the transistors, their **dwell time**, must be measured. The eigenstate-eigenvalue link would force us to place position detectors in the two ends of the active region. However, the quantum measurement, no matter how weak it is, introduces an effective collapse backaction in the system that disrupts its future evolution. Thus, the number given by these detectors would be meaningless to benchmark "unmeasured" transistors: no computer has position detectors at the ends of its transistors [30, 31]. Most two-time characterization attempts of "unmeasured" quantum systems face this same problem. For example, in thermodynamics: because **work** is by definition a dynamical property implying knowledge of the system (at least) at two different times, it seems there is no possible measurement-context-free definition for a quantum work operator [32–34]. More generally, **two-time correlations** of non-commuting observables, say F and B, cannot be defined within the Copenhagen school without including an explicit disturbance by a particular measurement scheme. For example, correlating the result of a strong measurement of F at time t_1 and a strong measurement of B at time t_2, clearly conveys the disturbing backaction of the measuring device, which collapses the state at t_1.[6] Thus, are we fundamentally forbidden to access dynamical information about the "unmeasured

[6] An alternative definition could be the real part of the (complex) expectation $\langle \hat{B}(t_2)\hat{F}(t_1)\rangle$ in the Heisenberg formalism, which turns out to be the correlation of a weak measurement [10] of F at time t_1 and a strong measurement of B at time t_2. Yet, as shown in Ref. [35], even an ideally weak measurement is in fact contextual (in the sense of footnote 8).

system"[7]? Or is there a way to consistently define non-*contextual*[8] properties (without contradicting the Bell-Kochen-Specker (BKS) theorem [36])? Bohmian mechanics, with its ontology of reality being persistent even when no measurement is taking place, appears to be the escape route. But is it?

Three impasses need to be clarified here. First: is there a (Bohmian) way to meaningfully talk about "unmeasured system" features, and even still be in accordance with phenomenology? Second: are these "unmeasured system" features experimentally accessible? If so, how can they be in agreement with the BKS theorem against non-contextual hidden variables? And third: can these features be employed to operationally compute practical information, or are they mere "philosophical reliefs"?

3.1 Breaking Impasse 1: Speakable Information of the "Unmeasured" System

Let us first clarify whether the information we obtain by measuring a quantum system is about the pre-measurement/"unmeasured" system or the post-measurement system. Consider an observable B of related operator $\hat{B} = \sum_b b|b\rangle\langle b|$, with $\{|b\rangle\}_b$ an orthonormal basis and $|\psi\rangle$ the wavefunction of the pre-measurement system. We have seen that (strongly) measuring B will lead the system to the **post**-measurement state $|b\rangle$ linked with the measured (eigen)value b, which will happen with a probability $|\langle b|\psi|^2$ due to the **pre**-measurement state. Thus, a single measurement tells us barely nothing about the **pre-**measurement system. But if a "measurement", as Bell pointed out [37], has the connotation of revealing information about the (**pre-**measurement) system, it seems that it would be more proper to name this process an "experiment" rather than a "measurement". We can try to recover the name "measurement" with an ensemble of these "experiments" over identically prepared pre-measurement states $|\psi\rangle$. With them, we could estimate the (squared) magnitudes of the **pre**-measurement projection-coefficients to each eigenstate $|\langle b|\psi|^2$ (e.g. using relative frequencies). Then, one could compute the expectation $\langle\psi|\hat{B}|\psi\rangle = \sum_b b|\langle b|\psi|^2$, which is also a number dependent on the pre-measurement state $|\psi\rangle$. However, from a Copenhagen point of view, this number (say, the average energy or position of an electron) can only be interpreted as a property of the **post**-measurement system, because by the eigenstate-eigenvalue link, only the post-measurement system can be attributed the observable b. When it comes to Bohmian mechanics, if $\hat{B}$ commutes with position $\hat{x}$, because the position x is "speakable" at all times, the number $\langle\psi|\hat{B}|\psi\rangle$ is the average property B of the **pre**-measurement system (as the simplest example, if $\hat{B} = \hat{x}$, it is the average Bohmian position of the unmeasured system). Yet, if $\hat{B}$ does not commute with $\hat{x}$ (e.g., for the momentum or the Hamiltonian operators), it is

[7] A system that is not being measured, e.g. a closed system evolving without quantum interaction with its environment.

[8] Contextual means it depends and implies the particular environment used to convey the information to the observer.

unclear if the expectation $\langle\psi|\,\hat{B}|\psi\rangle$ computed with the measured b, is a property of the pre-measurement system. In trying to clarify this, by linking the observable B to the position x of the Bohmian trajectories, which are "speakable", we can find a solution to the first impasse.

Given an arbitrary (Hermitian) operator $\hat{B}$, describing the observable B for the subsystem S, with normalized EWF $|\psi(t)\rangle$, let us define the position function $C^{\psi}(\vec{x},t) := \langle\vec{x}|\,\hat{B}|\psi(t)\rangle/\langle\vec{x}|\,\psi(t)$. If we write the expected value for $\hat{B}$ as a function of $C^{\psi}(\vec{x},t)$, we get that

$$\langle\hat{B}\rangle(t) = \langle\psi(t)|\,\hat{B}|\psi(t)\rangle = \int \langle\psi(t)|\,\vec{x}\,\langle\vec{x}|\,\hat{B}|\psi(t)\rangle dx = \int |\psi(\vec{x},t)|^2 C^{\psi}(\vec{x},t)dx. \quad (5)$$

This means that the spatial average of the (possibly complex) $C^{\psi}(\vec{x},t)$ yields, at all times, the same expected value for the observable B as that given by the Copenhagen theory. Now, let us define a real function $\mathfrak{B}^{\psi}(\vec{x},t) := \mathbb{R}e\{C^{\psi}(\vec{x},t)\}$. Because $\hat{B}$ is an observable, its expected value will be a real number, such that $\langle\hat{B}\rangle = \mathbb{R}e\{\langle\hat{B}\rangle\}$. Thus, taking the real part of equation (5), we get that

$$\langle\hat{B}\rangle(t) = \int |\psi(\vec{x},t)|^2 \mathfrak{B}^{\psi}(\vec{x},t)dx. \quad (6)$$

We can link this with the set of Bohmian trajectories $\{\vec{x}^{\,\xi}(t)\}_{\xi\in\Sigma}$ sampled in independent repetitions of the experiment, to get that $\langle\hat{B}\rangle(t) = \lim_{|\Sigma|\to\infty}\frac{1}{|\Sigma|}\sum_{\xi\in\Sigma}\mathfrak{B}^{\psi}(\vec{x}^{\,\xi}(t),t)$, by using the Quantum Equilibrium principle [13]. This means that the real number $\mathfrak{B}^{\psi}(\vec{x}^{\,\xi}(t),t)$, related to the $\vec{\xi}$-th Bohmian trajectory of the "unmeasured" system, when averaged over the ensemble of possible trajectories, gives the same value as the operator's expectation value. That is, irrespective of whether or not we give the observable B an ontological status, we can understand $\mathfrak{B}^{\psi}(\vec{x},t)$ as a mathematical feature related to B and linked to the Bohmian trajectory passing from $\vec{x}$ at time t (in the "unmeasured" system). This is why Holland gave the name *local expectation value* to position functions like $\mathfrak{B}^{\psi}$ [4]. However, we will just call them the "information linked to B and the Bohmian trajectory at $(\vec{x},t)$" or shortly "information $\mathfrak{B}^{\psi}$", to stress that we (still) mean nothing about their ontological or operational status.

For now, $\mathfrak{B}^{\psi}$ appears to be just an ad-hoc function of the trajectories for the operator expected value to be recovered from trajectories. Let us see though, that it can be more than this. What would this number be for each trajectory if the system state, $|\psi\rangle$, was an eigenstate of $\hat{B}$ of eigenvalue b?

$$\mathfrak{B}^{\psi}(\vec{x}) = \mathbb{R}e\left\{\frac{\langle\vec{x}|\,\hat{B}\psi}{\langle\vec{x}|\,\psi}\right\} = \mathbb{R}e\left\{\frac{\langle\vec{x}|\,\psi b}{\langle\vec{x}|\,\psi}\right\} = b \quad \forall\vec{x} \quad (7)$$

This means that a condition for $|\psi\rangle$ to be an eigenstate of $\hat{B}$ is that it is a state for which every Bohmian trajectory has the same value of the information $\mathfrak{B}^{\psi}$. On the

one hand, this tells us that the b indicated by the pointer of a projective measurement, can always be considered to be information linked to the Bohmian trajectory, even when its operator does not commute with position. On the other hand, in practice, it can be a tool to construct the operator $\hat{B}$ itself. One could define $\hat{B}$ in terms of $\mathfrak{B}^{\psi}$, as the collection of states $|b\rangle$ in which all Bohmian trajectories have the same value b for the information $\mathfrak{B}^{\psi}$ (and some condition to make the imaginary part of C^{ψ} negligible).

If the explicit shape of $\mathfrak{B}^{\psi}$ had nothing to do with Bohmian mechanics, this reverse definition of $\hat{B}$ would be a pointless definition. However, it turns out that if we set $\hat{B}$ to be the momentum operator $\hat{p}_k$ of the k-th degree of freedom, the trajectory information $\mathfrak{B}^{\psi}(\vec{x}, t)$ is exactly equal to the Bohmian momentum of the trajectory crossing $\vec{x}$: $m_k v_k(\vec{x}, t)$ [38]. If we set $\hat{B}$ to be the Hamiltonian operator $\hat{H}$, the information $\mathfrak{B}^{\psi}(\vec{x}, t)$ turns out to be exactly equal to the Bohmian energy (kinetic plus classical and quantum potentials [6]) of the trajectory crossing $\vec{x}$. One can see that the list of these "fortunate" matches for position functions that appeared to be designed only to satisfy the expectation values, goes on and on [4]. This suggests that we can employ Bohmian mechanics to derive the expression for $\mathfrak{B}^{\psi}$, thanks to its similarity with classical mechanics, and then define the related operator in those terms. Whether the information $\mathfrak{B}^{\psi}$ has an ontological status or not, whether it is operational or not, this is already (numerically) useful, because there are observables, like the total (particle plus displacement) current in a nano-device (plotted in Fig. 1b), for which there is no clear operator, but there is a clear Bohmian observable associated with it, as will be explained in detail later [7, 8].

In a nutshell, since we placed no restriction on $\hat{B}$, we are mathematically safe to assume that at all times, each Bohmian trajectory $\vec{\xi}$, has a simultaneously determined value $\mathfrak{B}^{\psi}(\vec{x}^{\,\xi}(t), t)$ linked to **every** observable operator $\hat{B}$. Whether the information $\mathfrak{B}^{\psi}(\vec{x}^{\,\xi}(t), t)$ reflects an **ontic property** (a property that the theory postulates to be part of the ontology) or not, is given by the quantum theory at hand. For example, we found that when $\mathfrak{B}^{\psi}$ is linked to the momentum operator $\hat{p}_k$, it is equal to the Bohmian particle's momentum, which is an ontic property in Bohmian mechanics, but not in the Copenhagen theory. The key is that when $\mathfrak{B}^{\psi}$ is equal to an ontic property, since the Bohmian trajectory exists in the absence of observation, B becomes "speakable" with a well-defined value at all times. Importantly though, we saw that the information $\mathfrak{B}^{\psi}$ is an equally well-defined number linked to each Bohmian trajectory, independently of the ontic character of B.[9] Then, the fact (we will show now) that the $\mathfrak{B}^{\psi}$ can

[9] The information $\mathfrak{B}^{\psi}$ will evolve continuously as long as the wavefunction evolves unitarily (which in Bohmian mechanics always does, as we saw). Then, if the system evolves from an eigenstate $|b_1\rangle$ to another $|b_2\rangle$ with eigenvalues $b_1 \neq b_2$, $\mathfrak{B}^{\psi}$ will take all the intermediate values not necessarily among the eigenvalues of $\hat{B}$. This suggests an interpretation in which the "quantization" of quantum mechanics is an apparent property, due to the fact that for a "proper" measurement, we require that a pointer saying b is compatible with a wavefunction $|b\rangle$ that yields for a strong measurement the result b with probability 1. That is, a wavefunction which has all its Bohmian trajectories with value b for $\mathfrak{B}^{\psi}$. Then, we would call it "quantum" because this delicate orchestration can only happen for a certain "quantized number" of wavefunctions (the eigenstates).

be operationally obtained in a laboratory, will make the information $\mathfrak{B}^{\psi}$ practically useful across the "unspeakables" and independently on the followed theory.

3.2 *Breaking Impasse 2: Is This "Unmeasured" System Information Operational?*

If we could only obtain the information $\mathfrak{B}^{\psi}$ in a laboratory when we used a strong von Neumann interaction, forcing it to become an eigenvalue of $\hat{B}$, all this would limit us in practice in the same manner as the eigenstate-eigenvalue link, even if we could now speak about these numbers in the absence of measurement. If so, we could not strictly say that the information $\mathfrak{B}^{\psi}$ is an **operational property**[10] of the unmeasured system. However, it turns out that the "unmeasured" $\mathfrak{B}^{\psi}$ is actually experimentally determinable even for a non-eigenstate pre-measurement system. The "how", explains the "cumbersome" definition $\mathfrak{B}^{\psi}(\vec{x}, t) = \mathbb{R}e\{\langle\vec{x}|\,\hat{B}|\psi\rangle / \langle\vec{x}|\,\psi\}$. It turns out to be the protocol that naive classical experimentalists [39] would follow if they thought the system had a defined position, initially uncertain to them, and the only quantum knowledge they had was that measurement interactions spoil the system's natural subsequent evolution. In order to know the property B of such a subsystem S (say, an electron) when it crosses $\vec{x}$, they would first couple an ancilla A to the subsystem S of EWF $|\psi\rangle$, through the measurement Hamiltonian $\bar{\mu}(t)\,\hat{p}_A \otimes \hat{B}$ but let the interaction strength μ be very small, such that the system state is only slightly perturbed. They would strongly measure the slightly entangled ancilla's position z_B with a certain probability density $P(z_B)$, getting a weak measurement about the property B of S. Before the slightly perturbed system S further evolved, they would couple the system to another ancilla and strongly measure its position z_x, with a certain conditional probability density $P(z_x|z_B)$. Finally, they would average the weak measurements of B for which the system S (the electron) was found at $\vec{x}$, in order to erase the noise introduced by the weakness of the coupling with A. If the averaged ensemble is large enough, the resulting conditional expectation will be equal to $\int z_B P(z_B|z_x) dz_B$, which as proven in [35], turns out to be roughly equal to $\mathfrak{B}^{\psi}(\vec{x}, t)$ (under feasible experimental conditions). This is called a *position post-selected weak value* [10].

A naive experimentalist would not be surprised at all by such a "coincidence". One can consider all this was juggling with results of several observations. But, when the information $\mathfrak{B}^{\psi}$ is an ontic property of the theory, one can legitimately say (under that theory), that the average weak measurements of B, for experiments in which the system (the electron) was at $\vec{x}$, gave $\mathfrak{B}^{\psi}(\vec{x}, t)$, because whenever the Bohmian trajectory (the electron) was at $\vec{x}$, it had indeed the property $\mathfrak{B}^{\psi}(\vec{x}, t)$. Be that as it

[10] A number that can be obtained in a laboratory with a well-defined protocol.

may, because we can follow this protocol in a lab for most observables B, irrespective of their ontic state, $\mathfrak{B}^{\psi}(\vec{x}, t)$ is (almost always[11]) an operational property [35, 38].

Let us clarify the non-contextuality of the information $\mathfrak{B}^{\psi}$. Because the Bohmian position and EWF of a subsystem immediately determine $\mathfrak{B}^{\psi}$ for any observable B, this apparently violates the BKS theorem [36], by which there can be no pre-existing variables that **non-contextually** determine the measurement outcomes for all observables (not even only for commuting groups of them). This does not preclude the weak values of the above protocol from being non-contextually pre-determined, because they deal with a different notion of "measurement": the hypotheses of the theorem refer to the Copenhagen quantum measurement (of Sect. 1), while the above weak value protocol "measuring" non-contextual information $\mathfrak{B}^{\psi}$, is an ensemble average of several Copenhagen quantum measurements, each of which is indeed contextual, since the Bohmian description of the measurement apparatus is necessary to determine their individual outcomes [14]. Moreover, as we saw (in footnote 9), the value of $\mathfrak{B}^{\psi}$ for a certain trajectory alone does **not** determine a von Neumann measurement outcome, since it is the coupling Hamiltonian (contextual) that forces the pre-measurement $\mathfrak{B}^{\psi}$ (a priori not even "quantized") to evolve to different ("quantized") eigenvalues of the operator $\hat{B}$. And even still, the weak value protocol, does produce a non-contextual $\mathfrak{B}^{\psi}$ value (through many contextual experiments). This is the reason why one might prefer to regard the post-selected averaging as an **uncontextualization protocol**. The clarification would be unnecessary, however, if history had preserved the original meaning of the word "measurement" as a **protocol** that unveils features of a system, existent before the interaction with the external probes. Unfortunately, according to standard quantum mechanics, as stated by Mermin [36], "the outcome of a measurement is brought into being by the act of measurement itself".

3.3 *Breaking Impasse 3: Is This Information Useful for a Non-Bohmian?*

Regardless of the followed quantum theory and whether one is ready to accept an ontological status for a certain information $\mathfrak{B}^{\psi}$, its relation with expected values

[11] There is a (quite important) exception. Identical particles are always ontologically distinguishable by their trajectories in Bohmian mechanics. In the laboratory however, there are no means to tag each individual particle under many-body wavefunctions with exchange symmetry. In consequence, if we follow our weak value protocol to "measure" the information $\mathfrak{B}^{\psi}_{(k)} := \mathbb{R}e\{\langle \vec{x}_1, \ldots, \vec{x}_M | \, \hat{Id}_{(1)} \cdots \hat{Id}_{(k-1)} \hat{B}_{(k)} \hat{Id}_{(k+1)} \cdots \hat{Id}_{(M)} | \psi \rangle / \langle \vec{x}_1, \ldots, \vec{x}_M | \, \psi \rangle\}$ related to the observable B of the k-th electron, in a system of M electrons of positions $\vec{x}_k$ with many-body wavefunction $|\psi\rangle$, what we will get instead is the average: $\sum_{k=1}^{M} \frac{1}{M} \mathfrak{B}^{\psi}_{(k)}(\vec{x}_1, \ldots, \vec{x}_M)$. Thus, the average $\mathfrak{B}^{\psi}_{(k)}$ for a multi-particle Bohmian trajectory is operational (say, the sum of the current contributions of the active region electrons, as discussed in the next paragraph), but the individual indistinguishable particle $\mathfrak{B}^{\Psi}_{(k)}$ (like the individual electron current contributions) are not, even if they might be ontic properties within Bohmian mechanics.

and the definition of the observable operator $\hat{B}$ is mathematically true. This has an important practical application that is also useful for a non-Bohmian. The information $\mathfrak{B}^{\psi}$ can be used to numerically predict the expected value of observables with no explicit definition of their formal operators. For this, one can express the observable B in the language of Bohmian mechanics to derive the shape of $\mathfrak{B}^{\psi}(x,t)$, and then get the expected value of the operator $\hat{B}$ (if there is any) by computing the trajectory ensemble average of $\mathfrak{B}^{\psi}$. For example, this is how we predict the expected total electrical current (including the displacement current) crossing the active region of a two-terminal nano-device operating at high frequencies (THz) in the BITLLES simulator [7, 8]. We can define the contribution to the total current through a surface σ, due to the Bohmian trajectory of a k-th electron $\vec{x}_k^{\,\xi}(t)$ of charge e, as $I_k^{(\xi)}(t) = \int_\sigma \vec{J}^{\,(\xi)}(\vec{r},t)\cdot d\vec{s} + \int_\sigma \varepsilon(\vec{r},t)\frac{\partial \vec{E}^{\,(\xi)}(\vec{r},t)}{\partial t}\cdot d\vec{s}$, where $\varepsilon(\vec{r},t)$ is the dielectric permittivity, $\vec{J}^{\,(\xi)}(\vec{r},t) = e\frac{\mathrm{d}\vec{x}_k^{\,\xi}(t)}{\mathrm{d}t}\delta(\vec{r}-\vec{x}_k^{\,\xi}(t))$ is the particle current density, and $\vec{E}^{\,(\xi)}(\vec{r},t)$ is the electric field generated by the electron, as a solution to Gauss' equation. The sum of these contributions, $I^{(\xi)}(t) = \sum_k I_k^{(\xi)}(t)$, will be the total Bohmian current at the surface σ for the ξ-th experiment in the ensemble $\{\vec{x}^{\,\xi}(t)\}_{\xi\in\Sigma}$. The phenomenological expectation of a total current operator $\hat{I}$ can then be estimated as the ensemble average of these currents, since by the Quantum Equilibrium principle, $lim_{|\Sigma|\to\infty}\frac{1}{|\Sigma|}\sum_{\xi\in\Sigma} I^{(\xi)}(t) = \langle\hat{I}\rangle(t)$, if such an operator exists.

In addition, it is also true that the information $\mathfrak{B}^{\psi}$ is a (typically) experimentally obtainable number that, no matter the followed interpretation of quantum mechanics, characterizes the theoretical pre-measurement wavefunction $|\psi\rangle$. This means it can be pragmatically employed to characterize an unmeasured quantum system, just like a tomography or momentum-postselected weak values are useful to obtain the pre-measurement wavefunction [40], no matter the ontological status or speakability of such a wavefunction. Following this, the $\mathfrak{B}^{\psi}$ that happen to be operational offer a natural solution to the puzzling search of non-contextuality for the metrics involving two different times [38].

For example, they provide a well-defined non-contextual two-time correlation function for general observables. Consider a big enough set of trajectories $\{\vec{x}^{\,\xi}(t)\}_{\xi\in\Sigma}$ sampled from the pre-measurement wavefunction $|\psi(t)\rangle$. Given the observables B, F, the ξ-th trajectory has associated informations $\mathfrak{B}^{\psi}(\vec{x}^{\xi}(t),t) := \mathbb{R}e\{\frac{\langle\vec{x}^{\,\xi}(t)|\hat{B}|\psi(t)\rangle}{\langle\vec{x}^{\,\xi}(t)|\psi(t)\rangle}\}$ and $\mathcal{F}^{\psi}(\vec{x}^{\xi}(t),t) := \mathbb{R}e\{\frac{\langle\vec{x}^{\,\xi}(t)|\hat{F}|\psi(t)\rangle}{\langle\vec{x}^{\,\xi}(t)|\psi(t)\rangle}\}$, which are well-defined even if the associated operators $\hat{B}, \hat{F}$ do not commute. This gives a natural correlation function defined as

$$\langle B(t_2)F(t_1)\rangle := \lim_{|\Sigma|\to\infty}\frac{1}{|\Sigma|}\sum_{\xi\in\Sigma}\mathfrak{B}^{\psi}(\vec{x}^{\,\xi}(t_2),t_2)\mathcal{F}^{\psi}(\vec{x}^{\,\xi}(t_1),t_1). \tag{8}$$

In a similar way, we can solve the problems concerning a quantum work definition, just as done by Refs. [41, 42]. First note that given a general system Hamiltonian $\hat{H} = \sum_k \frac{-\hbar^2}{2m_k}\frac{\partial^2}{\partial x_k^2} + V(\vec{x},t)$, we get

$$\mathcal{H}^{\psi}(\vec{x}^{\,\xi}(t),t) := \mathbb{R}e\left[\frac{\langle \vec{x}^{\,\xi}(t)|\,\hat{H}|\psi(t)\rangle}{\langle \vec{x}^{\,\xi}(t)|\,\psi(t)}\right] = \sum_{k=1}^{n}\frac{1}{2}m_k v_k(\vec{x}^{\,\xi}(t),t)^2 + V(\vec{x}^{\,\xi}(t),t) + Q(\vec{x}^{\,\xi}(t),t), \tag{9}$$

with Q the well-known Bohmian quantum potential [4–6]. This proves $\mathcal{H}^{\psi}(\vec{x}^{\,\xi}(t),t)$ is, as anticipated, the total Bohmian energy of the $\vec{\xi}$-th trajectory at time t. We can compute its associated Bohmian work with $\mathcal{W}^{(\xi)}(t_1,t_2) = \int_{t_1}^{t_2}\frac{\mathrm{d}\mathcal{H}^{\psi}(\vec{x}^{\,\xi}(t),t)}{\mathrm{d}t}dt = \mathcal{H}^{\psi}(\vec{x}^{\,\xi}(t_2),t_2) - \mathcal{H}^{\psi}(\vec{x}^{\,\xi}(t_1),t_1)$. As a result, a well-defined non-contextual definition of the quantum work could be the ensemble average of the trajectory works,

$$\langle W(t_1,t_2)\rangle = \lim_{|\Sigma|\to\infty}\frac{1}{|\Sigma|}\sum_{\xi\in\Sigma}\left(\mathcal{H}^{\psi}(\vec{x}^{\,\xi}(t_2),t_2) - \mathcal{H}^{\psi}(\vec{x}^{\,\xi}(t_1),t_1)\right). \tag{10}$$

Finally, we could give a reasonable Bohmian answer to the pathological search of an "unmeasured" dwell time, as the expected time spent by the Bohmian trajectory of the electron within the active region $\Gamma \subset \mathbb{R}^3$. Mathematically, the dwell time τ for the $\vec{\xi}$-th trajectory of the k-th electron with EWF $\psi^{\xi}(\vec{x}_k,t)$ is by definition given by the integral: $\tau^{(\xi)} = \int_0^{\infty} dt\int_{\Gamma}\delta(\vec{r}-\vec{x}_k^{\,\xi}(t))d\vec{r}$. This makes the expected time $\langle\tau\rangle$ to be given by the Quantum Equilibrium principle as an integral that is already employed to predict the dwell time,

$$\langle\tau\rangle = \lim_{|\Sigma|\to\infty}\frac{1}{|\Sigma|}\sum_{\xi\in\Sigma}\tau^{(\xi)} = \int_0^{\infty} dt\int_{\Gamma}|\psi^{\xi}(\vec{r},t)|^2 d\vec{r}. \tag{11}$$

It is worth noting that the Bohmian perspective allows to exclude in (11) the contribution of reflected trajecories, providing a more suitable metric for cutoff frequency estimates in electronic devices, since only the transmitted particles have a net contribution to the average electrical current. To conclude the section and link it with the discussion on non-Markovian SSEs, notice that, because in the non-Markovian case, the trajectory for the "unravelled" environment observable (what we denoted by $w(t)$), can no longer be interpreted as the result of a continuous measurement of the environment, it represents an unmeasured observable of the environment. Thus, this is readily a, perhaps unintended, application of $\mathfrak{B}^{\psi}$-like properties, which happen to be central to simulate the most general quantum systems that interact with many environmental degrees of freedom.

4 Conclusions

In this chapter, we have seen that inherently Bohmian concepts like the CWF or position post-selected weak values are indeed usable pragmatically as practical tools in the computation of phenomenologically accessible elements, such as the reduced density matrix, expectation values or time correlations. Therefore, with this chapter, we refute the main criticism to the Bohmian theory, by which the trajectories are

"unnecessary embellishments" of the orthodox theory, with no practical use. But then, if we can use Bohmian concepts as a tool, why not include them in the standard vocabulary? Not only for their problem-solving utility, but also because they can provide us ontological relief in front of the purely phenomenological Copenhagen view. As we said, this renewed appeal of the Bohmian theory is clearly motivated by a time when no engineer is really capable of accepting the "unspeakable" quantum reality [1, 2]. However, it must be noted that not even great parents of the quantum theory were ready to restrict themselves to the Copenhagen doctrine. For example, regarding the first section, von Neumann in his seminal book [16] explains that the collapse law is to be understood as an effective process that should be possible to be considered at an arbitrary point between the subsystem and the macroscopic device, instead of considering it to be a physical phenomenon [43]. Bohr himself assigned the collapse to the contextuality of experimental protocols in terms of macroscopic devices [44]. As we have reviewed, Bohmian mechanics satisfies the claims of both scientists. When it comes to the second section, it was J. M. Gambetta and H. M. Wiseman who pointed out that SSEs for non-Markovian systems tacitly implied the usage of CWFs from modal theories like Bohmian mechanics [11, 12] and who suggested the first formal position SSEs for such open quantum systems [19]. Finally, regarding the discussion on the unspeakables of the third section, Dirac himself was an exemplary physicist that employed "unspeakable unmeasured" system properties in the formulation of his major contributions to physics, leaving questioned the "observability doctrine" of the Copenhagen interpretation [44].

With all this, we might be wondering when will the mainstream decide to break the limiting walls around (non-relativistic) quantum mechanics, as taught to new generations of scientists every day. There is a pedagogical narrative (the Bohmian one) to explain it all while avoiding disjunctives with classical intuitions, a narrative that actually proves to be practically useful by offering additional tools to the Copenhagen theory. Will we someday include it in the standard program of quantum mechanics taught in our universities? Only time will tell.

References

1. X. Oriols, D.K. Ferry, Proc. IEEE **109**, 955 (2021)
2. D. Ferry, The Copenhagen Conspiracy (Jenny Stanford Publishing, 2019)
3. D. Bohm, Phys. Rev. **85**, 166 (1952)
4. P.R. Holland, The quantum theory of motion: an account of the de Broglie-Bohm causal interpretation of quantum mechanics (Cambridge University Press, 1993)
5. S. Teufel, D. Dürr, *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory* (Springer, Berlin, 2009)
6. X. Oriols, J. Mompart, *Applied Bohmian Mechanics: From Nanoscale Systems to Cosmology* (CRC Press, 2019)
7. D. Marian, N. Zanghì, X. Oriols, Phys. Rev. Lett. **116**, 110404 (2016)
8. G. Albareda, F. Traversa, A. Benali, X. Oriols, Fluct. Noise Lett. **11**, 1242008 (2012)
9. D. Pandey, E. Colomés, G. Albareda, X. Oriols, Entropy **21**, 1148 (2019)
10. Y. Aharonov, D.Z. Albert, L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988)

11. J. Gambetta, H. Wiseman, Phys. Rev. A **68**, 062104 (2003)
12. H.M. Wiseman, J.M. Gambetta, Phys. Rev. Lett. **101**, 140401 (2008)
13. D. Dürr, S. Goldstein, N. Zanghi, J. Stat. Phys. **67**, 843 (1992)
14. D. Dürr, S. Goldstein, N. Zanghì, J. Stat. Phys. **116**, 959 (2004)
15. X. Oriols, Phys. Rev. Lett. **98**, 066803 (2007)
16. J. Von Neumann, *Mathematische grundlagen der quantenmechanik*, vol. 38 (Springer, Berlin, 2013)
17. H.M. Wiseman, G.J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, 2009)
18. D. Dürr, S. Goldstein, R. Tumulka, N. Zanghì, Found. Phys. **35**, 449 (2005)
19. J. M. Gambetta, H.M. Wiseman, in *Fluctuations and Noise in Photonics and Quantum Pptics*, vol. 5111 (SPIE, 2003), pp. 313–324
20. H.M. Wiseman, Quant. Semiclass. Opt. Euro. Opt. Soc. Part B **8**, 205 (1996)
21. L. Li, M.J. Hall, H.M. Wiseman, Phys. Rep. **759**, 1 (2018)
22. K. Jacobs, D.A. Steck, Contemp. Phys. **47**, 279 (2006)
23. G. Albareda, H. López, X. Cartoixa, J. Suné, X. Oriols, Phys. Rev. B **82**, 085301 (2010)
24. G. Albareda, J. Suné, X. Oriols, Phys. Rev. B **79**, 075315 (2009)
25. H. López, G. Albareda, X. Cartoixá, J. Suñé, X. Oriols, J. Comput. Electron. **7**, 213 (2008)
26. E. Colomés, Z. Zhan, D. Marian, X. Oriols, Phys. Rev. B **96**, 075135 (2017)
27. M. Villani, G. Albareda, C. Destefani, X. Cartoixá, X. Oriols, Entropy **23**, 408 (2021)
28. Z. Zhan, X. Kuang, E. Colomés, D. Pandey, S. Yuan, X. Oriols, Phys. Rev. B **99**, 155412 (2019)
29. Z. Zhan, E. Colomes, X. Oriols, IEEE Trans. Electron Devices **64**, 2617 (2017)
30. G. Orlando, C. McDonald, N. Protik, G. Vampa, T. Brabec, J. Phys. B **47**, 204002 (2014)
31. E. Hauge, J. Støvneng, Rev. Mod. Phys. **61**, 917 (1989)
32. M. Perarnau-Llobet, E. Bäumer, K.V. Hovhannisyan, M. Huber, A. Acin, Phys. Rev. Lett. **118**, 070601 (2017)
33. J.M. Vilar, J.M. Rubi, Phys. Rev. Lett. **100**, 020601 (2008)
34. W. Niedenzu, M. Huber, E. Boukobza, Quantum **3**, 195 (2019)
35. D. Pandey, R. Sampaio, T. Ala-Nissila, G. Albareda, X. Oriols, Phys. Rev. A **103**, 052219 (2021)
36. N.D. Mermin, Rev. Mod. Phys. **65**, 803 (1993)
37. J. Bell, Phys. World **3**, 33 (1990)
38. D. Pandey, R. Sampaio, T. Ala-Nissila, G. Albareda, X. Oriols (2019)
39. H. Wiseman, New J. Phys. **9**, 165 (2007)
40. J.S. Lundeen, B. Sutherland, A. Patel, C. Stewart, C. Bamber, Nature **474**, 188 (2011)
41. D.H. Kobe, J. Phys. A: Math. Theor. **40**, 5155 (2007)
42. R. Sampaio, S. Suomela, T. Ala-Nissila, J. Anders, T. Philbin, Phys. Rev. A **97**, 012131 (2018)
43. L. Becker, Br. J. Philos. Sci. **55** (2004)
44. A. Oldofredi, M. Esfeld, Quantum **8**, 68 (2019)

Eigenstates in the Many Interacting Worlds Approach: Focus on 2D Ground States

Hannes Herrmann, Michael J. W. Hall, Howard M. Wiseman, and Dirk-André Deckert

Abstract The Many-Interacting-Worlds (MIW) approach to a quantum theory without wave functions proposed in [8] leads naturally to numerical integrators of the Schrödinger equation on comoving grids. As yet, little is known about concrete MIW models for more than one spatial dimension and/or more than one particle. In honour of Detlef Dürr, we report on a further development of the MIW approach to treat arbitrary degrees of freedom and provide a numerical proof of concept for ground states in 2d. The latter is part of a systematic numerical study [22] that includes also 1d ground and excited states. With this step towards the treatment of higher degrees of freedom we hope to stimulate their further study.

1 Introduction

The quantum dynamics of an N-particle system in d spatial dimensions is ruled by the Schrödinger equation. The latter defines the evolution of a field Ψ on the configuration space $\mathbb{R}^{Nd}$. A common method to represent such an object numerically is to sample it on a fixed grid. The grid dimensions scale as G^{Nd}, with G being the number of grid points along one degree of freedom. This exponential scaling makes numerics beyond $N = 3$ and $d = 3$ memory-wise infeasible even on modern computer architectures. Exceptional cases aside, the main two successful general approaches to deal with this quantum complexity problem are as follows. (1) For special initial values and in certain regimes (e.g., near product states and on certain density scales) one can find approximate solutions by solving non-linear one-particle equations such as the Hatree-Fock and Gross-Pitajevski equations; see, e.g., Ref. [1]. (2) One may employ a comoving grid that samples the wave functions with high resolution only where it has physically interesting features, while other regions are only covered with very

H. Herrmann · D.-A. Deckert (✉)
Mathematisches Institut, Ludwig-Maximilians-Universität München, München, Germany
e-mail: deckert@math.lmu.de

M. J. W. Hall · H. M. Wiseman
Centre for Quantum Dynamics, Griffith University, Brisbane, QLD 4111, Australia

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_10

few grid points; see, e.g., Refs. [2, 3]. The approach discussed in this paper belongs to class (2) and two central questions are how to find convenient locations for the grid-points and how to update them in parallel with the evolution of the wave function Ψ given by the Schrödinger equation:

$$i\partial_t \Psi_t(X) = \hat{H}\Psi_t(X), \qquad \hat{H} = -\frac{1}{2}\Delta + V(\hat{X}), \tag{1}$$

for $t \in \mathbb{R}$ and $X \in \mathbb{R}^{Nd}$, where Δ denotes the Laplace operator with respect to the configuration X, V is a classical potential, and we use units $t \mapsto t\hbar$ and $X \mapsto (\hbar/m^{1/2})X$ (for simplicity all particles are taken to have the same mass m).

One method is to distribute, say M, grid-points $Q^{(i)}_{t=0} \in \mathbb{R}^{Nd}$, $i = 1, \ldots, M$, at initial time $t = 0$ according to the $|\Psi_{t=0}|^2$-distribution. Thus, regions with larger contributions to the L^2-norm are sampled with higher resolution, while regions with smaller contributions are covered only by a few grid points. In order to ensure that the grid points follow the $|\Psi_t|^2$-distribution (a feature usually referred to as equivariance in the context of Bohmian mechanics [5]) one must transport them along the flux lines of the quantum probability current [4], i.e., along Bohmian trajectories $Q^{(i)}_t$, which obey the Bohmian law of motion [5]

$$\frac{dQ^{(i)}_t}{dt} = \Im \frac{\Psi^*_t(X)\cdot\nabla\Psi_t(X)}{\Psi^*_t(X)\cdot\Psi_t(X)}\bigg|_{X=Q^{(i)}_t}. \tag{2}$$

Hence, Bohm's velocity law (2) needs to be integrated simultaneously with the Schrödinger equation (1) on comoving coordinates $Q^{(i)}_t$, $i = 1, \ldots, M$. Using the decomposition $\Psi_t = \sqrt{P_t}e^{iS_t}$, the corresponding coupled set of equations (1)-(2) in comoving coordinates takes the form [2]:

$$\frac{d}{dt}P_t(Q^{(i)}_t) = -P_t(Q^{(i)}_t)\Delta S_t(Q^{(i)}_t), \qquad \frac{d}{dt}S_t(Q^{(i)}_t) = \frac{1}{2}\left(\frac{d}{dt}Q^{(i)}_t\right)^2 - V(Q^{(i)}_t) - U_t(Q^{(i)}_t), \tag{3}$$

$$\frac{d^2}{dt^2}Q^{(i)}_t = -\nabla V(Q^{(i)}_t) - \nabla U_t(Q^{(i)}_t) \tag{4}$$

with the initial velocity condition and the quantum potential U_t being, respectively

$$\frac{d}{dt}Q^{(i)}_t|_{t=0} = \nabla S_t(Q^{(i)}_t)|_{t=0}, \tag{5}$$

$$U_t(X) = -\frac{1}{2}\frac{\Delta P_t(X)^{1/2}}{P_t(X)^{1/2}}. \tag{6}$$

Note that constraint (5) together with (4) is equivalent to (2) while, thanks to (6), (3) are equivalent to (1), in the sense of the respective initial value problem. Numerical analysis of quantum systems with the help of trajectories has been studied in great depth and we refer the reader to the literature, e.g., [2, 3, 6, 7].

While mainly working on the mathematical foundations of physics, Detlef had an early interest in such practical applications of Bohmian Mechanics to the numerics of the Schrödinger equation and maintained contact with many colleagues in the quantum chemistry community working on trajectory-based approaches. It was this interest that led to the discussion between D.-A.D., M.J.W.H., and H.M.W., during a visit of D.-A.D. to Brisbane, from which Refs. [8, 9] and this work arose. H.M.W. noted that a further simplification might be obtainable when not only the grid points are distributed randomly according to $|\Psi|^2$ but when the $|\Psi|^2$-distribution itself can be approximately retrieved from the empirical distribution of the grid point locations $Q_t^{(i)}$, via some map $P(X; Q_t^{(1)}, \ldots, Q_t^{(M)})$ such that for all t, X the approximation

$$P_t(X) \approx P(X; \mathcal{Q}_t), \qquad \mathcal{Q}_t := (Q_t^{(1)}, \ldots, Q_t^{(M)}) \tag{7}$$

holds in a suitable sense as $M \to \infty$. In view of the weak law of large numbers, one may think of $P(X; \mathcal{Q}_t)$ as a smooth version of the empirical distribution

$$P(X; \mathcal{Q}_t) \approx \frac{1}{M} \sum_{i=1}^{M} \delta^{Nd}(X - Q_t^{(i)}), \tag{8}$$

as the $Q_t^{(i)}$, $1 = 1, \ldots, M$, stay approximately $|\Psi_t|^2$ distributed thanks to equivariance. Once a good candidate for $P(X; \mathcal{Q}_t)$ and its derivatives is identified, Eqs. (3)–(6) can be replaced by a closed system of equations for the trajectories $Q_t^{(i)}$, such as

$$\frac{d^2}{dt^2} Q_t^{(i)} = -\nabla \left[V(X) + U(X; \mathcal{Q}_t), \right] \Big|_{X = Q_t^{(i)}} \tag{9}$$

under the initial constraint (5), where now the density $P_t(X)$ in the quantum potential (6) is replaced by $P(X; \mathcal{Q}_t)$ so that the approximate quantum potential reads

$$U(X; \mathcal{Q}_t) = -\frac{1}{2} \frac{\Delta P(X; \mathcal{Q}_t)^{1/2}}{P(X; \mathcal{Q}_t)^{1/2}}. \tag{10}$$

Quantum expectation values of observable $f(\hat{X})$ can be recovered from the trajectories $Q_t^{(i)}$ by

$$\langle \Psi_t, f(\hat{X}) \Psi_t \rangle = \int d^{Nd}X \, P_t(X) f(X) \approx \int d^{Nd}X \, P(X; \mathcal{Q}_t) f(X) \approx M^{-1} \sum_{i=1}^{M} f(Q_t^{(i)}). \tag{11}$$

Due the possible interpretation of $Q_t^{(i)}$, $i = 1, \ldots, M$, as M coexisting "worlds" we follow Refs. [8, 9] in terming this the Many-Interacting-Worlds (MIW) approach.

Continuous versions of this general idea [10–13] predate our discrete MIW approach; see Refs. [14–17] for continuing interest.

The MIW approach stands or falls according to the possibility of finding a good candidate $P(X; \mathcal{Q}_t)$ and the ability to maintain the quality of the approximation (7) over time for M not too large. In Ref. [8] we have presented a surprisingly simple toy model for $N = 1$ particle in $d = 1$ dimension with the ansatz

$$P(Q^{(i)}; \mathcal{Q}_t) := \frac{1}{2}\left(\frac{1}{N(Q^{(i)} - Q^{(i-1)})} + \frac{1}{N(Q^{(i+1)} - Q^{(i)})}\right) \tag{12}$$

for $i = 1, \ldots, M$, with the ordering $Q_0^{(i)} < Q_0^{(i+1)}$ (setting $Q_t^{(0)} = -\infty$, $Q_t^{(M+1)} = +\infty$). The above ordering is preserved over time because the system (1) and (4) has a well-defined initial value problem [18], and hence, configuration space trajectories cannot cross. However, instead of approximating $U_t(X)$ directly via a smoothed $P(X; \mathcal{Q}_t)$, as in Eq. (10), the method in Ref. [8] approximates its average (proportional to the Fisher information of $P_t(X)$), via

$$\overline{U}_t = \int d^{Nd}dX\, P_t(X)U_t(X) = \frac{1}{8}\int P_t(X)\left|\frac{\nabla P_t(X)}{P_t(X)}\right|^2 \approx \frac{1}{8}\sum_{i=1}^{M}\left|\frac{\nabla P(Q^{(i)}; \mathcal{Q}_t)}{P(Q^{(i)}; \mathcal{Q}_t)}\right|^2. \tag{13}$$

Using Eq. (12) and the corresponding discrete approximation of $\nabla P_t(X)$ for $N = d = 1$, this leads to the replacement of Eq. (9) by the very similar form

$$\frac{d^2}{dt^2}Q_t^{(i)} = -\nabla_{Q_t^{(i)}}\left[V(Q_t^{(i)}) + U^{\mathrm{MIW}}(\mathcal{Q}_t)\right], \tag{14}$$

$$U^{\mathrm{MIW}}(\mathcal{Q}_t) = \frac{1}{8}\sum_{i=1}^{M}\left(\frac{1}{Q^{(i+1)} - Q^{(i)}} - \frac{1}{Q^{(i)} - Q^{(i-1)}}\right)^2. \tag{15}$$

The model defined by Eqs. (14) and (15) has the nice property of conserving total energy [8]

$$E = \sum_{i=1}^{M}\left[\frac{1}{2}\left(\frac{dQ_t^{(i)}}{dt}\right)^2 + V(Q_t^{(i)})\right] + U^{\mathrm{MIW}}(\mathcal{Q}_t), \tag{16}$$

and will be referred to as the 1d MIW model throughout this work. In Ref. [8], the 1d MIW model was shown to exhibit typical quantum behaviour such as superposition and tunnelling. In particular, numerical implementations of the model, with a very modest number of worlds, gave good qualitative agreement in the case of the time-evolution of two superposed Gaussians (representing double-slit interference). In addition, numerical testing showed good quantitative agreement for the compu-

tation of ground states, and convergence in the limit $M \to \infty$ has been proven for a harmonic potential [19].

The goal of this paper is to develop the MIW approach further, by explicitly treating more than one degree of freedom. Section 2 provides a general model for any finite number of degrees of freedom, i.e., finite particle numbers and spatial dimensions. In the spirit of this general approach we then present numerical algorithms for finding energy eigenstates. In [22], we benchmarked this general approach against the original 1d MIW model, using the harmonic and the Pöschl-Teller potentials, and furthermore presented an implementation capable of finding 1d eigenstates. Here, in a shortened version of that article, we focus on finding ground states to $d = 2$ dimensions in Sect. 3, in particular for the harmonic and Pöschl-Teller potentials. In comparison to the former, the latter is only weakly confining, which makes the lack of information at spatial infinity much more prominent in numerical simulations. We discuss how this problem can be addressed in our approach. The numerical methods and simulations reported here are based on results in [20]. Independent calculations by Sturniolo [23], for the ground states of higher dimensional systems in the framework of the MIW approach, will be commented on briefly in Sect. 4.

2 Generalization to Arbitrarily Many Degrees of Freedom

A formal extension of the 1d MIW model, to a system of N particles moving in d spatial dimensions, is given by retaining the equations of motion, Eq. (14), but generalizing Eq. (15) to

$$U^{\mathrm{MIW}}(\mathcal{Q}_t) = \frac{1}{8}\sum_{i=1}^{M}\left|\frac{\nabla P(Q_t^{(i)};\mathcal{Q}_t)}{P(Q^{(i)};\mathcal{Q}_t)}\right|^2, \tag{17}$$

for suitable approximations $P(Q_t^{(i)};\mathcal{Q}_t)$ and $\nabla P(Q_t^{(i)};\mathcal{Q}_t)$, of $P_t(Q_t^{(i)})$ and its derivative, respectively [8]. Here, we construct these approximations, based on two related approaches.

Triangulation method. The worlds or trajectories $Q_t^{(i)}$ lie in the D-dimensional configuration space $\mathbb{R}^{Nd}$ with $D := Nd$. This configuration space can be partitioned into a network of D-tetrahedra having the worlds as vertices, together with a single exterior region. For $D = 2$ such a triangulation is depicted in the left hand panel of Fig. 1 (purple lines), corresponding to a Delaunay triangulation [28]. Efficient algorithms are known for establishing such triangulations [28]. For a given triangulation, let $\{T_{i,j}\}$ denote the set of D-tetrahedra ('triangles') sharing $Q^{(i)}$ as a common vertex, at a given time t. Here we have dropped the explicit time label on $Q^{(i)}$ for convenience. Now, for a sufficiently smooth function $f(X)$ on configuration space one can, similarly to Eq. (11), approximate its average via

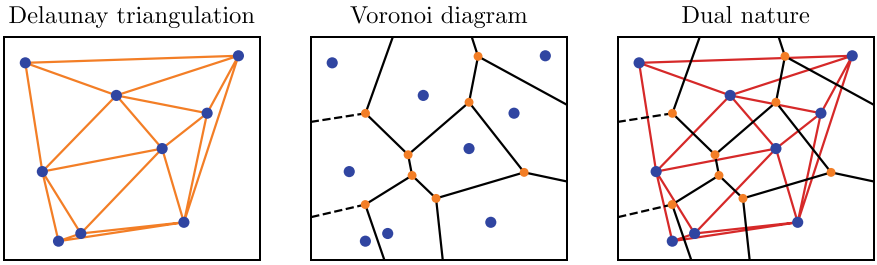


Fig. 1 The left hand panel shows a Delaunay triangulation of configuration space for a set of worlds (blue circles) at a given time. Each triangle (orange edges) is chosen in such a way that no worlds lie inside the circum-sphere of any triangle. The middle panel shows its dual graph, corresponding to partitioning configuration space into Voronoi cells (see definition (25) below). The corresponding cell boundaries are formed by hyperplanes (black lines) which bisect the triangulation lines. The right panel illustrates the duality of these graphs

$$\frac{1}{M}\sum_{i=1}^{M} f(Q^{(i)}) \approx \int dX\, P(X)\, f(X) = \frac{\sum_{i,j}\int_{T_{i,j}} dX\, P(X)\, f(X)}{D+1} \approx \frac{\sum_{i,j} |T_{i,j}|\, P(Q^{(i)})\, f(Q^{(i)})}{D+1},$$

where $|T_{i,j}|$ denotes the volume of the tetradron $T_{i,j}$, and the factor of $D+1$ arises because every tetrahedron is counted once for each of its $D+1$ vertices. Hence, equating these expressions for arbitrary $f(X)$, a suitable approximation for the probability density at $Q^{(i)}$ is given by

$$P_{\text{tri}}(Q^{(i)}; \mathcal{Q}) = \frac{D+1}{M\sum_j |T_{i,j}|}. \tag{18}$$

This reduces to Eq. (12) for the 1d MIW model when $D = Nd = 1$.

Cell method. An alternative to placing worlds at the vertices of a triangulation is to instead place each world within an individual cell, where the cells partition the configuration space. For example, for a Delaunay triangulation such cells can be chosen as the dual graph, corresponding to Voronoi cells [28]. An example is depicted in the right panel of Fig. 1. Note that some cells, corresponding to worlds at the edges, are infinite in extent. Such a partitioning leads to

$$P_{\text{cell}}(Q^{(i)}; \mathcal{Q}) = \frac{1}{M|\,\text{Cell}_i\,|} \tag{19}$$

an an alternative approximation for the probability density at $Q^{(i)}$, where $|\,\text{Cell}_i\,|$ denotes the configuration space volume of the cell containing trajectory $Q^{(i)}$.

Approximating the derivative of the probability density Eq. (17) further requires finding a suitable approximation for the derivative $\nabla P(Q^{(i)})$ at each time (where again we suppress the explicit dependence on time for convenience). This derivative has $D = Nd$ independent components, and hence we need to consider, for each trajectory $Q^{(i)}$ at that time, the change in probability in at least D different directions. These directions could be chosen, e.g., to be those which join $Q^{(i)}$ to its D closest neighbours, independently of the method used to estimate the density itself (e.g., via triangulation or cells). An alternative choice is to use the directions corresponding to all (or some) edges of a given triangulation which have $Q^{(i)}$ as a vertex. Yet another choice is to use the directions corresponding to all (or some) worlds that share a cell boundary with $Q^{(i)}$. Here we will remain general, and only suppose that $C_i \geq D$ neighbouring configurations or worlds are used to estimate $\nabla P(Q^{(i)})$, selected by any of the means above. Letting $\{Q^{(i,s)}\}$ denote these C_i configurations, we define the corresponding set of vectors $v^{(i,s)} := Q^{(i,s)} - Q^{(i)}$. By construction, these form a (typically overcomplete) basis set in configuration space. Now, writing $v^{(i,s)} = \sum_{k=1}^{D} A_{sk} e_k$ relative to some orthonormal basis set $\{e_k\}$, one has a corresponding set of dual basis vectors $\tilde{v}^{(i,s)} := \sum_k \tilde{A}_{sk} e_k$, with $\tilde{A} := A(A^\top A)^{-1}$. This dual basis satisfies the completeness property $\sum_s \tilde{v}^{(i,s)} (v^{(i,s)})^\top = I_D$, where $\top$ denotes the transpose and I_D is the $D \times D$ identity matrix. Using $P(Q^{(i,s)}) - P(Q^{(i)}) \approx v^{(i,s)} \cdot \nabla P(Q^{(i)})$, it follows immediately that one has

$$\nabla P(Q^{(i)}) = \sum_s \tilde{v}^{(i,s)} (v^{(i,s)})^\top \nabla P(Q^{(i)}) \approx \sum_s \left[P(Q^{(i,s)}) - P(Q^{(i)}) \right] \tilde{v}^{(i,s)}. \quad (20)$$

For any suitable approximation of $P(Q^{(i)}; \mathcal{Q})$, such as in Eqs. (18) or (19), one then has the corresponding approximation

$$\nabla P(Q^{(i)}; \mathcal{Q}) = \sum_s \left[P(Q^{(i,s)}; \mathcal{Q}) - P(Q^{(i)}; \mathcal{Q}) \right] \tilde{v}^{(i,s)} \quad (21)$$

of the derivative. This may now be inserted into Eq. (17) to obtain the corresponding MIW potential function $U^{\mathrm{MIW}}(\mathcal{Q}_t)$ at any given time t.

3 A Numerical Implementation for Ground States

While the approach given in Sect. 2 may be precise and general, we encountered several problems in its direct numerical implementation. For example, the construction of the dual basis set $\{\tilde{v}^{(i,s)}\}$ appearing in Eq. (21) requires computation of the inverse of the $C_i \times C_i$ matrix $(A^\top A)^{-1}$ for each of the M worlds. Recalling that $C_i \geq D = Nd$, this alone requires $O(MN^3d^3)$ calculations at each time step. Moreover, unlike the 1d case, the set of neighbouring configurations or worlds used to define triangulations, partitionings, derivatives, etc., can change over time, and thus requires constant updating. The selection of finitely many nearest neighbours inevitably provokes dis-

continuous changes throughout the dynamics. Though for very large M one may expect that these sudden jumps may have only little impact on the overall dynamics of worlds, for lower values of M these small discontinuities may cause oscillations in the world configurations $Q_t^{(i)}$ which, if not damped, propagate through the whole system until the numerical simulation breaks down. This forced us to replace this discrete notion of nearest neighbours, in computing the approximate $P_t(X)$ and its derivatives, by something more smooth. Enforcing some sort of smoothness may also come as no surprise: Even if the grid points may sample well regions in which $P_t(X)$, i.e., $|\Psi_t|^2$, is large, and thus, potentially increase the precision in the L^2-norm sense without the need of too many samples, the required precision in the pointwise sense in (9), i.e., (2) cannot be guaranteed, unless some prior knowledge on the smoothness is available.

It turns out that finding a smooth distribution that approximates the empirical distribution is an old problem, discussed thoroughly in the classical literature; see [24–27] for an overview. One general and, for many settings, very robust technique is so-called smooth kernel density estimation, which we introduce first. The density estimator for a given distribution of worlds $\mathcal{Q} = (Q^{(1)}, \ldots, Q^{(M)})$ is given by a sum of the form

$$P_h(X; \mathcal{Q}) := \frac{1}{\tilde{M}} \sum_{i=1}^{\tilde{M}} \frac{1}{h_i} K\left(\frac{X - \widetilde{Q}^{(i)}}{h_i}\right). \tag{22}$$

Here $\{\widetilde{Q}^{(i)}\}$ is a set of $\tilde{M}$ points in configuration space determined by $\mathcal{Q}$; the h_i are width parameters (usually referred to as bandwidths) similarly determined by $\mathcal{Q}$; and K is a smooth kernel function that fulfils $\int d^{Nd}X\, K(X) = 1$. Note that $P_h(X; \mathcal{Q})$ is automatically normalized. In this work we will focus on the Gaussian kernel $K(X) = (2\pi)^{-Nd/2} \exp(-\frac{1}{2} X^\top X)$, for which $\widetilde{Q}^{(i)}$ takes the role of a mean and $h_i^2 I_{Nd}$ defines a corresponding covariance matrix. Considering that the Schrödinger propagator is given by a Gaussian [5], this seems like a canonical choice.

The idea behind ansatz (22) is to allow for varying widths h_i, well-adapted to regions of high and low empirical density in the vicinity of suitably chosen locations $\widetilde{Q}^{(i)}$. If the empirical density is low in the neighbourhood of $\widetilde{Q}^{(i)}$, one chooses a large values of h_i (broad kernel function), and if the density is high, one chooses a small value (narrow kernel function). We will come back to the question of choosing optimal $\widetilde{Q}^{(i)}$ and h_i later, in Sect. 3.1. First we discuss how density estimation may be used in an algorithm for calculating ground state properties.

Gaussian kernel algorithm. Once the choice for the $\widetilde{Q}^{(i)}$ and h_i is settled, an algorithm for finding ground states can be given in terms of the following iteration:

1. Start with any initial distribution of M worlds $\mathcal{Q}_0 = \{Q_0^{(1)}, \ldots, Q_0^{(M)}\}$ and choose a suitably small time step $\Delta t > 0$.
2. From $\mathcal{Q}_0$, compute the approximate potential (10) in which the approximate density $P(X; \mathcal{Q}_t)$ is replaced by $P_h(X; \mathcal{Q}_0)$ given in (22).

3. Integrate the second order equation of motion (9) up to time Δt with zero initial velocities $\dot{Q}_0^{(i)} = 0$, to obtain a new empirical distribution $\mathcal{Q}_{\Delta t}$.
4. Replace $\mathcal{Q}_0$ by $\mathcal{Q}_{\Delta t}$ and go back to step 2 until a predefined stopping condition is met (e.g. given by an appropriate measure of convergence).

We shall refer to this algorithm as the *Gaussian kernel algorithm*. The numerical implementation used in this work is provided in [21].

A similar algorithm was discussed for the 1d MIW model in Ref. [8]. The reason why convergence can be expected is that in every integration step of (9) the initial velocities are set to zero. This introduces a loss of energy, as after each integration step 3 above the total energy

$$E_{\mathrm{kin}}(\Delta t) + E_{\mathrm{pot}}(\Delta t) = \sum_{i=1}^{M} \left[\frac{1}{2} \left(\dot{Q}_{\Delta t}^{(i)} \right)^2 - \int_0^{\Delta t} ds\, \dot{Q}_s^{(i)} \cdot \nabla [V(X) + U(X; \mathcal{Q}_s)]_{X=Q_s^{(i)}} \right] \tag{23}$$

is reduced by the positive kinetic energy E_{kin}. Hence, during the iteration of the algorithm the configuration of worlds $\mathcal{Q}$ will arrange itself to find a local minimum of $E_{\mathrm{pot}}(\Delta t)$. Providing that the potential $V(X)$ is confining, e.g., as in the case of a harmonic potential, it will work to focus the worlds, while the potential $U(X; \mathcal{Q})$ will work against clustering of worlds (cf. Ref. [8]). Since the integration time step Δt is small, and near a local minimum the velocities $\dot{Q}_s^{(i)}$ in Eq. (23) can also be expected to be small, a local minimum of E_{pot} should then fulfil

$$\nabla [V(X) + U(X; \mathcal{Q})]_{X=Q^{(i)}} \approx 0, \tag{24}$$

which according to the Bohmian equation of motion corresponds to a stationary state [5]. If $V(X)$ has only one local minimum one can therefore expect the algorithm to converge to a configuration of worlds $\mathcal{Q}$ that is distributed according to $|\Psi|^2$, where Ψ is the ground state wave function of the system with Hamiltonian $\hat{H}$ as per (1).

The main difference between the above algorithm, employing Gaussian kernels, and the MIW algorithm given in Ref. [8], is that the latter does not use a density estimator but instead computes forces as per (14), using the MIW potential (15), where the latter is conservative as per Eq. (16). One of the advantages of the Gaussian kernel model introduced here is that its form readily generalizes to any number of degrees of freedom Nd without sacrificing smoothness. In contrast, the form of (15) and its generalization via Eq. (17) depend on the use discrete derivatives, defined via finitely many neighbouring worlds, which leads to continuity issues as discussed at the beginning of this section.

3.1 Application to 2d Ground States

The Gaussian kernel algorithm defined in Sect. 3 is applicable to any number of degrees of freedom Nd but relies on good choices for the subdivisioning, $\widetilde{Q}^{(i)}$, and h_i. Any of the general methods in Sect. 2 is convenient in this regard, and we will follow the cell method based on Eq. (19). In order to use (19) we need to specify a subdivision of the configuration space into cells, and we shall use the Voronoi subdivision method [28] in our numerical implementation. For the configuration space $\mathbb{R}^{Nd}$ and a configuration of worlds $Q^{(i)} \in \mathbb{R}^{Nd}$, the Voronoi cell containing the world $Q^{(i)}$ is defined by

$$\text{Cell}_i := \left\{ X \in \mathbb{R}^{Nd} : \|X - Q^{(i)}\| < \|X - Q^{(j)}\| \, \forall j \neq i \right\} \tag{25}$$

We will call Cell_i an *inner cell* if it is bounded and an *outer cell* if it is unbounded. By definition the Voronoi cells form a subdivision of configuration space $\mathbb{R}^{Nd}$: $\bigcup_i \overline{\text{Cell}_i} = \mathbb{R}^{Nd}$ and $\text{Cell}_i \cap \text{Cell}_j = \emptyset$ for $j \neq i$. The Voronoi subdivision is well adapted to our problem of finding an *a priori* density such as (19) from an empirical distribution defined by $\mathcal{Q}$ as it naturally incorporates a measure of proximity in configuration space. We shall use

$$\widetilde{Q}^{(i)} = Q^{(i)} \text{ for } i = 1, \ldots, \widetilde{M} = M, \tag{26}$$

and in accordance with (19), enforce the corresponding *a priori* density constraint

$$P_h(\widetilde{Q}^{(i)}; \mathcal{Q}) \stackrel{!}{=} \frac{1}{M} \frac{1}{|\,\text{Cell}_i\,|} =: p_i \tag{27}$$

to find the widths h_i by means of the recurrence relation

$$h_i \leftarrow h_i P_h(\widetilde{Q}^{(i)}; \mathcal{Q})/p_i. \tag{28}$$

As a proof of concept, we tested this generalized Gaussian kernel algorithm for $N = 1$ particles and $d = 2$ spatial dimensions, in two cases: 1) A harmonic potential $V(X) = \frac{1}{2}\hbar^2\omega^2 X^\top X$ and 2) a Pöschl-Teller type potential. Note that, unlike the harmonic potential, the Pöschl-Teller potential has many more or less natural generalizations in more than one spatial dimension. For our proof of concept study we took the simple choice with $X = (x_1, x_2)$:

$$V(X) = V_1(x_1) + V_1(x_2), \text{ with } V_1(x) = \frac{\alpha^2}{2} \frac{\lambda(\lambda+1)}{\cosh^2(\alpha x)} \tag{29}$$

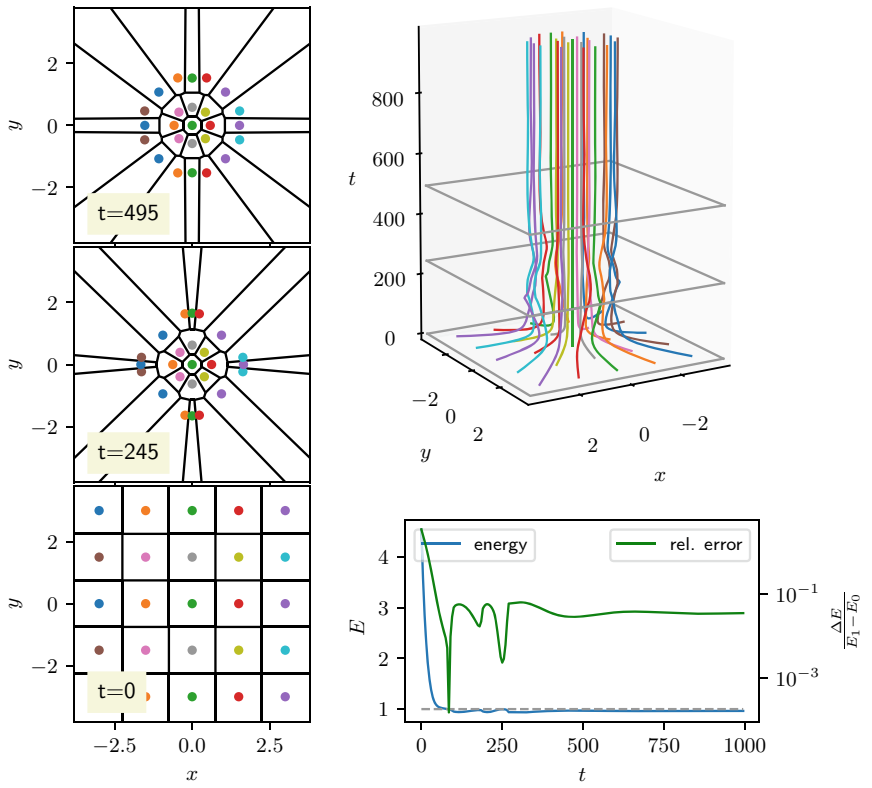


Fig. 2 Convergence of the Gaussian kernel algorithm for $M = 25$ worlds, $N = 1$ particle, and $d = 2$ spatial dimensions, in a harmonic potential $V(X) = \frac{1}{2}\hbar^2\omega^2 X^\top X$. The plot on the upper right shows the evolution of the configuration of worlds $\mathcal{Q}$ during the iteration of the algorithm. The integration time step was chosen to be $\Delta t = 5 \cdot 10^{-2}$ for $2 \cdot 10^4$ integration steps. The three plots on the left are configuration space snapshots for the respective times t shown in the lower left corner, respectively. The circles denote the worlds and the black lines illustrate their respective Voronoi cells. The plot in the lower right displays the convergence of the energy, again in units of $\hbar\omega$, as well as the relative error. The latter is computed as ratio of the difference of the difference w.r.t. the exact ground state energy (dashed line) and the difference of the exact energies of the first excited and the ground state. Note that in 2d the excited and ground state energies are given by $E_1 = 2\hbar\omega$, and $E_0 = \hbar\omega$

The results of the corresponding numerical simulations, for $M = 25$ worlds, are reported in Figs. 2 and 3, respectively, showing fairly good convergence to corresponding ground state configurations, and likewise, to the expected ground state energy.

As discussed in [22], for $Nd = 1$, there is a more convenient choice for (26) that helps to avoid crossings of world trajectories that are forbidden by their uniqueness according the Bohmian velocity law (2). This non-crossing property of worlds is, of course, far less of an issue for $Nd > 1$. However, one has to consider a potentially

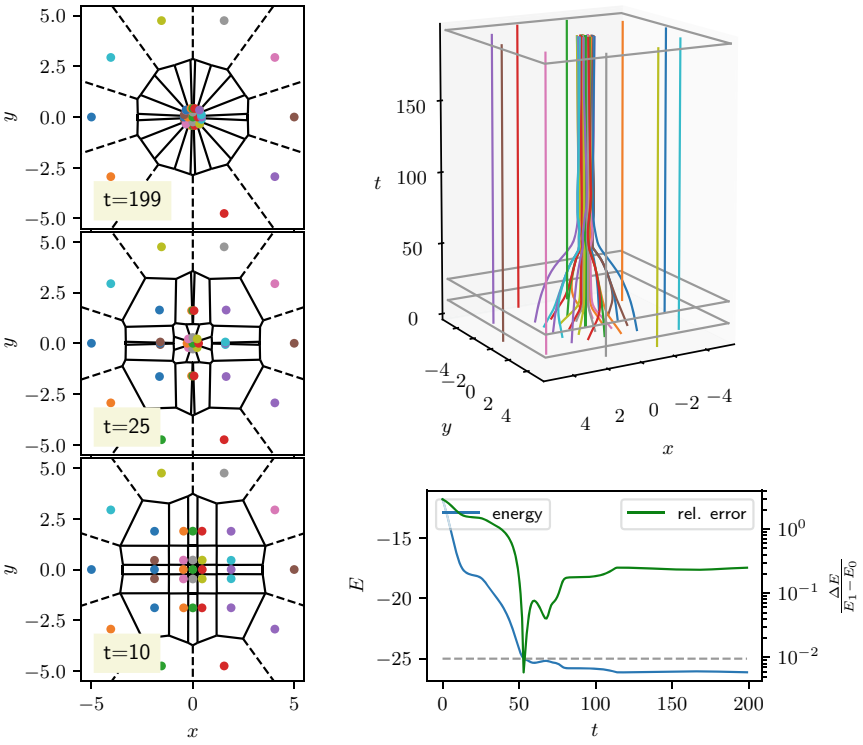


Fig. 3 Convergence of the Gaussian kernel algorithm for $M = 25$ worlds, $N = 1$ particle, and $d = 2$ spatial dimensions, in a Pöschl-Teller type potential defined as per Eq. (29) with $\lambda = 5$. The plot on the upper right shows the evolution of the configuration of worlds $\mathcal{Q}$ during the iteration of the algorithm. The integration time step was chosen to be $\Delta t = 2 \cdot 10^{-2}$ for 10^4 integration steps. The three plots on the left are configuration space snapshots for the respective times t shown in the lower left corner, respectively. The circles denote the worlds and the black lines illustrate their respective Voronoi cells. Note the boundary worlds which, as discussed in the text, have been fixed to lie on a circle of sufficiently large radius to stabilize the iteration. The plot in the lower right displays the convergence of the energy, again in units of $\alpha^2\hbar^2/m$, as well as the relative error. The latter is computed as ratio of the difference of the difference w.r.t. the exact ground state energy (dashed line) and the difference of the exact energies of the first excited and the ground state. Note that in 2d the excited and ground state energies are given by $E_1 = -20.5 \cdot \alpha^2\hbar^2/m$, and $E_0 = -25 \cdot \alpha^2\hbar^2/m$. The computed energy lies systematically below the exact ground state energy. As discussed, this systematic error is caused by the artificially fixed boundary worlds

more serious problem concerning the boundary worlds in the outer cells, for which the *a priori* distribution in Eq. (27) reduces to an uninformative value of $p_i = 0$, independently of the actual positions of the boundary worlds. In the case $Nd = 1$ there are only two boundary worlds $Q^{(1)}$ and $Q^{(M)}$ whereas, e.g., in our setup for the harmonic potential, with $Nd = 2$ and $M = 25$, we have 16 boundary worlds as depicted in Fig. 2.

In the case of the harmonic potential the boundary worlds were not found to be problematic, essentially because while their motion is not moderated by other worlds, the strongly-confining nature of the potential does not allow any of the worlds to escape to spatial infinity. In contrast, the Pöschl-Teller potential is asymptotically constant, and hence does not confine the boundary worlds. Due to that fact the corresponding numerical simulation easily becomes unstable. However, a straightforward solution to circumvent this problem is to introduce additional artificial worlds at fixed positions surrounding the actual worlds $\mathcal{Q}$; see the straight trajectories plotted in Fig. 3.

These artificial boundary worlds act to damp any unwanted oscillations of the outer worlds of $\mathcal{Q}$ but on the other hand encode a kind of boundary condition on the Hamiltonian at hand. Hence, these boundary worlds must be placed at sensible locations, having a sufficient distance to the actual worlds $\mathcal{Q}$, so that the accuracy of the world distribution is only changed in regions of configuration space where the density should in any case be very low. One may therefore expect that the accuracy of the numerically inferred moments of observables are not significantly affected. We have not tried to optimize the location of the boundary worlds in our first trial in Fig. 3, which is why the numerically determined value of the ground state is systematically smaller that the exact one.

All these troubles seem to be connected to the discrepancy between the two required approximation modes, i.e., in the L^2-norm sense, required for the statistics, and the one in the point-wise sense, required to obtain the world trajectories. Our choices made above in terms of subdivisioning methods, approximation kernels and their corresponding parameters, can be seen as forms of relieving this discrepancy through specification of a priori knowledge about the smoothness. These phenomena would of course have to be studied in more detail, however, our analysis already indicates that also in more than one spatial dimension one may expect our proposed approach to be applicable to ground and excited states for various potentials.

4 Conclusions

In this paper, we have generalized the original MIW algorithm [8] to treat a configuration space of more than one dimension, by using smooth kernel density estimation to approximate the empirical probability distribution. Our numerical implementation shows that this general algorithm performs well for calculating ground state energies and configurations for the harmonic and Pöschl-Teller potentials in two dimensions. Its application to both ground states and excited states of one-dimensional systems has been explored elsewhere [20, 22].

Kernel density estimators have also been recently applied by Sturniolo within the framework of the MIW approach [23], using a different method to construct the estimator, with numerical calculations yielding promising results for ground states of harmonic and Lennard-Jones potentials in two and three dimensions. Sturniolo further suggests that exponential kernels may perform better than Gaussian kernels

for finding ground state energies, but worse for finding ground state configurations, and that it may be possible to simulate temperature-dependent tunnelling effects in the MIW approach [23].

The original motivation in using the MIW approach in numerical computations was the hope of a generally-applicable method that reduced computational resources as compared to fixed-grid methods. At first sight, however, the provided numerical implementation still seems to be computationally expensive: even when neglecting the iteration that determines the bandwidths h_i in each integration step as per Eq. (28), the computational effort scales at least quadratically in the number of worlds M, since M contributions to the potential must be calculated for each world via Eq. (22). Fortunately there are many possible tricks to reduce this scaling, and indeed it has been claimed for general kernel density estimation approaches that this scaling can be reduced to one linear in M [27]. In our context the underlying idea for this claim corresponds to the fact that, due to the choice of bandwiths h_i, the Gaussian kernel functions are usually highly peaked in regions where many worlds cluster, so that the corresponding kernel functions fall off rapidly and the sum in the density estimator (22) can be truncated. Computational effort can be further reduced via a fixed lookup table for the kernel function, that is interpolated according to the scaling introduced by h_i. Finally, while the iteration to determine the h_i in Eq. (28) usually converges with very few iterations (since the configurations of the worlds $\mathcal{Q}$ change only slightly between the integration steps), it may also be possible to determine the h_i dynamically from the world configurations at each time step. This has been explored for kernel estimation with dynamics ruled by the heat equation [29], and similar techniques may apply to dynamics ruled by the Schrödinger equation.

In conclusion, therefore, development of the approach in this paper may lead to a more general and efficient numerical tool for ground state and other calculations, that tempers the exponential scaling of fixed-grid methods while maintaining the same numerical accuracy.

Acknowledgements We would like to thank M. Ghadimi and T. Gould for valuable discussions. Furthermore, D.-A.D. and H.H. would like to thank Griffith University for its hospitality, while M.J.W.H. and H.M.W. likewise thank the Ludwig Maximilian University. This work was partially funded by the Elite Network of Bavaria through the Junior Research Group "Interaction between Light and Matter" and by FQXi Grant FQXi-RFP-1519. Finally, since this is H.M.W.'s only contribution to this volume, I (H.M.W.) would like to take this opportunity to thank Detlef Dürr for being a source of wisdom and inspiration over many years, and for being a powerful friend and colleague for the unfortunately fewer years that I knew him personally.

References

1. E.H. Lieb, R. Seiringer, J.P. Solovej, and J. Yngvason, *The Mathematics of the Bose Gas and its Condensation.* Oberwolfach Seminars, Vol. 34, Birkhäuser (2005)
2. R.E. Wyatt, *Quantum Dynamics with Trajectories* (Springer, New York, 2005)
3. P. K. Chatteraj, *Quantum Trajectories* (CRC Press, 2010)

4. D.-A. Deckert, D. Dürr, P. Pickl, Quantum dynamics with bohmian trajectories. J. Phys. Chem. A **111**, 10325 (2007)
5. D. Duerr, S. Teufel, *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory* (Springer, Heidelberg 2009)
6. A.S. Sanz, S. Miret-Artés, Quantum mechanics with trajectories, in *A Trajectory Description of Quantum Processes. I. Fundamentals*. Lecture Notes in Physics (Springer, Berlin, Heidelberg, 2012), pp. 187–230
7. A. Benseny, G. Albareda, Á.S. Sanz, J. Mompart, X. Oriols, Applied bohmian mechanics. Eur. Phys. J. D **68**(10), 286 (2014)
8. M.J.W. Hall, D.-A. Deckert, H.M. Wiseman, Quantum phenomena modeled by interactions between many classical worlds. Phys. Rev. X **4**, 041013 (2014)
9. M. Ghadimi, M.J.W. Hall, H.M. Wiseman, Nonlocality in bell's theorem, in bohm's theory, and in many interacting worlds theorising. Entropy **20**, 567 (2018)
10. P. Holland, Computing the wavefunction from trajectories: particle and wave pictures in quantum mechanics and their relation. Ann. Phys. **315**, 505 (2005)
11. B. Poirier, Bohmian mechanics without pilot waves. Chem. Phys. **370**, 4 (2010)
12. G. Parlant, Y.C. Ou, K. Park, B. Poirier, Classical-like trajectory simulations for accurate computation of quantum reactive scattering probabilities. Computat. Theoret. Chem. **990**, 3 (2012)
13. J. Schiff, B. Poirier, Quantum mechanics without wavefunctions. J. Chem. Phys. **136**, 031102 (2012)
14. C. Sebens, Quantum mechanics as classical physics. Philos. Sci. **82**(2), 266–291 (2015)
15. K.J. Boström, Quantum mechanics as a deterministic theory of a continuum of worlds. Quantum Stud.: Math. Found. **2**, 315 (2015)
16. L. Smolin, Quantum mechanics and the principle of maximal variety. Found. Phys. **46**, 736 (2016)
17. P. Roser, M.T. Scoggins, Non-quantum behaviors of configuration-space density formulations of quantum mechanics. arXiv:2303.04959
18. K. Berndl, M. Daumer, D. Dürr, S. Goldstein, N. Zanghì, A survey of bohmian mechanics. Il Nuovo Cimento B (1971–1996) **110**, 737 (1995)
19. I.W. McKeague, B. Levin, Convergence of empirical distributions in an interpretation of quantum mechanics. Ann. Appl. Probab. **26**, 2540–2555 (2016)
20. H. Herrmann, Finding stationary states by interacting many worlds. Master Thesis, Mathematical Institute of the LMU Munich (2016)
21. H. Herrmann, D.-A. Deckert, Eigenstates in the many interacting worlds approach, GitLab repository (2023). https://gitlab.com/dirk-deckert-lmu/eigenstates-in-the-Many-Interacting-Worlds-approach
22. H. Herrmann, M.J.W. Hall, H.M. Wiseman, D.-A. Deckert, Eigenstates in the many interacting worlds approach: ground states in 1D and 2D and excited states in 1D (long version). arXiv:1712.01918
23. S. Sturniolo, Computational applications of the many-interacting-worlds interpretation of quantum mechanics. Phys. Rev. E **97**, 053311 (2018)
24. A.J. Izenman, Review papers: recent developments in nonparametric density estimation. J. Am. Stat. Assoc. **86**, 205–224 (1991)
25. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*, 2nd edn. (Hoboken, New Jersey: Wiley, Inc, 2015)
26. I.S. Abramson, On bandwidth variation in kernel estimates-a square root law. Ann. Stat. **10**, 1217–1223 (1982)
27. A. Elgammal, R. Duraiswami, L.S. Davis, Efficient kernel density estimation using the fast gauss transform with applications to color modeling and tracking. IEEE Trans. Pattern Anal. Mach. Intell. **25**, 1499–1504 (2003)
28. J.A. De Loera, J. Rambaau, F. Santos, Triangulations. Algorithms and Computation in Mathematics, vol. 25 (Springer, Berlin, 2010)

29. Z.I. Botev, J.F. Grotowski, D.P. Kroese, Kernel density estimation via diffusion. Ann. Stat. **38**(5), 2916–2957 (2010)
30. R. Courant, D. Hilbert, *Methoden der mathematischen Physik* (Springer, Berlin, 1924)
31. A. Ancona, B. Helffer, T. Hoffmann-Ostenhof, Nodal domain theorems a la courant. Doc. Math. **9**, 283–299 (2004)

Are There Observational Differences Between Bohmian Mechanics and Other Interpretations?

Lev Vaidman

Abstract While there is a consensus that leading interpretations of quantum mechanics cannot be distinguished with today's technology, it is unclear if a gedanken experiment which relies on unlimited technological power cannot do so. Another gedanken approach is considering sentient beings which have brains different from ours. Such gedanken situations will be analyzed with emphasis on a possible difference between Bohmian mechanics and the many-worlds interpretation.

1 Introduction

I can see a parallel between Detlef Dürr's and my own work on the interpretations of quantum mechanics. For both of us this was a central part of our research and we both believed that there is a single interpretation which is much better than others. However, while Detlef had no doubt about the superiority of Bohmian mechanics (BM) [1, 2], I am certain that the many-worlds interpretation (MWI) [3] is by far the best. The term "interpretation" might not be precise: different interpretations of quantum theory are sometimes actually different theories. In this paper, I want to shed light on possible observational differences between different interpretations and, in particular, between BM and MWI, although it is not clear that such differences exist.

We make our observations using our senses which provide our experiences, and the next section defines the connection between ontology and experience. In Sect. 3 I start the analysis by gedanken attributing conscious experience to a microscopic particle, a neutron. The advantage is that we can consider experiments which are performed in laboratories. Section 4 is devoted to similar experiments with macroscopic sentient beings. Here, the gedanken story is the possibility of performing such experiments. The next level of gedanken consideration in Sect. 5 is to consider macroscopic sentient beings with brains operating using spin states.

L. Vaidman (✉)
Raymond and Beverly Sackler School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv 69978, Israel
e-mail: vaidman@tauex.tau.ac.il

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_11

2 Experience Supervenience Postulates

In my view, a theory (interpretation) consists of two ingredients: an ontology described in mathematical terms and a rule which provides correspondence between the mathematical formalism and our experience. I do not presuppose any complicated/refined meaning of the "experience" beyond a basic physicalist notion, where there is no conceptual difference between my experience and the experience of a robot equipped with various sensors.

For some theories, the rule connecting ontology and experience seems trivial and is rarely explicitly mentioned. For example, Newtonian mechanics supervenience postulate:

> The experience of a sentient being supervenes on the position and velocity of the particles it consists of.

Standard textbook quantum mechanics (which includes collapse of the wave function at every quantum measurement) supervenience postulate:

> The experience of a sentient being supervenes on the wavefunction of its degrees of freedom.

For BM and especially for MWI, the supervenience rule is an essential part of the interpretation.

The ontology of the MWI is the universal wave function, period. The same universal wave function is also part of the ontology of BM (although Bohmians often attach to it a lower status [1]). The BM supervenience postulate is necessary to avoid multiple worlds in BM.

The MWI supervenience postulate:

> The experience of a sentient being supervenes on the wavefunction of its degrees of freedom within the world branch of the wave function of the universe.

It is the same postulate as in the textbook quantum mechanics, in which there is only one branch of the universal wave function.

For BM I suggest considering two possible postulates. The BM supervenience postulate I:

> The experience of a sentient being supervenes on the Bohmian positions of the particles the sentient being is made of.

The BM supervenience postulate II:

> The experience of a sentient being supervenes on the Bohmian collapsed wavefunction of its degrees of freedom.

Another possible proposal is supervenience on both Bohmian position and the Bohmian collapsed wave function. However, I feel that every one of the ingredients, Bohmian positions or Bohmian collapse wave function is enough to explain our experience, so this proposal seems to be unreasonably complicated.

3 Experiences of a Neutron

I want to start by analyzing a neutron Mach-Zehnder interferometer (MZI), a device that already four decades ago was used as a test bed for the strange behavior of quantum systems [4], see Fig. 1. In the past I analyzed such an interferometer attaching "consiousness" to the neutrons and arguing that we need the MWI to avoid schizophrenic experiences of neutrons in the interferometer [5]. Quantum physics attributes two paths for neutrons inside the interferometer, which are necessary for explaining the interference. The equations of quantum mechanics tell us that when a neutron reaches a beam splitter, its quantum wave splits into two parts. This is observed in numerous experiments. An experimentalist can tune the interferometer described in Fig. 1a such that no neutrons reach the detector D_2. Moreover, one of the ways to tune the neutron interferometer is to put a line of charges between the arms. This is the Aharonov-Casher effect [6, 7] the topological character of which leaves no other choice but to accept that the neutron must be in two arms and "experience" different forces in these two arms. To avoid schizophrenic neutrons, the MWI postulates that within a world neutrons cannot have distinct experiences, i.e., that from the moment a neutron enters the interferometer and until leaving it, there are two worlds for the neutron: in one world it takes arm A and in another, arm B.

Bohmian mechanics avoids schizophrenia of the neutron without multiple worlds by adopting one of the experience postulates. In the MWI framework, in the neutron MZI experiment there are two "neutron worlds", while in the BM there is only one neutron world. However, we cannot state that there is an observational difference. The experiences of the neutron in the BM are identical to those of the neutron in one of the MWI worlds. In the framework of the MWI we do not have direct observational evidence for the existence of multiple worlds, and both worlds of the MWI are possible Bohmian worlds. So there is no neutron passing an MZI, which has evidence for one interpretation and not the other.

Bell [8] was the first to recognize a strange behavior of the Bohmian trajectories in MZI without the second beamsplitter, see Fig. 1c. When the Bohmian particle moving in one arm of the MZI reaches the place O where the second beamsplitter has been, it is "caught" by the empty wave moving in the other arm and changes its velocity without any physical fields in this place. In this experiment, the history of experiences of the neutron in BM is different from any of the histories of experiences in the two neutron worlds of the MWI. The difference is in the BM jump from one MWI world to another. However, the jump and the history of experiences are not "written" in the memory of the neutron, so there is no moment of time in which the neutron can distinguish between the BM and the MWI. Indeed, although it is suggestive to assume that the neutron has different experiences when it accelerates, bouncing off a mirror or jumping from one wave packet to another at point O, the physicalism requires that the experience of a sentient being is given by a model of her brain (or the central processing unit). If the neutron does not have an internal structure, it cannot have experiences, while adding an internal structure of the neutron complicates the experiments. Indeed, a neutron, having memories of different experiences in arms

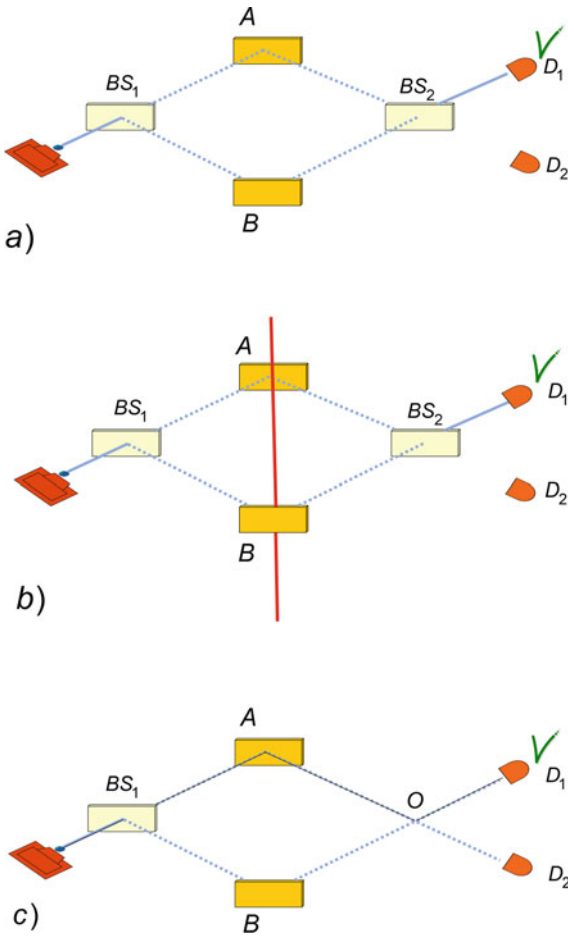


Fig. 1 Schematic picture of neutron Mach-Zehnder interferometer. (In laboratory implementations beamsplitters and mirrors are just parts of a symmetrically cut perfect crystal.) **a**) Neutron interferometer tuned to destructive interference toward detector D_2. **b**) Aharonov-Casher effect. The phase of the neutron interferometer is tuned by changing the charge density of the line of charges passing through the interferometer. **c**) Bohmian trajectory (continuous black line) in the interferometer without second beam splitter

A and B, will not interfere in the output ports of the MZI and we will not see the Aharonov-Casher effect. However, the interference was important when we argued for the necessity of the MWI but not for the question of observational differences between the MWI and the BM.

Consider a natural model of the neutron experience, its spin. (In BM this model requires accepting supervenience postulate II, i.e. that experience supervenes on the Bohmian collapsed spin wave function.) Let us put a magnetic field on path B which flips the spin of the neutron, and thus the spin will provide a memory of the experience in different arms. Adding interacting spin spoils the interference of the MZI, but does

not change the behavior of Bohmian particles as described in Fig. 1c. The Bohmian particle still jumps to the empty wave accelerating in a place without fields. The MWI world and the BM worlds have different histories, but again, there is no moment of time when the neutron, within a world, has evidence of the difference between BM and the MWI.

In order to consider the BM supervenience postulate I, we can replace the neutron by an atom and consider the experience at different paths recorded in different Bohmian positions of the particles the atom consists of. But if the internal quantum states of wavepackets moving on path A and path B differ, the jump of the Bohmian positions to the empty wave packet will not happen, and so the experience in the possible BM worlds will be identical to the experiences in the atom's world of the MWI. If using a supertechnological device we erase the memory in the wave packets of the neutron just before they reach the meeting point O, the histories in the BM world and the worlds of the MWI will be different, but the neutron will not have the memory to verify this. In all cases we do not have any observational difference.

In fact, for analyzing the MWI, even a sentient neutron or atom is a very problematic example because they are not macroscopic. Experience of a sentient being is defined only within a world, since in different worlds sentient beings have different experiences. According to my definition [3]:

> *A world is the totality of macroscopic objects: stars, cities, people, grains of sand, etc. in a definite classically described state.*

Important aspects of the problem cannot be considered with microscopic objects. Adding a microscopic object to the description of a world leads to a very different behavior [9, 10].

4 Wigner's Cat

Although there are (few) claims to the contrary, I am not aware of any realistic experiments which can distinguish between different interpretations of quantum mechanics (apart from constraining parameters of some physical collapse theories [11]). I, however, think that it is important to consider the possibility of having observational differences in gedanken experiments requiring technology that is not present today and might not even be present in any foreseeable future.

A gedanken experiment that has attracted renewed attention is the Wigner friend [12, 13]. Despite the alleged experimental demonstrations [14] I do not expect that the experiment will be performed in a real laboratory. Wigner was supposed to measure his macroscopic friend in superposition of macroscopically different states. In [14] the "friend" is a photon, so such an experiment is not better than the neutron interference experiment discussed above.

Let us consider Wigner's friend to be his cat. The cat is macroscopic enough and sentient enough, especially since Wigner trained his cat to be an observer in a spin experiment. The cat stands up if the detector corresponding to outcome "up" clicks

and lays down if detector "down" clicks. At time $t = 0$ the Stern-Gerlach experiment measuring spin in the z direction of a particle with initial state $|\uparrow_x\rangle$ is performed. The cat observes the result and acts according to her training.

Consider the following set of Wigner measurements. First, immediately after the procedure, he measures the observable of the lab with the cat which has two eigenstates: $|+\rangle$ and $|-\rangle$, where

$$|\pm\rangle \equiv \frac{1}{\sqrt{2}}(|cat\ stands\ up\rangle|\uparrow_z\rangle \pm |cat\ lays\ down\rangle|\downarrow_z\rangle. \tag{1}$$

Then, Wigner keeps the lab isolated and repeats the same measurement (with appropriate changes due to time evolution) every minute.

If the correct theory describing the universe is quantum mechanics with collapses when macroscopic objects are in superposition of macroscopically different states, then the possible results of Wigner are $+, +, -, +, -, ...$ This is because states $|+\rangle$ and $|-\rangle$ are superpositions of the cat standing up and laying down, so during the sixty seconds between Wigner's measurements they will collapse either to state $|cat\ stands\ up\rangle$ or to state $|cat\ lays\ down\rangle$ after which there is an equal probability for results "+" and "−" of Wigner's measurements.

If the correct theory describing the universe is the MWI or BM, then, Wigner's results are deterministic: $+, +, +, +, +, ...$ At time $t = 0$ the state $|+\rangle$ is prepared and evolves unitarily to its version at later times. So Wigner (given that he has supertechnology) can distinguish between collapse and non-collapse theories (see also Sect. 5 of [3]). He cannot distinguish between the BM and the MWI.

The cat can also be considered as an observer. In the framework of the MWI, at every moment there is a cat experiencing standing up and a cat experiencing laying down. In the framework of the BM one of these cats is an empty wave which has no Bohmian positions and does not have any experience. Depending on the way Wigner performs his measurements, in the BM we might have only one type of experience for the cat (say, standing up), or it might change due to the process of Wigner's measurement. (The latter can happen if Wigner performs an interference experiment after bringing the two wave packets of the cat to the same location.) However, Wigner, in order to perform his measurements, has to erase the memory of the cat, so there is no moment in time at which the cat has evidence about what is the right theory, the BM or the MWI.

5 Sentient Being with a Spin Brain

In Sect. 3 we already considered a sentient neutron with a spin brain, but a proper analysis requires macroscopic objects. Indeed, in BM one can talk about worlds that differ due to microscopic differences of the Bohmian positions of particles, but in the MWI the concept of a world requires macroscopic differences of macroscopic objects.

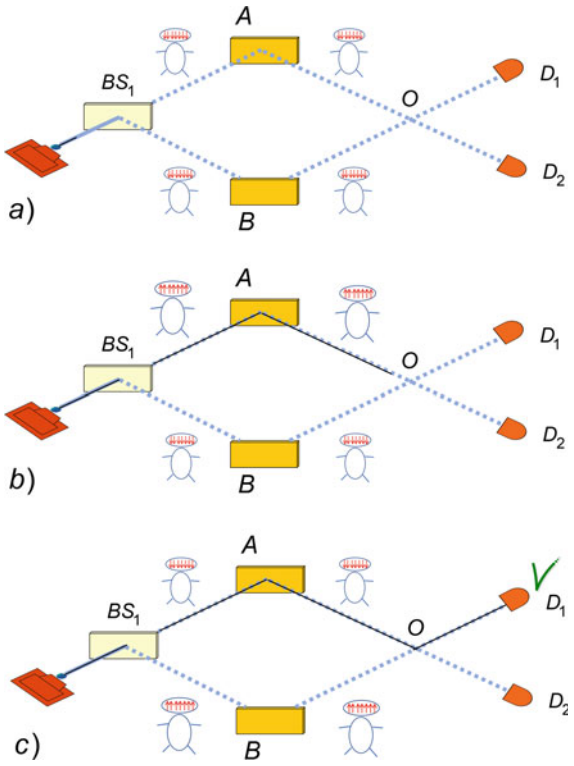


Fig. 2 Surrealistic trajectories of the neutron observed by sentient robots with spin brains. Analysis of the experience of robots according to the BM supervenience postulate II. **a**) The neutron enters the MZI interferometer without the second beamsplitter, and sentient spin-brain robots are ready to observe its trajectory. **b**) The neutron is inside the interferometer close to the point O. The robots on the path A detected the passing neutron (their spin-brain states flipped), while the robots on the path B remained in the ready state. **c**) The neutron passed the point O and was detected by a standard detector D_1. At the moment that the neutron passed the point O, the spin brain memories of the robots on the path A were erased and returned to "ready" while the spin brain states of the robots on the path B changed to incorrect records of the neutron passing there

Current brain studies do not suggest that our brain works with spin states, but we can imagine a sentient macroscopic robot with brain based on the *macroscopic* number of spin states. A particularly surprising situation will occur if this robot is placed in one arm of the MZI without the second beam splitter to observe the passing neutron there, as in Fig. 1c. Let us consider a world in which the neutron is detected on the detector D_1. Although the robot is macroscopic, if the observation of the neutron changes only its spin states, say flip them from "down" to "up", then from the Bohmian perspective no position measurement has been performed in the arms of the interferometer and, therefore, when the full and empty wave packets of the neutron meet at location O, the Bohmian particle has to jump from one wave packet to another. Therefore, the neutron (detected by D_1) had a Bohmian trajectory along

the path A. We get records of the robot telling us that the neutron took the lower path B, while the Bohmian trajectory of the neutron is the upper path A. This situation is called a surrealistic trajectory [15].

In fact, in this setup we can place more robots with spin brains, see Fig. 2 which will all agree after the click of the detector D_1, that the neutron passed through arm B. This does not fit the BM supervenience postulate I, according to which experience supervene solely on Bohmian positions, while here it is the spin states of the brains which "know" the result. We should say that the robots have no experience of seeing a neutron. If we accept the BM supervenience postulate II, we should say that the robots have experience of seeing the neutron in arm B, but they are all mistaken, because the Bohmian trajectory of the neutron was in path A, see Fig. 2c. This is also where the BM collapsed wave of the neutron *was* in the past. Indeed, when the neutron was inside the interferometer, the collapsed wave functions of the robot brains were different, they corresponded to the neutron passing through A, see Fig. 2b. The records of the brains changed to a neutron passing through B when the neutron passed point O, see Fig. 2c.

In my view, the change of brain records is a weakness of the interpretation, but it is not something that has observable consequences. At every moment in time the spin-brain records in BM corresponded to the records in one of the worlds of the MWI, so at no moment in time had the robot evidence for one interpretation against the other.

6 Summary

If we accept the starting point of BM that in the end of the day all quantum measurements are measurements of position of the pointers of the measuring devices (this approach leads to experience postulate I), then the theorem of BM about the robustness of the Born distribution of Bohmian positions under unitary evolution tells us that there cannot be an observational evidence distinguishing BM from the MWI. Ingenious proposals leading to surrealistic trajectories of the type described above can also be constructed without spins. Surrealistic trajectories appear when local interaction leads to a change in the quantum state (such as acquiring momentum) without an immediate significant change in Bohmian positions [16]. The general statement is that slow measuring devices, which show the results after the empty and full wave packets pass the intersection point, provide incorrect records of Bohmian positions of the particles. It might be slightly disturbing that only Bohmians would claim that such devices are not good measuring devices of position (in all other interpretations, which have no surrealistic trajectories, the measuring devices show correct histories), but no observational differences appear, only different interpretations.

Supertechnology, a la Wigner, cannot help. The basic supertechnology experiment, which allows distinguishing collapse and non-collapse interpretations involves interference between different worlds of the MWI, e.g. an interferometric device like an MZI, but with macroscopic objects (sentient observers) instead of particles, does

not distinguish between BM and the MWI. The MZI without a second beam splitter (the basis of surrealistic trajectories experiments) is simpler than a MZI. It is conceptually different from the interference experiment since there is no interference between the *A* and *B* branches. However, when the branches involve macroscopically different sentient beings, the technology for obtaining surrealistic trajectories is not simpler. Wigner needs to make the two branches identical in the spatial configuration space and then bring them back to their different macroscopic states. All these complicated manipulations will lead to histories in Bohmian mechanics different from those of the MWI, but these scenarios must involve memory erasure of the histories, so at no moment of time will Wigner, or any sentient being he makes the experiment with, have any evidence distinguishing one interpretation from the other.

The existence of sentient beings with spin brains does not change the conclusions. The BM supervenience postulate I does not allow such sentient beings. So, they will know, if they actually exist, that BM with postulate I is incorrect. They will have the option to accept BM with postulate II or MWI. If they will observe particles in experiment like in Fig. 2 (which, in fact, does not require supertechnology, the only gedanken part here is the existence of sentient being with spin brains) the histories of their experiences in the BM framework will be different from those of the MWI, but at no time will there be any evidence for a sentient being about the difference between BM with postulate II and the MWI.

I favor the MWI not because, but in spite of the plurality of worlds, so my motivation for BM is that it singles out one of the worlds of the MWI. I find the BM experience postulate I simple and natural and (in spite of featuring action at a distance) I find that the BM with postulate I is an attractive proposal. However, such a theory has to include at least implicitly a statement of nonexistence of sentient beings with spin brains or any other nonspatial degrees-of-freedom brains.

The alternative, BM with supervenience postulate II, seems less attractive. If experience supervenes on the wave function (the BM collapsed wave function), why not consider the MWI with essentially identical supervenience postulate? One might consider it as an advantage of BM that the BM collapsed wave function is better defined than the MWI branch wave function, but the MWI proponent can say that experience is not something that has to be precisely defined mathematically. Note also that even the BM collapsed wave function is rigorously defined only when the wave function is a superposition of spatially separated wave packets.

The research program of Detlef Dürr was not finished. We do not have a consensus about the interpretation of quantum mechanics. I believe that in my contribution I succeeded to shed some light on similarities and differences of BM and MWI and pointed to the direction which might lead to a progress: understanding better the connection between the formalism of quantum mechanics and our experience.

Acknowledgements This work has been supported in part by the Israel Science Foundation Grant No. 2064/19.

References

1. D. Dürr, S. Goldstein, N. Zanghi, Bohmian mechanics and the meaning of the wave function, in *Experimental Metaphysics: Quantum Mechanical Studies for Abner Shimony*, ed. by M.R.S. Cohen, J. Stachel, Boston Studies in the Philosophy of Science, Vol. 1 (Kluwer, 1997), pp. 25–38
2. D. Dürr, S. Teufel, Bohmian mechanics, in *Bohmian Mechanics* (Springer, 2009), pp. 145–171
3. L. Vaidman, Many-Worlds interpretation of Quantum mechanics, in *The Stanford Encyclopedia of Philosophy*. ed. by E.N. Zalta (Stanford University, Metaphysics Research Lab, 2002). https://urldefense.com/v3/__. https://plato.stanford.edu/entries/timethermo/–;!!NLFGqXoFfo8MMQ!vprs9m5l-WnJD262YutLOEjrctYKHQxmVxyqMmK3j9jx9VGPMCweSj3h5e1WNlCQJYochYf0E1PwDFXTbQsT4iw7TmlMeLmhMDzemjYoR2fFYoG1
4. D.M. Greenberger, The neutron interferometer as a device for illustrating the strange behavior of quantum systems. Rev. Mod. Phys. **55**(4), 875 (1983)
5. L. Vaidman, On schizophrenic experiences of the neutron or why we should believe in the many-worlds interpretation of quantum theory. Int. Stud. Philos. Sci. **12**, 245–261 (1998)
6. Y. Aharonov, A. Casher, Topological quantum effects for neutral particles. Phys. Rev. Lett. **53**, 319 (1984)
7. A. Cimmino, G. Opat, A. Klein, H. Kaiser, S. Werner, M. Arif, R. Clothier, Observation of the topological Aharonov-Casher phase shift by neutron interferometry. Phys. Rev. Lett. **63**, 380 (1989)
8. J.S. Bell, de Broglie–Bohm, delayed-choice, double-slit experiment, and density matrix, in *Quantum Mechanics, High Energy Physics and Accelerators: Selected Papers of John Bell* (with Commentary) (World Scientific, 1995), pp. 788–792
9. L. Vaidman, Time symmetry and the many-worlds interpretation, in *Many Worlds?: Everett, Quantum Theory, & Reality*, ed. by S. Saunders, J. Barrett, A. Kent, D. Wallace (Oxford University Press, 2010), pp. 582–596
10. K.J. McQueen, L. Vaidman, How the many worlds interpretation brings common sense to paradoxical quantum experiments, in *Scientific Challenges to Common Sense Philosophy*, ed. by R. Peels, J. de Ridder, R. van Woudenberg (Routledge, 2020), pp. 40–60
11. M. Carlesso, S. Donadi, L. Ferialdi, M. Paternostro, H. Ulbricht, A. Bassi, Present status and future challenges of non-interferometric tests of collapse models. Nat. Phys. **18**, 243–250 (2022)
12. E.P. Wigner, Remarks on the mind-body question, in *The Scientist Speculates*, ed. by I. Good (Heinemann, London, UK, 1961)
13. D. Frauchiger, R. Renner, Quantum theory cannot consistently describe the use of itself. Nat. Commun. **9**, 1–10 (2018)
14. M. Proietti, A. Pickston, F. Graffitti, P. Barrow, D. Kundys, C. Branciard, M. Ringbauer, A. Fedrizzi, Experimental test of local observer independence. Sci. Adv. **5**, eaaw9832 (2019)
15. B.G. Englert, M.O. Scully, G. Süssmann, H. Walther, Surrealistic Bohm Trajectories. Zeitschrift für Naturforschung A **47**, 1175–1186 (1992)
16. G. Naaman-Marom, N. Erez, L. Vaidman, Position measurements in the de Broglie-Bohm interpretation of quantum mechanics. Ann. Phys. **327**, 2522–2542 (2012)

A Tentative Completion of Quantum Mechanics

Jürg Fröhlich, Zhou Gang, and Alessandro Pizzo

Abstract We review a proposal of how to complete non-relativistic Quantum Mechanics to a physically meaningful, mathematically precise and logically coherent theory. This proposal has been dubbed *ETH-Approach* to Quantum Mechanics, "*E*" standing for "Events," "*T*" for "Trees," and "*H*" for "Histories." The *ETH*-Approach supplies the last one of three pillars Quantum Mechanics can be constructed upon in such a way that its foundations are solid and stable. Two of these pillars are well known. The third one has been proposed quite recently; it implies a general non-linear stochastic *law* for the time-evolution of states of *individual* physical systems.

1 What Is Missing in Text-Book Quantum Mechanics?

"It seems clear that the present quantum mechanics is not in its final form." (Paul Adrien Maurice Dirac)

In this paper we review some results on the foundations of quantum mechanics (*QM*). These results have appeared in papers already published. Various applications to concrete phenomena, such as radioactive decay of nuclei or fluorescence of atoms coupled to the quantized electromagnetic field, have been or will be presented elsewhere. Our main aim in this paper is to make a modest contribution towards removing some of the enormous jumble befuddling many people who attempt to work on the foundations of *QM*. We sadly miss the clear insights and useful comments our col-

Dedicated to the memory of Detlef Dürr.

J. Fröhlich (✉)
Institut für theoretische Physik, Eidgenössische Technische Hochschule, 8093 Zurich, Switzerland
e-mail: juerg@phys.ethz.ch

Z. Gang
Department of Mathematics and Statistics, Binghamton University, Binghamton, NY, USA
e-mail: gangzhou@binghamton.edu

A. Pizzo
Dipartimento di Matematica, Università di Roma "TorVergata", Rome, Italy
e-mail: pizzo@mat.uniroma2.it

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_12

league and friend *Detlef Dürr* would have contributed to the endeavor pursued in this paper. He thought about fundamental problems of quantum mechanics more deeply than most people and over many years [1].[1]

We start our review by explaining some of the shortcomings of text-book *QM* and the *"Copenhagen interpretation"*, with the purpose to highlight the need for a *completion* of the theory—as *Dirac* had anticipated. Text-book *QM* is a theory—alas, *incomplete*—of (*ensemble averages* over many identical) physical systems and of the time evolution of ensemble-averaged states based on the following two pillars:

(i) A physical system, S, is characterized by a list

$$\mathcal{O}_S = \{\widehat{X}_\iota = \widehat{X}_\iota^* \big| \iota \in \mathfrak{I}_S\}$$

of abstract bounded self-adjoint operators, where $\mathfrak{I}_S$ is a continuous index set. Every operator $\widehat{X} \in \mathcal{O}_S$ represents a (bounded function of a) *physical quantity* characteristic of S, such as the electromagnetic field in a bounded region of space-time, or the total momentum, energy or spin of all particles (e.g., atoms) in S localized in some bounded domain and interacting with the electromagnetic field. Different operators in $\mathcal{O}_S$ do in general *not* commute with one another. One assumes that if $\widehat{X} \in \mathcal{O}_S$ and F is a real-valued, bounded continuous function on $\mathbb{R}$ then $F(\widehat{X}) \in \mathcal{O}_S$, too. In general $\mathcal{O}_S$ does not have any additional structure (it is usually not a real linear space, let alone an algebra).
At every time t, there is a representation of $\mathcal{O}_S$ by bounded self-adjoint operators acting on a separable Hilbert space $\mathcal{H}$:

$$\mathcal{O}_S \ni \widehat{X} \mapsto X(t) = X(t)^* \in B(\mathcal{H}), \tag{1}$$

where $B(\mathcal{H})$ is the algebra of all bounded operators on $\mathcal{H}$.
Heisenberg picture time evolution: If S is an **isolated** system, i.e., one whose interactions with the rest of the Universe are negligibly weak, then the operators $X(t)$ and $X(t')$ representing a physical quantity $\widehat{X} \in \mathcal{O}_S$ at two times, t and t', are unitarily conjugated to one another. In an *autonomous* system,

$$X(t') = e^{i(t'-t)H_S/\hbar}\, X(t)\, e^{-i(t'-t)H_S/\hbar}, \tag{2}$$

where H_S is the Hamiltonian of S, which, for an autonomous system, is *time-independent*. For simplicity, we will henceforth assume that S is autonomous.

(ii) "States," ω, of S are assumed to be given by density matrices, Ω, i.e., by non-negative trace-class operators on $\mathcal{H}$ of trace one. The expectation at time t of an operator $\widehat{X} \in \mathcal{O}_S$ in the "state" ω of S is given by

$$\omega\big(X(t)\big) := \mathrm{Tr}\big(\Omega\, X(t)\big).$$

[1] We suspect, though, that our views of how to complete *QM* are likely to differ from what we think were his.

The state ω given by the density matrix Ω is *pure* iff Ω is a rank-1 projection $P = P^* = P^2$; otherwise it is a mixed state.

1.1 The Shortcomings of Text-Book Quantum Mechanics

In text-book *QM*, it is usually assumed, following *Schrödinger,* that, in the *Heisenberg picture*, "states" of an isolated physical system S are *independent* of time t, and, hence, that the Heisenberg picture is equivalent to the *Schrödinger picture*; namely

$$\omega(X(t)) = \mathrm{Tr}\big(\Omega X(t)\big) = \mathrm{Tr}\big(\Omega(t) X\big), \quad X := X(t_0),\ \Omega := \Omega(t_0),$$

where t_0 is an (arbitrarily chosen) initial time. In the Schrödinger picture, the Schrödinger (-von Neumann) equation

$$\dot{\Omega}(t) = -\frac{i}{\hbar}\big[H_S, \Omega(t)\big], \quad t \in \mathbb{R}. \tag{3}$$

describes the time evolution of *states* of S, while physical quantities of S are represented by *time-independent* bounded operators X on $\mathcal{H}$.

More generally, the time-dependence of "states" of a system S interacting with some environment is described by *linear, deterministic, trace-preserving, completely-positive maps,* $\big\{\Gamma(t,t')\big|t \geq t'\big\}$,

$$\Omega(t) = \Gamma(t,t')\big[\Omega(t')\big], \qquad \forall t \geq t', \tag{4}$$

where the operators $\Gamma(t,t')$ are defined on the linear space of trace-class operators on $\mathcal{H}$, and $\Gamma(t,t') = \Gamma(t,t'') \cdot \Gamma(t'',t')$, $t \geq t'' \geq t'$, with $\Gamma(t,t) = \mathbf{1}$; see [2, 3].

Thus, in text-book *QM*, the *time evolution of states* in the Schrödinger picture (see Eqs. (3), (4)) is *linear* and *deterministic*. Of course, this *cannot be the full story*! As already recognized by *Einstein* in 1916 in his paper on spontaneous and induced emission and absorption of light by atoms, which he described in probabilistic terms (introducing his A- and B-coefficients), *QM* is a *fundamentally probabilistic theory.* To anticipate an important fact about *QM*, we claim that the *linear deterministic evolution equations* (3) *and* (4) *only describe the evolution of* **ensemble averages** *of very many identical systems;* but that the time evolution of an ***individual*** *system* is *non-linear* and *stochastic*.

Thus, the fundamental problem arises to introduce an appropriate notion of *states of* **individual** *systems* and the correct general ***law*** determining the non-linear, stochastic evolution of these states. In other words, our task is to find the right *"ontology"* underlying *QM*.

According to the *Copenhagen interpretation* of *QM*, the deterministic evolution of the "state" of an individual system identical to S is *"interrupted"* at all times t

when an *"event"* happens, such as the emission or absorption of a photon by an atom, or when a measurement of the value of a physical quantity $\widehat{X} \in \mathcal{O}_S$ is carried out. In this latter case, the "state" of S is claimed to make a *"quantum jump"* to some "state" in the range of the spectral projection of $X(t)$ corresponding to the value of $\widehat{X}$ measured at time t, i.e., corresponding to the eigenvalue of $X(t)$ associated with the measured value of $\widehat{X}$. *QM* is claimed to predict the *probabilities* or *frequencies* of "quantum jumps" to eigenstates corresponding to *different possible values* of $\widehat{X}$ when measurements of the value of $\widehat{X}$ are repeated many times for identical, identically prepared systems. These frequencies are supposed to be given by the *Born Rule* applied to the state of S at the time when the measurement of $\widehat{X}$ begins.—This is, in words, the contents of *Lüders' measurement postulate* [4].

Critique of Text-Book QM

1. The notion of a *"measurement"* or *"observation"* appearing in the Copenhagen interpretation of *QM* is extremely vague. What is the difference between a period in the evolution of the state of a system *without* "measurement" and a period of evolution when a "measurement" is carried out?
   If the machinery used to measure the value of some physical quantity $\widehat{X}$ is *included* in what constitutes the *total* system S, now assumed to be *isolated*, one might expect—*erroneously*—that every event corresponding to a measurement of the value of $\widehat{X}$ could be viewed as the result of the *Schrödinger evolution of the state of the* **total** *system*. This would imply that *QM* is a *deterministic* theory—*which it obviously isn't*, as already noticed by Einstein; (see [5, 6] for more recent observations in this direction). So, what is going on?
2. If one takes Lüders' measurement postulate literally one is tempted to conclude that *QM* only makes useful predictions if it is known beforehand what measurements are planned by "observers" to be carried out, as well as what the times of their interventions are. One might then be misled to believe that the *free will* of "observers" plays a central role in *QM*.
3. The hypotheses implicit in the "Copenhagen interpretation" that one can freely choose the time when a measurement begins and that there are measurements that only take an arbitrarily small amount of time (which would actually imply that there are infinitely strong energy fluctuations associated with such measurements) strike us as totally absurd.
4. There are quantum phenomena, such as the radioactive decays of certain nuclei, as well as the precise decay times, or the fluorescence of atoms, that are intrinsically *random* and involve "quantum jumps." They are to be described by appropriate *stochastic* processes. But there are no "observers" involved to trigger them. So, where does the randomness of such phenomena originate from?

"Fake Cures" of Text-Book QM

- We think it is a mistake to imagine that the problems and paradoxes of text-book *QM* can be cured by some sort of *"interpretation"* of *QM*, such as "Relational QM," "QBism," "Consistent Histories," "Many-Worlds Interpretation" [7], "Information ontologies," etc.; see [8] and references given there.

As *David Mermin* put it: *New interpretations appear every year. None ever disappear.*

- We expect it to be equally unlikely that these problems and paradoxes can be eliminated by supplementing text-book *QM* with some *"ad-hoc mechanisms,"* such as ones based on decoherence [9, 10], spontaneous wave-function collapse [11] (which may remind one of electromagnetic or mechanical mechanisms used to explain *Lorentz contraction* before the advent of the theory of special relativity), or by attempting to reproduce the predictions of quantum mechanics by using cellular automata [12], etc.

Remark: *Bohmian mechanics*[2] *is* a logically coherent completion of (non-relativistic) quantum mechanics [1]. But it reminds one of "completing" classical electrodynamics by introducing a mechanical medium, the ether, thought to be the carrier of electromagnetic waves. The Bohmian particles are as "unobservable" as the ether, the most likely reason being that they are point-particles without any physical properties, such as electric charge or spin.—We do not expect that Dirac would have accepted it as a completion of *QM*.

In the following we attempt to convince the reader that the fundamental problem to solve in order to "complete" *QM* is to find a *universal quantum-mechanical* **law** that determines the non-linear stochastic time evolution of states of ***individual*** *systems,* with the properties that it correctly describes what is seen in experiments and that it reproduces the linear deterministic Schrödinger-von Neumann evolution of states averaged over an ensemble of very many identical isolated systems.

2 An Analogous Problem in the Theory of Diffusion Processes

An *analogous problem* in classical physics that may guide our thought process towards the right law is found in the theory of *diffusion* and *Brownian motion.* Consider a system consisting of a drop, $\mathfrak{E}$, of ink (e.g., eosin) in water. The *"state"* of $\mathfrak{E}$ at time t is given by its *density* ρ_t, which is a non-negative function on physical space $\mathbb{E}^3$. We normalize it such that $\int_{\mathbb{E}^3} d^3x\rho_t(x) = 1$. The time dependence of ρ_t is governed by the *diffusion equation*, viz. by a *deterministic linear law of evolution.*

$$\dot{\rho}_t(x) = D\,(\Delta\rho_t)(x), \qquad D : \text{diffusion constant.} \tag{5}$$

The well known solution of this equation is given by

$$\rho_t(x) = \int_{\mathbb{E}^3} d^3x'\,\Gamma_{t-t'}(x-x')\rho_{t'}(x'), \quad \Gamma_t(x) := (2\pi Dt)^{-\frac{3}{2}} e^{-\frac{|x|^2}{2Dt}},$$

[2] We have added a remark on Bohmian mechanics on request of one of the editors.

and the heat kernels Γ_t satisfy the *Chapman-Kolmogorov* equation. We will see that it is this property that distinguishes this classical model of a physical system from all quantum-mechanical models of physical systems.

According to the atomistic view of matter, $\mathfrak{E}$ really consists of very many eosin molecules, which, in an idealized description, can be viewed as point-like particles (far separated from one another, so that interactions among these particles can be neglected). The state of an individual particle is its *position* in physical space. The "state" of $\mathfrak{E}$, given by its density ρ, should then be interpreted as an *ensemble average* over the states of the particles constituting the ensemble $\mathfrak{E}$. An *individual system* in this ensemble consists of a *single particle*. According to *Einstein* and *Smoluchowski* (1905), the particles in $\mathfrak{E}$ exhibit *Brownian motion* arising from random collisions with lumps of water molecules. (From this they derived, for example, a formula for the diffusion constant, namely $D = \frac{k_B T}{6\pi\eta r}$.) We have learned from Einstein, Smoluchowski and *Wiener* that Brownian motion "unravels" the diffusion equation, with the following ***ontology***.

(i) At every time t, a particle is located in some point $x_\xi(t) \in \mathbb{E}^3$.

(ii) Its trajectory $\xi := \{x_\xi(t)\}_{t \geq t_0}$ is a random continuous curve—a Brownian path—in physical space $\mathbb{E}^3$; but the velocity of the particle is ill-defined at all times.

(iii) As shown by *Wiener*, there exists a ***probability measure***, $dW_{x_0}(\xi)$, on the space, Ξ, of particle trajectories, $\xi := \{x_\xi(t) \in \mathbb{E}^3 \mid t \geq t_0,\ x_\xi(t_0) = x_0\}$, starting from x_0 at time t_0; this measure is supported on trajectories ξ that are *Hölder continuous of index* $\frac{1}{2}$, etc.

(iv) An "event" at time t is the manifestation of the position, $x_\xi(t)$, of a particle. The trajectory ξ can thus be viewed as a *"history of events,"* a *random object*, and Ξ is the *"space of histories."*

Wiener measure $dW_{x_0}(\xi)$ allows us to predict probabilities of measurable sets of histories; for example,

$$\begin{aligned} &\mathrm{prob}\{\xi \in \Xi \,|\, x_\xi(t_i) \in \mathcal{O}_i,\ i = 1, 2, \ldots, n,\ t_0 < t_1 < \cdots < t_n\} \\ &\qquad = \int_\Xi dW_{x_0}(\xi) \prod_{i=1}^{n} \chi_{\{x_\xi(t_i) \in \mathcal{O}_i\}}(\xi)\,, \end{aligned} \tag{6}$$

where χ_Δ is the characteristic function of the set $\Delta \subset \Xi$.

The Chapman-Kolmogorov equation satisfied by the heat kernels implies that if regions $\mathcal{O}_i^{(\alpha)}$, $\alpha = 1, \ldots, N$, for some N, are chosen such that $\bigcup_{\alpha=1}^{N} \mathcal{O}_i^{(\alpha)} = \mathbb{E}^3$ then

$$\begin{aligned} &\sum_{\alpha=1}^{N} \mathrm{prob}\{\xi \,|\, x_\xi(t_1) \in \mathcal{O}_1, \ldots, x_\xi(t_i) \in \mathcal{O}_i^{(\alpha)}, \ldots, x_\xi(t_n) \in \mathcal{O}_n\} \\ &\qquad = \mathrm{prob}\{\xi \,|\, x_\xi(t_1) \in \mathcal{O}_1, \ldots, x_\xi(t_{i-1}) \in \mathcal{O}_{i-1}, x_\xi(t_{i+1}) \in \mathcal{O}_{i+1}, \ldots, x_\xi(t_n) \in \mathcal{O}_n\}\,. \end{aligned} \tag{7}$$

This property implies that if the position of a particle were measured at some intermediate time t_i and then a sum over all possible outcomes of this measurement were taken one would obtain the *same* predictions for the outcomes of measurements of the particle positions at times earlier than t_i and at times later than t_i as if no measurement had been made at time t_i. This means that the retrieval of information about the position of a particle does not affect its evolution. *QM* yields a totally different picture of reality (actually a more accurate one): A measurement *always* affects predictions on the evolution of a system *even* if a sum over all possible outcomes of the measurement were taken.

Using Wiener measure to take an average over the ensemble $\mathfrak{E}$ of very many identical particles, one recovers the *deterministic law* in Eq. (5) for the evolution of the "state" ρ_t,

$$\begin{aligned}\int_{\mathcal{O}} d^3x\, \rho_t(x) &= \int_{\mathcal{O}} d^3x \int d^3x_0\, \Gamma_{t-t_0}(x-x_0)\rho_{t_0}(x_0) \\ &= \int d^3x_0\, \rho_{t_0}(x_0) \int_{\Xi} dW_{x_0}(\xi)\, \chi_{\{x_\xi(t)\in\mathcal{O}\}}(\xi),\end{aligned} \tag{8}$$

for an arbitrary open subset $\mathcal{O} \subset \mathbb{E}^3$. We note that the Chapman-Kolmogorov equation for the heat kernels implies the Markov property for the Wiener measure dW_{x_0}, i.e., that a measurement of the particle position at some time t wipes out all memory of its trajectory at times earlier than t. In contrast, in quantum mechanics there usually are memory effects.

One might say that the Wiener measure *"unravels"* the diffusion equation (5). In the next section, we describe an "unraveling" of the linear, deterministic Schrödinger-von Neumann evolution of ensemble-averages of states of identical systems by a *non-linear, stochastic evolution* of states of ***individual*** systems inspired by the observations concerning diffusion and Brownian motion just sketched. This will yield a *completion of QM* and equip it with a plausible "ontology."

3 "Unraveling" the Schrödinger-Von Neumann Equation

The atoms or elementary particles themselves are not "real;" they form a world of potentialities or possibilities rather than one of things or facts. (Werner Heisenberg)

In this section we describe the ***third pillar*** to be added to the two conventional pillars of text-book quantum mechanics described in Sect. 1, in order to arrive at a complete theory. The ***ontology*** of our completion of *QM* will be found in *"random histories of events,"* defined appropriately; in analogy to histories of positions (Brownian paths) occupied by a point-like particle exhibiting Brownian motion. In *QM* one would like to equip the (non-commutative) space of histories of events with a *"quantum probability measure"*; in analogy with the Wiener measure of Brownian motion. Our task is to *find* this probability measure, or, more precisely, *to find*

an appropriate notion of states of physical systems in quantum mechanics and to describe their non-linear stochastic time evolution.

The *ETH-Approach to QM*, developed during the past decade (see [13–16]), accomplishes this task. Since this completion of *QM* may not be very widely known and appreciated, yet, we have to briefly sketch it again (in this paper for non-relativistic *QM*; but there also exists a *relativistic version* [17]). We follow the presentation in [18].[3]

3.1 Fundamental Ingredients of the ETH-Approach to Quantum Mechanics

In this section, we make use of the Heisenberg picture; and we consider *isolated* systems, i.e., systems, S, that have negligibly weak interactions with the rest of the Universe. For, only for isolated systems, the time-evolution of operators representing physical quantities of S has a conceptually clear description in the form of the *Heisenberg equations of motion*. The main ingredients of the *ETH*-Approach to the quantum theory of isolated systems are the following ones.

I. We define $\mathcal{E}_{\geq t}$ to be the (weakly closed) algebra[4] generated by all the operators

$$\left\{ X(t') \,\middle|\, t' \geq t,\ \widehat{X} \in \mathcal{O}_S \right\}, \tag{9}$$

Evidently,

$$\mathcal{E}_{\geq t'} \subseteq \mathcal{E}_{\geq t}, \quad \text{for} \quad t' > t\,.$$

For an *autonomous system* one has that

$$\mathcal{E}_{\geq t'} = e^{i(t'-t)H_S/\hbar}\, \mathcal{E}_{\geq t}\, e^{-i(t'-t)H_S/\hbar}, \quad \text{for} \quad t, t' \text{ in } \mathbb{R}\,. \tag{10}$$

II. An *isolated* **open** *physical system*, S, (i.e., an isolated system releasing *"events"*) is described by a "co-filtration," $\left\{\mathcal{E}_{\geq t} \,|\, t \in \mathbb{R}\right\}$, of von Neumann algebras (contained in the algebra, $B(\mathcal{H})$, of all bounded operators on $\mathcal{H}$) that satisfy the following

Principle of Diminishing Potentialities (*PDP*): *In an isolated* **open** *system S featuring events the following strict inclusions hold*

$$\mathcal{E}_{\geq t} \supsetneq \mathcal{E}_{\geq t'}\,, \quad \text{for arbitrary} \quad t' > t\,. \tag{11}$$

[3] It really does not make much sense to present this approach to *QM* in a new way each time it has to be recalled, because people have chosen not to take notice of it.

[4] i.e., a von Neumann algebra.

People tend to be perplexed when hearing about *PDP*, because they find it hard to believe that *PDP* is compatible with the unitary Heisenberg dynamics of operators described in Eqs. (2) and (10). However, in a relativistic local quantum (field) theory over an even-dimensional, flat space-time containing a massless "radiation field," such as quantum electrodynamics, and for an appropriate choice of the algebras $\mathcal{E}_{\geq t}$, $t \in \mathbb{R}$, *PDP* can be shown to be a consequence of *Huygens' Principle*, as formulated and proven in [21] in the context of algebraic quantum field theory. In [16], some concrete models, including models arising when the velocity of light tends to ∞, are shown to satisfy *PDP*.

III. The notion of ***"events"***[5] plays a central role in the *ETH*-Approach: A ***potential event*** in S setting in at time t is described by a partition of unity,

$$\mathfrak{P} := \{\pi_\xi \,|\, \xi \in \mathfrak{X}\} \subset \mathcal{E}_{\geq t}, \tag{12}$$

by orthogonal, mutually disjoint projections, π_ξ, with the properties that

$$\pi_\xi = \pi_\xi^*, \quad \pi_\xi \cdot \pi_\eta = \delta_{\xi\eta}\pi_\xi, \quad \forall\, \xi, \eta \in \mathfrak{X}, \quad \sum_{\xi\in\mathfrak{X}} \pi_\xi = \mathbf{1}\,, \tag{13}$$

where $\mathfrak{X}$ is a finite or countably infinite set of labels called the *spectrum* of the potential event $\mathfrak{P}$ and denoted by $\mathfrak{X} = \mathrm{spec}(\mathfrak{P})$.

IV. A *state* of an isolated system S at time t is given by a *quantum probability measure* on the lattice of orthogonal projections in $\mathcal{E}_{\geq t}$, i.e., by a functional, ω_t, with the properties that

(i) ω_t assigns to every orthogonal projection $\pi \in \mathcal{E}_{\geq t}$ a non-negative number $\omega_t(\pi) \in [0, 1]$, with $\omega_t(0) = 0$, and $\omega_t(\mathbf{1}) = 1$; and

(ii) ω_t is *additive*, i.e.,

$$\sum_{\pi\in\mathfrak{P}} \omega_t(\pi) = 1, \quad \forall \text{ potential events } \mathfrak{P} \subset \mathcal{E}_{\geq t}\,. \tag{14}$$

A generalization of *Gleason's theorem* due to *Maeda* [20] implies that states, ω_t, of S at time t, as defined above, are *positive, normal, normalized linear functionals* on $\mathcal{E}_{\geq t}$, i.e., *states* on $\mathcal{E}_{\geq t}$ in the usual sense of this notion employed in the mathematical literature. (Ignoring some mathematical subtleties) we henceforth identify ω_t with a density matrix on $\mathcal{H}$ denoted by Ω_t.

[5] in the sense the late *Rudolf Haag* used this terminology; see [19].

3.2 Consequences of the Principle of Diminishing Potentialities

The Principle of Diminishing Potentialities, when combined with the phenomenon of *entanglement,* implies that even if the state ω_t of S at time t were a "pure" state on the algebra $\mathcal{E}_{\geq t}$ its restriction to the algebra $\mathcal{E}_{\geq t'}$ must be expected to be *"mixed"* if $t' > t$. This observation opens the possibility to introduce the notion of "events *actualizing* at some time."

In accordance with the *"Copenhagen interpretation"* of *QM*, one might expect that a *potential event* $\mathfrak{P} = \{\pi_\xi \,|\, \xi \in \mathrm{spec}(\mathfrak{P})\} \subset \mathcal{E}_{\geq t}$, becomes actual (manifest) at some time $\geq t$ iff

$$\mathrm{tr}(\Omega_t A) = \sum_{\xi \in \mathfrak{X}} \mathrm{tr}(\pi_\xi \, \Omega_t \, \pi_\xi \, A), \quad \forall A \in \mathcal{E}_{\geq t} \,. \tag{15}$$

where Ω_t is the density matrix representing the state ω_t of S at time t. Notice that off-diagonal elements do *not* appear on the right side of (15), which thus describes an *incoherent* superposition of states in the images of disjoint orthogonal projections, i.e., a *"mixture."*

This expectation is made precise as follows. Given a state ω_t on $\mathcal{E}_{\geq t}$, we define $\mathcal{C}(\omega_t)$ to be the subalgebra of $\mathcal{E}_{\geq t}$ generated by *all* projections belonging to *all* potential events $\mathfrak{P} \subset \mathcal{E}_{\geq t}$ for which Eq. (15) holds. Further, $\mathfrak{P}(\omega_t)$ is the *finest potential event* contained in $\mathcal{C}(\omega_t)$ with the property that *all* its elements commute with *all* operators in $\mathcal{C}(\omega_t)$.[6] We then say that the potential event $\mathfrak{P}(\omega_t)$ *actualizes* at some time $\geq t$ iff $\mathfrak{P}(\omega_t)$ contains *at least* two non-zero orthogonal projections, $\pi^{(1)}, \pi^{(2)}$, which are disjoint, i.e., $\pi^{(1)} \cdot \pi^{(2)} = 0$, and have non-vanishing Born probabilities, i.e.,

$$0 < \omega_t(\pi^{(i)}) = \mathrm{tr}\big(\Omega_t \, \pi^{(i)}\big) < 1 \,, \quad \text{for } i = 1, 2 \,.$$

Equation (15) then holds true for $\mathfrak{P} = \mathfrak{P}(\omega_t)$, and the sum on the right side of (15) contains at least two distinct non-vanishing terms.

3.3 The State-Reduction Postulate and the Stochastic Evolution of States

The ***law*** describing the non-linear stochastic time evolution of states of an individual isolated open system S unraveling the linear deterministic evolution of ensemble averages of states is derived from a ***state-reduction postulate*** described next. This postulate makes precise mathematical sense as long as *time* is *discrete*.

[6] In more technical jargon, $\mathfrak{P}(\omega_t)$ generates the *center* of the centralizer $\mathcal{C}(\omega_t)$ of ω_t.

Let ω_t be the state of S at time t. Let dt denote a time step; (dt is strictly positive if time is discrete; otherwise one attempts to let dt tend to 0 at the end of the following constructions). We define a state $\overline{\omega}_{t+dt}$ on the algebra $\mathcal{E}_{\geq t+dt}$ ($\subsetneq \mathcal{E}_{\geq t}$) by restriction of ω_t to the algebra $\mathcal{E}_{\geq t+dt}$,

$$\overline{\omega}_{t+dt} := \omega_t\big|_{\mathcal{E}_{\geq t+dt}} .$$

As a manifestation of *PDP* and *entanglement,* the algebra $\mathcal{C}(\overline{\omega}_{t+dt})$ can be expected to be non-trivial (i.e., $\neq \mathbb{C} \cdot \mathbf{1}$) in general. This does, of course, *not* imply that the potential event $\mathfrak{P}(\overline{\omega}_{t+dt})$ actualizing at some time $\geq t + dt$ is non-trivial, too, i.e., $\neq \mathbf{1}$. But it is plausible that it will in general be non-trivial. (This is shown to be the case in a family of models studied in [16].)

Axiom CP: *Let*

$$\mathfrak{P}(\overline{\omega}_{t+dt}) = \big\{\pi_\xi \,|\, \xi \in \mathrm{spec}\big(\mathfrak{P}(\overline{\omega}_{t+dt})\big)\big\}$$

be the potential event actualizing at some time $\geq t + dt$, *given the state* $\overline{\omega}_{t+dt}$ *on* $\mathcal{E}_{\geq t+dt}$. *Then 'Nature' replaces the state* $\overline{\omega}_{t+dt}$ *on* $\mathcal{E}_{\geq t+dt}$ by a state $\omega_{t+dt} \equiv \omega_{t+dt,\pi}$ *represented by the density matrix*

$$\Omega_{t+dt,\pi} := \mathrm{tr}(\overline{\Omega}_{t+dt}\,\pi)^{-1} \cdot \pi\ \overline{\Omega}_{t+dt}\ \pi\,, \quad \text{for some} \quad \pi \in \mathfrak{P}(\overline{\omega}_{t+dt}), \tag{16}$$

with $\mathrm{tr}(\overline{\Omega}_{t+dt}\,\pi) \neq 0$. *The probability,* $\mathrm{prob}_{t+dt}(\pi)$, *for the state* $\omega_{t+dt,\pi}$, $\pi \in \mathfrak{P}(\overline{\omega}_{t+dt})$, *to be selected by 'Nature' as the state of* S *at time* $t + dt$ *is given by Born's Rule*

$$\mathrm{prob}_{t+dt}(\pi) = \mathrm{tr}(\overline{\Omega}_{t+dt}\ \pi)\,. \qquad \square \tag{17}$$

The projection $\pi(t + dt) := \pi \in \mathfrak{P}(\overline{\omega}_{t+dt})$ appearing in (16) and (17) is called ***actual event***, or *"actuality,"* at time $t + dt$.

The analogue of the initial position, x_0, of a Brownian path at time t_0 is the initial state ω_0 on $\mathcal{E}_{\geq t_0}$; the analogue of the Brownian trajectory $\xi = \big\{x_\xi(t)\,\big|\,t \geq t_0\big\}$ is given by a ***history***, $\mathfrak{h} := \big\{\pi(t_0 + dt), \pi(t_0 + 2dt), \ldots, \pi(t)\big\}$, of ***actual events*** originating from the initial state ω_0 of S at time t_0. With a history $\mathfrak{h}$ we associate a *"history operator"* defined by

$$H_\mathfrak{h}(t_0, t) := \prod_{t' \in \mathbb{Z}_{dt},\, t_0 < t' \leq t} \pi(t')\,.$$

In quantum mechanics, the role of the Wiener measure, dW_{x_0}, of Brownian motion is played by the probabilities

$$\begin{aligned}\mathrm{prob}_{\omega_0}\big[\mathfrak{h} = &\big\{\pi(t_0 + dt), \pi(t_0 + 2dt), \ldots, \pi(t)\big\}\big] := \\ &= \omega_0\big(H_\mathfrak{h}(t_0, t) \cdot H_\mathfrak{h}(t_0, t)^*\big) = \mathrm{tr}\big[H_\mathfrak{h}(t_0, t)^* \cdot \Omega_0 \cdot H_\mathfrak{h}(t_0, t)\big]\end{aligned} \tag{18}$$

of histories of events, where Ω_o is the density matrix representing the initial state ω_0 on the algebra $\mathcal{E}_{\geq t_0}$.

It follows from our discussion that the time-evolution of the *state* of an ***individual*** physical system S is described by a *stochastic branching process*, called *"quantum Poisson process"*, whose "state space" is referred to as the *non-commutative spectrum*, $\mathfrak{Z}_S$, of S and is defined as follows. By Eq. (10), all the algebras $\mathcal{E}_{\geq t}$ are isomorphic to one specific (universal) von Neumann algebra, which we denote by $\mathcal{M}$. The non-commutative spectrum, $\mathfrak{Z}_S$, of S is defined by

$$\mathfrak{Z}_S := \bigcup_{\omega} \Big(\omega \,, \mathfrak{P}(\omega) \Big) \,, \tag{19}$$

where the union over ω is a disjoint union, and ω ranges over *all* states on $\mathcal{M}$ of physical interest. ("States of physical interest" are normal states on $\mathcal{M}$ a concrete system can actually be prepared in.) The branching rules of a quantum Poisson process on $\mathfrak{Z}_S$ are uniquely determined by **Axiom CP**.
Comments.

- One may expect—and this can be verified in concrete models (see [16] for further details)—that, most of the time, the actual event, $\pi \in \mathfrak{P}(\overline{\omega}_{t+dt})$, which, according to the Born Rule, has the largest probability to happen, and hence is most likely to be chosen by 'Nature' (see (16)), has the property that

$$\omega_{t+dt} \equiv \omega_{t+dt,\pi} \approx \overline{\omega}_{t+dt} = \omega_t\big|_{\mathcal{E}_{\geq t+dt}} \,. \tag{20}$$

  This would imply that, most of the time, the evolution of the state is close to being trivial (as assumed in text-book *QM* in the absence of "measurements"). But, every once in a while, the state of the system makes a *"quantum jump"* corresponding to an actual event π in (16) that is very unlikely to materialize. Such "quantum jumps" happen for purely *entropic* reasons at *random times*.
- One may check that the non-linear stochastic evolution of states outlined above has the desirable feature that it reproduces the usual Schrödinger-von Neumann evolution when an ensemble-average over all possible histories of very many identical systems is taken.
- Our construction of the non-linear stochastic time evolution of individual systems is meaningful, mathematically, as long as $dt > 0$; but, for the time being, the limiting theory, as $dt \searrow 0$, is only understood precisely in examples.

4 Concluding Remarks

1. The *ETH-Approach* to Quantum Mechanics represents a *completion of QM* that provides a logically coherent description of the *stochastic time evolution of states* of ***individual*** *systems* in *QM* (unraveling Schrödinger-von Neumann evolution)

and of *events* and their recordings (see [14, 15, 17]). It has resemblences (albeit rather vague ones) with Everett's "Many Worlds" formalism [7] and spontaneous collapse models à la "GRW" [11]. But it supersedes these ad-hoc formalisms by a precise and more natural one. And it describes only **One World**: hopefully ours! Of course, it will have to stand the test of experiments.

2. As *Pauli* has put it: *If speculative ideas cannot be tested, they're not science; they don't even rise to the level of being wrong.* – We thus should ask whether the *Principle of Diminishing Potentialities* (*PDP*), which is a corner stone of the *ETH*-Approach to *QM*, is more than a speculative idea and whether it can be tested. It is clear that this principle can only be established in quantum theories of systems with infinitely many degrees of freedom. It has the status of a *theorem* in local relativistic quantum theory with massless particles on even-dimensional space-times; e.g., in 4D quantum electrodynamics (QED) [21], and in simple models of QED regularized at high energies by discretizing time [16]. It also holds in models emerging in the limit of the velocity of light tending to ∞ (see [16]). However, in this limit, the Hamiltonian is not bounded from below; i.e., the spectrum condition ($\nexists$ negative-energy states) is violated.
   We thus have strong reasons to expect that a completion of *QM* satisfying the spectrum condition and solving the *"measurement problem"* will succeed *only* in the guise of *local relativistic quantum theory* on even-dimensional space-times featuring massless bosons, photons and gravitons; (so that "Huygens Principle" [21] holds). The *ETH*-Approach to quantum mechanics sketched above *does have* an extension to *local relativistic quantum theory*; (see [17] for a preliminary account).
   Applications of the *ETH*-Approach to concrete models (e.g., models of the fluorescence of atoms, etc.) have been discussed [16] and will be presented elsewhere.
3. A quantum-mechanical analogue of the magic formula (6) for Brownian motion (see Sect. 1.2) has been proposed by *Lüders, Schwinger* and *Wigner* (see [22]). However, when applied to time-ordered series of measurements, their formula fails to satisfy an analogue of Eq. (7), because the non-commutativity of different potential events actualizing at different times leads to *interference effects*. Not surprisingly, this has been noticed by many people, who thought of various ways to rescue their formula. One formalism seemingly enabling one to come up with meaningful predictions that has become quite popular is known under the name of *"consistent histories"* [9, 10]. However, in our modest opinion, this formalism does **not** represent an acceptable completion of *QM*, because it talks about unpredictable and instantaneous interventions by "observers," a feature that extinguishes much of the predictive power of *QM*.
4. The *ETH*-Approach to *QM*, in particular *PDP*, introduces a fundamental *"arrow of time,"* i.e., a distinction between past and future into the theory: The past consists of *facts*, namely histories of *"actualities",* while the future consists of *"potentialities"* (much in the sense in which *Aristotle* originally conceived these notions).

Acknowledgements One of us (J. F.) thanks his collaborators, in particular *Baptiste Schubnel,* in earlier work on related problems for the pleasure of cooperation, and *Carlo Albert, Shelly Goldstein* and *Erhard Seiler* for their very encouraging interest in our efforts.

References

1. D. Dürr, S. Teufel, *Bohmian Mechanics* (Springer, Berlin and Heidelberg, 2009)
2. K. Kraus, *States, Effects, and Operations*, Lecture Notes in Physics, vol. 190. (Springer, Berlin, 1983)
3. V. Gorini, A. Kossakowski, E.C.G. Sudarshan, Completely positive semigroups of N-level systems. J. Math. Phys. **17**(5), 821 (1976) G. Lindblad, On the generators of quantum dynamical semigroups. Commun. Math. Phys. **48**, 119–130 (1976)
4. G. Lüders, Über die Zustandsänderung durch den Messprozess. Ann. Phys. (Leipzig) **443**(5–8), 322–328 (1950)
5. E.P. Wigner, Remarks on the mind-body question, in *Symmetries and Reflections* (Indiana University Press, Bloomington, 1967), pp. 171–184
6. J. Faupin, J. Fröhlich, B. Schubnel, On the probabilistic nature of quantum mechanics and the notion of closed systems. Ann. H. Poincaré **17**, 689–731 (2016)
7. H.D. Everett III., Relative state formulation of quantum mechanics. Rev. Mod. Phys. **29**, 454–462 (1957)
8. In: *Interpretations of quantum mechanics*, Wikipedia; see https://en.wikipedia.org/wiki/Interpretationshyphenofhyphenquantumhyphenmechanics
9. R.B. Griffiths, Consistent histories and the interpretation of quantum mechanics. J. Stat. Phys. **36**(1), 219–272 (1984)
10. M. Gell-Mann, J.B. Hartle, Classical equations for quantum systems. Phys. Rev. D **47**(8), 3345–3382 (1993)
11. G.C. Ghirardi, A. Rimini, T. Weber, Unified dynamics for microscopic and macroscopic systems. Phys. Rev. D **34**, 470–491 (1986)
12. G. 'tHooft, *The Cellular Automaton Interpretation of Quantum Mechanics*, Fundamental Theories of Physics, ed. by H. van Beijeren et al., vol. 185 (Springer, Cham, Heidelberg, New York, 2016)
13. J. Fröhlich, B. Schubnel, Quantum probability theory and the foundations of quantum mechanics, in *The Message of Quantum Science*, ed. by Ph. Blanchard, J. Fröhlich (Springer, Berlin, 2015)
14. Ph. Blanchard, J. Fröhlich, B. Schubnel, A Garden of forking paths - the quantum mechanics of histories of events. Nucl. Phys. B **912**, 463–484 (2016)
15. J. Fröhlich, A brief review of the "ETH - approach to quantum mechanics," in *Frontiers in Analysis and Probability*, ed. by N. Anantharaman, A. Nikeghbali, M. Rassias (Springer, Cham, 2020)
16. J. Fröhlich, A. Pizzo, The time-evolution of states in quantum mechanics according to the ETH-approach. Commun. Math. Phys. **389**, 1673–1715 (2022)
17. J. Fröhlich, Relativistic quantum theory, in *Do Wave Functions Jump? Perspectives of the Work of GianCarlo Ghirardi*, ed. by V. Allori, A. Bassi, D. Dürr, N. Zanghi, Fundamental Theories of Physics (Springer, Cham, 2020). and paper in preparation
18. J. Fröhlich, Z. Gang, On the Evolution of States in a Quantum-Mechanical Model of Experiments. arXiv:2212.02599. Submitted for publication in Ann. H. Poincaré
19. R. Haag, Fundamental irreversibility and the concept of events. Commun. Math. Phys. **132**, 245–251 (1990); See also: R. Haag, On quantum theory. Int. J. Quant. Inf. **17**, 1950037 (2019)
20. A.M. Gleason, Measures on the closed subspaces of a Hilbert space. J. Math. Mech. **6**, 885–893 (1957); S. Maeda Probability measures on projections in von Neumann algebras. Rev. Math. Phys. **1**, 235–290 (1989)

21. D. Buchholz, Collision theory for massless particles. Commun. Math. Phys. **52**, 147–173 (1977); see also: D. Buchholz, J. Roberts, New light on infrared problems: sectors, statistics, symmetries and spectrum. Commun. Math. Phys. **330**, 935–972 (2014)
22. G. Lüders, see [4]; J. Schwinger, The algebra of microscopic measurement. Proc. Natl. Acad. Sci. (USA) **45**(10), 1542–1553 (1959); E.P. Wigner, in *The Collected Works of Eugene Paul Wigner* (Springer, Berlin, 1993)

Relativistic Quantum Theory

Space–Time–Matter

Some Notes on the Localization Problem in Relativistic Quantum Theory

Christian Beck

Abstract This work aims to shed some light on the meaning of the positive energy assumption in relativistic quantum theory and its relation to questions of localization of quantum systems. It is shown that the positive energy property of solutions of relativistic wave equations (such as the Dirac equation) is very fragile with respect to state transformations beyond free time evolution. Paying attention to the connection between negative energy Dirac wave functions and pair creation processes in second quantization, this analysis leads to a better understanding of a class of problems known as the localization problem of relativistic quantum theory (associated for instance with famous results of Newton & Wigner, Reeh & Schlieder, Hegerfeldt or Malament). Finally, this analysis is reflected from the perspective of a Bohmian quantum field theory.

1 A Basic Theorem

We start with a result that follows from complex analysis of several complex variables:

Theorem 1 *Let λ be a complex measure*[1] *on $\mathbb{R}^4$ with support in the closure of the forward light cone $\overline{V}_+ = \{p \in \mathbb{R}^4 \mid p_\mu p^\mu = p_0^2 - \boldsymbol{p}^2 \geq 0,\ p_0 \geq 0\}$ of the origin. Consider the function $f : \mathbb{R}^4 \to \mathbb{C}$ given by*

$$f(x) = \int e^{ipx}\, d^4\lambda(p) \tag{1}$$

[1] A complex measure can be always understood as a collection of four ordinary measures, it has a real and an imaginary part which are signed measures. These in turn can each be decomposed into two normal finite measures using a Hahn-Jordan decomposition. The important thing here about a complex measure is that it is always finite (e.g. a finite ordinary measure is also a complex measure). That λ has support in $\overline{V}_+$ means that all integrals with respect to λ over subsets of $\mathbb{R}^4$ disjoint from $\overline{V}_+$ vanish, in particular $\int_{\mathbb{R}^4} d^4\lambda(p) = \int_{\overline{V}_+} d^4\lambda(p) \in \mathbb{C}$.

C. Beck (✉)
Department of Humanities and Arts, Technion-Israel Institute of Technology, Haifa 32000, Israel
e-mail: Christian_Beck@posteo.de

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_13

where $px = p_\mu x^\mu$ *is the Minkowski scalar product. If* f *vanishes on an open connected subset* $\mathcal{O} \subset \mathbb{R}^4$ *it follows that* $f \equiv 0$ *on all of* $\mathbb{R}^4$.

The proof can be found in [2] (see Corollary 4.6). It is based on the fact that f can be continued analytically to a region of $\mathbb{C}^4$ which has $\mathbb{R}^4$ as a part of its boundary. Thus, f in (1) can be regarded as the boundary value of an analytic function and the conclusion of Theorem 1 then follows with the help of generalizations of the Schwartz reflection principle and the identity theorem to functions of several complex variables.

Theorem 1 has a number of strong physical consequences for (relativistic) quantum mechanics, all of which are related in some sense and some of which will be discussed in this work. Physically x corresponds to a spacetime vector and p to the energy momentum four-vector. The condition $p \in \overline{V}_+$ is the so-called *spectrum condition*. It is adapted to relativistic considerations and says that the relativistic energy p_0 is positive in every Lorentz frame.

2 Implications for Wave Functions

In this section we think of f as (a component of) a relativistic wave function of positive energy, e.g., a positive energy solution of the free Klein-Gordon equation or a spinor component of a positive energy solution of the free Dirac equation. Such functions can be written in the form (see, e.g., [33, 34])

$$\psi(\boldsymbol{x}, t) = \int e^{i(p_0 t - \boldsymbol{p}\cdot\boldsymbol{x})}\, \delta\left(p_0^2 - (\boldsymbol{p}^2 + m^2)\right)\, \theta(p_0)\, \hat{\psi}(p)\, d^4 p \tag{2}$$

which is of the form (1) with the complex measure[2]

$$d^4\lambda(p) \;=\; \delta(p^2 - m^2)\, \theta(p_0)\, \hat{\psi}(p)\, d^4 p \tag{3}$$

(θ denotes the Heaviside step function). We shall switch in the following between the notations $\psi(\boldsymbol{x}, t) = \psi_t(\boldsymbol{x}) = \psi(x)$ (with $x \in \mathbb{R}^4$), depending on which is most appropriate for the current purpose.

2.1 *Causally Propagating Positive Energy Wave Functions Cannot Vanish in a Region*

Now suppose $\psi_t(\boldsymbol{x})$ vanishes at some time $t = t_0$ on an open, connected spatial subset (region) $\Delta \subset \mathbb{R}^3$, i.e., $\psi_{t_0}(\boldsymbol{x}) = 0$ for all $x \in \Delta$. If ψ propagates causally (which is

[2] To be precise, $\int \delta(p^2 - m^2)\, \theta(p_0)\, \hat{\psi}(p)\, dp_0$ must be in $L^1(\mathbb{R}^3, d^3 p)$ to define a complex measure.

the case for solutions of relativistic wave equations because of their hyperbolic form [24, 37]), for later (and earlier) t the support of ψ_t can spread at most with the speed of light as t evolves. Therefore, there must be an $\varepsilon > 0$ such that for each $s \in (-\varepsilon, \varepsilon)$, $\psi_{t_0+s}(\boldsymbol{x}) = 0$ on an open spatial set $\Delta_s \subset \mathbb{R}^3$. This way $\psi(x) = 0$ for all $x = (t, \boldsymbol{x})$ in an open subset $\mathcal{O} \subset \mathbb{R}^4$ (see Fig. 1, where the sets Δ_s are not depicted, but the dashed line at $t_0 + s$ indicates the complement of Δ_s). Theorem 1 thus entails that $\psi_t(\boldsymbol{x}) = 0$ for all $\boldsymbol{x}$ and t which contradicts the assumption that ψ is a wave function.

The conclusion is that a causally propagating wave function of the form (2) has at each time the property

$$\operatorname{supp}(\psi) = \mathbb{R}^3 \tag{4}$$

This implies in particular the often quoted statement that *relativistic wave functions of positive energy cannot have compact support* but have always *infinite tails*.

It is interesting to note that an analogous statement can also be made for non-relativistic Schrödinger wave functions. Theorem 1 has been formulated in a way that is well suited for relativistic analysis. However, a result analogous to Theorem 1 can be proved [4], which instead of the spectrum condition (that the relativistic energy is positive in each Lorentz frame) only needs the condition that the Hamiltonian (the generator of time translations), whose eigenvalues correspond to the allowed values of p_0 in (2), is bounded from below. This is true in particular for the Schrödinger Hamiltonian of non-relativistic quantum mechanics. Since Schrödinger wave functions can be zero on open connected sets (as Dirac wave functions can if contributions from negative energy eigenstates are allowed), this shows that Schrödinger wave functions spread instantaneously (with infinite propagation velocity) under the free time evolution.

In a sense, these interrelations can be seen as the core of *Hegerfeldt's theorem*[3] [20–22] .

[3] Hegerfeldt's theorem proves, roughly said, instantaneous spreading of any 'localization probabilities' in quantum theory (with Hilbert space $\mathcal{H}$) with positive energy, if there is a bounded spatial region $\Delta \subset \mathbb{R}^3$ and $\psi \in \mathcal{H}$, such that $\mathbb{P}^\psi(\Delta) = 1$ (perfect localization). The probabilities are assumed to be given by the quantum formalism, i.e., for any spatial region $\Delta \subset \mathbb{R}^3$ there is a positive bounded operator D_Δ, such that $\mathbb{P}^\psi(\Delta) = \langle \psi |\, D_\Delta\, \psi \rangle$. The connection to our discussion above becomes apparent when we choose the PVM of the standard position operator (indicator functions in position representation) $D_\Delta = \chi_\Delta$ and observe that $1 = \|\psi\|^2 = \int_{\mathbb{R}^3} |\psi(\boldsymbol{x})|^2 d^3x = \int_\Delta |\psi(\boldsymbol{x})|^2 d^3x = \langle \psi |\, \chi_\Delta\, \psi \rangle$ implies that $\psi(\boldsymbol{x}) = 0$ almost everywhere in the complement of Δ. For this choice, Hegerfeldt's theorem thus states that a compactly supported positive energy wave function cannot propagate causally. Hegerfeldt's theorem can be proven by application of Theorem 1 with the choice $f(x) = \langle \psi |\, U(x)\, \psi \rangle$, where $U(x)$ is a unitary representation of space-time translations (see [2]).

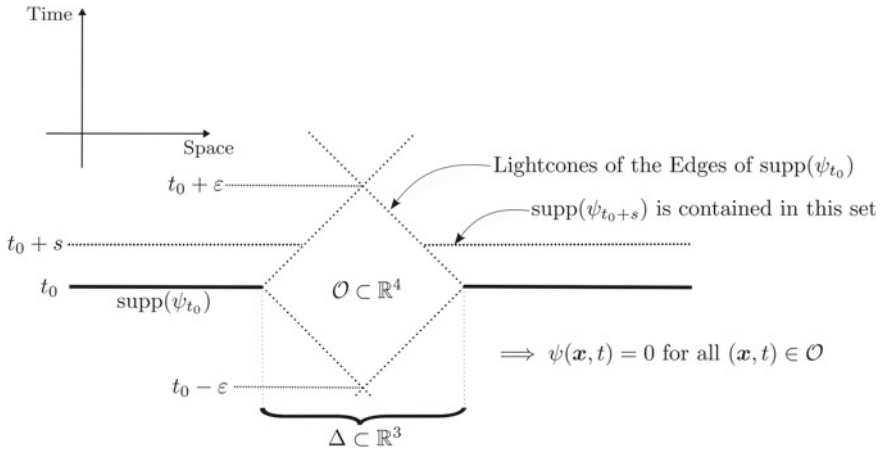


Fig. 1 Causal propagation of the support of a relativistic wave function: if the support of ψ_t can propagate at most at the speed of light and $\psi_{t_0}(\boldsymbol{x}) = 0$ for all $\boldsymbol{x}$ in a connected open spatial set $\Delta \subset \mathbb{R}^3$, then $\psi_t(\boldsymbol{x})$ also vanishes for $(t, \boldsymbol{x})$ in a connected open space-time set $\mathcal{O} \subset \mathbb{R}^4$ (the interior of the diamond in the middle). The diagonal dotted lines depict (the essential parts of) the forward and backward light cones of the edges of the support of ψ_{t_0}

2.2 *A Causally Propagating Positive Energy Wave Function is Completely Determined by Its Values in Any Region*

Consider two wave functions ψ and ψ' of the form (2) and suppose that at a certain time t_0 there exists an (arbitrarily small) open connected spatial set $\Delta \subset \mathbb{R}^3$ on which the wave functions coincide:

$$\psi(\boldsymbol{x}, t_0) = \psi'(\boldsymbol{x}, t_0) \quad \text{for all } x \in \Delta \tag{5}$$

Together with ψ and ψ', the wave function

$$\Phi(\boldsymbol{x}, t) := \psi(\boldsymbol{x}, t_0) - \psi'(\boldsymbol{x}, t_0) \tag{6}$$

is also of the form (2). However, at time t_0, Φ obviously vanishes on Δ so that Theorem 1 together with our discussion in Sect. 2.1 proves that Φ (and thereby either ψ or ψ' or both) cannot propagate causally. The other way around, this entails that two positive energy solutions of relativistic wave equations–which always propagate causally–cannot coincide on any open connected spatial set.

2.3 *Transformations of Causally Propagating Positive Energy Wave Functions are Very Special (e.g. Necessarily Nonlocal)*

Consider a solution ψ of a relativistic wave equation which is exposed to a local potential ϕ for some time, resulting in a transformed state $U_\phi\,\psi$, where U_ϕ is the unitary time evolution with potential ϕ. Let U be the free time evolution without potential corresponding to the same period of time as U_ϕ. Since the local potential can only locally perturb the wave function and since solutions of relativistic wave equations propagate causally, $U\,\psi$ and $U_\phi\,\psi$ can also differ only locally, i.e. $\Phi = U\,\psi - U_\phi\,\psi$ has compact support and thus cannot be a positive energy solution. Consequently, if $U\,\psi$ is a positive energy solution (which is the case if ψ has positive energy since the free time evolution leaves the positive energy property invariant), $U_\phi\,\psi$ must have contributions from the negative energy spectrum.

We can also formally set $U = 1$ to see that any local transformation of a relativistic positive energy state destroys its positive energy property. In other words, if we wiggle such a wave function just a little bit in the neighborhood of some point, immediately the whole function must change in a non-trivial way, if the resulting wave function shall continue to have positive energy. So a relativistic time evolution cannot be of this kind, it must either act on the wave function on the whole space (including the tails) in a very special way (as free time evolution does) or violate the positive energy property. Note the emphasis on 'very special': since the whole function is completely determined by its values in an arbitrarily small neighborhood, its global transformation must be perfectly concerted across all regions if it shall preserve the positive energy property!

2.4 *Discussion*

Tails: The infinite tails of positive energy wave functions do not contradict the fact that positive energy wave functions usually are, for all practical purposes, perfectly localized in bounded spatial regions. Various localization schemes for positive energy wave functions have been developed (most famously that of Newton and Wigner[4] [29], but see also e.g. [5, 6, 30] and the discussion of these schemes in [2]) which

[4] The Newton-Wigner (NW) scheme was originally developed in order to have a position operator in relativistic quantum theory, which leaves the positive energy property of a positive energy wave functions invariant. However, the price to pay turns out to be unacceptable: It leads to a deviation from Born's rule, a probability density which does not satisfy a continuity equation with respect to some probability current, the successful minimal coupling to an electromagnetic field does not work in the NW-representation and it violates Lorentz invariance in the sense that a NW-localized state in some Lorentz frame is not NW-localized in any other frame (see [2] and references therein). Nonetheless, the eigenstates of the NW-Operator are (in ordinary position representation) extremely localized Bessel-type functions of positive energy, which, beyond the characteristic length scale of the particle under consideration, virtually look like delta functions.

illustrate very nicely that such wave functions can be virtually zero already a few Compton wavelengths or even less away from their center. Moreover, it is straightforward to argue that an electron, for instance, which has interacted with an apparatus or, more generally, with its environment will have an extremely well localized wave function. Such localization processes are well understood in the context of decoherence theory (see, e.g., [23] and references therein).

Since wave functions in quantum theory are the amplitudes of a probability measure, this means that tails can be neglected for all practical purposes. Just as we disregard predictions of negligible probability in thermodynamics (such as rocks suddenly flying up instead of down due to a fluctuation in the thermal velocities of their molecules), we must of course do the same in quantum mechanics, so that we can safely assume wave functions to be compactly supported when making empirical predictions.

But it is good to be aware of the fact that in textbook quantum mechanics the probability interpretation of wave functions is a postulate and there is no statistical analysis (such as Boltzmann's statistical analysis of classical mechanics) to justify it. In Bohmian mechanics, on the other hand, a theory that describes matter as composed of literal particles that always have a position and whose motion is guided by their quantum mechanical wave function, such a statistical analysis can be performed [14–16]. That way, by analyzing the Bohmian equations of motion for measurement-like situations, the quantum probabilities can be derived as predictions for associated (typical) empirical relative frequencies by proving a law of large numbers. And the crucial assumption that goes into a proof of the law of large numbers (and thus, from the Bohmian point of view, establishes the quantum probabilities that are so successful for predictions) is that incredibly improbable events will not happen with empirical certainty (sometimes called Cournot's principle).

When the meaning and status of probabilities is less clear, the issue of infinite tails may be more problematic. This becomes particularly obvious in the Many-Worlds interpretation (MWI), where even the smallest probability events will (at least in a measurement context) actually be realized in some world. However one may interpret the quantum probabilities in MWI and however one may define its ontological content, one probably cannot avoid the fact that there are real worlds in which the infinite tails of positive energy wave functions are empirically relevant (see [26] for details and a remarkable example).

Transformations: The nonlocal nature of relativistic positive energy wave functions seems to be physically more interesting than infinite tails. Local transformations of relativistic positive energy wave functions necessarily lead to contributions of negative energy states in the resulting state. Moreover, even nonlocal transformations must be extremely special in order to rescue the positive energy property since the values of a relativistic positive energy wave function in any neighborhood already determines the whole function. And so it can be assumed that at the level of description of one-particle (or N-particle) wave function, transitions between negative and positive spectrum necessarily occur in physical processes (free time evolution as

well as trivial time evolution with a moderate stationary potential are perhaps the only non-trivial and obvious transformations that are special enough to leave spectral subspaces invariant).

Let us now commit ourselves to the special choice of the Dirac equation, which is the basis for the description of fermions and thus for the description of matter (electrons, quarks, etc.). However, the level on which the theory of fermions is not only empirically adequate but impressively successful in its predictions (antimatter, pair creation, Lamb shift etc.) is not that of one-particle or N-particle solutions of the Dirac equation but that of the associated quantum field theory (QFT), in case of the Dirac equation (external field) quantum electrodynamics (QED). This theory can be developed starting from the Dirac equation by second quantization (or more picturesquely from the Dirac sea picture) by allowing roughly speaking for a variable number of particles and interpreting negative energy wave functions by the operation of charge conjugation as positive energy wave functions of antiparticles. Transitions between negative and positive energies on the level of solutions of the Dirac equation thereby correspond to particle creation and annihilation processes with certain probabilities when lifted to the level of QED (see, e.g., [17, 18, 31, 37]).

Thus, the fragility of positive energy wave functions with respect to nontrivial (e.g., local) transformations, discussed above, suggests that interaction (causing such transformations) is intrinsically associated with particle creation and annihilation processes. Of course, it is to be expected that for everyday processes the corresponding probabilities are again negligibly small, only when high energies are involved this is no longer the case.

3 An Operational Implication

Now we come to an operational implication of Theorem 1. It shall be exemplified by a very general framework for describing a spatial detector experiment. The latter may be taken as only a representative of any local measurement (if any measurement device is triggered by a quantum system, the system was detected in the spatial region of the device).

3.1 Covariant Detector Formalism

Quantum Formalism: First, we assume that the probability that a detector covering a given spatial region is triggered by a quantum system at a given time (in the lab frame) can be expressed and calculated by the quantum formalism. This means that the click probability in the lab frame is given by an expression of the form

$$\mathbb{P}^{\psi}\left(\mathscr{D}_{(0,\Delta)}\right) = \langle\psi|\ D_{(0,\Delta)}\,\psi\rangle \tag{7}$$

Here, $\mathscr{D}_{(0,\Delta)}$ represents the event that a detector covering detector region $\Delta \subset \mathbb{R}^3$ is triggered at lab-time $t = 0$, the 'probability operator' $D_{(0,\Delta)}$ (sometimes called 'effect') has the property $0 \leq D_{(0,\Delta)} \leq 1$ and shall be an operator in the Heisenberg picture which acts on the Hilbert space of the measured system $\mathcal{H}$ and $\psi \in \mathcal{H}$ is the initial (pure[5]) state. For instance, in the standard ideal measurement scheme of textbooks $D_{(0,\Delta)}$ would be a projection but more generally and more adequate for realistic measurements it is an element of a (not necessarily projective) POVM.

Space-Time Translations: There is a unitary representation of space-time translations acting on $\mathcal{H}$ which has spectral representation[6]

$$U(x) = e^{i\widehat{P}x} = \int e^{ipx} d^4E(p) \tag{9}$$

Here $\widehat{P}x = \widehat{P}_\mu x^\mu$ and $px = p_\mu x^\mu$ are the Minkowski scalar products, and E is a PVM on $\mathbb{R}^4$ acting on $\mathcal{H}$, the PVM of the energy-momentum operator $\widehat{P}^\mu = \int p^\mu\, d^4E(p)$ which can be identified as the infinitesimal generator of space-time translations.

Space-Time Translation Covariance: We assume that space-time translations act naturally on the operators $D_{(0,\Delta)}$: If $x = (s, \boldsymbol{a}) \in \mathbb{R}^4$, the probability that a detector covering $\Delta + \boldsymbol{a}$ is triggered at time $t = s$ in the laboratory frame is given by

$$\mathbb{P}^{\psi}\left(\mathscr{D}_{(s,\Delta+\boldsymbol{a})}\right) = \langle\psi|\ D_{(0,\Delta)+x}\,\psi\rangle \tag{10}$$

where $D_{(0,\Delta)+x} = U(x)D_{(0,\Delta)}U^{-1}(x) \equiv D_{(s,\Delta+\boldsymbol{a})}$.

3.1.1 Additional Assumptions

To obtain the desired operational result (Theorem 2 below), the covariant detector formalism must satisfy some additional assumptions:

---

[5] We might also work with the more general expression

$$\mathbb{P}^{\rho}\left(\mathscr{D}_{(0,\Delta)}\right) = \mathrm{Tr}_{\mathcal{H}}\left[D_{(0,\Delta)}\,\rho\right] \tag{8}$$

where ρ is the initial density matrix, which need not be a pure state. However, since mixed states can always be expressed by (convex) linear combinations of pure states, we can build the following analysis on expression (7) without loss of generality.

[6] The fact that $U(x)$ can be written in this form is of course well known for concrete models of (relativistic) quantum theory and is ensured more generally by an immediate generalization of Stone's theorem from unitary strongly continuous representations of one parameter groups to unitary strongly continuous representations of general locally compact abelian groups, which is sometimes called the SNAG-theorem (according to Stone, Naimark, Ambrose and Godement) [27].

Spectrum Condition: We assume that the generator P_μ of space time translations (the energy-momentum operator) has its spectrum in the closed forward light cone: $\sigma(P_\mu) \subset \overline{V}_+ = \{p \in \mathbb{R}^4 \mid p_\mu p^\mu \geq 0,\ p_0 \geq 0\}$ (see Sect. 1).

Additivity: Now comes a very special assumption. We assume that for $\Delta \cap \Delta' = \emptyset$ and all $\psi \in \mathcal{H}$ there is a joint distribution of the events $\mathscr{D}_{(t,\Delta)}$ and $\mathscr{D}_{(t,\Delta')}$ such that

$$\mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)} \vee \mathscr{D}_{(t,\Delta')}\right) = \mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)}\right) + \mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta')}\right) \tag{11}$$

This assumption is not justified for general quantum systems; rather, it corresponds to a selection of very special quantum systems for which it appears to be a reasonable assumption. Indeed, the existence of a joint distribution alone only implies (see [2])

$$\mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)} \vee \mathscr{D}_{(t,\Delta')}\right) = \mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)}\right) + \mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta')}\right) - \mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)} \wedge \mathscr{D}_{(t,\Delta')}\right) \tag{12}$$

Therefore, Eq. (11) is equivalent to the requirement

$$\mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)} \wedge \mathscr{D}_{(t,\Delta')}\right) = 0 \tag{13}$$

i.e., it expresses the requirement that distant detectors cannot be triggered at the same time, given ψ is the initial state. Making this assumption for all $\psi \in \mathcal{H}$ seems to be justified if $\mathcal{H}$ is a Hilbert space of one particle wave functions, which might be taken to be also a subspace of a larger Hilbert space like the one particle sector of Fock space. If we set now

$$D_{(t,\Delta)\cup(t,\Delta')} := D_{(t,\Delta)} + D_{(t,\Delta')} \tag{14}$$

we thus obtain $\mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)} \vee \mathscr{D}_{(t,\Delta')}\right) = \langle\psi|\ D_{(t,\Delta)\cup(t,\Delta')}\ \psi\rangle$.

Additivity is actually not an independent assumption but rather a motivation for its relativistic generalization, causal additivity, which includes additivity as a special case:

Causal Additivity: In a relativistic theory, the natural generalization of additivity is the following: whenever (t, Δ) and (t', Δ') are spacelike separated

$$\mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)} \vee \mathscr{D}_{(t',\Delta')}\right) = \mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)}\right) + \mathbb{P}^\psi\left(\mathscr{D}_{(t',\Delta')}\right) \tag{15}$$

This condition is equivalent to the exclusion of joint detector clicks of two distant detectors at spacelike separation, i.e.,

$$\mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)} \wedge \mathscr{D}_{(t',\Delta')}\right) = 0 \tag{16}$$

By setting $D_{(t,\Delta)\cup(t',\Delta')} \equiv D_{(t,\Delta)} + D_{(t',\Delta')}$ we thus obtain $\mathbb{P}^\psi\left(\mathscr{D}_{(t,\Delta)} \vee \mathscr{D}_{(t',\Delta')}\right) = \langle\psi|\ D_{(t,\Delta)\cup(t',\Delta')}\ \psi\rangle$. This condition can be appropriately called *causal additivity* [2].

Local Commutativity: The last (relativity inspired) condition we need is the well known condition of local commutativity: whenever (t, Δ) and (t', Δ') are spacelike separated

$$\left[D_{(t,\Delta)}, D_{(t',\Delta')}\right] = 0 \tag{17}$$

This condition is usually demanded to exclude the possibility to use quantum nonlocality in order to send signals faster than light (Lüders theorem). For a detailed discussion of this condition and further physical motivations see Chap. 3 of [2].

3.2 A No-Go Theorem

Theorem 1 now implies the following result.[7]

Theorem 2 *A (non trivial) covariant detector formalism which satisfies the spectrum condition, local commutativity and causal additivity does not exist.*

The proof can be found in [2] (Theorem 4.25). Roughly speaking, it applies space-time translations to various detector arrangements[8] and thus shows that all click probabilities have an upper bound which can be made inductively arbitrarily small (in this sense, 'non-trivial' in Theorem 2 means 'with non-vanishing click probabilities'). The crucial step uses Theorem 1 by applying it to functions f of the form $f(x) = \langle \varphi | \, U(x) \, \psi \rangle$.

3.3 Discussion

Since there are detectors in the world which can be triggered by quantum systems,[9] Theorem 2 requires an explanation. One might question any of its assumptions, but of course the assumption of causal additivity is most questionable. Moreover, the discussion of the fragility of the positive energy property of relativistic wave functions with respect to nontrivial transformations together with the observation that spectral transitions of Dirac wave functions correspond to particle creation and

[7] Theorem 2 goes back to a theorem proved in its first version by Schlieder [32] and then gradually refined by Jankewitz [25], Malament [28] and Halvorson and Clifton [19], often known as Malament's Theorem.

[8] To be precise, the proof uses the obvious generalization of the causal additivity condition to arrangements with more than two detectors, but we skip that here for simplicity.

[9] As mentioned above, detector experiments are in this analysis only a representative of practically any quantum measurement (the measured system is detected in the laboratory). One might even argue somewhat drastically that our perception of matter is of this kind in the first place (given the measurement problem has been solved): When I see the table in front of me, I detect the position of a quantum system, given by a huge cluster of atoms, which together form a table.

annihilation processes in QED in Sect. 2.4 motivates also a closer look at the spectrum condition.

Thus, we shall not question here the assumption that the statistics of detector clicks can be predicted by a covariant detector formalism and that local commutativity is true. So according to Theorem 2, either the spectrum condition or causal additivity must be violated. Fortunately, these two options naturally complement each other. According to quantum theory, each measurement is associated with a state transformation[10] (also one with a negative outcome, like a switched on detector which was not (yet) triggered). And as argued in Sect. 2.4, most state transformations (in particularly if caused by a localized measuring device) cause spectral transitions on the one- or N-particle level of description, which in turn correspond for fermions to pair creation processes with certain probabilities in the associated QFT. This suggests to expect, for the quantum mechanical description of detector experiments, a violation of the spectrum condition at the level of Dirac wave functions and, when these processes have been lifted to the level of QED, corresponding transitions between the particle number sectors of the fermionic Fock space (while the spectrum condition is rescued in QED by charge conjugation of negative energy states). And the latter immediately destroys any basis for expecting causal additivity to hold in certain situations (one-particle initial states).

To see this, recall that causal additivity corresponds to the assumption that two distant detectors cannot be triggered at spacelike separation and its violation is therefore equivalent to the condition

$$\mathbb{P}^{\psi}\left(\mathscr{D}_{(t,\Delta)} \wedge \mathscr{D}_{(t',\Delta')}\right) > 0 \tag{18}$$

for spacelike separated (t, Δ) and (t', Δ'). For initial states ψ in the one-particle sector of Fock space[11] this appears to be against the spirit of relativity (a particle moving faster than light to trigger two detectors at spacelike separation). However, if the state transformations associated with such measurements do not leave the one particle sector of Fock space invariant, this violation appears quite natural. For instance, the state transformation caused by the potential of a switched on detector can create a particle by which this detector is being triggered. Since the state transformation associated with a probability operator D is encoded in a linear operator

[10] For instance, the probability operator D associated with a triggered detector (for simplicity we suppress the subscript (t, Δ) here) can be associated with a state transformation operator $\mathcal{R}$ so that an initial state ψ transforms according to $\psi \mapsto \frac{\mathcal{R}\,\psi}{\|\mathcal{R}\,\psi\|}$ and $D = \mathcal{R}^{\dagger}\mathcal{R}$ (for ideal measurements of textbooks, D and $\mathcal{R}$ would be one and the same projection operator, which corresponds to the projection postulate). For more general measurements which cannot be described on the level of pure states, the state transformation is associated with a set $\{\mathcal{R}_k\}$ of linear operators, so that an initial density matrix ρ transforms according to $\rho \mapsto \frac{\sum_k \mathcal{R}_k^{\dagger}\,\rho\,\mathcal{R}_k}{\mathrm{Tr}\sum_k \mathcal{R}_k^{\dagger}\,\rho\,\mathcal{R}_k}$ (Kraus representation). See [2] for a detailed development of the general measurement formalism.

[11] Theorem 2 can be generalized to an analogous assertion corresponding to any $N-$particle sector of Fock space: initial states for which it can be perfectly excluded that more than N detectors (for any $N \in \mathbb{N}$) are triggered at spacelike separation do not exist under the assumptions (see Corollary 4.27 in [2]).

$\mathcal{R}$ so that $\mathbb{P}^{\psi}(\cdot) = \langle \psi |\ D\ \psi \rangle = \|\mathcal{R}\,\psi\|^2$ (see Footnote 10), the state transformation, in a sense, enters into the probabilities: even if ψ was a state of a single particle, the predicted statistics can be statistics of many-particles if $\mathcal{R}$ does not leave the one-particle sector of Fock space invariant. Such operators are also well known in connection with observable quantities; the PVM of the local charge density operator in QED, for example, has this property [38].

This fits very well with a well-known result from the more abstract framework of axiomatic or algebraic quantum field theory (AQFT), the *Reeh-Schlieder theorem* (see, e.g., [39] for a comprehensive discussion), which can be also derived from a generalization of Theorem 1 (see [2]). The Reeh-Schlieder theorem implies (under the assumptions of AQFT) that the click probability of a local detector cannot be (exactly) zero even if the initial state is the vacuum state.

To conclude this discussion, note that the fact that causal additivity must be violated says nothing about the magnitude of this violation. The probabilities in (18) expressing this violation can be negligibly small, though not precisely zero. If no high energies are involved, negligibly small probabilities (18) are of course to be expected for one particle initial states ψ.

4 Towards a Spatial Distribution

The way the probability operators $D_{(t,\Delta)}$ were defined above, they belong in the first place to a two element POVM $\left\{D_{(t,\Delta)}\,,\ \mathbb{1}_{\mathcal{H}} - D_{(t,\Delta)}\right\}$ associated with two possible outcomes (say 'click $\equiv 1$' and 'no click $\equiv 0$'), which is the minimal structure to describe a detector experiment. However, one has in mind a more general structure, namely a general spatial distribution of a quantum system, which agrees with the click-probabilities given by this POVM for the detector regions.

Theorem 2 now also proofs the non-existence of a relativistically satisfying more general spatial POVM on physical space $\mathbb{R}^3$ (instead of $\{0\,,1\}$) under its assumptions (spectrum condition etc.). To see this, one can simply replace the meaning of the detector regions $\Delta \subset \mathbb{R}^3$, with arbitrary Borel sets $\Delta \subset \mathbb{R}^3$ of physical space. So consider now a spatial POVM in the Heisenberg picture acting on the considered Hilbert space, formed (at a fixed lab-time t) by positive operators $D_{(t,\Delta)}$ with Δ varying in the (measurable) subsets of $\mathbb{R}^3$. As a POVM, it must be additive, i.e., $D_{(t,\Delta)\cup(t,\Delta')} = D_{(t,\Delta)} + D_{(t,\Delta')}$ for all $\Delta \cap \Delta' = \emptyset$ and normalized, i.e., $\int_{\mathbb{R}^3} D_{(t,d^3x)} = \mathbb{1}_{\mathcal{H}}$ is the identity operator (normalization does not play any role for the present considerations). The additivity of such a POVM directly corresponds to the additivity condition (14) above and expressing it in terms of probabilities (i.e., $\mathbb{P}^{\psi}\left(\mathscr{D}_{(t,\Delta)}\right) = \langle \psi |\ D_{(t,\Delta)}\ \psi \rangle$ etc.) yields $\mathbb{P}^{\psi}\left(\mathscr{D}_{(t,\Delta)} \vee \mathscr{D}_{(t,\Delta')}\right) = \mathbb{P}^{\psi}\left(\mathscr{D}_{(t,\Delta)}\right) + \mathbb{P}^{\psi}\left(\mathscr{D}_{(t,\Delta')}\right)$ and thus again $\mathbb{P}^{\psi}\left(\mathscr{D}_{(t,\Delta)} \wedge \mathscr{D}_{(t,\Delta')}\right) = 0$ now for all disjoint spatial Borel sets $\Delta \cap \Delta' = \emptyset$. Calling the event $\mathscr{D}_{(t,\Delta)}$ sloppily 'the system is localized in Δ' we can thus phrase the additivity condition by 'the system cannot be localized in two disjoint regions at the same time' (a condition which is clearly false for, say, a two particle system). Causal additivity in this sense means that 'the

system cannot be localized in two spacelike separated regions', a condition which is the natural relativistic generalization of additivity.

Theorem 2, reformulated with respect to a spatial POVM, then says that such a POVM does not exist under the assumptions and thereby a corresponding probability distribution on physical space $\mathbb{R}^3$ does not exist. But what about the $|\psi(\boldsymbol{x})|^2-$distribution, which lays the foundation for the predictive success of quantum theory? If we want to express this distribution by a POVM, say for a positive energy solution ψ of the Dirac equation, there are two options at hand: One can use the indicator functions $\chi_\Delta(\boldsymbol{x})$ of (measurable) spatial subsets $\Delta \subset \mathbb{R}^3$ (the PVM of the standard position operator) or their projection[12] $P_+\,\chi_\Delta(\boldsymbol{x})\,P_+$ onto the positive energy subspace of the associated Hilbert space $\mathcal{H} = L^2(\mathbb{R}^3, d^3x) \otimes \mathbb{C}^4$ (both of which form a POVM on $\mathbb{R}^3$), since for $\psi \in \mathcal{H}_+ = P_+\mathcal{H}$ we have the $|\psi|^2-$weight of Δ

$$\mathbb{P}^\psi(\Delta) = \int_\Delta |\psi(\boldsymbol{x})|^2 d^3x = \langle\psi|\ \chi_\Delta(\boldsymbol{x})\,\psi\rangle = \langle\psi|\ P_+\,\chi_\Delta(\boldsymbol{x})\,P_+\,\psi\rangle \tag{19}$$

However, both of these POVMs violate assumptions of Theorem 2: multiplication of a positive energy wave function by the indicator functions obviously violates the spectrum condition by radically cutting off everything from the wave function outside of Δ which yields massive contributions from negative energy eigenstates (observe that an infinite potential well, i.e. an infinite amount of energy would be necessary to realize this operation physically) while their projection onto the positive energy subspace violates local commutativity. For the latter fact, Theorem 2 can be taken as a proof, but one may also prove it by direct calculation.

Nonetheless, the probability distribution given by $\mathbb{P}^\psi(\Delta) = \int_\Delta |\psi(\boldsymbol{x})|^2\, d^3x$ is well defined for positive energy states ψ, as long as we do not consider state transformations. A state transformation does not occur if a particle 'is there' (say in Bohmian mechanics) but occurs upon measurement.[13]

5 Particle Ontology

In view of the previous discussion it is clear, in principle, that the mentioned results do not pose a problem for a quantum theory with a particle ontology, provided it

[12] The projection operator P_+ onto the positive energy subspace of the Hilbert space of solutions of the Dirac equation can be written as $P_+ = \frac{1}{2}\left(\mathbb{1}_\mathcal{H} + \frac{\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta m}{\sqrt{\boldsymbol{p}^2+m^2}}\right)$, with the usual meaning of the symbols, see, e.g., [37].

[13] In particular, if D is an element of a POVM, the state transformation upon the associated measurement result is of the form $\psi \mapsto \mathcal{R}\,\psi = U\sqrt{D}\,\psi$ (or a generalization of this formula, if the measurement transforms pure states to mixed states), where U is a partial isometry. If $U = \mathbb{1}_\mathcal{H}$ and $\sqrt{D} = D = D^2$ is a projection, we recover the projection postulate for ideal measurements. See [2] for details, see also Footnote 10.

is able to describe particle creation and annihilation (which, of course, it should be for other reasons as well, if it is to reproduce the results of empirically successful relativistic quantum field theories).

There are several proposals for generalizing non-relativistic Bohmian mechanics to relativistic[14] QFT [3, 7, 8, 11–13, 35, 36]. The most elaborated of of these approaches is the so called *Bell type QFT* [11–13], which can be described in a very simplified way as follows: The configuration space is the collection of the configuration spaces (sectors) for each possible particle number[15] (and antiparticle number) and each sector is associated with a wave function (non-normalized and possibly zero) from the $N-$particle sector of the corresponding Fock space. The actual Bohmian configuration lives in a definite sector (N particles) at each instant and its distribution there is a $|\psi_N|^2-$distribution, ψ_N being the respective sector wave function. In the absence of jumps to other sectors (see below) the actual configuration is deterministically guided by the corresponding sector wave function through a Bohmian guiding equation (for the guiding equation of Dirac theory, see, e.g., [13]). An additional stochastic jump law provides us with probabilities for where and when particles may be created and/or annihilated (the jump process is driven by the interaction part of the second quantized Hamiltonian). For a given QFT (like regularized QED), these laws define a Markov process on the configuration space consisting of deterministic motion in an actual sector interrupted by stochastic jumps between the sectors, from which the empirical predictions (like cross sections, Lamb shift etc.) of this theory can be derived.

So the crucial question is how the $|\psi|^2-$distribution of Bohmian configurations fits with the absence of such a distribution for position measurements. In the non-relativistic case, the Bohmian positions of course agree with the results of (good) position measurements, at least to a good degree of accuracy.[16] In a relativistic Bohmian QED, this should be the case as well. However, there is a notable difference: the state transformation of a position measurement now not only localizes the wave function of a measured system (or suppresses it in regions where the measurement result is negative) in its actual particle sector, but also generates transitions in the particle number of the measured system with certain probabilities. The presence of a measuring apparatus can thus change the configuration of a measured system by changing the (actual sector of) configuration space. This in turn changes the

[14] In this context, relativistic QFT refers to QFT with particle creation and annihilation, based on a relativistic wave equation. The question of full Lorentz invariance is another question which is not treated in this work. Both regularization of the OFT and a description of N particles with nonlocal dynamics pose challenges for a fully Lorenzt invariant description; for treatment of the second point, cf. [1, 9, 10].

[15] Details of treatment of identical particles, different particle species etc. are skipped here.

[16] Empirical distributions of real world measurements must always minimally deviate from this prediction because measurements are never perfect but are always subject to certain errors with certain probabilities (such errors arise even at the fundamental level due to the quantum mechanical nature of measuring devices [2]). Implementing these uncertainties into the measurement scheme leads to an *approximate measurement POVM*, where the indicator functions of the standard position PVM are convoluted with an additional error distribution, cf. [2].

probabilities of outcomes of the measurement (even if this change will be negligible, if no too large energies are involved).

Therefore, one should expect that the POVM describing a Bohmian position measurement deviates to some degree from the POVM describing the actual distribution of Bohmain positions. While an electron's Bohmian position, for instance, is $|\psi|^2-$distributed, where ψ is a one particle wave function of positive energy, we should not expect the statistics of its position measurement to be given (precisely) by the corresponding POVM $\{P_+ \, \chi_\Delta(\boldsymbol{x}) \, P_+\}$ (possibly lifted to Fock space), which does not commute locally,[17] but by an operator which includes possible transitions in the particle number due to the intervention of the measuring device. A generic option for an operator describing position measurements[18] for fermions would be the PVM of the local charge density operator [38], which commutes locally but does not leave the one particle sector of Fock space invariant and hence violates causal additivity. When looking at a concrete position measurement, the question of the associated POVM of course depends on the theoretical modeling of the details of the measurement interaction (detector model).

It is interesting to note that another Bohmian dynamics corresponding directly to the statistics given by the PVM of the local charge density operator can also be defined quite naturally for fermionic Bell type QFT as shown in [38]. This theory is empirically equivalent to the fermionic Bell type QFT sketched above (as both are empirically equivalent to regularized QED of textbooks), but they are not equivalent on the ontological level. While in the absence of interaction in the latter case there is no particle creation and annihilation, in the former case configurations can jump between the sectors even under the free time evolution.

Acknowledgements I wish to thank Dustin Lazarovici for many productive discussions and for valuable comments on an earlier version of this article. I would also like to thank Roderich Tumulka for helpful discussions on this topic in the past. And I wish to express my deep gratitude to my teacher and friend Detlef Dürr, for the unique way he taught us and for his unique kindness and warmth.

[17] Observe that the motivation to require local commutativity is to exclude the possibility that a nonlocal state transformation upon measurement (see Footnotes 10 and 13) can be used to send superluminal signals (see Chap. 3 of [2] for a detailed analysis). For probability operators which are not associated with such state transformations, there is no justification for such a requirement.

[18] One might suggest to use the standard position PVM given by the indicator functions $\chi_\Delta(\boldsymbol{x})$. However, its violation of the spectrum condition is too strong so that any attempt to directly lift it to Fock space by second quantization will fail, because, roughly speaking, its action on Fock space would create infinitely many pairs, as can be estimated, e.g., from its Foldy-Wouthuysen representation (cf. [37]).

References

1. C. Beck, Wavefunctions and Minkowski Space-Time–On the Reconciliation of Quantum Theory with Special Relativity (2020). arXiv:2009.00440
2. C. Beck, *Local Quantum Measurement and Relativity* (Springer, 2021)
3. D. Bohm, B.J. Hiley, *The Undivided Universe: An Ontological Interpretation of Quantum Theory* (Routledge, London, 1995)
4. H.-J. Borchers, A remark on a theorem of B. Misra. Commun. Math. Phys. **4**(5), 315–323 (1967)
5. A. Bracken, J. Flohr, G. Melloy, Time-evolution of highly localized positive-energy states of the free Dirac electron, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 461 (The Royal Society, 2005), pp. 3633–3645
6. A. Bracken, G. Melloy, Localizing the relativistic electron. J. Phys. Math. Gen. **32**(34), 6127 (1999)
7. S. Colin, W. Struyve, A Dirac sea pilot-wave model for quantum field theory. J. Phys. Math. Theor. **40**(26), 7309 (2007)
8. D.-A. Deckert, M. Esfeld, A. Oldofredi, A persistent particle ontology for QFT in terms of the Dirac sea (2016). arXiv:1608.06141
9. D. Dürr, S. Goldstein, K. Münch-Berndl, N. Zanghì, Hypersurface bohm-dirac models. Phys. Rev. A **60**(4), 2729–2736 (1999)
10. D. Dürr, S. Goldstein, T. Norsen, W. Struyve, N. Zanghì, Can Bohmian mechanics be made relativistic? Proc. R. Soc. A. **470**(2162), 20130699 (2014)
11. D. Dürr, S. Goldstein, R. Tumulka, N. Zanghi, Trajectories and particle creation and annihilation in quantum field theory. J. Phys. Math. Gen. **36**(14), 4143 (2003)
12. D. Dürr, S. Goldstein, R. Tumulka, N. Zanghi, Bohmian mechanics and quantum field theory. Phys. Rev. Lett. **93**(9), 090402 (2004)
13. D. Dürr, S. Goldstein, R. Tumulka, N. Zanghi, Bell-type quantum field theories. J. Phys. Math. Gen. **38**(4), R1 (2005)
14. D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the origin of absolute uncertainty. J. Stat. Phys. **67**(5–6), 843–907 (1992)
15. D. Dürr, S. Goldstein, N. Zanghì, *Quantum Physics Without Quantum Philosophy* (Springer, Berlin Heidelberg, 2012)
16. D. Dürr, S. Teufel, *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory* (Springer, Berlin, 2009)
17. W. Greiner, *Relativistic Quantum Mechanics* (Springer, Wave equations, 2000)
18. W. Greiner, B. Müller, J. Rafelski, *Quantum Electrodynamics of Strong Fields: With an Introduction into Modern Relativistic Quantum Mechanics*. (Springer Science & Business Media, 2012)
19. H. Halvorson, R. Clifton, No place for particles in relativistic quantum theories? Philos. Sci. **69**(1), 1–28 (2002)
20. G.C. Hegerfeldt, Violation of causality in relativistic quantum theory? Phys. Rev. Lett. **54**(22), 2395–2398 (1985)
21. G.C. Hegerfeldt, Instantaneous spreading and Einstein causality in quantum theory (1998). arXiv: quant-ph/9809030
22. G.C. Hegerfeldt, S.N. Ruijsenaars, Remarks on causality, localization, and spreading of wave packets. Phys. Rev. D **22**(2), 377–384 (1980)
23. K. Hornberger, Introduction to decoherence theory, in *Entanglement and Decoherence*, (Springer, 2009), pp. 221–276
24. M. Ikawa, *Hyperbolic Partial Differential Equations and Wave Phenomena*, vol. 2 (American Mathematical Soc., 2000)
25. B. Jancewicz, Operator density current and relativistic localization problem. J. Math. Phys. **18**, 2487 (1977)
26. D. Lazarovici, C. Beck, Nonlocal vacuum phenomena. *In Preparation* (2023)

27. G.W. Mackey, Harmonic analysis and unitary group representations: the development from, to 1950. Cahiers du Séminaire d'histoire des mathématiques **2**(1992), 13–42 (1927)
28. D. Malament, In defense of dogma: why there cannot be a relativistic quantum mechanics of (localizable) particles, in *Perspectives on Quantum Reality: Non-Relativistic, Relativistic, and Field-Theoretic* (1966), pp. 1–10
29. T.D. Newton, E.P. Wigner, Localized states for elementary systems. Rev. Mod. Phys. **21**, 400–406 (1949)
30. T. Philips, Lorentz invariant localized states. Phys. Rev. **136**(3B), B893 (1964)
31. P. Pickl, Existence of spontaneous pair creation. Ph.D. thesis, Ludwig-Maximilians-Universität München, 2005
32. S. Schlieder, Zum kausalen Verhalten eines relativistischen quantenmechanischenSystems. *Quanten und Felder. Braunschweig: Friedrich Vieweg+ Sahn* (1971), pp. 145–160
33. F. Schwabl, *Advanced Quantum Mechanics* (Springer Science & Business Media, 2005)
34. S.S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Dover Publications, 2005)
35. W. Struyve, Pilot-wave theory and quantum fields. Rep. Prog. Phys. **73**(10), 106001 (2010)
36. W. Struyve, Pilot-wave approaches to quantum field theory, in *Journal of Physics: Conference Series*, vol. 306 (IOP Publishing, 2011), p. 012047
37. B. Thaller, *The Dirac Equation (Texts and Monographs in Physics)*, vol. 91 (Springer-Verlag, Berlin, 1992), pp. 1105–1115
38. R. Tumulka, Positron position operators. I. A natural option. Ann. Phys. **443**, 168988 (2022)
39. E. Witten, Notes on some entanglement properties of quantum field theory (2018). arXiv:1803.04993

Detection Time of Dirac Particles in One Space Dimension

A. Shadi Tahvildar-Zadeh and Stephanie Zhou

Abstract We consider particles emanating from a source point inside an interval in one-dimensional space and passing through detectors situated at the endpoints of the interval that register their arrival time. Unambiguous measurements of arrival or detection time are problematic in the orthodox narratives of quantum mechanics, since time is not a self-adjoint operator. By contrast, the arrival time at the boundary of a particle whose motion is being guided by a wave function through the deBroglie-Bohm guiding law is well-defined and unambiguous, and can be computationally feasible provided the presence of detectors can be modeled in an effective way that does not depend on the details of their makeup. We use an absorbing boundary condition for Dirac's equation (ABCD) proposed by Tumulka, which is meant to simulate the interaction of a particle initially inside a domain with the detectors situated on the boundary of the domain. By finding an explicit solution, we prove that the initial-boundary value problem for Dirac's equation satisfied by the wave function is globally well-posed, the solution inherits the regularity of the initial data, and depends continuously on it. We then consider the case of a pair of particles emanating from the source inside the interval, and derive explicit formulas for the distribution of first arrival times at each detector, which we hope can be used to study issues related to non-locality in this setup.

1 Introduction

In orthodox quantum mechanics, speaking of arrival/detection time of quantum particles is fraught with problems, since time does not lend itself to the self-adjoint operator formalism required for that approach.[1] Nevertheless, the arrival/detection

[1] It is possible to define time more generally as a *positive operator-valued measure* [13], but there is no unique way of doing that [18].

Dedicated to the memory of Detlef Dürr

A. S. Tahvildar-Zadeh (✉) · S. Zhou
Department of Mathematics, Rutgers University, Piscataway, NJ, USA
e-mail: shadit@math.rutgers.edu

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_14

time of a particle is something that is routinely measured in time-of-flight (TOF) experiments performed in labs [8, 9, 11]. There are numerous competing recipes in the literature for what the distribution of arrival/detection times should be (see [2–5] for a critique of some of these approaches, and the possibility of using experiments to distinguish them.) Many of these approaches ignore the presence of detection devices and consider the wave function of the particle to evolve unitarily, either under the free evolution or in presence of an external potential. As a result, they are prone to the "backflow" problem, which can cause their proposed candidates for the arrival time probability density to become negative at points close to the source. An alternative approach was taken by Tumulka [14–16], based on an idea of Werner [19], in which the presence of the detector is modeled through the imposition of an absorbing boundary condition on the Schrödinger flow (in the non-relativistic case) or the Dirac flow (in the relativistic case) of the wave function of the particle. Such a boundary condition ensures that particle velocities at the boundary are always outward. Tumulka showed that under such a boundary condition, his candidate for the probability density of the particle's arrival time is always non-negative. Subsequently, Teufel and Tumulka succeeded in showing [12] that the corresponding boundary value problem for the wave function has a unique square-integrable solution, in both the relativistic and the non-relativistic cases. Their existence proof uses techniques from functional analysis, and does not yield an explicit formula for the solution in either case.

In this note we show that in one space dimension, the initial-boundary-value problem for the Dirac equation satisfied by the wave function of a single spin-half particle, with absorbing boundary conditions corresponding to a pair of *ideal* detectors[2] placed at the two endpoints of an interval containing the particle source, is exactly solvable, and that the solution inherits the regularity of the initial data. Applying Bohmian Mechanics rules then makes it possible for actual particle trajectories to be computed for any particle whose initial position is distributed randomly according to any given initial wave function, thereby setting the stage for comparisons to be made with other proposals for arrival time distribution, and the possibility of experimental testing of this theory.

It is of interest to study how successful the ABCD method is in simulating actual detection, i.e. the interaction of the particle with the (presumably macroscopic) detecting apparatus that results in the device registering the presence of the particle at a particular time. In order to avoid faster-than-light signaling, it must be the case that shifting one of the detectors by a small amount does not alter the distribution of arrival times at the other detector, or at least not before any possible light signal from the shifted detector has time to reach the other one. (Recall that in Bell-type experiments, changing the direction of the magnets affects only the correlations between the distributions, not the distributions themselves.) For a single particle, this is easily verified.

We then derive explicit formulas for the distribution of arrival times at each detector, which in forthcoming work we plan to use to show that such superluminal

[2] See [15] for the definition of ideal versus non-ideal detectors.

signaling does not exist even when the initial wavefunction corresponds to a maximally entangled two-particle state.

2 Absorbing Boundary Conditions for the Dirac Equation

Using the proposed equations in [15], we let $\Omega = (-L, L)$ be an open interval in $\mathbb{R}$ and let $\psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} : [0, \infty) \times \Omega \to \mathbb{C}^2$ be the unique solution of the initial-boundary value problem (IBVP)

$$\begin{cases} ic\hbar\gamma^\mu \partial_\mu \psi = mc^2 \psi \\ \psi(0, s) = \mathring{\psi}(s); \quad s \in \Omega, \qquad \mathring{\psi} \in C_c^k(\Omega),\ k \geq 0 \\ \mathbf{n}(s) \cdot \alpha\ \psi(t, s) = \psi(t, s);\ t \geq 0, s \in \partial\Omega. \end{cases} \tag{1}$$

Here $\{\gamma^0, \gamma^1\}$ are Dirac gamma matrices, m is the rest mass of the spin-1/2 particle, c is the speed of light in vacuum, $\hbar$ is Planck's constant, $\mathbf{n}$ is the normal to $\partial\Omega$, and $\alpha = \alpha^1 := \gamma^0\gamma^1$.

The initial data $\mathring{\psi}$ is assumed to be C^k, for a fixed integer $k \geq 0$, and compactly supported inside the interval Ω. The data can therefore be extended outside Ω to be identically zero. In the following, when speaking of $\mathring{\psi}$ we always have this extension in mind.

The boundary of the spacetime domain for ψ is the set of points $\{(t, -L), (t, L)\}$. So we have $\mathbf{n} = 1$ at (t, L), $\mathbf{n} = -1$ at $(t, -L)$. Choosing $\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\gamma^1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, we have $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Plugging these into the boundary condition in (1), which we call an *Absorbing Boundary Condition for the Dirac equation* (ABCD), we get $\psi_+(t, L) = 0$ and $\psi_-(t, -L) = 0$ for all $t \geq 0$. Now plugging these two boundary conditions into the Dirac equation, we get the additional boundary condition $m\psi_\mp(t, \pm L) = \mp i\partial_s\psi_\pm(t, \pm L)$. This gives us an equivalent IBVP for the Klein-Gordon equation obtained by iterating the Dirac operator in (1):

$$\begin{cases} \partial_t^2\psi_\pm - \partial_s^2\psi_\pm + m^2\psi_\pm = 0 \\ \psi_\pm(0, s) = \mathring{\psi}_\pm(s) \in C_c^k(\Omega),\ k \geq 0 \\ \partial_t\psi_\pm(0, s) = -im\mathring{\psi}_\mp(s) \pm \partial_s\mathring{\psi}_\pm(s) \\ \psi_\pm(t, \pm L) = 0 \\ m\psi_\mp(t, \pm L) \pm i\partial_s\psi_\pm(t, \pm L) = 0 \end{cases} \tag{2}$$

(We have set $c = \hbar = 1$.)

The following theorem shows that (2) has an explicit solution by splitting the IBVP

into its initial and boundary value problem parts and finding the solutions to each. Although this solution is in the form of an infinite series, we shall see that at any fixed time there are only finitely many nonzero terms in it, so that there are no convergence issues.

Theorem 1 *The IBVP in* (2) *has a unique solution that is as regular as its initial data and depends continuously on it.*

Proof In order to solve these equations we set $\psi_\pm = \Phi_\pm + \chi_\pm$, where $\Phi_\pm$, defined on $[0, \infty) \times (-\infty, \infty)$, are the solutions to

$$\begin{cases} \partial_t^2 \Phi_\pm - \partial_s^2 \Phi_\pm + m^2 \Phi_\pm = 0 \\ \Phi_\pm(0, s) = \mathring{\psi}_\pm(s) \\ \partial_t \Phi_\pm(0, s) = -im\mathring{\psi}_\mp(s) \pm \partial_s \mathring{\psi}_\pm(s), \end{cases} \tag{3}$$

and where $\chi_\pm$ are functions defined on $[0, \infty) \times (-\infty, L)$, resp. $[0, \infty) \times (-L, \infty)$ that satisfy

$$\begin{cases} \partial_t^2 \chi_\pm - \partial_s^2 \chi_\pm + m^2 \chi_\pm = 0 \\ \chi_\pm(0, s) = 0 \\ \partial_t \chi_\pm(0, s) = 0 \\ \chi_\pm(t, \pm L) = f_\pm(t) \\ m\chi_\mp(t, \pm L) \pm i\partial_s \chi_\pm(t, \pm L) = i\partial_t f_\pm(t), \end{cases} \tag{4}$$

where $f_\pm(t) := -\Phi_\pm(t, \pm L)$.
(3) is the Cauchy problem for the one-dimensional Klein-Gordon equation, in older literature called the (lossless) *telegraph* equation (e.g. [1], p.544), whose solution is well-known (see e.g. [17]):

$$\begin{aligned} \Phi_\pm(t, s) = {} & \frac{1}{2}[\mathring{\psi}_\pm(s - t) + \mathring{\psi}_\pm(s + t)] - \frac{tm}{2} \int_{s-t}^{s+t} \frac{J_1(m\sqrt{t^2 - (s - \sigma)^2})}{\sqrt{t^2 - (s - \sigma)^2}} \mathring{\psi}_\pm(\sigma) d\sigma \\ & - \frac{1}{2} \int_{s-t}^{s+t} J_0(m\sqrt{t^2 - (s - \sigma)^2})(im\mathring{\psi}_\mp(\sigma) \mp \partial_\sigma \mathring{\psi}_\pm(\sigma)) d\sigma, \end{aligned} \tag{5}$$

where J_n is the Bessel function of order n. Using integration by parts, the above can be rewritten as follows:

$$\Phi_\pm(t, s) = \mathring{\psi}_\pm(s \pm t) - \frac{m}{2} \int_{s-t}^{s+t} Z_1(t, \pm(s - \sigma))\mathring{\psi}_\pm(\sigma) d\sigma - \frac{im}{2} \int_{s-t}^{s+t} Z_0(t, s - \sigma)\mathring{\psi}_\mp(\sigma) d\sigma, \tag{6}$$

where we define, for $\nu \geq 0$,

$$Z_\nu(t, s) := J_\nu(m\sqrt{t^2 - s^2}) \left(\frac{t - s}{t + s}\right)^{\nu/2}. \tag{7}$$

On the other hand, we can solve (4) using Laplace transform methods, as follows: We first find the general solution to the first three equations of (4):

$$\begin{aligned}\tilde{\chi}_+(p,s) &= c_1(p)e^{ks} + c_2(p)e^{-ks}\\ \tilde{\chi}_-(p,s) &= c_3(p)e^{ks} + c_4(p)e^{-ks},\end{aligned} \tag{8}$$

where

$$k := \sqrt{p^2+m^2}, \qquad \mathrm{Re}(p) > 0, \tag{9}$$

and tilde denotes the Laplace transform, i.e. $\tilde{f}(p) = \int_0^\infty f(t)e^{-pt}dt$. Applying this to the boundary conditions in (4), we obtain

$$\begin{aligned}\tilde{f}_+(p) &= c_1(p)e^{kL} + c_2(p)e^{-kL}\\ \tilde{f}_-(p) &= c_3(p)e^{-kL} + c_4(p)e^{kL}\\ ip\tilde{f}_+(p) &= m(c_3(p)e^{kL} + c_4(p)e^{-kL}) + i(kc_1(p)e^{kL} - kc_2(p)e^{-kL})\\ ip\tilde{f}_-(p) &= m(c_1(p)e^{-kL} + c_2(p)e^{kL}) + i(-kc_3(p)e^{-kL} + kc_4(p)e^{kL}).\end{aligned} \tag{10}$$

After solving the system (10) for c_1, c_2, c_3, c_4 and plugging back into (8), we arrive at

$$\begin{aligned}\tilde{\chi}_+(p,s) &= \frac{\tilde{f}_+(m^2e^{6kL} + 2k^2e^{2kL} - m^2e^{2kL} + 2kpe^{2kL}) + i\tilde{f}_-(kme^{4kL} - mpe^{4kL} + km + mp)}{D}e^{k(L+s)}\\ &+ \frac{\tilde{f}_+(2k^2e^{4kL} - 2kpe^{4kL} - m^2e^{4kL} + m^2) + i\tilde{f}_-(mpe^{6kL} - kme^{6kL} - mpe^{2kL} - kme^{2kL})}{D}e^{k(L-s)}\\ \tilde{\chi}_-(p,s) &= \frac{i\tilde{f}_+(mpe^{6kL} - kme^{6kL} - mpe^{2kL} - kme^{2kL}) + \tilde{f}_-(2k^2e^{4kL} - 2kpe^{4kL} - m^2e^{4kL} + m^2)}{D}e^{k(L+s)}\\ &+ \frac{i\tilde{f}_+(kme^{4kL} - mpe^{4kL} + km + mp) + \tilde{f}_-(m^2e^{6kL} + 2k^2e^{2kL} - m^2e^{2kL} + 2kpe^{2kL})}{D}e^{k(L-s)},\end{aligned}$$

where $D := m^2(e^{2kL} + i\frac{k+p}{m})(e^{2kL} - i\frac{k+p}{m})(e^{2kL} + i\frac{k-p}{m})(e^{2kL} - i\frac{k-p}{m})$.

After doing a partial fraction decomposition, the above can be rewritten as

$$\begin{aligned}\tilde{\chi}_\pm(p,s) = \frac{1}{2}\Bigg[&e^{-k(L\mp s)}\Big(\pm\frac{\tilde{f}_+(p) - \tilde{f}_-(p)}{1 + i\frac{k-p}{m}e^{-2kL}} + \frac{\tilde{f}_+(p) + \tilde{f}_-(p)}{1 - i\frac{k-p}{m}e^{-2kL}}\Big)\\ &+ i\frac{k-p}{m}e^{-k(L\pm s)}\Big(\pm\frac{\tilde{f}_+(p) - \tilde{f}_-(p)}{1 + i\frac{k-p}{m}e^{-2kL}} - \frac{\tilde{f}_+(p) + \tilde{f}_-(p)}{1 - i\frac{k-p}{m}e^{-2kL}}\Big)\Bigg].\end{aligned} \tag{11}$$

Noting that $\left|i\frac{k-p}{m}e^{-2kL}\right| < 1$, we can view (11) as the sum of four convergent geometric series, so that

$$\begin{aligned}\tilde{\chi}_\pm(p,s) = \frac{1}{2}\Bigg[&\pm\sum_{n=0}^{\infty}\Big(-i\frac{k-p}{m}\Big)^n e^{-k[(2n+1)L\mp s]}(\tilde{f}_+(p) - \tilde{f}_-(p)) + \sum_{n=0}^{\infty}\Big(i\frac{k-p}{m}\Big)^n e^{-k[(2n+1)L\mp s]}(\tilde{f}_+(p) + \tilde{f}_-(p))\\ &\mp\sum_{n=0}^{\infty}\Big(-i\frac{k-p}{m}\Big)^{n+1} e^{-k[(2n+1)L\pm s]}(\tilde{f}_+(p) - \tilde{f}_-(p)) - \sum_{n=0}^{\infty}\Big(i\frac{k-p}{m}\Big)^{n+1} e^{-k[(2n+1)L\pm s]}(\tilde{f}_+(p) + \tilde{f}_-(p))\Bigg]\end{aligned} \tag{12}$$

From a table of inverse Laplace transforms (e.g. [10], formula 14.52) we find

$$\mathcal{L}^{-1}[(\frac{(k-p)^\nu}{k})e^{-kx}](\tau,x) = m^\nu(\frac{\tau-x}{\tau+x})^{\frac{1}{2}\nu}J_\nu(m\sqrt{\tau^2-x^2})H(\tau-x) = m^\nu Z_\nu(\tau,x)H(\tau-x)$$

where H is the Heaviside function $H(t) = 1$ for $t > 0$ and 0 otherwise.
We therefore have the following explicit solution for $\chi_\pm$:

$$\begin{aligned}\chi_\pm(t,s) = \frac{1}{2}\sum_{n=0}^{\infty} i^n\Big[H(t-(2n+1)L\pm s)\Big((-1)^n\frac{d}{ds}\int_{(2n+1)L\mp s}^{t} d\xi\ F_-(t-\xi)Z_n(\xi,\pm((2n+1)L-s))\\ \pm\frac{d}{ds}\int_{(2n+1)L\mp s}^{t} d\xi\ F_+(t-\xi)Z_n(\xi,\pm((2n+1)L-s))\Big)\\ + iH(t-(2n+1)L\mp s)\Big((-1)^{n+1}\frac{d}{ds}\int_{(2n+1)L\pm s}^{t} d\xi\ F_-(t-\xi)Z_{n+1}(\xi,\pm((2n+1)L-s))\\ \pm\frac{d}{ds}\int_{(2n+1)L\pm s}^{t} d\xi\ F_+(t-\xi)Z_{n+1}(\xi,\pm((2n+1)L-s))\Big)\Big],\end{aligned} \tag{13}$$

where

$$F_\pm := f_+ \pm f_-.$$

Carrying out the differentiations in (13), we have

$$\begin{aligned}\chi_\sigma(t,s) = H(t-L+\sigma s)f_\sigma(t-L+\sigma s)\\ +\sum_{n=0}^{\infty} i^n\Big(H(t-(2n+1)L+\sigma s)\int_{(2n+1)L-\sigma s}^{t} f_{(-1)^n\sigma}(t-\xi)R_n(\xi,(2n+1)L-\sigma s)d\xi\\ - iH(t-(2n+1)L-\sigma s)\int_{(2n+1)L+\sigma s}^{t} f_{(-1)^{n+1}\sigma}(t-\xi)R_{n+1}(\xi,(2n+1)L+\sigma s)d\xi\Big),\end{aligned} \tag{14}$$

where $\sigma \in \{+,-\}$, and for $k \geq 0$,

$$R_k(\xi,\eta) := -m\eta(\xi-\eta)^k\frac{J_{k+1}(m\sqrt{\xi^2-\eta^2})}{(\xi^2-\eta^2)^{\frac{k+1}{2}}} + k(\xi-\eta)^{k-1}\frac{J_k(m\sqrt{\xi^2-\eta^2})}{(\xi^2-\eta^2)^{\frac{k}{2}}}.$$

Adding $\chi_\pm$ to $\Phi_\pm$, we arrive at the solution for $\psi_\pm$. This solution can clearly be put in terms of the initial data, as shown for example by the $n = 0$ term of χ_σ, which we denote by $\chi_{\sigma 0}$:

$$\begin{aligned}\chi_{\sigma 0}(t,s) = -H(t-L+\sigma s)\Big[\Phi_\sigma(t-L+\sigma s,\sigma L) + \int_{L-\sigma s}^{t} R_0(\xi,L-\sigma s)\Big(\mathring{\psi}_\sigma(\sigma(L+t-\xi))\\ -\frac{m}{2}\int_{\sigma L-t+\xi}^{\sigma L+t-\xi} Z_1(t-\xi,\sigma L-\varsigma)\mathring{\psi}_\sigma(\varsigma)d\varsigma - \frac{im}{2}\int_{\sigma L-t+\xi}^{\sigma L+t-\xi} Z_0(t-\xi,\sigma L-\varsigma)\mathring{\psi}_{-\sigma}(\varsigma)d\varsigma\Big)d\xi\Big]\\ + iH(t-L-\sigma s)\int_{L+\sigma s}^{t} R_1(\xi,L+\sigma s)\Big(\mathring{\psi}_{-\sigma}(-\sigma(L+t-\xi))\\ -\frac{m}{2}\int_{-\sigma L-t+\xi}^{-\sigma L+t-\xi} Z_1(t-\xi,-\sigma L-\varsigma)\mathring{\psi}_{-\sigma}(\varsigma)d\varsigma - \frac{im}{2}\int_{-\sigma L-t+\xi}^{-\sigma L+t-\xi} Z_0(t-\xi,\sigma L+\varsigma)\mathring{\psi}_\sigma(\varsigma)d\varsigma\Big)d\xi\end{aligned} \tag{15}$$

Furthermore, the Heaviside functions appearing in (14) show that χ_n, the n-th term in the summation, is supported in $\bigcup_{k=0}^{n}\mathcal{R}_k$, with regions $\mathcal{R}_k$ shown in Fig. 1.

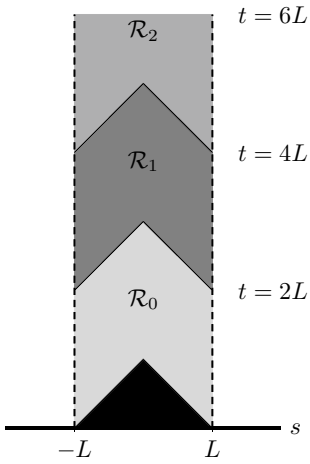


Fig. 1 Support of χ_n

This shows that for any fixed $t > 0$ there are only finitely many non-zero terms in (14), so convergence is never an issue.

Finally, the behavior of Bessel functions for small values of their argument, $J_n(x) = O(x^n)$ as $x \to 0$, imply that, despite appearances, the solution kernels Z_k and R_k are smooth and bounded on compact domains, which ensures that the solution is as regular as the data and depends continuously on it. □

3 Arrival Time of Bohmian Trajectories

Let $\psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}$ be the wave function of a spin-1/2 particle in $\mathbb{R}^{1,1}$. The *Dirac current*

$$j^\mu := \overline{\psi}\gamma^\mu\psi, \qquad \overline{\psi} := \psi^\dagger\gamma^0 \tag{16}$$

is the simplest Lorentz vector that can be constructed from the Dirac bispinor ψ. When ψ satisfies the Dirac equation $-i\gamma^\mu\partial_\mu\psi + m\psi = 0$, it follows that the vectorfield j is divergence free, i.e.

$$\partial_\mu j^\mu = 0. \tag{17}$$

Setting $\varrho(t, s) := j^0 = \psi^\dagger\psi$ and $J(t, s) := j^1 = \psi^\dagger\alpha^1\psi$, the above can be written as $\partial_t\varrho + \partial_s J = 0$, which has the form of an equation of continuity for a density ϱ. The quantity $v(t, s) = J/\varrho$ is thus naturally a velocity field defined on the 1-particle configuration space.

Let $Q(t)$ denote the actual position of the particle at time t. According to the principles of Bohmian Mechanics (e.g. [6], Chap. 9) the guiding equation for the motion of the particle is

$$\frac{dQ(t)}{dt} = \frac{J(t, Q(t))}{\varrho(t, Q(t))}. \tag{18}$$

The above ODE can be uniquely solved given initial data $Q_0 = Q(0) \in \mathbb{R}$ which is assumed to be distributed randomly according to the initial density $\varrho_0 := \mathring{\psi}^\dagger \mathring{\psi}$. The *arrival time* of the trajectory $s = Q(t)$ at the boundary $\partial\Omega$ of the domain is simply

$$T = \inf\{t > 0 \mid Q(t) \in \partial\Omega\} \tag{19}$$

(Recall that the infimum of the empty set is by definition $+\infty$.)

Let j^μ be as in (16). Let $Z = (T, \mathbf{X})$ where $\mathbf{X}$ is the location on $\partial\Omega$ where a particle gets first detected at time T, meaning there is a trajectory $Q(t)$ with $Q(0) \in \Omega$ that at time $t = T$ reaches the boundary for the first time, and $Q(T) = \mathbf{X}$. (If $T = \infty$, we write $Z = \infty$.) Z is a random variable defined on Ω, and according to [15] the distribution μ of Z satisfies

$$\mu(t_1 \le T < t_2, \boldsymbol{X} \in \partial\Omega) = \int_{t_1}^{t_2} dt \int_{\partial\Omega} \boldsymbol{n}(\boldsymbol{x}) \cdot \boldsymbol{j}^{\psi_t}(\boldsymbol{x}) d\sigma = \int_{t_1}^{t_2} \Big(j^1(t, L) - j^1(t, -L) \Big) dt. \tag{20}$$

It was shown in [15] that when the absorbing boundary conditions are satisfied, the integrand in (20) is a probability density function on $\mathbb{R}^+$. Moreover, since in this case the boundary is made up of just two points, call them A for Alice and B for Bob, the distribution density of arrival times at Alice's detector on the left and for Bob's detector on the right are

$$\varrho_A(t) = -j^1(t, -L) = |\psi_+(t, -L)|^2, \qquad \varrho_B(t) = j^1(t, L) = |\psi_-(t, L)|^2. \tag{21}$$

Remark 1 We note that these are not normalized probability density functions, since, as was shown in [15], a simple application of the Divergence Theorem implies that

$$\int_{-L}^{L} |\mathring{\psi}(s)|^2 ds - \int_{-L}^{L} |\psi(t, s)|^2 ds = \int_0^t \Big(j^1(T, L) - j^1(T, -L) \Big) dT = \int_0^t \varrho_B(T) + \varrho_A(T)\, dT. \tag{22}$$

Moreover, since $\mathring{\psi}$ is assumed to be compactly supported in $[-L, L]$, by the Born Rule the left hand side of (22) is

$$1 - \mathrm{Prob}(T = t, X \in \Omega).$$

Therefore, letting $t \to \infty$ we can only conclude that the *sum* of the total integrals of ϱ_A and ϱ_B is *at most* equal to one.

4 Two-Body Problem

Let us now imagine that the particle source in the middle of the interval emits *pairs* of particles. The two body multi-time wave function $\psi = (\psi_{--}, \psi_{-+}, \psi_{+-}, \psi_{++})$ for this problem has four components, each one of which is a function of two time and two space variables: $\psi_{\sigma_1\sigma_2} = \psi_{\sigma_1\sigma_2}(x_1, x_2)$ with $x_i^0 = t_i$, $x_i^1 = s_i$, $i = 1, 2$. It satisfies the multi-time system of Dirac equations

$$i\gamma_j^\mu \partial_{x_j^\mu}\psi = m\psi \qquad j = 1, 2 \tag{23}$$

where $\gamma_1^\mu := \gamma^\mu \otimes \mathbf{1}$ and $\gamma_2^\mu := \mathbf{1} \otimes \gamma^\mu$. The above equations are amended by the boundary conditions

$$\psi_{\pm\sigma_2}(t, \pm L, t, s) = 0, \qquad \psi_{\sigma_1\pm}(t, s, t, \pm L) = 0, \qquad \forall\sigma_1, \sigma_2 \in \{+, -\}, \forall t \geq 0, \forall s \in [-L, L] \tag{24}$$

The two-body tensor current j is defined as $j^{\mu\nu} = \overline{\psi}(\gamma^\mu \otimes \gamma^\nu)\psi$ where $\overline{\psi} := \psi^\dagger\gamma^0 \otimes \gamma^0$. It follows that

$$j^{00} = |\psi_{--}|^2 + |\psi_{-+}|^2 + |\psi_{+-}|^2 + |\psi_{++}|^2 \tag{25}$$

$$j^{01} = |\psi_{--}|^2 - |\psi_{-+}|^2 + |\psi_{+-}|^2 - |\psi_{++}|^2 \tag{26}$$

$$j^{10} = |\psi_{--}|^2 + |\psi_{-+}|^2 - |\psi_{+-}|^2 - |\psi_{++}|^2 \tag{27}$$

When ψ satisfies the multi-time Dirac system (23) the tensor current $j^{\mu\nu}$ satisfies a pair of conservation laws:

$$\partial_{t_1} j^{00} + \partial_{s_1} j^{10} = 0, \qquad \partial_{t_2} j^{00} + \partial_{s_2} j^{01} = 0. \tag{28}$$

Suppose that there is a particle detector at each endpoint of the interval Ω. The detector at left endpoint is controlled by Alice, and the detector at the right endpoint is controlled by Bob. Alice and Bob can change the position of their detectors independently of one another. Let $(x^\mu) = (t, s)$ denote a coordinate frame on physical spacetime with respect to which Alice and Bob are both stationary, and let Σ_t be the foliation of spacetime by constant t-slices. According to the Hypersurface Bohm-Dirac theory [7], with respect to Σ_t the guiding equation for each of the two particles is

$$\frac{dQ_1}{dt} = \frac{j^{10}}{j^{00}}(t, Q_1(t), t, Q_2(t)), \qquad \frac{dQ_2}{dt} = \frac{j^{01}}{j^{00}}(t, Q_1(t), t, Q_2(t)). \tag{29}$$

Suppose the source emits a pair of particles at $t = 0$. Given a trajectory of the two-body system $(Q_1(t), Q_2(t))$, its first arrival time chez Alice is the earliest time that either of the two particles reaches Alice's detector, and similarly for Bob. In other words

$$T_A := \min_{k=1,2} \inf\{t > 0 \mid Q_k(t) \leq -L\}, \qquad T_B := \min_{k=1,2} \inf\{t > 0 \mid Q_k(t) \geq L\}. \tag{30}$$

T_A and T_B are random variables on Ω (due to their dependence on the initial positions of the trajectories) describing the first of the two arrival times at Alice's detector, resp. Bob's. Let μ_A and μ_B denote their corresponding (un-normalized) probability densities. We have

Proposition 1

$$\begin{aligned}\mu_A(t) = &\int_{-L}^{L} \left(\sum_{\sigma_2} |\psi_{+\sigma_2}(t, -L, t, s')|^2 + \sum_{\sigma_1} |\psi_{\sigma_1 +}(t, s', t, -L)|^2 \right) ds' \\ &+ \int_0^t dt' \left(|\psi_{-+}(t', L, t, -L)|^2 + |\psi_{+-}(t, -L, t', L)|^2 \right) \end{aligned} \tag{31}$$

$$\begin{aligned}\mu_B(t) = &\int_{-L}^{L} \left(\sum_{\sigma_2} |\psi_{-\sigma_2}(t, L, t, s')|^2 + \sum_{\sigma_1} |\psi_{\sigma_1 -}(t, s', t, L)|^2 \right) ds' \\ &+ \int_0^t dt' \left(|\psi_{-+}(t, L, t', -L)|^2 + |\psi_{+-}(t', -L, t, L)|^2 \right). \end{aligned} \tag{32}$$

Proof Let T^f be the first time (according to Σ_t foliation) at which a particle is registered by *any* of the detectors at the boundary, let $Z^f = (T^f, S^f)$ with $S^f \in \{L, -L\}$ be the corresponding detection event, and let $I^f \in \{1, 2\}$ be the label of the registered particle. The proposed rule in [15], specialized to this particularly simple case, asserts that the joint probability distribution of I^f and Z^f is

$$\text{Prob}\left(I^f = 1, (t < T^f < t + dt, S^f = \pm L)\right) = ds\, dt\, \delta(s \mp L) \int_{-L}^{L} \pm j^{10}(t, s, t, s_2) ds_2 \tag{33}$$

$$\text{Prob}\left(I^f = 2, (t < T^f < t + dt, S^f = \pm L)\right) = ds\, dt\, \delta(s \mp L) \int_{-L}^{L} \pm j^{01}(t, s_1, t, s) ds_1. \tag{34}$$

Since these events are disjoint, it follows that the probability density for the arrival time of either particle at Alice *before the other one arrives at Bob* is

$$\mu_A^f(t) = \int_{-L}^{L} -j^{10}(t, -L, t, s') - j^{01}(t, s', t, -L)\, ds, \tag{35}$$

and similarly for Bob. We note that when the boundary conditions (24) are satisfied, the above becomes

$$\mu_A^f(t) = \int_{-L}^{L} \sum_{\sigma_2} |\psi_{+\sigma_2}(t, -L, t, s')|^2 + \sum_{\sigma_1} |\psi_{\sigma_1 +}(t, s', t, -L)|^2\, ds' \geq 0. \tag{36}$$

To account for all the particle arrivals at a given detector, we need to include the possibility of the particle arriving at that detector being the *second* one to arrive at any detector, e.g. one of the particles arrives at Bob's detector at a time $t' < t$ before

the other one arrives at Alice's at time t. In such a case, the Absorbing Boundary Rule proposed in [15, 16] stipulates that the detection of the particle by Bob causes the two-body wave function to undergo collapse, and the evolution of the remaining particles will henceforth be governed by the collapsed (i.e. the *conditional*) wave function. Let $\phi^{jB}(t', \cdot)$ denote the conditional wave function that results from Bob detecting particle labeled $j \in \{1, 2\}$ at time t'. Then

$$\phi^{1B}_{\sigma_2}(t', s_2) = \frac{\psi_{-\sigma_2}(t', L, t', s_2)}{\sqrt{\int_{-L}^{L} ds' \sum_{\sigma_2} |\psi_{-\sigma_2}(t', L, t', s')|^2}}, \quad \phi^{2B}_{\sigma_1}(t', s_1) = \frac{\psi_{\sigma_1 -}(t', s_1, t', L)}{\sqrt{\int_{-L}^{L} ds' \sum_{\sigma_1} |\psi_{\sigma_1 -}(t', s_1, t', L)|^2}}. \quad (37)$$

(Note that the leftover spin components corresponding to the absorbed particle are zero thanks to the boundary conditions (24).)

Let T_A^j be the time of arrival of particle labeled j at Alice's detector, and let $\hat{j}$ denote the label of the other particle, i.e. $\hat{1} = 2$ and $\hat{2} = 1$. By the formula for conditional probabilities, we have

$$\begin{aligned}\text{Prob}(t < T_A < t + dt) &= \sum_{j=1}^{2} \Big(\text{Prob}(t < T_A^j < t + dt, I^f = j) + \text{Prob}(t < T_A^j < t + dt, I^f = \hat{j})\Big) \\ &= \mu_A^f(t) + \sum_{j=1}^{2} \int_0^t dt' \text{Density}(t' < T_B^{\hat{j}} < t' + dt', I^f = \hat{j}) \text{Prob}(t < T_A^j < t + dt \mid T_B^{\hat{j}} = t', I^f = \hat{j}).\end{aligned}$$

To calculate the conditional probability that shows up on the last line, we need to use the conditional wave function, viz. for $j = 2$,

$$\text{Prob}(t < T_A^2 < t + dt \mid T_B^1 = t', I^f = 1) = \sum_{j=1}^{2} |\phi_-^{1B}(t, -L)|^2 + |\phi_+^{1B}(t, -L)|^2 = \frac{|\psi_{-+}(t', L, t, -L)|^2}{\sum_{\sigma_2} \int_{-L}^{L} ds' |\psi_{-\sigma_2}(t', L, t', s')|^2},$$

and similarly for $j = 1$, while

$$\text{Density}(t' < T_B^1 < t' + dt', I^f = 1) = \sum_{\sigma_2} \int_{-L}^{L} ds' |\psi_{-\sigma_2}(t', L, t', s')|^2, \quad (38)$$

which cancels the denominator in the above, establishing the claim. □

5 Detection Versus Arrival Time

How well does the ABCD method capture the detection phenomena, that is to say, the interaction of the particles emitted by the source inside the domain with the (presumably macroscopic) detectors places on the boundary of the domain? If we imagine that the detector at the left endpoint of the interval Ω is controlled by Alice, and the one on the right by Bob, it should for example be the case that the distribution of arrival times at Alice is not immediately affected if Bob decides to move his detector by an appreciable amount, or if he simply switches his detector off (meaning

the boundary condition is not imposed on Bob's side.) For a single particle, this is easily obtained (see below) but for two particles it is not at all obvious, since moving one of the detectors or switching it off changes the two-body wave function, and therefore the individual trajectories of an entangled pair of particles change as soon as one of them enters the domain of influence of Bob's boundary point (since the velocity of each particle depends on the actual positions of both particles). Nevertheless, one expects the distribution of arrival times not to change faster than what is allowed by relativity.

Proposition 2 *For a single particle, and for all $\varepsilon \in (0, L/2)$, the distribution of arrival times at $s = -L$ is unaffected for $t < 2L - \varepsilon$ when the detector at $s = L$ is shifted to $s = L + \varepsilon$.*

Proof By (20), the density of arrival times at the boundary point $s = -L$ is $\varrho_A(t) = -j^1(t, -L) = |\psi_+(t, -L)|^2$. We next note that due to the translation invariance of the equations, the situation with Bob's detector shifted to $L + \varepsilon$ while keeping the particle source and Alice's detector fixed, is equivalent to shifting Bob's detector $\varepsilon/2$ units to the right while shifting both the source and Alice's detector $\varepsilon/2$ units to the left. This observation allows us to use the solution formulas developed in the above for Ω that was symmetric with respect to the origin, also to the case where that symmetry is broken due to the shift in Bob's detector.

After such a shift, therefore, the interval is $\Omega' = [-L', L']$, where $L' = L + \varepsilon/2$, and the initial data is $\mathring{\psi}'_\pm(s) = \mathring{\psi}_\pm(s + \frac{\varepsilon}{2})$. If we denote by ψ' and Φ' the solutions to (2) and (3) after the shift, from (15) we would then have that

$$\psi'_+(t, s) = \Phi'_+(t, s) - H(t - L' + s)X(t, s) + iH(t - L' - s)Y(t, s), \qquad (39)$$

where X and Y are certain integral operators acting on the initial data $\mathring{\psi}'$. In particular,

$$Y(t, s) = \sum_{\sigma\in\{+,-\}} \int_{L'+s}^{t} d\xi \int_{-L'-t+\xi}^{-L'+t-\xi} d\zeta \; K_\sigma(t, \xi, L' + s, L' + \zeta)\mathring{\psi}'_\sigma(\zeta)$$

for certain explicit Kernels $K_\pm$. Assuming $0 < t < 2L$, upon evaluating at $s = -L'$ the first Heaviside function in (39) is zero, therefore we are left with $\psi'_+(t, -L') = \Phi'_+(t, -L') + iY(t, -L')$, and we can check that changing the inner variable of integration to $\zeta' = \zeta + \frac{\varepsilon}{2}$ we obtain

$$Y(t, -L') = \sum_\sigma \int_0^t d\xi \int_{-L'-t+\xi}^{-L'+t-\xi} d\zeta K_\sigma(t, \xi, 0, L' + \zeta)\mathring{\psi}'_\sigma(\zeta) = \sum_\sigma \int_0^t d\xi \int_{-L-t+\xi}^{-L+t-\xi} d\zeta' K_\sigma(t, \xi, 0, L + \zeta')\mathring{\psi}_\sigma(\zeta')$$

which is manifestly independent of ε. From the formula (6) one can similarly conclude that $\Phi'_+(t, -L')$ is independent of ε (or alternatively, one can use the domain of dependence property that holds for solutions to (3).) It thus follows that $\psi'_+(t, -L')$ is independent of ε, and therefore so is $\varrho_A(t)$. □

As a simple corollary of the above, we note that if for the 2-body problem we start with a pure product initial state $\mathring{\psi}^1 \otimes \mathring{\psi}^2$, the solution would remain a pure product as well (the equations (23) are for the non-interacting case), and therefore the same reasoning as above applies to show that the distribution of first arrival times at Alice will not be affected by a change in the position of Bob's detector.

Next we can try to answer the same question for two particles whose wavefunction is in a *maximally entangled state*, i.e., the solution ψ to the two-body problem (23) with boundary conditions (24) whose initial data is

$$\psi(0, s_1, 0, s_2) = \frac{1}{\sqrt{2}} \left(\mathring{\psi}^1 \otimes \mathring{\psi}^2 + \mathring{\psi}^3 \otimes \mathring{\psi}^4 \right), \tag{40}$$

where $\mathring{\psi}^i \in C_c^0(\Omega)$ for $i = 1, \ldots, 4$ are four normalized mutually orthogonal 1-body wavefunctions. Without much loss of generality, we can take these to be

$$\begin{cases} \mathring{\psi}^1_- \equiv 0 \\ \mathring{\psi}^1_+(s) = f_{\mu,\alpha}(s), \end{cases} \quad \begin{cases} \mathring{\psi}^2_-(s) = f_{\mu,\alpha}(s) \\ \mathring{\psi}^2_+ \equiv 0, \end{cases} \quad \begin{cases} \mathring{\psi}^3_- \equiv 0 \\ \mathring{\psi}^3_+(s) = g_{\mu,\alpha}(s), \end{cases} \quad \begin{cases} \mathring{\psi}^4_-(s) = g_{\mu,\alpha}(s) \\ \mathring{\psi}^4_+ \equiv 0. \end{cases} \tag{41}$$

where $f_{\mu,\alpha}$ and $g_{\mu,\alpha}$ are continuous functions compactly supported in $(\mu - \alpha, \mu + \alpha)$, with the property that $\int f^2 ds = \int g^2 ds = 1$ and $\int fg ds = 0$, and we can assume

$$f_{\mu,\alpha}(s) = f_{0,\alpha}(s - \mu), \qquad g_{\mu,\alpha}(s) = g_{0,\alpha}(s - \mu).$$

For $i = 1, \ldots, 4$, let ψ^i be the solution to the one-body problem (1) with initial data $\mathring{\psi}^i$. Clearly, the solution to the 2-body system (23–24) is $\psi = \frac{1}{\sqrt{2}} \left(\psi^1 \otimes \psi^2 + \psi^3 \otimes \psi^4 \right)$.

Recall that for $t < 2L$ the only non-zero term in the series (4) is the one with $n = 0$. Noticing that $\psi^i(t, -L) = 0$ due to the boundary condition, from (31) the distribution of first arrival times at Alice when Bob's detector is shifted to the right by the amount ε, with $\alpha << \varepsilon < L/2$ and letting $L' := L + \frac{\varepsilon}{2}$, becomes

$$\begin{aligned} \mu_A^\varepsilon(t) := \frac{1}{2} \int_{-L'}^{L'} ds \Big(& \left| \psi^1_+(t, -L')\psi^2_-(t, s) + \psi^3_+(t, -L')\psi^4_-(t, s) \right|^2 + \left| \psi^1_+(t, -L')\psi^2_+(t, s) + \psi^3_+(t, -L')\psi^4_+(t, s) \right|^2 \\ & + \left| \psi^1_+(t, s)\psi^2_+(t, -L') + \psi^3_+(t, s)\psi^4_+(t, -L') \right|^2 + \left| \psi^1_-(t, s)\psi^2_+(t, -L') + \psi^3_-(t, s)\psi^4_+(t, -L') \right|^2 \Big) \\ & + \frac{1}{2} \int_0^t dt' \Big(\left| \psi^1_-(t', L')\psi^2_+(t, -L') + \psi^3_-(t', L')\psi^4_+(t, -L') \right|^2 + \left| \psi^1_+(t, -L')\psi^2_-(t', L') + \psi^3_+(t, -L')\psi^4_-(t', L') \right|^2 \Big). \end{aligned} \tag{42}$$

In the above, $\psi^j_\pm$ are computed using initial data profiles $f_{-\varepsilon/2,\alpha}$ and $g_{-\varepsilon/2,\alpha}$. The goal would be to show that for $t < 2L - \varepsilon$ the above expression does not depend on ε.

It is fairly easy to verify this when both particles are *massless*, i.e. if we set $m = 0$. This is because in that case the initial data is just transported along characteristics, so that the boundary condition has no effect, and the integrals in (42) are easily seen to be independent of ε.

For massive particles, though, a more detailed analysis of the terms in (42) becomes necessary. We plan to carry this out in future work.

6 Summary and Outlook

In this paper we have demonstrated that there exists an explicit, unique solution to the initial-boundary-value problem for the 1-body Dirac equation in an interval in one-dimensional space, with absorbing boundary conditions as proposed in [15] modeling the presence of a detector at each endpoint, and that the solution inherits the regularity of the initial data and depends continuously on it.

We then studied the two-body problem corresponding to a pair of particles emanating from a source inside the interval, and obtained explicit formulas for the probability density of first arrival times at detectors situated at each end point of the interval. We posed the question whether it is possible to affect the distribution of arrival times at one detector by shifting the other detector, and to do so faster than allowed by relativity. We showed that this is not possible if the initial data is a pure product state, or if the particles are massless. In future work we plan to show that even for an entangled pair of massive particles this is still not possible, so that superluminal signaling is generally ruled out in this model.

Acknowledgements We would like to thank Shelly Goldstein, Roderich Tumulka, Markus Nöth, and Siddhant Das for helpful discussions. We especially thank Roderich Tumulka for pointing out an error in an earlier draft of our paper, and for suggesting to us the content of Proposition 1 and the outline of its proof.

References

1. R. Courant, D. Hilbert, *Methods of Mathematical Physics, Volume II: Partial Differential Equations* (John Wiley & Sons, 1962)
2. S. Das, D. Dürr, Arrival time distributions of spin-1/2 particles. Sci. Rep. **9**(1), 1–8 (2019)
3. S. Das, M. Nöth, Times of arrival and gauge invariance. Proc. R. Soc. A. **477**, (2021)
4. S. Das, M. Nöth, D. Dürr, Exotic Bohmian arrival times of spin-1/2 particles: an analytical treatment. Phys. Rev. A. **99**(5), 052124 (2019)
5. S. Das, W. Struyve, Questioning the adequacy of certain quantum arrival-time distributions. Phys. Rev. A. **104**, 042214
6. D. Dürr, S. Goldstein, N. Zanghì, *Quantum physics without quantum philosophy* (Springer Science & Business Media, 2013)
7. D. Dürr, S. Goldstein, K. Münch-Berndl, N. Zanghì, Hypersurface Bohm-Dirac models. Phys. Rev. A. **60**, 2729–2736
8. Ch. Kurtsiefer, J. Mlynek, A 2-dimensional detector with high spatial and temporal resolution for metastable rare gas atoms. Appl. Phys. B. **64**(1), 85–90 (1996)
9. Ch. Kurtsiefer, T. Pfau, J. Mlynek, Measurement of the Wigner function of an ensemble of Helium atoms. Nat. **386**(6621), 150–153 (1997)

10. F. Oberhettinger, L. Badii, *Tables of Laplace Transforms* (Springer-Verlag, Berlin Heidelberg, 1973)
11. T. Pfau, Ch. Kurtsiefer, Partial reconstruction of the motional Wigner function of an ensemble of Helium atoms. J. Mod. Opt. **44**(11–12), 2551–2564 (1997)
12. S. Teufel, R. Tumulka, Existence of Schrödinger evolution with absorbing boundary condition (2019). arXiv:1912.12057 [math-ph]
13. R. Tumulka, POVM (positive operator-valued measure), in *Compendium of Quantum Physics* (Springer, 2009), pp. 480–484
14. R. Tumulka, Distribution of the time at which an ideal detector clicks. Ann. Phys. **442**, 168910 (2022)
15. R. Tumulka, Detection time distribution for Dirac particles (2016). arXiv:1601.04571 [quant-ph]
16. R. Tumulka, Detection-time distribution for several quantum particles. Phys. Rev. A. **106**, 042220 (2022)
17. A.N. Tychonov, A.A. Samarski, *Partial Differential Equations of Mathematical Physics, Volume 1* (Holden-Day, 1964)
18. N. Vona, G. Hinrichs, D. Dürr, What does one measure when one measures the arrival time of a quantum particle?. Phys. Rev. Lett. **111**, 220404
19. R. Werner, Arrival time observables in quantum mechanics. Annales de l'Institute Henri Poincaré, Section A **47**, 429–449 (1987)

The Bohmian Solution to the Problem of Time

Ward Struyve

Abstract In canonical quantum gravity the wave function of the universe is static, leading to the so-called problem of time. We summarize here how Bohmian mechanics solves this problem.

1 Introduction

Canonical quantum gravity is an approach to quantum gravity which is obtained by applying the usual quantization techniques to the classical theory of gravity, general relativity [1]. These techniques have been successfully applied before to classical Yang-Mills theories to construct quantum theories describing the fundamental interactions other than gravity, leading to the Standard Model of elementary particle physics, and the hope is that they lead to similar success in the case of gravity. As usual, the quantization proceeds by bringing general relativity into Hamiltonian form and by replacing classical variables by operators acting on wave functions. This results in a set of wave equations for the wave function, one of which is known as the Wheeler-DeWitt equation. There is a host of problems with these wave equations, of both technical and conceptual nature. On the technical level, the main problem is that the wave equations are merely defined formally. Since the theory is non-renormalizable (unlike Yang-Mills theories), this problem cannot be bypassed by the usual perturbation methods. On the conceptual level, there is the measurement problem, which carries over from ordinary quantum theory, and which is especially severe in this context since the aim is to apply the theory to the whole universe, so

This paper is dedicated to Detlef Dürr with whom I had the pleasure to discuss quantum gravity and the problem of time on many occasions. For Detlef conceptual clarity was always quintessential to get a firm grip on this problem.

W. Struyve (✉)
Instituut voor Theoretische Fysica, Katholieke Universiteit Leuven, Leuven, Belgium

Centrum voor Logica en Filosofie van de Wetenschappen, Katholieke Universiteit Leuven, Leuven, Belgium
e-mail: ward.struyve@kuleuven.be

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_15

that there are no outside observers performing measurements. Another immediate problem is the problem of time. The wave equations entail a static wave function, which hence seems unable to describe our changing universe. Various ideas have been proposed on how to overcome this problem [1–4], without any consensus about the right approach. A recurring issue with many of these proposals is the lack of ontological clarity. What is the ontology in quantum gravity? What is the role of the wave function? Is it part of the ontology? Or merely a statistical representation of some sort? Does space-time itself exists? Or is it emergent? What are the objects in space-time?[1] A proper resolution of this problem requires clarity about these questions.

In Bohmian mechanics the ontology is clear. Non-relativistic Bohmian mechanics is about point-particles in Galilean space-time moving with a velocity that depends on the wave function which satisfies the usual Schrödinger equation [5–7]. In Bohmian quantum gravity [8–10], space-time itself is dynamical, like in general relativity. As in the classical Hamiltonian formulation, it can be written as a geometrodynamics, i.e., as a dynamics of a spatial 3-metric, whose time evolution determines a space-time metric. This time evolution, together with that of the matter degrees of freedom (particles or fields), is determined by the wave function which satisfies the wave equations of canonical quantum gravity. Even though the wave function is static, the spatial 3-metric and the matter degrees of freedom generically change over time. As such, the problem of time is immediately solved.

Before turning to Bohmian quantum gravity, it is instructive to consider the quantization of the non-relativistic particle. By putting the dynamics in parameterized form, it shares many features with the gravitational case, see e.g. [11–13]. We will give some details of the classical Hamiltonian formulation and how quantization leads to a stationary wave function. This also allows us to illustrate that classically there is no problem of time, contrary to what is sometimes claimed, see e.g. [14] and rebuttals [15, 16].

2 Non-relativistic Particle

Consider a single particle in Galilean space-time, whose possible trajectories $\mathbf{x}(t)$ satisfy the Newtonian equation

$$m\ddot{\mathbf{x}}(t) = -\nabla V(\mathbf{x}(t)), \tag{1}$$

with $V(\mathbf{x})$ a potential. The dynamics can be expressed in other ways, for example in the form of an action principle, with action $S = \int dt L$ and Lagrangian

$$L = \frac{m}{2}\dot{\mathbf{x}}^2 - V(\mathbf{x}). \tag{2}$$

[1] The wave function itself is not an object in space-time, but in some Hilbert space.

Extrema of the action satisfy (1). From the Lagrangian, also the Hamilton formulation can be obtained. With momentum $\mathbf{p} = \partial L/\partial\dot{\mathbf{x}} = m\dot{\mathbf{x}}$, the Hamiltonian is

$$H = \frac{1}{2m}\mathbf{p}^2 + V \tag{3}$$

and Hamilton's equations read

$$\dot{\mathbf{x}} = \frac{\mathbf{p}}{m}, \qquad \dot{\mathbf{p}} = -\nabla V. \tag{4}$$

Of course, by casting the dynamics this way, the ontology has not changed. The theory is still about a particle moving in space-time. It is not about, say, a point moving in phase space.

Another way to formulate the dynamics is by considering a different parameterization of the trajectory. Rather than parameterizing it by the time t, a parameter s can be introduced so that the trajectories are represented by $(t(s), \mathbf{x}(s))$. Denoting the derivatives with respect to s by primes, the dynamics reads

$$m\frac{1}{t'(s)}\left(\frac{1}{t'(s)}\mathbf{x}'(s)\right)' = -\nabla V(\mathbf{x}(s)). \tag{5}$$

This dynamics has a reparameterization symmetry. Namely, for any solution $(t(s), \mathbf{x}(s))$, also $(t(\tilde{s}(s)), \mathbf{x}(\tilde{s}(s)))$ will be a solution, for any different parameterization $\tilde{s}(s)$. But any two such solutions represent the same physical situation. Namely, they represent the same curve in space-time. The parameterization is just a part of the mathematical representation, without any physical significance. It is an instance of a gauge symmetry.

This symmetry is connected with indeterminism. Specifying $\mathbf{x}(t)$ and $\dot{\mathbf{x}}(t)$ at a particular time determines a unique solution to (1). On the other hand, specifying $(t(s), \mathbf{x}(s))$ at a certain value for s does not yield a unique solution to (5), which is an immediate consequence of the reparameterization symmetry. For a Lagrangian theory, indeterminism is often seen as the defining characteristic of gauge [17, 18].

Also the parameterized dynamics can be formulated in Hamiltonian form. Starting from the Lagrangian

$$\bar{L} = t'\left(\frac{m}{2}\frac{\mathbf{x}'^2}{t'^2} - V(\mathbf{x})\right), \tag{6}$$

the conjugate momenta are

$$\mathbf{p} = \frac{\partial\bar{L}}{\partial\mathbf{x}'} = m\frac{\mathbf{x}'}{t'}, \qquad p_t = \frac{\partial\bar{L}}{\partial t'} = -\frac{m}{2}\frac{\mathbf{x}'^2}{t'^2} - V(\mathbf{x}). \tag{7}$$

These momenta imply the constraint

$$C = p_t + \frac{1}{2m}\mathbf{p}^2 + V = 0, \tag{8}$$

which implies that the relations (7) cannot be inverted to yield the velocities in terms of the momenta. This means we have to resort to the formalism of constrained Hamiltonian dynamics [17, 18]. The canonical Hamiltonian vanishes,

$$\bar{H} = p_t t' + \mathbf{p}\mathbf{x}' - \bar{L} = 0, \tag{9}$$

but Hamilton's equations are derived from the total Hamiltonian

$$H_T = \bar{H} + NC = NC, \tag{10}$$

where N is an arbitrary (non-zero) function of the phase space variables. Together with the constraint $C = 0$, Hamilton's equation are

$$t' = N, \qquad p_t' = 0, \qquad \mathbf{x}' = N\frac{\mathbf{p}}{m}, \qquad \mathbf{p}' = -N\nabla V. \tag{11}$$

Because of the arbitrariness of N, the parameterization invariance has become explicit. Again, the dynamics was recast in a different way, without affecting the possible trajectories in space-time.

Another formulation of interest is the one in terms of the reduced phase-space [17, 18]. The reduced phase-space is obtained by eliminating the gauge variables. The gauge variables are identified as those variables that evolve completely freely. They are contrasted with what are sometimes called the *observables* [18] or *true degrees of freedom* [17], which evolve deterministically. In the present case, this means that the latter have zero Poisson brackets with the constraint and hence also that these variables must be static. Does this mean there is no change, as some have argued? This will depend on the ontological significance one attaches to those variables. Let us first make this formulation concrete. In general, the reduced phase space is hard to find, but in the case of $V = 0$, it is obtained by performing the following canonical transformation

$$\tilde{\mathbf{x}} = \mathbf{x} - \frac{1}{m}\mathbf{p}t, \qquad \tilde{t} = t, \qquad \tilde{\mathbf{p}} = \mathbf{p}, \qquad \tilde{p}_{\tilde{t}} = p_t + \frac{1}{2m}\mathbf{p}^2. \tag{12}$$

Along a trajectory $(t(s), \mathbf{x}(s))$, $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{p}}$ correspond to the position and momentum at $t = 0$. In terms of the new variables, the constraint reads $\tilde{p}_{\tilde{t}} = 0$ and the Hamiltonian is

$$\tilde{H} = N\tilde{p}_{\tilde{t}}. \tag{13}$$

The equations of motion are

$$\tilde{t}' = N, \qquad \tilde{p}_{\tilde{t}} = 0, \qquad \tilde{\mathbf{x}}' = 0, \qquad \tilde{\mathbf{p}}' = 0. \tag{14}$$

So the dynamics of $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{p}}$ decouples from that of $\tilde{t}$ and $\tilde{p}_{\tilde{t}}$. Moreover, because of the arbitrariness of N, the dynamics of $\tilde{t}$ is completely free and is identified as the gauge degree of freedom. By dropping the variables $\tilde{t}$ and $\tilde{p}_{\tilde{t}}$ from the description, the reduced phase space is obtained, which is parameterized by $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{p}}$. These variables $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{p}}$ are just static. (After all, they correspond to the initial state.)

What does all this buy us? The reduced phase space formulation is another way of formulating the dynamics of the Newtonian particle, where the unphysical parameterization has indeed been eliminated. But, as with the other formulations, the ontology has not changed by doing so. What is physically real is the particle whose position is changing over time according to $\mathbf{x}(t) = \tilde{\mathbf{p}}t/m + \tilde{\mathbf{x}}$. The path can indeed be fully specified by the initial state $(\tilde{\mathbf{x}}, \tilde{\mathbf{p}})$. But that is not to say that the ontology should comprise just that initial state. One could entertain the latter possibility, but then indeed it becomes problematic to account for a changing world.[2]

To finish this section, a word about clock variables. Consider n non-relativistic particles. Then a collection of particles may serve as a clock whenever its configuration is non-stationary. As a concrete example, suppose there is a particle for which one of the coordinates $z(t)$ is monotonically increasing with time.[3] The motion of the other particles might then be expressed with respect to the clock variable z, by inverting the relation, i.e., $t(z)$, and substituting that in the positions of the other particles: $\mathbf{x}_1(t(z)), \ldots, \mathbf{x}_n(t(z))$. Other variables could act as clock variables, but it is always the dynamics of the variables that determines whether they can serve as clock variables or not, and how one can switch between clock variables.

3 Non-relativistic Bohmian Mechanics

Canonical quantization is a recipe for getting a quantum theory from a classical theory, starting from the Hamiltonian formulation. In the previous section, we have seen three different Hamiltonian formulations. Let us now apply the quantization rules. Starting from the Hamiltonian (3) this results in the familiar Schrödinger equation[4]

---

[2] An ontology based on the reduced phase space is not always problematic. In other cases, it makes sense to consider such an ontology. For example, in the case of the free electromagnetic field, the reduced phase space can be parameterized by the transverse part of the field potential, together with its conjugate momentum, and one could entertain an ontology based on this field.

[3] Rather than considering the positions as functions of time t, we could equally well have done this discussion in terms of an arbitrary parameterization s. That there is an external time plays no role here, as long as there is change with respect to s.

[4] Throughout we assume units such that $\hbar = c = 1$.

$$\mathrm{i}\partial_t \psi(\mathbf{x}, t) = -\frac{1}{2m}\nabla^2 \psi(\mathbf{x}, t) + V\psi(\mathbf{x}, t) \tag{15}$$

for $\psi(\mathbf{x}, t)$. Starting from the Hamiltonian (10) yields

$$\mathrm{i}\partial_s \psi(\mathbf{x}, t, s) = 0, \qquad -\mathrm{i}\partial_t \psi(\mathbf{x}, t, s) - \frac{1}{2m}\nabla^2 \psi(\mathbf{x}, t, s) + V\psi(\mathbf{x}, t, s) = 0, \tag{16}$$

for the wave function $\psi(\mathbf{x}, t, s)$. (The second equation arises here as the operator equivalent of the constraint $C = 0$.) Here we see the appearance of a static wave function, i.e., static with respect to the parameterization s. As a wave equation, this is just the Schrödinger equation again.[5] Starting from the reduced phase space, we get

$$\mathrm{i}\partial_s \psi(\tilde{\mathbf{x}}) = 0, \tag{17}$$

for $\psi(\tilde{\mathbf{x}})$. This wave equation is also equivalent with (15), provided $\psi(\tilde{\mathbf{x}})$ is taken as the initial wave function $\psi(\mathbf{x}, 0)$ [12].

In the Bohmian theory, in addition to the wave function which satisfies the Schrödinger Eq. (15), there is an actual point-particle with position $\mathbf{x}$ whose velocity is given by

$$\dot{\mathbf{x}} = \mathbf{v}^{\psi}(\mathbf{x}, t), \tag{18}$$

where

$$\mathbf{v}^{\psi}(\mathbf{x}, t) = \frac{1}{m}\mathrm{Im}\frac{\boldsymbol{\nabla}\psi(\mathbf{x}, t)}{\psi(\mathbf{x}, t)}. \tag{19}$$

The velocity field is of the form $\mathbf{v}^{\psi} = \mathbf{j}^{\psi}/|\psi|^2$, where $\mathbf{j}^{\psi}$ is the usual current, satisfying the continuity equation

$$\partial_t |\psi|^2 + \boldsymbol{\nabla} \cdot \mathbf{j}^{\psi} = 0. \tag{20}$$

This form of the velocity field can be used to formulate a Bohmian dynamics in other contexts [19]. For example, in the formulation (16), we can consider trajectories in parameterized form $x(s) = (t(s), \mathbf{x}(s))$, with a velocity determined by the current $j^{\psi} = (|\psi|^2, \mathbf{j}^{\psi})$. Since the parameterization of the curve is arbitrary, the dynamics can be written as

$$x' \sim j^{\psi}, \tag{21}$$

which geometrically expresses that the tangent vector to the curve in space-time is always proportional to j^{ψ}. We can explicitly introduce an arbitrary (non-vanishing) factor $N(x, s)$ and write the dynamics as

[5] To show the equivalence as quantum theories, more work is needed, by also considering the associated Hilbert spaces.

$$x' = \frac{N}{|\psi|^2} j^{\psi} \tag{22}$$

or

$$t' = N, \qquad \mathbf{x}' = N\mathbf{v}^{\psi}(\mathbf{x}, t). \tag{23}$$

It is clear that this is the dynamics (18) written in parameterized form. In addition, even though the wave function is static with respect to s, the actual configuration x generically is not. (The fact that ψ still depends on t does not matter, see further below.) Applying the same recipe to the wave equation arising from reduced phase space quantization, suggest a Bohmian dynamics $\tilde{\mathbf{x}}' = 0$. This would only make sense if $\tilde{\mathbf{x}}$ is taken as the initial position with trajectories again given by (18).

Before turning to gravity, we want to further illustrate some aspects of the non-relativistic Bohmian dynamics that will be relevant in dealing with the problem of time in quantum gravity.

Consider an n-particle system, with particle positions $\mathbf{x}_k$, $k = 1, \ldots, n$, satisfying the guidance equations

$$\dot{\mathbf{x}}_k = \mathbf{v}_k^{\psi} = \frac{1}{m_k} \mathrm{Im} \frac{\nabla_k \psi}{\psi}, \tag{24}$$

with $\psi(\mathbf{x}_1, \ldots, \mathbf{x}_n, t)$ now satisfying the many particle Schrödinger equation

$$\mathrm{i}\partial_t \psi = -\frac{1}{2m_k} \sum_k \nabla_k^2 \psi + V\psi. \tag{25}$$

The first observation is that for a stationary state $\psi = \mathrm{e}^{-\mathrm{i}Et}\phi(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, the time dependence of the wave function is trival. For such a wave function, the velocity field $\mathbf{v}_k^{\psi}$ is time-independent, but still the trajectories could be highly non-trivial (see e.g. the simulations in [20]).

Second, even if the wave function of the total system is stationary, the wave function of a subsystem will typically be non-stationary [7, 8, 21]. To see this, we first need to define the subsystem wave function. Consider the wave function $\psi(\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y}_1, \ldots, \mathbf{y}_m, t)$, where the x-coordinates describe the subsystem and the y-coordinates its environment. The wave function of the subsystem — called the *conditional wave function* — can then be defined as

$$\psi_s(\mathbf{x}_1, \ldots, \mathbf{x}_n, t) = \psi(\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y}_1(t), \ldots, \mathbf{y}_m(t), t), \tag{26}$$

which is the total wave function evaluated for the actual particle positions $\mathbf{y}_1(t), \ldots, \mathbf{y}_m(t)$. One of the reasons this definition is natural is that the velocities of the particles of the subsystem can be written either in terms of ψ or ψ_s, i.e.,

$$\dot{\mathbf{x}}_k = \frac{1}{m_k} \mathrm{Im} \frac{\nabla_k \psi}{\psi} = \frac{1}{m_k} \mathrm{Im} \frac{\nabla_k \psi_s}{\psi_s}. \tag{27}$$

The conditional wave function ψ_s will generically evolve in time and under certain conditions it will satisfy an effective Schrödinger equation.

4 Quantum Gravity

General relativity describes how space-time, described by a manifold $\mathcal{M}$ and metric field $g_{\mu\nu}(x)$, interacts with matter. The equations of motion are given by the Einstein field equations, together with equations for the matter (say particles or a field). The theory exhibits a gauge invariance, namely the invariance under coordinate transformations, i.e., space-time diffeomorphisms. Two metrics connected by a diffeomorphism are considered physically equivalent and they are said to determine the same 4-geometry.

When a space-time is globally hyperbolic, it admits a foliation of $\mathcal{M}$ in terms of spatial hypersurfaces, because $\mathcal{M}$ then topologically equals $\mathbb{R} \times \Sigma$, with Σ a 3-surface. This allows a splitting of space-time into space and time, which is required for the Hamiltonian formulation.[6] Coordinates $x^\mu = (t, \mathbf{x})$ can be chosen such that the coordinate t labels time and $\mathbf{x}$ are coordinates on Σ. In terms of these coordinates, the space-time metric can be written as

$$g_{\mu\nu} = \begin{pmatrix} N^2 - N_i N^i & -N_i \\ -N_i & -h_{ij} \end{pmatrix}, \tag{28}$$

where $N > 0$ is the lapse function, $N_i = h_{ij} N^j$ are the shift functions, and h_{ij} is the induced Riemannian metric on the leaves of the foliation. In the Hamiltonian formulation, only the evolution of the 3-metric h_{ij} is non-trivial; the evolution of the lapse and shift functions is completely undetermined. This dynamics is referred to as *geometrodynamics*.

Canonical quantization of the theory leads to the following wave equation for the wave functional $\Psi(h_{ij}, \phi)$, where h_{ij} is again the 3-metric and ϕ is a scalar field. (Rather than having a scalar field describing the matter, there are of course other possibilities.) The wave functional satisfies the functional Schrödinger equation[7]

$$\mathrm{i}\partial_t \Psi = \int d^3x \left(N\widehat{\mathcal{H}} + N^i \widehat{\mathcal{H}}_i \right) \Psi, \tag{29}$$

where

[6] Just like in special relativity, this splitting is not unique.

[7] We used the same operator ordering as in [22].

$$\widehat{\mathcal{H}} = -\kappa\sqrt{h}\frac{\delta}{\delta h_{ij}}\left(\frac{1}{\sqrt{h}}G_{ijkl}\frac{\delta}{\delta h_{kl}}\right) - \frac{1}{2}\frac{1}{\sqrt{h}}\frac{\delta^2}{\delta\phi^2} + \mathcal{V}(h,\phi), \tag{30}$$

$$\widehat{\mathcal{H}}_i = -2h_{ik}D_j\frac{\delta}{\delta h_{jk}} + \frac{1}{2}\left(\partial_i\phi\frac{\delta}{\delta\phi} + \frac{\delta}{\delta\phi}\partial_i\phi\right), \tag{31}$$

with $\kappa = 16\pi G$, G the gravitational constant, G_{ijkl} the DeWitt metric (which depends on the 3-metric), h the determinant of h_{ij}, and $\mathcal{V}$ is a potential. In addition, the wave functional satisfies the constraints:

$$\widehat{\mathcal{H}}(\mathbf{x})\Psi = 0, \qquad \widehat{\mathcal{H}}_i(\mathbf{x})\Psi = 0, \quad i = 1,2,3, \tag{32}$$

the first of which is the Wheeler-DeWitt equation. These constraints immediately imply that

$$\partial_t\Psi = 0, \tag{33}$$

so that Ψ must be time-independent. This is the source of the problem of time.

In Bohmian quantum gravity, there is an actual 3-metric h_{ij} and a field ϕ, evolving according to

$$\dot{h}_{ij} = 2\kappa N G_{ijkl}\frac{\delta S}{\delta h_{kl}} + D_iN_j + D_jN_i, \tag{34}$$

$$\dot{\phi} = \frac{N}{\sqrt{h}}\frac{\delta S}{\delta\phi} + N^i\partial_i\phi, \tag{35}$$

with $\Psi = |\Psi|e^{iS}$ and D_i the 3-dimensional covariant derivative. Given a lapse and a shift function, a solution for the 3-metric defines a 4-metric using (28). There is an important difference with classical geometrodynamics, however. Whereas classically, any choice of lapse or shift defines the same 4-geometry, this is no longer the case for Bohmian geometrodynamics: different choices for the lapse function will lead to different 4-geometries (unless the difference is just a t-dependent factor). Different shift functions do not affect the 4-geometry. This implies that the theory is not invariant under general diffeomorphisms, but only under spatial diffeomorphisms. This is akin to the situation in special relativity, where the simplest Bohmian formulations employ a preferred reference frame (or more generally a preferred space-time foliation), which breaks the Lorentz symmetry (although not on the observational level). The reason is that the non-locality which is inherent to quantum theory (due to Bell's theorem) is hard to marry with Lorentz invariance, or with diffeomorphism invariance in the case of Bohmian quantum gravity. For some suggestions of how to formulate Lorentz invariant Bohmian theories, see [23].

There is no problem of time in Bohmian quantum gravity [8, 10, 24]. While the wave function is static, the 3-metric and the scalar field generically evolve in time. (For real wave functions, i.e. $S = 0$, such as the Hartle-Hawking wave function, there is no motion, and hence from the Bohmian point of view this wave function is inadequate to describe our universe [25, 26].) Unlike the wave function of the uni-

verse, the wave function of a subsystem is generically time dependent and in certain cases it (approximately) satisfies a Schrödinger equation. For example, in the study of cosmological perturbations (see [10] and references therein), relevant in for example cosmological inflation theory, one can consider a decomposition of the metric and the scalar field in terms of a homogeneous and isotropic background component and fluctuations: $h = h_0 + h_1$, $\phi = \phi_0 + \phi_1$. If the wave function is approximately of the form $\Psi(h, \phi) = \Psi_0(h_0, \phi_0)\Phi(h_0, \phi_0, h_1, \phi_1)$, where Ψ_0 yields the dominant contribution to the velocity field for h_0 and ϕ_0, the conditional wave function for the perturbations h_1 and ϕ_1 might satisfy (given again some approximations) some effective Schrödinger equation with a Hamiltonian depending on the actual Bohmian configuration $(h_0(t), \phi_0(t))$ (which is usually taken to be classical).

There is an approach to the problem of time that is very close to the Bohmian one [8, 24, 27]. In that approach, there is no time at the fundamental level, but it is said to emerge in a semi-classical approximation (see [1] for details and the history of this approach, see also [2, 28] for critical assessments). The starting point is to consider an approximation

$$\Psi(h_{ij}, \phi) \approx \Psi_0(h_{ij})\Phi(h_{ij}, \phi), \tag{36}$$

where Ψ_0 is approximately a "classical state", i.e., it can be approximated by the dominating contribution in its WKB expansion. Next, a classical trajectory $h_{ij}(t)$ is considered, determined by (34), where the phase is that of Ψ_0. (That the trajectory is indeed classical stems from the assumptions about Ψ_0.) Then the wave function

$$\psi(\phi, t) = \Phi(h_{ij}(t), \phi) \tag{37}$$

is defined and shown to approximately satisfy the time-dependent Schrödinger equation

$$\mathrm{i}\partial_t \psi = \int d^3x \left(N\widehat{\mathcal{H}}_M + N^i\widehat{\mathcal{H}}_{Mi}\right)\psi, \tag{38}$$

where $\widehat{\mathcal{H}}_M$ and $\widehat{\mathcal{H}}_{Mi}$ are the matter part of respectively (30) and (31), evaluated for $h_{ij}(t)$. This is the Schrödinger equation for a matter field in the external classical metric $h_{ij}(t)$. Granting the approximations, the introduction of the classical trajectory and of the wave function ψ is rather ad hoc and has no ontological basis. However, what is outlined here can be perfectly justified from the Bohmian point of view. The wave function ψ in (37) then basically amounts to the conditional wave function and in the case the Bohmian field h_{ij} evolves approximately classical, with a velocity field approximately determined solely by Ψ_0, the conditional wave function for the matter field will approximately satisfy (38). Unlike the postulation of (37) above, there is nothing ad hoc about the conditional wave function in Bohmian mechanics. It is the wave function that can be used to write the velocity field of the scalar field. So given that the approximations are justified (and the Bohmian theory makes precise the conditions under which they are), the wave function ψ can be used to write the velocity and hence the dynamics of the actual field $\phi(t)$.

Actually, as was noted before by Padmanabhan [29] and Greensite [30, 31], the assumption of classicality of Ψ_0 played no crucial role in the above analysis. That is, the analysis can be generalized to other wave functions Ψ, by considering a trajectory $(h_{ij}(t), \phi(t))$ satisfying (34) and (35), instead of a classical trajectory, and by replacing (37) by the wave function Ψ evaluated for one of the degrees of freedom of the metric $h_{ij}(t)$. Clearly, this amounts to just assuming the Bohmian dynamics and employing the conditional wave function to get a time-dependent wave equation.[8] In any case, this generalization allows for a much larger class of approximations than merely the one obtained by assuming classicality of Ψ_0 (see Sects. 7 and 9 of [10] and references therein).

There are also approaches to the problem of time that proceed by just postulating clock variables, see e.g. [32]. However, by merely postulating them rather than deriving them from the ontology, one tends to commit redundancy and risk inconsistency [33]. At the end of Sect. 2, it was discussed how clock variables can be defined in the context of classical mechanics. In Bohmian mechanics they are defined exactly the same way and as such this could provide an underpinning of these clock-based approaches to the problem of time.

We have only discussed canonical quantum gravity. Loop quantum gravity [13, 34], which is also obtained by quantizing general relativity, but based on a different representation than canonical quantum gravity, also suffers the problem of time. Here too, a Bohmian version could solve the problem. See [35] for the treatment of mini-superspace models (which are simplified models of quantum gravity, potentially applicable in cosmology).

Acknowledgements It is a pleasure to thank Detlef Dürr, Sheldon Goldstein, Stefan Teufel, Roderich Tumulka, Hans Westman and Nino Zanghì for discussions on this topic over the years. This work was presented at the OLOFOS Seminar in Leuven. I thank the audience and especially the respondent Alexandre Guay for the discussions. This work is supported by the Research Foundation Flanders (Fonds Wetenschappelijk Onderzoek, FWO), Grant No. G0C3322N.

References

1. C. Kiefer, *Quantum Gravity*, 2nd edn. (Oxford University Press, Oxford, 2007)
2. K.V. Kuchař, Time and interpretations of quantum gravity, in *Proceedings of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics*, eds. by G. Kunstatter, D. Vincent, J. Williams (World Scientific, Singapore, 1992). Reprinted in Int. J. Mod. Phys. D **20**, 3–86 (2011)
3. C.J. Isham, Canonical quantum gravity and the problem of time, in *Integrable Systems, quantum Groups, and quantum Field Theories*, eds. by L.A. Ibort and M.A. Rodriguez (Kluwer Academic Publishers, London, 1993), p. 157. arXiv:gr-qc/9210011
4. E. Anderson, *The Problem of Time* (Springer, 2017)

[8] The papers of Padmabhan and Greensite actually predate the development of Bohmian quantum gravity, which was initiated shortly afterwards by Vink [24]. Vink *also* emphasized that Bohmian quantum gravity yields a generalization of the semi-classical approach to time.

5. D. Bohm, B.J. Hiley, *The Undivided Universe* (Routledge, New York, 1993)
6. P.R. Holland, *The Quantum Theory of Motion* (Cambridge University Press, Cambridge, 1993)
7. D. Dürr, S. Teufel, *Bohmian Mechanics* (Springer-Verlag, Berlin, 2009)
8. S. Goldstein, S. Teufel, Quantum spacetime without observers: ontological clarity and the conceptual foundations of quantum gravity, in *Physics Meets Philosophy at the Planck Scale*, eds. by C. Callender, N. Huggett (Cambridge University Press, Cambridge, 2004), pp. 275–289. arXiv:quant-ph/9902018
9. N. Pinto-Neto, The Bohm interpretation of quantum cosmology. Found. Phys. **35**, 577–603 (2005). arXiv:gr-qc/0410117
10. N. Pinto-Neto, W. Struyve, Bohmian quantum gravity and cosmology, in *Applied Bohmian Mechanics: From Nanoscale Systems to Cosmology*, 2nd edn, eds. by X. Oriols, J. Mompart, pp. 607–656 (2019). arXiv:1801.03353 [gr-qc]
11. K.V. Kuchař, Canonical quantum gravity, in *General Relativity and Gravitation 1992*, eds. by R.J. Gleiser, C.N. Kozamah, O.M. Moreschi (Institute of Physics Publishing, Bristol, 1993), pp. 119–150. arXiv:gr-qc/9304012
12. G. Ruffini, Quantization of Simple Parametrized Systems. Ph.D. thesis, UC Davis, 1995. arXiv:gr-qc/0511088
13. C. Rovelli, F. Vidotto, *Covariant Loop Quantum Gravity* (Cambridge University Press, Cambridge, 2014)
14. J. Earman, Thoroughly modern McTaggart: or what McTaggart would have said if he had learned the general theory of relativity. Philosophers' Impr. **2**, 3 (2002)
15. T. Maudlin, Thoroughly muddled mctaggart: or how to abuse gauge freedom to generate metaphysical monstrosities. Philosophers' Impr. **2**, 4 (2002)
16. R. Healey, Can physics coherently deny the reality of time? R. Inst. Philos. Suppl. **50**, 293–316 (2002)
17. K. Sundermeyer, *Constrained Dynamics*. Lecture Notes in Physics, vol. 169 (Springer-Verlag, Berlin, 1982)
18. M. Henneaux, C. Teitelboim, *Quantization of Gauge Systems* (Princeton University Press, New Jersey, 1992)
19. W. Struyve, A. Valentini, de Broglie-Bohm guidance equations for arbitrary Hamiltonians. J. Phys. A **42**, 035301 (2009). arXiv:0808.0290v3 [quant-ph]
20. A. Cesa, J. Martin, W. Struyve, Chaotic Bohmian trajectories for stationary states. J. Phys. Math. Theor. **49**, 395301 (2016). arXiv:arXiv:1603.01387 [quant-ph]
21. D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the origin of absolute uncertainty. J. Stat. Phys. **67**, 843–907 (1992). arXiv:quant-ph/0308039. Reprinted in [36]
22. D. Dürr, W. Struyve, Quantum Einstein equations. Cl. Quantum Gravity **37**, 135002 (2020). arXiv:2003.03839 [gr-qc]
23. D. Dürr, S. Goldstein, T. Norsen, W. Struyve, N. Zanghì, Can Bohmian mechanics be made relativistic?. Proc. R. Soc. A **470**, 20130699 (2014). arXiv:1307.1714 [quant-ph]
24. J.C. Vink, Quantum potential interpretation of the wave function of the universe. Nucl. Phys. B **369**, 707–728 (1992)
25. E.J. Squires, An apparent conflict between the de Broglie-Bohm model and orthodoxy in quantum cosmology. Found. Phys. Lett. **5**, 71–75 (1992)
26. Y.V. Shtanov, Pilot wave quantum cosmology. Phys. Rev. D **54**, 2564–2570 (1996). arXiv:gr-qc/9503005
27. C. Callender, R. Weingard, Time, Bohm's theory, and quantum cosmology. Philos. Sci. **63**, 470–474 (1996)
28. E.Y.S. Chua, C. Callender, No time for time from no-time. Philos. Sci. **88**, 1172–1184 (2021)
29. T. Padmanabhan, A definition for time in quantum cosmology. Pramana J. Phys. **35**, L199–L204 (1990)
30. J. Greensite, Time and probability in quantum cosmology. Nucl. Phys. B **342**, 409–429 (1990)
31. J. Greensite, Ehrenfest's principle in quantum gravity. Nucl. Phys. B **351**, 749–766 (1991)
32. D.N. Page, W.K. Wootters, Evolution without evolution: dynamics described by stationary observables. Phys. Rev. D **27**, 2885–2892 (1982)

33. J.S. Bell, On the impossible pilot wave. Found. Phys. **12**, 989–999 (1982). Reprinted in J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987)
34. T. Thiemann, *Modern Canonical Quantum General Relativity* (Cambridge University Press, Cambridge, 2007)
35. W. Struyve, Loop quantum cosmology and singularities. Sci. Rep. **7**, 8161 (2017). arXiv:1703.10274 [gr-qc]
36. D. Dürr, S. Goldstein, N. Zanghì, *Quantum Physics Without Quantum Philosophy* (Springer-Verlag, Berlin, 2012)

Some Often Loosely Used Concepts with Potentially Problematic Implications

Daniel Sudarsky

Abstract We point out some concepts that appear rather frequently in physics discussions, which, despite a seemingly innocent initial appearance, turn out to have important implicit implications that put into question the very assumption of their meaningfulness. The message of this essay is that, in order to avoid the ensuing confusions, their usage should be accompanied with clarifications that make them meaningful, and then to confront the often uncomfortable underlying assumptions required to do so. In particular, we will visit the notions of novel physics occurring at a certain energy scale, or at certain distance from a black hole's event horizon, and the various meanings of the word "fluctuations".

1 Introduction

Physicists, as any other group devoted to a certain activity, are creatures of habits. We learn to employ certain concept within a certain realm and then, sometimes inadvertently, we push its usage into novel or extended realms, where such concept might lose its meaning or, alternatively, extra structure or additional characterization of the situation at hand would be required in order for it to retain it. As our discussions become more abstract and the intended context becomes larger, such attitude might lead to either serious confusion, or to outright erroneous conclusions.

Below we will discuss some of the most prevailing examples that arise in discussions concerning what I would refer to as *"fundamental physics"*, in contrast with, say, applied physics, where the rather direct connection of the discussions with the concreteness of specific situations that occupy the latter disciplines, often prevent the occurrence of the kind of problems I will be pointing out.

We will discuss some of the illustrative examples in all detail for the benefit of those readers who are not very familiar with the corresponding mathematical notions and manipulations, and want to see the conceptual discussion illustrated in very explicit

D. Sudarsky (✉)
Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Apartado Postal 70-543, 04510 Mexico City, Mexico
e-mail: sudarsky@nucleares.unam.mx

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_16

and precise terms. The present work, in fact, relies on some notes prepared for usage in summer schools that took place in the context of the activities of the John Bell Institute for the Foundations of Physics, on whose board Prof. Detlef sat, for the last couple of years of his life.

2 Novel Physics Occurring at a Certain Scale

This is one of the most commonly used notions and one that reflects, in a sense, a rather widespread intuition shared among most physicists working in fundamental aspects of high energy physics, gravitation, in general, and cosmology, in particular, namely the claim that our theories can be expected to work only up to certain scale, that new physics will be required to deal with situations involving higher energy and smaller length scales. The problem is, of course, that, according to one of the most cherished aspects of our understanding of physics, the very notions of energy and length are themselves relative, and thus the very statement, taken to have absolute validity, about the energy or length scales at which our theories might break down does not seem to be sensible at all.

The point is that the physical condition involving an arbitrary system cannot, in general, be characterized in a meaningful way by assigning to it, in a canonical manner, a characteristic energy or length. Such notions do not even make sense in the limited context of "special relativity", and much less in that of general relativity. Thus, in order to make things meaningful, we must either be more explicit regarding the notion of energy or length scale we are employing, indicating how is it to be extracted from the kind of situation under consideration, or, if that is the case, to make it clear that we are introducing additional structure into the discussion in order to make such statements meaningful. For instance, one could declare that in one's proposal there is a preferential frame, and that the energy one is referring to is, say, the maximal of the energies of the particles involved as "seen" in such preferential frame.

An illustrative example is provided by the statement, which arises in quantum field theory, that certain coupling constants, or other parameters of the theory *"run" with energy*. At first sight, this sounds as indicating a violation of Lorentz Invariance, which in fact, is not, as we will see.

It is often the case that something like the intrinsic effective strength of an interaction in a 2 particle scattering experiment (for instance electron-proton) depends, in a nontrivial manner, on the parameters characterizing the specific scattering situation. Such dependence can be separated into a part that is purely *kinematical* and tied to the frame in which it is being described, and a part that is frame independent, and which is considered as more *fundamental*. We might instead consider referring to the center of mass energy, but then, of course, we will have to face the fact that the value for a cup of coffee, not to say the Earth itself is much larger than the GUT (Grand Unified Theory) or Planck scales.

Consider the scattering of two particles (to be specific think of an electron and a muon) with "incoming" 4 momenta P_1^a & P_2^a and "outgoing" 4 momenta P_3^a and P_4^a. We have certain simple relations say, $P_1^2 = P_3^2 = M_\mu^2$ (for the muon) and $P_2^2 = P_4^2 = M_e^2$ (for the electron) which are just constants of nature, while overall conservation dictates $P_1^a + P_2^a = P_3^a + P_4^a$. One might define the 4 momentum exchange as $Q^a = P_3^a - P_1^a$, and compute the invariant Q^2.[1]

It is then not strange that the "actual intensity" of the interaction (a quantity that is reflected in the scattering probability) would depend on something like Q^2. Actually, one of the effects of quantum field theory and the perturbative treatment is that the so called radiative corrections induce such behavior. It is also the case that such dependence (i.e., the part that is frame independent) can be described by reabsorbing the effect in the value of a constant which is used in describing the strength of the interaction. This dependence is then described as *"running of the constant with energy"*.

For instance, in QED the effective coupling constant $\alpha \equiv \frac{e^2}{4\pi}$ "runs with energy" (as long as α remains small enough) according to:

$$\alpha(Q^2) = \frac{\alpha_0}{1 - (\alpha_0/3\pi) ln(Q^2/\mu^2)}, \qquad \alpha_0 = \alpha_(\mu^2) \tag{1}$$

i.e. the value of the quantity α_0 is defined in a scattering experiment with momentum exchange of magnitude $Q^2 = \mu^2$. For $Q \sim m_e/10$ we have $\alpha \sim 1/137$.

It is often precisely, in the sense discussed above, that high energy physicists talk about the possibility of physics beyond the standard model appearing, say, at the "GUT" energy scale. The problem is that without being explicit about the situation, to the point where the 4 momentum exchange can be identified, and a corresponding invariant magnitude be given a clear meaning, any such talk about *the running with energy* can become rather *"obscure"*. The "actual" energy of a particle always depends on the reference frame, while the magnitude of its 4-momentum is a constant of nature (m^2), so one can not make sense of something like a constant of nature running with energy unless it is in the kind of specific sense we described above (or something very similar), or, alternatively by declaring that one is considering a proposal in which there is a preferential frame, and, thus, one is contemplating giving up a basic postulate of special relativity.

Nonetheless, it is rather common to hear such talk in contexts where no appropriate notions can, or have, been introduced. For the sake of clarity, one should demand to know what specific energy are we supposed to be talking about when we are told that according to a certain proposal, at certain energy scale, some novel aspect of physics is expected to arise. It is also, in the sense discussed above, that one can meaningfully adopt the assumption that something dramatic would occur at, say the Planck scale. That is, if one does not want to renounce the basic postulates of special relativity, or even the underlying foundational principles behind general relativity, one might

[1] There are various other invariant characterizations of the scattering of two particles, but as a result of overall four momentum conservation they are not independent.

consider the possibility that, when the scale of some invariant quantity reaches a certain fraction of the Planck scale, some novel aspect of physics, that was otherwise negligible, becomes important. One might instead adopt the posture that those basic principles are, in fact, not valid at all in a strict sense, and that something like a preferred reference frame plays a fundamental role in physics. There is, of course, nothing wrong with contemplating such possibilities, in which case one should aim to be as specific as possible about what is being contemplated.

One relatively natural option is to make use of some contingent aspect of our world, as provided, for instance, by cosmology, or by the region of the universe that we happen to inhabit. Those could provide additional structure that might give meaning to some of the otherwise problematic concepts. The idea would be to postulate that such contingent facts define a preferential global (or relatively local) frame that plays a fundamental role in the micro physics one is considering and, thus, gives an invariant meaning to the notion of *a particle's energy*. For instance, one can take the frame defined by cosmology, more specifically a local frame defined in the following way: consider at any event P the set of four velocities of a possible observer $\{u^a /$ is future directed and $u^a u_a = -1, \}$ and single out the u^a that minimizes the anisotropy such observer would associate to the CMB.[2] Let us denote the corresponding, dipole minimizing u^a by W^a. That is, in fact, how we determine the so called peculiar velocity of our galaxy,[3] (with the Earth's peculiar velocity being, of course, affected by the Earth's rotation around the Sun, an effect that can be eliminated by appropriate year long averaging, and similar corrections to deal with the solar system's motion around the galactic center).

Such idea was, actually, being considered (and continues to be contemplated in other guises) a few years ago, when an avenue of research known as "Quantum Gravity Phenomenology" became extremely popular [1]. The basic idea was that Quantum Gravity effects would lead to certain kind of fundamental granularity of space-time, which would affect the behavior of particles, in a way that depended strongly on their energy E. If at that point one demanded to know which energy was being assigned to the particle in such discussions, the response would have been precisely 'that associated with the CMB" frame. The effect was supposed to be encoded in the expression for the modified dispersion relations (i.e., the relation between energy and momentum), namely $E^2 = \mathbf{P}^2 + m^2 + \xi E^3/M_P$ (with the Planck mass M_P indicating the quantum gravitational nature of the effect, and ξ an unknown parameter). That was clearly not Lorentz invariant and, thus, if the Lorentz transformation law between frames was preserved[4] it could only be valid in one frame at most. That frame was

[2] More precisely compute the dipole moment of $\delta T(\theta, \varphi)/\bar{T}$ (where $\delta T(\theta, \varphi) = T(\theta, \varphi) - \bar{T}$, with $T(\theta, \varphi)$ the CMB temperature observed in the direction specified by suitable angular coordinates on the sky, and $\bar{T}$ the sky average of such quantity), and define the preferred local frame at P selected by cosmology, by the condition that the dipole moment (as evaluated by an observer with four velocity u^a) vanishes, and where the CMB anisotropies are non rotating.

[3] Which is currently estimated to be about 600 km/sec.

[4] Some people working in the field have attempted to rely on what I see as a rather desperate move, involving a nonlinear modification of the Lorentz transformations, preserving the equivalence between all frames but where there is, nevertheless, a minimal length characterizing space-time,

taken to correspond to the four velocity W^a, so the modified dispersion relation could be expressed as $P^a P_a + m^2 + (\xi/M_P)(P^a W_a)^3 = 0$. Once expressed in this form, and after realizing that the Earth's four velocity is unlikely to coincide with W^a (at least not all the time), it becomes evident that in the Earth's frame that spacial four vector would have a spatial component **W**, and that, as a result of the Earth's rotation, that vector would rotate in the laboratory with a 24 hr. period. That makes those kinds of proposals susceptible to confrontation with experiments. In particular, with the so called Hughes-Drever experiments. In [2] such analysis was performed, setting what was at the time one of the most stringent bounds on ξ. The point is that the demand to be clear and explicit, in the sense we have been discussing, seems essential in order to be led to such kind of developments.

The idea of an essential granularity of spactime defining a globally preferential frame, as discussed above, was further explored in [3] focusing attention on radiative corrections which nominally involve *"virtual particles" of arbitrarily "high energies"*. Such statements are commonplace in the context of quantum field theory in flat space-times, but even then, the notions are not precise and needless is to say the theory is particularly obscure concerning its ontology. Nevertheless, the theory is extremely successful as attested by a large number of empirically confirmed results, among which we can point, for example, to the calculation of the numerical value of the electron's anomalous magnetic moment which matches experiments up to the fifteen or so digits available both experimentally and theoretically. The underlying theoretical framework behind the type of calculation involved can, perhaps, be best described in terms of the path integral approach in which the recipe for computing the probability for a transition involving specific initial and final states requires summing over all possible intermediate paths taking one state to the other. The point of the work we are discussing is that, if a granularity of space-time exists at any level, then the summation over all those intermediate paths ought to be restricted accordingly. One can expect such restriction to be associated with a preferred frame due to the following argument. At the particle energy levels (as defined in our laboratory's frame) that we have empirical access, there is a connection between momentum and wavelength (as per standard quantum mechanics, but reflected, also in other less obscure alternatives such as bohmian mechanics, spontaneous collapse theories or the so called many world type versions of the theory.[5]) Those wavelengths are, in the context of quantum field theory, associated with the various modes of the quantum field characterizing the basic degrees of freedom that are present in the theory. Let us be quite open about the possibility that all those notions would break down as we increase the lab. energies of the particles involved in our experiments. It seems natural

that is not subject to a Lorentz contraction. I call this move desperate, because it is accompanied by rather problematic consequences, including the fact that, in such proposals, the total energy of a pair of particles depend on the ordering we assign them, no object can have mass higher than a few grams, and a particle on the other side of the universe has profound effects on the dynamics of a collision occurring, for example, here on Earth. Those interested can look up the approach that goes by the name *"Double Special Relativity"*.

[5] The basic set of paths in which *the measurement problem* seems to be at least viably addressable [4].

to expect such break down to occur in a relative smooth manner, as those energies are increased. It can be expected to be accompanied by some novel behaviors that deviate from the prediction of the current theories. Let us imagine reaching the point where some quantitative measure of such deviation reaches the level of, say, 1% but where our current theory can be said to remain approximately valid. Now we can do the following: Perform tests about the energy (as measured in the laboratory) at which the anomalous level of 1% occurs for particles we prepare along different directions. If the resulting energy is the same for all directions, then our lab's reference frame coincides with the preferential reference frame selected by the new theory. If not, we can determine the velocity of the preferential frame with respect to the lab, by using the Lorentz transformations which presumably remain approximately valid at the relative low energies where the novel effects have been kept at the 1% level or lower. It is, of course, conceivable that when such experiments are carried out, the energies in question would turn out depend on direction in a manner that can not be corrected for by any Lorentz transformation. The theory will thus break even more symmetries than expected namely, the isotropy of physics in the preferred frame. However, a preferred frame can still be identified (as in the cosmological setting) looking for that frame in that the anisotropies are minimized (i.e. the frame in which the corresponding dipole term vanishes).

Now, having shown that in principle such preferred frame can be identified (and forgetting for the moment the complication implied by the possible breakdown of isotropy in that frame), it seems clear that we ought to consider the implications of the breakdown of our current theories for modes that are associated with energies (as defined in that frame) higher than a certain value Λ. That means that we should impose a frame dependent cut off on the modes which we include in the radiative corrections. Naturally, we do not know at this point what to do with modes that would have, according to the current theory, even higher energies. Perhaps no such modes exist. Perhaps something completely unexpected and indescribable in terms of our current theories takes place beyond such bounds.

However, we can estimate the effect of simply placing the above mentioned cut off. This analysis has been carried out in [3] with the surprising result that such move would result in Lorentz violating effects that are not suppressed by the scale at which the cut-off is placed. That is a truly surprising result that has had a substantial impact in the whole Quantum Gravity Phenomenology program. Intuitively, it can be regarded as the momentum-space version of the argument behind Olber's paradox, namely, that there is a fundamental difference, with strong empirical consequences, between an infinite (but static) universe and a finite one. For the result we have discussed, this corresponds to the fundamental difference between the result of integrating over the mode's 3-momentum on an unbounded 3-D euclidean space, and the result of restricting such integral to a bounded 3 sphere. In the latter case the main effect of the presence of a boundary does not change with the magnitude Λ (in momentum space) at which the boundary is located because, while the contribution of each area element of the boundary decreases as $1/\Lambda^2$ (as reflected in the functional form of the Feynman propagator), the total area of the boundary increases as Λ^2. This part of the argument pertains naturally only to the region in which the current theory is taken

as being approximately valid. Thus, in order to avoid such problem, the contribution arising from the novel elements of the new theory has to precisely compensate this effect. It is not clear how can such exact compensation come about without some substantial fine tuning, but, of course, that possibility has not been ruled out. In any event, it seems that proponents of such ideas do face a nontrivial challenge in this respect.

3 Novel Physics Occurring at a Certain Distance from a Black Hole Horizon

The discovery by Hawking that, according to quantum field theory in curved space-times, black holes ought to emit thermal radiation reflects a surprising connection between general relativity, quantum theory and thermodynamics. This result has taken a central place in discussions about the nature of the interface between quantum theory and gravitation, drawing the attention of many researchers working in various approaches towards Quantum Gravity. The main object of such discussions is connected with the natural expectation that, as a result of the emitted radiation, black holes can be expected to evaporate, leading to the so called black hole information paradox (better described as the black hole information puzzle, as a true paradox only appears if one makes some important additional assumptions, as discussed, for instance, in [5]). There are many works connected with this issue where one reads that something strange is supposed to happen at small distances (usually l_p) from the event horizon. We want to consider what, in the relativistic context, could possibly be the meaning of such statements.

Distance from a point to a null hypersurface in Minkowski space-time:
Let us study things explicitly in 2-D Minkowski space-time, where the metric is $ds^2 = -dT^2 + dX^2$. Consider the null hypersurface Σ given by $T = X + A$, and consider the distance to a point P with coordinates $(0, 0)$. The geodesics connecting P to a point Q on Σ are of the form $T = \lambda$, $X = m\lambda$ with a fixed m. The coordinates of Q are then $T_Q = A/(1-m)$, $X_Q = mA/(1-m)$, and the interval between P and Q is:

$$\Delta^2_{(m)} = A^2(1+m)/(m-1). \tag{1}$$

That can take any value, such as A^2 (for $m \to \pm\infty$) to $-A^2$ (for $m = 0$), In particular, it tends to $-\infty$ as $m \to 1^-$ and to $+\infty$ as $m \to 1^+$.

That is, the interval takes all values and can not be used to define an invariant separation. Of course, the value corresponding to $-A^2$ tied to a time interval of duration A, and that corresponding to the value A^2 tied to a space like interval of length A are associated with the particular frame associated coordinates X, T. Thus, in this case, there is no covariant notion of the separation of an event to a null hypersurface.

In 4-D Minkowski space-time with metric $ds^2 = -dT^2 + dX^2 + dY^2 + dZ^2$, the range of the values for the interval between the hypersurface Σ given by $T = X + A$ (and arbitrary values of Y, Z) is the same as before. No covariant notion of the separation of a point to a null hypersurface is available in this case either.

What happens with the related notions for space-like or time-like hypersurfaces? **The proper time from a point P to a spatial hypersuface Σ**: This can be defined in SR (and often in GR) as the maximum among the proper-times measured along causal geodesics connecting the two. We will not get into the details here, but considerations analogous to those we will use in the analysis of the next example can be easily applied to that case.

The Proper Distance from a point P to a time-like hypersuface Σ: This can be defined in SR as a mini-max problem.

Consider all flat spatial hypersurfaces $\{\Sigma_T\}$ containing P (corresponding to possible notions of an instant of time, according to some inertial observer, where P is part of the "present") and in each one of those Σ_T let's find the shortest geodesic connecting P and $\Sigma \cap \Sigma_T$, and denote its length by D_T. Now select the supremum among all those D_T, and identify that value with the proper distance between P and Σ. One can show that, in this case, the supremum is actually a maximum. There are, in fact, a multiplicity of Σ_T's corresponding to that value, and they all intersect along the corresponding geodesic.

Here, we offer a detailed **Proof:** Let us choose Minkowskian coordinates such that P corresponds to $(0, 0, 0, 0)$ and Σ to $(t, x, y, z = D)$. The collection of spatial hypersurfaces $\{\Sigma_T\}$ passing through p can be labeled by the time-like unit normal n^a_T so that the points X (with coordinates (t, x, y, z) are on Σ_T iff the "vector" $R^a = X^a - p^a = (t, x, y, z) - (0, 0, 0, 0)$ satisfies $\eta_{ab} R^a n^b_T = 0$. The general form of the unit normal is

$$n^a = \gamma(1, VCos(\theta), VSin(\theta)Cos(\phi), VSin(\theta)Sin(\phi)) \tag{2}$$

with $V < 1$ and $\gamma = (1 - V^2)^{-1/2}$. So the hypersurface Σ_T consists of points with coordinates such that $t = V(xCos(\theta) + ySin(\theta)Cos(\phi) + zSin(\theta)Sin(\phi))$.

A geodesic on Σ_T (parametrized by λ) and starting at p, has the form $t(\lambda) = d\lambda, x(\lambda) = a\lambda, y(\lambda) = b\lambda, z(\lambda) = c\lambda$, with $d = V(aCos(\theta) + bSin(\theta)Cos(\phi) + cSin(\theta)Sin(\phi))$. The intersection of such geodesic with Σ occurs when $c\lambda = D$ or $\lambda = D/c$. The length of such geodesic is

$$L = \int (-\eta_{ab} T^a T^b)^{1/2} d\lambda$$

$$= D[-V^2(a/cCos(\theta) + b/cSin(\theta)Cos(\phi) + Sin(\theta)Sin(\phi))^2 + (a/c)^2 + (b/c)^2 + 1^2]^{1/2} \tag{3}$$

Minimization of geodesic length coincides with minimization with respect to $(a/c, b/c)$, and the corresponding value of D_T is:

$$D_T = D[1 - \frac{V^2(Sin(\theta)Sin(\phi))^2}{1 - V^2[Cos^2(\theta) + Sin^2(\theta)Cos^2(\phi)]}]^{1/2} \quad (4)$$

This quantity is maximal when $V Sin(\theta)Sin(\phi)$, the z component of n^a, vanishes, and the result of the **mini-max** procedure gives $L_{min-max} = D$, which defines an invariant notion of canonical distance between P and Σ.

Let us get back to the distance between a null hypersurface and a point, but now in curved space-times. It seems quite natural to expect that, if certain notions can not be defined, even in the simple case of Minkowski space-time, it should be impossible to do so in the more general context involving curved space-times. For the most part, our intuition about the matter is correct. However, surprisingly sometimes one can define certain notions in a natural way, due to the presence of particular symmetries in some specific situations. Those might serve, for instance, to single out certain "special coordinates" adapted to such symmetries, and those, in turn, might be used to define the desired notions. Often we would be able to define the notions in purely geometric terms, making use of the symmetries, without relying on those specific coordinates. Alternatively, the coordinates might be constructed using the symmetries at hand and then be used in making the desired definitions. A noteworthy case: when there is a unique notion of *staticity*, i.e., a time-like Killing field (Kf) ξ^a that is hypersurface orthogonal. We might then define the distance between a point P and a specific null hypersurface Σ as the length of the shortest (space-like) geodesic connecting P to Σ, and lying on the space-like hyper-surface normal to ξ^a passing through P.

That sounds reasonable. Let us consider then the distance of a point P with coordinates $(0, R_p, 0, 0)$ (without loss of generality) on the exterior of Schwarzschild space-time, to the BH horizon. The metric is $ds^2 = -(1 - 2m/r)dt^2 + (1 - 2m/r)^{-1}$ $dr^2 + r^2(d\theta^2 + Sin^2(\theta)d\varphi^2$. The Kf has components $(1, 0, 0, 0)$. The hypersurface orthogonal to it containing P is the set of events with $t = 0$ and its metric is $d\sigma^2=(1 - 2m/r)^{-1}dr^2 + r^2(d\theta^2 + Sin^2(\theta)d\varphi^2$. The geodesic in question has coordinates $(0, r, 0, 0)$ with $r \in (2m, R_P)$. Its length is then $L = \int_{2m}^{R_P} \frac{1}{\sqrt{(1-2m/r)}}dr$ (which is finite).

However, note that for any P the geodesic in question intersects the Horizon at the bifurcation surface, which only exists for an *eternal black hole*. For a realistic black holes, formed by, say, the gravitational collapse of a star, all hypersurfaces *"heading"* to what would have corresponded to the bifurcation surface, actually intersect the collapsing star. That is, they enter the region where the space-time is no longer stationary, there is no time-like Kf, and thus our recipe is no longer applicable. Again, using this approach the quantity of interest is simply ill defined for realistic black holes.

Let us consider alternatives. In the case at hand, we have other special features, namely spherical symmetry. Here, we might consider to use the 2 dimensional surfaces S_r invariant under rotations, i.e., the spheres of constant r on hypersurfaces of staticity. In our case the one containing P has surface $A_{R_p} = 4\pi R_p^2$ while the horizon has surface $A_H = 4\pi(2m)^2$, so we might define the required distance "D" $= \sqrt{(A_{R_p} - A_H)/(4\pi)} = R_p - 2m$. This seems to do what we want. However, let

us note that this is not really any "distance" at all. Furthermore it is a quantity constructed out of global notions, which are available only in situations of very high symmetry.

Another proposal put forward by a colleague during a discussion was the following: Consider the value of some relevant curvature scalar at the event horizon, for instance, the Kretschmann scalar $K = W_{abcd}W^{abcd}$ where W_{abcd} is the Weyl tensor (which in this case coincides with the Riemann tensor, as the setting under consideration is vacuum general relativity, so the Ricci tensor and Ricci scalar vanish). We might consider then those points for which the value of the quantity K^{-4} (which has dimensions of length) differs from that at the horizon by, for example l_p or less. However, the strategy has some problems: (i) Except for very symmetric situations, K is not constant over the horizon. So we might have to rely on suitable averages. It is rather unclear how to find a suitable definition of those "averages" in general, in part because there the $3 - volume$ element of a null hypersurface is just zero. One might chose spatial sections and their area-elements, but except for static cases it is unclear how to proceed. (ii) The regions of space-time that have nothing to do with the black hole (and which we normally would not define as being close to it, say, some points over the surface of Earth) would, according to our scheme have to be declared to be closer to the horizon than l_p. The point is that no local characterization of the event horizon in terms of curvature or other local geometrical measure exists. Thus, it seems that no "deformation" of such idea can be used to yield the desired notion of *"proximity to the horizon"*, either. In short, no clear suitable candidate for the general notion of *the points that are closer to the horizon than* l_p (or any other specific length) seems to be available.

If the notion can not even be defined in general situations we must conclude that it can play no role in characterizing novel physics, unless we accept that our fundamental theories profoundly violate enshrined principles of Special and General Relativity. This is *conceivable*, of course, but then, in order to talk meaningfully about such theories (or even prospective theories involving that kind of features), we must clarify what principles of GR and SR we are giving up, and how should we look at GR or SR in such novel contexts, etc. It seems to me that this is essential if we seek clarity and avoidance of confusion, because it is quite likely that, while giving up some of those basic principles, the discussion of the novel proposals would nevertheless be carried out using notions and arguments that rely on those very theories we are supposedly discarding. In doing so, we should then seek to be careful in specifying which aspects are being discarded, and which are being preserved, hopefully examining the consistency of such choices.

Once again, going beyond the limited context of the discussion above, one might decide to rely on additional structure, like that provided by asymptotic infinity when making an idealization of an isolated black hole, and using it to identify an asymptotic rest frame of the black hole. But, of course, such strategy would become rather complicated if one wants to contemplate situations involving more than one black hole, or a black hole together with something else. One might again rely on structure provided by the cosmological context. In any event, one should seek to be clear about

which approach one is using, what additional structure one is relying on, and how exactly do those fit in the overall discussion.

The effects of quantum fluctuations:

Another notion that appears quite frequently in physics discussions, and that is prone to generate confusion, is that of *"fluctuations"*: We should distinguish at least 3 uses of the word:

(i) Variations through space-time of well defined attributes of an "extended" entity (e.g. the water level on the ocean).
(ii) Variations of well defined quantities within an ensemble of systems (e.g. the energy of a classical canonical ensemble of similar gas filled boxes).
(iii) Quantum indeterminacies or uncertainties in a single system (fluctuations in the position of a harmonic oscillator in its ground state).

The use of those notions interchangeably, and apparently encouraged by the unfortunate fact that the same word is reserved for such different things, is the source of several instances in which confusion overtakes the discussion of fundamental issues.

These notions are, of course, closely related but they are not really the same. For instance, within the context of classical statistical mechanics, the energy, at a certain time, of a single specific element of a canonical ensemble does not fluctuate. It can fluctuate in time due to its interaction with the thermal reservoir. If we focus on a given time, the energy can fluctuate as we move from one element of the ensemble to another. According to the *"standard"* interpretation of quantum theory, the position of an electron in the ground state of an hydrogen atom does not change in time, and thinking about its quantum fluctuations as reflecting such type of change is fundamentally wrong. According to the theory, the electron, in the ground state, simply does not have a well defined position. If one is not careful and misinterprets the situation, one can be easily led to conclude, for instance, that as the electron is jumping around it ought to be emitting photons, which is, of course, the wrong conclusion (and concerns precisely one of the fundamental issues that quantum theory was designed to address: the stability of matter). Nonetheless, it is quite common for such interpretational mistake to be made, as noted in the introduction, specially in discussions about the early universe, black hole physics etc.

As an extremely widespread example, we should point out to the usual accounts of the emergence of the primordial seeds of cosmic structure in the inflationary universe. In such discussions, which are quite generally claimed to be taking place within the context of standard quantum theory, one hears that the relevant fields of the problem are in the vacuum state (having been driven to such state by the diluting effect of the highly accelerated expansion associated with inflation). Specifically, the state in question is the so called Bunch Davies vacuum (which is the notion of vacuum associated with the stationary condition that would have prevailed in very early stages of inflation, more precisely, in the limit in which inflation was eternal into the past). The point, however, is that such state can be written explicitly, and can be easily shown to be 100% homogeneous and isotropic. If one adopts the posture that the complete description of the system in question is provided by the quantum state (i.e., one does no rely on any kind of hidden variables) and accepts that there

are no observers or measuring apparatuses that could be said to reduce the state "a la" Copenhaguen (or any similar posture), a point already emphasized in [7], one can not but conclude that the situation will remain forever homogeneous and isotropic, and thus that one has no real account for the emergence of any structure whatsoever (see [6] for critical appraisals of some of the most communly used arguments that attempt to evade such conclusions). Some people would go to extremes and argue that the Minkowski vaccum state of a simple scalar field, is also subject to quantum fluctuations, and thus not really homogeneous and isotropic, something that can be readily dismissed as in full contradiction with basic textbook constructions, and even axiomatic formulations of the theory [8].

Of course, there is room to modify that conclusion, but only if one adopts some alternative to the so called standard or text book version of quantum theory. The options available seem to be the Bohmian path, the spontaneous collapse theories path, and perhaps even some Everett-like scheme (see however [9] for discussion of problems with one such approach, and [11] for discussion of the issues and proposals to address those arising in the more popular so called Many Worlds Interpretations).

A recent attempt to account for the cosmological constant [12] was based precisely on that kind of confusion, which in turn, led to rather severe problems in the analysis [13].

One often hears claims that the black hole horizon is subject to quantum fluctuations and where the discussion then moves from one interpretation of the word to another. In my opinion, one can, of course, contemplate any idea along such lines, but should aim to do so while being clear about what exactly one is proposing, what notion of fluctuation one is using, and, more generally, what version of quantum theory one is adopting.

4 Conclusions

I should stress that I do not see a problem in contemplating giving up some principles of our *"well established theories"* (that is often how progress is made). But it seems, at least, careless, to give them up, and then keep pretending we are upholding considerations that rely on those very ideas, and using the discarded theory in our reasoning, as if no violence has been done to it.

The usage of notions (that might be well defined and sensible in a certain context), in rather different contexts, specially those where it is known that such notions lose their absolute meaning, is, surprisingly, one of the most frequent sources of confusion, among physicists.

I think it is fair to say that most of us have heard talks or read manuscripts in which it is argued that is natural to expect certain things to take place at certain energy scales, or at a certain distance from a black hole. That quantum fluctuations ought to be responsible for certain peculiar effects, and other similar expressions, which at first sight might seem to be very sensible. In this work, we point out why it is often the case that accepting at face value such claims, without asking for further

clarification, is a path towards utter confusion. It is not that proposals based on such ideas cannot be made in a sensible manner, but that in order to do so one should be quite specific about the context in which the ideas are supposed to make sense. That often would involve defining in a precise manner what energy one is referring to, what notion of distance one is employing, or what version of quantum theory (often referred to, as the interpretation) one is talking about, which, in fact, requires being quite specific about the ontology of the theory one is adopting. Failure to address, and to fully specify such aspects of the analysis can result in serious errors or misuses of the notions at hand, or in the oversight of some serious problems in the proposals.

In my view, the vagueness of usage of language in those cases is as problematic as, if not more than, some well known common language examples. For example: "Nothing is better than God, a sandwich is better than nothing, thus a sandwich is better than God".[6] If that kind of nonsense can arise in such a simple context, we should indeed be very worried about the level of confusion that can occur in much more abstract discussions.

This manuscript does not really contain anything novel, and is merely an attempt to call attention to various types of confusions that are often found in current discussions about fundamental aspects of physics, particularly so, in contexts involving novel proposals and ideas. Hopefully, this will contribute to clarify the occurrence and usage of the notions involved. Prof. Deltef Dürr was passionate about the quest for clarity and the endeavors to clean up confusion. He was also a truly devoted educator. Although I only had the privilege of interacting directly with him a few times, before his untimely departure, that much was very clear to me. I would thus like to think that he would look benevolently on this manuscript. In any event, any shortcomings present in this humble attempt to honor those particular virtues, among the many he had as a scientist and as a human being, are solely mine.

Acknowledgements We acknowledge partial financial support from CONACyT (México) project 140630.

References

1. D. Mattingly, Modern tests of Lorentz invariance. Living Rev. Relativ. **8**, (2005)
2. D. Sudarsky, L. Urrutia, H. Vucetich, Observational bounds on quantum gravity signals using existing data. Phys. Rev. Lett. **89**, 231 (2002)
3. J. Collins, A. Perez, D. Sudarsky, L. Urrutia, H. Vucetich, Lorentz invariance in quantum gravity: a new fine tuning problem?. Phys. Rev. Lett. **93**, 191301 (2004)
4. T. Maudlin, Three measurement problems. Topoi **14**, (1995)
5. E. Okón, D. Sudarsky, Losing stuff down a black hole. Found. Phys. **48**, 411 (2018)
6. D. Sudarsky, Shortcomings in the Understanding of Why Cosmological Perturbations Look Classical. Int. J. Mod. Phys. D **20**, 509 (2011); J. Berjon, E. Okón, D. Sudarsky, Critical review

[6] I should thank Vishnya Maudlin for giving me an example that works in English when I mentioned another one I knew, but which only works in Spanish.

of prevailing explanations for the emergence of classicality in cosmology. Phys. Rev. D **103**, 043521 (2021)
7. J.S. Bell, Quantum mechanics for cosmologists, in *Quantum Gravity II* (Oxford University Press, New York, 1981)
8. R. Haag, See for instance Axiom 2A , section II.1.2 of *Local Quantum Physics: Fields, Particles, Algebras* (Springer-Verlag, New York, Berlin, Heidelberg, 1992)
9. R.B. Griffiths, *Consistent Quantum Theory* (Cambridge University Press, Cambridge, 2003); M. Gell-Mann, J. Hartle, in *Proceedings of the 4th Drexel Conference on Quantum Non-Integrability: The Quantum-Classical Correspondence* (1998)
10. A. Kent, Consistent sets contradict. Phys. Rev. Lett. **78**, 2874 (1997); E. Okón, D. Sudarsky, On the consistency of the consistent histories approach to quantum mechanics. Found. Phys. **44**, 19 (2014)
11. D. Wallace, For a contemporary exposition see for instance. *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation* (Oxford University Press, 2014)
12. S.S. Cree, T.M. Davis, T.C. Ralph, Q. Wang, Z. Zhu, W.G. Unruh, Can the fluctuations of the quantum vacuum solve the cosmological constant problem?. Phys. Rev. D **98**(6), 063506 (2018)
13. G.R. Bengochea, G. León, E. Okón, D. Sudarsky, Can the quantum vacuum fluctuations really solve the cosmological constant problem?. Eur. Phys. J. C **80**, 8 (2020)

Probability Conservation for Multi-time Integral Equations

Matthias Lienert

This paper is dedicated to the memory of Detlef Dürr,

a wonderful person, scientist and mentor.

Abstract In relativistic quantum theory, one sometimes considers integral equations for a wave function $\psi(x_1, x_2)$ depending on two space-time points for two particles. A serious issue with such equations is that, typically, the spatial integral over $|\psi|^2$ is not conserved in time–which conflicts with the basic probabilistic interpretation of quantum theory. However, here it is shown that for a special class of integral equations with retarded interactions along light cones, the global probability integral is, indeed, conserved on all Cauchy surfaces. For another class of integral equations with more general interaction kernels, asymptotic probability conservation from $t = -\infty$ to $t = +\infty$ is shown to hold true. Moreover, a certain local conservation law is deduced from the first result.

1 Introduction

1.1 Motivation

An elegant but little-known approach to relativistic quantum theory involves wave functions $\psi(x_1, ..., x_N)$ depending on many space-time variables $x_k \in \mathbb{R}^4$ for many particles $k = 1, 2, ..., N$. In addition to other applications, such as by Detlef Dürr and coauthors [1, 2] in the foundations of relativistic quantum theory, these *multi-time wave functions* (see [3] for an introduction) make it possible write down closed integral equations which describe a fixed number of relativistic, interacting particles

M. Lienert (✉)
Marvel Fusion GmbH, Theresienhöhe 12, 80339 Munich, Germany
e-mail: lienertmat@gmail.com

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_17

in a manifestly covariant way. The best-known example is the Bethe-Salpeter (BS) equation [4] which has been used to describe bound states in quantum field theory (QFT). At the time of its discovery, it was hoped that the BS equation represented a fully relativistic–and interacting–generalization of the Schrödinger equation, at least for processes where fermion creation and annihilation are not relevant (such as bound state problems).

> The formulation of a completely relativistic wave equation for two-body systems has, in a certain sense, solved a long-standing problem of quantum mechanics. The natural and simple way in which relativistic invariance is achieved is, of course, very real progress, which may lead one to hope that the main features of the equation are more permanent than the solidity of its present field theoretic foundation might suggest. Furthermore, it is hardly necessary to recall that the usefulness of the equation has been amply demonstrated in several high-precision calculations of energy levels. –Wick, 1954 [5]

However, there is a serious issue with such integral equations. Due to their non-locality in configuration space (in the sense of PDEs), they typically do not imply (local) continuity equations, nor do they conserve the (global) probability integral. In the context of the Bethe-Salpeter equation, it has been said:

> [...] The absence of a positive-definite norm for the wave function and of any orthogonality theorem.–Wick, 1954, listing problems of the BS equation [5]

> Nakanishi (1965) explicitly calculated the normalization integrals in some special cases of the equal-mass Wick-Cutkosky model, and discovered that certain B-S amplitudes have negative or zero norm. – Nakanishi, 1969 [6]

Of course, these quotes mean nothing else than that the quantities proposed as a norm do not actually constitute one. They cannot have the physical meaning of a probability integral. Considering that quantum physics is based on the notion of probability, this seems rather problematic for the physical justification of the Bethe-Salpeter equation.

The motivation for integral equations for a multi-time wave function can also be approached from a second angle–one that was dear to Detlef Dürr: the quantization of Wheeler-Feynman (WF) electrodynamics [7–11]; see [12–14] for some of Detlef Dürr's works on the topic. This theory pursues the idea that the ultraviolet divergence problem of classical Maxwell-Lorenz electrodynamics can be avoided by "integrating out" the fields. The result is a dynamics where interactions between particles occur directly and exactly when particle world-lines are light-like separated.

The discussions with Detlef Dürr about finding a suitable quantum version of that theory sparked my personal curiosity about the subject. After studying previous proposals for quantizations of WF electrodynamics [15–18], [19, chap. 8], which all encounter their own difficulties, it seemed to me that integral equations for a multi-time wave function might be a more promising way forward. In [20], I laid out how these types of integral equations make it possible to transfer the principle of direct interactions along light cones, that forms the core of WF electrodynamics, to the quantum level. This was done in a way that retains the Dirac-Schrödinger equation with a spin-dependent Coulomb potential as the non-relativistic limit, thus

staying close to empirically successful models. In a series of papers [21–24], my co-authors and I were able to prove that multi-time integral equations provide a fully relativistic and interacting quantum dynamics which does not suffer from the ultraviolet divergence problem, even for singular light-cone interactions [24].

However, the question of probability conservation was left open and, as we have seen for the BS equation, there is reason for concern. Equations with interaction terms which are non-local in the configuration space of quantum mechanics do typically not imply local conservation laws. It is then at best unclear whether global probability conservation holds true. Historically, Feynman himself saw the problem of probability conservation as one of the central obstacles to quantizing WF electrodynamics, as he reports in his Nobel lecture:

> I found that if one generalized the action from the nice Lagrangian forms [...] to these forms [...] then the quantities which I defined as energy, and so on, would be complex. The energy values of stationary states wouldn't be real and probabilities of events wouldn't add up to 100%. –Feynman, 1965 [25]
>
> I don't think we have a completely satisfactory relativistic quantum-mechanical model, even one that doesn't agree with nature, but, at least, agrees with the logic that the sum of probability of all alternatives has to be 100%. –Feynman, 1965 [25]

In view of these difficulties, if one is not ready to dismiss multi-time integral equations altogether, one may conclude that the equations are not exactly the right ones and that some modification is in order. From the point of view that the interaction term in the BS equation which quantum electrodynamics suggests (an infinite series of Feynman diagrams in need of renormalization) is not the most simple and natural, such a modification seems easy to accept. In addition, the argument that non-local interaction terms usually preclude local conservation laws does not apply to global conservation laws, leaving room for logical possibilities which may not have been sufficiently explored. I am going to adopt these positions here. This makes it possible to prove that for certain classes of integral equations the global probability integral is, in fact, conserved.

2 The Integral Equation

For simplicity, we focus on the case of $N = 2$ Dirac particles. Moreover, we set $c = 1 = \hbar$. Then the class of integral equations we shall study reads:

$$\psi(x_1, x_2) = \psi^{\text{free}}(x_1, x_2) + i \int d^4x_1' \, d^4x_2' \, G_1(x_1 - x_1') G_2(x_2 - x_2') K(x_1', x_2') \psi(x_1', x_2'). \tag{1}$$

Here $\psi : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{C}^4 \otimes \mathbb{C}^4$ is a multi-time wave function with 16 complex components for two particles.[1] $\psi^{\text{free}}(x_1, x_2)$ is a solution of the free Dirac equation in x_1, x_2, i.e.:

[1] One can also study (1) on a sub-domain of $\mathbb{R}^4 \times \mathbb{R}^4$, e.g., on the set $\mathscr{S}$ of space-like configurations.

$$(-i\gamma_k^\mu \partial_{x_k^\mu} + m_k)\psi^{\text{free}}(x_1, x_2) = 0, \quad k = 1, 2. \tag{2}$$

The space-time integral in (1) extends over $\mathbb{R}^4 \times \mathbb{R}^4$, the entire configuration space-time. G_1 and G_2 are Green's functions of the Dirac equations of particles 1 and 2. We use the convention of [26, Appendix E]:

$$(-i\gamma_k^\mu \partial_{x_k^\mu} + m_k)G_k(x_k - x_k') = \delta^{(4)}(x_k - x_k'). \tag{3}$$

Here and in the following, particle indices in γ-matrices, Green's functions and propagators indicate on which spin index their matrix structure acts.

$K(x_1, x_2)$ is the so-called *interaction* kernel, a covariant, matrix-valued distribution. We require the following symmetry condition with respect to its matrix structure:

$$K^\dagger(x_1, x_2) = \gamma_1^0 \gamma_2^0 \, K(x_1, x_2) \gamma_1^0 \gamma_2^0. \tag{4}$$

As explained in [20], direct interactions along light cones in the spirit of Wheeler-Feynman electromagnetism can be expressed by the interaction kernel

$$K^{\text{sym}}(x_1, x_2) = \lambda \, \gamma_1^\mu \gamma_{2,\mu} \, \delta((x_1 - x_2)^2) \tag{5}$$

where $\lambda \in \mathbb{R}$ is a coupling constant and $(x_1 - x_2)^2 = (x_1^0 - x_2^0)^2 - |\mathbf{x}_1 - \mathbf{x}_2|^2$ denotes the Minkowski square. Note that (5) contains both retarded and advanced interaction terms, as can be seen by decomposing the delta distribution. The retarded part is given by:

$$K^{\text{ret}}(x_1, x_2) = \lambda \, \gamma_1^\mu \gamma_{2,\mu} \, \frac{1}{2|\mathbf{x}_1 - \mathbf{x}_2|} \delta(x_1^0 - x_2^0 - |\mathbf{x}_1 - \mathbf{x}_2|). \tag{6}$$

With these conventions, the factor i in the interaction term in (1) is required to obtain the correct non-retarded limit, i.e., a Schrödinger equation with spin-dependent Coulomb potential [20].

Relation to the Bethe-Salpeter equation. The BS equation is contained in the class of Eq. (1) for the case that G_k are Feynman propagators S_k^F for the two particles $k = 1, 2$, and for the case that $K(x_1, x_2)$ is given by an infinite series of Feynman diagrams. In the so-called ladder-approximation of the BS equation, only a certain sub-class of these Feynman diagrams (consisting of those exchanging only one virtual photon at a time) is considered. Then K simplifies to

$$K^{\text{BSL}}(x_1, x_2) = \lambda \, \gamma_1^\mu \gamma_2^\nu D_{\mu\nu}^F(x_1, x_2) \tag{7}$$

where $D_{\mu\nu}^F$ is the Feynman propagator of a photon (see [27, p. 331]). In Lorenz gauge:

$$D_{\mu\nu}^F(x_1, x_2) = \eta_{\mu\nu} \, D^F(x_1, x_2) \tag{8}$$

where D^F is the Feynman propagator of the wave eq. and $\eta_{\mu\nu}$ the Minkowski metric. As both D^F and $\frac{1}{4\pi}\delta((x_1 - x_2)^2)$ are Green's functions of the wave equation, (7) resembles (5). However, a crucial difference is that only (5) is supported on the light cone; (7) also has support outside.

Role of the dynamics. As discussed in [23], one can best understand the dynamics defined by (1) in the case of $G_k = S_k^{\text{ret}}$, the retarded Green's function of the Dirac equation for particle k. Then, for each incoming free wave function ψ^{free}, the integral equation defines a unique interacting solution ψ which agrees with ψ^{free} in the infinite past. Thus Eq. (1) can be viewed as a machinery which takes an incoming free solution and computes an interacting correction to it.

Notes on retarded Green's functions. We now collect useful properties of retarded Green's functions which are rooted in their simple relation to the propagator of the Dirac equation. These will play a crucial role in the upcoming arguments. Namely, we have:

$$S^{\text{ret}}(x - x') = \theta(x^0 - x'^0) S(x - x') \tag{9}$$

where θ is the Heaviside function and S the propagator of the Dirac equation. S can be used to time-evolve every free solution of the Dirac equation from one Cauchy surface Σ to another:

$$\psi^{\text{free}}(x) = -i \int_\Sigma d\sigma_\mu(x')\, S(x - x') \gamma^\mu \psi^{\text{free}}(x'). \tag{10}$$

Confusingly, both S^{ret} and S^F are called propagators, even though they do not have the property (10) for all wave functions and all Cauchy surfaces. From (10), one can deduce the composition property

$$\int_\Sigma d\sigma_\mu(x')\, S(x - x') \gamma^\mu S(x' - x'') = i S(x - x''). \tag{11}$$

Moreover, we have $[S(x - x')]^\dagger = -\gamma^0 S(x' - x)\gamma^0$. This allows us to compute the adjoint of the integral equation (1), denoting $\overline{\psi}(x_1, x_2) = \psi^\dagger(x_1, x_2)\gamma_1^0 \gamma_2^0$:

$$\overline{\psi}(x_1, x_2) = \overline{\psi}^{\text{free}}(x_1, x_2) - i \int d^4x_1'\, d^4x_2'\, \overline{\psi}(x_1', x_2') K(x_2', x_1') S_1(x_1' - x_1) S_2(x_2' - x_2) \times \theta(x_1^0 - x_1'^0)\theta(x_2^0 - x_2'^0). \tag{12}$$

3 Relativistic Probability Conservation

For Dirac particles, local probability conservation is expressed by the continuity equation $\partial_{x^\mu} j^\mu(x) = 0$ where $j^\mu = \overline{\psi}(x)\gamma^\mu\psi(x)$ denotes the probability current and $\overline{\psi}(x) = \psi^\dagger(x)\gamma^0$. Global probability conservation means that $\int_\Sigma d\sigma_\mu(x)\, j^\mu(x)$ does not depend on the choice of Cauchy surface $\Sigma \subset \mathbb{R}^4$.

For a multi-time wave function for N Dirac particles, these notions can be generalized as follows. Local probability conservation can be expressed by a set of N continuity equations,

$$\partial_{x_k^\mu} j^{\mu_1 \cdots \mu_N}(x_1, ..., x_N) = 0, \quad k = 1, 2, ..., N \tag{13}$$

where $j^{\mu_1 \cdots \mu_N} = \overline{\psi} \gamma_1^{\mu_1} \cdots \gamma_N^{\mu_N} \psi$ denotes the Dirac tensor current.

Equations (13) make it possible formulate a generalized version of the Born rule for all Cauchy surfaces. Let n be the future-directed unit normal vector field at Σ. Then

$$\rho(x_1, ..., x_N) = \overline{\psi}(x_1, ..., x_N) \gamma_1^{\mu_1} \cdots \gamma_N^{\mu_N} \psi(x_1, ..., x_N) n_{\mu_1}(x_1) \cdots n_{\mu_N}(x_N) \tag{14}$$

defines the probability density for N particles $k = 1, ..., N$ to cross Σ at the locations $x_1, ..., x_N \in \Sigma$. In fact, for theories with local interactions and finite propagation speed, it is possible to prove this rule using the usual Born rule in a distinguished frame [28].

The continuity Eq. (13) imply global probability conservation in the sense that

$$P(\psi, \Sigma) = \int_{\Sigma^N} d\sigma_{\mu_1}(x_1) \cdots d\sigma_{\mu_N}(x_N) \, \overline{\psi}(x_1, ..., x_N) \gamma_1^{\mu_1} \cdots \gamma_N^{\mu_N} \psi(x_1, ..., x_N) \tag{15}$$

does not depend on the choice of Cauchy surface Σ. In fact, this requires (13) only on the set of space-like configurations $\mathscr{S} \subset \mathbb{R}^{4N}$, not necessarily on the entire configuration-spacetime $\mathbb{R}^{4N}$. However, in the case that (13) hold true on $\mathbb{R}^{4N}$, one finds that the generalized probability integral

$$P(\psi, \Sigma_1, ..., \Sigma_N) = \int_{\Sigma_1 \times \cdots \times \Sigma_N} d\sigma_{\mu_1}(x_1) \cdots d\sigma_{\mu_N}(x_N) \, \overline{\psi}(x_1, ..., x_N) \gamma_1^{\mu_1} \cdots \gamma_N^{\mu_N} \psi(x_1, ..., x_N) \tag{16}$$

is independent of the choice of N (potentially different) Cauchy surfaces $\Sigma_1, ..., \Sigma_N$.

While $P(\psi, \Sigma)$ seems like the physically appropriate choice[2] for the probability integral, we shall consider $P(\psi, \Sigma_1, ..., \Sigma_N)$ as the more general notion in the following. This simplifies to investigate which local conservation laws follow from the global ones (see Sect. 4.3).

[2] The reason for this is that space-like configurations are the natural generalization of equal-time configurations. Non-space-like configurations can arise from multiple points on a single time-like (or light-like) world line. Thus, there is no physical reason to expect probability conservation on such configurations.

4 Results

4.1 Probability Conservation on All Cauchy Surfaces for Retarded Green's Functions and Retarded Interaction Kernels

As (1) is an integral equation with a non-local interaction term on configuration space, we do not expect it to imply local conservation laws. This means that a method to prove global probability conservation without first establishing local probability conservation is required. Conveniently, the propagator S of the Dirac equation provides such a method.

To see this, let $\psi^{\text{free}}(x)$ be a solution of the Dirac eq. and Σ, Σ' Cauchy surfaces. Then:

$$\begin{aligned} P(\psi^{\text{free}}, \Sigma) &= \int_{\Sigma} d\sigma_\mu(x)\, \overline{\psi}^{\text{free}}(x)\gamma^\mu \psi^{\text{free}}(x) \\ &\overset{(10)}{=} -i \int_{\Sigma} d\sigma_\mu(x)\, \overline{\psi}^{\text{free}}(x)\gamma^\mu \int_{\Sigma'} d\sigma_\nu(x') S(x-x')\gamma^\nu \psi^{\text{free}}(x') \\ &= \int_{\Sigma'} d\sigma_\nu(x') \left(-i \int_{\Sigma} d\sigma_\mu(x)\, \overline{\psi}^{\text{free}}(x)\gamma^\mu S(x-x') \right) \gamma^\nu \psi^{\text{free}}(x') \\ &\overset{(10)}{=} \int_{\Sigma'} d\sigma_\nu(x')\, \overline{\psi}^{\text{free}}(x')\gamma^\nu \psi^{\text{free}}(x') = P(\psi^{\text{free}}, \Sigma'). \end{aligned} \tag{17}$$

Using the relation of retarded Green's functions to the propagator S, we now prove our result.

Proposition 1 *Consider the integral Eq.* (1) *with retarded Green's functions,* $G_k = S_k^{\text{ret}}$, $k = 1, 2$ (9)*, and retarded interaction kernel* (6)*. Then for every solution* ψ *of* (1) *on* $\mathbb{R}^4 \times \mathbb{R}^4$*, the probability integral* $P(\psi, \Sigma_1, \Sigma_2)$ (16) *does not depend on the choice of Cauchy surfaces* $\Sigma_1, \Sigma_2 \subset \mathbb{R}^4$*.*

Proof Let ψ^{free} a solution of the free multi-time Dirac equations (2) and ψ a solution of the integral Eq. (1). Our strategy is to decompose $P(\psi, \Sigma_1, \Sigma_2)$ as

$$P(\psi, \Sigma_1, \Sigma_2) = P(\psi^{\text{free}}, \Sigma_1, \Sigma_2) + P_1(\psi, \Sigma_1, \Sigma_2) \tag{18}$$

and to show that P_1 vanishes for the retarded interaction kernel (6) and all Cauchy surfaces Σ_1, Σ_2. We already know that the free Dirac equations (2) imply the continuity equations (13). Thus:

$$P(\psi^{\text{free}}, \Sigma_1, \Sigma_2) = P(\psi^{\text{free}}, \Sigma_3, \Sigma_4) \tag{19}$$

for all Cauchy surfaces $\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4$. This allows us to deduce

$$P(\psi, \Sigma_1, \Sigma_2) = P(\psi^{\text{free}}, \Sigma_1, \Sigma_2) = P(\psi^{\text{free}}, \Sigma_3, \Sigma_4) = P(\psi, \Sigma_3, \Sigma_4) \tag{20}$$

which is the claim. The main work is to prove that $P_1(\psi, \Sigma_1, \Sigma_2)$ vanishes. Plugging the right hand side of Eq. (1) (for ψ) and its adjoint (12) (for $\overline{\psi}$) into $P(\psi, \Sigma_1, \Sigma_2)$, we find, considering (4):

$$\begin{aligned}
&P_1(\psi, \Sigma_1, \Sigma_2) = P(\psi, \Sigma_1, \Sigma_2) - P(\psi^{\text{free}}, \Sigma_1, \Sigma_2) = \\
&2\Im \int_{\Sigma_1 \times \Sigma_2} d\sigma_\mu(x_1)\, d\sigma_\nu(x_2) \int d^4x_1'\, d^4x_2'\, \overline{\psi}(x_1', x_2') K(x_1', x_2') S_1(x_1' - x_1) S_2(x_2' - x_2) \\
&\quad \times \theta(x_1^0 - {x_1'}^0)\theta(x_2^0 - {x_2'}^0) \gamma_1^\mu \gamma_2^\nu\, \psi^{\text{free}}(x_1, x_2) \\
&+ \int_{\Sigma_1 \times \Sigma_2} d\sigma_\mu(x_1)\, d\sigma_\nu(x_2) \int d^4x_1'\, d^4x_2'\, \overline{\psi}(x_1', x_2') K(x_1', x_2') S_1(x_1' - x_1) S_2(x_2' - x_2) \\
&\quad \times \theta(x_1^0 - {x_1'}^0)\theta(x_2^0 - {x_2'}^0) \gamma_1^\mu \gamma_2^\nu \int d^4x_1''\, d^4x_2''\, S_1(x_1 - x_1'') S_2(x_2 - x_2'') \\
&\quad \times \theta(x_1^0 - {x_1''}^0)\theta(x_2^0 - {x_2''}^0) K(x_1'', x_2'') \psi(x_1'', x_2'') \\
&=: P_{1,1}(\psi, \Sigma_1, \Sigma_2) + P_{1,2}(\psi, \Sigma_1, \Sigma_2)
\end{aligned} \tag{21}$$

with $P_{1,1}$ and $P_{1,2}$ defined as the two summands of the equation in the order of appearance.

It is crucial that due to the simple relation (9) of S^{ret} with the propagator S, the propagators S_1 and S_2 appear in the equation which, together with the hypersurface integrals $\int_{\Sigma_1 \times \Sigma_2} d\sigma_\mu(x_1)\, d\sigma_\nu(x_2)$, can be used to evolve ψ^{free}. In the first term $P_{1,1}$, we would like to use (10) in both x_1 and x_2. On first glance, this does not seem possible because $\theta(x_1^0 - {x_1'}^0)\theta(x_2^0 - {x_2'}^0)$ depends on the time variables x_1^0 and x_2^0 of the Cauchy surfaces Σ_1 and Σ_2, respectively. However, since the propagator S of the Dirac equation has only support inside of and on the light cone, we can write:

$$S(x' - x)\theta(x^0 - {x'}^0) = S(x' - x)\theta_\Sigma(x') \quad \text{and} \quad S(x - x')\theta(x^0 - {x'}^0) = S(x - x')\theta_\Sigma(x') \tag{22}$$

where Σ is a Cauchy surface that contains x and

$$\theta_\Sigma(x') = \begin{cases} 1 & \text{if } x' \in \text{past}(\Sigma) \\ 0 & \text{else.} \end{cases} \tag{23}$$

Here, $\text{past}(\Sigma) = \bigcup_{x \in \Sigma} \text{past}(x)$ denotes the part of space-time "below Σ". An important point is that $\theta_\Sigma(x')$ does not depend of x (as long as $x \in \Sigma$). Using (22), we obtain:

$$\begin{aligned}
P_{1,1}(\psi, \Sigma_1, \Sigma_2) = 2\Im \int_{\Sigma_1 \times \Sigma_2} d\sigma_\mu(x_1)\, d\sigma_\nu(x_2) \int d^4x_1'\, d^4x_2'\, \overline{\psi}(x_1', x_2') K(x_1', x_2') \\
\times S_1(x_1' - x_1) S_2(x_2' - x_2) \theta_{\Sigma_1}(x_1') \theta_{\Sigma_2}(x_2') \gamma_1^\mu \gamma_2^\nu\, \psi^{\text{free}}(x_1, x_2).
\end{aligned} \tag{24}$$

It is now possible to exchange the integrals, yielding

$$P_{1,1}(\psi, \Sigma_1, \Sigma_2) = 2\Im \int d^4x_1' \, d^4x_2' \, \overline{\psi}(x_1', x_2') K(x_1', x_2') \theta_{\Sigma_1}(x_1') \theta_{\Sigma_2}(x_2') \\ \times \int_{\Sigma_1 \times \Sigma_2} d\sigma_\mu(x_1) \, d\sigma_\nu(x_2) \, S_1(x_1' - x_1) S_2(x_2' - x_2) \gamma_1^\mu \gamma_2^\nu \, \psi^{\text{free}}(x_1, x_2). \quad (25)$$

This allows us to employ the propagation identity (10) twice to deduce:

$$P_{1,1}(\psi, \Sigma_1, \Sigma_2) = -2\Im \int d^4x_1' \, d^4x_2' \, \overline{\psi}(x_1', x_2') K(x_1', x_2') \theta_{\Sigma_1}(x_1') \theta_{\Sigma_2}(x_2') \psi^{\text{free}}(x_1', x_2'). \quad (26)$$

Now we use the integral equation (1) "backwards" to express ψ^{free} in terms of ψ:

$$\begin{aligned} P_{1,1}(\psi, \Sigma_1, \Sigma_2) &= -2\Im \int d^4x_1' \, d^4x_2' \, \overline{\psi}(x_1', x_2') K(x_1', x_2') \theta_{\Sigma_1}(x_1') \theta_{\Sigma_2}(x_2') \psi(x_1', x_2') \\ &+ 2\Im \, i \int d^4x_1' \, d^4x_2' \int d^4x_1'' \, d^4x_2'' \, \overline{\psi}(x_1', x_2') K(x_1', x_2') \theta_{\Sigma_1}(x_1') \theta_{\Sigma_2}(x_2') \\ &\times \theta(x_1'^0 - x_1''^0) \theta(x_2'^0 - x_2''^0) S_1(x_1' - x_1'') S_2(x_2' - x_2'') K(x_1'', x_2'') \psi(x_1'', x_2'') \\ &= 0 + 2\Re \int d^4x_1' \, d^4x_2' \int d^4x_1'' \, d^4x_2'' \, \overline{\psi}(x_1', x_2') K(x_1', x_2') \theta_{\Sigma_1}(x_1') \theta_{\Sigma_2}(x_2') \\ &\times \theta(x_1'^0 - x_1''^0) \theta(x_2'^0 - x_2''^0) S_1(x_1' - x_1'') S_2(x_2' - x_2'') K(x_1'', x_2'') \psi(x_1'', x_2''). \quad (27) \end{aligned}$$

In order to conclude that the first term vanishes, we have used the symmetry of K (4). Now we employ the identity $2\Re z = z + z^*$:

$$\begin{aligned} &P_{1,1}(\psi, \Sigma_1, \Sigma_2) = \\ &\int d^4x_1' \, d^4x_2' \, d^4x_1'' \, d^4x_2'' \, \overline{\psi}(x_1', x_2') K(x_1', x_2') S_1(x_1' - x_1'') S_2(x_2' - x_2'') K(x_1'', x_2'') \psi(x_1'', x_2'') \\ &\times \Big[\theta_{\Sigma_1}(x_1') \theta_{\Sigma_2}(x_2') \theta(x_1'^0 - x_1''^0) \theta(x_2'^0 - x_2''^0) + \theta_{\Sigma_1}(x_1'') \theta_{\Sigma_2}(x_2'') \theta(x_1''^0 - x_1'^0) \theta(x_2''^0 - x_2'^0) \Big] \end{aligned} \quad (28)$$

We compare this term to $P_{1,2}(\psi, \Sigma_1, \Sigma_2)$. Using (22) and exchanging the order of the integrals, we obtain:

$$\begin{aligned} P_{1,2}(\psi, \Sigma_1, \Sigma_2) &= \int d^4x_1' \, d^4x_2' \, d^4x_1'' \, d^4x_2'' \, \overline{\psi}(x_1', x_2') K(x_1', x_2') \\ &\times \int_{\Sigma_1 \times \Sigma_2} d\sigma_\mu(x_1) \, d\sigma_\nu(x_2) S_1(x_1' - x_1) S_2(x_2' - x_2) \gamma_1^\mu \gamma_2^\nu S_1(x_1 - x_1'') S_2(x_2 - x_2'') \\ &\times \theta_{\Sigma_1}(x_1') \theta_{\Sigma_2}(x_2') \theta_{\Sigma_1}(x_1'') \theta_{\Sigma_2}(x_2'') K(x_1'', x_2'') \psi(x_1'', x_2''). \end{aligned} \quad (29)$$

This allows us to utilize the composition property (11) for the propagators twice, yielding

$$P_{1,2}(\psi, \Sigma_1, \Sigma_2) = -\int d^4x_1' \, d^4x_2' \, d^4x_1'' \, d^4x_2'' \, \overline{\psi}(x_1', x_2') K(x_1', x_2') S_1(x_1' - x_1'') S_2(x_2' - x_2'') \\ \times K(x_1'', x_2'') \psi(x_1'', x_2'') \theta_{\Sigma_1}(x_1') \theta_{\Sigma_2}(x_2') \theta_{\Sigma_1}(x_1'') \theta_{\Sigma_2}(x_2''). \tag{30}$$

Comparing (28) and (30), we find that $P_{1,1}$ and $P_{1,2}$ cancel if:

$$\theta_{\Sigma_1}(x_1')\theta_{\Sigma_2}(x_2')\theta_{\Sigma_1}(x_1'')\theta_{\Sigma_2}(x_2'') \stackrel{!}{=} \\ \theta_{\Sigma_1}(x_1')\theta_{\Sigma_2}(x_2')\theta({x_1'}^0 - {x_1''}^0)\theta({x_2'}^0 - {x_2''}^0) + \theta_{\Sigma_1}(x_1'')\theta_{\Sigma_2}(x_2'')\theta({x_1''}^0 - {x_1'}^0)\theta({x_2''}^0 - {x_2'}^0). \tag{31}$$

To make further progress, note that

$$\theta_{\Sigma_k}(x_k')\theta({x_k'}^0 - {x_k''}^0) = \theta_{\Sigma_k}(x_k')\theta_{\Sigma_k}(x_k'')\theta({x_k'}^0 - {x_k''}^0) \tag{32}$$

as $x_k'' \in \text{past}(x_k')$ and $x_k' \in \text{past}(\Sigma_k)$ together imply that $x_k'' \in \text{past}(\Sigma_k)$. Thus, every term in (31) contains the factor $\theta_{\Sigma_1}(x_1')\theta_{\Sigma_2}(x_2')\theta_{\Sigma_1}(x_1'')\theta_{\Sigma_2}(x_2'')$ and (31) reduces to:

$$1 \stackrel{!}{=} \theta({x_1'}^0 - {x_1''}^0)\theta({x_2'}^0 - {x_2''}^0) + \theta({x_1''}^0 - {x_1'}^0)\theta({x_2''}^0 - {x_2'}^0). \tag{33}$$

In general, this condition does not hold. However, we have not used the special structure of the retarded interaction kernel K^{ret} (6) yet. Instead of (33), we will show that

$$K^{\text{ret}}(x_1', x_2') S_1(x_1' - x_1'') S_2(x_2' - x_2'') K^{\text{ret}}(x_1'', x_2'') \\ \times \Big[1 - \theta({x_1'}^0 - {x_1''}^0)\theta({x_2'}^0 - {x_2''}^0) - \theta({x_1''}^0 - {x_1'}^0)\theta({x_2''}^0 - {x_2'}^0)\Big] = 0. \tag{34}$$

In fact, including the first line adds several geometric conditions on those tuples $(x_1', x_2', x_1'', x_2'')$ which may actually contribute to $P_{1,1}(\psi, \Sigma_1, \Sigma_2) + P_{1,2}(\psi, \Sigma_1, \Sigma_2)$.

(i) ${x_2'}^0 = {x_1'}^0 - |\mathbf{x}_1' - \mathbf{x}_2'|$ (because of $K^{\text{ret}}(x_1', x_2')$),
(ii) ${x_2''}^0 = {x_1''}^0 - |\mathbf{x}_1'' - \mathbf{x}_2''|$ (because of $K^{\text{ret}}(x_1'', x_2'')$),
(iii) $|{x_1'}^0 - {x_1''}^0| \geq |\mathbf{x}_1' - \mathbf{x}_1''|$ (because of the support of S_1),
(iv) $|{x_2'}^0 - {x_2''}^0| \geq |\mathbf{x}_2' - \mathbf{x}_2''|$ (because of the support of S_2).

We now show that the above conditions only allow for the following two cases:

$$1.\ {x_k'}^0 > {x_k''}^0 \text{ for } k = 1, 2 \quad \text{or} \quad 2.\ {x_k''}^0 > {x_k'}^0 \text{ for } k = 1, 2.$$

The logic behind this claim is that if true, $K^{\text{ret}}(x_1', x_2') S_1(x_1' - x_1'') S_2(x_2' - x_2'')$ $K^{\text{ret}}(x_1'', x_2'') \neq 0$ implies $\theta({x_1'}^0 - {x_1''}^0)\theta({x_2'}^0 - {x_2''}^0) + \theta({x_1''}^0 - {x_1'}^0)\theta({x_2''}^0 - {x_2'}^0) = 1$. This, in turn, means that the square bracket in (34) vanishes. Hence, (34) is always

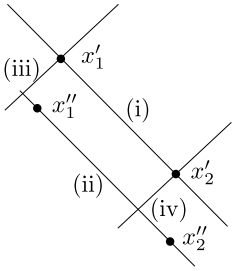


Fig. 1 Illustration of the proof of (34). Begin with x_1'. By (i), x_2' has to lie on the past light cone of x_1'. Next, consider x_1''. Assume that ${x_1'}^0 > {x_1''}^0$. By (iii), x_1'' has to be in the past of x_1'. The final point x_2'' has to lie on the past light cone of x_1'' by (ii). Moreover, by (iv), x_2'' has to be light-like or time-like to x_2'. As one can see, there is no other option for x_2'' than to lie in the past of x_2', which implies ${x_2'}^0 > {x_2''}^0$

satisfied, implying $P_1(\psi, \Sigma_1, \Sigma_2) = P_{1,1}(\psi, \Sigma_1, \Sigma_2) + P_{1,2}(\psi, \Sigma_1, \Sigma_2) = 0$, and thus probability conservation.

We begin with the first case and demonstrate that ${x_1'}^0 > {x_1''}^0$ implies ${x_2'}^0 > {x_2''}^0$ (see Fig. 1 for geometrical intuition). Consider the difference ${x_2'}^0 - {x_2''}^0$. Combining (i) and (ii) yields:

$$ {x_2'}^0 - {x_2''}^0 = {x_1'}^0 - {x_1''}^0 - |\mathbf{x}_1' - \mathbf{x}_2'| + |\mathbf{x}_1'' - \mathbf{x}_2''|. \tag{35} $$

Now, because of (iii) and ${x_1'}^0 > {x_1''}^0$, this implies:

$$ {x_2'}^0 - {x_2''}^0 \geq |\mathbf{x}_1' - \mathbf{x}_1''| - |\mathbf{x}_1' - \mathbf{x}_2'| + |\mathbf{x}_1'' - \mathbf{x}_2''|. \tag{36} $$

Next, we use the triangle inequality twice to obtain:

$$ |\mathbf{x}_1' - \mathbf{x}_2'| \;\geq\; |\mathbf{x}_1' - \mathbf{x}_1''| + |\mathbf{x}_1'' - \mathbf{x}_2'| \;\geq\; |\mathbf{x}_1' - \mathbf{x}_1''| + |\mathbf{x}_1'' - \mathbf{x}_2''| + |\mathbf{x}_2'' - \mathbf{x}_2'|. \tag{37} $$

Together with the previous inequality, this gives us:

$$ {x_2'}^0 - {x_2''}^0 \;\geq\; -|\mathbf{x}_2' - \mathbf{x}_2''| \quad \Leftrightarrow \quad {x_2''}^0 - {x_2'}^0 \;\leq\; |\mathbf{x}_2' - \mathbf{x}_2''|. \tag{38} $$

Note that relation (iv) implies that there are only two cases:

$$ \text{(a) } {x_2'}^0 - {x_2''}^0 \geq |\mathbf{x}_2' - \mathbf{x}_2''| \quad \text{or} \quad \text{(b) } {x_2''}^0 - {x_2'}^0 \geq |\mathbf{x}_2' - \mathbf{x}_2''|. $$

The crucial point now is that (38) contradicts (b) while being compatible with (a). As (a) and (b) are mutually exclusive, this implies (a) which in particular establishes ${x_2'}^0 \geq {x_2''}^0$.

The fact that ${x_2'}^0 > {x_2''}^0$ implies ${x_1'}^0 > {x_1''}^0$ follows from the same consideration and the fact that ${x_1''}^0 > {x_1'}^0$ is equivalent to ${x_2''}^0 > {x_2'}^0$ by exchanging x_k' with x_k'' for $k = 1, 2$. □

Remark. In special cases, the Feynman propagator S^F can also be used to propagate ψ^{free} (see [27, chap. 6.1]), e.g., for propagating positive energy wave functions towards the future. However, it does not vanish outside of the light cone. Since this property of S^{ret} is an integral part of the proof, the latter cannot be extended to the case of $G_k = S_k^F$. This suggests that probability conservation does not hold for the Bethe-Salpeter equation, in agreement with the literature.

4.2 Asymptotic Probability Conservation for Symmetric Green's Fns

One may wonder if probability conservation can be established for different classes of interaction kernels besides retarded ones. We now prove such a result, albeit a weaker one, for the case of symmetric Green's functions $G_k = S_k^{\text{sym}}$ with

$$S^{\text{sym}}(x - x') = \frac{1}{2}\varepsilon(x^0 - x'^0)S(x - x') \tag{39}$$

where $\varepsilon(y) = +1$ if $y \geq 0$ and $\varepsilon(y) = -1$ else.

Proposition 2 *Consider the integral Eq.* (1) *with symmetric Green's functions,* $G_k = S_k^{\text{sym}}$, $k = 1, 2$ (39)*, and interaction kernels with the matrix symmetry property* (4)*. Let* Σ_t *be an equal-time surface in any given Lorentz frame. Then for every solution* ψ *of* (1) *on* $\mathbb{R}^4 \times \mathbb{R}^4$*, the following statement* (asymptotic probability conservation) *holds true:*

$$\lim_{t\to-\infty} P(\psi, \Sigma_t) = \lim_{t\to+\infty} P(\psi, \Sigma_t). \tag{40}$$

Proof We proceed similarly to the retarded case. The main difference is that the Heaviside functions θ get replaced by $\frac{1}{2}\varepsilon$. As we focus on equal-time surfaces Σ_t, the propagation identities (10) and (11) can be used directly. One obtains:

$$P(\psi, \Sigma_t) = P(\psi^{\text{free}}, \Sigma_t) + P_1(\psi, \Sigma_t), \quad \text{with} \tag{41}$$

$$\begin{aligned} P_1(\psi, \Sigma_t) = \frac{1}{16}\int d^4x_1' \, d^4x_2' \, d^4x_1'' \, d^4x_2'' \; \overline{\psi}(x_1', x_2')K(x_1', x_2')S_1(x_1' - x_1'')S_2(x_2' - x_2'') \\ \times K(x_1'', x_2'')\psi(x_1'', x_2'')\Big[-\varepsilon(t - x_1'^0)\varepsilon(t - x_2'^0)\varepsilon(t - x_1''^0)\varepsilon(t - x_2''^0) \\ + \Big(\varepsilon(t - x_1'^0)\varepsilon(t - x_2'^0) + \varepsilon(t - x_1''^0)\varepsilon(t - x_2''^0)\Big)\,\varepsilon(x_1'^0 - x_1''^0)\varepsilon(x_2'^0 - x_2''^0)\Big] \end{aligned} \tag{42}$$

Only the square bracket depends on t–and is not constant in t, as can be seen by case differentiation. However, taking $t \to \pm\infty$ both leads to the same result

$$\lim_{t\to\pm\infty} [...] = -1 + 2\varepsilon(x_1'^0 - x_1''^0)\varepsilon(x_2'^0 - x_2''^0). \tag{43}$$

$$\text{Thus, we find:} \quad \lim_{t\to-\infty} P_1(\psi, \Sigma_t) = \lim_{t\to+\infty} P_1(\psi, \Sigma_t). \tag{44}$$

Interestingly, the term P_1 is non-zero here, in contrast to the retarded case.[3] Nevertheless, together with probability conservation for ψ^{free}, (44) allows us to deduce:

$$\begin{aligned}\lim_{t\to-\infty} P(\psi, \Sigma_t) &= P(\psi^{\text{free}}, -\infty) + P_1(\psi, -\infty) \\ &= P(\psi^{\text{free}}, +\infty) + P_1(\psi, +\infty) \\ &= \lim_{t\to+\infty} P(\psi, \Sigma_t) P(\psi, \Sigma_t). \end{aligned} \tag{45}$$

□

4.3 Implications for Local Conservation Laws

It is a well-known fact that in the single-particle case global probability conservation on all Cauchy surfaces, $\int_\Sigma d\sigma_\mu(x)\, j^\mu(x) = \int_{\Sigma'} d\sigma_\mu(x)\, j^\mu(x)\ \forall\, \Sigma, \Sigma'$, implies local probability conservation. This can be shown by constructing a small volume V around a space-time point x which is enclosed between two otherwise overlapping Cauchy surfaces Σ, Σ' (see Fig. 2) and then using the divergence theorem. As the enclosing volume V can be made arbitrarily small, it follows that $\partial_{x^\mu}\, j^\mu(x) = 0$.

Proposition 1 establishes global probability conservation on all Cartesian products $\Sigma_1 \times \Sigma_2$ of Cauchy surfaces, enabling a similar reasoning. Applying the argument to x_k and Σ_k yields

$$\int_{\Sigma_{3-k}} d\sigma_{\mu_{3-k}}(x_{3-k})\, \partial_{x_k^{\mu_k}} \left[\overline{\psi}(x_1, x_2)\gamma_1^{\mu_1}\gamma_2^{\mu_2}\psi(x_1, x_2)\right] = 0\ \ \forall \Sigma_{3-k}, \quad k = 1, 2. \tag{46}$$

Now, we apply the argument another time for x_{3-k} and Σ_{3-k} and obtain the result:

Proposition 3 *Let ψ be a solution of the integral Eq. (1) with $G_k = G_k^{\text{ret}}$, $k = 1, 2$ and $K = K^{\text{ret}}$ (6). Then (46) is satisfied. In addition, we have:*

$$\partial_{x_1^\mu}\partial_{x_2^\nu}\, \overline{\psi}(x_1, x_2)\gamma_1^\mu\gamma_2^\nu\psi(x_1, x_2) = 0. \tag{47}$$

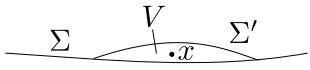


Fig. 2 Geometric construction for proving that global probability conservation on all Cauchy surfaces implies local probability conservation in the single-particle case

[3] In the symmetric case ψ^{free} does not need to agree with ψ in the infinite past nor in the infinite future [22].

Remarks.

1. Eq. (46) for $k = 1, 2$ implies global probability conservation on all space-like Cauchy surfaces (as can be seen using the divergence theorem), i.e., both notions are equivalent. Equation (47), however, is weaker. Its physical meaning is not clear.
2. (46) for $k = 1, 2$ *does not imply local probability conservation*. The reason is that even though (46) holds for all Σ_2, one cannot conclude that the integrand vanishes, as it might take negative values. This leads to the conclusion that *global probability conservation on all Cartesian products* $\Sigma \times \Sigma$ *is not equivalent to local probability conservation.*
3. To make it even clearer that local probability conservation does not hold, we calculate the four-divergence of the tensor current with respect to x_1:

$$\partial_{x_1^\mu} \left[\overline{\psi}\gamma_1^\mu \gamma_2^\nu \psi\right](x_1, x_2) = -2\Re\Big[\overline{\psi}(x_1, x_2)\gamma_2^\nu \int d^4x_2'\; S_2^{\text{ret}}(x_2 - x_2') K^{\text{ret}}(x_1, x_2')\psi(x_1, x_2')\Big]. \tag{48}$$

   Despite the spatio-temporal restrictions which S_2^{ret} and K_2^{ret} imply and the fact that it might be sufficient to restrict the derivation to space-like configurations $(x_1, x_2) \in \mathscr{S}$, the right hand side of (48) does not vanish in general.
4. Concerning the physical meaning of Eqs. (46) for $k = 1, 2$, one can rewrite them as:

$$\partial_{k,\mu} j_k^\mu(x_k, \Sigma) = 0 \;\; \forall\, \Sigma, \;\; k = 1, 2 \tag{49}$$

$$\text{where} \quad j_k^\mu(x_k, \Sigma) = \int_\Sigma d\sigma_\nu(x_{3-k})\, \overline{\psi}(x_1, x_2)\gamma_1^\mu \gamma_2^\nu \psi(x_1, x_2), \;\; k = 1, 2 \tag{50}$$

   One could try to use these currents to naively calculate the probability for the position of one particle, disregarding the position of the other. Viz, one might guess that the probability for particle k to be found in a small volume $d\sigma(x_k)$ around $x_k \in \Sigma$ is given by

$$\mathbb{P}(x_k \in d\sigma(x_k)) = j_k^\mu(x_k, \Sigma) n_\mu(x_k) d\sigma(x_k). \tag{51}$$

   where $n_\mu(x)$ is the future-directed unit normal vector field at $x \in \Sigma$.

5 Conclusion

Here it was shown that global probability conservation on all Cauchy surfaces holds for certain classes of multi-time integral equations. To make progress, it was necessary to deviate from the conventional wisdom about relativistic quantum-

mechanical integral equations given by the theory about the Bethe-Salpeter equation. The strongest result was obtained for retarded Green's functions $G_k = S_k^{\mathrm{ret}}$, $k = 1, 2$ and retarded interaction kernels $K = K^{\mathrm{ret}}$ (6). While retarded interactions are common in classical electrodynamics, a word of caution seems in order here. Using $K^{\mathrm{ret}}(x_1, x_2)$ implies that $x_1^0 > x_2^0$ has to hold for a configuration (x_1, x_2) to contribute to the interaction term. This seems unnatural as it breaks the symmetry between the particle labels. Related to this, for S_k^{ret}, K^{ret} and on space-like configurations $(x_1, x_2) \in \mathscr{S}$ one has

$$(-i\gamma_2^\nu \partial_{x_2^\nu} + m_2)\psi(x_1, x_2) = i \int d^4x_1' \, S_1^{\mathrm{ret}}(x_1 - x_1') K^{\mathrm{ret}}(x_1', x_2)\psi(x_1', x_2) = 0 \tag{52}$$

and therefore $\partial_{x_2^\nu} \overline{\psi}(x_1, x_2)\gamma_1^\mu \gamma_2^\nu \psi(x_1, x_2) = 0 \quad \forall (x_1, x_2) \in \mathscr{S}$. Equation (52) can be shown as follows. Consider the conditions for a term $\psi(x_1', x_2)$ to contribute to the integral. On the one hand, $K^{\mathrm{ret}}(x_1', x_2) \neq 0$ implies that x_2 lies on the past light cone of x_1'. On the other hand, $S_2^{\mathrm{ret}}(x_1 - x_1') \neq 0$ implies that x_1 lies in the future of x_1'. This is, however, incompatible with $(x_1, x_2) \in \mathscr{S}$.

Thus, on $\mathscr{S}$, one can take (52) to express that particle 2 is moving freely (while particle 1 is not) and interpret the interaction term in (1) as a single-sided action of particle 2 on particle 1.[4] This strengthens the concern that this type of interaction is physically not natural. In my opinion, the resulting dynamics represents a toy example and a first step towards a more natural result in the future, e.g., for K^{sym} (5).

In view of this situation, the second result, asymptotic probability conservation for the integral equation with symmetric Green's functions, $G_k = S_k^{\mathrm{sym}}$, $k = 1, 2$ seems particularly important. While weaker than probability conservation on all Cauchy surfaces, it holds for *arbitrary interaction kernels* respecting the basic symmetry property (4).

Moreover, it was shown that global probability conservation on all Cartesian products $\Sigma_1 \times \Sigma_2$ of Cauchy surfaces is equivalent to the semi-local property (46). Local probability conservation, however, does seem not hold.[5] While perfectly logical, this fact may seem surprising since in the single-particle case global probability conservation on all Cauchy surfaces Σ is equivalent to local probability conservation. Thus, we found that, *in the N-particle case, local probability conservation is stronger than global probability conservation on all (Cartesian products of) N Cauchy surfaces*.

In the future, it would be interesting to investigate if physical meaning can be given to the semi-local conservation law (46). For example, one may wonder if these properties are helpful to construct relativistic Bohmian laws of motion along the lines

[4] On configurations $(\mathbb{R}^4 \times \mathbb{R}^4) \setminus \mathscr{S}$, however, which the probability integral on $\Sigma_1 \times \Sigma_2$ for $\Sigma_1 \neq \Sigma_2$ uses, one in general has $(-i\gamma_2^\nu \partial_{x_2^\nu} + m_2)\psi(x_1, x_2) \neq 0$.

[5] Equation (52) implies $\partial_{x_2^\nu} \overline{\psi}(x_1, x_2)\gamma_1^\mu \gamma_2^\nu \psi(x_1, x_2) = 0$ only on space-like configurations $(x_1, x_2) \in \mathscr{S}$ while $\partial_{x_1^\mu} \overline{\psi}(x_1, x_2)\gamma_1^\mu \gamma_2^\nu \psi(x_1, x_2) \neq 0$ in general on $\mathbb{R}^4 \times \mathbb{R}^4$.

of [2] or [29], or to any other theory in the foundations of quantum mechanics which avoids the measurement problem. That would be in Detlef Dürr's spirit.

During the time when he was my PhD adviser, Detlef Dürr expressed that he thought further progress in the foundations of relativistic quantum theory required a clear and simple, mathematically solid underlying equation (free of the divergences that plague quantum field theory), be it only for a toy example. Once found, one could hope that such an equation would provide further guidelines for constructing a relativistic law of motion for Bohmian particles, as the Schrödinger and Dirac equations do in non-relativistic QM and relativistic single-particle QM, respectively. Perhaps, the integral equation discussed here can provide a starting point for such a consideration.

Overall, what has been achieved? I would say, a new way of constructing an interacting, relativistic equation which is now demonstrably compatible with global probability conservation. Do the examples provided constitute a full theory of relativistic quantum physics? Certainly not. Alas, one may conclude with a twist on the words of Feynman from the introduction:

> We now have a satisfactory relativistic quantum-mechanical model, one that doesn't agree with nature, but, at least, agrees with the logic that the sum of probability of all alternatives has to be 100%.

Disclaimer. This article reflects my personal work and does not represent scientific standpoints of my employer (Marvel Fusion GmbH).

Acknowledgements I would like to thank Siddhant Das and Roderich Tumulka for helpful discussions as well as Sascha Lill and an anonymous referee for constructive comments on the manuscript. Special thanks go to Markus Nöth for valuable remarks and a careful reading of an earlier version of the argument.

References

1. K. Berndl, D. Dürr, S. Goldstein, N. Zanghi, EPR-Bell Nonlocality, Lorentz invariance, and Bohmian quantum theory. Phys. Lett. A **53**, 2062–2073 (1996). https://arxiv.org/abs/quant-ph/9510027
2. D. Dürr, S. Goldstein, K. Münch-Berndl, N. Zanghi, Hypersurface Bohm-Dirac models. Phys. Rev. A **60**, 2729–2736 (1999). https://arxiv.org/abs/quant-ph/9801070v2
3. M. Lienert, S. Petrat, R. Tumulka, *Multi-time Wave Functions–An Introduction* (Springer, 2020)
4. E.E. Salpeter, H.A. Bethe, A relativistic equation for bound-state problems. Phys. Rev. **84**, 1232–1242 (1951)
5. G.C. Wick, Properties of Bethe-Salpeter wave functions. Phys. Rev. **96**, 1124–1134 (1954)
6. N. Nakanishi, A general survey of the theory of the Bethe-Salpeter equation. Prog. Theor. Phys. Suppl. **43**, 1–81 (1969)
7. K. Schwarzschild, Zur Elektrodynamik. II. Die elementare elektrodynamische Kraft. Göttinger Nachrichten **1903**, 132–141 (1903). http://resolver.sub.uni-goettingen.de/purl?PPN252457811_1903
8. H. Tetrode, Über den Wirkungszusammenhang der Welt. Eine Erweiterung der klassischen Dynamik. Z. Physik **10**, 317–328 (1922)

9. A.D. Fokker, Ein invarianter Variationssatz für die Bewegung mehrerer Massenteilchen. Z. Physik **58**, 386–393 (1929)
10. J.A. Wheeler, R.P. Feynman, Interaction with the absorber as the mechanism of radiation. Rev. Mod. Phys. **17**, 157–181 (1945)
11. J.A. Wheeler, R.P. Feynman, Classical electrodynamics in terms of direct interparticle action. Rev. Mod. Phys. **21**, 425–433 (1949)
12. G. Bauer, D.A. Deckert, D. Dürr, On the existence of dynamics in Wheeler-Feynman electromagnetism. Zeitschrift für angewandte Mathematik und Physik **64**, 1087–1124 (2013)
13. D.-A. Deckert, D. Dürr, N. Vona, Delay equations of the Wheeler–Feynman type. J. Math. Sci. **202**(5), 623–636 (2014). arXiv.org/abs/1212.6285
14. G. Bauer, D.-A. Deckert, D. Dürr, G. Hinrichs, Global solutions to the electrodynamic two-body problem on a straight line. Zeitschrift für angewandte Mathematik und Physik **68**(3), 71 (2017). arXiv.org/abs/1603.05115
15. P.C.W. Davies, A quantum theory of Wheeler-Feynman electrodynamics. Math. Proc. Camb. Philos. Soc. **68**(3), 751–764 (1970)
16. F. Hoyle, J.V. Narlikar, Cosmology and action-at-a-distance electrodynamics. Rev. Mod. Phys. **67**, 113–155 (1995)
17. A.O. Barut, The covariant many-body problem in quantumelectrodynamics. J. Math. Phys. **32**, 1091–1095 (1991)
18. G. Ludwig, Ansatz zu einer divergenzfreien Quantenelektrodynamik. Z. Naturforschung **5a**, 637–641 (1950)
19. D.-A. Deckert, Electrodynamic absorber theory–a mathematical study. Ph.D. thesis, Mathematisches Institut, Ludwig-Maximilians-Universität München (2009). http://nbn-resolving.de/urn:nbn:de:bvb:19-114215
20. M. Lienert, Direct interaction along light cones at the quantum level. J. Phys. Math. Theor. **51**(43), 435302 (2018). arXiv.org/abs/1801.00060
21. M. Lienert, R. Tumulka, A new class of Volterra-type integral equations from relativistic quantum physics. J. Integr. Equ. Appl. **31**(4), 535–569 (2019). arXiv.org/abs/1803.08792
22. M. Lienert, R. Tumulka, Interacting relativistic quantum dynamics of two particles on space-times with a Big Bang singularity. J. Math. Phys. **60**, 042302 (2019). arXiv.org/abs/1805.06348
23. M. Lienert, M. Nöth, Existence of relativistic dynamics for two directly interacting Dirac particles in 1+3 dimensions. Rev. Math. Phys. **33**, 2150023 (2021). arXiv.org/abs/1903.06020
24. M. Lienert, M. Nöth, Singular light cone interactions of scalar particles in 1+3 dimensions. arXiv.org/abs/2003.08677
25. R.P. Feynman, The development of the space-time view of quantum electrodynamics. Nobel Lect. (1965). http://www.nobelprize.org/nobel_prizes/physics/laureates/1965/feynman-lecture.html
26. I. Bialynicki-Birula, Z. Bialynicka-Birula, *Quantum Electrodynamics* (Pergamon Press, 1975)
27. W. Greiner, J. Reinhardt, *Quantum Electrodynamics*, 4th edn. (Springer, 2009)
28. M. Lienert, R. Tumulka, Born's rule for arbitrary Cauchy surfaces. Lett. Math. Phys. **110**(4), 753–804 (2020). https://doi.org/10.1007/s11005-019-01239-y
29. S. Goldstein, R. Tumulka, Opposite arrows of time can reconcile relativity and nonlocality. Cl. Quantum Gravity **20**(3), 557 (2003). arXiv.org/abs/quant-ph/0105040v5

Vacuum Polarisation Without Infinities

Dirk-André Deckert, Franz Merkl, and Markus Nöth

Abstract In honour of Detlef Dürr, we report on a mathematical rigorous computation of the electric vacuum polarisation current and extract the well-known expression for the second order perturbation. Intermediate steps in the presented calculation demonstrate, to the knowledge of the authors for the first time, mathematical rigorous versions of the combined dimensional and Pauli-Villars regularisation schemes. These are employed as computational tools to infer convenient integral representations during the computation. The said second order expression is determined up to a remaining degree of freedom of a real number–without ill-defined terms from start to end.

1 Introduction

The definition and the original computations of the electric vacuum polarisation current in quantum electrodynamics (QED), based on the pioneering works of Dirac, Heisenberg, and others, go way back to Schwinger, Feynman, and Dyson, and, in its original form, may still be best accessible from Dyson's manuscript [9]. Today, these are contained in various nuances in almost every textbook on advanced quantum mechanics. It might therefore appear that this topic has long been settled. Given that QED is such an important theory for the human understanding of nature and a century has passed, it better should have. But, at least mathematically, it has not. All of these computations start with an ill-defined equation of motion for the electric quantum vacuum, or worse, an ill-defined and far fetched scattering matrix coefficient, derive a similarly ill-defined expression for the electric current which is then massaged by means of several informal manipulations, such as "subtracting" an ill-defined zero-field electric current, "dropping" ill-defined expressions that do not appear gauge invariant, "introducing" diverging counter terms to absorb remaining infinities into the bare electric charge constant, etc. For everyone who has had the opportunity to get to know Detlef, it probably goes without saying that he was very unsatisfied

D.-A. Deckert (✉) · F. Merkl · M. Nöth
Mathematisches Institut, Ludwig-Maximilians-Universität München, München, Germany
e-mail: deckert@math.lmu.de

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_18

with this state of affairs. At the same time, of course, he could not have cared less whether the consensus opinion of the scientific community was otherwise. In fact, beside his renown work in the foundations of quantum mechanics, he dedicated a substantial part of his scientific work to the understanding of classical and quantum electrodynamics. In 2006, Detlef, Martin Schottenloher, D.-A.D. and F.M. started a seminar series on the mathematics of QED based on the books [9, 33, 35, 42], articles [8, 10, 34, 36, 39, 43], among others, entirely in the spirit of Dirac's quote [23, p. 184]:

> I must say that I am very dissatisfied with the situation, because this so-called *good theory* does involve neglecting infinities which appear in its equations, neglecting them in an arbitrary way. This is just not sensible mathematics. Sensible mathematics involves neglecting a quantity when it is small—not neglecting it just because it is infinitely great and you do not want it! *– Dirac, 1975*

which initiated the works [4–7]. Many years later, we are now under the impression of having a mathematical rigorous as well as non-perturbative understanding of this computation as well as the definition of the corresponding time evolution and scattering matrix for QED in an external, classical electrodynamic field, which will be published in a series of forthcoming articles. This first article shall provide an introduction by treating only the second order of perturbation of the electric vacuum polarisation current in our approach. A slightly different but also rigorous computation of the second order of perturbation was already done by Scharf [35] while the non-perturbative computation in [37] seems incomplete.

Along the way we have learned about many other series of works on various aspects of the mathematical rigorous description of the quantum vacuum, three of which we would like to mention here: First, the works of Mickelsson and collaborators [24, 26–28] developing a bundle theoretic, conceptual geometric theory of the phases in quantum field theory. Second, the works [29] and [31, 32] on adiabatic electron-positron pair-creation in an external field. Third, the long series of articles by Gravejat, Hainzl, Lewin, Séré, and Solovej which mainly studies the stationary solutions of a non-linear model of the quantum vacuum, among them [17, 18, 20–22] and, in particular, the overview in [25]. Those models are not only able to describe the polarisation of the quantum vacuum by an external field but also the back reaction of its quantum expectation and make contact to effective dynamics in terms of the Heisenberg-Euler Lagrangian [19]. Third, an entirely different approach was developed by Finster and his group which constructs the interacting fermionic projectors by means of a variational principle [11–16] and from which it was shown that many aspects of quantum field theory, also beyond QED, can be derived rigorously. Lastly, viewing QED in external fields from the perspective of algebraic quantum field theory in curved space-time [1, 2], the study of Hadamard states, see, e.g., [38], can be seen as a bridge between the algebraic formulation, causal fermion systems, and the approach followed here.

2 Bogoliubov's Electric Current Formula

Before choosing a second-quantised expression of the electric charge current, for whatever that entails physically, it is illustrative to discuss corresponding expressions in the one-particle quantum theory. For this purpose, we regard a one-particle Schrödinger evolution of the form

$$i\,\partial_t \psi(t) = H_A(t)\psi(t) \tag{1}$$

for times $t \in \mathbb{R}$, wave functions $t \mapsto \psi(t)$ with values in a Hilbert space $\mathcal{H}$, and an external, classical, time-dependent four-vector potential $A = (A_\mu)_{\mu=0,1,2,3} = (A_0, -\mathbf{A}) \in \mathcal{C}_c^\infty(\mathbb{R}^4, \mathbb{R}^4)$. In the whole paper, we impose natural units $1 = \hbar = c = \mu_0 = \epsilon_0$ which imply that charges are dimensionless, masses, momenta, energies and four-vector potentials have the dimension 1/length while the Fourier transformed four-vector potentials carry dimension $(\text{length})^3$. Furthermore, $(H_A(t))_{t\in\mathbb{R}}$ shall denote a family of Hamiltonians, i.e., possibly unbounded, self-adjoint operators with domain $D(H_A(t)) \subseteq \mathcal{H}$. In the physically relevant cases regarded below, the latter domain will be independent of time and coincide with the domain of the time-independent free Hamiltonian $H_0 = H_A|_{A=0}$. Suppose, (1) generates a strongly continuous unitary time-evolution $U_A = (U_A(t_1, t_0))_{t_0,t_1\in\mathbb{R}}$ on $\mathcal{H}$, such that for a given initial value $\psi(t_0)$ in a suitable domain at time t_0, the corresponding unique solution to (1) is given by the map $t \mapsto \psi(t) := U_A(t, t_0)\psi$, one may introduce the scattering operator

$$S_A = U_0(0, t_1)U_A(t_1, t_0)U_0(t_0, 0)\,, \tag{2}$$

where $t_0 \ll 0$ and $t_1 \gg 0$ are taken so small and large, respectively, that the temporal support of A is contained in the interval (t_0, t_1), and $U_0 = U_A|_{A=0}$ is short for the corresponding free time-evolution.

Using the notation $\partial_F g(F) = \frac{d}{d\epsilon} g(\epsilon F)|_{\epsilon=0}$ and viewing the charge current as being generated in response to a perturbation of the four-vector potential A, one may define the evaluation of the electric current at a test function $F = (F_0, -\mathbf{F}) \in \mathcal{C}_c^\infty(\mathbb{R}^4, \mathbb{R}^4)$ as the operator expression

$$J_A(F) := ie S_A^{-1}\partial_F S_{A+F}\,, \tag{3}$$

which we refer to as Bogoliubov's formula of the electric current. On a suitable domain, this derivative can be evaluated by employing the comparison of the dynamics of $A + F$ and A, i.e.,

$$\partial_F U_{A+F}(t_1, t_0) = -i \int_{t_0}^{t_1} ds\, U_A(t_1, s)\partial_F H_{A+F}(s)U_A(s, t_0)\,, \tag{4}$$

for times $t_0, t_1 \in \mathbb{R}$, and for which the appearance of the unbounded operators on the right-hand side will turn out unproblematically in the physical relevant cases regarded below. Provided the family of Hamiltonians $(H_A(t))_{t\in\mathbb{R}}$ are sufficiently regular in A, formulas in (3) and (4) imply for $t \ll 0$ earlier than the temporal support of A

$$J_A(F) = \int_{-\infty}^{+\infty} ds\, U_0(0,t)U_A(t,s)\partial_F H_{A+F}(s)U_A(s,t)U_0(t,0)\,. \tag{5}$$

Below we state the well-known evaluations of (5) for a minimally-coupled, non-relativistic, charged spin-0 Schrödinger particle and a relativistic Dirac particle both of electric charge $-e < 0$ and in an external four-vector potential A:

1. Schrödinger case, i.e., for $H_A(t) = H_A^S(t) := \frac{1}{2m}(-i\nabla + e\mathbf{A}(t))^2 - eA^0(t)$:

$$j_A(F) := \langle\psi, J_A(F)\psi\rangle = \int_{\mathbb{R}\times\mathbb{R}^3} dt\, d\mathbf{x}\, [\rho(t,\mathbf{x})F_0(t,\mathbf{x}) - \mathbf{j}(t,\mathbf{x})\cdot\mathbf{F}(t,\mathbf{x})]\,,$$
$$\rho(t,\mathbf{x}) = -e|\psi(t,\mathbf{x})|^2, \quad \text{and} \quad \mathbf{j}(t,\mathbf{x}) = -\frac{e}{m}\,\mathrm{Im}\,\psi(t,\mathbf{x})^*(\nabla + ie\mathbf{A}(t,\mathbf{x}))\psi(t,\mathbf{x})\,. \tag{6}$$

2. Dirac case, i.e., for $H_A(t) = H_A^D(t) := \boldsymbol{\alpha}\cdot(-i\nabla + e\mathbf{A}(t)) - eA^0(t) + \beta m$:

$$j_A(F) := \langle\psi, J_A(F)\psi\rangle = \int_{\mathbb{R}^4} j^\mu(x)F_\mu(x)d^4x, \qquad j^\mu(x) = -e\overline{\psi}(x)\gamma^\mu\psi(x)\,. \tag{7}$$

In conclusion, the electric currents of respective one-particle theories can be recovered from Bogoliubov's single current formula (3). By virtue of its generality, we will employ it as starting point to infer a second-quantised version.

3 Electric Current of a Dirac Sea in an External Field

In what follows, U shall denote a unitary operator on one-particle Hilbert space $\mathcal{H}$. Assuming some familiarity of the reader with the well-known second-quantisation formalism of the Dirac field, e.g. [42], we avoid a lengthy introduction and only caricature the respective Fock space vacuum expectation values. Let P_+, P_- denote the orthogonal projections of the one-particle Hilbert space onto the positive and negative spectral subspaces $\mathcal{H}^+, \mathcal{H}^-$ of the free Dirac Hamiltonian H_0^D, respectively. The corresponding splitting of $\mathcal{H}$ is denoted by $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C}^4) = \mathcal{H}^+ \oplus \mathcal{H}^-$ and the vacuum vector Ω in standard Fock space is taken to represent the Dirac sea of $\mathcal{H}^-$. As a first guess, one might try to define the second-quantised vacuum expectation value

$\langle\Omega, \widetilde{U}\Omega\rangle$ of a second-quantised version $\widetilde{U}$ of U by means of an infinite-dimensional Slater determinant, i.e., $\langle\Omega, \widetilde{U}\Omega\rangle \stackrel{?}{=} \det U_-\big|_{\mathcal{H}^-\to\mathcal{H}^-}$ for $U_- := P_- U P_-$. Interpreting the determinant as a Fredholm determinant and using the notation $I_1(\mathcal{H})$ and $I_2(\mathcal{H})$ to denote the trace class and Hilbert-Schmidt ideals, respectively, this would require having $U_- \in \mathrm{id}_- + I_1(\mathcal{H}^-)$. But this does not hold in general, hence, the overset '?'. A way to fix this is to utilise another unitary operator R_U on $\mathcal{H}^-$, in our notation $R_U \in U(\mathcal{H}^-)$, interpreted as a base change in the vacuum, to define

$$\left\langle\Omega, \widetilde{U}\Omega\right\rangle := \det U_- R_U\big|_{\mathcal{H}^-\to\mathcal{H}^-}\,. \tag{8}$$

This can indeed be shown to be the vacuum expectation value of some unitary lift $\widetilde{U}$ of the unitary one-particle operator U on $\mathcal{H}$ to some Fock space; cf. [4]. In the special case of

$$U_{+-} := P_+ U P_-,\ U_{-+} := P_- U P_+ \in I_2(\mathcal{H}), \qquad \|\,\mathrm{id} - U\| < 1 \Rightarrow U^*{}_- \in \mathrm{GL}(\mathcal{H}^-)\,, \tag{9}$$

it is possible to construct such an R along the following lines. First, we observe that

$$U_- U^*{}_- = (UU^*)_- - U_{-+}U^*{}_{+-} = \mathrm{id}_- - |U_{-+}|^2 \in \mathrm{id}_- + I_1(\mathcal{H}^-) \tag{10}$$

since the product of two Hilbert-Schmidt operators is in the trace class. Furthermore, the expression (10) is positive definite thanks to (9). However, $U^*{}_-$ is in general not unitary. Therefore, we employ a polar decomposition $U^*{}_- = R_U|U^*{}_-|$ with radial part $|U^*{}_-| = \sqrt{U_- U^*{}_-} \in \mathrm{id}_- + I_1(\mathcal{H}^-)$ having a positive Fredholm determinant. Exploiting the invertibility of $U^*{}_-$,

$$R_U = U^*{}_- |U^*{}_-|^{-1} \in U(\mathcal{H}^-) \tag{11}$$

renders the right-hand side of (8) well-defined. Furthermore, since $R_U \in U(\mathcal{H}^-)$, it turns out that the corresponding lift will also be unitary on the underlying Fock space; cf. [4].

The lift $\widetilde{U}$ of U is known to be unique up to a phase $e^{i\varphi} \in U(1)$; cf. [4, 39]. Hence, given $U \in U(\mathcal{H})$ that fulfils (9), the pair (U, R_U) characterises the equivalence class of lifts $[U, R_U] = \{(U, R_U Q) \mid Q \in U(\mathcal{H}^-) \cap \left(\mathrm{id}_{\mathcal{H}^-} + I_1(\mathcal{H}^-)\right)\}$ while two pairs $(U, R_U Q)$ and $(U, R_U Q')$ correspond to the same lift if and only if $\det Q^{-1}Q' = 1$. The lift of U characterised by (U, R_U), i.e., $Q = \mathrm{id}_{\mathcal{H}^-}$, shall be denoted by $\overline{U}$. This implies $\left\langle\Omega, \overline{U}, \Omega\right\rangle > 0$ for (8). For two $U, U' \in U(\mathcal{H})$ so close to $\mathrm{id}_{\mathcal{H}}$ such that all three operators U, U' and UU' fulfil (9), and two pairs (U, R) and (U', R'), characterising lifts $\widetilde{U}$ and $\widetilde{U}'$, respectively, $(UU', R'R)$ characterises a lift of UU' since $(UU')_- R'R = U_- U'_- R'R + U_{-+}U'_{+-}R'R \in \mathrm{id}_{\mathcal{H}^-} + I_1(\mathcal{H}^-)$. Note the reversed order in the second component $R'R$. Moreover, (U^{-1}, R^{-1}) characterises the lift $\widetilde{U}^{-1}$ of U^{-1}. We refer the interested reader to [4] and also to [33] for the underlying mathematical theory. We fix a reference vector potential $A_{\mathrm{ref}} \in \mathcal{C}_c^\infty(\mathbb{R}^4, \mathbb{R}^4)$ for

the rest of the article. We shall only regard vector potentials $A \in \mathcal{C}_c^\infty(\mathbb{R}^4, \mathbb{R}^4)$ such that the corresponding one-particle scattering operators fulfil $\| S_{A_{\mathrm{ref}}}^{-1} S_A - \mathrm{id}_{\mathcal{H}} \| < 1$, as a global construction will not be needed here. Moreover, it is well-known, e.g., [4, 34], that S_{A+-}, $S_{A-+} \in I_2(\mathcal{H})$ hold true for any $A \in \mathcal{C}_c^\infty(\mathbb{R}^4, \mathbb{R}^4)$. In view of (9), this ensures well-definedness of the lift

$$\widetilde{S}_{A+F}^{A} = \overline{S_{A_{\mathrm{ref}}}^{-1} S_A}^{-1} \overline{S_{A_{\mathrm{ref}}}^{-1} S_{A+F}} \tag{12}$$

and allows a first attempt in defining a vacuum expectation of the current in the spirit of (3):

$$\widetilde{j}_A(F) = \left\langle \Omega, \widetilde{J}_A(F)\Omega \right\rangle = i\,\partial_F \left\langle \Omega, \widetilde{S}_{A+F}^{A}\Omega \right\rangle = \mathrm{Re}\, i\,\partial_F \left\langle \Omega, \widetilde{S}_{A+F}^{A}\Omega \right\rangle . \tag{13}$$

We note that since $\widetilde{S}_{A+F}^{A}$ is a unitary lift on the underlying Fock space which depends smoothly on F in the relevant norm; cf. [4], the expression in the centre of (13) is real-valued.

However, taking different lifts $e^{i\varphi_A} \overline{S_{A_{\mathrm{ref}}}^{-1} S_A}$, for an arbitrary, A-dependent $\varphi_A \in \mathbb{R}$, gives

$$\widehat{S}_{A+F}^{A} = \left(e^{i\varphi_A} \overline{S_{A_{\mathrm{ref}}}^{-1} S_A} \right)^{-1} \left(e^{i\varphi_{A+F}} \overline{S_{A_{\mathrm{ref}}}^{-1} S_{A+F}} \right)$$

and yet another corresponding current

$$\widehat{j}_A(F) = i\,\partial_F \left\langle \Omega, \widehat{S}_{A+F}^{A}\Omega \right\rangle = \widetilde{j}_A(F) - d\varphi_A(F) . \tag{14}$$

It is therefore the task to select physically relevant candidates for the physical current j among those $\widehat{j}$ for the various phases φ. Morally, this non-uniqueness in the choice of the current reflects the ill-definedness of the current in the traditional formulation of QED.

In order to characterise this degree of freedom geometrically, we observe that the exterior derivative $c := d\widehat{j}$ does not dependent on the choice of the phase φ, i.e.,

$$d\widehat{j}_A(G, F) =: c_A(G, F) = d\widehat{j}_A(G, F) = \partial_G \widehat{j}_{A+G}(F) - \partial_F \widehat{j}_{A+F}(G) = d\widetilde{j}_A(G, F) \tag{15}$$

because $dd\varphi = 0$. By Poincaré's lemma and the fact that the space of permissible vector potentials A is star-shaped, the two-form c contains precisely the same information as the class of all $\widehat{j}$ with varying φ. By construction, the two-form c is closed, i.e., $dc = dd\widehat{j} = 0$. This will play a crucial role in the non-perturbative construction addressed in forthcoming papers.

The physically relevant j should now be selected with conditions C0–C4 in mind:

C0 *Exterior derivative*: Given (15), we require $dj = c$;

C1 *Causality:* For $F, G \in \mathcal{C}_c^\infty(\mathbb{R}^4, \mathbb{R}^4)$ such that the support of G does not overlap the closed causal past of the support of F, we require $\partial_G j_{A+G}(F) = 0$;

C2 *Relativistic invariance:* For any proper, orthochronous Lorentz-transformation Λ and any translation displacement $a \in \mathbb{R}^4$, we require $j_{\Lambda A}(\Lambda F) = j_A(F) = j_{A(\cdot - a)}(F(\cdot - a))$;

C3 *Gauge invariance:* For $\Gamma \in \mathcal{C}_c^\infty(\mathbb{R}^4, \mathbb{R})$, we require $j_{A+\partial\Gamma}(F) = j_A(F) = j_A(F + \partial\Gamma)$;

C4 *Reference current:* Should A_{ref} be sufficiently close to zero to allow for $A = 0$, then we require the vacuum expectation of the current to vanish in this case, i.e., $j_0(F) = 0$.

In order to derive an explicit expression for c for C0, cf. (15), we start by computing $\widetilde{j}_A(F)$ in (13). By (8), (11), and (12) we find

$$\begin{aligned}\widetilde{j}_A(F) = \;& \mathrm{Re}\, i\partial_F \det\left[(S_A^{-1}S_{A+F})_-(S_{A+F}^{-1}S_{A_{\text{ref}}})_-((S_A^{-1}S_{A_{\text{ref}}})_-)^{-1}\right] \\ & \times \det|(S_{A+F}^{-1}S_{A_{\text{ref}}})_-|^{-1} \det|(S_A^{-1}S_{A_{\text{ref}}})_-| \\ = \;& \mathrm{Re}\, i\partial_F \det\left[(S_A^{-1}S_{A+F})_-(S_{A+F}^{-1}S_{A_{\text{ref}}})_-({S_{A_{\text{ref}}}}^{-1}S_A)_-\right] \\ & \times \det|(S_{A+F}^{-1}S_{A_{\text{ref}}})_-|^{-1} \det|(S_A^{-1}S_{A_{\text{ref}}})_-|^{-1}\,, \end{aligned} \tag{16}$$

where we have used $|U_-|(U_-)^{-1} = |U_-|^{-1}(U^{-1})_-$ for $U = S_A^{-1}S_{A_{\text{ref}}}$ and have performed a cyclic permutation under the determinant by a corollary of Lidskii's theorem [40]. For $F = 0$ the product of the three determinants equals one because $\det(S_A^{-1}S_A)_-(S_A^{-1}S_{A_{\text{ref}}})_-({S_{A_{\text{ref}}}}^{-1}S_A)_-$ equals $\det|({S_{A_{\text{ref}}}}^{-1}S_A)_-|^2$. This allows to recast the above expression as $\widetilde{j}_A(F) = i\partial_F \log \Gamma_{S_{A+F}\, S_{A_{\text{ref}}}\, S_A}$, using the notation $\Gamma_{XYZ} := \arg\det[(Z^{-1}X)_-(X^{-1}Y)_-(Y^{-1}Z)_-]$. These terms Γ_{XYZ} have convenient properties summarised in the appendix in Lemma A.1. Exploiting those and the chain rule, the exterior derivative $c = d\widetilde{j}$, cf. (15), can be expressed as follows, using the notations $S_Y^X := S_X^{-1}S_Y = (S_X^Y)^{-1}$ and $\arg z := z/|z|$ for $z \in \mathbb{C} \setminus \{0\}$:

$$\begin{aligned} c_A(G,F) &= i\partial_G\partial_F \log \Gamma_{S_{A+F+G}\, S_{A_{\text{ref}}}\, S_{A+G}} - i\partial_F\partial_G \log \Gamma_{S_{A+F+G}\, S_{A_{\text{ref}}}\, S_{A+F}} \\ &= 2i\partial_F\partial_G \log \Gamma_{S_{A+F}\, S_{A_{\text{ref}}}\, S_{A+G}} \qquad \text{by Lem. A.1, prop. 2.} \\ &= 2i\partial_F\partial_G \log \Gamma_{S_{A+F}\, S_A\, S_{A+G}} \qquad \text{by Lem. A.1, prop. 4.} \\ &= 2i\partial_F\partial_G \log \arg\det[\,(S_{A+F}^{A+G})_-((S_{A+G}^{A+F})_- - (S_A^{A+F})_-+(S_{A+G}^A)_+-)\,] \\ &= 2i\partial_F\partial_G \log \arg\det[\mathrm{id}_- - |(S_{A+F}^{A+G})_+-|^2 - (S_{A+F}^{A+G})_-(S_A^{A+F})_-+(S_{A+G}^A)_+-\,] \\ &= -2\partial_F\partial_G \,\mathrm{Im}\,\mathrm{Tr}[(S_{A+F}^A)_-+(S_{A+G}^A)_+-] \end{aligned} \tag{17}$$

Recalling (15), in order to get our hands on a physically relevant candidate current j in the sense of above, we need to split c as follows

$$c_A(G,F) = \partial_G j_{A+G}(F) - \partial_F j_{A+F}(G)\,. \tag{18}$$

Note that the left-hand side is given as the non-perturbative expression (17). For the purpose of the splitting, we make the following ansatz. Suppose the vacuum

expectation value of the current $A \mapsto j_A$ is real analytic and has a power expansion of the form

$$j_A(F) = \sum_{n=2}^{\infty} j^{(n)}(F; \underbrace{A, \ldots, A}_{n-1 \text{ many}}), \tag{19}$$

where the n-th summand on the right-hand side is assumed to be linear in all n arguments and symmetric in the last $n-1$ arguments. The same summand corresponds to the $(n-1)$-st Taylor order in A. In view of condition C4 there is no $n = 1$ summand. This implies the expansion

$$c_A(G, F) = \sum_{n=2}^{\infty} c^{(n)}(G, F; \underbrace{A, \ldots, A}_{n-2 \text{ many}}), \qquad \text{for} \tag{20}$$

$$c^{(n)}(G, F; A, \ldots, A) = (n-1)\big(j^{(n)}(F; G, A, \ldots, A) - j^{(n)}(G; F, A, \ldots, A)\big). \tag{21}$$

In forthcoming works, we will justify the ansatz, i.e., the analyticity assumption, provide a non-perturbative form for the current and, on its basis, construct the corresponding scattering matrix and time-evolution. In this article, we will only demonstrate how to perform the splitting (18) for the lowest order term $n = 2$, in order to fulfil C0, and check the remaining C1–C4.

4 Second Order Perturbation Without Infinities

In view of C0, cf. (18) for the lowest order $n = 2$, we need to find an expression $j^{(2)}$ fulfilling

$$c^{(2)}(G, F) = j^{(2)}(F; G) - j^{(2)}(G; F) \tag{22}$$

together with conditions C1–C4. Not surprisingly, it will coincide with the well-known expression for the second order perturbation of the current of QED; e.g. [9]. However, in text-books, the latter is extracted from a mathematically non-rigorous computation involving infinities that are removed by hand. In what follows, we give this computation a mathematical sense.

We obtain the following finite Lebesgue integral for $c^{(2)}(G, F) = c_0(G, F)$, cf. (17), after employing (4) in order to compute the linearisation of S^A_{A+F}, and furthermore, a suitable partial integration in time to allow for the application of Fubini's theorem:

$$c^{(2)}(G,F) = i\,\partial_G\partial_F \operatorname{Tr}\left(P_- S^A_{A+F} P_+ S^A_{A+G} P_- - P_- S^A_{A+G} P_+ S^A_{A+F} P_-\right) = \int\limits_{\mathbb{R}^3\times\mathbb{R}^3} d^3p\, d^3q\, c^{(2)}(G,F;\mathbf{q},\mathbf{p})$$

$$\text{with } c^{(2)}(G,F;\mathbf{q},\mathbf{p}) := \sum_{\tau=\pm1} \tau\cdot(2\pi i)^2 \operatorname{Res}_{\substack{p^0=-\tau E(\mathbf{p}),\\ q^0=\tau E(\mathbf{q})}} \omega(G,F;q,p)\,,\ E(\mathbf{p}) := \sqrt{m^2+\mathbf{p}^2},$$

$$\omega(G,F;q,p) := \frac{ie^2}{(2\pi)^4}\operatorname{tr}[(\not p - m)^{-1}\not F(p-q)(\not q - m)^{-1}\not G(q-p)] = \omega(F,G;p,q)\,. \tag{23}$$

By abuse of notation, we denote the Fourier transform of $x \mapsto F_\mu(x)$ by the same symbol $k \mapsto F_\mu(k) = (2\pi)^{-4/2}\int d^4x\, e^{ik_\nu x^\nu} F_\mu(x)$. Furthermore, $\operatorname{Res}_{q^0=\cdot,\,p^0=\cdot}$ denotes the iterated residue operator.

In order to identify the current term $j^{(2)}$ in (22), we first regard the expression $\omega(q^0, p^0)$ for fixed $\mathbf{q}, \mathbf{p} \in \mathbb{R}^3$. In a first step, we add zero, written as a difference of two equal residues, i.e.,

$$\begin{aligned} c^{(2)}(G,F;\mathbf{q},\mathbf{p}) &= \sum_{\sigma=\pm1}(2\pi i)^2 \left(\operatorname{Res}_{\substack{p^0=-E(\mathbf{p}),\\ q^0=\sigma E(\mathbf{q})}} - \operatorname{Res}_{\substack{p^0=\sigma E(\mathbf{p}),\\ q^0=-E(\mathbf{q})}}\right)\omega(G,F;q,p) \\ &= (2\pi i)\Bigg(\int\limits_{[\mathbb{R}-i\delta]-[\mathbb{R}+i\delta]} dq^0 \operatorname{Res}_{p^0=-E(\mathbf{p})} - \int\limits_{[\mathbb{R}-i\delta]-[\mathbb{R}+i\delta]} dp^0 \operatorname{Res}_{q^0=-E(\mathbf{q})}\Bigg)\omega(G,F;q,p) \\ &= c^{(2)}_+(G;F;\mathbf{q},\mathbf{p}) - c^{(2)}_+(F;G;\mathbf{p},\mathbf{q}) \end{aligned} \tag{24}$$

for any fixed number $\delta > 0$ and

$$c^{(2)}_+(F;G;\mathbf{p},\mathbf{q}) := -2\pi i\Bigg(\int\limits_{[\mathbb{R}+i\delta]} dq^0 \operatorname{Res}_{p^0=-E(\mathbf{p})} + \int\limits_{[\mathbb{R}-i\delta]} dp^0 \operatorname{Res}_{q^0=-E(\mathbf{q})}\Bigg)\omega(F,G;p,q)\,. \tag{25}$$

The bracket notation $[\mathbb{R}\pm i\delta]$ denotes the standard parametrisation $\mathbb{R}\ni t\mapsto t\pm i\delta$ and their differences are understood as $\int_{[A]-[B]} = \int_{[A]} - \int_{[B]}$. We remark that $c^{(2)}_+$ has temporally causal support in the sense that, for $F, G \in \mathcal{C}^\infty_c(\mathbb{R}^4, \mathbb{R}^4)$ such that the support of F is temporally earlier than the support of G, we have $c^{(2)}_+(F;G;\mathbf{p},\mathbf{q}) = 0$. This can be seen by inserting the Fourier transformations of F and G in the time variable, observing that $\mathbb{C}\setminus\mathbb{R}\ni p^0\mapsto(\not p - m)^{-1}$ is holomorphic, and applying Cauchy's integral theorem in the limit $\delta\to+\infty$. Furthermore, it is convenient to introduce the substitution $k = (k^0,\mathbf{k}) := p - q$, i.e.,

$$c^{(2)}_+(F;G;\mathbf{p},\mathbf{q}) = -2\pi i \int\limits_{[\mathbb{R}-i\delta]} dk^0 \left(\operatorname{Res}_{p^0=-E(\mathbf{p})} + \operatorname{Res}_{p^0=k^0-E(\mathbf{p}-\mathbf{k})}\right)\omega(F,G;p,p-k) \tag{26}$$

Here, the contour $\mathcal{C}_{\mathrm{Wick}}(k)$ denotes any closed curve having winding number $+1$ around $p^0 = -E(\mathbf{p})$ and $p^0 = k^0 - E(\mathbf{p}-\mathbf{k})$, but winding number zero around $p^0 = +E(\mathbf{p})$ and $p^0 = k^0 + E(\mathbf{p}-\mathbf{k})$. We remark that the inner integral of (26), for fixed $\mathbf{p}, \mathbf{k}$ and taken as function of k^0, is holomorphic on

$$k^0 \in \mathbb{D} := \mathbb{C} \setminus ((-\infty, -2m] \cup [2m, \infty)). \tag{27}$$

In particular, for $k \in \mathbb{C}^4$ sufficiently close to zero, $\omega^{\mu\nu}(p,k)$ is well-defined for all $p \in i\mathbb{R} \times \mathbb{R}^3$. In this region the contour $\mathcal{C}_{\mathrm{Wick}}(k)$ can be replaced by $i\mathbb{R}$ oriented in positive imaginary direction, exploiting the $|p^0|^{-2}$ decay of $\omega^{\mu\nu}(p,k)$ for $|p^0| \to \infty$.

Considering (22), if $c_+^{(2)}(F;G;\mathbf{p},\mathbf{q})$ was integrable in $(\mathbf{p},\mathbf{q}) \in \mathbb{R}^3 \times \mathbb{R}^3$, its integral would have been a natural candidate for $j^{(2)}(F;G)$ due to the support properties discussed above. But this is in general not the case. To nevertheless find a suitable candidate by leveraging the knowledge of $c_+^{(2)}$, we take a different approach: It will turn out that second derivative $\partial_{m^2}^2 c_+^{(2)}(F;G;\cdot,\cdot)$ belongs to L^1, and therefore, with a function $\Pi^{\mu\nu}$ to be found, we shall search instead for a candidate of the form

$$j^{(2)}(F;G) = -\frac{ie^2}{(2\pi)^4} \int\limits_{[\mathbb{R}-i\delta]\times\mathbb{R}^3} d^4k\, F_\mu(k) G_\nu(-k) \Pi^{\mu\nu}(k)\,, \tag{28}$$

$$\text{fulfilling} \quad \partial_{m^2}^2 j^{(2)}(F;G) = \int\limits_{\mathbb{R}^3\times\mathbb{R}^3} d^3p\, d^3q\, \partial_{m^2}^2 c_+^{(2)}(F;G;\mathbf{p},\mathbf{q})\,. \tag{29}$$

Throughout the article, we have suppressed the m dependence in the notation. In the end, we shall integrate twice with respect to m^2; taking the second derivative with respect to m^2 rather than m is only a technical convenience. This can be seen as a variant of the Pauli-Villars regularisation scheme [30] with differences replaced by integrals of derivatives. The expression for $\Pi^{\mu\nu}$ will finally be identified as (42).

5 The Explicit Expression for the Second Order

Having in mind the goal (29) with $c_+^{(2)}$ as given in (26) we observe that even for k close to zero, $\omega^{\mu\nu}(p,k)$ in (**??**) is not Lebesgue-integrable in $p \in i\mathbb{R}\times\mathbb{R}^3$. However, we have the Lebesgue integrals

$$\int\limits_{\mathcal{C}_{\mathrm{Wick}}(k)\times\mathbb{R}^3} d^4p\, \partial_{m^2}^2 \omega^{\mu\nu}(p,k) \underset{\text{close to0}}{\overset{\text{for}k}{=}} \int\limits_{i\mathbb{R}\times\mathbb{R}^3} d^4p\, \partial_{m^2}^2 \omega^{\mu\nu}(p,k)\,. \tag{30}$$

Note that by dominated convergence, the right-hand side in (30) is a real-analytic function of $\mathbf{k} \in \mathbb{R}^3$ and a holomorphic function of $k^0 \in \mathbb{D}$; cf. (27).

In order to avoid the explicit computation of the second derivative, we introduce an artificial scaling parameter ϵ on which the integral shall depend meromorphically. The original integral (30) is then recovered for $\epsilon \to 0$. However, by virtue of the identity theorem for analytic functions, the value at $\epsilon = 0$ is determined by values for ϵ in any non-empty, open interval I of reals, which need not be in the vicinity of zero. It will turn out that I can be chosen so that the differential operator $\partial_{m^2}^2$ can be interchanged with the integral by dominated convergence without losing Lebesgue-integrability for $\epsilon \in I$. A scaling exponent ϵ can be introduced conveniently after recasting the integral by means of Feynman's parametrisation. This computation procedure can be seen as a mathematically rigorous version of the method of dimensional regularisation employed in physics [3, 41]. We interpret the attribute "dimensional" as a scaling exponent ϵ, which may be any real or even complex number, but not as the number of elements in a basis of a vector

space, which would be a natural number. Next, we shall demonstrate this procedure for the second order term, which can be seen as a rigorous version of the computation in [9, p. 70ff: Polarization of the Vacuum].

In order to arrive at an explicit expression for the second perturbation order of the current, we employ the Feynman parametrisation $(ab)^{-1} = \int_0^1 ((1-z)a + zb)^{-2}\, dz$, which holds for all $a, b \in \mathbb{C}$ such that the straight line from a to b does not contain 0. For $k \in \mathbb{C}^4$ sufficiently close to zero, $p \in i\mathbb{R} \times \mathbb{R}^3$ we may take $a = p^2 - m^2$ and $b = (p-k)^2 - m^2$ and recast (30) into

$$\int_{i\mathbb{R}\times\mathbb{R}^3} d^4p\, \partial_{m^2}^2 \frac{\mathrm{tr}[\gamma^\nu(\not{p}+m)\gamma^\mu(\not{p}-\not{k}+m)]}{(p^2-m^2)((p-k)^2-m^2)} = \int_{i\mathbb{R}\times\mathbb{R}^3} d^4p\, \partial_{m^2}^2 \int_0^1 dz \frac{\mathrm{tr}[\gamma^\nu(\not{p}+m)\gamma^\mu(\not{p}-\not{k}+m)]}{\left[(1-z)(p^2-m^2)+z((p-k)^2-m^2)\right]^2}$$
$$= \int_0^1 dz \int_{i\mathbb{R}\times\mathbb{R}^3} d^4p\, \partial_{m^2}^2 \frac{\mathrm{tr}[\gamma^\nu(\not{p}+m)\gamma^\mu(\not{p}-\not{k}+m)]}{\left[(1-z)(p^2-m^2)+z((p-k)^2-m^2)\right]^2}\,, \tag{31}$$

where in the last step we have used dominated convergence to interchange $\partial_{m^2}^2$ and $\int dz$ and the Lebesgue-integrability to interchange the integrals. For the sake of the computation, we may restrict ourselves to $k \in i\mathbb{R} \times \mathbb{R}^3$; the rest of the computation does not even require k to be close to zero. Next, employing the substitution $q = p - kz$, which is m-independent, to find

$$(31) = \int_0^1 dz \int_{i\mathbb{R}\times\mathbb{R}^3} d^4q\, \partial_{m^2}^2 \frac{\mathrm{tr}[\gamma^\nu(\not{q}+z\not{k}+m)\gamma^\mu(\not{q}+(z-1)\not{k}+m)]}{[m^2-(z-z^2)k^2-q^2]^2}, \tag{32}$$

and evaluating the trace, using $\mathrm{tr}\,\gamma^\nu\gamma^\mu = 4g^{\mu\nu}$, $\mathrm{tr}\,\gamma^\nu\gamma^\mu\gamma^\sigma = 0$, and $\mathrm{tr}\,\gamma^\nu\gamma^\sigma\gamma^\mu\gamma^\tau = 4(g^{\mu\sigma}g^{\nu\tau} - g^{\mu\nu}g^{\sigma\tau} + g^{\mu\tau}g^{\sigma\nu})$, as done, e.g., in [9, Equations (377)–(379)], and dropping terms of the form $\int d^4q\, q_\mu k^\mu f(q^2) = 0$, results in

$$(31) = \int_0^1 dz \int_{i\mathbb{R}\times\mathbb{R}^3} d^4q\, \partial_{m^2}^2 f_{k,m}^{\mu\nu}(q,z) \tag{33}$$

$$\text{with} \qquad f_{k,m}^{\mu\nu}(q,z) := 4\frac{2q^\mu q^\nu - 2k^\mu k^\nu(z-z^2) + g^{\mu\nu}(m^2 + k^2(z-z^2) - q^2)}{[m^2-(z-z^2)k^2-q^2]^2}\,. \tag{34}$$

Due to the Minkowski inner-product, we have $-k^2 \geq 0$, $-q^2 \geq 0$. Note further that $f_{q,k,z,m}^{\mu\nu}$ behaves as $O_{|q|\to\infty}(|q|^{-2})$ and $O_{|q|\to 0}(1)$ uniformly for $z \in [0,1]$, k in a compact domain in $i\mathbb{R}\times\mathbb{R}^3$, and also m in a compact domain in $\mathbb{R}^+$.

As announced above, we will now introduce an artificial scaling $(-q^2/u)^\epsilon$ with a complex exponent ϵ and another parameter u carrying the unit of m^2; recall that $\hbar = 1 = c$. The original term (30) shall be retrieved in the limit $\epsilon \to 0$.

$$(31) = \lim_{\epsilon\to 0} \int_0^1 dz \int_{i\mathbb{R}\times\mathbb{R}^3} d^4q \left(\frac{-q^2}{u}\right)^\epsilon \partial_{m^2}^2 f_{k,m}^{\mu\nu}(q,z) \tag{35}$$

Note that the integrand is an Lebesgue-integrable, holomorphic function of ϵ for $-2 < \mathrm{Re}\,\epsilon < +1$ which allows to interchange the limit as $\epsilon \to 0$ with the integral by dominated convergence. Even when dropping the $\partial_{m^2}^2$ operator, it remains so for the smaller domain $-2 < \mathrm{Re}\,\epsilon < -1$, which does, however, not contain zero any more. First, we will ignore this complication in the computations and work with smaller domain but, in the end, recover a domain that contains $\epsilon = 0$ by analytic continuation. On the smaller domain, using dominated convergence once again, we find

$$\int_0^1 dz \int_{i\mathbb{R}\times\mathbb{R}^3} d^4q \left(\frac{-q^2}{u}\right)^{\epsilon} \partial_{m^2}^2 f_{k,m}^{\mu\nu}(q,z) = \int_0^1 dz\, \partial_{m^2}^2 \int_{i\mathbb{R}\times\mathbb{R}^3} d^4q \left(\frac{-q^2}{u}\right)^{\epsilon} f_{k,m}^{\mu\nu}(q,z)\,. \tag{36}$$

We will now evaluate the inner integral for $-2 < \mathrm{Re}\epsilon < -1$. The following integral

$$I(\zeta,\eta) := u^{-\eta} \int_{i\mathbb{R}\times\mathbb{R}^3} d^4q \frac{(-q^2)^{\eta}}{(\zeta u - q^2)^2} = \int_{i\mathbb{R}\times\mathbb{R}^3} \frac{d^4q}{u^2} \frac{\left(\frac{-q^2}{u}\right)^{\eta}}{(\zeta - q^2/u)^2}\,, \quad \text{for } \zeta > 0, \mathrm{Re}\eta \in (-2,0)\,, \tag{37}$$

will play an important role in these calculations. The decoration by factors of powers of u renders the expression (37) dimensionless. By euclidean symmetry of the complexified Minkowski inner-product on $i\mathbb{R}\times\mathbb{R}^3$ and for $\zeta > 0$, we have

$$\int_{i\mathbb{R}\times\mathbb{R}^3} d^4q \left(\frac{-q^2}{u}\right)^{\epsilon} \frac{q^\mu q^\nu}{[\zeta u - q^2]^2} = \int_{i\mathbb{R}\times\mathbb{R}^3} d^4q \left(\frac{-q^2}{u}\right)^{\epsilon} \frac{\frac{1}{4} g^{\mu\nu} q^2}{[\zeta u - q^2]^2} = -\frac{u}{4}\, g^{\mu\nu} I(\zeta, 1+\epsilon)\,, \tag{38}$$

Using (36), (37), and $\zeta := (m^2 - (z - z^2)k^2)/u$, we recast the inner integral of (35) into

$$\begin{aligned} F_{k,m}^{\mu\nu}(z,\epsilon) &:= \int_{i\mathbb{R}\times\mathbb{R}^3} d^4q \left(\frac{-q^2}{u}\right)^{\epsilon} f_{k,m}^{\mu\nu}(q,z) \\ &= 2\, g^{\mu\nu} u I(\zeta, 1+\epsilon) + 4[(g^{\mu\nu}k^2 - 2k^\mu k^\nu)(z - z^2) + g^{\mu\nu} m^2] I(\zeta,\epsilon)\,. \end{aligned} \tag{39}$$

Next, we evaluate the integral $I(\zeta,\eta)$ for $\zeta > 0$ and $-2 < \mathrm{Re}\,\eta < 0$, i.e.,

$$\begin{aligned} I(\zeta,\eta) &= 2i\pi^2 u^{-\eta} \int_0^\infty \frac{r^{3+2\eta}}{(\zeta u + r^2)^2} dr = i\pi^2 \zeta^\eta\, \mathrm{B}(\eta+2, -\eta) = i\pi^2 \zeta^\eta \Gamma(2+\eta)\Gamma(-\eta) \\ &= i\pi^2 \zeta^\eta (1+\eta)\Gamma(1+\eta)\Gamma(-\eta) = i\pi^2 \zeta^\eta \frac{\pi(1+\eta)}{\sin(\pi(1+\eta))}\,, \end{aligned} \tag{40}$$

where B denotes the beta function and, in the last step, Euler's reflection formula was used. For all given $\zeta > 0$, the right-hand side of (40) implies that the function $(-2,-1) + i\mathbb{R} \ni \eta \mapsto I(\zeta,\eta)$ has a meromorphic extension on the whole complex plane with poles of first order at most on $\mathbb{Z}\setminus\{-1\}$. We denote it by the same symbol I. Hence, $F_{k,m}^{\mu\nu}(z,\epsilon)$ in (39) also has a holomorphic extension for $\epsilon \in \mathbb{C}\setminus\mathbb{Z}$ which will again be denoted by the same symbol. It is important to note that this extends the originally much smaller domain of $\epsilon \in (-2,-1) + i\mathbb{R}$, which did not even include a neighbourhood of $\epsilon = 0$. Given $\zeta = (m^2 - (z - z^2)k^2)/u > 0$, this extension now allows to expand for $\epsilon \to 0$ as follows:

$$I(\zeta,1+\epsilon)=2i\pi^2\zeta\left(\frac{1}{\epsilon}+\log\zeta+\frac{1}{2}+O^{\zeta}_{\epsilon\to 0}(\epsilon)\right),\quad I(\zeta,\epsilon)=-i\pi^2\left(\frac{1}{\epsilon}+\log\zeta+1+O^{\zeta}_{\epsilon\to 0}(\epsilon)\right),$$

$$\begin{aligned}F^{\mu\nu}_{k,m}(z,\epsilon)&=4i\pi^2[g^{\mu\nu}m^2-g^{\mu\nu}k^2(z-z^2)]\left(\frac{1}{\epsilon}+\log\zeta+\frac{1}{2}\right)\\&\quad-4i\pi^2[g^{\mu\nu}m^2+(g^{\mu\nu}k^2-2k^\mu k^\nu)(z-z^2)]\left(\frac{1}{\epsilon}+\log\zeta+1\right)+O^{z,k,m}_{\epsilon\to 0}(\epsilon)\\&=8i\pi^2(k^\mu k^\nu-g^{\mu\nu}k^2)(z-z^2)\left(\frac{1}{\epsilon}+\log\zeta\right)+2i\pi^2[4k^\mu k^\nu-g^{\mu\nu}(m^2+3k^2)](z-z^2)+O^{z,k,m}_{\epsilon\to 0}(\epsilon)\,.\end{aligned}$$

We emphasise that the remainder $O^{z,k,m}_{\epsilon\to 0}(\epsilon)$ is uniform in the parameters ζ and k,m,z, respectively, as long as they are restricted to compact domains. Moreover, it is smooth in m with $\partial^n_{m^2}O^{z,k,m}_{\epsilon\to 0}(\epsilon)=O^{z,k,m}_{\epsilon\to 0}(\epsilon)$, $n\in\mathbb{N}$. Hence, the expressions in (39) and, exploiting uniformity and smoothness of the remainder term, also the expression in (36), taken on the smaller domain $-2<\mathrm{Re}\,\epsilon<-1$, has a holomorphic extension for $\epsilon\in\mathbb{C}\setminus\mathbb{Z}$ given by

$$\int_0^1 dz\,\partial^2_{m^2}F^{\mu\nu}_{k,m}(z,\epsilon)=\partial^2_{m^2}\Pi^{\mu\nu}(k)+O^{k,m}_{\epsilon\to 0}(\epsilon)\quad\text{with}\tag{41}$$

$$\Pi^{\mu\nu}(k):=8i\pi^2(k^\mu k^\nu-g^{\mu\nu}k^2)\int_0^1 dz\,(z-z^2)\left[\log\left(1-(z-z^2)\frac{k^2}{m^2}\right)+\log\frac{m^2}{u}\right].\tag{42}$$

The last expression shows that the isolated singularity at $\epsilon=0$ is removable. Applying the identity theorem for holomorphic functions guarantees that the left-hand side of (36) is given by the expression on the left-hand side in (41) for $-2<\epsilon<+1$, which is a neighbourhood of $\epsilon=0$. Note that thanks to the derivative $\partial^2_{m^2}$, the term is independent of u, as it must be. Furthermore, the $1/\epsilon$-singularity has dropped out thanks to the same derivative, which is consistent with the fact that the left-hand side of (36) is holomorphic near $\epsilon=0$. Evaluating the holomorphic extension at $\epsilon=0$, we have verified that the quantity $\Pi^{\mu\nu}(k)$ defined in (42) fulfils

$$\partial^2_{m^2}\Pi^{\mu\nu}(k)=\int_{\mathcal{C}_{\mathrm{Wick}}(k)\times\mathbb{R}^3}d^4p\,\partial^2_{m^2}\omega^{\mu\nu}(p,k)\tag{43}$$

for $k\in i\mathbb{R}\times\mathbb{R}^3$. Interpreting the logarithm in (42) as its principal value $\log:\mathbb{C}\setminus\mathbb{R}^-_0\to\mathbb{C}$, the function $\Pi^{\mu\nu}$ on $i\mathbb{R}\times\mathbb{R}^3$ has a holomorphic extension to $\{k\in\mathbb{C}^4\,|\,k^2\notin\mathbb{R}\text{ or }k^2<4m^2\}$ which contains a neighbourhood of $0\in\mathbb{C}^4$ as well as the sets $\mathbb{D}\times\mathbb{R}^3$ and $\mathbb{R}^4+i\overline{\text{Causal}}$, where $\overline{\text{Causal}}=\{p\in\mathbb{R}^4:\,p^2\geq 0\}$ denotes the set of time-like or light-like four vectors. This implies (43) also for $k\in\mathbb{D}\times\mathbb{R}^3$ by virtue of the identity theorem for analytic functions. Note further that $\Pi^{\mu\nu}(k)$ is of the order $O^{m,u}_{|k|\to\infty}(|k|^2\log|k|)$ and thus allows the integral on the right-hand side of (28) to be well-defined and independent of $\delta>0$.

It is left to show (22). For this purpose, we claim

$$\partial^2_{m^2}\left(j^{(2)}(F;G)-j^{(2)}(G;F)-c^{(2)}(G,F)\right)=0\,.\tag{44}$$

To see this, we recall (28) and regard

$$\begin{aligned}\partial_{m^2}^2 j^{(2)}(F;G) &\overset{(44)}{=} -\int\limits_{[\mathbb{R}-i\delta]\times\mathbb{R}^3} d^4k\, F_\mu(k)G_\nu(-k) \int\limits_{\mathcal{C}_{\mathrm{Wick}}(k)\times\mathbb{R}^3} d^4p\, \partial_{m^2}^2\omega^{\mu\nu}(p,k)\\ &= -\int\limits_{\mathbb{R}^3\times\mathbb{R}^3} d^3k d^3p\, \partial_{m^2}^2 \int\limits_{[\mathbb{R}-i\delta]} dk^0\, F_\mu(k)G_\nu(-k) \int\limits_{\mathcal{C}_{\mathrm{Wick}}(k)} dp^0\, \omega^{\mu\nu}(p,k) = \int\limits_{\mathbb{R}^3\times\mathbb{R}^3} d^3p\, d^3q\, \partial_{m^2}^2 c_+^{(2)}(F;G;\mathbf{p},\mathbf{q}),\end{aligned} \tag{45}$$

where the employed commutation of the integrals and differential operators is justified by the integrability of the integrand, locally uniform in m. Inserting equation (24) proves the claim (44). Furthermore, (44) implies

$$j^{(2)}(F;G) - j^{(2)}(G;F) - c^{(2)}(G,F) = am^2 + b, \tag{46}$$

with two constants a, b that depend on F, G but are independent of m. Note that $j^{(2)}(F;G) - j^{(2)}(G;F) \to 0$ for $m \to \infty$ because the scaling term $\log(m^2/u)$ cancels in the difference; see (42). Moreover, we have $c^{(2)}(G,F) \to 0$ for $m \to \infty$, which can be seen from (23), the estimate $|p-q| \geq |p^0 - q^0| = E(\mathbf{p}) + E(\mathbf{q}) \geq 2^{-1/2}(|\mathbf{p}| + |\mathbf{q}| + 2m)$ for $p^0 = -\tau E(\mathbf{p}), q^0 = \tau E(\mathbf{q})$, and $\tau = \pm 1$, implied by Cauchy-Schwarz' inequality, and noting the fact that F, G decay super-algebraically fast in energy-momentum space. This guarantees $a = 0 = b$, i.e., (22).

Conclusion. Up to second order in perturbation, we have retrieved a physically relevant family of currents by (28) and (42), depending on an integration constant $C \in \mathbb{R}$, i.e.,

$$j^{(2)}(F;A) = \frac{e^2}{2\pi^2} \int\limits_{[\mathbb{R}-i\delta]\times\mathbb{R}^3} d^4k\, F_\mu(k)A_\nu(-k)(k^\mu k^\nu - g^{\mu\nu}k^2) \int_0^1 dz\, (z-z^2)\left[\log\left(1-(z-z^2)\frac{k^2}{m^2}\right) + C\right], \tag{47}$$

because it fulfils the physical relevant conditions C1–C4 in second order of perturbation:

- C2 *Relativistic invariance*: The Lorentz covariance is apparent from the Lorentz covariance of formula (42). Translation invariance follows from the fact that the test functions F, G appear in energy-momentum space only in the form $k \mapsto F_\mu(k)G_\nu(-k)$.
- C1 *Causality*: Temporal causality follows from the discussed temporal causality of $c_+^{(2)}$; see (25). Fully relativistic causality follows from the temporal causality and relativistic invariance.
- C3 *Gauge invariance*: Gauge invariance is apparent from formula (42) noting $k_\mu(k^\mu k^\nu - g^{\mu\nu}k^2) = 0$ and the symmetry in μ and ν.
- C4 *Reference current*: The vanishing reference current $j_{A=0} = 0$ is already build into the ansatz, see (19), because there is no summand corresponding to $n = 1$.

The family of currents (47) describes the second order term of the vacuum expectation of the electric current in a prescribed external four-vector potential $G = A$. Concerning the interpretation of the remaining constant C, it is helpful to regard the external current $j_{\mathrm{ext}}^\mu(k) = (k^\mu k^\nu - g^{\mu\nu}k^2)A_\nu^{\mathrm{ext}}(k)$, associated with the external field $G = A^{\mathrm{ext}}$ by Maxwell's equations. Changing the integration constant C, thus, changes the vacuum polarisation current $j^{(2)}$ in second order of perturbation by an additional current proportional to j_{ext}. In a self-consistently coupled theory, this mechanism can be interpreted as to leave the bare electric charge undefined.

Be that as it may. Finally and most importantly, we would like to express our gratitude to Detlef for being our teacher, dear colleague and friend. We miss him dearly.

A Tetrahedron Rule

Let A, B, C, D be unitary operators on $\mathcal{H}$ fulfilling (9) for all $U = X^{-1}Y$, $X, Y \in \{A, B, C, D\}$. For such operators we define: $\Gamma_{ABC} := \arg\det(A^{-1}B)_-(B^{-1}C)_-(C^{-1}A)_- \in U(1)$

Lemma A.1 *For such operators A, B, C, D we have:* ***1****. Γ_{ABC} is well-defined;* ***2****. $\Gamma_{ABC}^{-1} = \Gamma_{CBA}$;* ***3****. $\Gamma_{AAB} = 1$;* ***4****. $\Gamma_{ABC} = \Gamma_{BCD}\Gamma_{DCA}\Gamma_{ABD}$.*

Proof **1**: We observe that each pair of operators such as $(A^{-1}B)_-$ is invertible because of $\| \mathrm{id}_- - (A^{-1}B)_- \| < 1$. Hence, the Fredholm determinant is well-defined and non-zero because of $(A^{-1}B)_-(B^{-1}C)_-(C^{-1}A)_- = \big((A^{-1}C)_- - (A^{-1}B)_{-+}(B^{-1}C)_{+-}\big)(C^{-1}A)_-$
$\in (A^{-1}C)_-(C^{-1}A)_- + I_1(\mathcal{H}^-) = \mathrm{id}_{\mathcal{H}^-} - (A^{-1}C)_{-+}(C^{-1}A)_{+-} + I_1(\mathcal{H}^-) = \mathrm{id}_- + I_1(\mathcal{H}^-)$.
2: This can be seen by noting $\Gamma_{ABC}^{-1} = \overline{\Gamma_{ABC}}$ and $\overline{\det X} = \det X^*$ for $X \in \mathrm{id} + I_1$.
3: $\Gamma_{AAB} = \arg\det(A^{-1}B)_-(B^{-1}A)_- = \arg\det|(A^{-1}B)_-|^2 = 1$.
4: Applying Lidskii's theorem of cyclic permutation under the Fredholm determinant we find

$$\begin{aligned}\Gamma_{BCD}\Gamma_{DCA}\Gamma_{ABD} &= \arg\det(D^{-1}B)_-(B^{-1}C)_-(C^{-1}D)_- \arg\det(D^{-1}C)_-(C^{-1}A)_-(A^{-1}D)_- \\ &\times \arg\det(D^{-1}A)_-(A^{-1}B)_-(B^{-1}D)_- = \arg\det|(D^{-1}B)_-|^2\,(B^{-1}C)_-\,|(D^{-1}C)_-|^2 \\ &\circ (C^{-1}A)_-\,|(D^{-1}A)_-|^2\,(A^{-1}B)_- = \arg\det(B^{-1}C)_-\,(C^{-1}A)_-\,(A^{-1}B)_- = \Gamma_{ABC}\end{aligned}$$

because the operators under the square modulus, like $(B^{-1}D)_-(D^{-1}B)_- = |(D^{-1}B)_-|^2$ are positive definite operators in $\mathrm{id}_- + I_1(\mathcal{H}^-)$. □

References

1. C. Bär, K. Fredenhagen,*Quantum Field Theory on Curved Spacetimes: Concepts and Mathematical Foundations*. Lecture Notes in Physics, vol. 786 (Springer, 2009)
2. C. Bär, N. Ginoux, F. Pfäffle, *Wave Equations on Lorentzian Manifolds and Quantization*, vol. 3 (European Mathematical Society, 2007)
3. C.G. Bollini, J.J. Giambiagi, Dimensional renorinalization : The number of dimensions as a regularizing parameter. Il Nuovo Cimento B (1971–1996) **12**(1), 20–26 (1972)
4. D.-A. Deckert, D. Dürr, F. Merkl, M. Schottenloher, Time-evolution of the external field problem in Quantum Electrodynamics. J. Math. Phys. **51**(12), 122301 (2010)
5. D.-A. Deckert, F. Merkl, Dirac equation with external potential and initial data on Cauchy surfaces. J. Math. Phys. **55**(12), 122305 (2014)
6. D.-A. Deckert, F. Merkl, External field QED on Cauchy surfaces for varying electromagnetic fields. Commun. Math. Phys. **345**(3), 973–1017 (2016)
7. D.-A. Deckert, F. Merkl, A perspective on external field QED, in *Quantum Mathematical Physics: A Bridge between Mathematics and Physics*, pp. 381–399 (2016)
8. P.A.M. Dirac, Theorie du Positron, in *Selected Papers on Quantum Electrodynamics*, ed. by J. Schwinger (Dover Publications, Inc., New York, 1934)
9. F.J. Dyson, Lectures on Advanced Quantum Mechanics Second Edition (1951). (World Scientific, 2006)
10. H. Fierz, G. Scharf, Particle interpretation for external field problems in QED. Helvetica Physica Acta. Physica Theoretica **52**(4), 437– 453 (1980), (1979)
11. F. Finster, *The Principle of the Fermionic Projector*, vol. 35 (American Mathematical Society, 2006)
12. F. Finster, Causal variational principles on measure spaces. Journal für die reine undangewandte Mathematik **141–194**, 2010 (2010)

13. F. Finster, *Continuum Limit of Causal Fermion Systems*, vol. 186 (Springer, 2016)
14. F. Finster, J. Kleiner, Causal fermion systems as a candidate for a unified physical theory. J. Phys. Conf. Ser. **626**(1), 012020. (IOP Publishing, 2015)
15. F. Finster, M. Reintjes, A non-perturbative construction of the fermionic projector on globally hyperbolic manifolds I: Space-times of finite lifetime. Adv. Theor. Math. Phys. **19**(4), 761–803 (2015)
16. F. Finster, M. Reintjes, A non-perturbative construction of the Fermionic projector on globally hyperbolic manifolds II: Space-times of infinite lifetime. Adv. Theor. Math. Phys. **20**(5), 1007–1048 (2016)
17. P. Gravejat, C. Hainzl, M. Lewin, E. Séré, Construction of the Pauli-Villars-Regulated Dirac vacuum in electromagnetic fields. Arch. Rat. Mech. Anal. **208**(2), 603–665 (2013)
18. P. Gravejat, C. Hainzl, M. Lewin, É. Séré, Two Hartree-Fock models for the vacuum polarization. Journées équations aux dérivées partielles, 1–31 (2012)
19. P. Gravejat, M. Lewin, É. Séré, Derivation of the magnetic Euler–Heisenberg energy. Journal de Mathématiques Pures et Appliquées **117**, 59–93 (2018)
20. C. Hainzl, On the vacuum polarization density caused by an external field. Annales Henri Poincaré **5**(6), 1137–1157 (2004)
21. C. Hainzl, M. Lewin, E. Séré, Self-consistent solution for the polarized vacuum in a no-photon QED model. J. Phys. Math. Gen. **38**(20), 4483 (2005)
22. C. Hainzl, M. Lewin, E. Séré, J.P. Solovej, Minimization method for relativistic electrons in a mean-field approximation of quantum electrodynamics. Phys. Rev. A **76**(5), 052104 (2007)
23. H. Kragh, *Dirac: A Scientific Biography* (Cambridge University Press, 1990)
24. E. Langmann, J. Mickelsson, Scattering matrix in external field problems. J. Math. Phys. **37**(8), 3933–3953 (1996)
25. M. Lewin, A nonlinear variational problem in relativistic quantum mechanics, in *European Congress of Mathematics Kraków, 2–7 July, 2012*, pp. 627–641 (2014)
26. J. Mickelsson, Vacuum polarization and the geometric phase: gauge invariance. J. Math. Phys. **39**(2), 831–837 (1998)
27. J. Mickelsson, The phase of the scattering operator from the geometry of certain infinite-dimensional groups. Lett. Math. Phys. **104**(10), 1189–1199 (2014)
28. J. Mickelsson, S. Scott, Functorial QFT, gauge anomalies and the Dirac determinant bundle. Commun. Math. Phys. **219**(3), 567–605 (2001)
29. G. Nenciu, Existence of the spontaneous pair creation in the external field approximation of QED. Commun. Math. Phys. **109**(2), 303–312 (1987)
30. W. Pauli, F. Villars, On the invariant regularization in relativistic quantum theory. Rev. Mod. Phys. **21**, 434–444 (1949)
31. P. Pickl, D. Dürr, Adiabatic pair creation in heavy-ion and laser fields. Eur. Lett. **81**(4), 40001 (2008)
32. P. Pickl, D. Dürr, On adiabatic pair creation. Commun. Math. Phys. **282**(1), 161 (2008)
33. A. Pressley, G. Segal, *Loop Groups*, Oxford Mathematical Monographs (The Clarendon Press, Oxford University Press, New York, 1986)
34. S.N.M. Ruijsenaars, Charged particles in external fields. II. The quantized Dirac and Klein-Gordon theories. Commun. Math. Phys. **52**(3), 267–294 (1977)
35. G. Scharf, *Finite Quantum Electrodynamics*, 2nd edn. Texts and Monographs in Physics (Springer-Verlag, Berlin, 1995)
36. G. Scharf, W.F. Wreszinski, The causal phase in quantum electrodynamics. Il Nuovo Cimento A (1965–1970) **93**(1), 1–27 (1986)
37. G. Scharf, W.F. Wreszinski, The causal phase in quantum electrodynamics. Società Italiana di Fisica. Il Nuovo Cimento. A. Serie 11 **93**(1), 1–27 (1986)
38. J. Schlemmer, J. Zahn, The current density in quantum electrodynamics in external potentials. Ann. Phys. **359**, 31–45 (2015)
39. D. Shale, W.F. Stinespring, Spinor representations of infinite orthogonal groups. J. Math. Mech. **14**, 315–322 (1965)

40. B. Simon, *Trace Ideals and Their Applications*, Number 120 in Mathematical Surveys and Monographs (American Mathematical Soc., 2005)
41. G. 't Hooft, M. Veltman, Regularization and renormalization of gauge fields. Nucl. Phys. B **44**(1), 189–213 (1972)
42. B. Thaller, *The Dirac Equation, Texts and Monographs in Physics* (Springer-Verlag, Berlin, 1992)
43. T. Wurzbacher, Fermionic second quantization and the geometry of the restricted Grassmannian, in *Infinite Dimensional Kähler Manifolds* (Oberwolfach, 1995), vol. 31 of *DMV Sem.* (Birkhäuser, Basel, 2001), pp. 287–375

Mathematical Physics

The Arrow of Time

Peter Pickl and Aaron Schaal

Abstract Since Boltzmann's works, there is the question of explaining the arrow of time in statistical physics: While many microscopic physical theories do not know a preferred direction of time, one experiences an arrow of time in the macroscopic world. Glass shatters when it falls on the ground whereas we never see the reverse process. The usual explanations for that are based on a very special initial condition, a state of very high order which under the dynamics evolves in a natural way into a state of higher disorder. Here one can see (order → disorder) the asymmetry in the propagation in time. This explanation, however, shifts the problem to the question of where the initial state of high order came from. In the following we will discuss different possible explanations for the emergence of a macroscopic arrow of time, in particular we will discuss an idea going back to V. A. Antonov who found that the growth in disorder can be explained without assuming a special initial condition.

1 Introduction

> "Thou turn thy mind the more unto these bodies
> Which here are witnessed tumbling in the light:
> Namely, because such tumblings are a sign
> That motions also of the primal stuff
> Secret and viewless lurk beneath, behind.
> For thou wilt mark here many a speck, impelled
> By viewless blows, to change its little course,
> And beaten backwards to return again,
> Hither and thither in all directions round.

P. Pickl (✉) · A. Schaal
Fachbereich Mathematik, Eberhard-Karls-Universität Tübingen, Tübingen, Germany
e-mail: p.pickl@uni-tuebingen.de

A. Schaal
e-mail: aaron.schaal@uni-tuebingen.de

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_19

> Lo, all their shifting movement is of old,
> From the primeval atoms; for the same
> Primordial seeds of things first move of self,
> And then those bodies built of unions small
> And nearest, as it were, unto the powers
> Of the primeval atoms, are stirred up
> By impulse of those atoms' unseen blows,
> And these thereafter goad the next in size;
> Thus motion ascends from the primevals on,
> And stage by stage emerges to our sense,
> Until those objects also move which we
> Can mark in sunbeams, though it not appears
> What blows do urge them.
> Herein wonder not
> How'it is that, while the seeds of things are all
> Moving forever, the sum yet seems to stand
> Supremely still, except in cases where
> A thing shows motion of its frame as whole. [24]

Around the beginning of the twentieth century, three groundbreaking ideas manifestly changed the physical world view: relativity, quantum mechanics and kinetic gas theory. While relativity and quantum mechanics changed the fundamental laws of physics, kinetic gas theory is rather a unification than a modification of theories that were present before. It is based on the idea, respectively the re-discovery of the ancient idea of atomism (see the quote of Lucretius taken from the English translation of his famous poem "De rerum natura" $\sim$50 B.C. above), that the laws of macroscopic objects can be derived from the microscopic motion of atoms. Using the laws of Newtonian dynamics, it is in fact possible to derive the laws describing gases, for example the laws of Boyle-Mariotte or Gay-Lussac. Therefore, the ideas of kinetic gas theories provide a unification of kinematics and thermodynamics.

Ever since, the derivation of effective descriptions of macroscopic equations or effects from microscopic principles has been an active area of research. Next to the classical gas laws mentioned above, there is a large number of examples for which it is well understood how the collective behavior of atoms give rise to physical effects one experiences on the macroscopic level.

What makes kinetic gas theory special is the generality of its ideas. Kinetic gas theory is less a particular model describing our physical world with more accuracy, but a principle that allows for connecting the microscopic and macroscopic world. One is not limited to consider Newtonian motion for the dynamics of the "atoms". For example the macroscopic effect of Bose-Einstein condensation which is a collective effect of quantum mechanical particles is rather well understood. The ideas are not limited to physics, also biological and sociological systems of many individuals can be considered to connect the behavior of individuals to the collective behavior on the macroscopic level.

Despite their beauty and clarity, the ideas of atomism were subject to criticism, in particular in the early years after their formulation. The break-through of the kinetic theory of gases came with Einstein's explanation of Brownian motion [18]. Nevertheless there have been lasting objections against the theory until today.

The main reason for the confusion about Boltzmann's explanation is the fact that the macroscopic world has irreversible processes – an empirical fact one can not argue against – while the microscopic descriptions used by Boltzmann are reversible. In this paper we will discuss several attempts that give a possible explanation of this break in symmetry. By "possible" we mean that we are not going to discuss modern cosmological models and find the most plausible explanation for *our* universe, but rather provide an overview of the basic ideas.

2 Objections Against Kinetic Gas Theory

Despite their success, the ideas of Boltzmann and Maxwell [5, 27] – to mention the two most important founding fathers of kinetic gas theory – were, in the beginning, subject to criticism from contemporary physicists. Among the most prominent opponents was Ernst Mach, who thought that "The mechanical conception of the second law of thermodynamics, by distinguishing between orderly and disorderly motions and equating increase of entropy with increase of disorderly motions at the expense of orderly ones, seems a very artificial expedient" [26]. The positions of Mach and others were attacked by Einstein: "The antipathy of these scholars towards atomic theory can indubitably be traced back to their positivistic philosophical attitude. This is an interesting example of the fact that even scholars of audacious spirit and fine instinct can be obstructed in the interpretation of facts by philosophical prejudices. The prejudice – which has by no means died out in the meantime – consists in the faith that facts by themselves can and should yield scientific knowledge without free conceptual construction. Such a misconception is possible only because one does not easily become aware of the free choice of such concepts, which, through verification and long usage, appear to be immediately connected with the empirical material" [17]. Einstein addresses an important issue. His criticism also applies to other areas of science (see [7, 14]).

Next to Mach's repudiation, there were more substantial objections against the kinetic theory of gases: Loshmidt's "Umkehreinwand" (also called Loschmidt's paradoxon) formulated by Josef Loschmidt in 1876 [23] and Zermelo's "Wiederkehreinwand" (recurrence objection) phrased by Ernst Zemelo in 1896 [31]. The arguments of theses objections are as follows.

2.1 Umkehreinwand

The objection formulated by Loschmidt [23] aims to show a contradiction between atomism and the second law of thermodynamics, in particular a contradiction between Boltzmann's formula for the entropy and the fact that entropy is an non-decreasing function. Following Boltzmann, the entropy of a gas with phase-space density ρ is given by the integral $S^B = k^B \int \rho(x, v) \ln \rho(x, v) d^3x d^3v$.

Loschmidt's argument goes as follows: Think of a system which is at time $t = 0$ in a macro-state of low entropy $S(0)$ and which, after some time t has passed, evolves into a macro-state of larger entropy $S(t) > S(0)$. Assuming the correctness of the ideas of the kinetic theory of gases this behavior can be explained by the ballistic motion of a large number – let us say N – of atoms. In other words there has to be a point X_0 in the respective $6N$-dimensional phase space which evolves by the Newtonian laws of motions into a point X_t such that the Boltzmann entropy of the respective points in phase space increases ($S^B(X_0) < S^B(X_t)$).

Now consider the state $X(t_1)$ and reverse the velocities of all particles. Recall that Newtonian mechanics is invariant under the simultaneous reversal of time and velocities. Therefore, after time t has passed, that state will evolve into a state which is identical to X_0 with all velocities reversed. Since Boltzmann's entropy formula does not care about the flips of all velocities, it follows that one has found a microscopic state with the property that the entropy has now changed from the large value $S^B(X_t)$ to the smaller value $S^B(X_0)$. This contradicts the second law of thermodynamics.

It seems that the only assumption we made that is not based on strong empirical evidence is the validity of the ideas of kinetic gas theory, so this argument seems to disprove the validity of these ideas.

2.2 Wiederkehreinwand

The objection raised by Zermelo in 1896 has some similarities to Loschmidt's paradox. It also considers a micro-state X_0 as above that evolves into a state X_t of larger entropy. Now, Poincaré's recurrence theorem states that for any distance d, no matter how small it might be, there will be a time $s > t$ such that $| X_0 - X_s | < d$.[1] Now choose d such that the Boltzmann entropy of all states Y with $| X_0 - Y | < d$ satisfies $S^B(Y) < \frac{1}{2} \left(S^B(X_0) + S^B(X_t) \right) < S^B(X_t)$. Note that the the Boltzmann entropy is proportional to the logarithm of the phase-space volume occupied by the macro-state. This phase-space volume is continuous in the respective phase-space density of the macro-state and thus is continuous in the region $U_d(X_0)$. Hence, due to the intermediate value theorem, such a choice for Y is is always possible. It follows that we found a time $s > t$ with $S^B(X_s) < S(X_t)$ which again, contradicts the second law.

[1] In fact there is an infinite number of such times s. This, however, is irrelevant for the argument.

3 Refutation of These Objections

Unifying different physical theories respectively explaining the physics of a theory from another which is considered more fundamental, is the central goal of (theoretical) physics. The beauty of Boltzmann's works where he derives the laws of thermodynamics in very simple terms from the classical motion of particles is truly stunning. Given this, it is very surprising, that the kinetic theory of gases was subject to objections based on philosophical arguments, as Einstein put it.

Concerning the objections of Loschmidt and Zermelo, the answer of Boltzmann came rather prompt [4–6].

Umkehreinwand Boltzmann explains that the physical behavior of the system is not only determined by the laws of motion of the particles, but also by the initial condition. To argue for the growth of entropy of a given system, it is relevant to consider a typical initial micro-state of the system in its initial macro-state. The argument given by Loschmidt is based on a micro-state related to the macro-state at time t. Reversing all velocities one gets a very special state among all micro-states belonging to that macro-state, not a generic one ([5]).

In his reply to Loschmidt, Boltzmann moreover emphasizes that the second law of thermodynamics – in contrast to the first one – makes a statement that only holds with high probability: "Nevertheless Loschmidt's theorem seems to me to be of the greatest importance, since it shows how intimately connected are the second law and probability theory, whereas the first law is independent of it" [5].

Wiederkehreinwand Boltzmann agrees that there will be recurrence to the initial state up to arbitrary precision and that, following the trajectory, there will be a time interval where the entropy decreases. However the time scale of the recurrence is for the systems under consideration exorbitantly higher than the life time of the universe. Thus we will never experience such a decrease in entropy. Since the second law of thermodynamics is based on empirical considerations, it has to be formulated accordingly. Boltzmann roughly estimated the recurrence time of a system and wrote in reply to Zermelo: "Though this calculation makes no pretense to accuracy, it still shows that it cannot be proved from Poincaré's theorem that the theoretical existence of a recurrence time involves any contradiction with experience, since the length of this time makes any attempt to observe it ridiculous" [4].

3.1 Final Breakthrough

It was Einstein's work on Brownian motion that lead to a break-through of the kinetic theory of gases. Brownian motion was described in the first half of the 19th century by Robert Brown who found that the organelles of pollen are subject to an erratic movement when exposed to water. "After bursting, contained similar sub-cylindrieal par-tides, in reduced numbers however, and mixed with other particles, at least as numerous, of much smaller size, apparently spherical, and in rapid oscil-

latory motion" [8]. He found the same type of motion in other materials: "in a word, every mineral which I could reduce to a powder, sufficiently fine to be temporarily suspended in water, I found these molecules more or less copiously".

The first theoretic explanation of this effect goes back to Gouy [20], who argues that Brownian motion is a natural consequence of the kinetic theory of gases resp. atomism. In 1905 Einstein came up with a similar idea in [18] and calculates the drift of a Brownian particle of diameter $1\mu m$ at room temperature during one minute. Five years later Perrin tested Einstein's predictions experimentally [29]. It was this very accurate derivation of an formerly unexplained effect that lead to the break-through of the kinetic theory of gases (see also [15]).

3.2 The Second Law

Boltzmann gave a clear explanation on how the growth in entropy, i.e., the second law, can be argued for coming from a kinetic theory of gases. He defines entropy to be a constant times the logarithm of the number of micro-states associated to a certain macro state, respectively – for the non-discrete situation – as a constant times the Lebesgue-measure of the subset in phase-space that includes all micro-states that belong to a certain macro state. Boltzmann supports his argument by proving the famous H-theorem where he shows that all solutions of the Boltzmann equation have a non-decreasing entropy. Further, it is easily possible to find examples of solutions where the entropy in fact grows at least for some instance of time. Now, since Boltzmann's equation has been derived with mathematical rigor from the dynamics of many interacting, Newtonian particles (at least for short times; see [21]), the validity of Boltzmann's argument does not leave any room for doubts.

What is striking on the principal explanation is the generality with which it holds. Understanding Boltzmann's reasoning one sees that the validity of the second law is independent from the dynamics under consideration: Assume that there is a normalizable measure which is invariant under the dynamics of the system. Further assume some weak form of ergodicity, that is, assume that the trajectory moves eventually through the whole support of the measure. If the latter assumption does not hold, for example, if the system lives on a certain energy shell, reduce the underlying space accordingly. Then, starting in a macro-state of low measure (i.e. low entropy) the system will most likely propagate into a macro-state that occupies a larger measure (entropy) and finally arrive the macro-state of maximal measure, the so-called equilibrium state (see Fig. 1). The reverse process will not be seen for a typical trajectory. Starting at equilibrium one will stay in equilibrium practically forever and most likely not propagate into the tiny areas that stand for macro-states of lower entropy in reasonable times.

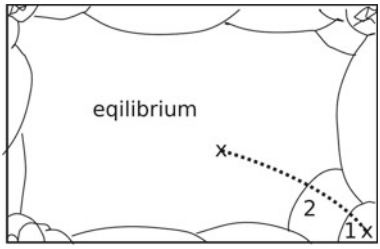


Fig. 1 This figure illustrates the growth of entropy of entropy under time evolution. Each point inside of the box stands for a micro-state of the system. The limitation of the box is made to illustrate that the static measure is normalizable. The different areas marked in the box stand for different macro-states. We assume that the system starts in a macro-state of low entropy, in the illustration this is the macro-state denoted by 1, the initial micro-state is marked with a cross. The propagation of the state goes along the dotted line. The trajectory thus goes through the macro-state number 2 and eventually reaches equilibrium, thus from boxes of smaller size to large ones. Since the Boltzmann entropy is defined as the logarithm of these sizes, one readily sees the growth in entropy, i.e. the second law. Note, that the sizes of the different boxes are drawn in a way to make them visible. For most practical situations one should think of the area occupied by macro-state number 2 a zillion times smaller than equilibrium, the area of macro-state 1 a zillion times smaller than that of macro-state 2. We wish to emphasize the simplicity and stability of the argument: Think of some trajectory starting in one of the regions of small size in this picture. It will practically always show an increase of entropy of the macro-states it visits

4 Emergence of Irreversibility

Since Boltzmann's, Clausius' and Maxwell's findings that the physics of the macroscopic world, in particular of gases, can be explained by the motion of atoms, there has been a vivid discussion about the emergence of irreversibility: While the dynamics of the atoms is generally assumed to be time-reversal invariant (observing Newton laws), the dynamics of macroscopic objects show in general a preferred direction of time. This seeming discrepancy has been coined the problem of irreversibility respectively of the thermodynamic arrow of time.

There have been several attempts to explain the emergence of the arrow of time. In the following we will mention the most important ones. Please note again, that we will in the following focus on in-principle possibilities of how a microscopic reversible system can lead to an irreversible behavior of the corresponding macroscopic world and not aim to explain the source of irreversibility in *our* universe.

4.1 Microscopic Irreversibility

One possibility is the trivial one: assuming an irreversible model for the microscopic system. This would not only lead to the trivial fact that there will be a preferred direction of time on the macroscopic level, it is also possible to argue for the less

obvious fact that the quality of irreversible systems is very much the way as we experience. The second law of thermodynamics says much more, not only that there is a preferred direction of time on the macroscopic level.

4.2 Fluctuation Hypothesis

"One has the choice of two kinds of pictures. One can assume that the entire universe finds itself at present in a very improbable state. However, one may suppose that the eons during which this improbable state lasts, and the distance from here to Sirius, are minute compared to the age and size of the universe. There must then be in the universe, which is in thermal equilibrium as a whole and therefore dead, here and there relatively small regions of the size of our galaxy (which we call worlds), which during the relatively short time of eons deviate significantly from thermal equilibrium. Among these worlds the state probability increases as often as it decreases. For the universe as a whole the two directions of time are indistinguishable, just as in space there is no up or down. However, just as at a certain place on the earth's surface we can call "down" the direction toward the center of the earth, so a living being that finds itself in such a world at a certain period of time can define the time direction as going from less probable to more probable states (the former will be the "past" and the latter the "future") and by virtue of this definition he will find that this small region, isolated from the rest of the universe, is "initially" always in an improbable state. This viewpoint seems to me to be the only way in which one can understand the validity of the second law and the heat death of each individual world without invoking an unidirectional change of the entire universe from a definite initial state to a final state". [6].

If there is a thing in Boltzmann's work which has not been carefully thought to the end, it is the idea he presents in this quote. Before being too harsh with the criticism one has to note that Boltzmann was not so much interested in explaining mechanisms on a global scale, but deriving the laws of thermodynamics from the kinetic theory of gases for situations that are empirically accessible. He was very skeptical with the idea of extrapolating his thoughts to a larger scale: "If perhaps this reduction of the second law to the realm of probability makes its application to the entire universe appear dubious, yet the laws of probability theory are confirmed by all experiments carried out in the laboratory" [5]. It is, of course, true that, for a system of Newtonian particles in a box, arbitrary large fluctuations in the entropy will occur if one just waits long enough. This fact is combined with an anthropic principle, i.e., the idea that a living being needs a large deviation from equilibrium to exist. Under the condition that living beings exist one arrives at states of very low entropy. The idea that this explains the arrow of time in our universe is called "fluctuation hypothesis". The hypothesis has been refuted by Eddington [16]: Assume that we are in a fluctuation which allows for the existence of living beings, for example a certain number of persons attending a conference. Still one should expect to have chaotic behavior outside of the venue of this conference, which is in contradiction

to what we in fact see in our world. In other words: conditioning on the existence of one person who thinks about nature, the existence of another being requires even more order and is thus very unlikely. Using random fluctuations as an argument and the anthropic principle, one should thus expect the existence of a single being, or even a single brain [10] without a body to exist in this completely chaotic world.

4.3 Past Hypothesis

One possible option is to use Boltzmann's argument that explain the second law from kinetic gas theory on the global scale. The idea, which is nowadays often referred to as "past hypothesis" in the literature [2], goes back to Boltzmann [6] "The second law of thermodynamics can be proved from the mechanical theory if one assumes that the present state of the universe, or at least that part which surrounds us, started to evolve from an improbable state and is still in a relatively improbable state. Hence.... the system will initially be in an improbable state, and ... will always precood to more probable states."

Boltzmann's idea was that one just has to assume that initially the system was in a macro-state of low entropy. He explains later, that there might be micro-states that in fact – as Zermelo explains – propagate into macro-states of even smaller entropy, however, these micro-state are extremely rare and can thus be excluded. To explain the assumption in more detail we give a slightly modified version proposed by Chen [13]:

"We can introduce this as an explicitly time-asymmetric postulate in the theory, by using the Past Hypothesis:

Past Hypothesis (PH) At the initial time of the universe, the micro-state of the universe is in a low-entropy macro-state.

Given that some micro-states are anti-entropic, it is standard to introduce a probability distribution over the micro-states compatible with the low-entropy macro-state:

Statistical Postulate (SP) The probability distribution of the initial micro-state of the universe is given by the uniform one (according to the natural measure) that is supported on the macro-state of the universe.

However, a detailed probability distribution may be unnecessary. In the typicality framework, we just need to be committed to a typicality measure:

Typicality Postulate (TP) The initial micro-state of the universe is typical inside the macro-state of the universe."

The "natural measure" the author refers to is of course the static measure, i.e., the measure that is invariant under the dynamics of the system.

The past Hypothesis thus lifts the assumptions Boltzmann made for laboratory systems to the global scale.

4.4 Non-normalizable Invariant Measure

In this manuscript we will focus on yet another possible explanation of the emergence of an arrow of time, namely an universe without any equilibrium state at all. The central idea of our investigation goes back to the works of Antonov ([3]), Lynden-Bell ([25]), Padmanabhan ([28]) and Carroll and Chen ([1, 9, 11]). It goes as follows: assume that for any macro-state which can be accessed by the system, there is another macro-state which can be also accessed by the system and covers a larger number of micro-states. No matter the macro-state of the system, there is always a macro-state of larger disorder which can be accessed as time evolves. Consequently, for most initial micro-states from any macro-state of the system, there will be propagation to a macro-state of higher entropy (see Fig. 2). Of course, all this is only possible under the assumption that the invariant measure for the system can not be normalized and in particular under the absence of an equilibrium state.

Note, that the goal of this paper is not to find a possible explanation for *our* universe but to explain the in-principle idea on how a time reversal system can give irreversibility on the macroscopic level. Therefore, we will define a model universe where the growth of entropy can be shown for typical initial conditions. To prove this with mathematical rigor for a suitable model is a work in progress, in the present manuscript we will summarize the ideas that lead to a suitable model and give some heuristic ideas to argue for its properties. We will restrict ourselves to classical Newtonian gravity, neither considering quantum cases nor general relativity.

There are already results in that direction. Our findings are in perfect agreement with [19] where a simple toy model – a non-interacting gas in $\mathbb{R}^3$ – is discussed. Considering the time dependence of the momentum of inertia I, the authors argue that, for any initial state, I has a parabolic form. Note that the number of micro-

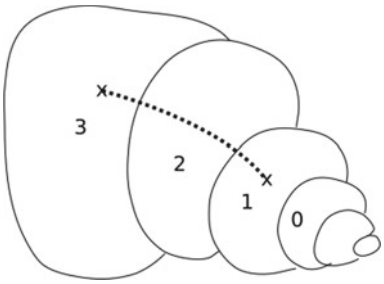


Fig. 2 This figure should be understood in the same way as Fig. 1 with the difference that there is no normalizable measure and thus no equilibrium state. Thus one has to think of an unbounded situation. We assume that there are ever growing regions depicting the different macro-states. No matter where we start with the time evolution, there will always be macro-states of larger volume and one might expect that a general trajectory will typically propagate from macro-states of ever growing volume. In contrast to 1 one needs not assume that the trajectory starts in a very special state, namely a state of low entropy. For practically any initial state one expects a growth in entropy

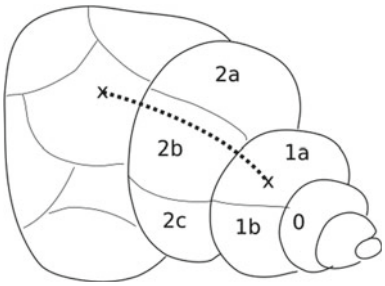


Fig. 3 This figure illustrates the importance of considering all relevant macro-variables when calculating the change in entropy of a system. Consider a system is described by the macro-variables A and B. Assume that the sets of all micro-states corresponding to different values of A behave like drawn in Fig. 2, that is that the number of micro-states changes for the different values of A and one might expect a growth in entropy while the value of the macro-variable changes from one to two and goes on to three. However, it might be, that considering the macro-variables A and B together, the respective number of micro-states stays constant. In the situation illustrated the values for A and B together changes from $(1, a)$ to $(2, b)$. The respective regions have identical volume. As long as the full description of the macro-state (up to irrelevant features) is not considered one can not judge whether the entropy is in fact growing or not

states grows with I. On a first view this seems to indicate that the entropy of the system also has a parabolic shape. However, one has to be careful. While this toy model is handy to explain the main features, it does in fact not result in a growth of entropy as the authors explain in a footnote of [19] "If we adopted the choice of macro-variables from the kinetic theory of gases ... then entropy would never change in the toy model".

The argument for this can found in the master thesis of Paula Reichert [30]: Not only the macro-variable I changes as time evolves, also the local temperature T (defined as the variance of the momenta) of the system does. Considering both macro-variables, i.e., taking the intersection of all micro-states with the right value of I and all micro-states with the right local temperature one finds that the volume of *this* set is in fact conserved under time evolution (Fig. 3).

4.4.1 Macro-Variables and Entropy

The toy model discussed above shows that defining the entropy by counting the number of micro-states for given values of macro-variables is dangerous. One has to be sure, that the macro-variables one considers give a complete description of the macro-state of the system. Assume that the macro-states a system is in during its time evolution are described by the macro-variables A, B, C. Considering a sub-set of those three and counting all micro-states related to this sub-set of macro-variables and taking the logarithm will not give the Boltzmann entropy. In the toy model example above, taking I or T alone this procedure would once give a value which

increases once a value that decreases. None of them can be seriously identified with entropy, which remains in fact constant for the system at hand.

But how can we be sure that I and T give a full description of the system? Or in general, that for the model under consideration a certain set of macro-variables suffices to uniquely describe the macroscopic state? This is hard to answer and it is not surprising that Boltzmann's entropy formula does not use macro-variables but rather the density on phase space to describe the macro-state of a gas instead: The entropy is given by the logarithm of the number of micro-states which approximate a certain phase-space density.

For the simple toy model, i.e. gas evolving freely, particles will stay independent. By Liouville's theorem it follows that the density of the gas will remain unchanged up to phase-space-volume-conserving transformations of the support of the function. It follows that the number of micro-states that describe this density will be a constant of motion, thus the entropy will not change. Thus to get a model which shows the behavior described in [11] one has to modify the toy-model.

In [22] Lazarovici and Reichert claim that for a many particle system evolving in $\mathbb{R}^3$ and subject to Newtonian gravity one will get an increase in entropy for practically all initial states of the system (up to a Lebesgue-zero set). To support their argument they calculate the number of micro-states for a given set of variables. However, they overlook that for very long times also such a system will cool down: It is known that any gravitating system which contains more than two particles will be unstable. So in the very long run one should not expect clusters, but only singles and pairs of particles. Their centers of masses will – on the very long time scale – move freely. Thus in the long run the system will behave exactly like the simple toy model discussed in [19, 30] and above.

This shows that temperature is in fact an important macro-variable of the system and not a simple function of the kinetic energy.[2]

4.4.2 Confined Universe

As explained above, a N-particle system with Newtonian gravity that is not confined will asymptotically approach the behavior of a free gas as time goes to infinity, hence the entropy of the system will stay bounded under its propagation. This might still result in an growth of the entropy for not too large times for a large class of initial states, however, the entropy increase will eventually slow down and entropy will remain bounded as times goes to infinity. Since we are in a situation of a non-normalizable static measure, typicality arguments are quite tricky. Thus the quality of the macro-states – in particular their probability – that show a growth in entropy is

[2] Note that temperature is a form of kinetic energy, but not kinetic energy per se. A stone falling in a gravitational field in the absence of friction will increase the total kinetic energy of the system while the temperature of the stone remains unchanged. While temperature is proportional to the variance in the momentum-direction and thus directly related to disorder of the system and thus entropy, kinetic energy is not. That is why temperature should be used as a macro-variable, not kinetic energy.

unclear. If the number of micro-states which belong to the macro-states with entropy close to the maximum is the vast majority, why should it be natural to expect that the initial state belonged to a macro-state whose entropy would grow under time-evolution?

This is a tricky problem and we suggest another modification of the models discussed so far. The keyword for this modification is mixing: High kinetic energy of the system will result in a high temperature if we manage to introduce a spatial mixing of the particles. This can be achieved by putting the system on a torus (or some other bounded region, preferably with periodic boundary conditions).

Therefore, the phase space of the system we wish to investigate is given by

$$\mathcal{P} = \mathbb{S}_1^{\times 3N} \times \mathbb{R}^{3N} \tag{1}$$

where $\mathbb{S}_1$ denotes the one-dimensional sphere. The dynamics of the system is given by the Newtonian equations of motion with attractive Coulomb interaction, i.e., for $X = (Q, P) \in \mathcal{P}$ we have

$$\begin{aligned} \dot{Q} &= P \\ \left(\dot{P}\right)_j &= \sum_{k \neq j} \frac{q_k - q_j}{|q_k - q_j|^3} \end{aligned}$$

where for any vector $A \in \mathbb{R}^{3N}$ the $(A)_j$ stands for the coordinates of the j^{th} three-vector component of A, i.e. $(A)_j = a_j \in \mathbb{R}^3$ for $A = (a_1, a_2, \ldots, a_N) \in \mathbb{R}^{3N}$.

Antonov ([3]) expected that for a similar system the entropy grows unboundedly. The heuristic argument for that is the following:

Since the potential energy of the system is unbounded from below due to the Coulomb singularity, the kinetic energy will exceed any limit. Due to the mixing of the particles' positions – caused by the evolution on the torus – the particles will not be ordered according to their velocity. Thus a growth in the kinetic energy results in an increase in temperature. Hence, the system will heat up more and more and the respective phase space volume will thus grow further and further.

To make the heuristic argument clearer, think of two particles which are very close to each other and, due to attractive Coulomb interaction and since the other particles are relatively far away, form a rather stable system. Like a double star these two particles will rotate around their center of mass. Forces (analogous to tidal forces) coming from these two particles will act on all other particles in the system and heat them up, while the two particles will get closer and closer. In the general setting, i.e., all the particles interacting with each other via Coulomb interactions, we expect that such clusters will form and the same argument as before applies.

To get a model in which the entropy can grow infinitely, it is crucial that the momenta of the particles are not bounded and hence the phase-space of the system has infinite volume. Otherwise, the invariant measure – which for this model is the Lebesgue measure on phase space – would be normalizable and entropy would be bounded from above (cf. [12]).

Note that the relevance of the fact that we consider a system on a torus (or some other region of finite volume) is twofold: First, it results in a mixing of the positions as described above and thus guarantees that an increase in the kinetic energy is translated in an increase in temperature. Second, on $\mathbb{R}^{6N}$, one shall not expect a growth of the kinetic energy above any threshold in the first place. As time evolves, the particles will increase their distances further and further and the transfer of potential energy into kinetic energy will eventually come to an end. This is different on the torus, where the gravitational interaction will be relevant on all time-scales.

There are no cosmological observations that indicate that our universe may not be finite in space. Hence, the proposed model might indeed give an explanation which is in principle valid for our universe.

Acknowledgements The authors are most grateful to Detlef Dürr for everything he did for us, and in particular for starting the collaboration on the topic and for all the inspiring meetings we had over the years.

We also wish to thank Julian Barbour, Tim Koslowski, Dustin Lazarovici, Tim Maudlin and Paula Reichert for helpful discussions on the topic.

The authors acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project-number 417613818 ("A microscopic model for explaining the arrow of time").

References

1. A. Aguirre, S.M. Carroll, C Johnson, Out of equilibrium: understanding cosmological evolution to lower-entropy states. J. Cosmol. Astropart. Phys. **2012**(02), 024 (2012)
2. D. Albert. *Time and Chance* (Harvard University Press, 2000). https://doi.org/10.2307/j.ctvjsf57g.
3. V. Antonov, Most probable phase distribution in spherical star systems and conditions for its existence, in *Dynamics of Star Clusters*, ed. by J. Goodman, P. Hut. Vol. 113. (1985), p. 525. Originally published: V.A. Antonov. In: Vast. Leningrad University **7**(135) (1962)
4. L. Boltzmann, Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo. Annalen der Physik und Chemie **293**(4), 773–784 (1896). https://doi.org/10.1002/andp.18962930414. English translation in: S.G. Brush, *Kinetic Theory: Irreversible Processes* (Elsevier, 2016)
5. L. Boltzmann, Über die Beziehung eines allgemeinen mechanischen Satzes zum zweiten Hauptsatze der Wärmetheorie (1970), pp. 240–247. https://doi.org/10.1007/978-3-322-84986-1_5. English translation in: S.G. Brush, *Kinetic Theory: Irreversible Processes* (Elsevier, 2016)
6. L. Boltzmann, Zu Hrn. Zermelo's Abhandlung "Über die mechanische Erklärung irreversibler Vorgänge". Annalen der Physik und Chemie **296**(2), 392–398 (1897). https://doi.org/10.1002/andp.18972960216. English translation in: S.G. Brush, *Kinetic Theory: Irreversible Processes* (Elsevier, 2016)
7. J. Bricmont, Science of chaos or chaos in science?. Ann. N. Y. Acad. Sci. **775**(1), 131–175 (1995). https://doi.org/10.1111/j.1749-6632.1996.tb23135.x
8. R. Brown, A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. Philos. Mag. **4**(21), 161–173 (1828). https://doi.org/10.1080/14786442808674769.
9. S. Carroll, *From Eternity to Here: The Quest for the Ultimate Theory of Time* (Dutton, 2010)

10. S. Carroll, Why Boltzmann Brains Are Bad, in *Current Controversies in Philosophy of Science*, ed. by S. Dasgupta, B. Weslake, R. Dotan (Routledge, 2020), pp. 7–20
11. S. Carroll, J. Chen, Does inflation provide natural initial conditions for the universe?. Int. J. Mod. Phys. D **14**(12), 2335–2339 (2005). https://doi.org/10.1142/s0218271805008054
12. S. Carroll, J. Chen, Spontaneous inflation and the origin of the arrow of time (2004). arXiv.hep-th/0410270
13. E.K. Chen, The past hypothesis and the nature of physical laws, in *The Probability Map of the Universe: Essays on David Albert's Time and Chance*, ed. by B. Loewer, E. Winsberg, B. Weslake (Harvard University Press, Cambridge, Masschusetts, 2023), pp. 204–248
14. D. Dürr, S. Goldstein, N. Zanghí, *Quantum Physics Without Quantum Philosophy* (Springer, Berlin, Heidelberg, 2013). https://doi.org/10.1007/978-3-642-30690-7.
15. D. Dürr, S. Teufel, *Bohmian Mechanics* (Springer, 2009)
16. A. Eddington, The end of the world: from the standpoint of mathematical physics. Nat. **127**(3203), 447–453 (1931). https://doi.org/10.1038/127447a0.
17. A. Einstein, The advent of the quantum theory. Sci. **113**(2926), 82–84 (1951). https://doi.org/10.1126/science.113.2926.82.
18. A. Einstein, Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Annalen der Physik **322**(8), 549–560 (1905). https://doi.org/10.1002/andp.19053220806.
19. S. Goldstein, R. Tumulka, N. Zanghí, Is the hypothesis about a low entropy initial state of the Universe necessary for explaining the arrow of time?. Phys. Rev. D **94**(2), 023520 (2016). https://doi.org/10.1103/PhysRevD.94.023520., https://link.aps.org/doi/101103/PhysRevD.94.023520
20. L.G. Gouy, Note sur le mouvement brownien. Journal de Physique Théorique et Appliquée **7**(1), 561–564. https://doi.org/10.1051/jphystap:018880070056101.
21. O. Lanford, Time evolution of large classical systems, in *Dynamical Systems, Theory and Applications: Battelle Seattle 1974 Rencontres* (2005), pp. 1–111
22. D. Lazarovici, P. Reichert, Arrow(s) of time without a past hypothesis, in *Statistical Mechanics and Scientific Explanation* (World Scientific, 2020), pp. 343–386. https://doi.org/10.1142/9789811211720_0010
23. J. Loschmidt, *Über den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft: I - IV*. (aus der KK Hof-und Staatsdruckerei, 1876)
24. T.L. Carus, *De rerum natura*. Vol. Book II. Ca. 50 BC. Chap. 2. English translation: William Ellery Leonard et al. On the Nature of Things. (Courier Corporation, 2012)
25. D. Lynden-Bell, Negative specific heat in astronomy, physics and chemistry. Phys. A: Stat. Mech. Its Appl. **263**(1–4), 293–304 (1999)
26. E. Mach, *Die Principien der Wärmelehre. Historisch-kritisch entwickelt*. Leipzig: Barth, 1896. VIII, 472. English translation: B. McGuinness (ed.), *Principles of the Theory of Heat: Historically and critically elucidated.* (Springer, 2012). https://doi.org/10.1007/978-94-009-4622-4.
27. J.C. Maxwell, *Matter and Motion* (Cambridge University Press, 2010). https://doi.org/10.1017/cbo9780511709326
28. T. Padmanabhan, Antonov instability and gravothermal catastrophe-revisited. Astrophys. J. Suppl. Ser. **71**, 651–664 (1989)
29. J. Perrin, Mouvement Brownien et molécules. J. Phys. Theor. Appl. **9**(1), 5–39 (1910)
30. P. Reichert, Can a parabolic evolution of the entropy of the universe provide the foundation for the second law of thermodynamics?. MA thesis, 2012. http://www.mathematik.uni-muenchen.de/~bohmmech/theses/Reichert_Paula_MA_v2.pdf
31. E. Zermelo, Über einen Satz der Dynamik und die mechanische Wärmetheorie. In: *Annalen der Physik und Chemie* **293**(3), 485–494 (1896). https://doi.org/10.1002/andp.18962930314.

The Ergodic Hypothesis: A Typicality Statement

Paula Reichert

Abstract This paper analyzes the ergodic hypothesis in the context of Boltzmann's late work in statistical mechanics, where Boltzmann lays the foundations for what is today known as the typicality account. I argue that, based on the concepts of stationarity (of the measure) and typicality (of the equilibrium state), the ergodic hypothesis, as an idealization, is a consequence rather than an assumption of Boltzmann's account. More precisely, it can be shown that every system with a stationary measure and an equilibrium state (be it a typical state with respect to the phase space or the time average) behaves essentially as if it were ergodic. I claim that Boltzmann was aware of this fact as it grounds both his notion of equilibrium, relating it to the thermodynamic notion of equilibrium, and his estimate of the fluctuation rates.

1 Introduction

The ergodic hypothesis has been formulated by [1, 2] and has famously been discussed by [3] in their influential encyclopedia article on statistical mechanics, where they provide an overview of and comment on Boltzmann's work in statistical physics.

Ever since, the ergodic hypothesis has been debated controversially. This refers not only to the status of the ergodic hypothesis within Boltzmann's work (see, e.g., [4]), but more generally to its applicability with respect to realistic systems (see, e.g., [5, 6]) and its relevance for physics as such (see, e.g., [7–9]).

Despite its debatable status, the concept of ergodicity has attracted a lot of attention. Today there even exists a proper branch of mathematics, so-called ergodic theory, with a plentitude of rigorous mathematical results (most notably, the results of [10–12]; see [13] for an overview).

P. Reichert (✉)
Mathematisches Institut, Ludwig-Maximilians-Universität München, München, Germany
e-mail: reichert@math.lmu.de

Department for Humanities & Arts, Technion-Israel Institute of Technology, Haifa, Israel

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_20

Interestingly enough, though, Boltzmann himself never highlighted the ergodic hypothesis. Although he introduces it in his early work, he mentions it not even once in his two volumes on gas theory, which constitute his *opus magnum* on statistical mechanics (cf. [14]). Still, he seems to rely on ergodicity, at least as an idealization, also in his later work like, for instance, when he estimates the rate of fluctuations in the letter to Zermelo (cf. [15]).

This said, has ergodicity been a fundamental assumption of Boltzmann as the Ehrenfests suggest? If so, why didn't he make this more explicit? This seems the more surprising as he does emphasize the explanatory value of other concepts. For instance, he stresses the fact that equilibrium is a typical state, i.e., a state which is realized by an overwhelming number of micro configurations, at several points throughout his work (see, e.g., [14–16]).

In this paper, I argue that ergodicity, as an idealization, or essential ergodicity, in the strict sense (as defined in Sect. 3.3 below), is a consequence rather than an assumption of Boltzmann's approach. Based on this, I claim that the ergodic hypothesis should be read as a typicality statement, in a way analogous to how Boltzmann taught us to read the H-theorem (see [15, 16]). That is, just as a dynamical system of many particles doesn't approach equilibrium for all, but for typical initial conditions (given a low-entropy initial macrostate) and stays there not for all, but for most times, in the case of ergodicity, not all, but typical systems behave not strictly, but essentially, that is qualitatively, as if they were ergodic.

To make this point precise, what can be shown is the following: *On typical trajectories, the time and phase space averages of physical macrostates coincide in good approximation*. This property of the dynamics, which I call 'essential ergodicity', follows from the stationarity of the measure and the typicality of the equilibrium state alone.

2 The Ergodic Hypothesis

To discuss the ergodic hypothesis, we need to introduce the realm of Boltzmann's statistical mechanics: the theory of measure-preserving dynamical systems.

2.1 Measure-Preserving Dynamical Systems

Let $(\Gamma, \mathcal{B}(\Gamma), T, \mu)$ denote a Hamiltonian system. For N particles, $\Gamma \cong \mathbb{R}^{6N}$ is called phase space. It is the space of all possible microstates X of the system, where a point $X = (q, p)$ in Γ represents the positions and momenta of all the particles: $(q, p) = (q_1, ..., q_{3N}, p_1, ..., p_{3N})$.

The Hamiltonian flow T is a one-parameter flow $T^t(q, p) = (q, p)(t)$ on Γ with t representing time. It is connected to the Hamiltonian vector field v_H as follows: $v_H(T^t(q, p)) = dT^t(q, p)/dt$. In other words, the flow lines are the integral curves along the Hamiltonian vector field, where the latter is specified by $v_H = (\partial H/\partial p, -\partial H/\partial q)$. This is the physical vector field of the system, generated by the Hamiltonian H, and the flow lines represent the possible trajectories of the system. Finally, μ refers to the Liouville measure,

$$d\mu = \prod_{i=1}^{3N} dq_i dp_i, \tag{1}$$

or to any other stationary measure derived thereof.

Note that we call a measure μ *stationary* (with respect to T) if and only if the flow T is measure-preserving (with respect to μ). Given a Hamiltonian system, it follows from Liouville's theorem that the Liouville measure is conserved under the Hamiltonian phase flow. That is, for every $A \in \mathcal{B}(\Gamma)$,

$$\mu(T^{-t}A) = \mu(A). \tag{2}$$

Since the Liouville measure is just the $6N$-dimensional Lebesgue measure, this says that phase space volume is conserved under time evolution.

If we introduce the notion of the time-evolved measure, $\mu_t(A) := \mu(T^{-t}A)$, we can reformulate the condition of stationarity as follows. A measure μ is *stationary* if and only if, for every $A \in \mathcal{B}(\Gamma)$,

$$\mu_t(A) = \mu(A). \tag{3}$$

According to this equation, the measure itself is invariant under time translation, which is the main reason for physicists to accept it as *the* measure grounding a statistical analysis in physics (see, e.g., [3, 17, 18]). In practice, we are not concerned with the Liouville measure *per se*, but with appropriate stationary measures derived thereof.[1]

[1] Consider, for instance, an isolated system. Within that system, total energy E is conserved. Hence, trajectories are restricted to the constant-energy hypersurface $\Gamma_E = \{(q, p) \in \Gamma | H(q, p) = E\}$, from which it follows that the microcanonical measure

$$d\mu_E = \prod_{i=1}^{3N} dq_i dp_i \, \delta(H(q, p) - E)$$

is the appropriate stationary measure of the dynamics in that case.

2.2 Variants of the Ergodic Hypothesis

Within the framework of Hamiltonian systems or, more generally, measure-preserving dynamical systems, we can analyze Boltzmann's ergodic hypothesis.

Let again $(\Gamma, \mathcal{B}(\Gamma), T, \mu)$ be a measure-preserving dynamical system and $A \in \mathcal{B}(\Gamma)$. Let, in what follows, $\mu(\Gamma) = 1$.[2] We call

$$\mu(A) = \int_{\Gamma} \chi_A(x) d\mu(x) \tag{4}$$

the 'phase space average' of A with χ_A being the characteristic function which is 1 if $x \in A$ and 0 otherwise. Further we call

$$\hat{A}(x) = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \chi_A(T^t x) dt \tag{5}$$

the 'time average' of A for some $x \in \Gamma$. Here it has been proven by [10] that the infinite-time limit exists pointwise almost everywhere on Γ and the limit function $\hat{A}(x)$ is integrable.

A dynamical system is called *ergodic* if and only if, for all $A \in \mathcal{B}(\Gamma)$ and almost all $x \in \Gamma$ (i.e. for all x except a measure-zero set), the time and phase averages coincide:

$$\mu(A) = \hat{A}(x). \tag{6}$$

In other words, a system is called ergodic if and only if, for almost all solutions, the fraction of time the system spends in a certain region in phase space (in the limit $t \to \infty$!) is precisely *equal* to the phase space average of that region.

Historically, the ergodic hypothesis has been formulated differently. In its original version due to [1] (cited by [3]), it refers to the assertion that a trajectory literally has to *go through every point* in phase space (more precisely, in the constant-energy hypersurface). But this would imply that there is only *one* solution with all possible microstates belonging to one and the same solution. This has been proven impossible by [21, 22].

In a weaker formulation, the so-called 'quasi-ergodic hypothesis' demands that a trajectory has to *come arbitrarily close to every point* in phase space (see [3]). Later, the results of [10, 11] established the precise conditions under which equality of the time and phase space average is obtained.[3]

---

[2] Throughout this paper, we deal with systems where Γ is finite and, hence, μ is normalizable. In that case, we can set $\mu(\Gamma) = 1$ without loss of generality. The hard case of infinite phase spaces has to be discussed elsewhere (see [19, 20] for a first discussion).

[3] Reference [10] gives a definition of ergodicity in terms of invariant sets (where a set $A \in \mathcal{B}(\Gamma)$ is called invariant if and only if $T^{-1}A = A$). If, for all sets $A \in \mathcal{B}(\Gamma)$ with $T^{-1}A = A$,

$$\mu(A) = 0 \qquad \text{or} \qquad \mu(A) = 1,$$

For realistic physical systems, this equality of the time and phase space average—that is, ergodicity—turned out to be extremely hard to prove, if it can be proven at all (it took almost 50 years to extend the proof of [23] which held for the model of one billiard ball on a 2-dimensional table to the full model of $N \geq 2$ hard spheres in a container with periodic boundary conditions, i.e. a torus, of dimension $d \geq 2$; see [24]).

At this point, the question arises: What if we were not interested in the exact coincidence of the time and phase space average in the first place? What if all we need is an approximate equality of the time and phase space average on typical trajectories? The point I want to make is the following: Boltzmann, being concerned with the analysis of realistic physical systems, need not be and presumably was not interested in ergodicity in the strict sense. According to [3], Boltzmann used ergodicity to estimate the fraction of time a system spends in a certain macrostate. To obtain such an estimate, however, it suffices to establish a result qualitatively comparable to ergodicity: *an almost equality of the time and phase space average of physical macrostates on typical trajectories.* This is precisely where the notion of essential ergodicity comes into play.

3 Essential Ergodicity

We need one last ingredient to grasp the notion of essential ergodicity and that is the notion of typicality of macro- and microstates. We will then find that, given a stationary measure and a typical macrostate, that is, an equilibrium state in Boltzmann's sense, a typical system behaves essentially as if it were ergodic.

3.1 Typicality and Boltzmann's Notion of Equilibrium

Given a measure on the space of possible states of the system—like a volume measure on phase space—this is naturally a measure of probability or typicality.[4] Let again μ denote the volume measure on Γ. We call a measurable set $A \subset \Gamma$ 'typical' (with respect to Γ) if and only if

$$\mu(A) = 1 - \varepsilon \tag{7}$$

---

then the system is called 'ergodic'. Thus a system is called 'ergodic' if and only if all invariant sets are of full or zero measure. In other words, there exist no two (or more) disjoint invariant sets of non-zero measure. The two definitions of ergodicity relate to one another via Birkhoff's theorem.

[4] There is a little caveat to this statement. While it is definitely true whenever phase space is finite and the measure is normalizable, one has to be careful with infinite phase spaces and non-normalizable measures. For problems related to the latter, see [25] or [19]. The distinction between the notions of probability and typicality has been drawn and discussed elsewhere (see, e.g., [26, 27] or [28]).

for $0 < \varepsilon << 1$. This definition of 'typical sets' directly entails a definition of 'typical points' (cf. [28]). We say that a point x is 'typical' (with respect to Γ) if and only if $x \in A$ and A is typical with respect to Γ.

In Boltzmann's statistical mechanics, we are concerned with 'points' (microstates) and 'sets' (macro-regions). Macro-regions are regions of phase space corresponding to physical macrostates of the system. More precisely, every microstate X, represented by a point (q, p) on Γ, belongs to respectively determines a certain macrostate $M(X)$, represented by an entire region $\Gamma_M \subset \Gamma$—the set of all microstates realizing that particular macrostate. While a microstate comprises the exact positions and velocities of all the particles, $X = (q_1, ..., q_N, p_1, ..., p_N)$, a macrostate $M(X)$ is specified by the macroscopic, thermodynamic variables of the system, like volume V, temperature T, and so on. By definition, any two macrostates M_i and M_j are macroscopically distinct, hence there are only finitely many macrostates M_i, and all macrostates together provide a partition of phase space into disjoint 'macro regions' Γ_{M_i} with $\Gamma = \bigcup_{i=1}^{n} \Gamma_{M_i}$. Here it is a consequence of the large number of particles that every macrostate $M(X)$ is realized by a huge number of microstates X and, hence, the precise way of partitioning doesn't matter.

In this set-up, Boltzmann defined 'equilibrium' precisely as the *typical* macrostate of the system.

Definition 1 (*Boltzmann equilibrium*) Let $(\Gamma, \mathcal{B}(\Gamma), T, \mu)$ be a dynamical system. Let Γ be partitioned into finitely many disjoint, measurable subsets $\Gamma_{M_i}, i = 1, \ldots, n$ by some (set of) physical macrovariable(s) M_i, i.e., $\Gamma = \bigcup_{i=1}^{n} \Gamma_{M_i}$. Then a set $\Gamma_{Eq} \in \{\Gamma_{M_1}, ..., \Gamma_{M_n}\}$ with phase space average

$$\mu(\Gamma_{Eq}) = 1 - \varepsilon \tag{8}$$

where $\varepsilon \in \mathbb{R}, 0 < \varepsilon << 1$, is called the 'equilibrium set' or 'equilibrium region'. The corresponding macrostate M_{Eq} is called the 'Boltzmann equilibrium' of the system.

Be aware that this definition is grounded on a particular, physical macro partition of phase space. In other words, it is not an arbitrary value of ε which, when given, determines an equilibrium state—such a definition would be meaningless from the point of physics. Instead, it is a partition determined by the physical macrovariables of the theory, which is given, and it is with respect to that partition that a region of overwhelming phase space measure, if it exists, defines an equilibrium state in Boltzmann's sense (and by the way determines the value of ε).

At this point, it has been Boltzmann's crucial insight that, for a realistic physical system of $N \approx 10^{24}$ particles (where, for a medium-sized object, we take Avogadro's constant) and a partition into macroscopically distinct states, there always exists a region of overwhelming phase space measure (see, e.g., [15]).[5] This follows essentially from the vast gap between micro and macro description of the system and the

[5] [29] proves the existence of a region of overwhelming phase space measure for a large class of realistic physical systems.

fact that, for a large number of particles, small differences at the macroscopic level translate into huge differences in the corresponding phase space volumes.

To obtain an idea of the numbers, consider a gas in a medium-sized box. For that model, [30, 31] estimates the volume of all non-equilibrium regions together as compared to the equilibrium region to be:

$$\frac{\mu(\bigcup_{i=1}^{n} \Gamma_{M_i} \setminus \Gamma_{Eq})}{\mu(\Gamma_{Eq})} = \frac{\mu(\Omega \setminus \Gamma_{Eq})}{\mu(\Gamma_{Eq})} \approx 1 : 10^N \tag{9}$$

with $N \approx 10^{24}$. This implies, with $\mu(\Gamma_{Eq}) \approx \mu(\Omega)$, that ε is of the order $1 : 10^N \approx 1 : 10^{10^{24}} = \frac{1}{10^{10000000000000000000000000}}$.

Both Boltzmann's realization that equilibrium is a typical state and his understanding that any two distinct macrostates relate to macro-regions that differ vastly in size provided the grounds for his explanation of irreversible behaviour (cf. [14, 15]; see [9, 32–34]) for further elaboration of this point). In the following, however, we are only concerned with ergodicity and, related to that, a system's long-time behaviour.

3.2 *Precise Bounds on the Time and Phase Space Average of the Equilibrium State*

In what follows, we give precise bounds on the time average of the equilibrium state. Therefore, consider a dynamical system with a stationary measure μ and an equilibrium state Γ_{eq} in the sense of Boltzmann. That is, $\mu(\Gamma_{Eq}) = 1 - \varepsilon$.

To be able to formulate the bound on the time average and, later, the notion of 'essential ergodicity', we have to distinguish between a 'good' set G and a 'bad' set B of points $x \in \Gamma$. Let, in what follows, B be the 'bad' set of points for which the time average of equilibrium $\hat{\Gamma}_{Eq}(x)$ is smaller than $1 - k\varepsilon$ (with $1 \leq k \leq 1/\varepsilon$). All points in this set determine trajectories which spend a fraction of less than $1 - k\varepsilon$ of their time in equilibrium. Let further G be the 'good' set of points with a time average $\hat{\Gamma}_{Eq}(x)$ of at least $1 - k\varepsilon$. All points in this set determine trajectories that spend a fraction of at least $1 - k\varepsilon$ of their time in equilibrium. To be precise,

$$B := \{x \in \Gamma | \hat{\Gamma}_{Eq}(x) < 1 - k\varepsilon\}, \qquad G := \{x \in \Gamma | \hat{\Gamma}_{Eq}(x) \geq 1 - k\varepsilon\}. \tag{10}$$

While, for a realistic physical system, ergodicity is hard to prove – if it can be proven at all –, essential ergodicity is not. In fact, it follows almost directly from the stationarity of the measure and the typicality of the equilibrium state. To be precise, with respect to the two sets B and G the following can been shown. For all $\varepsilon, k \in \mathbb{R}$ with $0 < \varepsilon << 1$ and $1 \leq k \leq 1/\varepsilon$:

$$\mu(B) < 1/k, \qquad \mu(G) > 1 - 1/k. \tag{11}$$

The proof can be found in [35]. An essential ingredient entering the proof is the pointwise existence and integrability of the time average (cf. [10]). Hence, in the case of non-ergodic systems, the time average of equilibrium need not attain a fix value on (almost all of) Γ—in fact, it may have different values on different trajectories –, but still it *exists* (pointwise almost everywhere) and this suffices to estimate the size of the set of trajectories with a time average smaller (or larger) than a particular value.

To grasp the full meaning of Eq. 11, consider a physically relevant value of k. Recall that, for a medium-sized macroscopic object, ε is tiny: $\varepsilon \approx 10^{-N}$ with $N \approx 10^{24}$. In that case, one can choose k within the given bounds ($1 \leq k \leq 1/\varepsilon$) large enough for $\mu(B)$ to be close to zero and $\mu(G)$ to be close to one. Consider, for example,

$$k = 1/\sqrt{\varepsilon}. \tag{12}$$

In that case, we distinguish between the 'good' set G of trajectories which spend at least $1 - \sqrt{\varepsilon}$ of their time in equilibrium and the 'bad' set B of trajectories which spend less than $1 - \sqrt{\varepsilon}$ of their time in equilibrium. And we obtain:

$$\mu(B) < \sqrt{\varepsilon}, \qquad \mu(G) > 1 - \sqrt{\varepsilon}. \tag{13}$$

Given the value of ε from above, $\varepsilon \approx 10^{-10^{24}}$, it follows that $\sqrt{\varepsilon} \approx 10^{-10^{23}}$. Consequently, the equilibrium region is of measure $\mu(\Gamma_{Eq}) \approx 1 - 10^{-10^{24}}$ and the measures of the sets B and G are

$$\mu(B) < 10^{-10^{23}}, \qquad \mu(G) > 1 - 10^{-10^{23}}. \tag{14}$$

Note that B is now the set of trajectories which spend less than $1 - 10^{-10^{23}}$ and G the set of trajectories which spend at least $1 - 10^{-10^{23}}$ (!) of their time in equilibrium. We thus find that trajectories which spend almost all of their time in equilibrium are typical whereas trajectories which spend less than almost all of their time in equilibrium are atypical!

The converse statement has be proven as well ([36, 37]; see [35] for a different proof). It says that if there exists a region $\Gamma_{Eq'} \subset \Gamma$ in which by far most trajectories spend by far most of their time, then this region has very large phase space measure. To be precise, if there exists a region G' with $\mu(G') = 1 - \delta$ such that $\forall x \in G'$: $\hat{\Gamma}_{Eq'}(x) \geq 1 - \varepsilon'$, then the following holds:

$$\mu(\Gamma_{Eq'}) \geq (1 - \varepsilon')(1 - \delta). \tag{15}$$

Here we are again interested in those cases where δ and ε' are very small, $0 < \delta << 1$ and $0 < \varepsilon' << 1$ (while the result holds for other values of δ and ε' as well).

This converse result tells us that, if there exists a state in which a typical trajectory spends by far most of its time, then this state is of overwhelming phase space measure.

Why is this converse statement interesting? It doesn't start from Boltzmann's notion of equilibrium. Instead, it starts from a thermodynamic or thermodynamic-like notion of equilibrium.

According to a standard thermodynamics textbook (like, e.g., [38] or [39]), a thermodynamic equilibrium is a state in which a system, once it is in that state, stays for all times. In what follows, we give a definition which relaxes that standard definition a little bit in that it allows for rare fluctuations out of equilibrium and for some atypical trajectories (all $x \notin G'$) that don't behave thermodynamic-like.[6]

Definition 2 (*Thermodynamic equilibrium*) Let $(\Gamma, \mathcal{B}(\Gamma), T, \mu)$ be a dynamical system. Let Γ be partitioned into finitely many disjoint, measurable subsets $\Gamma_{M_i} (i = 1, ..., n)$ by some (set of) physical macrovariable(s) M_i, i.e., $\Gamma = \bigcup_{i=1}^{n} \Gamma_{M_i}$. Let $G' \subset \Gamma$ with $\mu(G') = 1 - \delta$ and $0 < \delta << 1$. Let $0 < \varepsilon' << 1$. A set $\Gamma_{Eq'} \in \{\Gamma_{M_1}, ..., \Gamma_{M_n}\}$ (connected to a macrostate $M_{Eq'}$) with time average

$$\hat{\Gamma}_{Eq'}(x) \geq 1 - \varepsilon' \tag{16}$$

for all $x \in G'$ is called a 'thermodynamic equilibrium'.

To summarize, we obtain that, for every dynamical system with a stationary measure and a state of overwhelming phase space measure, almost all trajectories spend almost all of their time in that state, and the other way round, given a state in which almost all trajectories spend almost all of their time, that state is of overwhelming phase space measure. *Hence, an equilibrium state in Boltzmann's sense is a thermodynamic equilibrium and the other way round!*[7]

The only two assumptions which enter the proofs in [35] are:

(a) that the measure is stationary (resp. the dynamics is measure-preserving), i.e., $\mu_t(A) = \mu(A)$ for all $A \in \mathcal{B}(\Gamma)$ and

(b) that there is a macrostate of overwhelming phase space measure, i.e., a Boltzmann equilibrium Γ_{Eq} with $\mu(\Gamma_{Eq}) = 1 - \varepsilon$, or, for the reverse direction, a) and

(c) that there is a state in which typical trajectories spend by far most of their time, i.e., a thermodynamic equilibrium $\Gamma_{Eq'}$ with $\hat{\Gamma}_{Eq'} \geq 1 - \varepsilon'$.

Ergodicity doesn't enter the proofs, nor do we get ergodicity out of it. However, we get something similar to ergodicity, what we call 'essential ergodicity'.

---

[6] References [40, 41] would call this a 'thermodynamic-like equilibrium' to draw the distinction between this notion and the standard textbook definition.

[7] Based on the apparently missing connection between the time and the phase space average of equilibrium, Frigg and Werndl assert that Boltzmann's account of thermodynamic behaviour, which has later become known as the 'typicality account', is simply 'mysterious' [42, p. 918]. In follow-up papers (cf. [36, 37]) they even claim that the typicality account doesn't relate to thermodynamics *at all* because it doesn't draw the connection between Boltzmann's definition of equilibrium (in terms of the phase space average) and the thermodynamic definition of equilibrium (in terms of the time average). Here essential ergodicity counters the critique and closes the explanatory gap as it connects the time and phase space averages of the equilibrium state in a mathematically precise way.

3.3 Essential Ergodicity

While, for an ergodic system, the time and phase space averages *exactly coincide for all but a measure-zero set of solutions*, for an essentially ergodic system, the time and phase space averages *approximately coincide on typical solutions*. To be precise, the following definition applies.

Definition 3 (*Essential ergodicity*) Let $(\Gamma, \mathcal{B}(\Gamma), T, \mu)$ be a dynamical system. Let Γ be partitioned into finitely many disjoint, measurable subsets $\Gamma_{M_i} (i = 1, ..., n)$ by some (set of) physical macrovariable(s) M_i, i.e., $\Gamma = \bigcup_{i=1}^{n} \Gamma_{M_i}$. Let $0 < \varepsilon << 1$. A system is called 'essentially ergodic' if and only if

$$|\hat{\Gamma}_{M_i}(x) - \mu(\Gamma_{M_i})| \leq \varepsilon \tag{17}$$

$\forall i = 1, ..., n$ and $\forall x \in G$ with $\mu(G) \geq 1 - \delta$, $0 < \delta << 1$.

For a measure-preserving system with an equilibrium state (a Boltzmann or thermodynamic equilibrium), the Eq. 17 follows in a straightforward way from the two definitions of equilibrium given in Eqs. 8 and 16 and the corresponding results on the time and phase space average, Eqs. 14 and 15, respectively.[8]

Theorem 1 (FAPP ergodic hypothesis) *Let* $(\Gamma, \mathcal{B}(\Gamma), T, \mu)$ *be a measure-preserving dynamical system. Let there be an equilibrium state* M_{Eq} *(a Boltzmann or thermodynamic equilibrium) with corresponding equilibrium region* $\Gamma_{Eq} \subset \Gamma$.

Then the system is essentially ergodic. In particular, there exists an $\varepsilon \in \mathbb{R}$ *with* $0 < \varepsilon << 1$ *such that*

$$|\hat{\Gamma}_{Eq}(x) - \mu(\Gamma_{Eq})| \leq \varepsilon \tag{18}$$

$\forall x \in G$ *with* $\mu(G) \geq 1 - \delta$, $0 < \delta << 1$.

Proof We only prove Eq. 18. From that, the Eq. 17 follow directly. Let $0 < \delta', \varepsilon', \varepsilon'' << 1$. For the first direction of proof, consider a thermodynamic equilibrium, i.e., $\hat{\Gamma}_{Eq}(x) \geq 1 - \varepsilon'$ for all $x \in G'$ with $\mu(G') = 1 - \delta'$. It follows from Eq. 15 that $\mu(\Gamma_{Eq}) \geq (1 - \varepsilon')(1 - \delta')$ and, hence,

$$|\hat{\Gamma}_{Eq}(x) - \mu(\Gamma_{Eq})| \leq \varepsilon' + \delta' - \varepsilon'\delta'.$$

Now set $G = G'$, $\delta = \delta'$ and $\varepsilon = \varepsilon' + \delta' - \varepsilon'\delta'$.

[8] Note that the hard part of the proof of essential ergodicity is the proof of Eq. 14, which is the estimate on the time average of the Boltzmann equilibrium. This proof can be found in [35].

For the other direction, consider a Boltzmann equilibrium, i.e., $\mu(\Gamma_{Eq}) = 1 - \varepsilon''$. It follows from Eq. 11 that $\mu(G'') > 1 - \sqrt{\varepsilon''}$ with $G'' = \{x \in \Gamma | \hat{\Gamma}_{Eq}(x) \geq 1 - \sqrt{\varepsilon''}\}$. Hence, for all $x \in G''$,

$$|\hat{\Gamma}_{Eq}(x) - \mu(\Gamma_{Eq})| \leq \varepsilon''.$$

Now set $G = G''$, $\delta = \sqrt{\varepsilon''}$ and $\varepsilon = \varepsilon''$.

3.4 Scope and Limits of (Essential) Ergodicity

Although the notion of essential ergodicity is weaker than the notion of ergodicity, it predicts qualitatively the same long-time behaviour. In particular, it tells us that a typical trajectory spends by far most of its time in equilibrium, where equilibrium is defined in Boltzmann's way in terms of the phase space average, and it makes this notion of 'by far most' mathematically precise.[9] This justifies, in a rigorous way, Boltzmann's assumption of ergodicity as an idealization or FAPP truth in analyzing the system's long-time behaviour (as done, e.g., in his estimate of the fluctuation rate [15]). In other words, based on Boltzmann's account, the ergodic hypothesis is well-justified. It is a good working hypothesis for those time scales on which it begins to matter that trajectories wind around all of phase space.

Let us, at this point, use the above result on essential ergodicity to estimate the rate of fluctuations out of equilibrium. Recall that, according to Eq. 14, typical trajectories spend at least $1 - 10^{-10^{23}}$ of their time in equilibrium, when equilibrium is of measure $\mu(\Gamma_{Eq}) = 1 - 10^{-10^{24}}$ (which is a reasonable value for a medium-sized object). In other words, they spend a fraction of less than $10^{-10^{23}}$ of their time out of equilibrium, that is, in a fluctuation. If we assume that fluctuations happen randomly, in accordance with a trajectory wandering around phase space erratically, we obtain the following estimate for typical trajectories: a fluctuation of 1 second occurs about every $10^{10^{23}}$ seconds. But this means that a typical medium-sized system spends trillions of years in equilibrium as compared to one second in non-equilibrium, a time larger than the age of the universe![10]

So far we argued that essential ergodicity substantiates Boltzmanns assertions about the long-time behaviour of macroscopic systems. What about the short-time behaviour? In physics and philosophy, several attempts have been made to use ergodicity in some way or the other to explain a system's evolution from non-equilibrium to equilibrium (see [43] or [44, 45]; for earlier attempts as well as a thorough critique, see [9] and the references therein).

---

[9] Goldstein makes a similar point when he asserts that, even without ergodicity, the value of any thermodynamic variable is constant 'to all intents and purposes' [34, p. 46].

[10] This agrees with the time estimate Boltzmann presents in his letter to Zermelo [15, p. 577].

In this paper, I argue that ergodicity—just like epsilon-ergodicity, essential ergodicity, or any other notion involving an infinite-time limit—*does not* and *cannot* tell us anything about the approach to equilibrium, which is a behaviour within *short times*. This is simply due to the fact that the notion of ergodicity (or any notion akin to that) involves an *infinite-time limit*. Because of that limit, ergodicity can, at best, tell us something about the system's *long-time* behaviour where 'long-time' refers to time scales comparable to the recurrence times, where it begins to matter that the system's trajectory winds around all of phase space. For those short time scales on which the system evolves from non-equilibrium to equilibrium, ergodicity (or any notion akin to that) doesn't play any role. In fact, for a realistic gas, the equilibration time scale (i.e. the time scale of a system's approach to equilibrium) is fractions of a second as compared to trillions of years for the recurrence time!

Boltzmann's explanation of the irreversible approach to equilibrium is a genuine typicality result (see the discussion and references at the end of Sect. 3.1)—ergodicity doesn't add to nor take anything from that.

At this point, a quote of the mathematician Schwartz fits well.[11] Schwartz writes with respect to Birkhoff's ergodic theorem and the widely-spread conception that ergodicity might help to explain thermodynamic behaviour [8, pp. 23–24]:

> The intellectual attractiveness of a mathematical argument, as well as the considerable mental labor involved in following it, makes mathematics a powerful tool of intellectual prestidigitation – a glittering deception in which some are entrapped, and some, alas, entrappers. Thus, for instance, the delicious ingenuity of the Birkhoff ergodic theorem has created the general impression that it must play a central role in the foundations of statistical mechanics. [...] The Birkhoff theorem in fact does us the service of establishing its own inability to be more than a questionably relevant superstructure upon [the] hypothesis [of typicality].

4 Conclusion

Based on typicality and stationarity as the two basic concepts of Boltzmann's approach, it follows that ergodicity, as an idealization, or essential ergodicity, in the strict sense, is a consequence rather than an assumption of Boltzmann's account.

I believe that Boltzmann was aware of this fact. In my opinion, he simply didn't highlight the precise mathematical connection between the concepts of typicality, stationarity, and essential ergodicity because it was absolutely clear to him that, given a state of overwhelming phase space volume and a stationary measure, by far most trajectories would stay in that state by far most of their time—just like by far most trajectories starting from non-equilibrium would move into equilibrium very quickly. He didn't need a mathematical theorem to make this more precise.

[11] This quote was one of the first quotes (and essays) that were given to me by Detlef Dürr, to whom this memorial volume is dedicated. It is the style of writing that Detlef liked and that he himself employed on similar occasions.

Let me now end this paper with a variation of the both picturesque and paradigmatic example of Tim Maudlin, about typicality incidents occurring in the Sahara desert.[12] In what follows, I will adapt this example to the case of essential ergodicity.

A person wandering through the Sahara is typically surrounded by sand by far most of her time. In other words, she is typically hardly ever in an oasis. This fact is independent of the exact form of her 'wandering about', if she changes direction often, or not, if she moves fast, or not, and so on. Even if she doesn't move at all, she is typically surrounded by sand (in that case, for all times). In other words, independent of the dynamics, the long-time average of 'being surrounded by sand' is close to one on typical trajectories. This follows solely from the fact that all oases together constitute a vanishing small part of the Sahara desert and remain to do so throughout all times.

References

1. L. Boltzmann, Einige allgemeine Sätze über Wärmegleichgewicht. Wiener Berichte **63**(2), 679 (1871)
2. J.C. Maxwell, On Boltzmann's theorem on the average distribution of energy in a system of material points 1878. Trans. Camb. Philos. Soc. **12**, 547 (1879)
3. P. Ehrenfest, T. Ehrenfest-Afanassjewa, Begriffliche Grundlagen der statistischen Auffassung in der Mechanik, in *Encyklopädie der mathematischen Wissenschaften Mit Einschluss ihrer Anwendungen*, vol. 4. (Teubner, Leipzig, 1911), pp. 4–90
4. S.G. Brush, Foundations of statistical mechanics 1845–1915. Arch. Hist. Exact Sci. **4**(3), 145–83 (1967)
5. J. Earman, M. Rédei, Why ergodic theory does not explain the success of equilibrium statistical mechanics. Br. J. Philos. Sci. **47**(1), 63–78 (1996)
6. S. Smale, On the problem of reviving the ergodic hypothesis of Boltzmann and Birkhoff, in *The Collected Papers of Stephen Smale*, vol. 2, ed. by F. Cucker, R. Wong. (Singapore University Press, Singapore, 1979), pp. 823–30
7. L. Sklar, Statistical explanation and ergodic theory. Philos. Sci. **40**(2), 194–212 (1973)
8. J. Schwartz, The pernicious influence of mathematics on science, in *Discrete Thoughts: Essays on Mathematics, Science and Philosophy*. ed. by M. Kac, G.C. Rota, J. Schwartz. (Birkhäuser, Boston, 1992), pp. 19–25
9. J. Bricmont, Science of chaos or chaos in science? Phys. Mag. **17**, 159–208 (1995)
10. G.D. Birkhoff, Proof of the ergodic theorem. Proc. Natl. Acad. Sci. U. S. A. **17**, 656–60 (1931)
11. J. von Neumann, Proof of the quasi-ergodic hypothesis. Proc. Natl. Acad. Sci. **18**(1), 70–82 (1932)
12. A.I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover Publications, Richmond, 1949)
13. K. Petersen, *Ergodic Theory*. Cambridge Studies in Advanced Mathematics 2. (Cambridge University Press, Cambridge, 1983)

[12] Known to the author from private conversation. The original version is about a person's approach from non-equilibrium (here: an oasis) to equilibrium (here: the remainder of the desert), where it is the atypical initial condition, the special fact of 'being in an oasis' *in the very beginning*, which is in need of explanation. The fact that a person, walking around in an unspecific and maybe even random way, walks out of the oasis into the desert is merely typical (we call it *typical within atypicality*; see [46] for this phrasing). According to [34], it is the explanation of the atypical initial condition which constitutes the hard part of any explanation of thermodynamic irreversibility.

14. L. Boltzmann, Vorlesungen über Gastheorie. Verlag v. J. A. Barth (Nabu Public Domain Reprint), Leipzig (1896)
15. L. Boltzmann, Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo. Wiedemann's Annalen **57**, 773–84 (1896)
16. L. Boltzmann, Zu Hrn. Zermelos Abhandlung "über die mechanische Erklärung irreversibler Vorgänge'. Wiedemann's Ann. **60**, 392–98 (1897)
17. G.W. Gibbons, S.W. Hawking, J.M. Stuart, A natural measure on the set of all universes. Nucl. Phys. B **281**, 736–51 (1987)
18. D. Dürr, A. Frümel, M. Kolb, Einführung in die Wahrscheinlichkeitstheorie als Theorie der Typizität. (Springer, Berlin, 2017)
19. S. Goldstein, R. Tumulka, N. Zanghì, Is the hypothesis about a low entropy initial state of the universe necessary for explaining the arrow of time? Phys. Rev. D **94**(023520) (2016)
20. D. Lazarovici, P. Reichert, Arrow(s) of time without a past hypothesis, in *Statistical Mechanics and Scientific Explanation*. Determinism, Indeterminism and Laws of Nature. (World Scientific, Singapore, 2020), pp. 343–86
21. A. Rosenthal, Beweis der Unmüglichkeit ergodischer Gassysteme. Annalen der Physik **347**(14), 796–806 (1913)
22. M. Plancherel, Beweis der Unmöglichkeit ergodischer mechanischer Systeme. Annalen der Physik **347**(15), 1061–63 (1913)
23. Y.G. Sinai, Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards. Russ. Math. Surv. **25**(2), 137–89 (1970)
24. N. Simányi, Further developments of Sinai's ideas: The Boltzmann-Sinai hypothesis. To appear in the Abel Prize volume dedicated to Ya. G. Sinai. Arxiv version (2015). https://doi.org/10.48550/arXiv.1512.08272
25. J.S. Schiffrin, R.M. Wald, Measure and probability in cosmology. Phys. Rev. D **86**(023521) (2012)
26. S. Goldstein, Typicality and notions of probability in physics, in *Probability in Physics*, ed. by Y. Ben-Menahem, M. Hemmo. (Springer, Berlin, Heidelberg, 2012), pp. 59–71
27. D. Lazarovici, P. Reichert, Typicality, irreversibility and the status of macroscopic laws. Erkenntnis **80**(4), 689–716 (2015)
28. I. Wilhelm, Typical: a theory of typicality and typicality explanations. Br. J. Philos. Sci. **73**(2) (2019). https://doi.org/10.1093/bjps/axz016
29. O.E. Lanford, Entropy and equilibrium states in classical statistical mechanics, in *Statistical Mechanics and Mathematical Problems*, ed. by A. Lenard. Lecture Notes in Physics, vol. 20 (1973), pp. 1–113
30. R. Penrose, *The Emperor's New Mind* (Oxford University Press, Oxford, 1989)
31. R. Penrose, *The Road to Reality* (Vintage, London, 2004)
32. J. Lebowitz, Macroscopic laws, microscopic dynamics, time's arrow and boltzmann's entropy. Phys. A **194**, 1–24 (1993)
33. J. Lebowitz, Boltzmann's entropy and time's arrow. Phys. Today, 32–38 (1993)
34. S. Goldstein, Boltzmann's approach to statistical mechanics, in *Chance in Physics. Foundations and Perspectives*, ed. by J. Bricmont, D. Dürr, M.C. Galavotti, G. Ghirardi, F. Petruccione, N. Zanghì. (Springer, Berlin, 2001), pp. 39–54
35. P. Reichert, Essentially ergodic behaviour. Br. J. Philos. Sci. **74**(1) (2020). https://doi.org/10.1093/bjps/axaa007
36. R. Frigg, C. Werndl, Reconceptualising equilibrium in Boltzmannian statistical mechanics and characterising its existence. Stud. Hist. Philos. Mod. Phys. **49**(1), 19–31 (2015)
37. R. Frigg, C. Werndl, Rethinking Boltzmannian equilibrium. Philos. Sci. **82**(5), 1224–35 (2015)
38. H.B. Callen, *Thermodynamics and an Introduction to Thermostatistics* (Wiley, New York, 1960)
39. H. Reiss, *Methods of Thermodynamics* (Dover, Minneaola, New York, 1996)
40. D. Lavis, Boltzmann and Gibbs: an attempted reconciliation. Stud. Hist. Philos. Mod. Phys. **36**, 145–73 (2005)
41. D. Lavis, Boltzmann, Gibbs, and the concept of equilibrium. Philos. Sci. **75**(5), 682–96 (2008)

42. R. Frigg, C. Werndl, Demystifying typicality. Philos. Sci. **79**, 917–29 (2012)
43. P. Vranas, Epsilon-ergodicity and the success of equilibrium statistical mechanics. Philos. Sci. **65**(4), 688–708 (1998)
44. R. Frigg, C. Werndl, Explaining thermodynamic-like behavior in terms of epsilon- ergodicity. Philos. Sci. **78**, 628–52 (2011)
45. R. Frigg, C. Werndl, A new approach to the approach to equilibrium, in *Probability in Physics*, ed. by Y. Ben-Menahem, M. Hemmo. (Springer, Berlin, 2012), pp. 99–113
46. D. Dürr, S. Teufel, *Bohmian Mechanics* The Physics and Mathematics of Quantum Theory. (Springer, Berlin, 2009)

A Simple Mathematical Framework for Learning and Teaching Probability Theory

Günter Hinrichs

Abstract Random variables with arbitrary distributions as well as large classes of stochastic processes can be constructed on $(0, 1)^d$ with uniform distribution. Treating topics such as the law of large numbers or the central limit theorem entirely within this probability space, one can avoid to expose students to an unsound amount of general measure theory without giving up any mathematical rigor. In the light of an understanding of probability theory as a theory of typical physical behaviour, such as taught by Detlef Dürr, I argue that such a procedure entails no loss of generality. Observing the limitations of this procedure, one is led to general measure theory as a tool to overcome them.

1 Introduction

Among the usual topics taught to undergraduate students of mathematics, probability theory is certainly one with particularly striking differences from teacher to teacher. There seems to be little agreement about the appropriate degree of abstraction and generality. Formally, the decision someone makes here will usually manifest itself in the role and amount of measure theory.

On the one hand, already a quick look on the historical development (e.g. [2] contains a nice survey) shows that central topics for introductory courses like explicit calculations with standard distributions, law of large numbers, central limit theorem and statistical inference have developed even beyond a stage accessible to beginners without the measure-theoretical framework that emerged in the late 19th and 20th century. (Hilbert formulated in 1906 his 6th problem that probability theory is now so well understood that it should be axiomatized...) Therefore it should clearly be possible also nowadays to understand and teach them without the latter. On the other hand, one might wish that introductory courses also prepare for dealing with more advanced "20th and 21st century topics" with a more intrinsic role of measure theory.

G. Hinrichs (✉)
Universität Augsburg, Augsburg, Germany
e-mail: guenter.hinrichs@math.uni-augsburg.de

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_21

If one wants to include measure theory for the latter reason, one should find a compromise concerning its amount, otherwise time might be over before one comes to anything else... A common strategy that one can guess behind many such compromises is something like the following: "Imagine an introduction that starts with the Kolmogorov (or other) axioms and then builds everything up step by step a formally perfect way. Accept this as the plan according to which one would have to proceed if one wanted the best possible theoretical understanding. Make the necessary (very large) omissions and changes in order to get by with the time and to get room for pedagogical parts, applications to real-world problems etc. The result can be used as a concept for a good lecture."

Such an approach is likely to keep parts of measure theory as a "black box" and to treat it in an ambiguous way, conveying that it is on the one hand very essential to use it, but on the other hand too unimportant and boring to explain it properly. As part of a possible way out of this unsatisfactory situation, I would like to recall that quite a large part of probability theory can be formulated on the particularly simple probability space $(0, 1)^d$ with uniform distribution. I will explain how this approach relates to the view of probability theory as a theory of typical physical happenings. Moreover, I will emphasize that, up to some point, this approach requires no advanced measure theory on this space, but only basic analysis skills, and serves as propaedeutics for what follows beyond this point.

2 Typical Physical Behaviour

I am going to build on the following view of probability theory, which is to a very large extent influenced by Detlef Dürr's view as elaborated e.g. in [1]:

The empirical law of large numbers tells us what it means that a coin shows head with probability $\frac{1}{2}$. This law is also sufficient to deduce the usual rules to calculate with that sort of probabilities, how to incorporate independence into a formula etc. One will presumably not get a deduction that counts as mathematically rigorous nowadays—who cares?[1]—, but a deduction in the sense that one understands things that are further away from everyday knowledge and not so clear by logical reasoning on the basis of things that are clearer and closer to everyday knowledge.

However, one can ask for different things apart from rules for practical calculations, e.g. whether this probability $\frac{1}{2}$, somehow a physical property of the coin, and the resulting regular behaviour can itself be reduced to something more elementary. Then, of course, the everyday understanding of probabilities via the empirical law of large numbers is no longer a suitable starting point, and a similarly inappropriate one would be axioms for probabilities which are informally motivated by this law. A starting point that leads to a noncircular argumentation is to count in how many among all possible results of a long coin tossing series the relative frequency of

[1] An unforgettable quote by Detlef Dürr—he used it frequently, but not in this context.

heads is as close to $\frac{1}{2}$ as one usually observes it in practice. One finds out that this is so in "almost all" results and has thus reduced the regular behaviour of the relative frequencies to its simpler and thus maybe more fundamental property of being "typical".

In order that simple counting is appropriate, all outcomes have to be in some sense on equal footing. At first sight this seems to be a restrictive requirement, but at second sight, one can recognize many experiments with nonequal outcomes as coarse-grainings of happenings that, on a finer scale, have equal outcomes. Ultimately, this refers to all experiments that can be reasonably well understood within classical mechanics, thus all such experiments admit a typicality analysis: The set of equal outcomes is the microscopic physical phase space—some subset of $\mathbb{R}^{6n}$ which contains the possible initial positions and velocities of the n involved microscopic particles—and the volume or Lebesgue measure, by the Liouville theorem about its stationarity under the phase flow, qualifies as "continuous counting" of possible cases. Many different microscopic states lead to the same macroscopic state (e.g. a coin showing head or tails) observed in the experiment, i.e. there is a mapping ("random variable") from phase space to a much lower-dimensional space which attributes to each microscopic state how it looks macroscopically.

Of course, one can give no explicit complete microscopic description of any real experiment starting from initial conditions and phases flow, but at least one "knows" from this picture that, e.g., an unfair coin should be described as a random variable with values 0 and 1, say, defined on a subset of some $\mathbb{R}^d$ equipped with the uniform distribution. One can also "derive" from the physical picture that the chaotic dependence of the coin tossing result on the microscopic initial conditions leads to independence of results of repeated experiments, i.e. such results can be described as a sequence of random variables on the same space which have identical distributions and satisfy the equations defining independence. Now one can do the following: Try to model every "random" experiment (single or repeated) in a mathematically simple and convenient way as a random variable X or a family of (independent) random variables $X_1, \ldots$ defined on some subset Ω of some $\mathbb{R}^d$ as sample space equipped with the uniform distribution. Then the details of the random variables have no physical meaning, but one has incorporated the principal situation. Therefore, if one mathematically proves a law of large numbers on this space which says that for a large share of elements ω of the sample space the relative frequency with which $X_1(\omega), \ldots$ are one stays close to its probability, then one can still, to some extent, attribute the meaning to this proof that one has tried to describe the unfair coin in terms of its equal microscopic states and, by counting, has found out, that the law of large numbers is a typical phenomenon.

The situation is in apparent contrast to the usual mathematical approach, in which one defines the random variables on an arbitrary abstract "probability space", not necessarily with uniform distribution, and can only prove that the regular relative frequencies appear "with high probability"—with respect to this abstract "probability space", whatever that means...

3 Working on $(0, 1)^d$ with Uniform Distribution

What is a simple way to model random experiments as random variables X_i on a subset Ω of $\mathbb{R}^d$ with uniform distribution? Building on the previous considerations rather than prior knowledge about stochastics, one might come up with the following solution: For an experiment with one number as outcome, take a one-dimensional sample space, e.g. $\Omega = (0, 1)$. If the details of the mapping X from "microscopic" $\omega \in \Omega$ to macroscopic values are not physical anyway and it should be simple, then order its values, i.e. admit only monotonically increasing $X : \Omega \to \mathbb{R}$. By varying the slope or skipping values, one should be able to realize all conceivable ratios of different values (i.e. one could take the set of all image measures of monotonically increasing X as a definition of all possible probability distributions on $\mathbb{R}$). Similarly, for an arbitrary experiment with $d \in \mathbb{N}$ steps one might suggest $(0, 1)^d$ (with $\mathbb{P}$ denoting the d-dimensional volume) and random variables $(X_1, \ldots, X_d) : \Omega \to \mathbb{R}^d$ with "ordered values", which means in this case that each X_i should increase monotonically in every coordinate ω_j. d independent repetitions of the "same" experiment, one of which would be described by $X : (0, 1) \to \mathbb{R}$, can be conveniently integrated into this scheme by setting $X_i(\omega) := X(\omega_i)$.

As it is well-known, this scheme comprises everything that one usually defines as "probability distribution on the Borel sigma algebra of $\mathbb{R}$" because a random variable with any such distribution can be realized as the quantile function $X := F^{-1}$ of its cumulative distribution function. A little more thought shows that also the d-dimensional scheme covers all probability distributions (in the usual sense) on the Borel sigma algebra of $\mathbb{R}^d$: Let $\mathbb{Q}$ be any such distribution and $P_1, \ldots, P_d : \mathbb{R}^d \to \mathbb{R}$ the coordinate projections. In order to construct random variables $(X_1, \ldots, X_d)$ on $(0, 1)^d$ which have the distribution $\mathbb{Q}$ and are monotonically increasing in every coordinate, start with $X_1 := F_{P_1}^{-1}$. According to general measure theory, $\mathbb{Q}(P_i \in \cdot \mid P_1 = p_1, \ldots, P_{i-1} = p_{i-1})$ has a regular version and therefore in particular a conditional distribution function $\mathbb{Q}(P_i \leq p_i \mid P_1 = p_1, \ldots, P_{i-1} = p_{i-1}) = F_i(p_1, \ldots, p_{i-1}, p_i)$. Set recursively $X_i(\omega) := F_i^{-1}(X_1(\omega), \ldots, X_{i-1}(\omega), \omega_i)$ (where the quantile function is formed w.r.t. ω_i, the other variables being fixed), this will do the job.

So, as long as one deals only with distributions on $\mathbb{R}^d$, one loses no generality by using the suggested framework. It has the big advantage that events e.g. of the form $\{X_i \in (a, b]\}$ are always Jordan measurable, i.e. their volume is the "laymen volume" computed by approximating from inside and outside with finite unions of cubes, and expectation value and moments of X_i, if existent, can be computed via the (possibly improper) Riemann integral. The same holds true for $X_1 + \ldots + X_N$ because this random variable is also monotone in every coordinate. Therefore, questions which are "not too complicated" can be treated without advanced measure and integration theory, even without the concept of sigma-additivity. This comprises in particular central results for sums of independent random variables like the weak law of large numbers and the central limit theorem.

(The strong law of large numbers in the form $\mathbb{P}\left(\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} X_i = \mathbb{E}X\right) = 1$ can not be treated in this way. However, in the usual framework, it is equivalent to the statement $\forall \varepsilon > 0 : \lim_{N\to\infty} \sup_{M:M\geq N} \mathbb{P}\left(\max_{j\in\{N,\ldots,M\}} \left|\frac{1}{j}\sum_{i=1}^{j} X_i - \mathbb{E}X\right| > \varepsilon\right) = 0$. The latter works because the event inside $\mathbb{P}$ is a subset of $(0,1)^M$ constructed by finitely many operations. I prefer this formulation anyway because, in contrast to the usual one which emphasizes that the strong law uses a mathematically stronger version of convergence, this one gives a more practical distinction between weak and strong law, namely: According to the weak law, for sufficiently large N it is typical that $\frac{1}{N}\sum_{i=1}^{N}$ and $\mathbb{E}X$ are close together. According to the strong law, it is typical that they are not only closed together for one fixed large N, but that they also remain close together if N is increased further arbitrarily.)

To the experienced reader it will be more or less clear that things can be proved as stated. This text is intended as a sketch for such readers. If someone wants to see all details worked out, I ask him to wait for [3] to appear...

4 Proceeding to a More General Framework

The described framework allows to design introductory courses in which all central results are proved completely, transparently, elementarily and without "black boxes". On the introductory level, one will hardly encounter serious limitations of this approach. Of course, one might wish to consider $f(X_i)$ with nonmonotonic f, but for practical purposes, piecewise monotonic f are sufficient and it is clear that then, all arguments only need slight modifications. Products $X_1 \cdots X_N$ of monotonic random variables are another example of a construct that might be of interest and, in general, is only piecewise monotonic.

A topic that advanced learners will encounter and that brings more serious complications are time-continuous stochastic processes (e.g. Wiener or Poisson process). Here infinite families of random variables are intrinsically important (in contrast to the situation in limit theorems for N random variables with $N \to \infty$ like the law of large numbers). As long as the paths are at least continuous from one side, it suffices to consider rational times. How can a process $(Y_t)_{t\in\mathbb{Q}}$ with prescribed joint distribution of the Y_t be constructed on $\Omega = (0,1)^d$ in order to keep touch to physics and typicality analysis? Suppose you have already succeded to construct a sequence $(Z_n)_{n\in\mathbb{N}}$ of independent random variables with uniform distribution on $(0,1)$, then it goes as follows: Enumerate $\mathbb{Q}$ as $\{t_1, t_2, \ldots\}$. Construct iteratively $(X_{t_1}, \ldots, X_{t_d})$ on $(0,1)^d$ with the desired joint distribution as described in Sect. 3. Then set $Y_{t_d} := X_{t_d}(Z_1, \ldots, Z_d)$. (For the Wiener process, this strategy actually yields the Lévy construction if the enumeration is appropriate.)

So the only missing part to construct arbitrary such processes are the Z_n. They cannot be constructed as monotonic functions on $(0,1)$ order $(0,1)^d$. A different way to construct them on $(0,1)$ is: For $\omega \in (0,1)$, let $d_n(\omega)$ be the nth dyadic digit (after the comma). Then the d_n are independent and uniformly distributed on $\{0,1\}$ and

$Z(\omega) := \sum_n 2^{-n} d_n(\omega)$ is the identity and thus uniformly distributed on $(0, 1)$. The distribution of Z will not change if one replaces the d_n by an other i. i. d. sequence with the same distribution, e.g. a subsequence (d_{i_n}). Grouping the d_n into countably many subsequences and forming "Z's" out of each one, one arrives at the desired Z_n.

Now sets like $\{Z_n \in (a, b]\}$ or, even worse, $\{X_t \in (a, b]\}$ are quite complicated and need no more be accessible to the simple Jordan measure because of the infinite construction procedure. Still, by construction, the first set should, in some sense, be as simple as $\{Z \in (a, b]\} = (a, b]$. In other words: Now one has arrived at a point where one would like to have a more refined measure theory on $(0, 1)$ and a theory of integration which depends only on the "frequency", not on the "order" of function values. And since the subsets of $(0, 1)$ are so complicated and have no practical relevance, one would prefer that the results are, as far as possible, formulated directly in terms of the image space without permanent need to track things back to $(0, 1)$. In this way, one arrives at rather general probability spaces—but only as soon as one needs them and without depreciating by their introduction anything that one has done without the general theory...

5 Kommst du mit zum Bahnhof?

When I was one of Detlef's doctoral students in Munich, it was a particular pleasure to walk from time to time from the mathematical institute to the train station with him. We did not mainly talk about mathematics and physics then, and if we did, we used the atmosphere that was unlike in the institute to talk about different topics than there.

For my contribution to this memorial volume, I tried to choose a mathematical (physical? metaphysical? philosophical?) topic that I might have brought up and we might have started to discuss on the way to the train station if we had the opportunity to continue this tradition...

References

1. Detlef Dürr, Anne Frömel und Martin Kolb, *Einführung in die Wahrscheinlichkeitstheorie als Theorie der Typizität.* Springer Spektrum (2017)
2. Boris Wladimirowitsch Gnedenko, *Lehrbuch der Wahrscheinlichkeitstheorie* (Verlag Harry Deutsch, 1997)
3. G. Hinrichs, *Grundlagen der Wahrscheinlichkeitsrechnung.* Work in progress, to be published

Asymptotic Analysis of the Weakly Interacting Bose Gas: A Collection of Recent Results and Applications

Lea Boßmann, Nikolai Leopold, David Mitrouskas, and Sören Petrat

To Detlef Dürr,

a great teacher and wonderful person

Abstract We consider a gas of N bosons with interactions in the mean-field scaling regime. We review a recent proof of the asymptotic expansion of its spectrum and eigenstates and two applications of this result, namely the derivation of an Edgeworth expansion for fluctuations of one-body operators and the computation of the binding energy of an inhomogeneous Bose gas to any order. Finally, we collect related results for the dynamics of the weakly interacting Bose gas and for the regularized Nelson model.

L. Boßmann (✉)
Mathematisches Institut, Ludwig-Maximilians-Universität München, Theresienstr. 39, 80333 München, Germany
e-mail: bossmann@math.lmu.de

N. Leopold
Department of Mathematics and Computer Science, University of Basel, Spiegelgasse 1, 4051 Basel, Switzerland
e-mail: nikolai.leopold@unibas.ch

D. Mitrouskas
Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400 Klosterneuburg, Austria
e-mail: mitrouskas@ist.ac.at

S. Petrat
School of Science, Constructor University, Campus Ring 1, 28759 Bremen, Germany
e-mail: spetrat@constructor.university

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_22

1 Introduction

Bose gases have been studied from many different perspectives since the discovery of Bose–Einstein condensation (BEC), which, after the theoretical prediction in 1924 by Bose [8] and Einstein [17, 18], was first experimentally realized in 1995 by the groups of Cornell/Wieman [5] and Ketterle [16]. In a typical experiment, the bosons are initially caught in an external trap, where they are cooled down to a superposition of low-energy eigenstates; subsequently, they are released and their behavior is observed. If the number of particles in the gas is large, neither an analytical nor a numerical analysis of the system is feasible, which makes the use of appropriate approximations indispensable.

The resulting evolution equations are sometimes broadly called *effective equations*. The study of their emergence from a microscopic theory of interacting particles is a typical question in mathematical and statistical physics. In a different context, namely that of conductivity, and also of Brownian motion, this field is where Detlef started his career as a mathematical physicist. We therefore like to think that he would have enjoyed the kind of results we are presenting here, and we dedicate this article to him.

Over the last two decades in particular, there have been many contributions in the mathematical physics community devoted to a rigorous derivation of suitable effective equations for different models of BEC. In this review, we restrict ourselves to the weakly interacting Bose gas, also known as the mean-field or Hartree regime, which describes trapped bosons with weak and long-range interactions. The corresponding Hamiltonian for the N-body system is given by

$$H_N = \sum_{j=1}^{N} \left(-\Delta_j + V^{\mathrm{trap}}(x_j)\right) + \frac{1}{N-1} \sum_{1 \le i < j \le N} v(x_i - x_j)\,, \tag{1}$$

acting on the Hilbert space $L^2_{\mathrm{sym}}((\mathbb{R}^d)^N)$ of square integrable, permutation symmetric functions on $(\mathbb{R}^d)^N$. We assume the two-body interaction potential $v : \mathbb{R}^d \to \mathbb{R}$ to be bounded, symmetric and—for our spectral results—of positive type, i.e., to have a non-negative Fourier transform. The confining potential $V^{\mathrm{trap}} : \mathbb{R}^d \to \mathbb{R}$ is assumed to be measurable, locally bounded, non-negative, and such that $V^{\mathrm{trap}}(x)$ tends to infinity as $|x| \to \infty$. Instead of using an external potential in $\mathbb{R}^d$, one often restricts the particles to the d-dimensional unit torus $\mathbb{T}^d$, which usually simplifies the analysis since the resulting system is homogeneous.

The spectral and dynamical properties of the model (1) have been subject to extensive research; for more recent results, see, e.g., [27–29, 34–36, 38, 39, 41, 51], respectively. Let us also refer to [37] for a more general review of BEC.

In this article, we start in Sect. 2 by reviewing results related to the spectrum and eigenfunctions based on [14]. In Sect. 3, we review the Edgeworth expansion from [12] and the binding energy expansion from [10]. Finally, in Sect. 4, we review the dynamical results from [13, 21].

In the following Sects. 2 and 3, we consider the ground state Ψ_N of H_N and the ground state energy $\mathscr{E}_N$, i.e.,

$$\mathscr{E}_N = \inf \operatorname{spec}(H_N)\,, \qquad H_N \Psi_N = \mathscr{E}_N \Psi_N\,. \tag{2}$$

Under appropriate conditions on v and V^{trap}, it is well known that Ψ_N is unique and exhibits complete asymptotic BEC in the minimizer $\varphi \in L^2(\mathbb{R}^d)$ of the Hartree energy functional, which is given by

$$\mathcal{E}_{\mathrm{H}}[\phi] := \int_{\mathbb{R}^d} \left(|\nabla\phi(x)|^2 + V^{\mathrm{trap}}(x)|\phi(x)|^2\right) \mathrm{d}x + \tfrac{1}{2} \int_{\mathbb{R}^{2d}} v(x-y)|\phi(x)|^2|\phi(y)|^2 \,\mathrm{d}x\,\mathrm{d}y\,. \tag{3}$$

We denote its minimum under the constraint $\|\phi\| = 1$ by $e_{\mathrm{H}} := \mathcal{E}_{\mathrm{H}}[\varphi]$. Complete asymptotic BEC in the state φ means that Ψ_N is determined by φ in the sense of reduced densities, i.e.,

$$\lim_{N\to\infty} \operatorname{Tr} \left|\gamma_N^{(1)} - |\varphi\rangle\langle\varphi|\right| = 0\,, \tag{4}$$

where $\gamma_N^{(1)} := \operatorname{Tr}_{2\ldots N} |\Psi_N\rangle\langle\Psi_N|$ denotes the one-particle reduced density matrix of Ψ_N. Heuristically, this implies that $N - o(N)$ particles occupy the condensate state φ. Consequently, the leading order of $\mathscr{E}_N$ is given by the condensate energy $N e_{\mathrm{H}}$.

2 Asymptotic Expansion of the Ground State

2.1 Main Result

The first result we review in these notes is an expansion of the N-body ground state Ψ_N and of the ground state energy $\mathscr{E}_N$ in powers of $N^{-1/2}$, which is proven in [14].

Theorem 1 *Let $a \in \mathbb{N}_0$ and let N be sufficiently large. Then there exists a constant $C(a)$ such that*

$$\left\| \Psi_N - \sum_{\ell=0}^{a} N^{-\frac{\ell}{2}} \psi_{N,\ell} \right\|_{L^2((\mathbb{R}^d)^N)} \leq C(a) N^{-\frac{a+1}{2}} \tag{5}$$

and

$$\left| \mathscr{E}_N - N e_{\mathrm{H}} - \sum_{\ell=0}^{a} N^{-\ell} E_\ell \right| \;\leq\; C(a) N^{-(a+1)}\,. \tag{6}$$

The coefficients $\psi_{N,\ell} \in L^2_{\mathrm{sym}}((\mathbb{R}^d)^N)$ and $E_\ell \in \mathbb{R}$ are computed in [14] in full generality. As an example, $\psi_{N,1}$ and E_1 are given in (16) and (18).

To leading order ($a = 0$), this was proven in [27, 36, 38, 51]. The higher orders ($a > 0$) were rigorously derived in [14], and related results were obtained in [45–47].

The coefficients e_{H} and E_ℓ are independent of N. The N-body wave functions $\psi_{N,\ell}$ are of the structure $\psi_{N,\ell} = \sum_{k=0}^{N} \varphi^{(N-k)} \otimes_s \chi_\ell^{(k)}$, where $\chi_\ell^{(k)}$ are k-body wave functions which are independent of N. Hence, the N-dependence of $\psi_{N,\ell}$ is trivial, and the computational effort to obtain physical quantities such as expectation values with respect to the N-body state, does not scale with N.

The constants $C(a)$ grow rapidly in a, which means that (5) and (6) are *asymptotic* expansions (and not converging series): given any order a of the approximation, one can choose N sufficiently large that the estimates are meaningful.

Theorem 1 extends to the low-energy excitation spectrum of H_N and to a certain class of unbounded interaction potentials v, including the repulsive three-dimensional Coulomb potential (see [14] for the full statement). Moreover, it implies an asymptotic expansion of the corresponding one-body reduced density matrices [9].

2.2 Idea of Proof

The contributions to the ground state energy beyond the leading order are caused by particles which are excited from the condensate due to the interactions. To describe these excitations, one decomposes Ψ_N as

$$\Psi_N = \sum_{k=0}^{N} \varphi^{\otimes(N-k)} \otimes_s \chi^{(k)}\,, \qquad \chi^{(k)} \in \bigotimes_{\mathrm{sym}}^{k} \{\varphi\}^\perp\,, \qquad \boldsymbol{\chi} := \big(\chi^{(k)}\big)_{k=0}^{N} \in \mathcal{F}_{\perp\varphi}^{\leq N} \subset \mathcal{F}_{\perp\varphi} \tag{7}$$

with $\otimes_s$ the symmetric tensor product and where $\{\varphi\}^\perp$ denotes the orthogonal complement of φ in $L^2(\mathbb{R}^d)$ [36]. The excitations form a vector in the (truncated) excitation Fock space over $\{\varphi\}^\perp$, which is denoted by $\mathcal{F}_{\perp\varphi}$ (resp. $\mathcal{F}_{\perp\varphi}^{\leq N}$). The creation/annihilation operators a^*/a and the number operator $\mathcal{N}_{\perp\varphi}$ on this Fock space are defined in the usual way. The relation between Ψ_N and the corresponding excitation vector $\boldsymbol{\chi}$ is given by the unitary map

$$U_{N,\varphi} : L^2((\mathbb{R}^d)^N) \to \mathcal{F}_{\perp\varphi}^{\leq N}\,, \quad \Psi_N \mapsto U_{N,\varphi}\Psi_N = \boldsymbol{\chi}\,. \tag{8}$$

Conjugating H_N with $U_{N,\varphi}$ and subtracting the condensate energy Ne_{H} yields the operator

$$\mathbb{H} \;:=\; U_{N,\varphi}\,(H_N - Ne_{\mathrm{H}})\,U_{N,\varphi}^* \tag{9}$$

on $\mathcal{F}_{\perp\varphi}^{\leq N}$, whose ground state is denoted by $\boldsymbol{\chi}$. Hence, the ground state energy E of $\mathbb{H}$,

$$E = \langle \boldsymbol{\chi}, \mathbb{H}\boldsymbol{\chi}\rangle_{\mathcal{F}_{\perp\varphi}^{\leq N}} = \mathscr{E}_N - Ne_{\mathrm{H}}\,, \tag{10}$$

gives us precisely the corrections to the condensate energy Ne_{H} in (6). After extending $\mathbb{H}$ trivially to the full excitation Fock space $\mathcal{F}_{\perp\varphi}$, computing (9) as in [36, Proposition 4.2] yields an expansion of $\mathbb{H}$ in powers of $N^{-1/2}$,

$$\mathbb{H} \;=\; \mathbb{H}_0 + \sum_{j=1}^{a} N^{-\frac{j}{2}} \mathbb{H}_j + N^{-\frac{a+1}{2}} \mathbb{R}_a \tag{11}$$

for any $a \in \mathbb{N}_0$. The coefficients $\mathbb{H}_j$ and the remainders $\mathbb{R}_a$ in this expansion are unbounded operators on $\mathcal{F}_{\perp\varphi}$ which depend on v, V^{trap} and φ. The operators $\mathbb{H}_j$ are independent of N.

The leading order term $\mathbb{H}_0$ in (11) is the well-known Bogoliubov Hamiltonian, which is a very useful approximation of $\mathbb{H}$ because it is quadratic in the number of creation/annihilation operators. Under the given assumptions on v, it can therefore be diagonalized by a Bogoliubov transformation $\mathbb{U}_0$, in the sense that $\mathbb{U}_0 \mathbb{H}_0 \mathbb{U}_0^* = E_0 + \int \mathrm{d}x\, a_x^* D(x, y) a_y$ for some positive one-body operator D. The unique ground state of $\mathbb{H}_0$ is thus given by

$$\boldsymbol{\chi}_0 = \mathbb{U}_0^* |\Omega\rangle \,, \tag{12}$$

where $|\Omega\rangle$ is the vacuum state, and its ground state energy is E_0. It is well known [27, 36, 38, 51] that

$$\lim_{N\to\infty} E = \lim_{N\to\infty} (Ne_{\mathrm{H}} - \mathscr{E}_N) = E_0 \,, \qquad \lim_{N\to\infty} \|\boldsymbol{\chi} - \boldsymbol{\chi}_0\|_{\mathcal{F}_{\perp\varphi}} = 0 \,, \tag{13}$$

where we trivially extended $\boldsymbol{\chi}$ to a vector in $\mathcal{F}_{\perp\varphi}$. Consequently, E_0 gives the leading order ($a = 0$) correction to $\mathcal{E}_N - Ne_{\mathrm{H}}$ in (6); analogously, the leading order contribution in (5) is given by $\psi_{N,0} = U_{N,\varphi}^* \boldsymbol{\chi}_0 |_{\mathcal{F}_{\perp\varphi}^{\leq N}}$.

Assuming that $\boldsymbol{\chi}$ and E have expansions in $N^{-1/2}$, an expansion of the eigenvalue equation yields

$$(\mathbb{H}_0 - E_0)\boldsymbol{\chi}_1 + (\mathbb{H}_1 - E_{1/2})\boldsymbol{\chi}_0 = 0, \tag{14}$$

where $E_{1/2}$ is the coefficient of $N^{-1/2}$ in the expansion of E. Projecting this equation on $\boldsymbol{\chi}_0$ with the projector $\mathbb{P}_0 := |\boldsymbol{\chi}_0\rangle\langle\boldsymbol{\chi}_0|$ and then using $\mathbb{H}_0\boldsymbol{\chi}_0 = E_0\boldsymbol{\chi}_0$, we find

$$E_{1/2} = \langle \boldsymbol{\chi}_0, \mathbb{H}_1 \boldsymbol{\chi}_0 \rangle = 0, \tag{15}$$

where the last equality follows since $\mathbb{H}_1$ is cubic in the number of creation and annihilation operators and $\mathbb{U}_0$ is a Bogoliubov transformation, i.e., it maps linear combinations of a^*/a into linear combinations of a^*/a. (Alternatively, one can argue that $\boldsymbol{\chi}_0$ is quasi-free, and thus the left-hand side of (15) vanishes due to Wick's rule.) Therefore, no $N^{-1/2}$ order appears in the energy expansion (6); in fact, similar arguments can be used to show that every half-integer power of N^{-1} vanishes. Projecting Equation (14) on the orthogonal complement using $\mathbb{Q}_0 = 1 - \mathbb{P}_0$, we find

$$\begin{aligned}\chi_1 &= \frac{\mathbb{Q}_0}{E_0 - \mathbb{H}_0}\mathbb{H}_1\chi_0 = \mathbb{U}_0^*\left(\mathbb{U}_0\frac{\mathbb{Q}_0}{E_0 - \mathbb{H}_0}\mathbb{U}_0^*\right)\mathbb{U}_0\mathbb{H}_1\mathbb{U}_0^*|\Omega\rangle \\ &= \mathbb{U}_0^*\left(\int_{\mathbb{R}^d} \mathrm{d}x\, \Theta_1(x)a_x^*|\Omega\rangle + \int_{\mathbb{R}^{3d}} \mathrm{d}x^{(3)}\Theta_3(x^{(3)})a_{x_1}^*a_{x_2}^*a_{x_3}^*|\Omega\rangle\right),\end{aligned} \tag{16}$$

where we abbreviate $x^{(3)} = (x_1, x_2, x_3)$. Note that the last equality follows again from the facts that $\mathbb{H}_1$ is cubic in a^*/a and that $\mathbb{U}_0$ is a Bogoliubov transformation, as well as using that $\mathbb{U}_0\frac{\mathbb{Q}_0}{E_0-\mathbb{H}_0}\mathbb{U}_0^*$ is particle-number conserving; the functions $\Theta_1 \in L^2(\mathbb{R}^d)$ and $\Theta_3 \in L^2((\mathbb{R}^d)^3)$ can then be explicitly computed. Finally, the coefficients $\psi_{N,\ell}$ in the expansion (5) of the N-body ground state Ψ_N (Theorem 1) are constructed from (16) by (7). The functions $\psi_{N,\ell}$ depend on N by construction. However, this N-dependence is trivial, since it comes only from the splitting into condensate φ and excitations χ. The coefficients χ_ℓ in the expansion of χ are completely independent of N.

To prove Theorem 1, we follow a different route than using the eigenvalue equation: we expand $\mathbb{P} := |\chi\rangle\langle\chi|$ around $\mathbb{P}_0$ in a (Rayleigh-Schrödinger) perturbation series. By (13), the projectors $\mathbb{P}$ and $\mathbb{P}_0$ can be expressed as

$$\mathbb{P}_0 = \frac{1}{2\pi\mathrm{i}}\oint_\gamma \frac{1}{z - \mathbb{H}_0}\,\mathrm{d}z\,, \qquad \mathbb{P} = \frac{1}{2\pi\mathrm{i}}\oint_\gamma \frac{1}{z - \mathbb{H}}\,\mathrm{d}z\,, \tag{17}$$

for any $\mathcal{O}(1)$-contour γ whose interior contains both E and E_0 but no other point from the spectra of $\mathbb{H}$ and $\mathbb{H}_0$; this is possible by (13) and since $\mathbb{H}$ and $\mathbb{H}_0$ have a spectral gap of $\mathcal{O}(1)$. Now, one uses the expansion (11) of $\mathbb{H}$ to expand $(z - \mathbb{H})^{-1}$ around $(z - \mathbb{H}_0)^{-1}$. Since $\mathbb{P}$ is a rank-one projector, this immediately implies an expansion of the corresponding vector χ. After some lengthy computations using (11), the identity (12), the fact that $\mathbb{H}_j$ for j odd (even) is odd (even) in the number of creation and annihilation operators, and that $\mathbb{U}_0$ is a Bogoliubov transformation diagonalizing $\mathbb{H}_0$, one obtains the expansion (16) and the higher orders by using Cauchy's integral formula.

The main work in the proof of Theorem 1 is to estimate the error terms in the expansions above. For example, to control the error for $a = 1$, we bound $\mathbb{H}_1$, $\mathbb{R}_0$ and $\mathbb{R}_1$ by powers of $(\mathcal{N}_{\perp\varphi} + 1)$, prove a uniform bound on finite moments of the number operator with respect to χ, and provide suitable estimates for the commutators of powers of $\mathcal{N}_{\perp\varphi}$ with resolvents of $\mathbb{H}_0$. The expansion of the ground state energy $\mathcal{E}_N$ is then another consequence of the expansion of $\mathbb{P}$. For example, the next order term after the Bogoliubov energy is given by

$$E_1 = \langle\chi_0, \mathbb{H}_2\chi_0\rangle + \left\langle\chi_0, \mathbb{H}_1\frac{\mathbb{Q}_0}{E_0 - \mathbb{H}_0}\mathbb{H}_1\chi_0\right\rangle. \tag{18}$$

3 Applications

3.1 Edgeworth Expansion

Let the Bose gas be in its ground state Ψ_N and consider the statistics of experiments described by self-adjoint one-body operators on $L^2((\mathbb{R}^d)^N)$, i.e., operators of the form

$$B_j = \underbrace{\mathbb{1} \otimes \ldots \otimes \mathbb{1}}_{j-1} \otimes B \otimes \underbrace{\mathbb{1} \otimes \ldots \otimes \mathbb{1}}_{N-j} . \tag{19}$$

By the Born rule and since Ψ_N is permutation symmetric, the family $\{B_j\}_{j=1}^N$ defines a family of identically distributed random variables: the probability that the random variable B_j takes values in $A \subset \mathbb{R}$ is given by

$$\mathcal{P}_{\Psi_N}(B_j \in A) = \left\langle \Psi_N, \mathbb{1}_A(B_j)\Psi_N \right\rangle , \tag{20}$$

where $\mathbb{1}_A$ denotes the characteristic function of the set A. Since we consider N indistinguishable bosons, we are interested in describing the statistics of experiments described by symmetrized operators $\sum_{j=1}^N B_j$. Centering and rescaling leads us to consider operators

$$\mathcal{B}_N := \frac{1}{\sqrt{N}} \sum_{j=1}^{N} (B_j - \mathbb{E}_{\Psi_N}[B]) , \tag{21}$$

where $\mathbb{E}_{\Psi_N}[B] = \langle \Psi_N, B_1 \Psi_N \rangle$. From Theorem 1 we know that Ψ_N is not a product state, which implies that the random variables B_j are not independent. However, their dependency is weak, and on the level of the excitation Fock space, the correlations are described to leading order by a quasi-free state, i.e., a Bogoliubov transformation acting on the vacuum as in (12). Quasi-free states satisfy a Wick rule in analogy to Gaussian random variables, hence to leading order the statistics of (21) can be expected to be Gaussian. We need to make the additional assumption that the variance of that Gaussian does not vanish, which is equivalent to assuming that the Hartree minimizer φ is not an eigenstate of B. Indeed, it is shown in [12] that the fluctuations satisfy a weak Edgeworth expansion.

Theorem 2 *Let $a \in \mathbb{N}_0$ and $g \in L^1(\mathbb{R})$ such that its Fourier transform $\widehat{g} \in L^1(\mathbb{R}, (1 + |k|^{3a+4}))$. Then, for any self-adjoint bounded operator B on $L^2(\mathbb{R}^d)$ which does not have the Hartree minimizer φ as an eigenstate, there exists $C_B(a, g) > 0$ such that*

$$\left| \mathbb{E}_{\Psi_N}[g(\mathcal{B}_N)] - \sum_{j=0}^{a} N^{-\frac{j}{2}} \int_{\mathbb{R}^d} \mathrm{d}x \, g(x) p_j(x) \, \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \right| \leq C_B(a, g) N^{-\frac{a+1}{2}} . \tag{22}$$

The functions p_j are N-independent real polynomials of degree $3j$ which are even/odd for j even/odd. In particular,

$$p_0(x) = 1\,, \tag{23a}$$

$$p_1(x) = \frac{\alpha}{6\sigma^3} H_3\left(\frac{x}{\sigma}\right)\,, \tag{23b}$$

where $H_3(x) = x^3 - 3x$ is the third Hermite polynomial. The N-independent parameters $\sigma, \alpha \in \mathbb{R}$ are given in (25) and in [12].

The leading order ($a = 0$) of the expansion is a central limit theorem, which was proven in [48] (see also [7] for the related dynamical result). Analogously to Theorem 1, the constant $C_B(g, a)$ in Theorem 2 grows in a, hence (22) is an asymptotic expansion. It constitutes a weak Edgeworth expansion in the sense of [22], which, in particular, does not imply an asymptotic expansion of the probability $\mathcal{P}_{\Psi_N}(\mathcal{B}_N \in A)$ for $A \subset \mathbb{R}$. Also note that Edgeworth expansions give us a detailed picture of the probability distribution near the expectation value. A more detailed description of the tails are large deviation results, see, e.g., [31, 49].

Theorem 2 extends to a class of low-energy excited states of H_N. In this case, one does not obtain a Gaussian central limit theorem, because these excited states are not quasi-free. However, they are still given by some polynomial of creation operators acting on a quasi-free state, hence the limiting distribution is a Gaussian multiplied with a polynomial. This leads to a generalized Edgeworth-type expansion with different polynomials of higher degree (see [12] for the details).

To prove Theorem 2, we show an expansion of the characteristic function of the random variable $\mathcal{B}_N$. Making use of the expansion $\boldsymbol{\chi} = \boldsymbol{\chi}_0 + \mathcal{O}(N^{-1/2})$ from Theorem 1, we obtain

$$\begin{aligned}\left\langle \Psi_N, \mathrm{e}^{\mathrm{i}k\mathcal{B}_N}\Psi_N\right\rangle &= \left\langle \boldsymbol{\chi}, \mathrm{e}^{\mathrm{i}kU_{N,\varphi}\mathcal{B}_N U_{N,\varphi}^*}\boldsymbol{\chi}\right\rangle = \left\langle \Omega, \mathrm{e}^{\mathrm{i}k\mathbb{U}_0(a^*(qB\varphi)+a(qB\varphi))\mathbb{U}_0^*}\Omega\right\rangle + \mathcal{O}(N^{-\frac{1}{2}}) \\ &= \mathrm{e}^{-\frac{1}{2}\sigma^2k^2} + \mathcal{O}(N^{-\frac{1}{2}})\end{aligned} \tag{24}$$

with $q := 1 - |\varphi\rangle\langle\varphi|$ and where

$$\sigma := \|\nu\|\,, \qquad \nu := U_0 qB\varphi + \overline{V_0 qB\varphi}\,. \tag{25}$$

for certain bounded operators U_0, V_0 on $\{\varphi\}^\perp \subset L^2(\mathbb{R}^d)$, Here, we used that

$$\mathbb{U}_0\left(a^*(qB\varphi) + a(qB\varphi)\right)\mathbb{U}_0^* = a^*(\nu) + a(\nu) \tag{26}$$

since $\mathbb{U}_0$ from (12) is a Bogoliubov transformation. By Fourier transformation, this yields (22) for $a = 0$. The higher orders in (22) are computed along the same lines, making use of higher orders in Theorem 1.

Finally, let us compare Theorem 2 (which concerns the fluctuations $\mathcal{B}_N$ of dependent random variables distributed according to Ψ_N) with the corresponding result for the fluctuations $\mathcal{B}_N^{\mathrm{iid}}$ of i.i.d. random variables distributed according to the prod-

uct state $\varphi^{\otimes N}$. Standard probability theory (e.g. [43]) yields for $\mathcal{B}_N^{\mathrm{iid}}$ an Edgeworth expansion which is of the same structure as (22), i.e., a Gaussian multiplied with polynomials of degree $3j$ with the same even/odd structure. However, there are important differences: First, the variance of the Gaussian in the i.i.d. case is given by $\sigma_{\mathrm{iid}}^2 = \|qB\varphi\|^2 = \langle\varphi, B^2\varphi\rangle - \langle\varphi, B\varphi\rangle^2 \neq \sigma^2$, which can be seen analogously to (24) noting that $U_{N,\varphi}\varphi^{\otimes N} = |\Omega\rangle$. Moreover, the first polynomial p_1^{iid} is of the same functional form as p_1, but $\alpha_{\mathrm{iid}} \neq \alpha$. In the higher orders, also the functional form of the polynomials p_j^{iid} is different from p_j; for example, p_2^{iid} contains the Hermite polynomials H_4 and H_6 while p_2 has an additional contribution from H_2. This can be understood as follows: an Edgeworth expansion is an expansion in terms of the cumulants κ_ℓ of the distribution. In the i.i.d. situation, the cumulants satisfy the scaling relation $\kappa_\ell[\mathcal{B}_N^{\mathrm{iid}}] = N^{1-\frac{\ell}{2}}\kappa_\ell[\widetilde{B}]$ for $\widetilde{B} = B - \langle\varphi, B\varphi\rangle$. In contrast, in the interacting case, each cumulant has a full series expansion, which leads to the additional contributions (see [12] for a detailed discussion).

3.2 Binding Energy

Another application of Theorem 1 concerns the binding energy, i.e., the energy it takes to remove one particle from the Bose gas in its ground state. Let us introduce the unscaled Hamiltonian

$$H(N, v) = \sum_{j=1}^{N} \left(-\Delta_j + V^{\mathrm{trap}}(x_j)\right) + \sum_{1\leq i<j\leq N} v(x_i - x_j). \tag{27}$$

We now consider this Hamiltonian for N particles and for $N-1$ particles, both with the same weak interaction $(N-1)^{-1}v =: \lambda_N v$, i.e., we consider the N-body Hamiltonian

$$H(N, \lambda_N v) = \sum_{j=1}^{N} \left(-\Delta_j + V^{\mathrm{trap}}(x_j)\right) + \lambda_N \sum_{1\leq i<j\leq N} v(x_i - x_j), \tag{28}$$

which is the Hamiltonian from (1), and the $(N-1)$-body Hamiltonian

$$H(N-1, \lambda_N v) = \sum_{j=1}^{N-1} \left(-\Delta_j + V^{\mathrm{trap}}(x_j)\right) + \lambda_N \sum_{1\leq i<j\leq N-1} v(x_i - x_j). \tag{29}$$

If we denote the corresponding ground state energies by $E(N)$ and $\widetilde{E}(N-1)$, the binding energy is defined as

$$\Delta E(N) := E(N) - \widetilde{E}(N-1). \tag{30}$$

Theorem 1 gives us an expansion of $E(N)$. But note that in our expansion we have not separated the contributions in N coming from the number of particles and those from the coupling constant $\lambda_N = (N-1)^{-1}$. Hence, in order to obtain an expansion of $\widetilde{E}(N-1)$, we need to replace in E_N first the N by $N-1$ and then v by $\frac{N-2}{N-1}v$. The resulting series for $\widetilde{E}(N-1)$ then needs to be rewritten as a power series in N^{-1}, just as in Theorem 1. The result is a power series expansion of $\Delta E(N)$ in powers of N^{-1}.

Theorem 3 *Under the assumptions of Theorem 1, the binding energy $\Delta E(N)$ can be expanded as*

$$\Delta E(N) = \sum_{\ell=0}^{a} N^{-\ell} E_\ell^{\text{binding}} + \mathcal{O}(N^{-(a+1)}) \tag{31}$$

for any $a \in \mathbb{N}$. The coefficients E_ℓ^{binding} are stated explicitly in [10].

We know from [27] (or from Theorem 1 for $a = 0$) that the leading order contribution is given by

$$E_0^{\text{binding}} = N e_{\text{H}}(v) - (N-1) e_{\text{H}}\big((N-2)(N-1)^{-1}v\big) = e_{\text{H}} + \frac{1}{2}\langle\varphi, \big(v * |\varphi|^2\big)\varphi\rangle, \tag{32}$$

where $e_{\text{H}}(v)$ is the Hartree energy with potential v. The next order E_1^{binding} was derived in [40] for the Bose gas on the torus. Note that [40] discusses the extension to the inhomogeneous case as a conjecture, which we address here with Theorem 3 for $a = 1$. For $a = 2$ we compute the coefficient E_2^{binding} explicitly on the torus in [10].

4 Dynamics

4.1 Two-Body Interaction

Let us assume that the Bose gas has initially been prepared in the ground state Ψ_N of H_N. Now we switch off the trap and let the gas propagate. Hence, the N-body wave function $\Psi_N(t)$ at time $t > 0$ is given by the solution of the time-dependent Schrödinger equation, generated by H_N with $V^{\text{trap}} \equiv 0$. It is well known (see, e.g., [2–4, 6, 15, 19, 23–26, 30, 32, 44, 50, 52]) that the property of BEC is preserved by the time evolution, and that the time evolved condensate wave function $\varphi(t)$ is a solution of the Hartree equation,

$$\mathrm{i}\partial_t \varphi(t) = \left(-\Delta + v * |\varphi(t)|^2 - \mu^{\varphi(t)}\right)\varphi(t)\,, \tag{33}$$

for some conveniently chosen phase $\mu^{\varphi(t)} \in \mathbb{R}$. The main result of [13] is an asymptotic expansion of the resulting dynamics.

Theorem 4 *Let $a \in \mathbb{N}_0$ and $t \in \mathbb{R}$. Then there exists $C(a) > 0$ such that*

$$\left\| \Psi_N(t) - \sum_{\ell=0}^{a} N^{-\frac{\ell}{2}} \psi_{N,\ell}(t) \right\|_{L^2((\mathbb{R}^d)^N)} \leq \mathrm{e}^{C(a)t} N^{-\frac{a+1}{2}} . \tag{34}$$

The coefficients $\psi_{N,\ell}(t) \in L^2_{\mathrm{sym}}((\mathbb{R}^d)^N)$ are given in [13] in full generality.

The leading order ($a = 0$) of (34) was proven in [35, 39]. Related results for the higher orders ($a > 0$) were obtained in [11, 25, 26, 42]. Theorem 4 extends to a more general class of initial data. Besides, it implies an expansion of the reduced densities as well as a generalized Wick rule for the correlation functions (see [13] for the full statement).

Analogously to (7), the N-body wave functions $\Psi_{N,\ell}(t)$ are constructed by combining the time-evolved condensate $\varphi(t)$ with orthogonal excitations $\boldsymbol{\chi}(t) \in \mathcal{F}_{\perp\varphi(t)}$ and deriving a series expansion

$$\left\| \boldsymbol{\chi}(t) - \sum_{\ell=0}^{a} N^{-\frac{\ell}{2}} \boldsymbol{\chi}_\ell(t) \right\|_{\mathcal{F}_{\perp\varphi(t)}} \leq \mathrm{e}^{C(a)t} N^{-\frac{a+1}{2}} \tag{35}$$

for the time-evolved excitations. The leading order $\boldsymbol{\chi}_0(t)$ is given by the solution of the Bogoliubov equation, i.e., the time-dependent Schrödinger equation generated by the time-dependent analogue $\mathbb{H}_0(t)$ of the leading operator in (11). This is a very useful approximation because the time evolution $U_0(t, t_0) : \mathcal{F}_{\perp\varphi(t_0)} \to \mathcal{F}_{\perp\varphi(t)}$ generated by $\mathbb{H}_0(t)$ acts as a Bogoliubov transformation. As a consequence, solving the Bogoliubov equation essentially reduces to the problem of solving a 2×2 matrix differential equation, which is a huge simplification in complexity compared to the full N-body problem. Given the solution of the Bogoliubov equation, the first order correction is

$$\begin{aligned} \boldsymbol{\chi}_1(t) = &\sum_{j \in \{-1,1\}} \int_{\mathbb{R}^d} \mathrm{d}x \, \mathfrak{C}_1^{(j)}(t; x) a_x^{\sharp_j} \boldsymbol{\chi}_0(t) \\ &+ \sum_{(j_1,j_2,j_3) \in \{-1,1\}^3} \int_{\mathbb{R}^{3d}} \mathrm{d}x^{(3)} \, \mathfrak{C}_3^{(j_1,j_2,j_2)}(t; x^{(3)}) a_{x_1}^{\sharp_{j_1}} a_{x_2}^{\sharp_{j_2}} a_{x_3}^{\sharp_{j_3}} \boldsymbol{\chi}_0(t) , \end{aligned} \tag{36}$$

where we denoted $a^{\sharp_1} := a^*$ and $a^{\sharp_{-1}} := a$. The N-independent functions $\mathfrak{C}_1^{(j)}(t) \in L^2(\mathbb{R}^d)$ and $\mathfrak{C}_3^{(j_1,j_2,j_3)}(t) \in L^2((\mathbb{R}^3)^d)$ are explicitly given in terms of the initial data and the solution of the 2×2 matrix differential equation mentioned above (see, e.g., [9, Equations (3.22)]).

4.2 Regularized Nelson Model

The techniques from the previous subsection can also be applied to non-relativistic quantum field models such as the regularized Nelson model in a many-particle mean-field limit. This model describes N bosons that are linearly coupled to a quantized scalar (Klein–Gordon) field. The wave function $\Psi_N(t) \in L^2_{\mathrm{sym}}((\mathbb{R}^3)^N) \otimes \mathcal{F}$ evolves according to the Schrödinger equation with Hamiltonian

$$H_N^{\mathrm{Nelson}} = \sum_{j=1}^{N} \left(-\Delta_j + N^{-1/2}\widehat{\Phi}(x_j)\right) + \int_{\mathbb{R}^3} \mathrm{d}k\, \omega(k) a_k^* a_k. \tag{37}$$

Here, $\omega(k) = \sqrt{k^2 + m^2}$, with mass $m \geq 0$, is the dispersion relation of the field bosons, a_k^*/a_k are the bosonic pointwise creation/annihilation operators, and

$$\widehat{\Phi}(x) = \int_{\mathbb{R}^3} \mathrm{d}k\, \frac{g(k)}{\sqrt{2\omega(k)}} e^{-2\pi i k x} \left(a_k^* + a_{-k}\right) \tag{38}$$

denotes the field operator with even cutoff function $g : \mathbb{R}^3 \to \mathbb{R}$ such that $g/\sqrt{\omega}$ and g/ω are square integrable. If the particle-field state is initially prepared as a Bose–Einstein condensate of N particles with condensate wave function $\varphi_0 \in L^2(\mathbb{R}^3)$ and a coherent state of field bosons with classical field $\sqrt{N}\alpha_0 \in L^2(\mathbb{R}^3)$, then the condensation/coherent state structure is preserved under the time evolution generated by (37), see [1, 20, 33]. The corresponding mean-field equations describe the coupled evolution of the condensate wave function $\varphi(t)$ and the classical field $\alpha(t)$. They are known as Schrödinger–Klein–Gordon equations and given by

$$\begin{cases} i\partial_t \varphi(t) &= \Big(-\Delta + \Phi(t) - \frac{1}{2}\langle \varphi(t), \Phi(t)\varphi(t)\rangle \Big)\varphi(t) \\ i\partial_t \alpha(t) &= \omega\alpha(t) + \frac{g}{\sqrt{2\omega}} \widehat{|\varphi(t)|^2} \\ \Phi(t,x) &= \int_{\mathbb{R}^3} dk\, \frac{g(k)}{\sqrt{2\omega(k)}} e^{-2\pi i k x} \left(\overline{\alpha(t,k)} + \alpha(t,-k) \right) \end{cases} \tag{39}$$

with initial datum (φ_0, α_0), where $\widehat{|\varphi(t)|^2}$ is the Fourier transform of $|\varphi(t)|^2$. In [21] it was shown that (for suitably chosen initial states) the time evolution of the regularized Nelson model satisfies an asymptotic expansion in the spirit of Theorem 4. The main difference to the previous subsection, where N bosons interact via pair potentials, is that the system consists now of two types of particles. In order to study the fluctuations around the mean-field dynamics it is therefore necessary to factor out the Bose–Einstein condensate as well as the coherent state of the field bosons, which can be done in a similar manner as in (7). The resulting orthogonal excitations $\boldsymbol{\chi}(t)$ are elements of the double Fock space $\mathcal{F}_{\perp\varphi(t)} \otimes \mathcal{F}$ and the corresponding quadratic Bogoliubov Hamiltonian $\mathbb{H}_0(t)$ (and its higher-order corrections) are operators on this space. The Bogoliubov dynamics captures not only correlations among the particles and the field excitations themselves but also between the particles and field excitations.

Acknowledgements It is our pleasure to thank Marco Falconi, Nataša Pavlović, Peter Pickl, Robert Seiringer and Avy Soffer for the collaboration on the works [11, 13, 14, 21, 33, 39]. L.B. was supported by the German Research Foundation within the Munich Center of Quantum Science and Technology (EXC 2111). N.L. acknowledges support from the Swiss National Science Foundation through the NCCR SwissMap and funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N° 101024712. S.P. acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - project number 512258249.

References

1. Z. Ammari, M. Falconi, Wigner measures approach to the classical limit of the Nelson model: convergence of dynamics and ground state energy. J. Stat. Phys. **157**(2), 330–362 (2014)
2. Z. Ammari, M. Falconi, B. Pawilowski, On the rate of convergence for the mean field approximation of bosonic many-body quantum dynamics. Commun. Math. Sci. **14**(5), 1417–1442 (2016)
3. Z. Ammari, F. Nier, Mean field limit for bosons and propagation of Wigner measures. J. Math. Phys. **50**(4), 042107 (2009)
4. Z. Ammari, F. Nier, Mean field propagation of Wigner measures and BBGKY hierarchies for general bosonic states. J. Math. Pures Appl. **95**(6), 585–626 (2011)
5. M.H. Anderson, J.R. Ensher, M.R. Matthews, C.E. Wieman, E.A. Cornell, Observation of Bose-Einstein condensation in a dilute atomic vapor. Science **269**(5221), 198–201 (1995)
6. C. Bardos, F. Golse, N.J. Mauser, Weak coupling limit of the n-particle Schrödinger equation. Methods Appl. Anal. **7**(2), 275–294 (2000)
7. G. Ben Arous, K. Kirkpatrick, B. Schlein, A central limit theorem in many-body quantum dynamics. Commun. Math. Phys. **321**, 371–417 (2013)
8. Bose, Plancks Gesetz und Lichtquantenhypothese. Z. Phys. **26**, 178–181 (1924)
9. L. Boßmann, Low-energy spectrum and dynamics of the weakly interacting Bose gas. J. Math. Phys. **63**, 061102 (2022)
10. L. Boßmann, N. Leopold, D. Mitrouskas, S. Petrat, A Note on the Binding Energy for Bosons in the Mean-field Limit. Preprint (2023) arXiv:2307.13115
11. L. Boßmann, N. Pavlović, P. Pickl, A. Soffer, Higher order corrections to the mean-field description of the dynamics of interacting bosons. J. Stat. Phys. **178**(6), 1362–1396 (2020)
12. L. Boßmann, S. Petrat, Weak Edgeworth expansion for the mean-field Bose gas. Lett. Math. Phys. **113**(4), 77 (2023)
13. L. Boßmann, S. Petrat, P. Pickl, A. Soffer, Beyond Bogoliubov dynamics. Pure Appl. Anal. **3**(4), 677–726 (2021)
14. L. Boßmann, S. Petrat, R. Seiringer, Asymptotic expansion of low-energy excitations for weakly interacting bosons. Forum Math. Sigma **9**, e28 (2021)
15. L. Chen, J.O. Lee, B. Schlein, Rate of convergence towards Hartree dynamics. J. Stat. Phys. **144**(4), 872–903 (2011)
16. K.B. Davis, M.-O. Mewes, M.R. Andrews, N.J. van Druten, D.S. Durfee, D.M. Kurn, W. Ketterle, Bose-Einstein condensation in a gas of sodium atoms. Phys. Rev. Lett. **75**(22), 3969–3973 (1995)
17. A. Einstein, Quantentheorie des einatomigen idealen Gases. Sitzber. Kgl. Preuss. Akad. Wiss., 261–267 (1924)
18. A. Einstein, Quantentheorie des einatomigen idealen Gases. Zweite Abhandlung. Sitzber. Kgl. Preuss. Akad. Wiss., 3–14 (1925)
19. L. Erdős, H.-T. Yau, Derivation of the nonlinear Schrödinger equation from a many-body Coulomb system. Adv. Theor. Math. Phys. **5**(6), 1169–1205 (2001)

20. M. Falconi, Classical limit of the Nelson model with cutoff. J. Math. Phys. **54**(1), 012303 (2013)
21. M. Falconi, N. Leopold, D. Mitrouskas, S. Petrat, Bogoliubov dynamics and higher-order corrections for the regularized Nelson model. Rev. Math. Phys. **35**(4), 2350006 (2023)
22. K. Fernando, C. Liverani, Edgeworth expansions for weakly dependent random variables. Ann. Inst. H. Poincaré, Probab. Stat. **57**(1), 469–505 (2021)
23. J. Fröhlich, A. Knowles, A. Pizzo, Atomism and quantization. J. Phys. A **40**(12), 3033 (2007)
24. J. Fröhlich, A. Knowles, S. Schwarz, On the mean-field limit of bosons with Coulomb two-body interaction. Commun. Math. Phys. **288**(3), 1023–1059 (2009)
25. J. Ginibre, G. Velo, The classical field limit of scattering theory for non-relativistic many-boson systems. I. Commun. Math. Phys. **66**(1), 37–76 (1979)
26. J. Ginibre, G. Velo, The classical field limit of scattering theory for non-relativistic many-boson systems. II. Commun. Math. Phys. **68**(1), 45–68 (1979)
27. P. Grech, R. Seiringer, The excitation spectrum for weakly interacting bosons in a trap. Commun. Math. Phys. **322**(2), 559–591 (2013)
28. M. Grillakis, M. Machedon, D. Margetis, Second-order corrections to mean field evolution of weakly interacting bosons. I. Commun. Math. Phys. **294**(1), 273 (2010)
29. M. Grillakis, M. Machedon, D. Margetis, Second-order corrections to mean field evolution of weakly interacting bosons. II. Adv. Math. **228**(3), 1788–1815 (2011)
30. K. Hepp, The classical limit for quantum mechanical correlation functions. Commun. Math. Phys. **35**(4), 265–277 (1974)
31. K. Kirkpatrick, S. Rademacher, B. Schlein, A large deviation principle in many-body quantum dynamics. Ann. Henri Poincaré **22**(8), 2595–2618 (2021)
32. A. Knowles, P. Pickl, Mean-field dynamics: singular potentials and rate of convergence. Commun. Math. Phys. **298**(1), 101–138 (2010)
33. N. Leopold, P. Pickl, Mean-field limits of particles in interaction with quantized radiation fields, in *Macroscopic Limits of Quantum Systems*. ed. by D. Cadamuro, M. Duell, W. Dybalski, S. Simonella (Springer International Publishing, Cham, 2018), pp.185–214
34. M. Lewin, P.T. Nam, N. Rougerie, Derivation of Hartree's theory for generic mean-field Bose systems. Adv. Math. **254**, 570–621 (2014)
35. M. Lewin, P.T. Nam, B. Schlein, Fluctuations around Hartree states in the mean field regime. Am. J. Math. **137**(6), 1613–1650 (2015)
36. M. Lewin, P.T. Nam, S. Serfaty, J.P. Solovej, Bogoliubov spectrum of interacting Bose gases. Commun. Pure Appl. Math. **68**(3), 413–471 (2015)
37. E.H. Lieb, R. Seiringer, J.P. Solovej, J. Yngvason, *The Mathematics of the Bose Gas and its Condensation* (Birkhäuser, 2005)
38. D. Mitrouskas, Derivation of mean field equations and their next-order corrections: Bosons and fermions. Ph.D. thesis (2017)
39. D. Mitrouskas, S. Petrat, P. Pickl, Bogoliubov corrections and trace norm convergence for the Hartree dynamics. Rev. Math. Phys. **31**(8) (2019)
40. P.T. Nam, Binding energy of homogeneous Bose gases. Lett. Math. Phys. **108**(1), 141–159 (2018)
41. P.T. Nam, M. Napiórkowski, Bogoliubov correction to the mean-field dynamics of interacting bosons. Adv. Theor. Math. Phys. **21**(3), 683–738 (2017)
42. T. Paul, M. Pulvirenti, Asymptotic expansion of the mean-field approximation. Discret. Contin. Dyn. Syst. A **39**(4), 1891–1921 (2019)
43. V.V. Petrov. *Sums of Independent Random Variables* (De Gruyter, 2022)
44. P. Pickl, A simple derivation of mean field limits for quantum systems. Lett. Math. Phys. **97**(2), 151–164 (2011)
45. A. Pizzo, Bose particles in a box I. A convergent expansion of the ground state of a three-modes Bogoliubov Hamiltonian (2015). arXiv:1511.07022
46. A. Pizzo, Bose particles in a box II. A convergent expansion of the ground state of the Bogoliubov Hamiltonian in the mean field limiting regime (2015). arXiv:1511.07025

47. A. Pizzo, Bose particles in a box III. A convergent expansion of the ground state of the Hamiltonian in the mean field limiting regime (2015). arXiv:1511.07026
48. S. Rademacher, Central limit theorem for Bose gases interacting through singular potentials. Lett. Math. Phys. **110**(8), 2143–2174 (2020)
49. S. Rademacher, R. Seiringer, Large deviation estimates for weakly interacting bosons. J. Stat. Phys. **188**(1), 1–21 (2022)
50. I. Rodnianski, B. Schlein, Quantum fluctuations and rate of convergence towards mean field dynamics. Commun. Math. Phys. **291**(1), 31–61 (2009)
51. R. Seiringer, The excitation spectrum for weakly interacting bosons. Commun. Math. Phys. **306**(2), 565–578 (2011)
52. H. Spohn, On the Vlasov hierarchy. Math. Meth. Appl. Sci. **3**(1), 445–455 (1981)

A Study of the Radiation-Reaction on a Point Charge that Moves Along a Constant Applied Electric Field in an Electromagnetic B_{opp}–$L_{andé}$–T_{homas}–$P_{odolsky}$ Vacuum

Holly K. Carley and Michael K.-H. Kiessling

Abstract The relativistic problem of motion of a classical electrical point charge that has been placed between the plates of a charged capacitor and then released from rest is well-posed in Bopp–Landé–Thomas–Podolsky (BLTP) electrodynamics. That theory introduces a single new parameter, Bopp's $\varkappa$, a reciprocal length. The present article concerns the small-$\varkappa$ regime. Radiation-reaction effects on the motion are shown to appear at order $\varkappa^3$. It is found that in the initial phase the motion is accurately accounted for by test particle theory, with the inertia determined by the bare mass of the particle. Subsequently, radiation-reaction effects cause substantial deviations from the test particle motion.

1 Introduction

In this tribute to Detlef Dürr we focus on a lesser known scientific passion of Detlef, i.e., "lesser known in comparison to Bohmian Mechanics," and this is the classical electromagnetic radiation-reaction problem. Here is how Detlef once characterized the situation:

IN REMEMBRANCE OF
Detlef Dürr

H. K. Carley
Department of Mathematics, New York City College of Technology, City University of New York, 300 Jay Street, Brooklyn, New York, NY 11201, USA

M. K.-H. Kiessling (✉)
Department of Mathematics, Rutgers University, 110 Frelinghuysen Rd., Piscataway, NJ 08854, USA
e-mail: miki@math.rutgers.edu

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_23

> When people realized that there is a problem, quantum physics was invented. Then everyone began to work on quantum physics and eventually the problem was forgotten. But it still exists. (Detlef Dürr, private communication to the senior author, sometime in the mid 1980s.)

The problem, in a nutshell, is this: The symbolic system of equations of Lorentz electrodynamics with point charges is notoriously ill-defined. The energy and momentum densities of the electromagnetic Maxwell–Lorentz fields of a point charge source are not integrable over any neighborhood of the point charge, and also the Lorentz formula for the electromagnetic "self"-force on such a point charge source is ill-defined. More recently [DeHa2016] it was noted that also the Lorentz formula for the electromagnetic force of one point charge source onto another becomes ill-defined after relatively short times.

Undeterred by infinities [PPV2011], physicists have tried to extract the force of radiation-reaction on an accelerated point charge by analyzing the power emitted by it towards $|\boldsymbol{s}| = \infty$, per the retarded Liénard–Wiechert fields. Von Laue [Lau1909] obtained the expression $\frac{2e^2}{3c^3} P^{\perp}_{\mathbf{u}(\tau)} \cdot \frac{d^2}{d\tau^2}\mathbf{u}(\tau)$ for the Minkowski force four-vector due to the radiation (cf. [Jac1975]); here, τ denotes proper time and $P^{\perp}_{\mathbf{u}(\tau)}\cdot$ the four-projection onto the subspace that is four-orthogonal to the four-velocity $\mathbf{u}(\tau)$ $(= \frac{d}{d\tau}\mathbf{q}(\tau))$. Its third derivative of the particle's spacetime location $\mathbf{q}(\tau)$ has been the cause of consternation. In particular, it vanishes during intervals of constant four-acceleration when the Larmor formula predicts radiation. Moreover, when non-zero, it changes the initial-value problem from second to third order, and almost all solutions then display unphysical behavior. It has been argued that a way out of this "third-order" dilemma is offered by the fact that the radiation-reaction on the particle should cause only a small correction term to the test particle type force. The co-variant version of the equations of test particle motion reads $\frac{d^2}{d\tau^2}\mathbf{q}(\tau) = \frac{e}{mc}\mathbf{F}(\mathbf{q}(\tau)) \cdot \mathbf{u}(\tau)$, where $\mathbf{F}(\mathbf{s})$ is the Faraday tensor of the applied fields at a spacetime point $\mathbf{s}$. Hence the expression $\frac{d^2}{d\tau^2}\mathbf{u}(\tau)$ in von Laue's four vector should be interpreted as stand-in for the first-order proper time derivative of $\frac{e}{mc}\mathbf{F}(\mathbf{q}(\tau)) \cdot \mathbf{u}(\tau)$, which will not have derivatives of $\mathbf{q}(\tau)$ higher than second order. The so radiation-reaction-corrected equation of motion, sometimes called the Eliezer–Ford–O'Connell (EFC) equation of motion, reads

$$\frac{d^2}{d\tau^2}\mathbf{q}(\tau) = \frac{e}{mc}\mathbf{F}(\mathbf{q}(\tau)) \cdot \mathbf{u}(\tau) + \frac{2}{3}\frac{e^3}{mc^4} P^{\perp}_{\mathbf{u}(\tau)} \cdot \frac{d}{d\tau}\big(\mathbf{F}(\mathbf{q}(\tau)) \cdot \mathbf{u}(\tau)\big). \tag{1.1}$$

Landau and Lifshitz approximated (1.1) further by replacing all first-order proper time derivatives of $\mathbf{u}(\tau)$ obtained from $\frac{d}{d\tau}\big(\mathbf{F}(\mathbf{q}(\tau)) \cdot \mathbf{u}(\tau)\big)$ by $\frac{e}{mc}\mathbf{F}(\mathbf{q}(\tau)) \cdot \mathbf{u}(\tau)$. This approximation to (1.1) is known as the Landau–Lifshitz (LL) equation of motion.

Equation (1.1), and also its Landau–Lifshitz approximation, enjoy some practical successes. Interestingly, this practical success story has a serious blemish. Namely, for a point charge that moves along a constant applied electrostatic field the LL equation of motion simply reproduces the test particle motion, for its radiation-reaction force term vanishes in this situation; cf. [PMD2006]. One may hope to

obtain the radiation-reaction effects on a point charge that moves along a constant electric field by pushing the expansion of the EFC equation further until a non-vanishing radiation-reaction force term is obtained. However, all higher-order terms obtained from such an expansion vanish also.

As an aside, we mention that the LL equation has been derived rigorously as an effective equation of motion in a spacetime adiabatic limit, not for a point charge but for an extended charge distribution's geometric center, and with m standing for $m_{\mathrm{b}} + m_{\mathrm{f}}$, where m_{b} is a bare mass and m_{f} a field energy contribution; see [Spo2004]. Thus, if one works with extended charge distributions, as in the Abraham–Lorentz-type classical electron theory [Lor1904], and endows the particle with a non-vanishing m_{b}, then one may realistically hope to obtain higher-order radiation-reaction corrections that do not vanish for motion along a constant electric field. The LL equation of motion was also obtained in an asymptotic expansion about a "vanishing particle limit," which captures the motion of a particle with extended charge and mass distributions in the limit of vanishing size, mass, and charge, yet with nonzero charge-to-mass ratio, see [GHW2009]. However, higher-order terms in such expansions will eventually depend on largely arbitrary assumptions about the structure of the extended distributions. Moreover, the formulation of a properly Lorentz co-variant model with extended charged particles [ApKi2001] involves a non-trivial foliation of Minkowski spacetime and poses conceptual challenges for the initial value problem.

In this paper we are interested in the classical theory of motion for a true point charge that interacts with the electromagnetic fields it generates. In [Kie2019] the first well-posedness result of the joint initial value problem for the evolution of the electromagnetic fields and the relativistic motion of N point charges was announced, not for the ill-defined Lorentz electrodynamics with point charges, but for an electrodynamic model that goes back to work by Bopp [Bop1940,Bop1943], Landé–Thomas [Lan1941,LaTh1941], and Podolsky [Pod1942] (BLTP). The BLTP model replaces Maxwell's law of the electromagnetic vacuum (viz. $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{B}}$ and $\boldsymbol{\mathcal{E}} = \boldsymbol{\mathcal{D}}$) with a linear differential relation. A well-defined equation of motion was supplied in [Kie2019]. The proof of well-posedness of the joint Cauchy problem will be published in [KTZ2023]. Also the scattering problem for a single particle that encounters a localized potential is well-posed [Hetal2021]. These authors showed that in this problem the "self"-force formula of [Kie2019] can be converted into a formal Lorentz-type expression that involves integration over the whole past of the particle motion, first proposed in [LaTh1941] and further studied in [Zay2014,GPT2015].

In the following we demonstrate that BLTP electrodynamics captures the radiation-reaction on a point charge that is released from rest in a constant applied electric field. We understand our work as part of a proof-of-concept. In principle our approach can handle also more realistic models than BLTP electrodynamics, see [Kie2019].

2 BLTP Electrodynamics with a Single Point Charge

The electromagnetic vacuum in BLTP electrodynamics is defined by the two equations

$$\boldsymbol{\mathcal{H}}(t,\boldsymbol{s}) = \big(1+\varkappa^{-2}\Box\big)\boldsymbol{\mathcal{B}}(t,\boldsymbol{s})\,, \tag{2.1}$$

$$\boldsymbol{\mathcal{D}}(t,\boldsymbol{s}) = \big(1+\varkappa^{-2}\Box\big)\boldsymbol{\mathcal{E}}(t,\boldsymbol{s})\,; \tag{2.2}$$

in (2.1) and (2.2), the parameter $\varkappa^{-1}$ is the "Bopp length" [Bop1940,Bop1943], and $\Box \equiv c^{-2}\partial_t^2 - \Delta$ is the d'Alembertian, with c the vacuum speed of light. The evaluations $\boldsymbol{\mathcal{H}}(t,\boldsymbol{s}), \boldsymbol{\mathcal{B}}(t,\boldsymbol{s}), \boldsymbol{\mathcal{E}}(t,\boldsymbol{s})$, and $\boldsymbol{\mathcal{D}}(t,\boldsymbol{s})$ of the fields at the space point $\boldsymbol{s} \in \mathbb{R}^3$ and instant of time $t \in \mathbb{R}$ are defined in any convenient flat foliation of Minkowski spacetime into space & time. These fields satisfy the familiar system of pre-metric Maxwell field equations, which consist of two evolution equations

$$\tfrac{\partial}{\partial t}\boldsymbol{\mathcal{B}}(t,\boldsymbol{s}) = -c\boldsymbol{\nabla}\boldsymbol{\times}\boldsymbol{\mathcal{E}}(t,\boldsymbol{s})\,, \tag{2.3}$$

$$\tfrac{\partial}{\partial t}\boldsymbol{\mathcal{D}}(t,\boldsymbol{s}) = +c\boldsymbol{\nabla}\boldsymbol{\times}\boldsymbol{\mathcal{H}}(t,\boldsymbol{s}) - 4\pi e\delta_{\boldsymbol{q}(t)}(\boldsymbol{s})\boldsymbol{v}(t)\,, \tag{2.4}$$

and two constraint equations

$$\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{B}}(t,\boldsymbol{s}) = 0\,, \tag{2.5}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{D}}(t,\boldsymbol{s}) = 4\pi e\delta_{\boldsymbol{q}(t)}(\boldsymbol{s})\,. \tag{2.6}$$

Here, $e(>0)$ is the elementary electric charge, $\boldsymbol{q}(t) \in \mathbb{R}^3$ its position and $\boldsymbol{v}(t) \in \mathbb{R}^3$ its velocity at time t.

The particle's velocity is defined as usual to be the time-derivative of its position vector,

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{q}(t) =: \boldsymbol{v}(t). \tag{2.7}$$

In the relativistic generalization of Newton's point mechanics by Einstein, Lorentz, and Poincaré, the velocity $\boldsymbol{v}(t)$, in turn, changes with time according to

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\boldsymbol{v}(t)}{\sqrt{1-\frac{1}{c^2}|\boldsymbol{v}(t)|^2}} = \frac{1}{m_{\mathrm{b}}}\boldsymbol{f}(t); \tag{2.8}$$

here, $m_{\mathrm{b}} \neq 0$ is the *bare inertial rest mass* of the particle, and $\boldsymbol{f}(t)$ is the total electromagnetic force acting on it. Following Poincaré (cf. [Mil1998]) we define it as (cf. [Kie2019])

$$\boldsymbol{f}(t) := e\boldsymbol{\mathcal{E}}^{\mathrm{hom}} - \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^3}\boldsymbol{\Pi}^{\mathrm{field}}(t,\boldsymbol{s})\mathrm{d}^3\,s, \tag{2.9}$$

where $\boldsymbol{\mathcal{E}}^{\text{hom}}$ is a constant applied electric field (an idealization of the field between the plates of a capacitor), and $\boldsymbol{\Pi}^{\text{field}}(t, \boldsymbol{s})$ is the momentum vector-density of the Maxwell-BLTP fields

$$4\pi c\boldsymbol{\Pi}^{\text{field}} = \boldsymbol{\mathcal{D}}\times\boldsymbol{\mathcal{B}} + \boldsymbol{\mathcal{E}}\times\boldsymbol{\mathcal{H}} - \boldsymbol{\mathcal{E}}\times\boldsymbol{\mathcal{B}} - \varkappa^{-2}\big(\nabla\cdot\boldsymbol{\mathcal{E}}\big)\Big(\nabla\times\boldsymbol{\mathcal{B}} - \tfrac{1}{c}\tfrac{\partial}{\partial t}\boldsymbol{\mathcal{E}}\Big). \tag{2.10}$$

3 The Initial Data

As announced in [Kie2019] and shown in [KTZ2023], BLTP electrodynamics is well-posed as a joint initial value problem for the fields and the point charge, requiring initial data $\boldsymbol{\mathcal{B}}(0, \boldsymbol{s}), \boldsymbol{\mathcal{D}}(0, \boldsymbol{s}), \boldsymbol{\mathcal{E}}(0, \boldsymbol{s}), (\frac{\partial}{\partial t}\boldsymbol{\mathcal{E}})(0, \boldsymbol{s})$ for the fields, and $\boldsymbol{q}(0)$ and $\boldsymbol{v}(0)$ for the particle. The data for $\boldsymbol{\mathcal{B}}$ and $\boldsymbol{\mathcal{D}}$ are constrained by the divergence equations, and $\boldsymbol{v}(0)$ by $|\boldsymbol{v}(0)| < c$.

In the ensuing sections we discuss these BLTP-dynamical equations for a single point charge moving along the constant applied electric field $\boldsymbol{\mathcal{E}}^{\text{hom}}$, starting from rest, with the initial fields the sum of the external field and the electrostatic field of the point charge. Thus, for the particle we have

$$\boldsymbol{q}(0) = \boldsymbol{0} \qquad \text{and} \qquad \boldsymbol{v}(0) = \boldsymbol{0}. \tag{3.1}$$

For the fields we have

$$\boldsymbol{\mathcal{D}}(0, \boldsymbol{s}) \equiv \boldsymbol{\mathcal{E}}^{\text{hom}} + e\frac{\boldsymbol{s}}{|\boldsymbol{s}|^3} \tag{3.2}$$

and

$$\boldsymbol{\mathcal{E}}(0, \boldsymbol{s}) \equiv \boldsymbol{\mathcal{E}}^{\text{hom}} + e\frac{1 - (1 + \varkappa|\boldsymbol{s}|)e^{-\varkappa|\boldsymbol{s}|}}{|\boldsymbol{s}|^2}\frac{\boldsymbol{s}}{|\boldsymbol{s}|}. \tag{3.3}$$

We also have $\big(\partial_t\boldsymbol{\mathcal{E}}\big)(0, \boldsymbol{s}) \equiv \boldsymbol{0}$, as well as $\boldsymbol{\mathcal{B}}(0, \boldsymbol{s}) \equiv \boldsymbol{0}$.

4 The Solution of the Field Equations

For the initial data of our problem the electromagnetic fields outside the forward light cone of the initial location of the particle remain precisely the electrostatic fields, i.e., the magnetic field $\boldsymbol{\mathcal{H}}$ and the magnetic induction field $\boldsymbol{\mathcal{B}}$ vanish, while the electric displacement field $\boldsymbol{\mathcal{D}}(t, \boldsymbol{s})$ is given by (3.2) and the electric field $\boldsymbol{\mathcal{E}}(t, \boldsymbol{s})$ is given by (3.3), for all $t \geqslant 0$.

Inside the forward light cone of the initial particle location, but away from the particle position at t, the fields $\boldsymbol{\mathcal{D}}(t, \boldsymbol{s})$ and $\boldsymbol{\mathcal{H}}(t, \boldsymbol{s})$ are for all $t \geqslant 0$ given by $\boldsymbol{\mathcal{D}} = \boldsymbol{\mathcal{E}}^{\text{hom}} + \boldsymbol{\mathcal{D}}^{\text{ret}}_{\text{LW}}$ & $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{H}}^{\text{ret}}_{\text{LW}}$, with (the acceleration vector of the point charge is highlighted in red)

$$\mathcal{D}^{\text{ret}}_{\text{LW}}(t,\boldsymbol{s}) = e\frac{c^2-|\boldsymbol{v}|^2}{|\boldsymbol{s}-\boldsymbol{q}|^2}\frac{c\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\boldsymbol{v}}{\big(c-\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\cdot\boldsymbol{v}\big)^3}\bigg|_{\text{ret}} + e\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\big[\big(c\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\boldsymbol{v}\big)\times\boldsymbol{a}\big]}{|\boldsymbol{s}-\boldsymbol{q}|\big(c-\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\cdot\boldsymbol{v}\big)^3}\bigg|_{\text{ret}}, \tag{4.1}$$

$$\mathcal{H}^{\text{ret}}_{\text{LW}}(t,\boldsymbol{s}) = \boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})|_{\text{ret}}\times\mathcal{D}^{\text{ret}}_{\text{LW}}(t,\boldsymbol{s}) \tag{4.2}$$

the retarded Liénard–Wiechert fields. Here, $\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s}) = \frac{\boldsymbol{s}-\boldsymbol{q}}{|\boldsymbol{s}-\boldsymbol{q}|}$ is a *normalized* vector from $\boldsymbol{q}$ to $\boldsymbol{s}$, the notation "$|_{\text{ret}}$" means that $(\boldsymbol{q},\boldsymbol{v},\boldsymbol{a}) = (\boldsymbol{q},\boldsymbol{v},\boldsymbol{a})(t^{\text{ret}})$ to the left of "$|_{\text{ret}}$," with $t^{\text{ret}}(t,\boldsymbol{s})$ being defined implicitly by $c(t-t^{\text{ret}}) = |\boldsymbol{s}-\boldsymbol{q}(t^{\text{ret}})|$; inside the initial forward light cone, $0 < t^{\text{ret}} < t$. The terms $\propto \boldsymbol{a}$ in (4.1) and (4.2) account for the radiation.

Note that the electromagnetic Liénard–Wiechert fields $\mathcal{H}^{\text{ret}}_{\text{LW}}$ and $\mathcal{D}^{\text{ret}}_{\text{LW}}$ exhibit both a $\propto 1/r^2$ singularity and a $\propto 1/r$ singularity, where r denotes $|\boldsymbol{s}-\boldsymbol{q}(t)|$; they each have a directional singularity at the location of the point charge source, too.

Similarly, inside — and on — the forward light cone of the initial particle location, but away from the particle position at t, the MBLTP field solutions $\boldsymbol{\mathcal{B}}(t,\boldsymbol{s})$ and $\boldsymbol{\mathcal{E}}(t,\boldsymbol{s})$ for $t \geqslant 0$ are given by $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{B}}_0 + \boldsymbol{\mathcal{B}}_1$ and $\boldsymbol{\mathcal{E}} = \boldsymbol{\mathcal{E}}_0 + \boldsymbol{\mathcal{E}}_1$, with $\boldsymbol{\mathcal{B}}_0 \equiv \boldsymbol{0}$ and $\boldsymbol{\mathcal{E}}_0 \equiv \boldsymbol{\mathcal{E}}^{\text{hom}}$, and

$$\begin{aligned}\boldsymbol{\mathcal{E}}_1(t,\boldsymbol{s}) = {}& e\varkappa^2\Big(\frac{1-(1+\varkappa|\boldsymbol{s}|)e^{-\varkappa|\boldsymbol{s}|}}{\varkappa^2|\boldsymbol{s}|^2} - \tfrac{1}{2}\Big)\frac{\boldsymbol{s}}{|\boldsymbol{s}|} + e\varkappa^2\int_0^{ct-|\boldsymbol{s}|}\frac{J_2\big(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}|^2}\big)}{c^2(t-t')^2-|\boldsymbol{s}|^2}\boldsymbol{s}\,\mathrm{d}(ct')+\\ & e\varkappa^2\tfrac{1}{2}\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\boldsymbol{v}/c}{1-\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\cdot\boldsymbol{v}/c}\bigg|_{\text{ret}} - \\ & e\varkappa^2\int_0^{t^{\text{ret}}(t,\boldsymbol{s})}\frac{J_2\big(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\big)}{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\big(\boldsymbol{s}-\boldsymbol{q}(t')-\boldsymbol{v}(t')(t-t')\big)\,c\mathrm{d}t',\end{aligned} \tag{4.3}$$

$$\begin{aligned}\boldsymbol{\mathcal{B}}_1(t,\boldsymbol{s}) = {}& e\varkappa^2\tfrac{1}{2}\frac{\boldsymbol{v}\times\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})/c}{1-\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\cdot\boldsymbol{v}/c}\bigg|_{\text{ret}} - \\ & e\varkappa^2\int_0^{t^{\text{ret}}(t,\boldsymbol{s})}\frac{J_2\big(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\big)}{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\boldsymbol{v}(t')\times\big(\boldsymbol{s}-\boldsymbol{q}(t')\big)\,\mathrm{d}t'.\end{aligned} \tag{4.4}$$

The fields $\boldsymbol{\mathcal{B}}(t,\boldsymbol{s})$ and $\boldsymbol{\mathcal{E}}(t,\boldsymbol{s})$ are globally bounded in $\boldsymbol{s}$ for each t, and away from the point charge they are Lipschitz-continuous in $\boldsymbol{s}$, including across the initial forward light cone.

Similarly,

$$\begin{aligned}\nabla\cdot\boldsymbol{\mathcal{E}}(t,\boldsymbol{s}) = {}& e\varkappa^2\frac{e^{-\varkappa|\boldsymbol{s}|}-1}{|\boldsymbol{s}|} + e\varkappa^3\int_0^{ct-|\boldsymbol{s}|}\frac{J_1\big(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}|^2}\big)}{\sqrt{c^2(t-t')^2-|\boldsymbol{s}|^2}}\mathrm{d}(ct')+\\ & e\varkappa^2\frac{1}{\big(1-\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\cdot\boldsymbol{v}/c\big)}\frac{1}{|\boldsymbol{s}-\boldsymbol{q}|}\bigg|_{\text{ret}} - \\ & e\varkappa^3\int_0^{t^{\text{ret}}(t,\boldsymbol{s})}\frac{J_1\big(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\big)}{\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}}c\mathrm{d}t',\end{aligned} \tag{4.5}$$

and

$$\left(\nabla\times\boldsymbol{\mathcal{B}}-\tfrac{1}{c}\tfrac{\partial}{\partial t}\boldsymbol{\mathcal{E}}\right)(t,\boldsymbol{s})=e\varkappa^2\frac{1}{\left(1-\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\cdot\boldsymbol{v}/c\right)}\frac{1}{|\boldsymbol{s}-\boldsymbol{q}|}\frac{\boldsymbol{v}}{c}\Big|_{\rm ret}- \tag{4.6}$$
$$e\varkappa^3\int_0^{t^{\rm ret}(t,\boldsymbol{s})}\frac{J_1\left(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\right)}{\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}}\boldsymbol{v}(t')\mathrm{d}t'.$$

5 Evaluation of the Radiation-Reaction Force

With the help of these solution formulas, the electromagnetic force of the MBLTP field on its point charge source can be computed as follows. Since each electromagnetic field component is the sum of a vacuum field and a sourced field, the bilinear $\boldsymbol{\Pi}^{\rm field}$ decomposes into a sum of three types of terms: the vacuum-vacuum terms, the source-source terms, and the mixed vacuum-source terms. In our problem the vacuum field is $\boldsymbol{\mathcal{E}}^{\rm hom}$; it does not contribute to $\boldsymbol{\Pi}^{\rm field}$, but appears separately at rhs(2.9). As explained in [Kie2019], this term is not put in by hand but is a contribution to the momentum balance due to a surface integral at "$|\boldsymbol{s}|=\infty$." Hence the only contribution to rhs(2.9) from $\boldsymbol{\Pi}^{\rm field}$ is the source-source contribution, a "self"-field force in BLTP electrodynamics. Thus, (2.9) is given by

$$\boldsymbol{f}(t)=e\boldsymbol{\mathcal{E}}^{\rm hom}+\boldsymbol{f}^{\rm self}[\boldsymbol{q},\boldsymbol{v};\boldsymbol{a}](t), \tag{5.1}$$

where $e\boldsymbol{\mathcal{E}}^{\rm hom}$ is the Lorentz force evaluated with the vacuum field (i.e., a "test particle contribution" to the total force), and (after taking advantage of hyperbolicity; cf. [Kie2019])

$$\boldsymbol{f}^{\rm self}[\boldsymbol{q},\boldsymbol{v};\boldsymbol{a}](t)\equiv-\frac{\mathrm{d}}{\mathrm{d}t}\int_{B_{ct}(\boldsymbol{q}_0)}\Big(\boldsymbol{\Pi}^{\rm field}_{\rm source}(t,\boldsymbol{s})-\boldsymbol{\Pi}^{\rm field}_{\rm source}(0,\boldsymbol{s}-\boldsymbol{q}_0-\boldsymbol{v}_0t)\Big)d^3s \tag{5.2}$$
$$=-\frac{\mathrm{d}}{\mathrm{d}t}\int_{B_{ct}(\boldsymbol{q}_0)}\boldsymbol{\Pi}^{\rm field}_{\rm source}(t,\boldsymbol{s})d^3s, \tag{5.3}$$

with $\boldsymbol{\Pi}^{\rm field}_{\rm source}$ given by (2.10) with $(\boldsymbol{\mathcal{B}},\boldsymbol{\mathcal{D}}-\boldsymbol{\mathcal{E}}^{\rm hom},\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{E}}^{\rm hom},\boldsymbol{\mathcal{H}})$ in place of $(\boldsymbol{\mathcal{B}},\boldsymbol{\mathcal{D}},\boldsymbol{\mathcal{E}},\boldsymbol{\mathcal{H}})$. To go from (5.2) to (5.3) we made use of the initial data $\boldsymbol{q}_0=\boldsymbol{0}$ and $\boldsymbol{v}_0=\boldsymbol{0}$, and $\boldsymbol{\Pi}^{\rm field}_{\rm source}(0,\boldsymbol{s})\equiv\boldsymbol{0}$.

The "self"-field force can be evaluated using retarded spherical coordinates (r,ϑ,φ) to carry out the d^3s integrations over the ball $B_{ct}(\boldsymbol{q}_0)$, after which one can differentiate w.r.t. t. For this very special problem of straight line motion of a charge starting from rest at the origin, this yields

$$\begin{aligned} \boldsymbol{f}^{\text{self}}[\boldsymbol{q},\boldsymbol{v};\boldsymbol{a}](t) = \frac{e^2}{4\pi}\Bigg[&- \mathbf{Z}_{\boldsymbol{\xi}}^{[2]}(t,t) \\ &- \sum_{0\leqslant k\leqslant 1} c^{2-k}(2-k)\int_0^t \mathbf{Z}_{\boldsymbol{\xi}}^{[k]}\big(t,t^{\mathrm{r}}\big)(t-t^{\mathrm{r}})^{1-k}\mathrm{d}t^{\mathrm{r}} \\ &- \sum_{0\leqslant k\leqslant 2} c^{2-k}\int_0^t \tfrac{\partial}{\partial t}\mathbf{Z}_{\boldsymbol{\xi}}^{[k]}\big(t,t^{\mathrm{r}}\big)(t-t^{\mathrm{r}})^{2-k}\mathrm{d}t^{\mathrm{r}}\Bigg]. \end{aligned} \tag{5.4}$$

Here, $\boldsymbol{\xi}(t) \equiv (\boldsymbol{q},\boldsymbol{v},\boldsymbol{a})(t)$ and $\mathbf{Z}_{\boldsymbol{\xi}}^{[2]}(t,t) := \lim_{t^{\mathrm{r}}\to t}\mathbf{Z}_{\boldsymbol{\xi}}^{[k]}\big(t,t^{\mathrm{r}}\big)$, where

$$\mathbf{Z}_{\boldsymbol{\xi}}^{[k]}\big(t,t^{\mathrm{r}}\big) = \int_0^{2\pi}\int_0^{\pi}\big(1-\tfrac{1}{c}v(t^{\mathrm{r}})\cos\vartheta\big)\,\boldsymbol{\pi}_{\boldsymbol{\xi}}^{[k]}\big(t,\boldsymbol{q}(t^{\mathrm{r}})+c(t-t^{\mathrm{r}})\boldsymbol{n}\big)\sin\vartheta\mathrm{d}\vartheta\mathrm{d}\varphi\,, \tag{5.5}$$

with $v(t)$ defined by $v(t)|\boldsymbol{\mathcal{E}}^{\text{hom}}| \equiv \boldsymbol{v}(t)\cdot\boldsymbol{\mathcal{E}}^{\text{hom}}$, and with $\boldsymbol{n} = (\sin\vartheta\cos\varphi,\ \sin\vartheta\sin\varphi,\ \cos\vartheta)$ a normal vector to the retarded sphere of radius $r = c(t-t^{\mathrm{r}})$, where we measure ϑ from the $\boldsymbol{\mathcal{E}}^{\text{hom}}$ direction and φ from an arbitrary axis $\perp \boldsymbol{\mathcal{E}}^{\text{hom}}$.

Moreover, the $\boldsymbol{\pi}_{\boldsymbol{\xi}}^{[k]}(t,\boldsymbol{s})$ with $k\in\{0,1,2\}$ and $\boldsymbol{s}\neq\boldsymbol{q}$ are defined as follows. We set

$$\mathrm{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s}) := \frac{J_1\Big(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\Big)}{\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}}, \tag{5.6}$$

$$\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s}) := \frac{J_2\Big(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\Big)}{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\big(\boldsymbol{s}-\boldsymbol{q}(t')-\boldsymbol{v}(t')(t-t')\big), \tag{5.7}$$

and note that

$$\int_0^{t_{\xi^\circ}^{\text{ret}}(t,\boldsymbol{s})}\mathbf{K}_{\xi^\circ}(t',t,\boldsymbol{s})c\mathrm{d}t' = \int_0^{ct-|\boldsymbol{s}|}\frac{J_2(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}|^2})}{c^2(t-t')^2-|\boldsymbol{s}|^2}\boldsymbol{s}\,d(ct'), \tag{5.8}$$

$$\int_0^{t_{\xi^\circ}^{\text{ret}}(t,\boldsymbol{s})}\mathrm{K}_{\xi^\circ}(t',t,\boldsymbol{s})c\mathrm{d}t' = \int_0^{ct-|\boldsymbol{s}|}\frac{J_1(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}|^2})}{\sqrt{c^2(t-t')^2-|\boldsymbol{s}|^2}}\,d(ct'). \tag{5.9}$$

We will use $\big|_{\text{ret}}$ to mean that $\boldsymbol{q}(\tilde{t}),\boldsymbol{v}(\tilde{t}),\boldsymbol{a}(\tilde{t})$ are evaluated at $\tilde{t} = t_{\boldsymbol{\xi}}^{\text{ret}}(t,\boldsymbol{s})$, *not* $t_{\xi^\circ}^{\text{ret}}(t,\boldsymbol{s})$. Then

$$\begin{aligned} \boldsymbol{\pi}_{\boldsymbol{\xi}}^{[0]}(t,\boldsymbol{s}) = &-\varkappa^4\frac{1}{4}\left[\frac{(\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\frac{1}{c}\boldsymbol{v})\times(\frac{1}{c}\boldsymbol{v}\times\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s}))}{\big(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\big)^2}\right]_{\text{ret}} \\ &+\varkappa^4\frac{1}{2}\left[\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\frac{1}{c}\boldsymbol{v}}{1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}\right]_{\text{ret}}\times\int_0^{t_{\boldsymbol{\xi}}^{\text{ret}}(t,\boldsymbol{s})}\boldsymbol{v}(t')\times\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t' \\ &-\varkappa^4\frac{1}{2}\left[\frac{\frac{1}{c}\boldsymbol{v}\times\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}{1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}\right]_{\text{ret}}\times\int_0^{t_{\boldsymbol{\xi}}^{\text{ret}}(t,\boldsymbol{s})}c\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t' \end{aligned} \tag{5.10}$$

$$
\begin{aligned}
&+\varkappa^4\frac{1}{2}\left[\frac{\frac{1}{c}\boldsymbol{v}\times\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}{1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}\right]_{\mathrm{ret}}\times\left(\frac{1-(1+\varkappa|\boldsymbol{s}|)e^{-\varkappa|\boldsymbol{s}|}}{\varkappa^2|\boldsymbol{s}|^2}-\frac{1}{2}\right)\frac{\boldsymbol{s}}{|\boldsymbol{s}|}\\
&+\varkappa^4\frac{1}{2}\left[\frac{\frac{1}{c}\boldsymbol{v}\times\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}{1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}\right]_{\mathrm{ret}}\times\int_0^{t_{\xi^\circ}^{\mathrm{ret}}(t,\boldsymbol{s})}c\mathbf{K}_{\xi^\circ}(t',t,\boldsymbol{s})\mathrm{d}t'\\
&-\varkappa^4\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}c\mathbf{K}_{\xi}(t',t,\boldsymbol{s})\mathrm{d}t'\times\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}\boldsymbol{v}(t')\times\mathbf{K}_{\xi}(t',t,\boldsymbol{s})\mathrm{d}t'\\
&+\varkappa^4\left(\frac{1-(1+\varkappa|\boldsymbol{s}|)e^{-\varkappa|\boldsymbol{s}|}}{\varkappa^2|\boldsymbol{s}|^2}-\frac{1}{2}\right)\frac{\boldsymbol{s}}{|\boldsymbol{s}|}\times\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}\boldsymbol{v}(t')\times\mathbf{K}_{\xi}(t',t,\boldsymbol{s})\mathrm{d}t'\\
&+\varkappa^4\int_0^{t_{\xi^\circ}^{\mathrm{ret}}(t,\boldsymbol{s})}c\mathbf{K}_{\xi^\circ}(t',t,\boldsymbol{s})\mathrm{d}t'\times\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}\boldsymbol{v}(t')\times\mathbf{K}_{\xi}(t',t,\boldsymbol{s})\mathrm{d}t'\\
&-\varkappa^4c\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}\mathrm{K}_{\xi}(t',t,\boldsymbol{s})\mathrm{d}t'\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}\mathrm{K}_{\xi}(t',t,\boldsymbol{s})\boldsymbol{v}(t')\mathrm{d}t',\\
&-\varkappa^3\frac{1-e^{-\varkappa|\boldsymbol{s}|}}{|\boldsymbol{s}|}\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}\mathrm{K}_{\xi}(t',t,\boldsymbol{s})\boldsymbol{v}(t')\mathrm{d}t',\\
&+\varkappa^4\int_0^{t_{\xi^\circ}^{\mathrm{ret}}(t,\boldsymbol{s})}\mathrm{K}_{\xi^\circ}(t',t,\boldsymbol{s})c\mathrm{d}t'\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}\mathrm{K}_{\xi}(t',t,\boldsymbol{s})\boldsymbol{v}(t')\mathrm{d}t',
\end{aligned}
$$

and

$$
\begin{aligned}
\pi_{\xi}^{[1]}(t,\boldsymbol{s})=&-\varkappa^2\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\frac{(\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\boldsymbol{a}])\cdot\frac{1}{c}\boldsymbol{v}}{c^2\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^4}+\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\boldsymbol{a}}{2c^2\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^3}\right]_{\mathrm{ret}}\\
&-\varkappa^2\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\boldsymbol{a}}{c^2\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^3}\right]_{\mathrm{ret}}\times\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}\boldsymbol{v}(t')\times\mathbf{K}_{\xi}(t',t,\boldsymbol{s})\mathrm{d}t'\\
&+\varkappa^2\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\boldsymbol{a}}{c^2\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^3}\right]\right]_{\mathrm{ret}}\times\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}c\mathbf{K}_{\xi}(t',t,\boldsymbol{s})\mathrm{d}t'\\
&-\varkappa^2\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\boldsymbol{a}}{c^2\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^3}\right]\right]_{\mathrm{ret}}\times\left(\frac{1-(1+\varkappa|\boldsymbol{s}|)e^{-\varkappa|\boldsymbol{s}|}}{\varkappa^2|\boldsymbol{s}|^2}-\frac{1}{2}\right)\frac{\boldsymbol{s}}{|\boldsymbol{s}|}\\
&-\varkappa^2\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\boldsymbol{a}}{c^2\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^3}\right]\right]_{\mathrm{ret}}\times\int_0^{t_{\xi^\circ}^{\mathrm{ret}}(t,\boldsymbol{s})}c\mathbf{K}_{\xi^\circ}(t',t,\boldsymbol{s})\mathrm{d}t'\\
&+\varkappa^3\left[\frac{1}{1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}\right]_{\mathrm{ret}}\int_0^{t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})}\mathrm{K}_{\xi}(t',t,\boldsymbol{s})\left[\boldsymbol{v}(t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s}))+\boldsymbol{v}(t')\right]\mathrm{d}t',\\
&+\varkappa^2\left[\frac{1}{1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}\right]_{\mathrm{ret}}\frac{1-e^{-\varkappa|\boldsymbol{s}|}}{|\boldsymbol{s}|}\frac{1}{c}\boldsymbol{v}(t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s}))\\
&-\varkappa^3\left[\frac{1}{1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}\right]_{\mathrm{ret}}\int_0^{t_{\xi^\circ}^{\mathrm{ret}}(t,\boldsymbol{s})}\mathrm{K}_{\xi^\circ}(t',t,\boldsymbol{s})c\mathrm{d}t'\frac{1}{c}\boldsymbol{v}(t_{\xi}^{\mathrm{ret}}(t,\boldsymbol{s})),
\end{aligned}
\tag{5.11}
$$

and

$$\boldsymbol{\pi}_{\xi}^{[2]}(t,s) = -\varkappa^2\left[\frac{1}{\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},s)\right)^2}\frac{1}{c}\boldsymbol{v} - \left[1-\frac{1}{c^2}|\boldsymbol{v}|^2\right]\frac{\left(\boldsymbol{n}(\boldsymbol{q},s)-\frac{1}{c}\boldsymbol{v}\right)\times\left(\frac{1}{c}\boldsymbol{v}\times\boldsymbol{n}(\boldsymbol{q},s)\right)}{\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},s)\right)^4}\right]_{\mathrm{ret}} \tag{5.12}$$

$$+\varkappa^2\left[\left[1-\frac{1}{c^2}|\boldsymbol{v}|^2\right]\frac{\frac{1}{c}\boldsymbol{v}\times\boldsymbol{n}(\boldsymbol{q},s)}{\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},s)\right)^3}\right]_{\mathrm{ret}}\times\int_0^{t_\xi^{\mathrm{ret}}(t,s)} c\mathbf{K}_\xi(t',t,s)\mathrm{d}t'$$

$$-\varkappa^2\left[\left[1-\frac{1}{c^2}|\boldsymbol{v}|^2\right]\frac{\frac{1}{c}\boldsymbol{v}\times\boldsymbol{n}(\boldsymbol{q},s)}{\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},s)\right)^3}\right]_{\mathrm{ret}}\times\left(\frac{1-(1+\varkappa|s|)e^{-\varkappa|s|}}{\varkappa^2|s|^2}-\frac{1}{2}\right)\frac{s}{|s|}$$

$$-\varkappa^2\left[\left[1-\frac{1}{c^2}|\boldsymbol{v}|^2\right]\frac{\frac{1}{c}\boldsymbol{v}\times\boldsymbol{n}(\boldsymbol{q},s)}{\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},s)\right)^3}\right]_{\mathrm{ret}}\times\int_0^{t_{\xi^\circ}^{\mathrm{ret}}(t,s)} c\mathbf{K}_{\xi^\circ}(t',t,s)\mathrm{d}t'$$

$$-\varkappa^2\left[\left[1-\frac{1}{c^2}|\boldsymbol{v}|^2\right]\frac{\boldsymbol{n}(\boldsymbol{q},s)-\frac{1}{c}\boldsymbol{v}}{\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},s)\right)^3}\right]_{\mathrm{ret}}\times\int_0^{t_\xi^{\mathrm{ret}}(t,s)}\boldsymbol{v}(t')\times\mathbf{K}_\xi(t',t,s)\mathrm{d}t'.$$

Although this is an intimidating list of integrals, we can already extract an important conclusion: The equation of motion for our point charge does not feature time-derivatives of the particle position $\boldsymbol{q}(t)$ higher than second order. This result holds also for BLTP electrodynamics in general [Kie2019]. Hence, BLTP electrodynamics with point charges does not suffer from the $\dddot{\boldsymbol{q}}(t)$ problem.

5.1 The Small-$\varkappa$ Regime

To make further progress in the evaluation of the integrals we will concentrate our efforts on an asymptotic analysis of the small-$\varkappa$ regime. We make a formal power series expansion about $\varkappa = 0$ given by $\boldsymbol{f}^{\mathrm{self}}[\boldsymbol{q},\boldsymbol{v};\boldsymbol{a}](t) = \sum_{n=0}^{\infty}\boldsymbol{F}_0^{(n)}[\boldsymbol{q},\boldsymbol{v};\boldsymbol{a}](t)$, where $\boldsymbol{F}_0^{(n)} \propto \varkappa^n$; the subscript $_0$ at $\boldsymbol{F}_0^{(n)}$ indicates that we are expanding about $\varkappa = 0$. It is manifest that the terms $O(\varkappa^0)$ and $O(\varkappa^1)$ vanish identically, so we need to discuss terms $O(\varkappa^n)$ for $n \geqslant 2$. Several of the spherical integrations can been carried out explicitly in terms of well-known functions. In particular, the contributions $\boldsymbol{F}_0^{(2)}$ and $\boldsymbol{F}_0^{(3)}$ can be computed explicitly.

5.1.1 Radiation-Reaction at $O(\varkappa^2)$

To arrive at the $O(\varkappa^2)$ contribution we divide the expressions for $\boldsymbol{\pi}_{\xi}^{[k]}$ by $\varkappa^2$ and take $\varkappa \to 0$. The only two terms that survive in the limit are those in the first line of rhs(5.11) and rhs(5.12), respectively (indicated below by a superscript ^{,1}; later, also superscripts ^{,3}, ^{,4}, ^{,7} will appear). Carrying out the pertinent integrations in (5.5) one notes that the result only depends on t^{r}, not on t, so that the third line of rhs(5.4) vanishes at $O(\varkappa^2)$. Thus,

$$\tfrac{4\pi}{e^2}\boldsymbol{F}_0^{(2)}(t) = -\mathbf{Z}_{\boldsymbol{\xi}}^{[2],1}(t,t) - c\int_0^t \mathbf{Z}_{\boldsymbol{\xi}}^{[1],1}\big(t,t^{\mathrm{r}}\big)\mathrm{d}t^{\mathrm{r}}. \tag{5.13}$$

Explicitly, (5.13) reads

$$\begin{aligned}\boldsymbol{F}_0^{(2)}(t) =& -\frac{1}{2}e^2\varkappa^2\frac{\boldsymbol{v}(t)}{v(t)}\left[2\frac{c}{v(t)} - \frac{c^2}{v(t)^2}\ln\frac{1+\frac{1}{c}v(t)}{1-\frac{1}{c}v(t)}\right] \\ &+ e^2\varkappa^2\int_0^t \frac{c^3}{v(t^{\mathrm{r}})^3}\left[\tfrac{1}{c}v(t^{\mathrm{r}})\frac{2-\frac{v(t^{\mathrm{r}})^2}{c^2}}{1-\frac{v(t^{\mathrm{r}})^2}{c^2}} - \ln\frac{1+\frac{1}{c}v(t^{\mathrm{r}})}{1-\frac{1}{c}v(t^{\mathrm{r}})}\right]\tfrac{1}{c}\boldsymbol{a}(t^{\mathrm{r}})\mathrm{d}t^{\mathrm{r}}\end{aligned} \tag{5.14}$$

The term in the first line of rhs(5.14) is the contribution from the first line of rhs(5.12), the term in the second line of rhs(5.14) is the contribution from the first line of rhs(5.11). Since for straight-line motion $\boldsymbol{v}(t)$ and $\boldsymbol{a}(t)$ are collinear, and $\boldsymbol{a}(t) = \dot{\boldsymbol{v}}(t)$, one can carry out the time integration in the second line of rhs(5.14) in terms of elementary functions of $\boldsymbol{v}/c$, and a few algebraic manipulations then give (see the erratum in [Kie2019])

$$\boldsymbol{F}_0^{(2)}(t) = \mathbf{0}. \tag{5.15}$$

In this problem of straight line motion in a constant external electric field, with the particle starting from rest, the BLTP radiation-reaction force vanishes *exactly* at $O(\varkappa^2)$.

5.1.2 Radiation-Reaction at $O(\varkappa^3)$

We next evaluate the $O(\varkappa^3)$ contribution to the radiation-reaction force for small $\varkappa$. To arrive at the $O(\varkappa^3)$ contribution, subtract the $O(\varkappa^2)$ terms from the expressions for $\boldsymbol{\pi}_{\boldsymbol{\xi}}^{[k]}$, divide the result by $\varkappa^3$ and take the limit $\varkappa \to 0$. This yields the contributions from three $\boldsymbol{\pi}_{\boldsymbol{\xi}}^{[k]}$ terms $\propto \varkappa^3$, namely $\boldsymbol{\pi}_{\boldsymbol{\xi}}^{[1],4}$, $\boldsymbol{\pi}_{\boldsymbol{\xi}}^{[1],7}$, and $\boldsymbol{\pi}_{\boldsymbol{\xi}}^{[2],3}$. They contribute the following force $\propto \varkappa^3$,

$$\begin{aligned}\boldsymbol{F}_0^{(3)}(t) =& -e^2\varkappa^3\boldsymbol{q}(t) \\ &- e^2\varkappa^3\tfrac{4}{3}\int_0^t \frac{c^2}{v(t^{\mathrm{r}})^2}\left\{1+\tfrac{1}{2}\frac{c}{v(t^{\mathrm{r}})}\ln\frac{1-\frac{1}{c}v(t^{\mathrm{r}})}{1+\frac{1}{c}v(t^{\mathrm{r}})}\right\}(t-t^{\mathrm{r}})\boldsymbol{a}(t^{\mathrm{r}})\mathrm{d}t^{\mathrm{r}} \\ &-e^2\varkappa^3\tfrac{2}{3}\int_0^t\left(1-\frac{v(t^{\mathrm{r}})^2}{c^2}\right)\frac{c}{v(t^{\mathrm{r}})}\left\{1+\tfrac{1}{2}\frac{c}{v(t^{\mathrm{r}})}\ln\frac{1-\frac{1}{c}v(t^{\mathrm{r}})}{1+\frac{1}{c}v(t^{\mathrm{r}})}\right\}c\mathrm{d}t^{\mathrm{r}}\frac{\boldsymbol{\mathcal{E}}^{\mathrm{hom}}}{|\boldsymbol{\mathcal{E}}^{\mathrm{hom}}|}\end{aligned} \tag{5.16}$$

Integration by parts yields for the integral in the second line of rhs(5.16)

$$\int_0^t \frac{c^2}{v(t^{\mathrm{r}})^2}\left\{1+\tfrac{1}{2}\frac{c}{v(t^{\mathrm{r}})}\ln\frac{1-\frac{1}{c}v(t^{\mathrm{r}})}{1+\frac{1}{c}v(t^{\mathrm{r}})}\right\}(t-t^{\mathrm{r}})\boldsymbol{a}(t^{\mathrm{r}})\mathrm{d}t^{\mathrm{r}} = \tag{5.17}$$

$$\int_0^t c\int_0^{v(t^{\mathrm{r}})/c}\frac{1}{x^2}\left\{1+\tfrac{1}{2}\frac{1}{x}\ln\frac{1-x}{1+x}\right\}\mathrm{d}x\mathrm{d}t^{\mathrm{r}} = \tag{5.18}$$

$$-\frac{1}{2}\int_0^t\frac{c}{v(t^{\mathrm{r}})}\left[1+\left(1-\frac{v(t^{\mathrm{r}})^2}{c^2}\right)\tfrac{1}{2}\frac{c}{v(t^{\mathrm{r}})}\ln\frac{1-\frac{1}{c}v(t^{\mathrm{r}})}{1+\frac{1}{c}v(t^{\mathrm{r}})}\right]c\mathrm{d}t^{\mathrm{r}}$$

Comparison with the third line of rhs(5.16) reveals cancellations, and we end up with

$$\boldsymbol{F}_0^{(3)}(t) = -\frac{1}{3}e^2\varkappa^3\boldsymbol{q}(t). \tag{5.19}$$

This is a very surprising result: the $O(\varkappa^3)$ term of the radiation-reaction force in our initial value problem is a harmonic oscillator force! This result relies on the particular setup of the initial data and the geometry of the problem, but not more.

6 The Volterra Equation for the Acceleration

The equation of motion can be recast as a Volterra integral equation for the acceleration,

$$\boldsymbol{a}(t) = W[\boldsymbol{v}]\cdot\left(e\boldsymbol{\mathcal{E}}^{\mathrm{hom}} + \boldsymbol{f}^{\mathrm{self}}[\boldsymbol{q},\boldsymbol{v};\boldsymbol{a}]\right)(t). \tag{6.1}$$

Here,

$$W[\boldsymbol{v}] := \tfrac{1}{m_{\mathrm{b}}}\sqrt{1-\tfrac{|\boldsymbol{v}|^2}{c^2}}\left[\boldsymbol{I}_{3\times3} - \tfrac{1}{c^2}\boldsymbol{v}\otimes\boldsymbol{v}\right], \tag{6.2}$$

which for motion along $\boldsymbol{\mathcal{E}}^{\mathrm{hom}}$ is the same as

$$W[\boldsymbol{v}] := \tfrac{1}{m_{\mathrm{b}}}\left(1-\tfrac{|\boldsymbol{v}|^2}{c^2}\right)^{3/2}\boldsymbol{I}_{3\times3}. \tag{6.3}$$

In [KTZ2023] we show that the Volterra equation can be uniquely solved to yield $\boldsymbol{a}$ as a nonlinear expression in $\boldsymbol{q}$ and $\boldsymbol{v}$, posing a second-order initial value problem for $\boldsymbol{q}(t)$.

6.1 The Volterra Equation to $O(\varkappa^3)$ (Small $\varkappa$)

With the radiation-reaction force evaluated to $O(\varkappa^3)$ we obtain the equation of motion

$$\boldsymbol{a}(t) = \tfrac{1}{m_{\mathrm{b}}}\left(1-\tfrac{1}{c^2}|\boldsymbol{v}(t)|^2\right)^{3/2}\left(e\boldsymbol{\mathcal{E}}^{\mathrm{hom}} - \tfrac{1}{3}e^2\varkappa^3\boldsymbol{q}(t)\right). \tag{6.4}$$

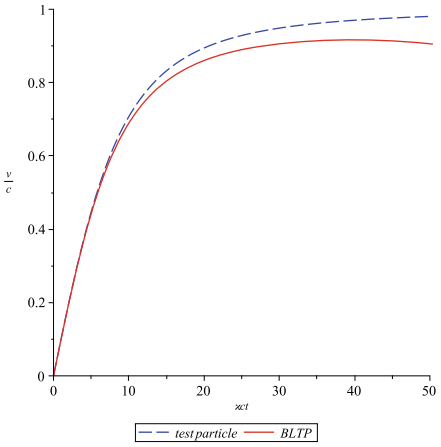


Fig. 1 The velocity of a point charge, starting from rest in a constant applied electrostatic field $\boldsymbol{\mathcal{E}}^{\text{hom}} = 10e\varkappa^2$, versus time, as per test particle theory (dashed curve), resp. BLTP electrodynamics with radiation-reaction included to $O(\varkappa^3)$ (continuous curve), when $\varkappa e^2/m_\text{b}c^2 = 0.01$. The period of the velocity of the BLTP motion is $\varkappa cT = 160$. The test particle's velocity asymptotes to c

Equation (6.4) is equivalent to the problem of special-relativistic test particle motion in a harmonic oscillator potential, featuring time-periodic solutions with conserved particle energy

$$U = \sqrt{m_\text{b}^2c^4 + |\boldsymbol{p}|^2c^2} - e\boldsymbol{\mathcal{E}}^{\text{hom}} \cdot \boldsymbol{q} + \tfrac{1}{6}e^2\varkappa^3|\boldsymbol{q}|^2. \tag{6.5}$$

It is not clear to us whether this means that the validity of the $O(\varkappa^3)$ approximation is restricted to short times $\varkappa ct \ll 1$ (see Fig. 1, which looks reasonable) or whether such periodic motion over longer times is a genuine feature of BLTP electrodynamics as long as $\varkappa e^2/m_\text{b}c^2 \ll 1$. In the latter case BLTP electrodynamics would presumably be eliminated for good from the list of contenders for a realistic classical electrodynamics.

7 Summary and Outlook

We have shown that BLTP electrodynamics, as defined in [Kie2019], accounts for radiation-reaction effects on the point charge motion along a constant applied electric field, thereby passing a litmus test that other models (in particular, the Landau–Lifshitz and Eliezer–Ford–O'Connell equations of motion) fail. Our results are based on a small-$\varkappa$ expansion of the BLTP force expression. This is acceptable for our "proof-of-concept" demonstration. However, BLTP electrodynamics is physically viable at most in the large $\varkappa$ regime [CKP2019]—if at all. An assessment of the large-$\varkappa$ regime we leave to some future work.

Acknowledgements We thank the referee for interesting comments that have prompted us to improve the presentation.

References

[ApKi2001] W. Appel, M.K.-H. Kiessling, Mass and spin renormalization in Lorentz electrodynamics. Ann. Phys. **289**, 24–83 (2001)
[Bop1940] F. Bopp, Eine lineare Theorie des Elektrons. Ann. Phys. **430**, 345–384 (1940)
[Bop1943] F. Bopp, Lineare Theorie des Elektrons II. Ann. Phys. **434**, 573–608 (1943)
[BoIn1934] M. Born, L. Infeld, Foundation of the new field theory. Proc. Roy. Soc. Lond. A **144**, 425–451 (1934)
[CKP2019] H.K. Carley, M.K.-H. Kiessling, V. Perlick, On the Schrödinger spectrum of a Hydrogen atom with electrostatic Bopp–Landé–Thomas–Podolsky interaction between electron and proton. Int. J. Mod. Phys. A **34**, 1950146 (23pp.) (2019)
[DeHa2016] D.-A. Deckert, V. Hartenstein, On the initial value formulation of classical electrodynamics. J. Phys. A: Math. Theor. **49**, 445202 (19pp.) (2016)
[GHW2009] S.E. Gralla, A. Harte, R.M. Wald, A rigorous derivation of electromagnetic self-force. Phys. Rev. D **80**, 024031 (2009)
[GPT2015] J. Gratus, V. Perlick, R.W. Tucker, On the self-force in Bopp–Podolsky electrodynamics. J. Phys. A: Math. Theor. **48**, 435401 (28pp.) (2015)
[Hetal2021] V. Hoang, M. Radosz, A. Harb, A. DeLeon, A. Baza, Radiation reaction in higher-order electrodynamics. J. Math. Phys. **62**, 072901 (31pp.) (2021)
[Jac1975] J.D. Jackson, *Classical Electrodynamics*, 2nd edn. (Wiley, New York, 1975)
[Kie2019] M.K.-H. Kiessling, Force on a point charge source of the classical electromagnetic field. Phys. Rev. D **100**, 065012 (2019). *Erratum* ibid. **101**, 109901(E) (2020)
[KTZ2023] M.K.-H. Kiessling, A.S. Tahvildar-Zadeh, Bopp-Landé-Thomas-Podolsky electrodynamics as initial value problem, in preparation (2023)
[Lan1941] A. Landé, Finite Self-Energies in Radiation Theory. Part I. Phys. Rev. **60**, 121–126 (1941)
[LaTh1941] A. Landé, L.H. Thomas, Finite Self-Energies in Radiation Theory. Part II. Phys. Rev. **60**, 514–523 (1941)
[Lau1909] M. Laue, Die Wellenstrahlung einer bewegten Punktladung nach dem Relativitätsprinzip. Ann. Phys. **28**, 436–442 (1909)
[Lor1904] H.A. Lorentz, Weiterbildung der Maxwell'schen Theorie: Elektronentheorie., Encyklopädie d. Mathematischen Wissenschaften **V**2, Art. 14, pp. 145–288 (1904)
[Mil1998] A.I. Miller, *Albert Einstein's Special Theory of Relativity* (Springer, New York, 1998)
[PMD2006] A. de Parga, R. Mares, S. Dominguez, An unphysical result for the Landau-Lifshitz equation of motion for a charged particle. Rev. Mex. Fis. **52**, 139–142 (2006)
[Pod1942] B. Podolsky, A generalized electrodynamics. Part I. Non-quantum, Rev. **62**, 68–71 (1942)
[PPV2011] E. Poisson, A. Pound, I. Vega, The motion of point particles in curved spacetime. Living Rev. Rel. **14**(190), 7 (2011)
[Spo2004] H. Spohn, *Dynamics of Charged Particles And Their Radiation Fields* (Cambridge, UP, 2004)
[Zay2014] A.E. Zayats, Self-interaction in the Bopp-Podolsky electrodynamics: can the observable mass of a charged particle depend on its acceleration? Ann. Phys. (NY) **342**, 11–20 (2014)

Revisiting Quantum Mechanical Zero-Range Potentials

Rodolfo Figari and Alessandro Teta

Abstract In this contribution we make a brief overview of history and recent results in the theory of many quantum particles interacting via zero-range forces. We recall the regularisation mechanism suggested by several authors in the past in order to avoid the "fall to the center" problem in three-body systems. Following those suggestions a family of three-body point interaction Hamiltonians bounded from below were made available recently. We conclude showing that a similar kind of ultraviolet problem is already present in the theory of point interaction Hamiltonians in one-body Quantum Mechanics. A careful look to the entire family of many center point interaction Hamiltonians shows that the great majority of them do not become either singular or trivial when the positions of two or more scattering centers tend to coincide. In this sense, those Hamiltonians appear to be renormalised by default as opposed to the "local" point interaction Hamiltonians usually considered in the literature since the early days of Quantum Mechanics. The renormalization mechanism turns out to be very similar to the one used in the three-body problem.

1 Introduction

Both authors of this contribution had the privilege and the pleasure of having had Detlef for many years as a friend. Becoming friends with Detlef has always been simple and engaging because he has been a joyful, smiling and generous person. He always loved Italy where every year he came on holiday with family and dog. He has also collaborated with many Italian researchers in Rome, L'Aquila, Naples, Genoa etc.

At the turn of the century, in connection with his interest on the foundations of Quantum Mechanics, he began to analize the decoherence effects induced on a

R. Figari (✉)
Gran Sasso Science Institute, Via F. Crispi, 7, 67100 L'Aquila, Italy
e-mail: figari@unina.it

A. Teta
Dipartimento di Matematica, La Sapienza Università di Roma, P.le A. Moro, 5, 00185 Roma, Italy
e-mail: teta@mat.uniroma1.it

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_24

quantum system by interaction with its environment. In particular, Herbert Spohn and Detlef analized the coupling to the radiation field of a charged particle as a source of decoherence [8]. In the same book [8] and in the two editions of [14] many theoretical and experimental physicists presented their ideas and results about the emergence of "classicality" triggered by interaction in open quantum systems. A few years later the group led by M. Arndt and A. Zeilinger performed experiments of matter wave quantum interference using different kind of fullerenes, giant macromolecules made of many thousands of atoms. Outputs of one of these experiments and their theoretical interpretation were published in [15] where the authors analized the observed loss of spatial coherence in the particle wave function due to collisions with background gases present between the slits and the particle detector. In their words "From the gradual suppression of quantum interference with increasing gas pressure we are able to support quantitatively both the predictions of decoherence theory and our picture of the interaction process. We thus explore the practical limits of matter wave interferometry at finite gas pressures and estimate the required experimental vacuum conditions for interferometry with even larger objects". The theoretical techniques they made use of were a mixture of quantum and classical approaches. In particular, they used a classical time of flight picture between scattering processes in order to avoid the difficulties of a many-body quantum treatment. The single collision process was analized quantum mechanically using the Joos and Zeh formula for the scattering matrix in a two particle system one of which is much lighter than the other.

Together with Detlef we decided to perform a more rigorous investigation of the dynamics of the collisional decoherence process using zero-range forces as interaction model. In fact, Spohn himself suggested the idea to consider models of environment simpler than the radiation field to give a detailed quantitative estimate of the decoherence effect.

The first step was to compute the loss of visibility of the interference fringes in a single scattering process or, which is the same, to give a more rigorous derivation of the Joos and Zeh formula.

The second step turned out to be much more difficult to do. In order to investigate the loss of spatial coherence due to an increasing number of scattering processes one needs a detailed description of the dynamics of a heavy particle travelling in a "gas" of light particles in three dimensions which is a project of considerable difficulty with any choice of short range interactions between the heavy particle and the light ones. In particular, there was a long lasting problem connected to zero-range interactions: all the proposed Hamiltonians for a system of three quantum particles interacting via zero-range forces were unbounded from below, a severe sign of instability of the physical system under investigation.

In the next section we will summarize the history of the successes and failures in the theory of many quantum particles interacting via zero-range forces, the conjectures which were formulated in order to solve the ultraviolet problem and recent results in the field.

In the last two sections we will show that the short distance singularities in the many-particle systems are already observed at the level of one quantum particle in a many-center array of the so-called local zero-range scatterers. We will clarify that

there is a large family of point interaction Hamiltonians (the so-called non local ones) that do not show any ultraviolet pathology and can be considered effective models of well behaved short range interactions. As a relevant application we detail their use in the Born-Oppenheimer approximation of a three-body system made of one light and two heavy particles.

2 Zero-Range Interactions in Many-Particle Systems

Here we recall the difficulties arising in the construction of Hamiltonians for many-particle systems with zero-range interactions in dimension three and describe some recent attempts to obtain a physically reasonable version of such Hamiltonians.

In order to explain the problem we first consider a system composed of three identical spinless bosons of mass $1/2$. Let us fix the center of mass reference frame so that x_1, x_2 and $x_3 = -x_1 - x_2$ represent the Cartesian coordinates of the three particles. We also introduce the Jacobi coordinates $x = x_2 - x_3$ and $y = x_1 - \frac{1}{2}(x_2 + x_3)$ and the Hilbert space of the system

$$L^2_{\mathrm{sym}}(\mathbb{R}^6) := \left\{\psi \in L^2(\mathbb{R}^6) \,\middle|\, \psi(x,y) = \psi(-x,y) = \psi\left(\tfrac{1}{2}\,x + y, \tfrac{3}{4}\,x - \tfrac{1}{2}\,y\right)\right\}. \tag{2.1}$$

Indeed, notice that the symmetry conditions in (2.1) corresponds to the exchange of particles $2,3$ and $1,2$ that implies also the condition $\psi(x,y) = \psi\left(\frac{1}{2}\,x - y, -\frac{3}{4}\,x - \frac{1}{2}\,y\right)$, associated with the exchange of particles $3,1$. If the bosons interact via zero-range forces, then the system is described, at least formally, by the Hamiltonian

$$\hat{\mathcal{H}}^{3b} = \mathcal{H}^{3b}_0 + \nu\,\delta(x) + \nu\,\delta(\tfrac{1}{2}\,x + y) + \nu\,\delta(\tfrac{1}{2}\,x - y) \tag{2.2}$$

where $\nu \in \mathbb{R}$ is a (formal) coupling constant and $\mathcal{H}^{3b}_0$ is the free Hamiltonian of the system, i.e.

$$\mathcal{H}^{3b}_0 = -\Delta_x - \tfrac{3}{4}\,\Delta_y\,. \tag{2.3}$$

Here and in the following we set $\hbar = 1$. In order to define a rigorous counterpart of $\hat{\mathcal{H}}^{3b}$, one needs to build a perturbation of the free Hamiltonian supported on the coincidence hyperplanes

$$\pi_1 := \left\{(x,y) \in \mathbb{R}^6 \mid x = 0\right\}, \quad \pi_2 := \left\{(x,y) \in \mathbb{R}^6 \mid y = -\tfrac{1}{2}\,x\right\}, \quad \pi_3 := \left\{(x,y) \in \mathbb{R}^6 \mid y = \tfrac{1}{2}\,x\right\}. \tag{2.4}$$

In other words, we look for a self-adjoint and bounded from below extension in $L^2_{\mathrm{sym}}(\mathbb{R}^6)$ of the following symmetric and densely defined operator

$$\dot{\mathcal{H}}^{3b}_0 := \mathcal{H}^{3b}_0\big|_{\mathscr{D}(\dot{\mathcal{H}}^{3b}_0)}, \qquad \mathscr{D}(\dot{\mathcal{H}}^{3b}_0) := H^2_0(\mathbb{R}^6 \setminus \cup_{i=1}^3 \pi_i)\,. \tag{2.5}$$

Roughly speaking, any such extension acts as the free Hamiltonian in $\mathbb{R}^6 \setminus \cup_{i=1}^3 \pi_i$ and is characterized by a (singular) boundary condition on each hyperplane π_i. The choice of a physical reasonable extension, or equivalently of the boundary condition, is not easy due to the fact that the defect spaces associated to the operator (2.5) are infinite dimensional. A possible approach is to proceed by analogy with the case of a point interaction in the one-body case ([1]). Following such analogy one is led to the boundary condition on π_1

$$\psi(x, y) = \frac{\xi(y)}{|x|} - \frac{1}{a}\,\xi(y) + \ o(1), \qquad |x| \to 0, \tag{2.6}$$

where a is the two-body scattering length and $\xi(y) = \lim_{|x|\to 0} (|x|\psi(x, y))$. Notice that, due to the symmetry constraint, (2.6) implies the analogous boundary conditions on π_2, π_3. Unfortunately, as a matter of fact the extension defined via the boundary condition (2.6) is symmetric but not self-adjoint and its self-adjoint extensions are all unbounded from below. This instability phenomenon is known as Thomas effect and it has been rigorously proved in [17, 18]. Therefore the natural problem arises of figuring out if and how one can modify the boundary condition (2.6) to obtain a bounded from below Hamiltonian. In a comment on this point, at the end of the paper [17] Minlos and Faddeev claim that it is possible to find a self-adjoint and bounded from below Hamiltonian by replacing

$$-\frac{1}{a}\,\xi(y) \ \to \ -\frac{1}{a}\,\xi(y) + (K\xi)(y) \tag{2.7}$$

in the boundary condition (2.6), where K is a convolution operator in the Fourier space with a kernel $K(p - p')$ satisfying

$$K(p) \sim \frac{\gamma}{|p|^2}\,, \qquad \text{for} \quad |p| \to \infty \tag{2.8}$$

with the positive constant γ sufficiently large. The authors do not explain the reason of their assertion neither they clarify the physical meaning of the boundary condition (2.7). They only conclude: "A detailed development of this point of view is not presented here because of lack of space" and, strangely enough, their idea has never been developed in the literature.

Almost twenty years later Albeverio, Høegh-Krohn and Wu ([2]) have proposed the same kind of recipe formulated in position space, i.e.

$$-\frac{1}{a}\,\xi(y) \ \to \ -\frac{1}{a}\,\xi(y) + \frac{\gamma}{|y|}\xi(y)\,.$$

Also in this case the proof has been postponed to a forthcoming paper which has never been published. In recent years the proposal of Minlos and Faddeev has been reconsidered ([3, 9, 10, 12], see also [16] for a different approach) and it has been proved that a self-adjoint and bounded from below Hamiltonian can actually be

constructed. More precisely, one considers the extensions characterized by the boundary condition

$$\psi(x, y) = \frac{\xi(y)}{|x|} + \alpha(y)\xi(y) + \mathrm{o}(1), \qquad |\mathbf{x}| \to 0, \tag{2.9}$$

where α is a position dependent parameter given by

$$\alpha : y \longmapsto -\frac{1}{a} + \frac{\gamma}{|y|}\,\theta(|y|) \tag{2.10}$$

with γ a positive parameter representing the strength of the regularization and

$$\theta \in L^\infty(\mathbb{R}^+), \qquad |\theta(r) - 1| \leqslant c\, r \quad \text{for some } c > 0. \tag{2.11}$$

Simple choices for the function θ are

$$\theta(r) = 1\,, \qquad \theta(r) = \begin{cases} 1 & r \leqslant b \\ 0 & r > b \end{cases}, \qquad \theta(r) = e^{-r/b}\,, \qquad b > 0\,. \tag{2.12}$$

We stress that, due to the symmetry constraints of $L^2_{\mathrm{sym}}(\mathbb{R}^6)$, the boundary condition (2.9) implies the analogous boundary conditions for $y \to -\frac{1}{2}x$ and for $y \to \frac{1}{2}x$.

Taking the formal operator characterized by the boundary condition (2.9) as a starting point, one can construct a quadratic form defined on a suitable domain and one can prove that for γ larger than a threshold value $\gamma_c^{3b} > 0$ the quadratic form is closed and bounded from below. Therefore it uniquely defines a self-adjoint and bounded from below operator which, by definition, represents the Hamiltonian for the three boson system with regularized contact interactions (see [3] for details).

Let us comment on the physical meaning of the regularization introduced with the boundary condition (2.9). Comparing (2.6) and (2.9), we can interpret $-\alpha(y)^{-1}$ as an effective, position dependent scattering length

$$a_{\mathrm{eff}}(|y|) := -\frac{1}{\alpha(y)} = \frac{a|y|}{|y| - a\,\gamma\,\theta(|y|)}$$

associated to the zero-range interaction between the particles 2, 3 taking place when $x_2 = x_3$, i.e., for $x = 0$. On the other hand the coordinate $|y|$ is the distance between the third particle 1 and the common position of particles 2, 3. Then we have

$$\lim_{|y|\to 0} a_{\mathrm{eff}}(|y|) = 0\,, \qquad \lim_{|y|\to\infty} a_{\mathrm{eff}}(|y|) = a\,,$$

i.e., the effective scattering length associated to the interaction of particles 2, 3 converges to zero if the third particle 1 is close to the common position of particles 2, 3 while it converges to a if the third particle 1 is sufficiently far away. In other words,

we have introduced a three-body interaction which reduces to zero the two-body interaction when the third particle approaches the common position of the first two. On the other hand, the usual two-body interaction is restored if the third particle is at large distance from the first two. This is precisely the mechanism that prevents in our model the fall to the center phenomenon, i.e., the Thomas effect.

The approach outlined for the Hamiltonian of the three boson system can be generalized to the case of the Hamiltonian for N identical bosons interacting with a test particle via zero-range forces. It is worth to stress that such a model is particularly interesting for a rigorous analysis of the decoherence effect induced on the test particle by multiple collisions with the light particles of the Bose gas.

In the rest of this section we give a sketch of the construction of such Hamiltonian (for all the technical details we refer to [9]). Let us consider a system in dimension three made of N identical bosons with mass m_0 interacting via zero-range forces with a test particle with mass m. The Hilbert space of the system is $L^2(\mathbb{R}^3) \otimes L^2_{\mathrm{sym}}(\mathbb{R}^{3N})$, $N \geqslant 2$, and, at a formal level, the Hamiltonian reads

$$\hat{\mathcal{H}} = \mathcal{H}_0 + \nu \sum_{i=1}^{N} \delta(x - x_i), \tag{2.13}$$

where $\mathcal{H}_0$ is the free Hamiltonian, given by

$$\mathcal{H}_0 := \frac{1}{2m} \Delta_x - \frac{1}{2m_0} \sum_{i=1}^{N} \Delta_{x_i} \tag{2.14}$$

Proceeding as in the three bosons case, we define the rigorous counterpart of (2.13) as a self-adjoint and bounded from below extension of

$$\dot{\mathcal{H}}_0 := \mathcal{H}_0|_{\mathscr{D}(\dot{\mathcal{H}}_0)}, \qquad \mathscr{D}(\dot{\mathcal{H}}_0) := H^2_0(\mathbb{R}^{3(N+1)} \setminus \pi) \tag{2.15}$$

where

$$\pi := \bigcup_{i=1}^{N} \pi_i \,, \qquad \pi_i := \left\{ (x, x_1, \ldots, x_N) \in \mathbb{R}^{3(N+1)} \,\middle|\, x = x_i \right\} \tag{2.16}$$

and in particular we consider the extension characterized by the following boundary condition on π_i

$$\psi(x, x_1, \ldots, x_N) = \frac{\xi(\frac{x+\eta x_i}{1+\eta}, x_1, \ldots, \check{x}_i, \ldots, x_N)}{|x - x_i|} + \alpha_i(x_1, \ldots, x_N)\xi(x_i, \ldots, \check{x}_i, \ldots, x_N) + o(1) \tag{2.17}$$

for $|x - x_i| \to 0$, where

$$\eta = \frac{m}{m_0} \,, \tag{2.18}$$

the symbol $\check{x}_i$ indicates that the variable x_i is omitted and

$$\alpha_i(x_1,\dots,x_N) = -\frac{1}{a} + \gamma \sum_{j=1,\, j\neq i}^{N} \frac{\theta(|x_i - x_j|)}{|x_i - x_j|}\,. \tag{2.19}$$

Recall that a is the two-body scattering length, $\gamma > 0$ and θ satisfies (2.11). We also stress that, due to the bosonic symmetry, the function ξ is invariant under the exchange of the last $N-1$ arguments.

Using a quadratic form method one can show that for $\gamma > \gamma_c$, with

$$\gamma_c = \frac{2(1+\eta)}{\pi} \sin^{-1}\left(\frac{1}{1+\eta}\right) - \frac{2\sqrt{\eta(\eta+2)}}{\pi(N-1)(1+\eta)}\,, \tag{2.20}$$

the formal extension of (2.15) characterized by the boundary condition (2.17) can be actually constructed as a self-adjoint and bounded from below Hamiltonian in $L^2(\mathbb{R}^3) \otimes L^2_{\mathrm{sym}}(\mathbb{R}^{3N})$.

Here we only give the final result of the construction, i.e., the rigorous definition of such Hamiltonian. Let us introduce the "potential" produced by the "charge density" ξ distributed on π_i

$$\widehat{\mathcal{G}_i^\lambda \xi}(k, p_1, \dots, p_N) = \sqrt{\frac{2}{\pi}}(1+\eta)\frac{\hat{\xi}(k+p_i, p_1, \dots \check{p}_i, \dots p_N)}{k^2 + \eta \sum_{j=1}^N p_i^2 + 2m\lambda}\,, \qquad \lambda > 0 \tag{2.21}$$

and the corresponding total potential

$$\mathcal{G}^\lambda \xi = \sum_{i=1}^{N} \mathcal{G}_i^\lambda \xi\,. \tag{2.22}$$

Then the Hamiltonian is given by

$$\mathscr{D}(\mathcal{H}) = \left\{\psi \mid \psi = w^\lambda + \mathcal{G}^\lambda \xi,\ w^\lambda \in H^2(\mathbb{R}^{3(N+1)}),\ \xi \in H^1(\mathbb{R}^{3N}),\ (O^\lambda \xi)_i = w^\lambda|_{\pi_i},\ \forall i\right\}, \tag{2.23}$$

$$\mathcal{H}\psi = \mathcal{H}_0 w^\lambda - \lambda \mathcal{G}^\lambda \xi \tag{2.24}$$

where

$$\begin{aligned}
&(O^\lambda \xi)_i(x_1,\dots,x_N) = \alpha_i(x_1,\dots,x_N)\xi(x_i,\dots,\check{x}_i,\dots,x_N) \\
&+\tfrac{1}{(2\pi)^{3N/2}} \int dp_1 \dots dp_N\, e^{i\sum_{j=1}^N x_j\cdot p_j} \Bigg[\sqrt{\tfrac{\eta}{(1+\eta)^2} p_i^2 + \tfrac{\eta}{1+\eta} \sum_{j=1,\, j\neq i}^{N} p_j^2 + \tfrac{2m\lambda}{1+\eta}}\, \hat{\xi}(p_i, p_1, \dots \check{p}_i, \dots p_N) \\
&-\tfrac{1+\eta}{2\pi^2} \sum_{j=1,\, j\neq i}^{N} \int dk\, \frac{\hat{\xi}(k+p_j, p_i - k, \dots, \check{p}_i, \dots, \check{p}_j, \dots, p_N)}{(1+\eta)k^2 - 2\eta\, p_i \cdot k + \eta \sum_{l=1}^N p_l^2 + 2m\lambda} \Bigg].
\end{aligned} \tag{2.25}$$

We stress that the last equality in (2.23) plays the role of a boundary condition on the hyperplane π_i and it is not hard to see that it is an equivalent formulation of the boundary condition (2.17).

We conclude this section describing at formal level the limit of the Hamiltonian for $\eta \to 0$, i.e., when the bosons are infinitely heavy and then become fixed centers at the positions $x_1, \ldots, x_N$. To simplify the notation with consider the case $N = 2$, with $2m = 1$. We will see that in the limit one obtains the one-particle Hamiltonian with non local point interactions placed at x_1, x_2.

Let us introduce the notation

$$R = |x_1 - x_2|, \quad q_1 = 4\pi\xi(x_1, x_2), \quad q_2 = 4\pi\xi(x_2, x_1), \quad \hat{\alpha}(R) = 4\pi\left(-\frac{1}{a} + \gamma\,\frac{\theta(R)}{R}\right). \tag{2.26}$$

Then, from (2.21), (2.22), we find for $\eta \to 0$

$$\begin{aligned}
&\mathcal{G}^\lambda\xi(x, x_1, x_2) \to \tfrac{1}{(2\pi)^{9/2}}\int dk dp_1 dp_2\; e^{ix\cdot k + ix_1\cdot p_1 + ix_2\cdot p_2}\sqrt{\tfrac{2}{\pi}}\;\frac{\hat{\xi}(k+p_1, p_2) + \hat{\xi}(k+p_2, p_1)}{k^2+\lambda}\\
&= \sqrt{\tfrac{2}{\pi}}\;\xi(x_1, x_2)\,\tfrac{1}{(2\pi)^{3/2}}\int dk\;\frac{e^{i(x-x_1)\cdot k}}{k^2+\lambda} + \sqrt{\tfrac{2}{\pi}}\;\xi(x_2, x_1)\,\tfrac{1}{(2\pi)^{3/2}}\int dk\;\frac{e^{i(x-x_2)\cdot k}}{k^2+\lambda}\\
&= \sum_{i=1}^{2}\frac{e^{-\sqrt{\lambda}|x-x_i|}}{4\pi|x-x_i|}\;q_i\,.
\end{aligned} \tag{2.27}$$

Moreover, from (2.25) we have for $\eta \to 0$

$$\begin{aligned}
&(O^\lambda\xi)_1(x_1, x_2) \to \Big(\hat{\alpha}(R) + \frac{\sqrt{\lambda}}{4\pi}\Big)q_1 - \tfrac{1}{(2\pi)^3}\int dp_1 dp_2\; e^{ix_1\cdot p_1 + ix_2\cdot p_2}\,\tfrac{1}{2\pi^2}\int dk\;\frac{\hat{\xi}(k+p_2, p_1-k)}{k^2+\lambda}\\
&= \Big(\hat{\alpha}(R) + \frac{\sqrt{\lambda}}{4\pi}\Big)q_1 - \xi(x_2, x_1)\;\tfrac{1}{2\pi^2}\int dk\;\frac{e^{i(x_1-x_2)\cdot k}}{k^2+\lambda}\\
&= \Big(\hat{\alpha}(R) + \frac{\sqrt{\lambda}}{4\pi}\Big)q_1 - \frac{e^{-\sqrt{\lambda}R}}{4\pi R}\;q_2
\end{aligned} \tag{2.28}$$

and $(O^\lambda\xi)_2(x_1, x_2) \to \Big(\hat{\alpha}(R) + \frac{\sqrt{\lambda}}{4\pi}\Big)q_2 - \frac{e^{-\sqrt{\lambda}R}}{4\pi R}\;q_1$. Let us also observe that from (2.20) we have $\gamma_c \to 1$ for $\eta \to 0$. In conclusion we find that $\mathcal{H}$ reduces in the limit $\eta \to 0$ to the following one-particle Hamiltonian in $L^2(\mathbb{R}^3)$

$$\mathscr{D}(H_{\hat{\alpha}(R)}) = \Big\{u \mid u = \phi^\lambda + \sum_{i=1}^{2}\frac{e^{-\sqrt{\lambda}|\cdot - x_i|}}{4\pi|\cdot - x_i|}\;q_i,\;\; \phi^\lambda \in H^2(\mathbb{R}^3),\;\; q_i \in \mathbb{C}, \sum_{j=1}^{2}\Sigma_{ij}^\lambda q_j = \phi^\lambda(x_i)\Big\}, \tag{2.29}$$

$$H_{\hat{\alpha}(R)}u = -\Delta\phi^\lambda - \lambda\sum_{i=1}^{2}\frac{e^{-\sqrt{\lambda}|\cdot - x_i|}}{4\pi|\cdot - x_i|}q_i \tag{2.30}$$

where

$$\Sigma^{\lambda} = \begin{pmatrix} \hat{\alpha}(R) + \frac{\sqrt{\lambda}}{4\pi} & -\frac{e^{-\sqrt{\lambda}R}}{4\pi R} \\ -\frac{e^{-\sqrt{\lambda}R}}{4\pi R} & \hat{\alpha}(R) + \frac{\sqrt{\lambda}}{4\pi} \end{pmatrix}. \tag{2.31}$$

Notice that the behavior of $u \in \mathscr{D}(H_{\hat{\alpha}(R)})$ near x_i is given by

$$u(x \sim x_i) = \frac{q_i}{4\pi|x - x_i|} + \hat{\alpha}(R)q_i + o(1)\,. \tag{2.32}$$

It is immediate to realize that we have obtained the Hamiltonian of a particle subject to non local point interactions placed at the centers x_1 and x_2 with a strength $\hat{\alpha}(R)$ depending on the distance between the centers according to the last formula in (2.26). We shall see in the next section that, in contrast with the standard local point interactions, such Hamiltonian exhibits a regular behaviour for $R \to 0$, i.e., when the two centers approach each other.

3 Point Potentials in One-Body Quantum Mechanics in $\mathbb{R}^3$

A many-center point interaction Hamiltonian in one-body quantum theory is defined to be one of the self-adjoint extensions of the symmetric operator $\dot{H}_0 = -\Delta \restriction C_0^{\infty}(\mathbb{R}^d \setminus \{y_1, \ldots, y_n\})$ where $\{y_1, \ldots, y_n\}$ are the positions in $\mathbb{R}^d$ of the point interaction centers. As is well known, there are non trivial self-adjoint extensions of $\dot{H}_0$ in dimension d equal one, two and three, whereas in dimensions higher or equal to four $\dot{H}_0$ is already essentially self-adjoint and its closure is the free Laplacian.
In [5] L. Dabrowskj and H. Grosse classified via the von Neumann construction all the self-adjoint extensions of $\dot{H}_0$ in $d = 1, 2, 3$. In order to distinguish them from the point interaction Hamiltonians since long extensively used in various fields of applied and mathematical physics (see [1]) they used for them the name non-local in spite of the fact that the family of extensions they found should also include all the previously defined Hamiltonians.
Below we summarize their results and the subsequent computations made by us for the case $d = 3$ and $n = 2$ in order to simplify the comparison with the three-body problem that we examined in the previous section. Furthermore, we are going to consider Hamiltonian operators having in their domains only functions which are symmetric in the exchange of the positions of the interaction centers.
The defect spaces N_i and N_{-i} of the symmetric operator $\dot{H}_0 = -\Delta \restriction C_0^{\infty}(\mathbb{R}^d \setminus \{y_1, y_2\})$ are two dimensional and are respectively the linear span of the solutions of fundamental equation $(-\Delta \mp i)G^j_{\pm i} = \delta_{y_j} \quad j = 1, 2$ which are the only eigenfunctions of the adjoint of $\dot{H}_0$ relative to the eigenvalues $\pm i$. Explicitly

$$G^j_{\pm i} = \frac{e^{i\sqrt{\mp i}|x-y_j|}}{4\pi|x-y_j|}\,, \qquad j=1,2$$

with $\mathrm{Re}\,\sqrt{\mp i} > 0$. Due to the assumption of symmetry with respect to the exchange of the positions y_1 and y_2 the only admissible states in the ith defect space are complex multiples of the sum $G^1_{\pm i} + G^2_{\pm i}$. The dimensions of the defect spaces are then reduced to one. As a consequence, according with the von Neumann construction any extension H_U of $\dot{H}_0$ will have domain and action given by

$$\begin{aligned} D(H_U) &= \left\{ f \in L^2(\mathbb{R}^3) | f = f_0 + C\sum_{k=1}^{2} G^k_i + C\sum_{k,j=1}^{2} U_{kj} G^j_{-i} \right\} \\ H_U f &= -\Delta f_0 + i\,C\sum_{k=1}^{2} G^k_i - i\,C\sum_{k,j=1}^{2} U_{kj} G^j_{-i} \end{aligned} \tag{3.1}$$

where $C \in \mathbb{C}$, $f_0 \in H^2(\mathbb{R}^3)$ is symmetric under the exchange of the interaction centers positions and the 2×2 matrix U must have the form

$$U = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} \qquad \text{where} \qquad 0 \leqslant \theta < 2\pi\,. \tag{3.2}$$

We will collect the main results concerning the properties of the entire family of self-adjoint extensions of $\dot{H}_0$ in the following proposition.

Proposition 3.1 *Each H_U, with U of the form (3.2), has the following properties.*
(i) The resolvent of H_U is

$$(H_U - z)^{-1} = G_z + \sum_{m,n=1}^{2} [\Gamma_U(z)]^{-1}_{mn} \left(\overline{G_z(\cdot - y_m)}\,, \cdot \right) G_z(\cdot - y_n) \tag{3.3}$$

with

$$\Gamma_U(z) = 2i \begin{pmatrix} s & S \\ S & s \end{pmatrix} \begin{pmatrix} \frac{e^{-i\theta/2}}{2\cos\frac{\theta}{2}} & 0 \\ 0 & \frac{e^{-i\theta/2}}{2\cos\frac{\theta}{2}} \end{pmatrix} + \begin{pmatrix} c_z & C_z \\ C_z & c_z \end{pmatrix} \tag{3.4}$$

where,

$$s = \frac{1}{4\sqrt{2}\pi} \qquad S = \frac{1}{4\pi R} e^{-R/\sqrt{2}} \sin(R/\sqrt{2}) \qquad c_z = \frac{\sqrt{-z} - \sqrt{i}}{4\pi} \qquad C_z = -\frac{e^{-\sqrt{-z}R} - e^{-\sqrt{i}R}}{4\pi R} \tag{3.5}$$

and $R = |y_1 - y_2|$. Notice that for $\theta = \pi$ the matrix $\Gamma_U(z)$ becomes infinite and the resolvent of H_U converges for $\theta \to \pi$ to the resolvent of the free Laplacian.

(ii) Every function $\Psi \in D(H_U)$ has the following behaviour around each interaction center

$$\Psi(x \sim y_i) = \frac{q}{4\pi|x - y_i|} + \alpha(R)\, q + o(1) \tag{3.6}$$

where

$$\alpha(R) = -\frac{1 - t_\theta}{4\sqrt{2}\pi} + \frac{1}{4\pi R} e^{-R/\sqrt{2}} \cos\left(R/\sqrt{2}\right) + \frac{t_\theta}{4\pi R} e^{-R/\sqrt{2}} \sin(R/\sqrt{2}) \tag{3.7}$$

and

$$t_\theta = \tan\frac{\theta}{2}\,. \tag{3.8}$$

(iii) The essential spectrum of H_U is

$$\sigma_{ess}(H_U) = \sigma_{ac}(H_U) = [0, \infty) \tag{3.9}$$

and it is absolutely continuous. The discrete spectrum of H_U consists of at most one eigenvalue in $(-\infty, 0)$. More precisely, $-\lambda$ for $\lambda > 0$ is an eigenvalue of H_U if and only if

$$4\pi\alpha(R)\frac{R}{\sqrt{2}} = -\sqrt{\lambda}R + e^{-\sqrt{\lambda}R} \tag{3.10}$$

(iv) The functions

$$\Psi_{\alpha(R),Y}(k, x) = e^{ik\cdot x} + \sum_{k,l=1}^{2} [\Gamma_U(k^2)]^{-1}_{mn}\, e^{ik\cdot y_n} \frac{e^{i|k||x-y_m|}}{4\pi\,|x - y_m|}\,, \qquad k \in \mathbb{R}^3 \tag{3.11}$$

make up the set of all the generalized eigenfunctions of H_U.

(v) The scattering length of H_U is

$$a_U = -\frac{1}{4\pi}\sum_{m,n=1}^{2} [\Gamma_U(0)]^{-1}_{mn} = \frac{2R}{\frac{R}{\sqrt{2}}(1 - t_\theta) - e^{-\frac{R}{\sqrt{2}}}(t_\theta \sin\frac{R}{\sqrt{2}} + \cos\frac{R}{\sqrt{2}}) + 1} = -2\left(4\pi\alpha(R) - \frac{1}{R}\right)^{-1}. \tag{3.12}$$

The scattering length for $R = 0$ becomes $a_U = \frac{\sqrt{2}}{1-t_\theta}$ while for $R \to \infty$ one has $a_U \to \frac{2\sqrt{2}}{1-t_\theta}$.

Below, we make few remarks in order to highlight the main consequences of the proposition stated above.

Remark 3.2 For each fixed θ the function $\alpha(R)$ appearing in the "boundary condition" around each interaction point becomes singular when $R \to 0$. In fact, $\alpha(R)$ has exactly the same form of the renormalisation term introduced in the previous section to obtain a point interaction Hamiltonian bounded from below in the three-body problem.

Remark 3.3 Inverting the relation between $\alpha(R)$ and t_θ one obtains

$$t_\theta(R) = \frac{4\pi R\alpha - e^{-R/\sqrt{2}}\cos R/\sqrt{2} + \frac{R}{\sqrt{2}}}{\frac{R}{\sqrt{2}} + e^{-R/\sqrt{2}}\sin R/\sqrt{2}}\,. \tag{3.13}$$

Equation (3.13) shows that for each fixed $\alpha \in \mathbb{R}$ and distance $R > 0$ there is a t_θ (and in turn a self-adjoint extension) but it also shows that $t_\theta \to -\infty$ when the distance of the two interaction centers approaches zero implying that the Hamiltonian tends to the free one. We recall that the self-adjoint extensions characterized by a constant α have been historically referred to as the "local" ones and always considered the only acceptable models of point interaction Hamiltonians in Quantum Mechanics. As is evident from what has been presented above the multi-center local self-adjoint extensions of $\dot{H}_0$ are not additive as opposed to what happens for the sum of two short range potential functions which is always well defined regardless of how much their supports overlap (in particular, they do not disappear when the interaction centers tend to coincide).
In contrast, all the extensions relative to a fixed t_θ do not become trivial when the distance between the interaction points approaches zero. In fact, they converge to a one center point interaction Hamiltonian with scattering length different from zero and depending on the extension parameter t_θ.

Remark 3.4 From (3.12) we deduce that the scattering length of H_U increases to infinity when the distance of the interaction centers approaches zero if U corresponds to the parameter $t_\theta = 1$. This implies that $t_\theta = 1$ characterizes the self-adjoint extension showing scale invariance when the two scattering centers tend to coincide.

Remark 3.5 The resolvent of the two centers Hamiltonian H_U for $y_2 \to y_1 \equiv y$ tends to the operator

$$G_z + \frac{1}{\frac{t_\theta - 1}{4\pi\sqrt{2}} + \frac{\sqrt{-z}}{4\pi}} \left(\overline{G_z(\cdot - y)}, \cdot\right) G_z(\cdot - y)$$

which is the resolvent of a single point interaction Hamiltonian with "strength" $\dfrac{t_\theta - 1}{4\pi\sqrt{2}}$.

The eigenvalue problem for H_U amounts to solve Eq. (3.10). Denoted by W_0 the Lambert function (see e.g. [4]), the eigenvalue $E(R)$ is given by

$$E(R) = -\lambda\,, \qquad \text{with} \quad \sqrt{\lambda} = \frac{1}{R}(W_0(e^u) - u) \tag{3.14}$$

where

$$u = 4\pi\alpha(R)R = \left(-(1 - t_\theta)\frac{R}{\sqrt{2}} + e^{-R/\sqrt{2}}(t_\theta \sin R/\sqrt{2} + \cos R/\sqrt{2})\right). \tag{3.15}$$

For $t_\theta = 1$ it easy to check that the energy eigenvalue for $R \to 0$ and $R \to \infty$ have the following behaviour

- For $R \to 0$

$$E(R) \sim -\frac{R^2}{16} \tag{3.16}$$

- For $R \to \infty$

$$E(R) \sim -\frac{W_0(1)^2}{R^2} \sim -\frac{0.32}{R^2}\,. \tag{3.17}$$

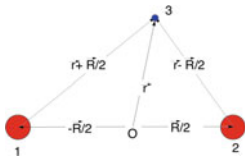
For $0 < t_\theta < 1$, the function $E(R)$ remains bounded away from zero and

$$\lim_{R\to 0} E(R) = \lim_{R\to\infty} E(R) = -\frac{(1-t_\theta)^2}{2} \tag{3.18}$$

and $E(R) \sim -\frac{(1-t_\theta)^2}{2} - \frac{\sqrt{2}(1-t_\theta)\,e^{-(1-t_\theta)R/\sqrt{2}}}{R} - \frac{e^{-2(1-t_\theta)R/\sqrt{2}}}{R^2}$ for $R \to \infty$.

4 Two Heavy and One Light Particle—The Born Oppenheimer Approximation

Following the suggestions given at the end of Sect. 2 we want to further investigate the dynamics of a three-body system in the limit in which it is possible to separate the slow dynamics of two heavy non interacting bosons and the fast dynamics of a light particle interacting with the two bosons via zero-range forces.



The Hamiltonian in the Jacobi coordinates formally reads (see e.g. [13])

$$H = -\frac{1}{\mu}\triangle_{\vec{R}} - \frac{1}{\nu}\triangle_{\vec{r}} + \delta(\vec{r} + \vec{R}/2) + \delta(\vec{r} - \vec{R}/2)$$

where

$$\vec{r} = \vec{r}_3 - \frac{1}{2}(\vec{r}_1 + \vec{r}_2), \quad \vec{R} = \vec{r}_1 - \vec{r}_2, \quad \nu = \frac{2M}{2M+m}, \quad \mu = \frac{M}{2m}\,.$$

For this purpose, we are going to use a bottom up strategy based on the Born-Oppenheimer approximation using only what we learnt about point interaction Hamiltonians.

In fact, formula (2.31) suggests that the fast dynamics of the light particle is generated by a two center point interaction Hamiltonian of the type we analized in the previous section.

The mutual interaction between the two bosons acquired as a consequence of their common interaction with the light one is examined through the Born-Oppenheimer approximation. In this approximation the analysis of the eigenvalue problem for the three body system is performed assuming eigenfunctions of the form

$$\Psi(\vec{r}, \vec{R}) = \psi(\vec{r}; \vec{R})\Phi(\vec{R})$$

where $\psi(\vec{r}; \vec{R})$ is the solution of the time independent Schrödinger equation for the light particle depending parametrically on $\vec{R}$

$$\left[-\frac{1}{\nu}\Delta_{\vec{r}} + \delta(\vec{r} + \vec{R}/2) + \delta(\vec{r} - \vec{R}/2)\right]\psi(\vec{r}; \vec{R}) = E(R)\psi(\vec{r}; \vec{R}) \tag{4.1}$$

and

$$\left[-\frac{1}{\mu}\Delta_{\vec{R}} + E(R)\right]\Phi(\vec{R}) = \mathcal{E}\Phi(\vec{R}) \tag{4.2}$$

where $R = |\vec{R}|$ and $\mathcal{E}$ is the approximate eigenvalue of the three-body system.

Remark 4.1 If the function $E(R)$ is computed for $\alpha = 0$ (see Eq. (3.10)) one finds that $E(R) \sim -W_0(1)^2/\nu R^2$ both for R near zero and for R tending to infinity. This turns out to imply that the system of the two bosons is unstable for the presence of energy eigenvalues going to $-\infty$. In fact all the self-adjoint realizations of a Schrödinger operators with potential $-\gamma/R^2$, $\gamma > 1/4$, share this pathology. It is difficult not to notice the similarities with the instability problems discussed in Sect. 2 relative to boundary conditions with fixed scattering length.

Using the two-center point interaction Hamiltonian described in the previous section with $t_\theta = 1$ the effective potential $E(R)$ turns out to be regular, bounded everywhere and decaying as $-W_0(1)^2/\nu R^2$ at infinity. In the case $W_0(1)^2/\nu > 1/4$ the energy eigenvalues are bounded from below and there are infinitely many low energy eigenstates with eigenvalues accumulating at zero energy. They can be shown to satisfy the "Efimov scale". All the results relative to problem (4.2) are stated in the following theorem, whose proof will be given in a forthcoming paper ([11]).

Theorem 4.2 *Let $t_\theta = 1$. There exists an infinite sequence of negative eigenvalues $\mathcal{E}_n$ with $\mathcal{E}_n \to 0$ for $n \to \infty$. Moreover,*

$$\mathcal{E}_n = -C\, e^{\frac{2}{\beta}\left(\tan^{-1}(2\beta) + \phi_\beta - n\pi\right)}\left(1 + \zeta_n\right)$$

where $C > 0$, $\zeta_n \to 0$, *for* $n \to \infty$ *and* $\beta = \sqrt{\mu W(1)^2/\nu - 1/4}$. *The Efimov geometrical law holds*

$$\lim_n \frac{E_n^i}{E_{n+1}^i} = e^{\frac{2\pi}{\beta}}.$$

Let $t_\theta < 1$. *The number of eigenvalues is finite.*

Dedication. Like all those in this volume our contribution is dedicated to the memory of Detlef. He would probably appreciate it but he certainly wouldn't refrain from joking with us with "aren't you tired of being involved in this point interaction business?" Sure Detlef, sometimes we are a little bit tired but we can't miss your sweet and irrepressible smile when you play with us. Thanks Detlef for having been our friend.

References

1. S. Albeverio, F. Gesztesy, R. Høegh-Krohn, H. Holden, *Solvable Models in Quantum Mechanics*, 2nd edn. (AMS Chelsea Publishing, Providence, RI, 2005). p. 2105735. With an appendix by Pavel Exner
2. S. Albeverio, R. Hoegh-Krohn, T.T. Wu, A class of exactly solvable three-body quantum mechanical problems and the universal low energy behaviour. Phys. Lett. A **83**(3), 105–109 (1981)
3. G. Basti, C. Cacciapuoti, D. Finco, A. Teta, Three-body Hamiltonian with regularized zero-range interactions in dimension three. Ann. H. Poincaré **24**, 223–276 (2023)
4. R. Corless, G. Gonnet, D. Hare, D. Jeffrey, D. Knuth, On the Lambert W function. Adv. Comput. Math. 5:329–359 (1996)
5. L. Dabrowski, H. Grosse, On nonlocal point interactions in one, two, and three dimensions. J. Math. Phys. **26**(11), 2777–2780 (1985)
6. T.M. Dunster, Bessel functions of purely imaginary order, with an application to second-order linear differential equations having a large parameter. SIAM J. Math. Anal. **21**(4), 995–1018 (1990)
7. D. Dürr, R. Figari, A. Teta, Decoherence in a two-particle model. J. Math. Phys. **45**(4), 1291–1309 (2004)
8. D. Dürr, H. Spohn, Decoherence through coupling to the radiation field, in *Decoherence: Theoretical, Experimental and Conceptual Problems*, ed. by P. Blanchard, D. Giulini, E. Joos, C. Kiefer, I.O. Stamatescu, Lecture Notes in Physics, vol. 538 (Springer, 2000), pp. 77–86
9. D. Ferretti, A. Teta, Regularized zero-range Hamiltonian for a Bose gas with an impurity. Complex Anal. Oper. Theory **17**, 55 (2023)
10. D. Ferretti, A. Teta, Some remarks on the regularized Hamiltonian for three bosons with contact interactions. arXiv:2207.00313v1 [math-ph]. Quantum Mathematics I, ed. by M. Correggi, M. Falconi, Springer INdAM Series vol. 57, 2023
11. R. Figari, H. Saberbaghi, A. Teta, A new look at the theory of point interactions. arXiv:2306.10292
12. F. Figari, A. Teta, On the Hamiltonian for three bosons with point interactions, in *Quantum and stochastic mathematical physics. Sergio Albeverio, Adventures of a mathematician*, ed. by A. Hilbert, E. Mastrogiacomo, S. Mazzucchi, B. Rudiger, S. Ugolini, Springer proceedings in mathematics statistics, vol. 377 (Springer, 2023), pp. 141–159
13. A.C. Fonseca, E.F. Redish, P.E. Shanley, Efimov effect in an analytically solvable model. Nucl. Phys. A **320**, 273–288 (1979)

14. E. Joos, C. Kiefer, D. Giulini, J. Kupsch, I.O. Stamatescu, H.D. Zeh, *Decoherence and the Appearance of a Classical World in Quantum Theory*, 1st edn. (Springer, 1996), 2nd edn. 2003
15. K. Hornberger, S. Uttenthaler, B. Brezger, L. Hackermueller, M. Arndt, A. Zeilinger, Collisional decoherence observed in matter wave interferometry. Phys. Rev. Lett. **90** (2003)
16. A. Michelangeli, Models of zero-range interaction for the bosonic trimer at unitarity. Rev. Math. Phys. **33**, 2150010 (2021)
17. R.A. Minlos, L. Faddeev, On the point interaction for a three-particle system in Quantum Mechanics. Sov. Phys. Dokl. **6**(12), 1072–1074 (1962)
18. R.A. Minlos, L. Faddeev, Comment on the problem of three particles with point interactions. Sov. Phys. Jetp. **14**(6), 1315–1316 (1962)
19. M. Schechter, *Operator Methods in Quantum Mechanics* (Dover, 1981)
20. M. Reed, B. Simon, *Methods of Modern Mathematical Physics: Analysis of Operators*, vol. 4 (Academic, New York, 1978)

Philosophy of Physics

Situated Observation and the Quantum Measurement Problem

Jeffrey Barrett

The merit of Bohmian mechanics is not determinism, but the refutation of all claims that quantum mechanics cannot be reconciled with a realistic description of reality. In physics, one needs to know what is going on. Bohmian mechanics tells us what is going on and it does so in the most straightforward way imaginable.

Teufel and Dürr (2009, 9–10)

Abstract A situated observer is an observer as modeled within the world characterized by one's physical theory. A physical theory arguably only makes empirical predictions if it makes predictions for the records of a situated observer. In this spirit, one has a satisfactory solution to the measurement problem only if one has a formulation of quantum mechanics that makes the right empirical predictions for the records of a situated observer. Bohmian mechanics addresses the measurement problem by explaining what measurement records are, how a situated observer might produce them, and why a situated observer should expect the standard quantum predictions. The notion of an effective wave function is crucial in understanding how the theory makes empirical predictions.

1 The Measurement Problem

Our best physical theories should describe both us and the world we inhabit. After all, we are physical systems like any other, and the outcomes of our measurements are physical records. As John Bell put the point, "For me it is absolutely unquestioned in my analysis of things that the real world is out there, and that I am an incident in it" (Bell 1989).

J. Barrett (✉)
Logic and philosophy of science, University of California Irvine, Irvine, CA 92697-5100, USA
e-mail: j.barrett@uci.edu

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_25

In order to be consistently applicable to all systems, a physical theory must be able to explain how the process of observation leads to determinate physical records. And for it to be empirically adequate, it must accurately predict our records. A physical theory is empirically adequate in this sense only if it accurately predicts the physical records of situated observers, observers as modeled within the world characterized by the theory.[1]

If a physical theory exhibits this strong sort of empirical adequacy, it explains why the world looks the way it does to the observers who inhabit it. We shall say that such a theory is *strongly adequate*.[2] A strongly adequate theory successfully predicts and explains our physical records.

There are clear virtues to having a strongly adequate theory. Unfortunately, the standard von Neumann formulation of quantum mechanics is not strongly adequate. The quantum measurement problem arises from trying to describe the process of measurement for a physically situated observer in the theory.

Hugh Everett III presented the measurement problem in the form of a "hypothetical drama" in his Ph.D. thesis (1956, 73–5). He argued that the standard formulation of quantum mechanics is "untenable" because it does not allow for a consistent account of how a physically situated observer produces the physical records that represent the outcomes of her measurements (1956, 75). Eugene Wigner (1961) later retold Everett's story.

Consider an object system S that begins in the superposition of x-spin states represented by

$$\alpha|\uparrow_x\rangle_S + \beta|\downarrow_x\rangle_S$$

and suppose that an observer F and her measuring device M begin ready to make an x-spin measurement of S as in Fig. 1. The composite system, then, begins in the state:

$$|\text{“r”}\rangle_F|\text{“r”}\rangle_M(\alpha|\uparrow_x\rangle_S + \beta|\downarrow_x\rangle_S).$$

To keep things simple, suppose that M is a perfect measurement device and F a perfect observer. That is, suppose that M's pointer reading becomes perfectly correlated with the x-spin of S so that $|\text{“r”}\rangle_M|\uparrow_x\rangle_S$ would evolve to $|\text{“}\uparrow_x\text{”}\rangle_M|\uparrow_x\rangle_S$ and $|\text{“r”}\rangle_M|\downarrow_x\rangle_S$ would evolve to $|\text{“}\downarrow_x\text{”}\rangle_M|\downarrow_x\rangle_S$ when M interacts with S. And suppose that F's measurement record becomes perfectly correlated with M's pointer so that $|\text{“r”}\rangle_F|\text{“}\uparrow_x\text{”}\rangle_M|\uparrow_x\rangle_S$ would evolve to $|\text{“}\uparrow_x\text{”}\rangle_F|\text{“}\uparrow_x\text{”}\rangle_M|\uparrow_x\rangle_S$ and $|\text{“r”}\rangle_F|\text{“}\downarrow_x\text{”}\rangle_M|\downarrow_x\rangle_S$ would evolve to $|\text{“}\downarrow_x\text{”}\rangle_F|\text{“}\downarrow_x\text{”}\rangle_M|\downarrow_x\rangle_S$ when F interacts with M.

Assuming such perfect correlating interactions, if we consider the situated observer F to be a physical system like any other, the resultant state of the

[1] The idea of a situated observer is, as we shall see, one that has played an important role in the conceptual foundations of quantum mechanics for at least the last seventy years. Jenann Ismael (2007 and 2016) provides examples of how one might appeal to the notion for philosophical purposes.

[2] See Barrett (2020) and (2021) for a characterization of various sorts of empirical adequacy one might want a formulation of quantum mechanics to exhibit.

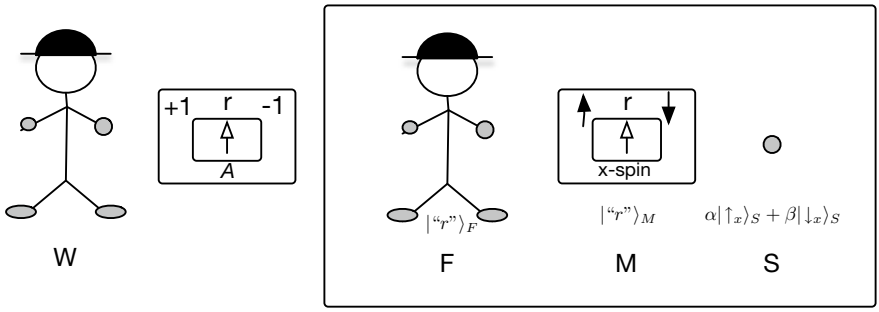


Fig. 1 Everett's hypothetical drama

composite system on the standard linear (unitary) dynamics after she looks at the pointer will be:

$$\alpha|\text{“}\uparrow_x\text{”}\rangle_F|\text{“}\uparrow_x\text{”}\rangle_M|\uparrow_x\rangle_S + \beta|\text{“}\downarrow_x\text{”}\rangle_F|\text{“}\downarrow_x\text{”}\rangle_M|\downarrow_x\rangle_S. \qquad (\Diamond)$$

The state ($\Diamond$) describes an entangled superposition of two record states, one where F has the physical record "$\uparrow_x$" and one where she has the physical record "$\downarrow_x$". The problem is that this state does not characterize her as having any determinate record at all on the standard interpretation of states. And if there is no record, the theory clearly fails to predict her measurement result with the standard quantum probabilities. So if F is treated as a situated observer subject to the linear dynamics, then on the standard interpretation of states, one fails to get the right empirical predictions for the simple reason that F fails to have any determinate physical record.

The standard formulation of quantum mechanics seeks to address the problem by stipulating that either M's interaction with S or F's interaction with $M + S$ causes a collapse of the composite system so that the final state is either $|\text{“}\uparrow_x\text{”}\rangle_F|\text{“}\uparrow_x\text{”}\rangle_M|\uparrow_x\rangle_S$ with probability $|\alpha|^2$, a state where F has the determinate record "$\uparrow_x$", or $|\text{“}\downarrow_x\text{”}\rangle_F|\text{“}\downarrow_x\text{”}\rangle_M|\downarrow_x\rangle_S$ with probability $|\beta|^2$, a state where F has the determinate record "$\downarrow_x$". Each of these states is logically inconsistent with state ($\Diamond$), the state that would obtain if F is a physically situated observer, but the problem is yet worse. While von Neumann thought that he had shown that it did not matter empirically whether it was M's interaction with S or F's interaction with $M + S$ that caused the collapse, as Wigner emphasized in his retelling of the story, precisely when the collapse occurs has empirical consequences.

Suppose that the external observer W measures an observable A of the composite system $F + M + S$ that has state ($\Diamond$) as a eigenstate with eigenvalue $+1$ and a state orthogonal to ($\Diamond$) and in the subspace spanned by the same two determinate record states as an eigenstate with eigenvalue -1. If the situated observer F and other systems in the box obey the standard linear dynamics, then the composite system $F + M + S$ will end up in state ($\Diamond$), and the result of W's A-measurement will be $+1$ with probability 1. But if there is a collapse of the state, resulting in one of the two

possible determinate record states $|\text{“}\uparrow_x\text{”}\rangle_F|\text{“}\uparrow_x\text{”}\rangle_M|\uparrow_x\rangle_S$ or $|\text{“}\downarrow_x\text{”}\rangle_F|\text{“}\downarrow_x\text{”}\rangle_M|\downarrow_x\rangle_S$, then the result of W's A-measurement will have a positive probability of resulting in -1.

Wigner took this to mean that one might, at least in principle, perform an experiment that determines what interactions cause collapses. He conjectured that one would find that a collapse of the physical state occurs when a conscious observer, an observer associated with a non-physical mind, apprehends the state of a physical system. With this, Wigner gave up on treating F as a physically situated observer.

In contrast, Everett held that observers are physical systems like any other and that all physical systems are subject to the standard linear dynamics. He believed that a formulation of quantum mechanics is satisfactory only if it makes the right empirical predictions for the records of situated observers. He modeled such observers as "automata" capable of correlating the states of their records with the properties of the physical systems they observe (1956, 118–9 and 1957, 183–5). In the context of his hypothetical drama, Everett believed that F recorded both "$\uparrow_x$" and "$\downarrow_x$" as relative records. And he was committed to W's A-measurement of $F + M + S$ resulting in $+1$ with probability 1. More generally, he sought to show that the physical records of a typical relative observer will exhibit the standard quantum statistics in his formulation of pure wave mechanics. While there is much to say about the sense in which pure wave mechanics explains the records of situated observers, we are concerned here with another approach to situated observation.[3]

Everett knew that David Bohm had already described a no-collapse formulation of quantum mechanics that made predictions for the records of situated observers. He discussed Bohmian mechanics in two places in the long version of his dissertation (1956, 75, 153–5). Everett's complaint was that Bohm's theory was "more cumbersome than the conceptually simpler theory based on pure wave mechanics." That said, he took Bohmian mechanics to be of "great theoretical importance" because it "showed that 'hidden variable' theories are indeed possible" contrary to received wisdom. Concerning hidden-variable theories like Bohm's, Everett further acknowledged that "[i]t cannot be disputed that these theories are often appealing and might conceivably become important should future discoveries indicate serious inadequacies in the present scheme [pure wave mechanics] (i.e. they might be more easily modified to encompass new experience)" (1956, 155). While Everett thought that he could adequately explain (or better, explain away) quantum probabilities, given the subsequent difficulty in understanding forward-looking quantum probabilities in pure wave mechanics, that Bohmian mechanics treats quantum probability in an entirely straightforward way is a manifest virtue, one that makes the approach of central importance to the foundations of quantum mechanics.

The measurement problem is a problem in providing a consistent account of situated observation, something to which Bohmian mechanics is particularly well suited. Supposing that F, M, and S are all subject to the linear dynamics is incompatible

[3] See Saunders, Barrett, Kent, and Wallace (2010) and Barrett (2020) for discussions of the challenges Everett's approach faces and how they might be addressed. Wallace's (2012) representation-theorem is the best worked-out proposal.

with supposing that F ends up with one or the other determinate physical record on the standard interpretation of states. Hence, in order to understand F as a physically situated observer either (i) the linearly dynamics at least sometimes fails to describe how the state of a physical system evolves (collapse theories like GRW opt for this[4]), (ii) there is more than one situated observer described by state ($\Diamond$) (Everett opts for this in his relative-state formulation of pure wave mechanics[5]), or (iii) state ($\Diamond$) is incomplete inasmuch as it fails to specify the value of F's record.[6] On Bohmian mechanics, state ($\Diamond$) provides an incomplete description of F. In order to specify the value of F's record, one must also give the determinate particle configuration.[7] As a result, the theory explains why F has a perfectly determinate measurement record and how she got it, and it predicts that W's A-measurement will result in $+1$ with probability 1.

Being able to tell Everett's hypothetical drama consistently provides a litmus test for whether one's formulation of quantum mechanics provides a coherent account of situated observation. The thought is that a theory is a serious contender for resolving the measurement problem only if it can do so. Bohmian mechanics addresses the measurement problem by explaining what measurement records are, how a situated observer might produce them, and why a situated observer should expect the standard quantum predictions. The notion of an effective wave function is crucial to understanding how each step works.

2 Bohmian Mechanics

Bohmian mechanics can be characterized as follows[8]:

1. *representation of states*: The complete physical state of a system S at time t is given by the wave function $\psi(q,t)$ over configuration space and a point in configuration space $Q(t)$.
2. *interpretation of states*: The position of every particle is always determinate and is given by the current configuration $Q(t)$.

3I. *linear dynamics*: The wave function evolves in the standard unitary way. In the simplest case

$$i\hbar\frac{\partial\psi(q,t)}{\partial t} = \hat{H}\psi(q,t)$$

[4] See Ghirardi, Rimini, and Weber (1986).

[5] See Everett (1956) and (1957).

[6] See Everett (1956) and Albert (1992) for similar classifications of options.

[7] See Bohm (1952).

[8] This is Bell's (1987) formulation of the theory. Bell also showed how to make local field properties determinate instead of particle positions. See Vink (1993) for a description of how one might similarly make any physical observable determinate and Barrett (2005) for a discussion.

3II. *particle dynamics*: Particles move according to

$$\frac{dQ_k(t)}{dt} = \frac{1}{m_k} \frac{\mathrm{Im}\ \psi^*(q,t)\nabla_k\psi(q,t)}{\psi^*(q,t)\psi(q,t)}\Big|_{Q(t)}$$

where m_k is the mass of particle k and $Q(t)$ is the current configuration.

4. *distribution postulate*: The probability density of the configuration $Q(t_0)$ is $|\psi(q,t_0)|^2$ at an initial time t_0.

One gets empirical predictions for a situated observer by supposing that an observer's experience supervenes on her measurement records as characterized by the effective wave function. The notion of the effective wave function of a system was introduced by Dürr, Goldstein, and Zanghì (DGZ) (1992).[9] We will consider how this notion helps to explicate the process of situated observation in the theory.

In Bohmian mechanics, both the wave function $\psi(q,t)$ and the particle configuration $Q(t)$ evolve in $3N$-dimensional configuration space, where N is the number of particles in the system one wishes to describe. The $3N$-coordinates of the configuration $Q(t)$ give the position of each particle at time t. One can think of the probability density $|\psi(q,t)|^2$ as describing the density of a compressible fluid in configuration space. The wave function evolves deterministically according to the linear dynamics (rule 3I), and as the compressible fluid flows about in configuration space, it carries the point representing the particle configuration $Q(t)$ along as described by the particle dynamics (rule 3II). The configuration moves as if it were a massless particle carried by the probability current.

While the theory characterizes the dynamics for the entire universe, one can often assign a wave function to a proper subsystem. We will say that a physical system has an *effective wave function* at a time if its configuration can be understood as moving in accord with a wave function describing just that system. If so, one can give the dynamics for both the wave function and the system's configuration in a subspace of the full configuration space. Dürr, Goldstein, and Zanghì (1992, Sect. 5) specify sufficient conditions for a physical system having a well-defined effective wave function in terms of the properties of that system and its complement. In the case of real physical systems, satisfying such conditions is typically a matter of degree. Importantly, one can expect the DGZ conditions to hold well for a post-measurement object system if different measurement outcomes correspond to macroscopically distinct pointer positions. That the object system has a well-defined effective wave function in this case explains why it behaves as if a collapse had occurred.

We will consider in some detail situations where different measurement outcomes correspond to sharply individuated regions of wave function support in configuration space. In such cases, the configuration will select the component of the wave function associated with one of these regions as the effective wave function characterizing

[9] DGZ also wanted to explain why the distribution postulate (rule 4) is satisfied. Here we shall simply assume that it is satisfied as a boundary condition stipulated by the theory. This way of specifying the theory has, I believe, the virtues clarity, precision, and simplicity. Of course, it remains open for one to argue for the plausibility of the rule on dynamical grounds.

both the effective state of the observer's object system and her measurement record. Under physical conditions where her measurement record is stable, she will be able to make reliable inferences regarding the subsequent behavior of the object system from the value of her record.

Quantum probabilities in Bohmian mechanics are purely epistemic. They result from one not knowing the initial particle configuration. The dynamics has the property that if the epistemic probability density for the particle configuration is ever given by the standard epistemic quantum probabilities $|\psi(q,t)|^2$, then it will continue to be until one makes a measurement. After a measurement, it will be given by the effective wave function characterizing one's object system and measurement record. The distribution postulate (rule 4) stipulates that the epistemic probability density for the particle configuration at time t_0 is $|\psi(q,t_0)|^2$. For our purposes, we will take the justification for this constraint to be empirical. If one sets one's priors in accord with the distribution postulate, then the theory predicts the standard quantum probabilities for particle configurations. But if one fails to set one's priors in this way, the posterior probabilities one gets by conditioning on one's measurement records yield demonstrably wrong expectations.

If the distribution postulate is satisfied, particle positions will in fact be distributed relative to the effective wave function precisely as quantum mechanics predicts with probability 1. But a physically situated observer never has epistemic access to these precise positions. If she did, she would be able to use the theory's deterministic dynamics to predict the results of her future observations more precisely than allowed by the standard quantum probabilities.

An observer's records are determined by the effective wave function, the wave function selected by the particle configuration, not by particle positions themselves. The theory is empirically adequate because it predicts the right effective wave functions and hence the right records for situated observers. A simple example will illustrate how the theory predicts a situated observer's measurement records.

3 Measurement Records in Bohmian Mechanics

Suppose that an observer wants to measure the x-spin of an electron e that is initially in an eigenstate of z-spin up. To this end, she sends the electron through an appropriately-oriented inhomogeneous magnetic field then correlates the position of a recording particle p with the direction e was deflected. The interaction between e and p is arranged so that the recording particle p shifts from position a to position b if and only if the electron e travels path B as in Fig. 2. One might think of the state of recording particle p as representing the situated observer's physical record, but precisely how this works is subtle.

To say how the state of the system evolves, we need to say how the wave function evolves in configuration space and how the particles move in response to the resulting probability current. We shall consider what this looks like in both ordinary 3-dimensional space and in configuration space.

Given the specified interaction between e and p, the two-particle wave function evolves as follows as the particles interact:

$$
\begin{aligned}
&|\uparrow_z\rangle_e|0\rangle_e|a\rangle_p = \\
&\frac{1}{\sqrt{2}}|\uparrow_x\rangle_e|0\rangle_e|a\rangle_p + \frac{1}{\sqrt{2}}|\downarrow_x\rangle_e|0\rangle_e|a\rangle_p \\
&\qquad\qquad \downarrow \\
&\frac{1}{\sqrt{2}}|\uparrow_x\rangle_e|A\rangle_e|a\rangle_p + \frac{1}{\sqrt{2}}|\downarrow_x\rangle_e|B\rangle_e|a\rangle_p \\
&\qquad\qquad \downarrow \\
&\frac{1}{\sqrt{2}}|\uparrow_x\rangle_e|A\rangle_e|a\rangle_p + \frac{1}{\sqrt{2}}|\downarrow_x\rangle_e|B\rangle_e|b\rangle_p \\
&\qquad\qquad \downarrow \\
&\frac{1}{\sqrt{2}}|\uparrow_x\rangle_e|0'\rangle_e|a\rangle_p + \frac{1}{\sqrt{2}}|\downarrow_x\rangle_e|0'\rangle_e|b\rangle_p
\end{aligned}
$$

The wave function evolves in $3N$-dimensional configuration space (since $N = 2$, this is a 6-dimensional space). As it does, the resulting probability current pushes the single point representing the two-particle configuration.

If the initial wave function is symmetric about the line from 0 to $0'$, then the distribution postulate entails that e is equally likely to be in the top-half and the bottom-half of the wave function. It is this feature of the theory that leads to the right quantum probabilities.

If the electron begins in the top half of the initial wave packet, then the two-particle system evolves as indicated in Fig. 2 in ordinary three-dimensional space. This is because the single point representing the positions of each of the two particles is pushed along by probability currents in configuration space as indicated in Fig. 3. The resultant particle configuration selects the $|\uparrow_x\rangle_e$ wave packet as the effective wave function. This means that e is now effectively x-spin up.

While the two x-spin wave packets overlap in ordinary 3-space, they do not overlap in configuration space. Since the motion of the particles depends on the probability current in configuration space, the two-particle configuration will remain associated with the x-spin-up wave packet as long as there is no splitting of that wave packet or interference with other components of the state in configuration space. As a result, e will move precisely as that packet moves and hence behave as if it now has the property of being determinately x-spin up.

If the electron begins in the bottom half of the initial wave packet, then the two-particle system evolves as indicated in Fig. 4 in ordinary three-dimensional space. And the single point representing the positions of each of the particles is pushed along by probability currents as indicated in Fig. 5. Here the resultant particle configuration selects the $|\downarrow_x\rangle_e$ wave packet as the effective wave function.

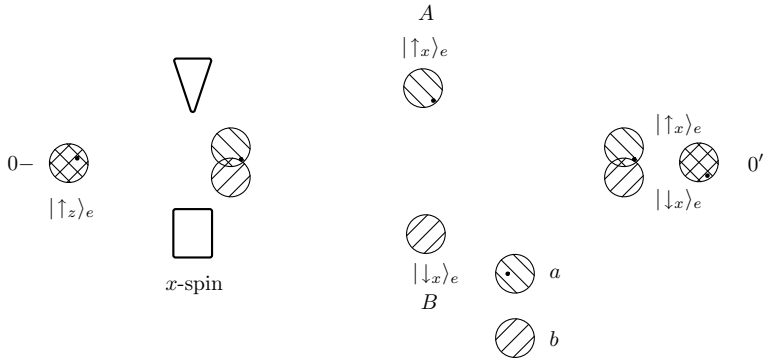


Fig. 2 Situated observation of position in three-dimensional space producing an x-spin up (path A) record

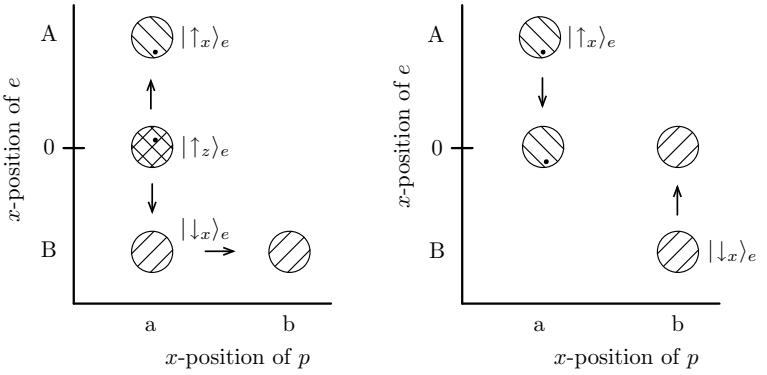


Fig. 3 Situated observation of position in configuration space producing an x-spin up (path A) record

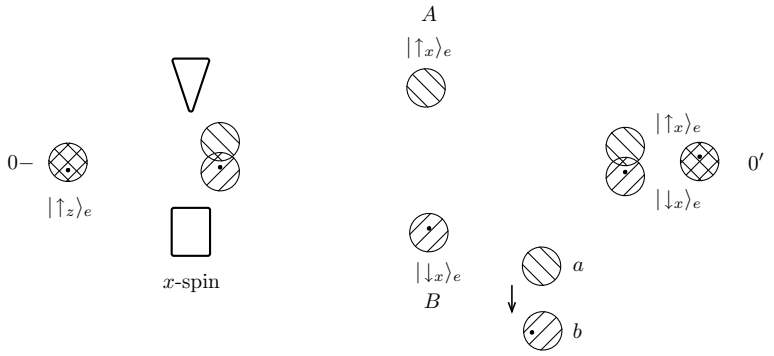


Fig. 4 Situated observation of position in three-dimensional space producing an x-spin down (path B) record

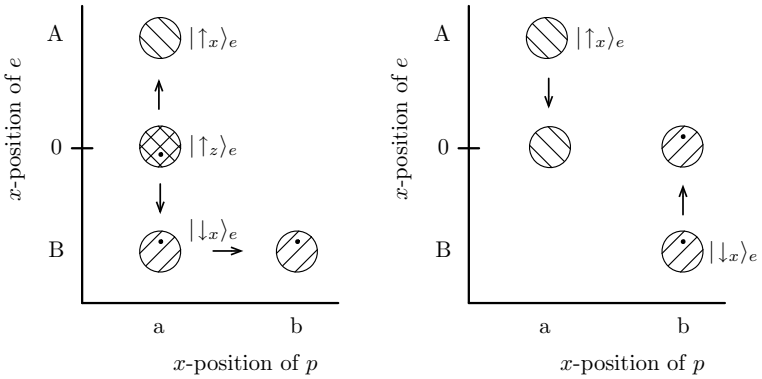


Fig. 5 Situated observation of position in configuration space producing an x-spin down (path B) record

Again, while the two x-spin wave packets end up in the same region of three-dimensional space, they do not overlap in configuration space. In this case, the configuration remains associated with the x-spin-down flavored wave packet which now serves as the effective wave function.

In each case, the two-particle configuration selects an effective wave function that reliably records the path taken by the electron. In the first case, e ends up effectively x-spin up, and it ends up effectively x-spin down in the second. It is because of the displacement generated by the position correlation between e and p that the two x-spin wave packets do not overlap in configuration space. The situated observer's record ensures that e will exhibit the same effective x-spin if one repeats the measurement. The presence of a reliable record causes e to behave as if it had an intrinsic, fully-determinate x-spin. The more degrees of freedom involved in the record, the harder it will be to get interference between wave packets corresponding to different x-spins in configuration space, and the more stable the effective wave function governing the motion of e will be.[10]

Seeing how a situated observer might record the x-spin of e allows us to specify bounds on the empirical content of her record. The observer's record reliably tells her which path e traveled. It also tells her e's effective x-spin. But her record does not tell her what e's effective x-spin was *before* the measurement as it had no effective x-spin. If the measuring apparatus had split the initial z-spin up wave packet in such a way as to capture the two-particle configuration in the x-spin-down wave packet in the first experiment, e's effective x-spin would have be x-spin down instead of x-spin up. What the observer can infer from her record is that e was in the top half of the initial z-spin up wave packet.

While the situated observer's record tells her that e was in the top half of the initial wave packet before the measurement and is now associated with the x-spin

[10] See Barrett (2000), (2020, 208–214), and (2021) for more detailed discussions this sort of two-path experiment and surreal trajectories in Bohmian mechanics.

up wave packet, it does not tell her where e is in the new x-spin up wave packet. If it did, she would be able to predict the results of future experiments more precisely than allowed by the standard quantum probabilities. So the observer's record does not provide epistemic access to the precise position of e. And while the apparatus is designed to record the position of e in the position of p, her record does not provide epistemic access to the precise position of p either. If it did, she would again be able to predict the results of future experiments more precisely than allowed by the standard quantum probabilities.

One might be tempted to say that the empirical content of the situated observer's record is the approximate positions of e and p, but the theory allows her to infer more than this from her record. Knowing that the two-particle configuration is associated with the x-spin up wave packet means that the observer can, among other things, predict what would happen on a subsequent x-spin measurement, something that is not entailed by even the precise positions of the particles. She also knows something about the probabilistic *distribution* of the particles' positions relative to the wave function. Namely, if the distribution postulate is satisfied, she can infer that they are distributed with an epistemic probability density given by the norm squared of the effective wave function.[11]

One might put the point differently by considering what the situated observer most directly *sees*. She does not see the electron e, nor can she infer precisely where it is. And she does not see the particle p or precisely where it is either, even when p is the physical system she most directly counts on to record her result. Rather, there is an important sense in which what she most directly sees is the effective wave function selected by the two-particle configuration. Specifically, her experience supervenes on the value of her physical record, and the value of her physical record supervenes on the effective wave function selected by the post-measurement configuration. That is, the observer's experience might be different only if the value of her record were different, and the value of her record might be different only if the effective wave function were different. In this regard, note that a different particle configuration might only change the value of her record, and hence experience, if it no longer selected the same effective wave function.

The role of the determinate particle configuration, then, is not to give the situated observer particles to see. Rather, it is to provide an effective wave function on which her records might supervene. The effective wave function places an upper bound on what she can infer from her records regarding the properties of the systems she has observed. Importantly, it is a consequence of the theory's dynamics that if the initial configuration satisfies the standard quantum statistics, then the physical records determined by the effective wave function will as well. Bohmian mechanics is strongly adequate because it predicts that such records will exhibit the standard quantum probabilities.

We have assumed idealized wave functions and measurement interactions that yield perfect correlations. Telling the story in this way allows one to see the basic

[11] See Sect. 5 of Dürr, Goldstein, and Zanghì's (1992) for the setup and a detailed discussion of this point.

structure of the theory. We also supposed that the situated observer knows the initial effective wave function of the system she is measuring. A situated observer is able to infer correspondingly less from her records as these conditions are weakened.

4 Discussion

Bohmian mechanics makes the standard quantum predictions for a situated observer's records as characterized by the effective wave function. And it does so even in the context of Everett's hypothetical drama.

Observer F might obtain an x-spin measurement record for S by means of a measuring device that works like the apparatus described above. The measurement both creates a determinate effective x-spin for S and produces a record of its effective x-spin. The full content of the record is given by the effective wave function. If the distribution postulate is satisfied, such records will exhibit the standard quantum statistics. The theory explains why S will behave as if it had a x-spin corresponding to the record until an incompatible measurement is made. Indeed, it is the presence of the record that makes the effective spin of S stable. If F's experience supervenes on her physical measurement records as characterized by the effective wave function, the theory accounts for her seeing determinate outcomes in accord with the standard quantum probabilities.

Regarding nested measurement, the theory also predicts that W's A-measurement will result in $+1$ with probability 1. Since the composite system $F + M + S$ ends up in state $(\Diamond)$ after F's x-spin measurement of S, and since $(\Diamond)$ is an eigenstate of the observable A with eigenvalue $+1$, when W's measuring device correlates the position of its pointer with the value of A, all of the probability will move from the region in configuration space representing the pointer being in position "r" to the region representing the pointer being in position "$+1$". As a result, if the configuration begins in a region of positive wave function support, the probability current is sure to push it to the region where the pointer records "$+1$". So if W's experience supervenes on his measurement records as characterized by the effective wave function, he will see the pointer pointing at "$+1$".

Inasmuch as Bohmian mechanics provides a consistent account of how a situated observer's measurements produce physical records that exhibit the standard quantum probabilities, it is strongly adequate. As such, Bohmian mechanics provides a compelling resolution to the measurement problem.[12]

[12] I would like to thank Roderich Tumulka, Shelly Goldstein, and an anonymous reviewer for helpful comments. This paper is in honor of Detlef Dürr, one of the kindest people I have known.

References

D.Z. Albert, *Quantum Mechanics and Experience* (Harvard University Press, Cambridge, MA, 1992)

J.A. Barrett, P. Byrne (eds.), *The Everett Interpretation of Quantum Mechanics: Collected Works 1955–1980 with Commentary* (Princeton University Press, Princeton, 2012)

J.A. Barrett, *The Quantum Mechanics of Minds and Worlds* (Oxford University Press, Oxford, 1999)

J.A. Barrett, The persistence of memory: Surreal trajectories in Bohm's theory. Philos. Sci. **67**(4), 680–703 (2000)

J.A. Barrett, The preferred basis problem and the quantum mechanics of everything. Br. J. Philos. Sci. **56**(2), 199–220 (2005)

J.A. Barrett, *The Conceptual Foundations of Quantum Mechanics* (Oxford University Press, Oxford, 2020)

J.A. Barrett, Situated observation in Bohmian mechanics. Stud. Hist. Philos. Sci. **88**, 345–357 (2021)

J.S. Bell, "Excerpts of 1989 interview of John Bell by Renée Weber," as reported by Shelly Goldstein (1989), https://sites.math.rutgers.edu/~oldstein/bell-weber.html

J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987)

D. Bohm, "A suggested interpretation of quantum theory in terms of 'Hidden Variables"', Parts I and II. Phys. Rev. **85**, 166–179, 180–193 (1952). J.A. Wheeler, W.H. Zurek (eds.), *Quantum Theory and Measurement* (Princeton University Press, Princeton, NJ, 1983), pp. 369–396

D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the origin of absolute uncertainty. J. Stat. Phys. **67**, 843–907 (1992)

H. Everett III 'Relative state' formulation of quantum mechanics. Rev. Mod. Phys. **29**, 454–462 (1957). In Barrett and Byrne (eds) (2012, 173–196)

H. Everett III, *The Theory of the Universal Wave Function* (1956). In Barrett and Byrne (eds) (2012, 72–172)

G.C. Ghirardi, A. Rimini, T. Weber, Unified dynamics for microscopic and macroscopic systems. Phys. Rev. D **34**, 470–491 (1986)

J.T. Ismael, *The Situated Self* (Oxford University Press, New York, 2007)

J.T. Ismael, *How Physics Makes Us Free* (Oxford University Press, New York, 2016)

S. Saunders, J. Barrett, A. Kent, D. Wallace (eds.), *Many Worlds?: Everett, Quantum Theory, and Reality* (Oxford University Press, Oxford, 2010)

S. Teufel, D. Dürr, *Bohmian Mechanics?: The Physics and Mathematics of Quantum Theory* (Springer, Berlin, Heidelberg, 2009)

J.C. Vink, Quantum mechanics in terms of discrete beables. Phys. Rev. A **48**, 1808–1818 (1993)

D. Wallace, *The Emergent Multiverse: Quantum Theory according to the Everett Interpretation* (Oxford University Press, Oxford, 2012)

E. Wigner, Remarks on the mind-body problem, in *The Scientist Speculates*. ed. by I.J. Good (Basic Books, New York, 1961), pp.284–302

Why the Book of Nature is Written in the Language of Mathematics

Dustin Lazarovici

Abstract The essay traces the following idea from the presocratic philosopher Heraclitus, to the Pythagoreans, to Newton's *Principia*: Laws of nature are laws of proportion for matter in motion. Proportions are expressed by numbers or, as the essay proposes, even identical to real numbers. It is argued that this view is still relevant to modern physics and helps us understand why physical laws are mathematical.

1 The "Unreasonable" Effectiveness of Mathematics

Why is mathematics so successful in describing the natural world? More profoundly, why are the fundamental laws of nature–as far as we know them today–expressed in mathematical language?

The puzzle can present itself in different ways, depending on what one takes mathematics to be. If one believes that abstract mathematical objects or structures exist in some Platonic heaven, one may wonder why they should have anything to do with the physical world and how we, as material beings in space and time, are able to acquire knowledge of them. With such questions in mind, some authors have gone as far as to suggest that the universe we live in is itself mathematical (Tegmark (2014); see also Tumulka (2017)).

If one believes that mathematics is a human invention, one must marvel at the confluence of human genius and nature's kindness that makes it so successful. One may try to deflate the "unreasonable effectiveness of mathematics" (Wigner 1960) by attributing some of it to selection bias (Wenmackers 2016), pointing to pieces of mathematics that, so far, have no use in natural science. One may also argue that our cognitive apparatus, which allowed us to invent mathematics, is the product of natural evolution and therefore well-adapted to the world (as if the traits that prevented our ancestors from being eaten by a tiger would naturally lead to the invention of complex analysis). But none of these arguments explain why the language we have been successful with is precisely that of mathematics rather than, say, biblical Hebrew or

D. Lazarovici (✉)
Humanities and Arts Department, Technion–Israel Institute of Technology, Haifa, Israel
e-mail: dustin@technion.ac.il

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_26

instructions for a Turing machine. And at the end of the day, they do little to address Wigner's sentiment that "[t]he miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve" (1960, p.14).

2 On the Rationality of the Cosmos in Presocratic Philosophy

To understand what mathematics has to do with natural laws–not just in practice, but in principle–it helps to go back to a time when the idea of a lawful cosmos awakened; the time of the Presocratic natural philosophers around the 6th and 5th century BCE. We must imagine an intellectual period marked by a profound insight: that we are living in a *cosmos* (lit. *order*), that the world is organized according to rational principles, and that the human intellect has, at least in principle, access to them. In short, it is a period animated by the idea that the world is comprehensible.

2.1 Parmenides

We have to start with Parmenides, the great ontologist, because much of the philosophy of the following centuries unfolds in the dialectic that he begins. Parmenides teaches, nay, proves, that What Is (*to eon*) must be uncreated, unchanging, and indivisible–one eternal whole:

> One path only is left for us to speak of, namely, that It is. In this path are very many tokens that what is is uncreated and indestructible; for it is complete, immovable, and without end. Nor was it ever, nor will it be; for now it is, all at once, a continuous one. For what kind of origin for it wilt thou look for? In what way and from what source could it have drawn its increase? ... I shall not let thee say nor think that it came from what is not; for it can neither be thought nor uttered that anything is not. (Poem of Parmenides; fr. 28 B8.1-13 DK)[1]

Recognizing What Is is the Way of Truth (*alêtheia*). It is not the world presented to us by our senses but something accessible by rational thought. Indeed, "it is the same thing that can be thought and that can be" (fr. 28 B3.1 DK).

Parmenides was also a great natural philosopher. "A whole series of important astronomical discoveries is credited to him: that the morning star and the evening star are one and the same; that the earth has the shape of a sphere ... that the phases of the moon are due to the changing way in which the illuminated half-sphere of the moon is seen from the earth" (Popper, 1992, p. 14). But as Popper argues, these discoveries–in particular, that the moon merely *appears* to be changing–only contribute to his mistrust of the senses. They pertain to the Way of *doxa*, of human beliefs or seemings, not true knowledge of What Is.

[1] Unless stated otherwise, Presocratic fragments are quoted in the translation by Burnet (1920).

It remains unclear how the two relate to one another. Parmenides' rationalism goes so far that, on his Way of Truth, little attempt is made to save the phenomena. Something about the holistic *Being* has to give if it is supposed to explain the cosmos we experience.

2.2 Anaxagoras

In response to Parmenides, Anaxagoras separates mind and matter, leaving a cosmic intelligence–the *Nous*–as a moving principle to act upon the material world. The *Nous* causes change and diversification by creating a cosmic vortex through which matter begins to separate into its constituent elements. *Nous* is also in *us*, as our minds that control our bodies. The implication is that we can understand the world because we share in the cosmic intelligence that shapes it. The testimony of the senses is not entirely dismissed, but its tentative character is expressed in the doctrine that "appearances are a sight of the unseen" (fr. 59 B21a DK). True knowledge requires the refinement of sense experience by rational thought.

It remains unclear how to understand the *Nous* when it comes to the subjective or individual aspects of mind, what we might call consciousness or, less anachronistically, soul (*psyche*). While the Presocratics don't always get a fair shake from Aristotle, his criticism of Anaxagoras as conflating mind and soul (*De anima* 1.2) seems pertinent.

2.3 Heraclitus

Heraclitus, "the Dark One," is very clear on one point, that he speaks about something which is *common to all* (frr. 22 B2, B80, B89, B113, B114 DK). For example:

> The waking have one common world, but the sleeping turn aside each into a world of his own. (B89 DK)

Erwin Schrödinger sees therein the idea of an external reality emerging "from the fact that part of our sensations and experiences overlap" (2014, p. 73). We can put it in a different way. While Anaxagoras separates the all-encompassing BEING of Parmenides into matter and mind, Heraclitus splits off the cognizing subject, leaving an external world as the object of cognition (cf. Dürr and Lazarovici (2012)).

Common to all is also the *logos*, the ordering and unifying principle of the world. Since recovering from the influence of Hegel, it has become widely accepted that cosmology, not logic or dialectic, is the right starting point for understanding this central concept of Heraclitean philosophy (Kurtz 1971). *Logos* does not rule some abstract realm of thought; it rules the universe we all inhabit. We may start with fragments like the following:

> This world order (*kosmos*), the same for all, none of the gods or humans made it, but it always was and is and will be: fire ever-living, kindled in measures (*métra*) and extinguished in measures. (B30 DK; translated by Laks and Most (2016))

Heraclitus is sometimes presented as the great antagonist of Parmenides because the reality he describes appears like the opposite of static being. It is a world in flux, an endless process of *becoming*, opposites united in a ceaseless cycle of transmutation. And yet, in this flow of change, Heraclitus recognizes something constant, something that manifests order and reflects the underlying *logos*. Fire, which Heraclitus takes to be the most fundamental element, transforms *in measures*, that is, in certain regular *proportions*:

> Turnings of fire: first sea; then half of the sea, earth; and the other half, lightning storm. [...] It spreads out as sea and its measure reaches the same *logos* as it was before it became earth. (B31 DK; translated by Laks and Most (2016))

If one wants to settle on a translation for "logos," the best fit here is indeed *proportion* (Kurtz 1971). Compared to the *Nous* of Anaxagoras, the Heraclitean *logos* is a more abstract and impersonal concept, coming closer to that of *natural law*.

The last quote is one of the notoriously obscure fragments of Heraclitus, whose precise meaning is hard to reconstruct. The meaning of "lightning storm" (*prêstêr*) is disputed–is it a form of fire, or a fourth element, viz. air? Also ambiguous is the subject of the second sentence and hence what transformation it describes (maybe of water back into fire; almost certainly, Heraclitus describes a kind of cycle process).[2] These issues notwithstanding, it seems clear enough that the fragment expresses a law of the form *water : earth = water : storm*, and presumably also *fire : water = water : earth*.

3 Mathematical Interlude

It may not be obvious to us today that the term "measures" already points to something mathematical. Perhaps we need a definition:

Definition 1 Two magnitudes A and B of the same kind, are *commensurable* if there exists a third magnitude ϵ and natural numbers n, m such that $A = n \cdot \epsilon$ and $B = m \cdot \epsilon$. In this case, ϵ is a *measure* of A and B and the *ratio* $A : B$ corresponds to $n : m$.

It must be emphasized that magnitudes are not numbers, but physical or geometrical quantities (lengths, areas, masses, etc.). Only the *ratio* of two commensurable magnitudes corresponds to the ratio of two numbers–or what we now recognize as a (rational) number in its own right. It is important to keep this in mind, especially when we talk about the Pythagoreans, because the above definition anticipates a fundamental motive of their science and philosophy.

[2] On these questions, see, e.g., Kurtz (1971); Jones (1972); Schadewaldt (1978); Kirk et al. (1983).

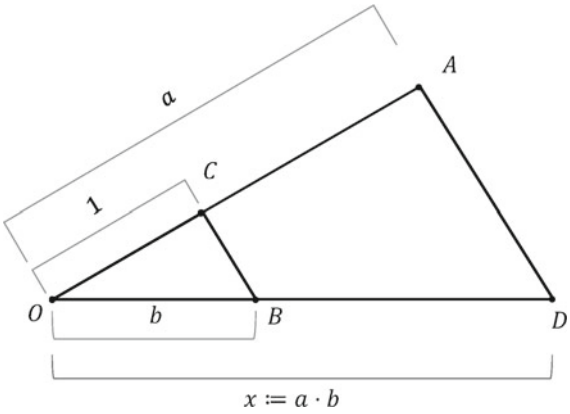


Fig. 1 Descartes' construction of the product of $a = \overline{OA}$ and $b = \overline{OB}$. The segment $\overline{OC}$ is (arbitrarily) chosen as unity. By the intercept theorem, the constructed $x = \overline{OD}$ satisfies $a : 1 = x : b$

Magnitudes by themselves have only some of the structure of numbers (see Maudlin (2014, pp. 9–25)). Two magnitudes of the same kind can be added and subtracted but they cannot be multiplied or divided (to yield a third magnitude of the same kind). A workaround, at least for line segments, is introduced much later in Descartes' *La Géométrie* (1637) and requires some arbitrary length to be designated as unity (see Fig. 1). This allowed for the very powerful algebraization of geometrical problems that paid off immediately with a precise characterization of (im)possible constructions with compass and straightedge.

But the Cartesian solution is very non-Pythagorean and indeed nonsensical from a strictly geometric point of view. It corresponds to defining the product of 2m and 3m as 6m, when we would could have just as well chosen a different unit, say cm, and multiplied the same two lengths to 60000.

It is rarely noticed that we are committing the same sin when we represent numbers as points on the number line. There is nothing numerical about a linear continuum *per se*. An arbitrary segment must be designated as a unit length, say between two points marked "0" and "1". Only relative to the scale thus introduced can we say that points (or their distances from 0) correspond to numbers.

4 Pythagoreanism and Platonism

4.1 Plato

We saw that Heraclitus, in his cosmological fragments, describes the *logos* as a law of proportion for the transformations of elements. It is this *logos* that unifies the different elements in cycles of change. While the context differs, we find the same kind of calculation in the creationist cosmogony of Plato's *Timaeus*:

> God in the beginning of creation made the body of the universe to consist of fire and earth. But two things cannot be rightly put together without a third; there must be some bond of union between them. And the fairest bond is that which makes the most complete fusion of itself and the things which it combines; and proportion is best adapted to effect such a union. [...] God placed water and air in the mean between fire and earth, and made them to have the same proportion so far as was possible (as fire is to air so is air to water, and as air is to water so is water to earth); and thus he bound and put together a visible and tangible heaven. And for these reasons, and out of such elements which are in number four, the body of the world was created, and it was harmonized by proportion [...]. (Tim. 31b–32c; translated by Jowett (1892))

Plato also makes explicit what we can only surmise for Heraclitus, that numbers (expressing proportions) are a reflection of the eternal in a world in motion:

> Now the nature of the ideal being was everlasting, but to bestow this attribute in its fullness upon a creature was impossible. Wherefore he [the creator] resolved to have a moving image of eternity, and when he set in order the heaven, he made this image eternal but moving according to number, while eternity itself rests in unity; and this image we call time. (37d)

True knowledge is knowledge of the eternal forms. In the world of change, we can only deal in likelihood. This epistemological principle is itself expressed as a law of proportion: "As being is to becoming, so is truth to belief." (*Tim.* 29c; cf. *Rep.* VII 534a). The genesis of the soul explains the possibility of knowledge. It was created out of the divisible and material on the one hand and the indivisible and unchangeable on the other, and therefore partakes of the nature of both. It is noteworthy that soul and number are ascribed a similar status as intermediates between the physical world and the realm of the eternal (cf. Plato's analogy of the divided line in *Rep.* VI 509d–511e).

4.2 The Pythagoreans

Between Heraclitus and Plato, we have the Pythagoreans, and among them a group known as the *mathēmatikoi*.[3] They developed four sciences or *mathemata*, which would come to form the classical *quadrivium* of education: arithmetic, geometry, astronomy, and music (or harmonics).

The study of musical harmony began with the observation that the simultaneous striking of different chords produces consonance when the cord lengths stand in certain ratios: 2:1 (the octave), 3:2 (the perfect fifth), 4:3 (the perfect fourth), etc. It later turned into a more axiomatic science of harmonic proportions and the musical scale. The Pythagorean astronomers recognized the same harmonic proportions in

[3] The Pythagorean influence on Plato is undeniable (the Platonic character Timaeus is commonly identified as a Pythagorean). Placing Heraclitus in the same lineage is more contentious. Plato criticizes Heraclitus on the basis that if everything were in flux, truth and knowledge would not be possible (*Cratylus* 402a ff.). Heraclitus calls Pythagoras an "imposter" (fr. B129 DK), someone who has studied many things but lacks understanding (B40 DK). Heraclitus was not an easy fellow. Nonetheless, a reconciliation of these great thinkers is not only possible but plausible, and I set forth the connections as they seem correct to me.

the motions of celestial bodies, postulating that the sun, the moon, and the planets (including Earth) move uniformly in circular orbits around a "central fire". The idea of a "music of the spheres" would culminate 2000 years later in Kepler's *Harmonice Mundi* (1619). Geometry was the study of proportions in their purest form, the discovery of mathematical laws in the relations of lengths, areas, and angles. The Pythagorean theorem is just the most obvious example.

The idea of Pythagoreanism as holding that all things are literally made out of numbers is a caricature based on the school's mystical currents. Undoubtedly, though, Number was considered divine, the universal principle behind harmony, rationality, and beauty in the skies and on Earth. Only through Number is it possible to understand the cosmos:

> Indeed, it is the nature of Number which teaches us all things which would otherwise remain impenetrable and unknown to every man. For there is nobody who could get a clear notion about things in themselves, nor in their relations, if there was no Number or Number-essence. By means of sensation, Number instills a certain proportion, and thereby establishes among all things harmonic relations [...]; it incorporates intelligible reasons of things, separates them, individualizes them, both in limited and unlimited things. (Philolaus, fr. B11 DK, cited by Guthrie and Fideler (1987))

4.3 The Discovery of Incommensurability

If one can still sense how sublime and fulfilling the Pythagorean worldview must have seemed to believers, it helps to understand the shock caused by the discovery of *incommensurability*. The Pythagoreans had an algorithm–today, we call it the Euclidean algorithm–to find the greatest common measure of two like magnitudes. Subtract the smaller magnitude as often as possible from the greater, and then the remainder from the smaller, and so on. *Hippasus* is usually credited with the discovery that, for certain line segments–such as the diagonal and side of a square or a regular pentagon–this algorithm never terminates (see Fig. 2 below). The discovery of incommensurability thus also marks the beginning of the mathematical struggle with infinity. Legend has it that Hippasus was drowned at sea as punishment for his blasphemy (fr. 18 A4 DK).

Euclid (Book X, Def. 1.3) and before him Plato (*Rep.* VII 534d, VIII 546c) already refer to incommensurable line segments as *irrational*,[4] Plato in a way that suggests the term had been established before, maybe by the Pythagoreans themselves. It is still a big conceptual leap from here to understanding the proportions of incommensurable magnitudes as (irrational) *numbers*, but the step seems almost inevitable.

The Pythagoreans had more immediate concerns. They had to save their sciences, in particular geometry, whose arithmetic foundation crumbled with Hippasus' discovery of incommensurability. A fundamental question that arose is what it means for different magnitudes to stand in the same proportion if this proportion no longer

[4] *árritos*, which translates more literally to *ineffable* or *inexpressible*; also *alogon*.

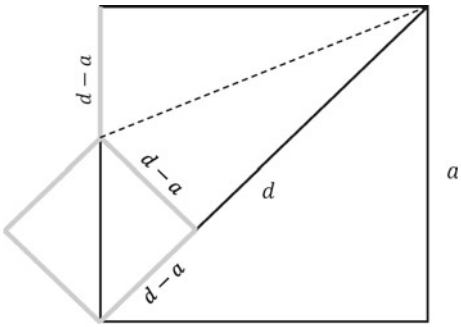


Fig. 2 Euclidean algorithm for the side and diagonal of a square. Trying to find a common measure leads to an infinite regress. In the next step, we have to subtract the side of the new square ($a_2 = d - a$) from its diagonal ($d_2 = 2a - d$)

corresponds to a ratio of natural numbers. In other words, one needs an identity criterion for proportions that applies also in the incommensurable case. The following solution is attributed to Eudoxus of Cnidus, a student of Plato. It provides the basis for the theory of proportions presented in Book V of Euclid's *Elements*.

Definition 2 Let A, B and C, D be magnitudes of the same kind. The ratio $A : B$ is equal to $C : D$ if for all natural numbers n, m one of the following three cases holds:

$$\begin{aligned} m \cdot A < n \cdot B \quad &\text{and} \quad m \cdot C < n \cdot D \\ m \cdot A = n \cdot B \quad &\text{and} \quad m \cdot C = n \cdot D \\ m \cdot A > n \cdot B \quad &\text{and} \quad m \cdot C > n \cdot D \end{aligned}$$

The second case can only occur for commensurable magnitudes (for which then $A : B = C : D = n : m$). But if we take the step of recognizing rational numbers (and thus license writing $\frac{n}{m}$), we can see from Eudoxus' definition that any proportion *partitions* the rationals such that either $\frac{n}{m} < A : B$ or $\frac{n}{m} \geq A : B$. This is precisely the idea behind Richard Dedekind's construction of the real numbers, though the fact that it took 2000 years (and the invention of set theory) to carry it out shows the magnitude of the achievement. While Dedekind (1872) makes a point of looking for arithmetic as opposed to geometric principles for the continuum, it is straightforward to translate his account into a completion of Eudoxus' theory of proportions with all the structure of the real numbers. For instance, multiplication: Given two proportions $A : B$ and $C : D$, their product is the smallest proportion $E : F$ such that, for all $k, l, m, n \in \mathbb{N}$, $nB < mA$ and $kD < lC$ implies $(lm)E \leq (kn)F$. Non-positive numbers can be included by admitting magnitudes of positive or negative orientation.

With the discovery of incommensurability, we lose the crutch of saying that the ratio of two magnitudes *is like* the ratio of two (natural) numbers. Instead, we are led to recognize proportions as numbers in their own right–those forming the continuum of reals.

4.4 The Birth of Modern Physics

The Pythagorean influence was still very present at the time of the scientific revolution, where it combined with the right amount of empirical methodology (not too little, but also not too much). Modern physics was born with the discovery of mathematical laws that are laws of proportion for the motion of matter.

Galileo performed his acceleration experiments and reported that "the spaces traversed were to each other as the squares of the times" (1638/1954, p. 179). Expressed here is not the algebraic formula $s = \frac{1}{2}at^2$, which relates dimensionful quantities on both sides, but the fact that, for pairs (s_1, t_1), (s_2, t_2) of times and corresponding distances, s_1 and s_2 have the same ratio as the squares of t_1 and t_2 (see Thm. II, Prop. II on naturally accelerated motion in the *Discorsi*). That Galileo thought geometrically is also evident in the famous passage from *Il Saggiatore* (1623) that inspired the title of this essay:

> Philosophy is written in this grand book, the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth. (Quoted from Drake (1957, p. 238))

Around the same time, Kepler combined Pythagorean ideas with Copernican heliocentrism and found his *harmonic law* for planetary motion: *The square of the orbital period is proportional to the cube of the semi-major axis of its orbit.* Two generations later, Newton proved that this law follows from a centripetal force inversely proportional to the squares of the distances (Prop. XV, Thm. VII in the *Principia*). In the *Principia*, one still looks in vain for differential equations or even the famous formula $F = \frac{GmM}{r^2}$. Classical mechanics is developed geometrically, including "the method of the first and last ratios of quantities" introduced to apply results of Euclidean geometry to curve segments as they become vanishingly small.[5]

5 Why Laws of Nature are Mathematical

In the preface to the first edition of the *Principia*, Newton made explicit how he saw the relationship between mathematics and natural philosophy in the task of "reduc[ing] the phenomena of nature to mathematical laws" (Newton 1687/1999, p. 381). The practical side of mechanics involves the manual art of measuring magnitudes and carrying out geometrical constructions. "[G]eometry is founded on mechanical practice and is nothing other than that part of universal mechanics which reduces the art of measuring to exact propositions and demonstrations"

[5] Although Newton had developed a more abstract differential calculus in his *Method of Fluxions* (completed 1671, but not published until 1736), it was not used in the *Principia* (first published 1687).

(p. 382). In essence, mechanics, as an empirical science, falls short of exact geometry only through practical limitations, particularly the inaccuracies of measurements.

It would be an overstatement to call Newton a Pythagorean. But he is part and pinnacle of a long tradition of thought that recognizes geometry–in the sense of the rational investigation of relations between magnitudes–as the nexus between physics and mathematics. The understanding we can gather from this tradition is that the appropriacy of mathematics for the formulation of the laws of nature is neither accidental nor merely a matter of convenience. There is something genuinely mathematical about the very concept of natural laws.

Why are the laws of physics mathematical? Because physics is the science of matter in motion. Regularities of motion manifest themselves in proportions of times, distances, and other geometric or perhaps kinematic quantities. Proportions are numbers. And numbers are mathematical.

I believe this answer is still relevant today, as our physics and mathematics have become so much more sophisticated. A physical theory can involve whatever kind of abstract calculus and higher-order mathematical structures we need. At the end of the day, the theory must link up to *matter in motion*, and this is where mathematics meets the physical world, both conceptually and metaphysically. This presupposes, however, two things that can no longer be taken for granted in contemporary physics: the laws must be mathematically consistent and precise. And the theory must postulate a clear ontology of matter as that to which the mathematical formalism ultimately refers.[6]

There would thus be another story to tell about how the Pythagorean understanding has been lost in more recent times; perhaps completely when Bohr declared that the formalism of quantum mechanics "represents a purely symbolic scheme" (in Schilpp (1949, p.110)). What a fall from grace for *theory*, from a vision of the divine *logos* to a meaningless manipulation of symbols that refers to nothing in the world. But I'll leave this tragedy for another time.

5.1 Numbers as Proportions

When I say that (real) numbers are proportions, I mean that they are relations between magnitudes. Magnitudes themselves are not numerical (only relative to a chosen unit of measurement) and include spatiotemporal relations as well as concrete physical properties. The metaphysical details of this proposal remain to be spelled out elsewhere.[7] Here, I want to make the point that the understanding of numbers as proportions (rather than abstract objects of set theory) narrows the gap between what we now call Platonism and nominalism.

[6] Ideally, it needs what Dürr, Goldstein, and Zanghì (1992) named *primitive ontology* (see Lazarovici and Reichert (2022) for a recent discussion) or what John Bell (2004, Chap. 7) called *local beables*.

[7] I will also not discuss the ontological status of other mathematical objects. Both a selective realism and full-blown Platonism are consistent with the view I propose in regard to numbers.

The ratio of the diagonal to the side of a square is $\sqrt{2}$, as is the ratio of the sides of two squares where the first has twice the surface area of the second. These are true identity statements, necessarily and a priori. Numbers are universals transcending their various instantiations since everything that is particular to given lengths or areas or other magnitudes quite literally cancels out when we consider their proportions. This is why the Pythagoreans insisted, as Proclus reports, that "numbers are purer and more immaterial than magnitudes" and appear "to every mind as one and not many, and as free of any extraneous figure or form" (1992, p. 78).

On the other hand, if space and time are actually continua, all real numbers are instantiated in the physical world–in space-time itself and (if this is still too abstract) in the motions of material entities. This requires less than a metric structure since we don't need absolute distances. In fact, the nominalist program of Hartry Field (*Science without Numbers*, 2016), which builds on Hilbert's axiomatization of Euclidean geometry, can be read as an exploration of how far one can get with only intrinsic structure, such as relations of congruence. I just don't think it thereby "eliminates" numbers in any metaphysically interesting sense. One can debate the question of ontological priority (if one is so inclined). But if, say, a circle exists in your universe, then the number π exists as well.

5.2 Conclusion

I am certainly not advocating a return to the mathematics of the early 18th century or dismissing the awesome progress we have made ever since. We have explored so much more of the mathematical universe, set our inquiries on solid logical foundations, and developed powerful concepts and mathematical methods without which modern science and technology would not be possible. We have gained tremendous knowledge, but we have also lost some of the wisdom of past giants.

It is easy to get lost in formalism and mathematical abstraction, to the point that it seems a great mystery how any of it could have anything to do with the natural world, let alone a *logos* that is not of our own making. This often combines with a tendency to make us humans both too small and too large at the same time: it seems inconceivable that we could have insight into either mathematical truth or the laws of nature unless we are somehow the engineers of both. Detlef Dürr strictly rejected such thinking. For him, the purpose of doing science was not only to understand the cosmos, but also to recognize our proper place in it.

I believe–and this is one of the many insights I owe to Detlef–that the understanding of numbers as proportions is at least the beginning of an answer to why the laws of nature are mathematical. The more profound mystery is why laws of nature exist in the first place; what explains the very rationality and comprehensibility of the universe. With this, I leave the final word to my teacher:

> What is the origin of physical law? We could answer: there is no origin; it is a brute fact that everything can be described by a law, and in the end, it is our human law because our senses experience regularities. And we are looking for a code to describe these experiences. And

> mathematics is a good code that we have developed in a process of trial and error. This is wrong. We are not working like that, at least not as physicists. If we did, we would pile up all kinds of mathematical garbage just as moles pile up mounds of earth. Galilei didn't do that, Newton didn't do that, and least of all Einstein. To better understand why this idea is wrong, you must understand the mathematical formulation of the law, or rather of the laws that we have discovered so far. It is not a summary of our observations that all bodies fall to the ground; it is not said that some bodies do this and others do that; it is not a bookkeeper's order that we write down. We are looking purposefully for the underlying law of everything. We are guided by ideas of beauty, simplicity, elegance that the law should satisfy, and with these categories, we are successful. There is no good explanation for our successes [...].
>
> — Detlef Dürr (2007): *Was heißt und zu welchem Ende studiert man Physik?*[8]

Acknowledgements I owe most of the insights explored in this essay to Detlef Dürr, who owed just as much to his friend Reinhard Lang. I thank Yoav Beirach and Enrico Piergiacomi for very helpful comments and Stephen Lyle for his excellent copy editing.

References

J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, 2nd edn. (Cambridge University Press, 2004)
J. Burnet, *Early Greek Philosophy*, 3rd edn. (Adam & Charles Black, London, 1920)
R. Dedekind, *Stetigkeit und irrationale Zahlen* (Vieweg, 1872)
R. Descartes, La Géométrie. In *Discours de la méthode* (Jan Maire, Leyden, 1637)
S. Drake (ed.), *Discoveries and Opinions of Galileo* (Knopf Doubleday, New York, 1957)
D. Dürr, Was heißt und zu welchem Ende studiert man Physik? Weihnachtsvorlesung WS 07/08 (2007), http://www.mathematik.uni-muenchen.de/~duerr/WeihnachtsvorlWS0708.pdf
D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the origin of absolute uncertainty. J. Stat. Phys. **67**(5–6), 843–907 (1992)
D. Dürr, D. Lazarovici, Der Dialog des Demokrit. In *Welt Der Gründe*, Deutsches Jahrbuch Philosophie. (Felix Meiner Verlag, 2012), pp. 1207–1217
H. Field, *Science Without Numbers*, 2nd edn. (Oxford University Press, Oxford, New York, 2016)
G. Galilei, *Dialogues Concerning Two New Sciences. Translated by H. Crew and A. de Salvio* (Courier Corporation, Original work published 1638, 1954)
K.S. Guthrie, D. Fideler, *The Pythagorean Sourcebook and Library: An Anthology of Ancient Writings Which Relate to Pythagoras and Pythagorean Philosophy* (Red Wheel/Weiser, 1987)
H. Jones, Heraclitus: Fragment 31. Phronesis **17**(3), 193–197 (1972)
B. Jowett (ed.), *The Dialogues of Plato* (Oxford University Press, 1892)
G. Kirk, J.E. Raven, M. Schofield, *The Presocratic Philosophers: A Critical History with a Selection of Texts* (Cambridge University Press, 1983)
E. Kurtz, *Interpretationen zu den Logos-Fragmenten Heraklits* (Georg Olms Verlag, 1971)
A. Laks, G. Most, (eds.), *Early Greek Philosophy. 9 Volumes*. Loeb Classical Library (Harvard University Press, Cambridge, MA, London, 2016)
D. Lazarovici, P. Reichert, The point of primitive ontology. Found. Phys. **52**(6), 120 (2022)
T. Maudlin, *New Foundations for Physical Geometry: The Theory of Linear Structures* (Oxford University Press, Oxford, New York, 2014)

[8] *What is physics and to what end does one study it?* Christmas lecture at the University of Munich. Translation by D.L.

I. Newton, *The Principia: Mathematical Principles of Natural Philosophy. A new translation by I.B. Cohen and A. Whitman, assisted by J. Budenz* (University of California Press, 1999)
K.R. Popper, How the Moon might throw some of her Light upon the Two Ways of Parmenides. Class. Q. **42**(1), 12–19 (1992)
Proclus, *A Commentary on the First Book of Euclid's Elements. Translated by Glenn R. Morrow* (Princeton University Press, 1992)
W. Schadewaldt, *Die Anfänge der Philosophie bei den Griechen: die Vorsokratiker und ihre Voraussetzungen*. Tübinger Vorlesungen Band I (Suhrkamp, 1978)
P.A. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist*, 1st edn. Number VII in The Library of Living Philosophers (The Library of Living Philosophers Inc., Evanston, Illinois, 1949)
E. Schrödinger, *'Nature and the Greeks' and 'Science and Humanism'* (Cambridge University Press, Cambridge, 2014)
M. Tegmark, *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality* (Knopf Doubleday Publishing Group, 2014)
R. Tumulka, On the question why there exists something rather than nothing, in *The Philosophy of Cosmology*. ed. by J.D. Barrow, J. Silk, K. Chamcham, S. Saunders (Cambridge University Press, Cambridge, 2017), pp.76–82
S. Wenmackers, Children of the cosmos, in *Trick or Truth? The Mysterious Connection Between Physics and Mathematics, The Frontiers Collection*. ed. by A. Aguirre, B. Foster, Z. Merali (Springer International Publishing, Cham, 2016), pp.5–20
E.P. Wigner, The unreasonable effectiveness of mathematics in the natural sciences. Richard courant lecture in mathematical sciences delivered at New York University, May 11, 1959. Commun. Pure Appl. Math. **13**(1), 1–14 (1960)

Bohmian Mechanics as Cartesian Science

Michael Esfeld

(for volume in honour of Detlef Dürr)

Abstract The paper shows how Bohmian mechanics fits into modern science as conceived by Descartes. The primitive ontology of particle positions only vindicates the Cartesian conception of matter as *res extensa* also in the domain of quantum physics. Finally, I briefly sketch out how Bohmian mechanics can also be construed as respecting the limits that Descartes set for modern science: there is no conflict between the ontology and dynamics of matter as set out in Bohmian mechanics and a Cartesian stance that takes the mind to be irreducible to matter.

1 From de Broglie and Bohm to Bohmian Mechanics

For many years, I've misconceived the de Broglie-Bohm quantum theory. Louis de Broglie's (1928) original proposal is one of a wave-particle dualism. Thus, in the double slit experiment, the particle goes through one slit and the wave travels through both slits. But this doesn't make sense: As soon as one considers a system that consists of more than one particle, the wave function cannot be construed as a wave that spreads in physical space. Whatever it may be, the wave *function* isn't a wave.

David Bohm's (1952) revival of de Broglie's (1928) proposal is of another sort: Bohm associates the wave function with a specifically quantum force that is known as the quantum potential. Accordingly, he writes the guiding equation down as a second order equation. But this isn't convincing either. Whatever it may be, the wave function cannot be a Newtonian force. It doesn't satisfy Newton's third law. That is precisely the reason why position is an additional, so called hidden variable in Bohm's theory: the wave function acts on the particles, but the particles don't act back on the wave function. Hence, knowing the wave function does in general not

M. Esfeld (✉)
Department of Philosophy, University of Lausanne, Lausanne, Switzerland
e-mail: Michael.Esfeld@unil.ch
URL: http://www.michaelesfeld.com

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_27

provide information about the actual particle positions. Consequently, it isn't possible to conceive the *dynamics* of quantum objects along the lines of classical dynamics by adding a specifically quantum force. Trying to do so draws the attention away from the crucial point, namely that it is possible—and reasonable—to conceive the *ontology* of quantum objects on a par with the ontology of classical objects, namely as particles that always have a determinate position and hence move on trajectories. The confusion about a quasi-classical dynamics in Bohm's (1952) proposal often provokes the misconception that throws the baby away with the bathwater, namely to abandon the classical ontology of particles together with the classical dynamics.

Although the wave function can neither be a wave nor a force spreading in physical space, the view is widespread that in Bohm's quantum theory, both the particles and the wave function make up the basic ontology of physical things. Thus, David Albert writes in his seminal book of 1992 that was the first book that cleared up all the fuzz about measurement in quantum mechanics:

> What the physical world consists of besides particles and besides force fields, on this theory, is (oddly) *wave functions.* … The *quantum–mechanical wave functions* are conceived of in this theory as genuinely physical *things* … (Albert 1992, p. 135).

This is odd indeed. For me, I couldn't make sense of that and therefore didn't take Bohm's theory seriously for many years. For if indeed that were the truth of the matter, the Everettian had a good argument to convince the Bohmian: If we need the wave function anyway in our ontology of physical things, there's no point to admit particles in addition to the wave function as so called hidden variables. Harvey Brown (who used to be a Bohmian) and David Wallace (who is an outspoken Everrettian) drove that point home in a famous paper in 2005 (Brown and Wallace 2005). However, Everettian quantum mechanics hardly is a convincing alternative through its denial of empirical reality, that is, the existence of a distribution of matter in space and time that manifests itself in measurement outcomes. By way of consequence, for lack of a better alternative, the legacy of Copenhagen still is influential and framed also my thoughts about quantum mechanics for many years, albeit purged of positivism and the fuzz about measurement.

However, everything is clear since Detlef Dürr, Shelly Goldstein and Nino Zanghì have developed Bohmian mechanics since the end of the 1980s. In their seminal paper published in 1992 (Dürr et al. 1992), they write:

> What we regard as the obvious choice of primitive ontology—the basic kinds of entities that are to be the building blocks of everything else (except, of course, the wave function)—should by now be clear: Particles, described by their positions in space, changing with time—some of which, owing to the dynamical laws governing their evolution, perhaps combine to form the familiar macroscopic objects of daily experience. (Quoted from the reprint in Dürr, Goldstein and Zanghì 2013, p. 29).

Hence, the conclusion from the fact that an ontology of both the particles and the wave function being physical things is unconvincing is not that there only is the wave function. On the contrary, the ontology—that is, the primitive ontology of matter in space and time—is particles only. John Bell has laid the ground for clearing up the

confusion in this way since he has coined the term "local beables" in 1975 (Bell 1975).

What, then, is the wave function? According to Dürr, Goldstein and Zanghì, the wave function is nomological. That is not to say that it is a law on a par with the guiding equation or Newton's three laws. It is to say that it plays a nomological role: It consists in the role that it exercises for the motion of the particles—instead of being a physical entity over and above the particles. Again, John Bell makes this point without using the term "nomological" in the paper in which he introduces the notion of local beables in 1975:

> One of the apparent non-localities of quantum mechanics is the instantaneous, over all space, 'collapse of the wave function' on 'measurement'. But this does not bother us if we do not grant beable status to the wave function. We can regard it simply as a convenient but inessential mathematical device for formulating correlations between experimental procedures and experimental results, i.e., between one set of beables and another. Then its odd behaviour is acceptable as the funny behaviour of the scalar potential of Maxwell's theory in Coulomb gauge. (Quoted from the reprint in Bell 1987, p. 53).

Bell's wording here is unnecessarily instrumentalist. The crucial point is this one: If the particles are recognized as local beables, then we have all the beables that are required in physics. There is no need to admit the wave function as a further beable. Refusing to do so has moreover the advantage that it helps to dissipate concerns about quantum non-locality being some kind of "spooky action at a distance" (Born and Einstein 1971, p. 158) in Bohmian mechanics.

Bohmian mechanics satisfies Einstein's principle of separability (see Einstein 1948; English translation of the decisive passage in Howard 1985, pp. 187–188): Each particle always has a definite position of its own, and position is the only property of the particles. There are no superpositions of anything in physical space. Furthermore, in Bohmian mechanics, the violation of Einstein's principle of local action is not tied to—and hence not limited to—wave function collapse; the wave function never collapses. The rejection of superpositions and the respect of the principle of separability have the consequence that quantum non-locality is generic in Bohmian mechanics. For instance, in the double slit experiment, the trajectory of the particle after it has passed through one slit depends on whether or not the other slit is open. That dependence is immediate and independent of the distance between the two slits. The non-locality hence arises well before the particle hits the screen and a measurement occurs. But there's no action between the two slits.

In treating the wave function as nomological instead of as a physical entity on a par with the particles, Bohmian mechanics brings out that the nomological structure of the world is non-local. This means that the dynamics of the universe is holistic: The time evolution of any one particle in the universe is correlated with strictly speaking the time evolution of any other particle in the universe via the universal wave function figuring in the guiding equation. These non-local correlations are a nomological primitive. There is no further explanation or account of why the dynamics of the universe is non-local rather than local. In particular, there is no action in the sense of something spreading or being transmitted instantaneously across physical space.

And there's no holism that blurs the distinction between the particles: They always have a determinate position, and position is their only property.

If the wave function is nomological in this sense, all the major stances in the metaphysics of laws of nature can be applied to it. In particular, it admits a treatment in the framework of what is known as Humeanism. On Humeanism, the primitive ontology is the *entire* ontology. Laws are the axioms or theorems of the best system, that is, the system that strikes the best balance between being simple and being informative in representing the evolution of the particular matters of fact, the particles in the case of Bohmian mechanics. In philosophical shoptalk, the laws supervene on the particle trajectories throughout space and time. The trajectories determine the laws, instead of the laws predetermining, governing or even producing the trajectories.

Applying this stance to the wave function leads to quantum Humeanism, which was elaborated on by several researchers independently of one another in the 2010s (Callender 2015, Miller 2014, Esfeld 2014, Bhogal and Perry 2017). The wave function is a parameter that figures in the best system. Hence, first come the particle trajectories as a matter of fact, then comes the universal wave function as dynamical parameter that belongs to the system that yields the best representation of the particle trajectories in terms of being simple and informative. This stance makes evident again that there is no conclusive reason to admit any sort of a necessary connection between distinct particles. As a matter of fact, when one seeks for a simple and informative representation of their motion, this will be one that represents their motion as being correlated via an entangled wave function figuring in the law of motion. That is all that there is to quantum non-locality.

Furthermore, quantum Humeanism implies that not only the wave function is nomological in being a parameter in the best system, but so are all the other physical magnitudes apart from the primitive variable of position. In Bohmian mechanics, all the magnitudes apart from position are situated on the level of the wave function, including the classical magnitudes of mass and charge (see Brown et al. 1995 and 1996). They're not intrinsic properties of the particles (see Esfeld et al. 2017). Hence, the primitive ontology, the mosaic of matter in space and time, consists of naked particles so to speak that are characterized only by their positions. This stance has subsequently become known as Super-Humeanism (see Esfeld and Deckert 2017, Sect. 2.3).

Detlef Dürr wasn't a Humean about the nomological structure for sure. But this isn't a big deal to my mind: Humeanism is helpful to illustrate a number of issues such as the significance of non-locality as sketched out in this section. The big deal is to be clear about the distinction between on the one hand ontology in the sense of a primitive ontology of matter in space and time and on the other hand dynamical or nomological structure in the sense of what enters our physical theories in virtue of the function that it exercises for the motion of matter as captured by the primitive ontology.

2 *Matter as* Res Extensa

In the first discussion that I had with Detlef Dürr in September 2011, Detlef emphasized that all there is to matter is particles standing in spatial relations and the change of the relations—in short, matter moves. Everything else is dynamical or nomological structure consisting in the role that it plays for the motion of matter. This means that Bohmian mechanics is Cartesian science. Bohmian mechanics illustrates that Cartesianism remains valid in quantum mechanics. By "Cartesian science", I mean natural science as construed by Descartes.

The main feature of the Cartesian conception of matter is objectivity: The sensory qualities such as colours, sounds, tastes, smells and the like do not belong to the things in nature, but to our way of gaining knowledge of them by using our senses. If one abstracts from all these features, what remains of the natural world is extension and change in extension—that is, motion (see notably *Principia Philosophiae*, part 2, § 4). As physics then shows, extension, in turn, is in the last resort point particles standing in distance relations to each other.

There is a good reason for admitting that objectivity boils down to such a meagre treatment of matter: When examining a knowledge claim in science, all the empirical evidence that can be obtained to confirm or invalidate the claim in question consists in observations of the positions and changes of position of discrete objects. Accordingly, all measurement outcomes are recorded as relative positions within configurations of discrete objects—such as, for instance, pointer positions or digital numbers on a screen. John Bell emphasizes this point:

> … in physics the only observations we must consider are position observations, if only the positions of instrument pointers. It is a great merit of the de Broglie–Bohm picture to force us to consider this fact. If you make axioms, rather than definitions and theorems, about the 'measurement' of anything else, then you commit redundancy and risk inconsistency. (Quoted from the reprint in Bell 1987, p. 166).

The qualification "in physics" is appropriate: Common sense observations typically involve colours, sounds or scents of spatially arranged objects. In common sense, the positions of objects are discerned by means of these sensory qualities. Science abstracts from the sensory qualities. What then remains are the relative positions of discrete objects and their change. These are correlated with the sensory qualities, in the sense that science can account for changes in sensory qualities on the basis of changes in position.

According to physics, macroscopic objects are composed of microscopic objects that ultimately are point particles standing in distance relations. Consequently, if a theory describes the spatial arrangement of the particles and its change in time correctly—that is, the arrangement and evolution of fermionic matter (see Bell 1987, p. 175)—, it has got everything right that can ever be checked in scientific experiments (see also Maudlin 2019, pp. 49–50). Two theories that agree on the spatio-temporal arrangement of the elementary particles defined in terms of the positions of these particles only cannot be distinguished by any empirical means, whatever else they

may otherwise say and disagree on. By the same token, two possible worlds with the same spatio-temporal arrangement of the elementary particles are indiscernible by any scientific means.

This is the strongest argument for treating position as the only basic or primitive physical magnitude, and thus for the natural world, insofar as it is accessible to science, being *res extensa* only: Admitting anything else over and above positions as basic or primitive would imply treating empirically indiscernible situations or worlds as being nonetheless different in some matters of fact. Obviously, this is a generalization of Leibniz' famous argument against Newton's ontological commitment to absolute space and time (see notably Leibniz' third letter to Clarke, § 5, in Leibniz 1890, pp. 363–364): The argument applies, in fact, to anything that is admitted as ontologically primitive in the scientific description of the world beyond relative positions and their change.

That notwithstanding, if all that there is to matter is distance relations between sparse point particles and the changes in these distances, it may seem that their material nature fades away upon inquiry. However, this concern is unfounded. There is nothing incoherent in the notion of point particles as elaborated on in Bohmian mechanics.

If there is a plurality of objects, there has to be something that individuates them—that is, something that answers the question why *this* is one object, *that* another, etc., so that there is a plurality of objects instead of just one. Furthermore, there also has to be something that unites these objects so that they make up a world. In other words, there has to be a world-making relation; that is, a relation that binds all and only those objects together that belong to a world. It is evident that the distance relation fulfils the latter task: All and only those objects that are spatially related constitute a world. If there were objects not at a distance from each other, they would inhabit different worlds. If they are related by distance, they are in one and the same world (see Lewis 1986, pp. 69–81).

Moreover, the distance relations—and only they—individuate the objects: What distinguishes each object in a configuration of objects is the position that it has relative to all the other objects. Even if a configuration is partially symmetrical, there always is at least one object in the real world outside that symmetry relative to which all the other objects can be distinguished. Magnitudes that are attributed to physical objects over and above their relative positions—such as mass or charge—cannot distinguish the latter: they differentiate between various kinds of particles, such as the particle species admitted in today's standard model of elementary particles. But they cannot distinguish between the individual particles within a species or kind, because all the particles of a given species—such as, for instance, all electrons—have the same values of mass, charge, etc. The demand for something that individuates the physical objects is fulfilled by the distance relations, and by them only. Therefore, there is no need for anything more than distance relations to both individuate the objects and have a relation that binds them together so that they constitute a world.

Indeed, one can regard these considerations as confirming the Cartesian metaphysics of nature, and vindicating it also in the context of contemporary science including quantum mechanics: Nature, insofar as it is accessible to scientific enquiry,

is *res extensa.* That is to say, there is nothing more to matter than extension in the guise of distance relations individuating point particles and the change in these relations. Against this background, Dirk-André Deckert and myself have set out to show in our book published in 2017 how modern physics can be construed on the basis of a primitive ontology of matter that is defined by the following two axioms or principles:

(1) There are distance relations that individuate simple objects—namely, matter points.
(2) The matter points are permanent, with the distances between them changing. (Esfeld and Deckert 2017, p. 21).

At this general level, the geometry into which the configuration of matter has to be conceived as being embedded in order to achieve a representation of its dynamics remains completely open. Accordingly, the viability of this primitive ontology has to be vindicated for each physical theory separately, as outlined in the book Esfeld and Deckert (2017).

3 *Mind as* Res Cogitans

When I asked Detlef Dürr about the nature of the mind, he tended to reply that we shall tackle this subject when we will have a better understanding of physics. It seems to me that this attitude is based on a misunderstanding: Everything becomes clear when one realizes that Bohmian mechanics is Cartesian science. Unfortunately, our discussions about these issues came to an abrupt end through Detlef's untimely death.

Cartesian science abstracts from all subjective judgements and seeks objectivity. That is how one gets to Bohmian mechanics: matter being featureless point particles in motion. However, this very method of seeking objectivity implies that it can in principle not be applied to subjective features. If the scientific viewpoint consists in abstracting from the latter in order to reach objectivity, then it simply follows that those same subjective features are not accessible to the scientific viewpoint.

This limitation concerns in the first place sense experience. A being that has sense experience is not merely an object that moves according to certain laws of motion; rather, it has a subjective perception and feeling of what it is like to be in the world, having certain qualitative experiences. Science can discover sufficient physiological conditions for having sense experience, and the content of the experience may supervene on certain brain states, given certain conditions in the environment. Nevertheless, any scientific theory misses the qualitative character of the experience, the subjective perspective on the world. It cannot account for what it is like to see colours, taste cheese, smell smoke, jump for joy, etc. Accordingly, the issue of how to account for subjective experience has come to be known as the hard problem of consciousness following Chalmers (1996).

Subjective experience pertains to many higher-level animals. Thought and action—which, as far as we know, characterize only humans—presuppose a subjective perspective on the world, and thus experience, but are still categorically different from it. The obvious argument against human thought and action not being accessible to the method of Cartesian science is that in the case of these, the issue is not what the objective facts are, but how human subjects assess them in forming beliefs and intentions for action. This brings in free will. A being forms beliefs if and only if she has the capacity to position herself with respect to her sensory impressions, desires and needs and to make up her mind by deliberating about reasons for her beliefs and actions.

This freedom is also a presupposition for science. The very formulation, endorsement and testing in experience of a scientific theory presupposes the freedom of scientists to position themselves with respect to the evidence. The referents of the theory—whatever the theory poses as existing in the world—cannot impose acceptance of the theory on persons or justify the theory. In that sense—as beings that formulate and justify theories—persons are as ontologically primitive as are the point particles.

As one can spell out Cartesian *res extensa* in terms of point particles being individuated by distance relations, so one can spell out Cartesian *res cogitans* in terms of relations that are characteristic of the mind, namely thinking relations. That is to say: Standing in distance relations (extension) makes it that points are matter points (point particles), whereas standing in thinking relations makes it that points are minds. In neither case is there a substance in the sense of a thing with an intrinsic essence (see Esfeld 2020, Chap. 3, for details).

Being clear about the distinction between primitive ontology and dynamical structure is a big deal also because it dissipates concerns that one may have about the theory of matter being incompatible with the central features of the mind such as notably free will. Bohmian mechanics is a deterministic theory. However, on Bohmian Humeanism, the trajectories that the particles take in fact fix what the wave function of the universe is, instead of the wave function governing or even producing the trajectories. Hence, on this stance, there is no conflict between a deterministic dynamics and humans having free will because the laws being deterministic doesn't imply that anything that happens in the world is predetermined in the sense of being necessitated by something else. Against this background, there is the prospect of making a precise proposal how there can be—even libertarian—free will in a Bohmian universe (see Esfeld 2020, Sect. 2.3, and Esfeld 2022 for details, although the proposal still needs elaboration).

Furthermore, Bohmian mechanics leaves many options in the philosophy of time open, including even presentism. As Tim Maudlin notes.

> Bohm's theory is deeply congenial to an ontology which maintains that all which exists is that which exists *now*, i.e. at a point in time classically conceived. … Those puzzled about the status of velocities in an ontology in which only an instant of time *exists* can happily adopt a Bohmian ontology of particles (with position) and the wave-function. (Maudlin 2011, p. 113 note 22).

Bohmian mechanics is not even committed to absolute time. It can be conceived in a Leibnizean, relationalist framework. As Detlef Dürr (2020) showed in one of the last research projects to which he contributed, Bohmian mechanics can be done on shape space, that is, by working only with relational and thus scale invariant quantities, instead of conceiving the particle configuration as being inserted in an absolute space and time. Hence, in Bohmian mechanics, the ontology of space and time can be as parsimonious as the ontology of matter. In particular, if one endorses Bohmian presentism (which can be construed as relationalist presentism) and Bohmian Humeanism, there is no problem to admit an open future, which one may consider as being required for free will.

This is the big deal: epistemic humility about science without compromising the scientific enterprise. Bohmian mechanics shows how one can be committed to a fully objective, fully scientific realist and fully deterministic theory of matter and its evolution and yet be clear about the limits of objective science. There's nothing in this theory that imposes anything on us that might be considered as coming into conflict with central features of the mind such as the openness of the future or human free will even on a libertarian conception of free will. To avoid the pitfalls of scientism—that is the view that science in the sense of modern natural science is unlimited—, there is no need to compromise physics in any way (as done when one takes quantum mechanics to undermine deterministic laws of nature or even realism). One just has to understand how Bohmian mechanics vindicates Cartesian science also in the domain of quantum physics.

References

D.Z. Albert, *Quantum Mechanics and Experience* (Harvard University Press, Cambridge (Massachusetts), 1992)

J.S. Bell, The theory of local beables. *Th-2053-CERN*, 28 July 1975 (1975)

J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987)

H. Bhogal, Z. Perry, What the Humean should say about entanglement. Noûs **51**, 74–94 (2017)

D. Bohm, A suggested interpretation of the quantum theory in terms of 'hidden' variables. I and II. Phys. Rev. **85**, 166–179, 180–193 (1952)

H. Born, M. Born, A. Einstein, *The Born-Einstein Letters: Translated by Irene Born.* Macmillan, London (1971)

H.R. Brown, C. Dewdney, G. Horton, Bohm particles and their detection in the light of neutron interferometry. Found. Phys. **25**, 329–347 (1995)

H.R. Brown, A. Elby, R. Weingard, Cause and effect in the pilot-wave interpretation of quantum mechanics, in *Bohmian Mechanics and Quantum Theory: An Appraisal*. ed. by J.T. Cushing, A. Fine, S. Goldstein (Springer, Dordrecht, 1996), pp.309–319

H.R. Brown, D. Wallace, Solving the measurement problem: de Broglie-Bohm loses out to Everett. Found. Phys. **35**, 517–540 (2005)

C. Callender, One world, one beable. Synthese **192**, 3153–3177 (2015)

D.J. Chalmers, *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, Oxford (1996)

L. de Broglie, La nouvelle dynamique des quanta, in *Electrons et photons: rapports et discussions du cinquième Conseil de physique tenu à Bruxelles du 24 au 29 octobre 1927 sous les auspices de l'Institut international de physique Solvay*. Gauthier-Villars, Paris, pp. 105–132. English translation in G. Bacciagaluppi, A. Valentini (2009): *Quantum theory at the crossroads. Reconsidering the 1927 Solvay conference.*Cambridge University Press, Cambridge, pp. 341–371 (1928)

D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the origin of absolute uncertainty. J. Stat. Phys. **67**, 843–907 (1992)

D. Dürr, S. Goldstein, N. Zanghì, *Quantum Physics Without Quantum Philosophy* (Springer, Berlin, 2013)

D. Dürr, S. Goldstein, N. Zanghì, Quantum motion on shape space and the gauge dependent emergence of dynamics and probability in absolute space and time. J. Stat. Phys. **180**, 1–43 (2020)

A. Einstein, Quanten–mechanik und wirklichkeit. Dialectica **2**, 320–324 (1948)

M. Esfeld, Quantum humeanism. Philosophical Quarterly **64**, 453–470 (2014)

M. Esfeld, *Science and Human Freedom* (Palgrave-Macmillan, London, 2020)

M. Esfeld, Super-humeanism and mental causation, in *Powers, Time and Free Will*. ed. by C.J. Austin, A. Marmodoro, A. Roselli (Springer, Cham, 2022), pp.257–272

M. Esfeld, D.-A. Deckert, *A Minimalist Ontology of the Natural World* (Routledge, New York, 2017)

M. Esfeld, D. Lazarovici, V. Lam, M. Hubert, The physics and metaphysics of primitive stuff. Br. J. Philos. Sci. **68**, 133–161 (2017)

D. Howard, Einstein on locality and separability. Stud. Hist. Philos. Sci. **16**, 171–201 (1985)

G.W. Leibniz, *Die philosophischen Schriften von G.W. Leibniz. Band 7*, ed. by C.I. Gerhardt. Weidmannsche Verlagsbuchhandlung, Berlin (1890)

D. Lewis, *On the Plurality of Worlds* (Blackwell, Oxford, 1986)

T. Maudlin, *Quantum Non-locality and Relativity* (3rd edn. 1st ed. Wiley-Blackwell, Chichester, 1994, 2011)

T. Maudlin, *Philosophy of Physics: Quantum Theory*. Princeton University Press, Princeton (2019)

E. Miller, Quantum entanglement, Bohmian mechanics, and Humean supervenience. Australas. J. Philos. **92**, 567–583 (2014)

Who's Afraid of the Measurement Problem?

Valia Allori

Abstract Scientific realists usually claim that quantum mechanics can be made compatible with scientific realism by solving the measurement problem, even if there is disagreement about which solution is best. In this paper I argue this is due to having different views about what it means to make quantum theory compatible with scientific realism: 'relaxed' realists think it is enough to solve the adequacy problem, 'modest' realists believe that there is also a precision problem, while 'robust' realists insist that quantum theory still needs to be suitably completed. These attitudes are connected with the type of explanation one favors: while relaxed realists favor principle theories, robust realists prefer constructive theories, and modest realists provide non-constructive dynamical hybrids as long as they preserve locality and separability.

1 Introduction

Quantum theory suffers from the *realism problem*: it is incompatible with scientific realism, as it cannot provide a clear picture of reality. Traditionally scientific realists think that one can solve the realism problem by solving the *measurement problem.* Solutions of the measurement problem include the GRW theory,the pilot-wave theory, and the many-worlds theory. People disagree about which solution is best. In this paper I focus on the disagreement between primitive ontologists, who favor the pilot-wave theory, and wavefunction realist, who do not. To better understand their differences, I also discuss the information-theoretic (IT) approach. I argue (Sect. 2) that the disagreement between these approaches boils down to their disagreement about how standard quantum theory should solve the realism problem. Proponents of the IT approach think that the realism problem is an adequacy problem: the theory needs to make empirically adequate predictions. Wavefunction realists instead focus on a precision problem (which is the measurement problem): a theory has to have an

V. Allori (✉)
University of Bergamo, Bergamo, Italy
e-mail: valia.allori@fastmail.com

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_28

ontology whose dynamics is unique and precisely defined. Finally, primitive ontologists maintain that the problem to solve is a completeness problem: the theory has to have a clear and precise microscopic spatiotemporal ontology. Also, I argue that the reason for these differences is that they have diverging realist commitments (Sect. 3), because they think of satisfactory theories differently (Sect. 4). Those who think the realism problem is the adequacy problem endorse a relaxed realist attitude: a theory should systematize the phenomena with principles constraining the phenomena. Instead, those who think that quantum theory is incomplete have in mind a robust realist understanding, where a theory explains the phenomena constructively. Finally, those who think that the precision problem is the one to be solved endorse a modest version of realism, which combines an interest in the reality behind the phenomena with a non-constructive understanding guided by the desire of keeping a local and separable ontology. I start Sect. 5 by showing how it is unclear why relaxed realists are really realists, given their focus on appearances. Moreover, I argue that it does little sense for primitive ontologists to discuss the other solutions of the measurement problem other than the pilot-wave theory. As for the wavefunction realists, first, I show that the type of explanations they provide seems to undermine their motivation for looking for a better theory than standard quantum theory. I conclude in Sect. 6.

2 Quantum Problems

Standard quantum mechanics is merely a recipe for predicting experimental results. As such, it is incompatible with scientific realism, the view that theories can give us information about the world beyond the appearances. Let us call this incompatibility *the realism problem*. Traditionally, it is argued that the realism problem is the measurement problem. This problem arises when we assume that the wavefunction, the main object of the theory evolves in time according to the linear Schrödinger equation: since a solution of this equation describes possible states of affairs of the world, because of linearity any sum of solutions (superpositions) describes possible situations too. If one also assumes that every system is completely described by such wavefunction, then there will be 'superpositions of states' at all scales, such as a cat in a superposition state of 'living' and 'dead.'

2.1 The Adequacy Problem

Thus, the first problem of standard quantum theory is that it is *not empirically adequate*: we observe no macroscopic superpositions, which instead are predicted by the theory. This problem is solved in physics books by postulating the collapse rule: when we make a measurement the wavefunction randomly and instantaneously collapses into one of the terms of the superposition. There are some realists who think that there is no realism problem left: standard quantum theory with collapse

lays out a set of constraints imposed on the empirical data. These data exist objectively and mind-independently, so one could be realist about those, and not give too much importance of what is going on at the microscopic level. Proponents of the information-theoretic (IT) approach, QBists and Rainforest Realists can be thought as being realists of this sort.[1]

2.2 The Precision Problem

However, many find the collapse rule unsatisfactory: it is unclear when one is supposed to apply it, and why it applies. In other words, the realism problem is a *precision problem*: the wavefunction should evolve according to one single precise dynamics, whose definition should not appealing to vague concepts like measurement or observer. This is the measurement problem as usually understood. Its solutions need to be more than just empirically adequate: they also have to have a unique and precisely defined dynamics. As a consequence, standard quantum mechanics, which is empirically adequate but imprecise, is not adequate. The most promising solutions of this problem are recognized to be: the many-worlds theory[2]; the pilot-wave theory[3]; the GRW theory,[4] because they all have a unique and precise dynamics. The wavefunction is an ingredient in all these theories. There are different views about how to think about the wavefunction. One of them is *wavefunction realism*, according to which the wavefunction is not a field in spacetime but it lives in the high-dimensional configuration space.[5]

2.3 The Completeness Problem

Others instead have argued that one also needs the ontology to be in spacetime. Since the wavefunction is not a spatiotemporal object, then *all 'wavefunction-only'*

[1] See, respectively: Bub and Pitowsky (2010); Fuchs (2010), Fuchs, Mermin, and Schack (2014), DeBrota, and Stacey (2019); Ladyman and Ross (2007, 2013), Ladyman, (2016).

[2] Everett (1957).

[3] Bohm (1952).

[4] Ghirardi et al. (1986).

[5] Albert (1996, 2013, 2015), Ney (2012, 2013, 2015, 2017, 2021) and references therein. There are other views about how to think of the wavefunction. When the wavefunction is taken to be a vector in Hilbert space, it is called *vector space realism or Hilbert space fundamentalism* (see Carroll 2022 and references therein). If the wavefunction, or better the quantum state, is instead taken to be in its own category of existence, then one talks about *quantum state fundamentalism* (Maudlin 2019). In the many-worlds framework, some also have endorsed a view called *spacetime state realism*, which takes the states associated to spacetime regions as fundamental (Wallace and Timpson 2009). For a comparison between spacetime state realism, the many-world theory, and the primitive ontology approach, see Allori (2023).

quantum theories are fundamentally incomplete. If so, pure wavefunction theories like GRW and many-worlds are satisfactory: they need to be completed by a spatiotemporal ontology.[6] This idea has recently resurfaced in the primitive ontology framework.[7] In their view, material entities are not represented by the wavefunction but by the *primitive ontology*, which represents a suitable spatiotemporal ontology.[8] There are different ways of completing quantum theory with different spatiotemporal ontologies (particles, waves, spatiotemporal events or 'flashes', and so on), the simplest of which is given by the pilot-wave theory: the simplest ontology (particles), and the simplest evolutions (deterministic). One could propose different ontologies governed by a Schrödinger-evolving wavefunction as well as by a GRW-evolving wavefunction.[9]

[6] This attitude can be historically tracked down for instance to Lorentz, who objected to Schrödinger that his wavefunction was physically unacceptable because it is a field in configuration space, rather than in three-dimensions like electromagnetic fields (Lorentz in Przibram, 1967). Similar concerns were raised by Einstein (see Einstein's letters to Schrödinger and Ehrenfest in Howard 1990) (at least) an early Schrödinger (1926), and de Broglie (1927, reprinted in de Broglie 1956). Even Heisenberg expressed similar worries, which arguably pushed him toward anti-realism (Bloch 1976).

[7] Dürr, Goldstein, and Zanghí (1992); Dürr, Goldstein, and Zanghì (1997); Allori et al. (2008); Allori (2013a, 2013b). For a review, see Allori (2015); Allori (2019).

[8] Others have emphasized the importance of spacetime for a satisfactory ontology. For instance, Maudlin (e.g. 2016) has argued that satisfactory theories have local beables: "those which (unlike for example the total energy) can be assigned to some bounded space–time region" (Bell 1987). Also, Norsen (2010) has proposed that we should actively look for a theory entirely formulated in terms of spatiotemporal ontologies, without a wavefunction in high dimensional space. This turns out to be technically difficult, but perhaps the essence of this can be saved by understanding the wavefunction as a multi-field, or poly-wave, in three-dimensional space. This multi-field is an extension of the concept of field, as it assigns a number to a set of locations, rather than only one location, in the three-dimensional space (Forrest 1988; Belot 2012; Chen 2017; Hubert and Romano 2018; Romano 2020). Arguably, one could think of these views as solving the completeness problem, as all of them effectively add a spatiotemporal ontology to quantum theory (The disagreement between these views and primitive ontologists has to do with symmetries properties, as discussed in Allori 2021a).

[9] Examples of the latter include GRWp (in which particles evolves stochastically and nonlinearly with a law induced by a GRW-evolving wavefunction, see Allori 2020a); GRWm (in which the ontology is a matter field defined in terms of a GRW-evolving wavefunction, which therefore inherits a dynamics with the same stochastic and nonlinear features, see Benatti, Ghirardi and Grassi 1995); GRWf (which has a flashes ontology whose distribution is governed by a GRW-evolving wavefunction; this theory was first proposed by Bell (1987), and then adopted in Tumulka (2006), who developed a relativistic extension). Examples of the latter turn out to be theories with a many-worlds character: for instance, in Sm the matter field inherits the superpositions generated by the Schrödinger-evolving wavefunction (Allori et al. 2008, 2011; Allori 2019, and references therein).

3 Quantum Realisms

I have argued in the previous section that there is a disagreement about what the realism problem is supposed to be. The disagreement is about what dynamics, ontology and explanation count as satisfactory. In this section, I discuss different requirements for the ontology, while in the next section, I discuss different explanatory strategies.

3.1 Relaxed Realism

The approaches focusing on the adequacy problem think that quantum theory with collapse as more than just a recipe for predicting results. They think of it as a theory with a macroscopic ontology of measurement outcomes. Thus, the imprecision of the collapse rule does not matter: measurements are unanalyzable primitives[10] or there are no primitives.[11] Usually, these attempts are taken to be not realist enough: realists are supposed to care about more than just empirical adequacy. However, the proponents of these approaches maintain that either one should remain agnostic about which of the various possible microscopic processes is correct, or even that a microscopic description is needed. Let's call this perspective *relaxed realism:* a satisfactory theory needs a *spatiotemporal objective ontology,* which does not have to be microscopic.

3.2 Modest Realism

Those who care about the precision problem think that one needs all physical objects and processes being governed by a unified, precise dynamics. For instance, wave-function realists care about having the same dynamics for all scales. Nonetheless, they do not insist on the fundamental ontology to be in spacetime. Rather, they favor are open to pure-wavefunction theories, such as GRW and many-worlds, for reasons connected with locality and separability. Locality is a property of the interaction: influence travels at finite velocity. Instead, separability is a property of the ontology: the whole can be seen as the sum of its parts. These assumptions seems both commonsensical and needed to do physics.[12] However, quantum nonlocality has threatened them both.[13] Ney has argued that only a pure-wave ontology is local and

[10] As is the IT approach or in QBism.

[11] As in Rainforest Realism.

[12] See Ney (2021) for more.

[13] See, e.g., Goldstein et al. 2011. and references therein.

separable in the fundamental space.[14] Consequently, wavefunction realists endorse a *modest* form of realism: a satisfactory theory has *a local and separable ontology in the fundamental space.* They want more than empirical adequacy, but they do not require an ontology is spacetime.

3.3 Robust Realism

Primitive ontologists think that wavefunction-only quantum theories are incomplete: they require a fundamental spatiotemporal ontology, like relaxed realists, but they also want it to be microscopic because they want to think of macroscopic objects as composed of such fundamental ontology, in a Lego brick picture of reality.[15] Thus, they are *robust realists*: a satisfactory quantum theory has a suitably *microscopic (spatiotemporal) fundamental ontology.*

4 Quantum Explanations

Wavefunction realists argue that GRW, say, needs no completion because it solves the measurement problem by changing the dynamics. Primitive ontologists disagree because of the type of realism they endorse: robust realism requires an effectively microscopic (spatiotemporal) ontology, which GRW does not have. In contrast, wavefunction realists being modest realist, are willing to sacrifice a spatiotemporal ontology for a local and separable theory. Proponents of the IT interpretation are even more flexible, as they do not even care about having a unique and precise dynamics or a microscopic ontology. That's what makes them relaxed realists. In this section I discuss how the type of realism one endorses goes hand in hand with a specific type of explanatory strategy.

---

[14] If you have a spatiotemporal ontology, like in the pilot-wave theory, objects may be though as separable (they are made of particles), but they interact nonlocally (through the wavefunction). If instead you have a spatiotemporal ontology in a many-worlds theory (like in the case of spacetime state realism), the theory may be seen as local (the state of any spatiotemporal region depends only on the state of some cross-section of its past light cone) but it is non-separable (the intrinsic properties of a localized region of space are represented by a density operator, and the density operators of two subsystems do not determine the density operator of their union).

[15] This is explicit in Allori (2013a, 2013b, 2015, 2019). In the case of a particle ontology, the individual Lego bricks (the particle fundamental ontology) used to build a castle (a macroscopic object) are in the same space as the castle (spacetime) and they are smaller than the castle (they are microscopic).Notice that waves can be a suitable ontology for robust realists too, even if they are delocalized objects and, strictly speaking, they are not microscopic, as long as the following conditions are met. First, they are oscillating in (three-dimensional) space, evolving in time, and also they superimpose to form stable and localized wave-packets to reproduce particle-like observed behavior. These conditions being satisfied would make waves effectively microscopic. As a consequence, one could think of the Lego brick being (three-dimensional) wave-packets instead of particles.

According to Einstein, theories are either constructive or they are theories of principles.[16] *Principle theories* are formulated in terms principles, which are used as constraints on physically possible processes. Instead, *constructive theories involve the dynamical reduction* of macroscopic objects in terms of the motion and interactions of their microscopic constituents. An example of a principle theory is thermodynamics (it has principles such as e.g. "energy is conserved"), and an example of (the corresponding) constructive theory is statistical mechanics. Another example of principle theory is the 1905 theory of special relativity, as it was formulated in terms of the two principles: the equivalence of inertial frames for all physical laws, and the principle of constancy of the velocity of light.

4.1 Principle Theories

Some have argued that principle theories are preferable. For instance, Bub and Pitowsky (2010) have argued that quantum theory is a principle theory, and that to explain is to constrain the phenomena without the need of a dynamical account.[17] One could argue that a principle type of explanation is to be preferred because it provides a framework which is independent on the detailed assumption about the nature of matter. Moreover, Flores (1999) has argued that principle theories are explanatory because they unify.[18] Be that as it may, the preference of the proponents of the IT approach for principle theories fits well with their relaxed realism: a satisfactory theory systematizes the phenomena, and this can be successfully done using principles.

4.2 Hybrid Theories

Wavefunction realism has three steps to complete: first, recover three-dimensional space from high-dimensional space; then recover a microscopic non-fundamental ontology from the wavefunction; and finally, account for the macroscopic behavior. For the first step, wavefunction realists argue that three-dimensional space suitably emerges. There are various strategies, but they all have in common that they are based on principles. For example, Albert uses the principle that the privilege the dimensions in which the Hamiltonian is written. Ney instead uses the principle that privilege

[16] Einstein (1919).

[17] "There is no deeper explanation for the quantum phenomena of interference and entanglement than that provided by the structure of Hilbert space, just as there is no deeper explanation for the relativistic phenomena of Lorentz contraction and time dilation than that provided by the structure of Minkowski space–time"(*ibid*.).

[18] Friedman (1974), Kitcher (1989).

the dimensions that respect the fundamental symmetries of the dynamics.[19] These approaches use principles in the second step as well, namely for the emergence of non-fundamental microscopic objects. Albert and Loewer (1996) propose to modify the so-called EER (eigenvalue-eigenstate rule) of standard quantum mechanics as to define particles as suitably emergent. Later, Albert (2015) proposed that particles as we experience them are to be understood as emerging as 'functional shadows' of the high-dimensional fundamental wavefunction. Once we have this microscopic non-fundamental ontology, we can understand macroscopic objects as composed of them. So, to complete step two, Albert uses principles, but in step three he employs a constructive explanation. Ney (2021) instead first derives microscopic (spatiotemporal) particles as the derivative parts of the wavefunction: there is a particle when there is a 'bump' in the squared of the wavefunction. If so, a particle location is indeterminate, as the wavefunction may be spread out. Particles may partially instantiate different locations to different degrees, given by the squared amplitude of the wave function in that point. Then, in contrast with Albert, Ney does not think of macroscopic objects as composed of microscopic particles. Rather, she thinks that decoherence, namely the interaction of the environment, is responsible for the emergence of macroscopic, classically behaving patterns, along the lines of strategies adopted by supporters of the many-worlds theory.[20]

These explanatory strategies are not constructive: the macroscopic phenomena are not explained in terms of the fundamental spatiotemporal microscopic dynamics. Rather, they seem much closer to the type of explanation provided by principle theories. Albert uses the Hamiltonian and functionalism, Ney uses symmetries and partial instantiation to establish the derivative reality of three-dimensional reality, thereby providing principles that constrain the phenomena. On the other hand, these are not entirely principle explanations. In fact, wavefunction realists care about the dynamics in virtue of wanting to solve the precision problem. Moreover, Albert's functionalist account takes the dynamics into account using the form of the Hamiltonian (rather than its solutions) to recover three-dimensional objects. Similarly, Ney's focus on symmetries can be understood as taking the dynamics seriously, as well as her appeal to decoherence. So, at the end, the type of explanation provided by wavefunction realism is a hybrid between principle, compositional and dynamical explanation.

4.3 Constructive Theories

Primitive ontologists favor constructive theories.[21] Brown (2005) has argued that constructive theories are more explanatory than principle theories because not only

[19] A similar strategy is employed in vector space realism, in which three-dimensionality is recovered in terms of the energy eigenvalues of the Hamiltonian. See Carroll (2022).

[20] See Wallace (2012).

[21] . They follow Einstein, who believed that physics should look for constructive theories, and accept principle theories only when one has no other option. He thus expressed his own dissatisfaction for the theory of special relativity at the time. However, he could have said something similar for

they accounts for what we should expect to happen, but also for why it happens.[22] The essence of constructive explanation is to explain compositionally and dynamically. So, primitive ontologists treat standard quantum theory as thermodynamics. They are both principle theories: the quantum recipes describe the phenomena by specifying the statistics of the experimental results, just as thermodynamics provides constraints on macroscopic phenomena. As such, they have a constructive counterpart, to which we can reduce them. The constructive counterpart of thermodynamics is classical mechanics, by thinking of gases as collections of particles. What is the constructive counterpart for standard quantum mechanics? This theory would allow us to understand not only which quantum principles hold, but why they do. Notice therefore that, as it would be absurd to use a gas ontology for classical mechanics to constructively explain thermodynamics, *one should not use the wavefunction as the ontology* of the constructive quantum theory: since the wavefunction 'belongs' to the principle theory, it does not make sense to use it as the ontology for the constructive one. In addition, since constructive explanation requires a spatiotemporal fundamental ontology, the obvious choice is the one of particles. For once, particles seem more compatible with the empirical evidence of tracks in detectors. If so, the straightforward constructive counterpart of standard quantum mechanics is the pilot-wave theory.

5 Final Remarks

For each approach I have identified the driving motivation: preserving principle explanation, preserving locality and separability, and preserving constructive explanation. In this section, I wish to evaluate these accounts in their own terms.

5.1 Puzzles for Relaxed Realists

As anticipated, relaxed realist attempts are taken to be *not realist enough*.[23] In my opinion, however, this type of objections is driven by the idea that to be realist is to look for a constructive rather than a principle explanation. In response, proponents of the IT approach deny, as we have seen, that we need constructive understanding. However, many remain unconvinced that we should settle for principle type of explanations, especially if one could have also a constructive understanding.

---

quantum theory: his preference for constructive theories is compatible with his idea that quantum mechanics is incomplete. Moreover, it fits well with his statistical interpretation of quantum theory, as it is a principle theory by constraining the phenomena with suitable rules.

[22] See also Brown (2005), Brown and Pooley (2004), Brown and Timpson (2006). See Flores (1999) and Felline (2011) for a connection with Salmon's mechanistic view of explanation (1984).

[23] For example, Egg (2019) has put forward a set of arguments that some implementations of this type of realism do not deserve to be labelled realist.

Another feature of these approaches that contributed to some confusion is that the wavefunction is not seen as 'ontic,' namely as describing some physical facts, because only measurement results are taken ontologically seriously. As a consequence, these approaches make the move of thinking of *the wavefunction as epistemic*: roughly, it encodes our knowledge of the system. Nonetheless, this move makes the approach mind dependent. If you have a theory containing something expressing your lack of complete information, you are acknowledging that the theory is incomplete. However, while it is true that in these approaches the wavefunction does not represent physical objects, *it does not follow that it cannot be ontic.* In fact, the label 'ontic' just means that it is an objective ingredient of the framework. So, the wavefunction could for instance represent a fact describing the interaction between systems, like the one expressed by the Hamiltonian. This is straightforwardly compatible with these principle approaches, without making the wavefunction epistemic.[24]

5.2 Puzzles for Modest Realists

A problem for wavefunction realists has to do with their motivation. How do they justify their attachment to separability and locality? As anticipated, their answer is that they are intuitive. However, the locality in wavefunction realism is high-dimensional locality, so quantum theory is still nonlocal in three-dimensional space. The wavefunction realist points out that locality is needed to explain physical action. Nonetheless, *this move is effective only if we think of three-dimensional locality.* Moreover, people want locality because this is compatible with relativity: since in relativity there is a maximum velocity there cannot be instantaneous action at a distance. However, again, this is three-dimensional locality. Similarly, the wavefunction realist cares about separability because it preserves our intuitive notion of compositionality: if separability is true, composed systems can be broken down into simpler ones. Nonetheless, the separability we care about is in three-dimensions, as in the case of locality.

Another problem is that within wavefunction realism phenomena are not accounted from entirely in a constructive way. In fact, one first has to derive three-dimensionality, then particles, and then macroscopic objects. Each of these steps (except the last, in the case of Ney's approach) involves some sort of principle, like relaxed realists. But didn't they claim that they wanted more? Both approaches essentially systematize the phenomena, with the only difference that wavefunction realists systematizes the data precisely. So, one may wonder why we should care about precision at all: what is it to be gained by systematizing the phenomena precisely? The point of solving the measurement problem was that there seemed to be value in having a precise dynamics. But what is this value if one only cares about reproducing the appearances? Why do wavefunction realist value the dynamics *if they provide a non-dynamical explanation?* Having a non-constructive explanation and giving

[24] See Allori (2020b, 2021b).

importance to the dynamics seem to pull in opposite directions: the former pushes towards relaxed realism, while the latter towards a robust realism, making wavefunction realism a peculiar hybrid. There is a tension between the desire of the wavefunction realist of a robust kind of realism, and the kind of explanation wavefunction realism actually provides, which is not constructive: the wavefunction realist starts as off a robust realist, but she ends up (too) close to the relaxed realist. One can defend wavefunction realism (at least in the case of Albert) by observing that, as a matter of fact, the principles are needed to recover three-dimensionality and microscopic objects, but once these emerge, we can think of macroscopic objects effectively as if they are composed of microscopic entities. That is, constructive explanation still holds. That is, it is not the only step needed to recover the phenomena, but one of the two: principles first, constructive explanation next.

Notice that having at least one constructive step is essential. In fact, consider thermodynamics and statistical mechanics. According to many, the laws of thermodynamics could be constructively accounted for in terms of classical statistical mechanics. If wavefunction realism does not have a constructive step, then it becomes difficult to see how they can accept statistical mechanics: we can think of gases as composed by particles only if these particles emerge microscopically. The principles need to be used only to go from the high-dimensional space to the three-dimensional one, and not after the microscopic particles have emerged. In other words: if someone wishes quantum theories to be explanatory in terms of principles only, then they should not be too attached to a constructive understanding. However statistical mechanics constructively explains thermodynamics, and this arguably extends also to quantum statistical mechanics. But if the explanation provided by wavefunction realism is non-constructive, then there is a tension. How are these two explanations compatible?

5.3 Puzzles for Robust Realists

In a robust realist quantum theory, the description provided by the wavefunction is never complete, and the fundamental ontology of the theory is spatiotemporal and microscopic. As we have seen, this is straightforwardly accomplished by the pilot-wave theory. Nonetheless, robust realists such as primitive ontologists engage with other quantum theories, and they seem to treat them as equally acceptable robust realist alternatives. That is, *instead of starting from the completeness problem, they focus on the measurement problem.* But this is puzzling, as they also argue that all robust realist quantum theories have to have a spatiotemporal ontology, and two out of three ways of solving the measurement problems do not respect such requirement: GRW and many-worlds do not solve the completeness problem. These theories solve the measurement problem in ways other than completing standard quantum theory. So why are they even an option for the primitive ontologists? Admittedly, primitive ontologists insist that these theories need to be supplemented by a spatiotemporal ontology. However, they are bound to be artificial, as *they were meant to solve the*

measurement problem without solving the completeness problem.[25] What is it to be gained in taking these theories and then turning them into solutions of the completeness problem by supplementing them with a spatiotemporal ontology, especially given that one already has the simplest way of doing that, namely the pilot-wave theory?

Primitive ontologists reply that these theories are mere examples of constructive theories, not as real alternatives: they lack either simplicity or motivation.[26] Moreover, one could argue that the value of looking at GRW-type theories with a spatiotemporal ontology is to explore the *compatibility with quantum mechanics and relativity*.[27] In fact, in addition to the Lorentz invariance of the laws, people disagree about what it means that a theory is compatible with relativity. Some have argued that a theory is compatible with relativity if it is formulated only with relativistic spatiotemporal structure. Lorentz invariant extensions of the pilot-wave theory all have a preferred slicing of spacetime, namely a foliation, and since this is not a relativistic spatiotemporal structure these theories are all in tension with relativity.[28] Instead some Lorentz invariant extension of GRW theories with a spatiotemporal ontology (like GRWm and GRWf) require no preferred foliation. This is the sense in which they are more compatible with relativity than the pilot-wave theory.[29] This, GRW-type theories seem worth exploring. However, it is unclear whether these theories can truly accomplish this task.[30]

Additionally, the status of the wavefunction in this framework is unclear. According to Maudlin (2019), the quantum state, represented by the wavefunction, represents some objective fact about the world which is best understood as unanalyzable. However, this primitivist categorization does not straightforwardly fit with the constructive schema. What is the role of the quantum state within the constructive schema? A better fit with the constructive understanding, I think, are the approaches which regard the wavefunction as nomological: "the wave function tells the matter

[25] GRW was developed to unify a wavefunction dynamics which could eliminate unobserved macroscopic superpositions without appealing to measurements, while the many-worlds theory was developed to maintain the Schrodinger dynamics, which consequently lead to recognize the existence of macroscopic superpositions.

[26] For instance, take a many-worlds theory like theory like Sm. This theory predicts macroscopic superpositions because the matter density field inherits the superpositions of the wavefunction. There is no reason for the primitive ontologists to endorse a theory with a many-worlds character, as it is in contrast with the spirit of constructive explanation. So, Sm is not a real contender. Since the many-worlds character of Sm results from the linearity of the Schrodinger evolution, which is inherited by the evolution of the matter field, arguably a theory like GRWm has not the same many-worlds character, at least not macroscopically. Anyway, less controversial is the case for GRWp of GRWf, since particle or spatiotemporal events cannot superimpose.

[27] Allori (2020a).

[28] See most notably Berndl et al. (1996), Dürr et al. (1999), Dürr et al. (2013).

[29] See Tumulka (2006, 2020) for a Lorentz invariant GRWf without and with interaction; see Bedingham et al. (2014) for a Lorentz invariant GRWm theory.

[30] Allori (2022).

how to move."[31] My suggestion, dubbed *wave-functionalism*, is that the wavefunction should be understood as one way of realizing one of the ingredients needed to implement the dynamics for the fundamental ontology. In a slogan, the wavefunction is the functions it plays in the theory.[32] This is compatible with a constructive explanation: the wavefunction does not represent the Lego bricks the castle is made of, but it represents the rules in the booklet to guide us constructing the castle.

As a final remark, being the approach constructive, laws of nature should be seen as describing the evolution of a physical system through time.[33] In fact, a constructive approach has two components: matter, understood in terms of Lego bricks, and the laws, understood as instructions about how these bricks can stick together. Instead, Chen and Goldstein (2022) have recently proposed an account of laws of nature as constraint: "laws explain, but not by accounting for the dynamic production of successive states of the universe from earlier ones." In this view, laws substantially play the role of principles: they exclude what cannot happen without further explaining why that is. If so, this view is in direct contrast with the constructive understanding. This creates a tension for Goldstein, who in the past have defended the primitive ontology approach. If he endorses a principle understanding of laws, then he undermines his case for a spatiotemporal ontology on the basis of constructive explanation. Conversely, if he finds laws as constrains as not problematical, then he has to provide additional reasons to favor constructive explanation and to require a microscopic spatiotemporal ontology.

6 Conclusion

In this paper I have argued that, contrarily of the common understanding, it is not obvious what problem one needs to solve to make standard quantum mechanics amenable to a realist interpretation. Wavefunction realists think it is the measurement problem, namely the problem of precisely eliminating unobserved macroscopic superpositions. In their quest for a local and separable ontology, they require a unique and precise dynamics, but not a spatiotemporal ontology. However, not everyone agrees. For instance, the proponents of the IT approach argue that standard quantum mechanics with collapse can be interpreted from a realist perspective, as a realist theory for them only needs to be empirically adequate. In this relaxed type of realism, a theory explains in terms of principles. In contrast, primitive ontologists require a constructive explanation in terms of a microscopic fundamental ontology. This leads them to endorse a robust type of realism and to reject all theories without a microscopic ontology.

Even if I am partial to a constructive understanding, the purpose of this paper was not to argue for it. Rather it was to understand better where the disagreement

[31] Allori et al. (2008).

[32] Allori (2021b).

[33] Maudlin (2007).

was coming from. In fact, by clarifying that, one can better fine-tune their arguments. For instance, to say to a proponent of the IT approach that they need a microscopic ontology because constructive explanation fails in their approach will certainly not change their mind: they do not care about constructive explanation. Rather, to convince them to adopt another view one would have to argue that, for instance, principle theories are not explanatory enough. Moreover, after having tried to distill the motivations behind the IT approach, I could make a better case for it by showing that one can have principle theories with a non-epistemic wavefunction. Similarly, it is ineffective to point out to the wavefunction realists that they need a spatiotemporal ontology: such an ontology will not preserve locality, which is their guiding principle. Instead, to convince them that they approach fails, for instance, one could propose a local constructive theory (which would therefore have a spatiotemporal ontology) presumably even if such proposal were retrocausal or superdeterministic.[34] Moreover, a better case for wavefunction realism could be made by motivating locality and separability differently, or justifying a high-dimensional ontology in some other way, other than via separability and locality. Finally, pointing out that the pilot-wave theory has a foliation is not going to convince primitive ontologists: they are going to argue that other constructive theories do not have it. Instead, a better argument against the primitive ontology approach would be one which shows that constructive explanation is impossible for some reason. Conversely, one could improve the arguments for the primitive ontology approach by showing that this constructive understanding can be extended to more general theories like relativistic quantum field theories.

Be that as it may, the bottom line is that only after there is mutual appreciation of the alternative views, one can start making progress efficiently, avoiding situations in which the two sides talk past one another. Hopefully, with this paper I have contributed to this.

Acknowledgements I am grateful to the participants of the *Reading Group in Quantum Foundations* organized by the Quantum Engineering group at the University of Grenoble Alpes (March 2021), of the *VII International Workshop on Quantum Mechanics and Quantum Information: Quantum Ontology and Metaphysics* (March 2021), and to an anonymous reviewer for their insightful feedback on some of the material presented in this paper.

References

D.Z. Albert, Elementary quantum metaphysics, in *Bohmian Mechanics and Quantum Theory: An Appraisal*, ed. by J. Cushing, A. Fine, S. Goldstein (Kluwer, 1996), pp. 277–284

D.Z. Albert, Wave-function realism, in *The Wave-Function.* ed. by D.Z. Albert, A. Ney (Oxford University Press, New York, 2013), pp.52–57

D.Z. Albert, *After Physics* (Harvard University Press, Cambridge, 2015)

D.Z. Albert, B. Loewer, Tails of schrödinger's cat, in *Perspectives on Quantum Reality.* ed. by R. Clifton (Kluwer, Dordrecht, 1996), pp.81–91

[34] See, for instance Ciepielewski, Okon and Sudarsky (2020).

V. Allori, On the metaphysics of quantum mechanics, in *Precis de la Philosophie de la Physique*, ed. by S. Lebihan (Vuibert, 2013a), pp. 116–151
V. Allori, Primitive ontology and the structure of fundamental physical theories, in *The Wavefunction: Essays in the Metaphysics of Quantum Mechanics*, ed. by D.Z. Albert, A. Ney (Oxford University Press, 2013b), pp. 58–75
V. Allori, Primitive ontology in a nutshell. Int. J. Quantum Found. **1**(3), 107–122 (2015)
V. Allori, Scientific realism without the wave-function, in *Scientific Realism and the Quantum*, ed. by J. Saatsi, S. French (Oxford University Press, 2019)
V. Allori, Spontaneous localization theories with a particle ontology, in *Do Wave Functions Jump? Perspectives on the Work of GianCarlo Ghirardi*, ed. by V. Allori, A. Bassi, D, Dürr, N. Zanghì (Springer, Berlin, 2020a)
V. Allori, Why scientific realists should reject the second dogma of quantum mechanics, in *Quantum, Probability, Logic: The Work and Influence of Itamar Pitowsky*, M. Hemmo, O. Shenker (Springer, Berlin, 2020b), pp. 19–48
V. Allori, Primitive beable is not local ontology: on the relation between primitive ontology and local beables. Crítica **53**(159), 15–43 (2021a)
V. Allori, Wave-functionalism. Synthese **199**, 12271–12293 (2021b)
V. Allori, What is it like to be a GRW theory? quantum mechanics and relativity, still in conflict after all these years. Found. Phys. **52**, 79 (2022)
V. Allori, "Many-Worlds: Why there is No Consensus." *Quantum Reports* (2023).
V. Allori, S. Goldstein, R. Tumulka, N. Zanghí, On the common structure of Bohmian mechanics and the Ghirardi-Rimini-Weber theory. Br. J. Philos. Sci. **59**(3), 353–389 (2008)
V. Allori, S. Goldstein, R. Tumulka, N. Zanghí, Many-worlds and Schrödinger's first quantum theory. Br. J. Philos. Sci. **62**(1), 1–27 (2011)
J.S. Bell. Speakable and unspeakable in quantum mechanics: collected papers on quantum philosophy (Cambridge University Press 1987)
D. Bedingham, D. Dürr, G.C. Ghirardi, S. Goldstein, R. Tumulka, N. Zanghí, Matter density and relativistic models of wavefunction collapse. J. Stat. Phys. **154**(1–2), 623–631 (2014)
G. Belot, Quantum states for primitive ontologists. Eur. J. Philos. Sci. **2**(1), 67–83 (2012)
F. Benatti, G.C. Ghirardi, R. Grassi, Describing the macroscopic world: closing the circle within the dynamical reduction program. Found. Phys. **25**, 5–38 (1995)
K. Berndl, D. Dürr, S. Goldstein, N. Zanghì, Nonlocality, Lorentz invariance, and Bohmian quantum theory. Phys. Rev. A **53**, 2062–2073 (1996)
D. Bohm, A suggested interpretation of the quantum theory in terms of "Hidden" Variables" I and II. Phys. Rev. **85**, 166–193 (1952)
F. Bloch, Heisenberg and the early days of quantum mechanics. Phys. Today (1976)
H.R. Brown, *Physical Relativity: Space-Time Structure from a Dynamical Perspective* (Oxford University Press, Oxford, 2005)
H.R. Brown, O. Pooley, Minkowski space-time: a glorious non-entity, in *The Ontology of Spacetime*, ed. by D. Dieks (Elsevier, 2004), pp. 67–89
H.R. Brown, C. Timpson, Why special relativity should not be a template for a fundamental reformulation of quantum mechanics, in *Physical Theory and its Interpretation: Essays in Honor of Jeffrey Bub*, ed. By W. Demopoulos, I. Pitowsky (Springer, Berlin, 2006), pp. 29–42
J. Bub, I. Pitowsky, Two Dogmas about Quantum Mechanics, in *Many Worlds? Everett, Quantum Theory, and Reality*, S. Saunders, J. Barrett, A. Kent, D. Wallace (Oxford University Press, 2010), pp. 433–459
S. Carroll, Reality as a vector in Hilbert space, in *Quantum Mechanics and Fundamentality: Naturalizing Quantum Theory between Scientific Realism and Ontological Indeterminacy*, ed. by V. Allori (Springer Synthese Library, 2022)
E.K. Chen, Our fundamental physical space: an essay on the metaphysics of the wave function. J. Philos. **114**(7), 333–365 (2017)

E.K. Chen, S. Goldstein, Governing without a fundamental direction of time: minimal primitivism about laws of nature, in *Rethinking the Concept of Law of Nature*, ed. by Y. Ben-Menahem. Jerusalem Studies in Philosophy and History of Science (Springer, Cham, 2022)
G.S. Ciepielewski, E. Okon, D. Sudarsky, On superdeterministic rejections of settings independence. Br. J. Philos. Sci. (2020). https://doi.org/10.1086/714819
L. De Broglie, La mécanique ondulatoire et la structure atomique de la matière et du rayonnement. J. de Phys. et Le Radium **8**(5), 225–241 (1927)
L. De Broglie, *Une Tentative d'Interprétation Causale et Nonlinéaire de la Mécanique Ondulatoire.* Paris: Gauthier-Villars. English translation: *Non-linear Wave Mechanics: A Causal Interpretation* (Elsevier, Amsterdam, 1956, 1960)
J.B. DeBrota, B.C. Stacey FAQbism (2019). https://arxiv.org/pdf/1810.13401.pdf
D. Dürr, S. Goldstein, K. Münch-Berndl, N. Zanghì, Hypersurface bohm-dirac models. Phys. Rev. A **60**, 2729–2736 (1999)
D. Dürr, S. Goldstein, T. Norsen, W. Struyve, N. Zanghì, Can bohmian mechanics be made relativistic? Proc. R. Soc. A **470**, 20130699 (2013)
D. Dürr, S. Goldstein, N. Zanghí, Quantum equilibrium and the origin of absolute uncertainty. J. Stat. Phys. **67**, 843–890 (1992)
D. Dürr, S. Goldstein, N. Zanghì, Bohmian mechanics and the meaning of the wave function, in *Boston Studies in the Philosophy of Science*, ed. by R.S. Cohen, M. Horne, J. Stachel, vols. 1, 193. Experimental Metaphysics—Quantum Mechanical Studies for Abner Shimony (Kluwer Academic Publishers, Boston, 1997), pp. 25–38
A. Einstein, What is the Theory of Relativity? The London Times; reprinted in A. Einstein, *Ideas and Opinions: 227–232* (Crown Publishers Inc., New York, 1919, 1982)
M. Egg, Dissolving the measurement problem is not an option for the realist. Stud. Hist. Philos. Sci. Part B: Stud. Hist. Philos. Mod. Phys. **66**, 62–68 (2019)
H. Everett, Relative state formulation of quantum mechanics. Rev. Mod. Phys. **29**, 454–462 (1957)
L. Felline, Scientific explanation between principle and constructive theories. Philos. Sci. **78**(5), 989–1000 (2011)
F. Flores, Einstein's theory of theories and types of theoretical explanation. Int. Stud. Philos. Sci. **13**(2), 123–134 (1999)
M. Friedman, Explanation and scientific understanding. J. Philos. 5–19 (1974)
P. Forrest, *Quantum Metaphysics* (Blackwell, Oxford, 1988)
C. Fuchs, QBism, the Perimeter of Quantum Bayesianism (2010). https://arxiv.org/pdf/1003.5209.pdf
C. Fuchs, N.D. Mermin, R. Schack, An introduction to QBism with an application to the locality of quantum mechanics. Am. J. Phys. **82**(8), 749–754 (2014)
G.C. Ghirardi, A. Rimini, T. Weber, Unified dynamics for microscopic and macroscopic systems. Phys. Rev. D **34**, 470 (1986)
S. Goldstein, T. Norsen, D.V. Tausk, N. Zanghì, Bell's theorem. Scholarpedia **6**(10), 8378 (2011)
D. Howard, Nicht Sein Kann Was Nicht Sein Darf, or the prehistory of EPR, 1909–1935: Einstein's early worries about the quantum mechanics of composite systems, in *Sixty-Two Years of Uncertainty*, ed. by A.I. Miller (Plenum Press, New York, 1990)
M. Hubert, D. Romano The wave-function as a multi-field. Eur. J. Philos. Sci. 1–17 (2018)
P. Kitcher, Explanatory unification and the causal structure of the world, in *Minnesota Studies in the Philosophy of Science*, vol. 13, ed. by P. Kitcher, W. Salmon (University of Minnesota Press, 1989), pp. 410–503
J. Ladyman, Are there individuals in physics, and if so, what are they? In *Individuals Across the Sciences*, ed. by A. Guay, T. Pradeu (Oxford University Press, Oxford, 2016), pp. 193–206
J. Ladyman, D. Ross, *Every Thing Must Go: Metaphysics Naturalized* (Oxford University Press, Oxford, 2007)
J. Ladyman, D. Ross, The world in the data, in *Scientific Metaphysics*. ed. by D. Ross, J. Ladyman, H. Kincaid (Oxford University Press, Oxford, 2013), pp.108–150
T. Maudlin, *The Metaphysics Within Physics* (Oxford University Press, 2007)

T. Maudlin, Local beables and the foundations of physics, in *Quantum Nonlocality and Reality 50 Years of Bell's Theorem*, ed. by M. Bell, S. Gao (Cambridge University Press, 2016), pp. 317–330
T. Maudlin, *Philosophy of Physics: Quantum Theory* (Cambridge University Press, 2019)
A. Ney, The status of our ordinary three-dimensions in a quantum universe. Noûs **46**, 525–560 (2012)
A. Ney, Ontological reduction and the wave-function ontology, in D.Z. Albert, A. Ney op. cit.: (2013), pp. 168–183
A. Ney, Fundamental physical ontologies and the constraint of empirical coherence. Synthese **192**(10), 3105–3124 (2015)
A. Ney, Finding the world in the wave-function: some strategies for solving the macro-object problem. Synthese 1–23 (2017)
A. Ney, *Finding the World in the Wave Function* (Oxford University Press, 2021)
T. Norsen, The theory of (exclusively) local beables. Found. Phys. **40**(12), 1858–1884 (2010)
K. Przibram, *Letters on Wave Mechanics: 44; translated, ed. by M. Klein* (Philosophical Library, NY, 1967)
D. Romano, Multi-field and Bohm's theory. Synthese (2020). https://doi.org/10.1007/s11229-020-02737-6
W.C. Salmon, *Scientific Explanation and the Causal Structure of the World* (Princeton University Press, 1984)
E. Schrödinger, Quantisierung als Eigenwertproblem (Zweite Mitteilung). Annalen der Phys. **79**, 489–527 (1926). English translation: Quantisation as a Problem of Proper Values. Part II
R. Tumulka, A relativistic version of the Ghirardi-Rimini-Weber model. J. Stat. Phys. **125**, 821–840 (2006)
R. Tumulka, A relativistic GRW flash process with interaction, in *Do Wavefunctions Jump? Perspectives of the Work of GianCarlo Ghirardi*, ed. V. Allori, A, Bassi, D. Dürr, N. Zanghí (Springer, Berlin, 2020), pp. 321–348
D. Wallace, *The Emergent Multiverse: Quantum Theory according to the Everett Interpretation* (Oxford University Press, 2012)
D. Wallace, C. Timpson, Quantum mechanics on spacetime i: spacetime state realism. Br. J. Philos. Sci. **61**(4), 697–727 (2009)

The Changing Bell View of Beables: A Forgotten Story

Federico Laudisa

Abstract John S. Bell is known, among other things, for the introduction of the notion of *beable*. The development of this notion inspired the so-called primitive ontology (PO) approach to the foundations of quantum mechanics, proposed for the first time by Detlef Dürr, Sheldon Goldstein and Nino Zanghì in 1992. It is not very well known, however, that the Bell theory of beables had an early formulation, in which Bell curiously adopts some Bohr-reminiscent insights to attack exactly the standard Copenhagen version of quantum mechanics. Here I reconstruct the two stages of the Bell theory of beables, showing that the first stage is in fact unable to adequately confront the foundational problem it was designed to address. Only the second stage of the Bell theory could represent a motivation for the PO approach: in this respect, it may be of some interest to compare the two-stage reconstruction of the Bell theory with recent analyses of the PO approach in terms of beables. I dedicate this paper to the dear memory of Detlef Dürr, a leading figure of the international community of the foundations of physics and a lovely man. He will be long remembered for his inspiring contributions: it was a privilege to enjoy his company and his doctrine.

1 Introduction

John Stewart Bell is unanimously recognized as one of the leading figures, if not *the* leading figure, of the foundational debate on quantum mechanics (QM) since the second half of the twentieth-century. He is also acknowledged as a fierce and relentless enemy of Copenhagenish approaches to QM: as is well known, his critical attitude toward any purely operational and instrumental understanding of quantum principles led him to encourage alternative views, ranging from Bohmian mechanics (starting from Bell 1966, his first work concerning the hidden variables' issue) to (idiosyncratic) forms of the Everett interpretation (Bell 1976), up to an explicit support to

F. Laudisa (✉)
Department of Humanities and Philosophy, University of Trento, Via Tommaso Gar 14, 38122 Trento, Italy
e-mail: federico.laudisa@unitn.it

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_29

the so-called dynamical reduction model, or GRW version, of QM in the latest part of his career (1987, 1989, 1990). One of the most provocative proposals on the Bell part has been the introduction of the notion of *beable*, a term first introduced in 1973 with the specific aim of addressing what Bell took to be an intrinsic ambiguity in the quantum description of observation:

> This terminology, *be*-able as against *observ*able, is not designed to frighten with metaphysic those dedicated to realphysic. It is chosen rather to help in making explicit some notions already implicit in, and basic to ordinary quantum theory." (Bell 1975, in Bell 2004[2], p. 52).

The claim that a 'theory of' beables was needed, and its connection with the issue of locality, were the focus of the seminal papers of the Seventies in which Bell started to elaborate on the notion of beable. At the time the suggestion had not been taken too seriously, but the foundational role of beables has surfaced again in more recent times, when this notion turned out to be at the source of a true research program, the *primitive ontology* (PO) approach, in the area of the foundations and interpretations of quantum mechanics.

This approach was originally proposed by Detlef Dürr, Shelly Goldstein and Nino Zanghì (Dürr et al. 1992). It emphasizes the need, for a well-founded theory, to specify in ontologically clear terms the kind of entity the theory itself is primarily supposed to account for. As to the notion of beable, it was proposed for the first time as the expression of an attitude toward the foundations of quantum mechanics inspired to (some form of) scientific realism, but at that time no PO approach was available yet. It was clear that the proposal of a notion of beable by Bell was an expression of dissatisfaction toward the standard formulation of quantum mechanics, but it is far from transparent what the anti-instrumentalistic role, assigned by Bell to that notion, should have been exactly. I wish to show that there are at least *two* different readings of the notion of beable in the development of Bell's foundational analyses, corresponding to an evolution in time of the interpretation that Bell provides for the notion itself. At an early stage, the concept of beable emerges as the consequence of a peculiar Bohrian-sounding view of the status and role of measurement in QM: within this view Bell, across several of his papers devoted to the foundations of QM, repeatedly and instrumentally exploits Bohr in different places, in order to support claims that in fact are meant to undermine the Copenhagen formulation of quantum mechanics. In this sense, Bell appears ironically to be using a Bohrian insight as a weapon *against* standard QM! I will stress that this early formulation of the notion of beable, in spite of Bell's aspirations toward a less unsatisfactory interpretation of QM, is in fact unable to improve upon the ambiguity of standard QM concerning, for instance, the description of the measurement process. Only later the Bell interpretation of the notion of beable evolves more explicitly into a second, more focused formulation. I will emphasize that it is this new formulation that is apt to intertwine with the locality/non-locality issue arising from the formulation of the 1964 Bell theorem. In retrospect, therefore, we can recognize in this *second* stage of the Bell formulation of the notion of beable one strong motivation for the PO approach to the foundations of quantum mechanics. Then the possibility arises to assess the relation between the two-stage development of the Bell notion of beable

and the further evolution of the PO approach, in light of the complex relationships between the latter and the former.

2 The Early History of Beables: Bell and Bohr

The first occurence of the term *beable* can be found in a short, programmatic Bell paper entitled "Subject and object" and published in 1973 (Bell 1973, in Bell 2004[2], pp. 40–44). First of all, the paper has a telling title. Bell decides to address the central role assigned to measurement in the standard formulation of quantum mechanics in terms of a distinction—that between 'subject' and 'object'—that has a *philosophical* flavor.[1] By pairing an *object* with a *measured system* and a *subject* with a *measurer*, Bell charges the standard formulation of quantum mechanics with a kind of subjectivism, according to which the theory is bound to retain a fundamental vagueness and ambiguity on where the boundary between subject and object is supposed to be located, no matter how good for practical use the theory is:

> The subject-object distinction is indeed at the root of the unease that may people still feel in connection with quantum mechanics. […] In extremis the subject-object division can be put somewhere at the 'macroscopic' level, where the practical adequacy of classical notions makes the precise location quantitatively unimportant. But although quantum mechanics can account for these classical features of the macroscopic world as very (very) good approximations, it cannot do more than that. The snake cannot completely swallow itself by the tail. This awkward fact remains: the theory is only *approximately* unambiguous, only *approximately* self-consistent. (Bell 1973, in Bell 2004[2], pp. 40–41, emphasis in the original).

It is in expressing his hope in a less-and-less ambiguous formulation that Bell introduces for the first time the term *beable*:

> […] it should again become possible to say of a system not that such and such may be *observed* but that such and such *be* so. The theory would not be about '*observ*ables' but about '*be*ables'." (Bell 1973, in Bell 2004[2], p. 41).

Here in "Subject and object", Bell does not elaborate as precisely as one could wish a *theory* of beables, but we can interpret his wording as suggesting at least two conditions that such a theory should satisfy:

(i) although the use of the notion of beable cannot simply amount to make quantum mechanics a classical theory in any sense, a theory of beables should account for "an image of the everyday classical world", namely they should enable us—as middle-size natural systems—to recover our subjective experience;

(ii) at the same time, a theory of beables should justify the idea that beables somehow *ground*, or, even better, *constitute* observables: as Bell says with a sort of

[1] According to a Bell biographer, the very title was a choice of the organizers of the conference in which the paper was first presented (Whitaker 2016, p. 290), but Bell employs the distinction with a conscious purpose.

'metaphysical' tone, "the idea that quantum mechanics is primarily about 'observables' is only tenable when such beables are taken for granted. Observables are *made out of* beables." (Bell 1973, in Bell 2004[2], p. 41).

Surprisingly, in order to support the plausibility of beables Bell appears to rely on a well-known passage of Niels Bohr, taken from the Bohr contribution to the 1949 celebrated volume *Albert Einstein Philosopher-Scientist*:

> […] it is decisive to recognize that, *however far the phenomena transcend the scope of classical physical explanation, the account of all evidence must be expressed in classical terms.*" (Bohr 1949, p. 209, emphasis in the original).

Bell suggests not only that his notion of beable does justice to the Bohr plea for an account of evidence in classical terms, but also that—if formulated in terms of the beables' theory—such plea can be put to work in order to solve the above mentioned problem generated by the inherent ambiguity and approximation of standard quantum mechanics. The Bell suggestion is ironical, since it uses a major claim of the patriarch of the Copenhagen interpretation as a weapon *against* the Copenhagen interpretation itself: the theory of beables is introduced here clearly as an 'antidote' to the tendency to adopt an axiomatic formulation of quantum mechanics that relies essentially on an ill-defined (according to Bell) notion of measurement.

The use of the name of Bohr in the 1973 paper is not new to Bell, though. It occurs in the very first section of the first article devoted by Bell to the issue of hidden variables, namely the paper *On the problem of hidden variables in quantum mechanics*, written in 1963 but published in 1966. It is the path-breaking article in which Bell reviews the existing impossibility proofs for a hidden variable re-interpretation of quantum mechanics—from von Neumann 1932 to Jauch-Piron 1963, through the work of Gleason in 1957—only to find them all wanting. As is well known, according to Bell all these proofs—no matter what the internal variants were—shared a common drawback, that of requiring assumptions that it was not reasonable to require from *any* possible, hypothetical hidden variable completion of quantum theory.[2] It is in the context of anticipating, in the first section, the core of the article that Bell exploits the name of Bohr, in order to support his claim and make the unreasonableness of the existing impossibility proofs even more apparent:

> It will be urged that these analyses [i.e. the above mentioned proofs] leave the real question untouched. In fact it will be seen that these demonstrations require from the hypothetical dispersion free states, not only that appropriate ensembles thereof should have all measurable properties of quantum mechanical states, *but certain other properties as well*. These additional demands appear reasonable when results of measurement are loosely identified with properties of isolated systems. They are seen to be quite unreasonable when one remembers with Bohr 'the impossibility of any sharp distinction between the behaviour of atomic objects and the interaction with the measuring instruments which serve to define the conditions under which the phenomena appear'. (Bell 1966, in Bell 2004[2], pp. 1–2, my emphasis).

The Bohr view, referred to by Bell, is that in a quantum measurement process a peculiar, non-classical form of non-separability emerges between object system and

[2] For recent re-assessments of the Bell arguments against von Neumann-Gleason and Jauch-Piron see, respectively, Acuna 2021, and (Laudisa 2023).

apparatus. In his 1966 paper Bell appears to exploit this Bohrian non-separability in support of his critical attitude toward the no-hidden variable theorems by von Neumann, Gleason and Jauch-Piron. In other words, Bell presents the Bohr view as an early instance of would have been called 'contextuality', suggesting at the same time that this should have long taught von Neumann, Gleason, Jauch and Piron that any serious hypothetical hidden-variable completion of quantum mechanics was bound to incorporate a form of context-dependence in the first place.[3]

In the same vein, the name of Bohr emerges in the 1971 Bell paper *Introduction to the hidden-variable question*, where Bell first introduces the family of stochastic hidden variable theories. In discussing "the very essential role of apparatus" in the quantum–mechanical description of the measurement process, Bell argues that.

> The result of the measurement does not actually tell us about some property previously possessed by the system, but about something which has come into being in the combination of system and apparatus. *Of course, the vital role of the complete physical set-up we learned long ago, especially from Bohr*." (Bell 1971, in Bell Bell 2004[2], p. 35).

Bell returns to the same point in his later 1982 article "The impossible pilot wave". In recalling once again the lack of generality of the early no-hidden variable theorems, Bell writes about what he calls 'the Gleason-Jauch argument':

> For a given operator P_1 it is possible (when the dimension N of the spin space exceeds 2) to find more than one set of other orthogonal projection operators to complete it:
>
> $$1 = P_1 + P_2 + P_3 \ldots$$
> $$= P_1 + P'_2 + P'_3 \ldots$$
>
> where P'$_2$ … commute with P_1 and with one another, but not with P_2 …. And the extra assumption is this: the result of 'measuring' is independent of which complementary set… or… is 'measured' at the same time. The de Broglie-Bohm picture does not respect this. […] In denying the Gleason-Jauch independence hypothesis, the de Broglie-Bohm picture illustrates rather the importance of the experimental set-up as a whole, *as insisted on by Bohr. The Gleason-Jauch axiom is a denial of Bohr's insight.* (Bell 1982, in Bell 2004[2], p. 165).

We have evidence, then, that Bohr has a place in the Bell line of thought about the foundations of quantum mechanics already in the early Sixties, as a forerunner of the idea of contextuality.

But let us return to what we called the Bell theory of beables, as expressed by the conditions (i) and (ii). These conditions appear far from uncontroversial, when

[3] For the meaning and role of contextuality in the Bohr philosophy of quantum mechanics is a relevant issue in the Bohrian scholarship: see for instance (Dieks 2017). Given that Bohr was standardly conceived as the major representative of an approach to the foundations of quantum mechanics that could not be more alien to Bell in many respects, Abner Shimony has playfully described Bell's use of the Bohr claim: "Bell, *by a judo-like manoeuvre*, cited Bohr in order to vindicate a family of hidden variables theories in which the values of observables depend not only upon the state of the system but also upon the context." (Shimony 1984, in Shimony 1993, p. 121, my emphasis).

referred to the early, Bohrian characterization of beables by Bell. If condition (i) sounds milder, since it seems to require just compatibility with common sense, condition (ii) is more puzzling. What sort of 'constitution' property is supposed to be involved in the claim that observables are 'made out of' beables? What are beables supposed to be in order to 'make up' observables? And what is the exact relation of such intuition of 'constitution' with the Bohrian view of quantum measurements? The use of Bohr *against* Copenhagen quantum mechanics would do no harm as such, but the Bell strategy is dubious by a *conceptual* point of view. I wish to argue that the Bohrian requirement to express experimental evidence in 'classical' terms, in order for linguistic communications among scientists to be consistenly preserved, can hardly be put usefully to work to provide the unambiguous description of the quantum measurement process that Bell was searching for.

The extent to which the reference to Bohr may really play the role of dissolving the ambiguity deplored by Bell is a matter of dispute, since it concerns the status of an issue that is still debated in the reconstructions of the Bohr attitude toward quantum mechanics: the issue of whether, according to Bohr, quantum mechanics should be taken as universal—i.e. applicable to *all* physical systems, including measuring instruments—or not. The problem of the universality of quantum mechanics in principle emerged since the very origins of quantum theory, due to the increasing divergence from all preceding classical physics that was apparent in the experimental development of the theory already in the first decades of the twentieth century. In the early days of the debate on the foundations of quantum mechanics, it was far from clear what the relation between the classical and the quantum regimes was supposed to be, until the mathematical treatment of the theory in the 1932 von Neumann treatise allowed physicists to put the problem in a clearer light in terms of the notorious 'measurement problem', raising for the first time the universality issue for quantum mechanics. The von Neumann treatment, and the place occupied by this problem in his first formally rigorous formulation of quantum theory, already revealed how controversial the status of measurement in quantum mechanics would have been, to the extent that the very notion of measurement would turn out to be the *locus classicus* for emphasizing the lack of consensus on the interpretation of the theory: von Neumann explicitly confronts the implications of the assumption that—in the context of a measurement of a physical quantity on a quantum system S with an apparatus A—the laws of QM govern *both S and A*. This view has acquired with time the status of a commonplace: 'quantum fundamentalism'—this is how, for instance, Zinkernagel (2015) calls it—is the claim that "Everything in the universe (if not the universe as a whole) is fundamentally of a quantum nature and ultimately describable in quantum–mechanical terms." In Zinkernagel's words:

> In this formulation, quantum fundamentalism contains both an ontological and an epistemological thesis: that everything is of a quantum nature is an ontological claim, whereas the idea that everything can (at least in principle) be *described* in quantum terms is epistemological. The ontological component of quantum fundamentalism can also be expressed as the idea that we live in a quantum world. (Zinkernagel 2015, p. 419, emphasis in the original).

In fact Bohr never discussed explicitly the measurement problem in the von Neumann formal context. A wide consensus was established among most Bohr

scholars, however, according to which his overall philosophical outlook legitimates a *non*-universalistic reading of quantum mechanics, mainly due to the special role attributed to classical categories in accounting for the experimental evidence in quantum measurements. For instance in a recent, qualified defense of this consensus, Zinkernagel (2015) refers to a 1938 paper in which Bohr argues that.

> […] in each case some ultimate measuring instruments, like the scales and clocks which determine the frame of space-time coordination – on which, in the last resort, even the definitions of momentum and energy quantities rest – must always be described *entirely* on classical lines, and consequently kept outside the system subject to quantum mechanical treatment." (Bohr 1938, p. 104, emphasis in the original).

One can make sense of this argument, according to Zinkernagel, only under the assumption that quantum mechanics actually *fails* to be universal:

> A way to understand Bohr's requirement is that we need a reference frame to make sense of, say, the position of an electron (in order to establish with respect to what an electron has a position). And, by definition, a reference frame has a well-defined position and state of motion (momentum). Thus the reference frame is not subject to any Heisenberg uncertainty, and it is in this sense (and in this context) classical. This does not exclude that any given reference system could itself be treated quantum mechanically, but we would then need another – classically described – reference system e.g. to ascribe position (or uncertainty in position) to the former. (Zinkernagel 2015, p. 430).[4]

This view has been challenged. Already (Landsman 2007), for instance, had argued that the Bohr texts would not justify an interpretation of his thought to the effect that there exists an independent natural realm of an intrinsic classical character. Let us consider the following passage, contained in a famous Bohr paper entitled "On the notions of causality and complementarity", published in 1948 on the philosophical journal *Dialectica*:

> The construction and the functioning of all apparatus like diaphragms and shutters, serving to define geometry and timing of the experimental arrangements, or photographic plates used for recording the localization of atomic objects, will depend on properties of materials which are themselves essentially determined by the quantum of action. (Bohr 1948, p. 145).

On the basis of texts like this, Landsman claims that the division system/apparatus, in which the former is described quantum-mechanically whereas the latter is described classically, has no *ontological* import:

> there is no doubt that both Bohr and Heisenberg believed in the fundamental and universal nature of quantum mechanics, and saw the classical description of the apparatus as a purely *epistemological* move, which expressed the fact that a given quantum system is being used as a measuring device" (Landsman 2007, p. 437, emphasis added).

In a recent contribution Dieks reinforces this challenge, defending an exclusive *epistemic* reading of the role of the classical notions in the Bohr view of the quantum measurement process, denying any *ontological* quantum non-universalism by Bohr (Dieks 2017).

[4] A more sustained defense of this view is contained in Zinkernagel (2016).

This dispute on the ontological or epistemological flavor of the quantum/classical divide, however, leaves the ambiguity point that concerns us here untouched. We do not need to take a stance on whether the boundary between the classical and the quantum world concerns our knowledge or the ultimate structure of Nature to see that we are forced anyway, within the Heisenberg-Bohr Copenhagen framework, to acknowledge that, on one hand, we cannot but locate somewhere the infamous 'cut' between system and apparatus, and on the other hand there is no rigorous recipe even on a pragmatic level about where *exactly* we should put it. As Dieks himself remarks, in the very first section of the seminal complementarity paper published in 1927, Bohr emphasizes that.

> The circumstance […] that in interpreting observations use has always to be made of theoretical notions entails that for every particular case *it is a question of convenience* at which point the concept of observation involving the quantum postulate with its inherent "irrationality" is brought in" (Bohr 1934, p. 54, emphasis added),

Wolfgang Pauli echoed the same point in a 1949 paper, entitled "The Philosophical Significance of the Idea of Complementarity":

> […] modern physics generalizes the old placing in opposition of apprehending subject on one hand and object apprehended on the other to the idea of the cut between the observer or instrument of observation and the system observed. While the existence of such a cut is a necessary condition of human cognition, modern physics regards the position of the cut as to a certain extent arbitrary, and as the result of a choice partly determined by considerations of expediency, and therefore partly free. (Pauli 1950, p. 41, emphasis added).

As a consequence, the 'ambiguity' and 'approximation' of the standard formulation of quantum mechanics cannot be removed by the use of the Bohrian framework, and Bell needed to say (and later did say) more to characterize the kind of solution he envisioned. In particular, the Bohrian model of the quantum measurement may at most satisfy the Bell condition (i), namely, the 'functionalistic' recovery of subjective experience, but fails to satisfy unambiguously the 'constitutive' Bell condition (ii), since the concrete individuation of the relevant beables depends on *arbitrary* criteria: with the resources allowed by the Bohr framework, quantum observables simply *cannot* be 'made out' of beables.

3 Beyond Bohr: The New Life of Beables

In the first appearance of the notion of beable, the early Bell move—use Bohr *against* Copenhagen quantum mechanics—looks then rather unfortunate. But the role that we have analyzed in the previous section starts to be replaced in the subsequent development of the notion itself. For Bell returns to beables in a 1975 paper, whose title ("The theory of local beables") this time mentions explicitly the need for a *theory* of these 'objects', whatever they are meant to be. At first sight, the very opening of the paper is in line with the Bohrian attitude we have alluded to:

> This is a pretentious name for a theory which hardly exists otherwise, but which ought to exist. The name is deliberately modelled on 'the algebra of local observables'. This terminology, *be*-able as against *observ*able, is not designed to frighten with metaphysic those dedicated to realphysic. It is chosen rather to help in making explicit some notions already implicit in, and basic to ordinary quantum theory. For, in the words of Bohr, ' it is decisive to recognize that, however far the phenomena transcend the scope of classical physical explanation, the account of all evidence must be expressed in classical terms.' It is the ambition of the theory of local beables to bring these 'classical terms' into the equations, and not relegate them entirely to the surrounding talk. (Bell 1975, in Bell 2004[2], p. 52).

In clarifying what beables are supposed, or meant, to be, Bell refers again to macroscopic pieces of experimental settings in a broad sense—and this is, once again, entirely Bohrian in spirit—but, *this time*, he expresses explicitly the need for a clear *theory* of them, in terms of a more robust sense of physical reality:

> The beables must include the settings of switches and knobs on experimental equipments, the current in coils, and the readings of instruments. 'Observables' must be *made*, somehow, out of beables. The theory of local beables should contain, and give precise physical meaning to, the algebra of local observables. (Bell 1975, in Bell 2004[2], p. 52).

This appears to be a turning point in the Bell characterization of beables. Not only Bell refers to the difference in electromagnetism between 'physical' entities (like the electric and magnetic fields) and 'unphysical' entities (like potentials), in order to set up a distinction according to which beables should be clearly located on the 'physical' side. He also points here to what we have called above a condition of 'constitution', a more fundamental status that beables should be endowed with: it is this status that in principle justifies the observables being *made out* of beables. This conjunction of realism—beables *are out there*—and constitution—beables are what *make up* observables and all that gravitates around observation—characterizes the new Bell theory of beables, and his later paper "Beables for quantum field theory" (1984) testifies it:

> There is nothing in the mathematics to tell what is 'system' and what is 'apparatus', nothing to tell which natural processes have the special status of 'measurements'. Discretion and good taste, born of experience, allow us to use quantum theory with marvelous success, despite the ambiguity of the concepts named above in quotation marks. But it seems clear that in a serious fundamental formulation such concepts must be excluded. In particular we will exclude the notion of 'observable' in favour of that of '*be*able'. The beables of the theory are those elements which might correspond to elements of reality, to things which exist. Their existence does not depend on 'observation'. Indeed observation and observers must be made out of beables. (Bell 1984, in Bell 1987, p. 174).

That beables should correspond "to elements of reality, to things which exist" might still sound compatible with the Bell early, Bohrian-sounding formulation that we analyzed in the previous section, but clearly this is not the case with the claim that the existence of beables *does not depend on 'observation'*: in Bohrian terms, on the contrary, it is exactly the reference to the context of observation that allows macroscopic pieces of experimental settings (namely, what Bell takes as beables in his early formulation) to be part of a scientifically meaningful experience.

In connection with this emphasis *both* on the 'reality' of beables and their 'constitutive' nature, Bell introduces for the first time a connection with an intuitive sense of *locality*, called here *local causality*[5]:

> We will be particularly concerned with *local* beables, those which (unlike the total energy) can be assigned to some bounded space-time region. […] It is in terms of local beables that we can hope to formulate some notion of local causality. (Bell 1975, in Bell 2004[2], p. 53, emphasis in the original).

It is *this* focus on locality—I argue—that determines a new twist for the formulation of a theory of beables, a formulation which starts to diverge from the Bohrian-sounding notion reviewed in the previous section and receives a more distinctive 'fundamental' status in somewhat ontological terms. Bell attempts to figure out a definition of local causality that can work also in an indeterministic setting, an attempt that leads him to introduce an expression like $\{A|\Lambda\}$, that stands for the probability of a particular value A, given particular values Λ (Bell 1975, in Bell 2004[2], p. 54). An interesting point to note here is that, in introducing this expression, Bell employs the term 'beable' to denote a *value* (of a physical quantity), something very different from "settings of switches and knobs on experimental equipments", which was the original, Bohrian-sounding meaning attached to the term. On this new background Bell operates in a much more explicitly 'realistic' (and much less 'Bohrian') vein—a background in which it is perfectly sensible to conceive an observer-independent world whose unveiling is a major task for fundamental physics—and the new reading of beables in terms of values is immediately put to work in an EPR-kind of context:

> Let A be localized in a space-time region 1. Let B be a second beable localized in a second region 2 separated from 1 in a spacelike way. Now my intuitive notion of local causality is that events in 2 should not be 'causes' of events in 1, and viceversa. But this does not mean that the two sets of events should be uncorrelated, for they could have common causes in the overlap of their backward light cones. It is perfectly intelligible then that if Λ in 1 does not contain a complete record of events in that overlap, it can be usefully supplemented by information from region 2. So in general it is expected that
>
> $$\{A|\Lambda, B\} \neq \{A|\Lambda\}$$
>
> However, in the particular case that Λ contains already a *complete* specification of beables in the overlap of the two light cones, supplementary information from region 2 could reasonably be expected to be redundant. (Bell 1975, in Bell 2004[2], p. 54, emphasis in the original).

It is quite clear, then, that the above mentioned specification of beables makes sense in the Bell second, ontologically-loaded formulation of the notion of beable (and not in the old, Bohrian-sounding one). Moreover, this new formulation is immediately put to work in the investigation on whether, in the Bell language, quantum mechanics might be shown to be 'locally causal' if reformulated as a sub-theory of a 'more complete' theory:

---

[5] As already remarked by others, this expression is likely to be misleading in suggesting that the influence at stake *should* have a direction, which in fact is not necessarily the case.

> But could it not be that quantum mechanics is a fragment of a more complete theory, in which there are other ways of using the given beables, or in which there are additional beables – hitherto 'hidden' beables? And could it not be that this more complete theory has local causality? Quantum mechanical predictions would then apply not to given values of all the beables, but to some probability distribution over them, in which the beables recognized as relevant by quantum mechanics are held fixed. We will investigate this question, and answer it in the negative. (Bell 1975, in Bell 2004[2], p. 55).

Thus, the notion of beable (in his second, mature sense) appears to have been a major factor for motivating the development of the PO research program in the foundations of QM. After the initial proposal by Detlef Dürr, Shelly Goldstein and Nino Zanghì, in more recent times the scientific literature on the evaluation of the PO approach has been growing significantly in quantity and depth (Allori 2013, 2015). Our aim here was just to provide an attempt of reconstructing the Bell own conceptual evolution on the notion of beable over the years: an interesting, open question is to investigate how the Bell theory of beables fares with respect to some recent claims concerning the relationship between beables and the PO approach, and whether the evolution from the first to the second formulation of the Bell theory of beables can shed some light on the study of this relationship.

References

V. Allori, Primitive ontology and the structure of fundamental physical theories, in *The Wave Function: Essays on the Metaphysics of Quantum Mechanics*. ed. by D.Z. Albert, A. Ney (Oxford University Press, Oxford, 2013), pp.58–75

V. Allori, Primitive ontology in a nutshell. Int. J. Quantum Found. **1**, 107–122 (2015)

J.S. Bell, On the problem of hidden variables in quantum mechanics. Rev. Mod. Phys. **38**, 447–452 (1966), reprinted in Bell 2004[2], pp. 1–13.

J.S. Bell, Introduction to the hidden-variable question, in *Foundations of Quantum Mechanics*, ed. by B. D'Espagnat (Academic Press, New York, 1971), pp. 171–181, reprinted in Bell 2004[2], pp. 29–39.

J.S. Bell, Subject and object, in *The Physicist's Conception of Nature*, ed. by J. Mehra (Reidel, Dordrecht, 1973), pp. 697–690, reprinted in Bell 2004[2], pp. 40–44.

J.S. Bell, The theory of local beables, TH-2053-CERN (1975), reprinted in Bell 2004[2], pp. 52–62

J.S. Bell, The measurement theory of Everett and de Broglie's pilot wave, in *Quantum Mechanics, Determinism, Causality and Particles*, ed. by M. Flato et al. (Reidel, Dordrecht, 1976), pp. 11–17, reprinted in Bell 2004[2], pp. 93–99.

J.S. Bell, On the impossible pilot wave, *Foundations of Physics*, vol. 12, pp. 989–999 (1982), reprinted in Bell 2004[2], pp. 159–168.

J.S. Bell, Beables for quantum field theory, CERN-TH 4035/84 (1984), reprinted in Bell 2004[2], pp. 173–180.

J.S. Bell, Are there quantum jumps?, in *Schrödinger: Centenary Celebration of a Polymath*, ed. by C.W. Kilmister (Cambridge University Press, Cambridge, 1987), pp. 41–52, reprinted in Bell 2004[2], pp. 201–212.

J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 2004[2])

N. Bohr, *Atomic theory and the description of nature* (Cambridge University Press, Cambridge, 1934)

N. Bohr, The causality problem in atomic physics, in *New Theories in Physics*, International Institute of Intellectual Cooperation, Paris, pp. 11–45 (1938) (reprinted in ed. by J. Faye, H.J. Folse, *The Philosophical Writings of Niels Bohr*, *Vol. IV: Causality and Complementarity,* (Ox Bow, Woodbridge, 1998), pp. 94–121).

N. Bohr, On the notions of causality and complementarity, *Dialectica* 2, pp. 312–319 (1948) (reprinted in ed. by J. Faye, H.J. Folse, *The Philosophical Writings of Niels Bohr*, *Vol. IV: Causality and Complementarity* (Ox Bow, Woodbridge, 1998), pp. 141–148).

N. Bohr, Discussion with Einstein on the epistemological problems in atomic physics, in *Albert Einstein Philosopher-Scientist*, ed. by P.A. Schilpp (Harper & Torchbooks, New York, 1949), pp. 199–243

D. Dieks, Niels Bohr and the formalism of quantum mechanics, in 2017, *Niels Bohr and the Philosophy of Physics. Twenty-First Century Perspectives*, ed. by J. Faye, H.J. Folse (Bloomsbury, London-New York, 2017), pp. 303–334.

D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the origin of absolute uncertainty. J. Stat. Phys. **67**, 843–907 (1992)

N.P. Landsman, Between classical and quantum, in *Handbook of the Philosophy of Science. Philosophy of Physics*, ed. by J. Butterfield, J. Earman (Elsevier, Amsterdam, 2007), pp. 417–553.

F. Laudisa, How and why did locality become 'local realism'? A historical and critical analysis (1963–1978). Stud. Hist. Philos. Sci. **97**, 44–57 (2023)

W. Pauli, Die philosophische Bedeutung der Idee der Komplementarität, in *Experientia*, vol. 6, p. 72–75 (1950) (Engl. ed. The philosophical significance of the idea of complementarity, in ed. by W. Pauli, *Writings on Physics and Philosophy*, ed. by C.P. Enz, K. von Meyenn) (Springer, Berlin 1994), pp. 35–42 (page references to the English edition)

A. Shimony, Contextual hidden variables theories and Bell's inequalities. *British Journal for the Philosophy of Science* 35, pp. 25–45 (reprinted in *Search for a Naturalistic World View. Natural Science and Metaphysics*, vol. II (Cambridge University Press, Cambridge, 1993, 1984), pp. 104–129

A. Whitaker, *John Bell and Twentieth-Century Physics* (Oxford University Press, Oxford, Vision and Integrity, 2016)

H. Zinkernagel, Are we living in a quantum world? Bohr and quantum fundamentalism, in *One Hundred Years of the Bohr Atom*, *Proceedings from a Conference. Scientia Danica. Series M: Mathematica et Physica*, ed. by F. Aaserud & H. Kragh, vol. 1 (Royal Danish Academy of Sciences and Letters, Copenhagen, 2015), pp. 419–434.

H. Zinkernagel, Niels Bohr on the wave function and the classical/quantum divide. Stud. Hist. Philos. Mod. Phys. **53**, 9–19 (2016)

Probability and Typicality in Statistical Mechanics

Barry Loewer

Abstract Detlef Dürr was inspirational to many who write about issues in the philosophical foundations of physics and probability. For many years I have been interested in his work on statistical mechanics and Bohmian mechanics and especially by the role of typicality in these theories. In my contribution I will say a few words comparing typicality and probability approaches to statistical mechanics and ask whether the approaches are friends or foes.

Since it is difficult to understand the nature of probability, especially when a theory's dynamics are deterministic, Dürr and others proposed that typicality is the more fundamental notion and that it is crucial to understanding statistical mechanics and Bohmian mechanics. His work inspired an approach to statistical mechanics that has many connections to the earlier history of the subject. Among the recent contributors to this approach in addition to Detlef have been Allori, Goldstein (2011), Hubert (2022), Lazarovici (2020), Lebowitz (1981), Maudlin (2020), Tumulka, Wilhelm (2022), and Zanghì.

The original developers of statistical mechanics, Maxwell, Boltzmann, and Gibbs, thought of the probabilities that occur in classical statistical mechanics as epistemic, but they also thought that they are in some way objective and physical. They are not merely subjective degrees of belief since there seems to be something objective about the world that makes certain probability distributions over microstates the correct ones. Boltzmann argued that they derive from the chaotic behavior of molecules. There also seems to be something not merely epistemic about them since they underlie and explain objective regularities like the melting ice cubes and other manifestations of the second law of thermodynamics. But they cannot be objective physical probabilities that derive from indeterministic dynamical laws, usually called "propensities," since the dynamics assumed in classical statistical mechanics and Bohmian mechanics are deterministic. So, how should we understand the probabilities in these theories?

B. Loewer (✉)
Department of Philosophy, Rutgers University, New Brunswick, NJ, USA
e-mail: loewer@philosophy.rutgers.edu

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_30

One proposal is that logic or rationality plays a role in dictating the (or those) objectively correct probability distributions. The usual version of this idea is based on the principle of indifference (PI), which proposes that if nature or you have no reason to prefer one possibility over another, you should assign them equal probabilities. Notoriously, if there are infinitely many possibilities each possibility has probability equal to 0 so the principles must be formulated using probability densities. But then PI may lead to alternative probability densities and so to contradictions. To avoid contradictions, a measure needs to be selected and since there are infinitely many measures it is not clear which measure is correct and why it is correct. Jaynes (1968) famously attempted to address this issue with his maximum entropy principle. But the real issue is that there is no justification for promoting not knowing which possibility is the actual one into knowledge of a probability distribution over possibilities? As David Albert likes to say, "Not having a clue as to which possibility is actual is not the same as having equal confidence in each."

Typicality accounts avoid this problem by explaining thermodynamic behavior not in terms of probabilities but in terms of the concept of typicality; a concept that its advocates consider more basic than probability.[1] They then characterize probability in terms of typical frequencies. The idea is that thermodynamic behavior (e.g., the melting of ice cubes), while not following from dynamical laws alone, is typical behavior. This means, roughly, that thermodynamic behavior occurs as a matter of dynamical law *almost all the time* for *almost all possible initial conditions.* By appealing to "typicality" this approach seeks to avoid and solve the problems encountered by attempts to understand what probability is.

The basic idea behind the concept is that *B*s are typical among *A*s if and only if almost all *A*s are *B*. "Almost all" means that the number of *B*s that are not *A*s is negligible. Typicality is a purely numerical (e.g., "how many") notion. If *A* is a finite set, then it is just a matter of counting. Unlike probability, typicality is not a matter of degree. *B*s may be typical among *A*s, atypical among *A*s or neither. Typicality has no conceptual connection with randomness and is alleged to not run into the same difficulties that the usual interpretations of probability encounter.

If *A* is infinite but *A* and *B* have the same cardinality a measure is required in order to specify what "almost all" means when claiming that almost all *A*s are *B*s. Unlike objective probability typicality is not connected with degree of belief by Lewis' Principal Principle but it is connected with full belief/disbelief by a similar principle. If you think *B*s are typical among *A*s then if you have no defeating reason you should believe the next *B* will be an *A*.

The questions immediately arise: What measures are appropriate and how to justify that a particular measure is appropriate? Advocates of the typicality approach say that some measures are "natural" and others not "not natural." For example, for the situation of determining typical outcomes of coin flips the Liouville measure over microstates compatible with the macro state of the coin and flipping mechanism is natural while a measure that assigns positive measure only to finitely many of the

[1] For accounts of typicality and its application to understanding probability and statistical mechanics, see Goldstein (2011), Hubert (2022), Lazarovici (2020), Maudlin (2020), and Wilhelm (2022).

microstates compatible with this macro state is unnatural. The idea is that a natural measure better captures the intuitive meaning of "almost all." It is also claimed that whether or not a measure is natural depends on how it is related to a system's dynamics. For example, the Liouville measure is natural as a measure of typicality because it is stationary (i.e., is the same over time as the dynamics evolves a system). These features are said to justify its choice as "the correct measure."

Advocates of typicality don't understand the typicality measure as contingent hypothesis in addition to the dynamical laws but something more along the lines of an a priori consequence of the dynamical laws. This seems to make it close to being nomologically necessary. It seems then that even in worlds with our dynamical laws, but pervasive anti-thermodynamic behavior the typicality account will say that thermodynamic behavior is typical, and one should believe e.g., ice cubes melt. This may be a problem since thermodynamic behavior is atypical in these worlds. These two justifications—that a natural measure captures the intuitive meaning of "almost all" actual *A*s are *B* and that it plays nice with the dynamics—are in tension with each other. The intuitive meaning of *A*s are typically *B*s is that almost all the actual *A*s are *B*. But since the typicality measure doesn't necessarily follow from the dynamics there are nomologically possible worlds in which not all *A*s are *B*s. In fact, there are worlds in which most *A*s are not *B*s. It is true that in most possible worlds (as specified by the measure) most *A*s are *B*s but since there are nomologically possible worlds in which most *A*s are not *B*s the claim that the dynamical laws bring with them what it means for most *A*s to be *B*s can't be correct. It doesn't obviously follow from the dynamical laws. The correct measure is the appropriate one seems to have to be considered to be an additional law.

The typicality approach usually endorses a principle connecting believing that *A* is typical to believing *A*.

CP: If one believes that *B*s are typical among *A*s and believes that *x* is an *A* then one is rationally licensed to believe that *x* is a *B* as long as she has no "undermining" belief that *x* is a *C* and *B*s are not typical among *A*s that are *C*.

CP is the typicality version of what is known as "Cournot's Principle."[2] It is akin to Lewis' Principal Principle (PP):

$$\text{(PP) } \mathrm{Cred}(B|\mathrm{P}(B) = x \,\&\, A) = x.$$

(Cred($B|A$) is a subject's degree of credence (belief) that *B* given *A*.) *A* is any proposition that is admissible with respect to *B*. A proposition is admissible if and only if it provides information about *B* only by providing information about P(*B*).

CP connects typicality with belief or acceptance while the PP connects probability with degree of belief. I think of Cournot's Principle as "a poor man's Principal Principle."

[2] See Shafer (2006) for a discussion of Cournot's principle. Shafer points out that many of the developers of probability including Bernoulli, Cournot, Kolmogorov endorsed the probabilistic version and held that it was the way that probability makes contact with non-probabilistic facts and justified beliefs about them.

On the typicality account probabilities can be identified with typical long run frequencies.[3] The idea is that given the Liouville measure over initial micro conditions of the coin tossing system if the coin is fair for typical initial conditions typical sequences of outcomes of repetitions of coin tosses under conditions *C* yield heads about half the time. Further, typical sequences will appear random in the sense that they meet statistical tests for randomness even though the dynamics are completely deterministic (Hubert 2022; Lazarovici 2020).

The typicality account of probability, although it is a frequency account, is very different from the usual actual and hypothetical frequency accounts which are open to many objections (see Hajek 2002). Like other frequency accounts it assigns probabilities to type not token events. A token event (e.g., the outcome of the next game between the Celtics and the Knicks) satisfies many distinct types (e.g., game between Celtics and Knicks, game played on Saturday, game played in Madison Square Garden etc.), and the typical frequency of Knicks winning may be different for each of these types. If the complete micro condition of the state of the universe prior to the game is *A* then it will determine whether the Knicks win or lose. But the complete macro state may lead to a typical frequency of Knicks winning different from 1 or 0 which we can think of as the probability of the token event. I will return to this point later.

The typicality approach provides an account of why it is that thermodynamic behavior including the second law is typical and an account of all thermodynamic probabilities. Its Achilles' heel it is that it depends on a particular measure as the correct measure and faces the problems of explaining the epistemic and metaphysical status of the measure. That the Liouville measure is the correct typicality measure doesn't follow either analytically from the meaning of "typicality" nor is it entailed by the dynamical laws. Further, one may wonder why *G*s being typical among *F*s licenses believing that a particular *G* is *F* even in the finite case with an assumption that the *F*s are equally likely. Something like the Principal of Indifference seems to be being assumed. Without an account of its metaphysical status and an epistemic justification the typicality account is at best, incomplete.

I now want to contrast the typicality account with a probabilistic version of Boltzmannian statistical mechanics where the probabilities in terms of a development of Lewis' Humean best systems account of laws and chances. Humean accounts of laws understand laws to be propositions that describe significant regularities among events and states. They are called "Humean" since they avoid any appeal to laws necessitating or governing regularities or to necessary connections among properties. On Lewis' Humean best system account the fact that a proposition is a law supervenes on the distribution of events throughout space–time.

Lewis' best systems account (BSA) says that laws are certain propositions entailed by the axioms that best systematize the totality of the world's fundamental events. According to Lewis the goodness of a system is measured in terms of its informativeness and simplicity. He mentions these because they are criteria physicists appeal to in deciding among fundamental theories, but it is best to think of a best system as

[3] See Hubert (2022) for an account of probability in terms of typicality.

a system that optimally balances satisfaction of whatever criteria in addition to these that physicists use when deciding among theories.

I now want to contrast the typicality account with a probabilistic version of Boltzmannian statistical mechanics where the probabilities in terms of a development of Lewis' Humean best systems account of laws and chances. Humean accounts of laws understand laws to be propositions that describe significant regularities among events and states. They are called "Humean" since they avoid any appeal to laws necessitating or governing regularities or to necessary connections among properties. On Lewis' Humean best system account the fact that a proposition is a law supervenes on the distribution of events throughout space–time.

On the BSA objective probabilities are not propensities or frequencies but probability axioms are added to a system as a way of providing information simply. They are related to actual frequencies since these are simple and informative but not identical to them. For example, a sequence of outcomes of coin flips (*HTHTTTHTHTTHHT*…) can be described by saying that its members are the outcomes of independent tosses with $P(H) = P(T) = 1/2$. Lewis proposes that a system containing probability axioms is evaluated in terms of "fit" as well as simplicity and informativeness. The degree of fit of a system to a world is measured by the likelihood of the world given the candidate system. This is an instance of the likelihood principle that is used in statistics in comparing evidential support for competing theories.[4]

We can think of probability propositions as providing information about non-probabilistic propositions by way of Lewis' Principal Principle (PP). A useful version connects objective conditional probabilities with conditional degrees of belief:

$$\left(\mathrm{PP}^*\right) \mathrm{Cred}(B|A) = x \text{ if and only if } \mathrm{P}(B|A) = x.$$

This is an externalist version of the PP that has no need for admissibility.

The BSA account of objective probabilities is similar to an actual frequency account since its probabilities supervene on the actual distribution of non-probabilistic events. But it differs in that it doesn't identify probabilities with actual frequencies and allows for the possibility that while the probability of an event E equals x, repetitions of independent occurrences of E may result in a frequency different from x.

Lewis thought that probabilistic best systems are restricted to those whose dynamical laws are indeterministic and specify the probabilities at a time of a complete microstate evolving into alternative future states.[5] But the basic idea behind the BSA can be developed as to include objective probabilities even when a system's dynamical laws are deterministic (Loewer 2004). This is accomplished by allowing candidate systems to include a probability distribution over initial conditions or over alternative histories compatible with the deterministic laws. The chance of A at t can be identified with the objective probability of A given the micro history of the

[4] The likelihood principle is built into Bayesian inference.

[5] The dynamical equation of GRW that describes how the world's or an isolated system's quantum state evolves is an example of such an indeterministic dynamical law.

world up until *t*. If the dynamical laws are deterministic then chances will all trivially be 0 or 1. But even if the dynamical laws are deterministic there may be objective conditional probabilities that differ from 0 and 1.[6]

The Mentaculus account of statistical mechanics is based on the BSA of laws and objective probabilities. It is the claim that the best system for our world has as its axioms:

(1) The dynamical laws
(2) A uniform probability distribution (PROB) given by the Liouville measure over the phase space.
(3) The Past Hypothesis (PH) that says that the macro state of the universe soon after the time of the big bang is one whose entropy is very low and satisfies certain other cosmological macro conditions.

It has been argued (Albert 2000) that this system implies that as long as the universe or any of its isolated subsystems are not in their equilibrium macro states (states of maximum entropy) it is very likely that their entropies will increase. More generally, the Mentaculus entails all the usual claims of statistical mechanics.[7] The arguments for this are based on Boltzmann's original arguments developed over the history of statistical mechanics and are the same ones employed by the typicality account.

Both the typicality and the Mentaculus accounts can explain the success of statistical mechanics. However, there are important differences between the Mentaculus account and the other accounts. Here are some:

1. Where the typicality account is based on typicality and that is supposed to be akin to a cardinality or counting measure the Mentaculus account is based on a best system construal of probability. The Mentaculus makes no use of typicality or subjective probabilities in accounting for statistical mechanics. The notion of typicality plays no role in characterizing probability, but typical behavior can be characterized as highly probable behavior.

2. The typicality account is based on assuming the Liouville measure as the typicality measure and is faced with the problems of explaining and justifying its epistemic and metaphysical status. It can be understood either as an a priori claim explicating the meaning or "typical" for such and such dynamics or as an additional law. The typicality account is not committed to a particular metaphysics of laws.

The Mentaculus doesn't assume a measure but rather claims that one of the contingent axioms of the best system of the world is a probability distribution specified by the Liouville measure. The probability measure doesn't follow a priori from the dynamical laws. That the two fit together so that, for example, the probability measure is stationary is a consequence of their being so related makes for a better systematization. Because of this there are worlds compatible with the dynamical laws

[6] Some (e.g., Shaffer) object to calling these objective probabilities "chances" since they think of chances as conceptually connected with indeterminism. No matter what they are called. The important point is that they are objective probabilities.

[7] See Albert (2000) and Loewer (2020).

whose probabilities differ from the statistical probabilities and even worlds with no probabilities at all.[8] The reasons to believe that PROB entails the correct probabilities for the actual world are usual reasons for believing a scientific theory. It is the simplest probability distribution that fits the evidence that has so far been obtained and satisfies other scientific criteria for a theory.

3. The Mentaculus assigns an objective probability to every physically specifiable proposition (proposition that can be identified with a set of worlds compatible with the dynamical laws) and a conditional probability P($B|A$) for every physically specifiable A and B where P(A) > 0. Since the Mentaculus assigns P($B|A$) for all B and A, it is a "probability map of the universe."[9] While the typicality measure also assigns a measure to every proposition, only M(A), M($A|B$) near 1 or 0 are meaningful. When near 1, these mean that A is typical and Bs are typical among As, and when near 0, that A is atypical and Bs are atypical among As.

4. Since the Mentaculus assigns objective conditional probabilities P($B|A$) to all physical propositions B, A and assumes the PH it is able to account for various temporal asymmetries ("time's arrows") including the increase in entropy, that we can influence the future but not the past, that we can know much more about the past than the future, the temporal asymmetry of counterfactuals, that causes precede their effects. The typicality and hybrid accounts make no such claims. If the PH is added to these accounts, it can do some but not all of this work.

5. The typicality and Mentaculus accounts do not conflict, but they make different assumptions and have somewhat different but non-conflicting consequences. I think of them as friends not foes.

References

D. Albert, *Time and Chance* (Harvard University Press, 2000)
S. Goldstein, Typicality and notions of probability in physics, in *Probability in Physics,* ed. by Y. Ben-Menahem, M. Hemmo (Springer, Heidelberg, 2011), pp. 59–71
A. Hajek, Interpretations of probability, in *Stanford Encyclopedia of Philosophy* (2002, 2019). https://plato.stanford.edu/entries/probability-interpret/
M. Hubert, Reviving frequentism. Synthese **99**, 5255–5584 (2022)
E.T. Jaynes, Prior probabilities. IEEE Trans. Syst. Sci. Cybern. **4**(3), 227–241 (1968)
D. Lazarovici, *Typicality as a Way of Reasoning in Physics and Metaphysics.* Doctoral Thesis, Université de Lausanne, Faculté des Lettres, Section de Philosophie, Lausanne, Switzerland (2020)
J. Lebowitz, Microscopic dynamics and macroscopic laws. Ann. N. Y. Acad. Sci. **373**, 220–233 (1981)
B. Loewer, David Lewis' humean theory of objective chance. Philos. Sci. **71**(5), 1115–1125 (2004)

[8] There are also worlds compatible with the actual world's dynamical laws whose laws differ from the actual world. This is sometimes taken to be an objection to the BSA, but I think they can be answered.

[9] The name "Mentaculus" is taken with a nod to self-irony from the film "A Serious Man" by the Coen brothers where a delusional character says he is composing a "probability map of the universe" he calls "the Mentaculus".

B. Loewer, The mentaculus vision, in *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, ed. by V. Allori (World Scientific, 2020)
T. Maudlin, The grammar of typicality, in *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, ed. by V. Allori (World Scientific, 2020)
G. Shafer, Why Did Cournot's Principle Disappear?" (2006). http://www.glennshafer.com/assets/downloads/disappear.pdf
I. Wilhelm, Typical: a theory of typicality and typicality explanation. British J. Philos. Sci. **73**(2), 561–581 (2022)

The History of *Moral Certainty* as the Pre-history of *Typicality*

Mario Hubert

> *In the whole conduct of the understanding, there is nothing of more moment than to know when and where, and how far to give assent.*
> — John Locke, Of the Conduct of the Understanding

Abstract This paper investigates the historical origin and ancestors of typicality, which is now a central concept in Boltzmannian Statistical Mechanics and Bohmian Mechanics. Although Ludwig Boltzmann did not use the word typicality, its main idea, namely, that something happens *almost always* or is valid for *almost all* cases, plays a crucial role for his explanation of how thermodynamic systems approach equilibrium. At the beginning of the 20th century, the focus on *almost always* or *almost everywhere* was fruitful for developing measure theory and probability theory. It was apparently Hugh Everett III who first mentioned typicality in physics in 1957 while searching for a justification of the Born rule in his interpretation of quantum mechanics. The historically closest concept before these developments is *moral certainty*, which was invented by the medieval French theologian Jean Gerson, and it became a standard concept at least until the Age of Enlightenment, when Jakob Bernoulli proved the Law of Large numbers.

In Memory of my Teacher Detlef Dürr

M. Hubert (✉)
Department of Philosophy, The American University in Cairo, Cairo, Egypt
e-mail: mario.hubert@aucegypt.edu

A. Bassi et al. (eds.), *Physics and the Nature of Reality*, Fundamental Theories of Physics 215, https://doi.org/10.1007/978-3-031-45434-9_31

1 A Brief History of Typicality

Typicality[1] is a statistical concept that has been developed in the context of the Boltzmannian approach to Statistical Mechanics and the arrow of time (see, for example, Lebowitz 1993, 2008; Goldstein 2001; Goldstein et al. 2020; Bricmont 2022) and in the context of Bohmian mechanics (Dürr et al. 1992; Dürr and Teufel 2009; Oldofredi et al. 2016). In Statistical Mechanics, one big problem is to properly reduce the Second Law of Thermodynamics to microphysical processes (Myrvold 2020; Robertson 2021). The Boltzmannian approach says that microstates within a specific macrostate *typically* evolve to a macrostate of higher entropy; in other words, *most* or *almost all* microstates within a specific macrostate evolve to a macrostate of higher entropy. In a similar vein, typicality has been used to explain the quantum mechanical Born rule from the microscopic behavior of Bohmian particles (Dürr, Goldstein, and Zanghì, 1992).

In both cases, there is a debate about how to mathematically formalize *almost all* properly and how explanatorily successful typicality is (Volchan 2007; Frigg 2009, 2011; Frigg and Werndl 2012; Lazarovici and Reichert 2015). Typicality is usually formalized by a measure on configuration space or phase space. The main use of the measure is to distinguish between "big" sets (those with measure close to 1) and "small" sets (those with measure close to 0). A problem with using standard measure theory to formalize typicality is that these measures assign sizes to all kinds of sets, also those that are neither "big" nor "small". One way out would be to consider equivalence classes of absolutely continuous measures or to replace the measure by a new predicate (Maudlin 2020).

These issues should not bother us here. Instead, my aim is to trace the history of typicality. Where does it come from? How did it evolve? Have there been similar concepts?

The idea of typicality in Statistical Mechanics goes back to the work of Ludwig Boltzmann (1844–1906) on the Second Law of Thermodynamics (see Darrigol 2018, for a recent exegesis of Boltzmann's original papers). The concrete problem was to explain under which (microphysical) circumstances a box of gas in non-equilibrium approaches equilibrium. Boltzmann provided this explanation with his H-Theorem in 1872. The way he wrote his paper *Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen* (*engl.* Further Studies on Heat Equilibrium among Gas Molecules),[2] however, has created some confusion. Although Boltzmann knew about potential exceptions to the Second Law of Thermodynamics as early as 1868,[3]

[1] Detlef Dürr's work on typicality has influenced much of current research of this notion. In his book *Bohmsche Mechanik als Grundlage der Quantenmechanik* written in German, he used the adjective *typisch* or the noun *das Typische* (Dürr 2001); later he changed the noun to *Typizität* (Dürr, Froemel, and Kolb, 2017).

[2] My translation.

[3] James Clerk Maxwell, whose results Boltzmann developed further, was also aware of those exceptions already in 1867 when he wrote in a letter to his friend Peter Tait about what has become known as *Maxwell's Demon*.

he wrote in 1872 that the H function (the negative of the entropy) "must necessarily decrease" (Darrigol and Renn 2013, Sect. 25.2.7).

It was only in 1877 that Boltzmann did correct his proof of the H-theorem after Loschmidt's reversibility objection in 1876, and so he emphasized that the Second Law of Thermodynamics allows for exceptions considering an appropriate time scale:

> One cannot prove that for every possible initial positions and velocities of the spheres, their distribution must become more uniform after a very long time; one can only prove that the number of initial states leading to a uniform state is infinitely larger than that of initial states leading to a non-uniform state after a given long time; in the latter case the distribution would again become uniform after an even longer time. (Boltzmann, 1877, quoted in Darrigol and Renn 2013, p. 775)

The reception of Boltzmann's work was already turbulent at his time. Either scholars did not read his work, misunderstood his work, or downplayed his work (Brush 1976, Chap. 14). In any case, from 1877 on, Boltzmann was explicit in his writings and his lectures that a box of gas in non-equilibrium will approach equilibrium for *almost all* initial condition, or, in modern parlance, *typically.*

Similar ideas of typicality have been applied in mathematics a little earlier in the 19th century. The German mathematician Carl Friedrich Gauß (1777–1855) and his Belgian colleague Adolphe Quetelet (1796–1874) used the "most probable value" in their works on the normal distribution in the period between 1809 and 1857 (Wagner 2020). Being aware of these developments in mathematics, the German sociologist Max Weber (1864–1920) used the word *typical* in his lectures on *General ("Theoretical") Political Economy* between 1894 and 1898.

In the first half of the 20th century, notions of *almost all*, *almost everywhere*, and *almost always true* were used in the developments of topology, measure theory, probability theory, and classical and quantum statistical mechanics (von Plato 1994; Hald 1998; Goldstein et al. 2010).

Given these developments, it seems that typicality was first explicitly mentioned in physics by Hugh Everett (1957, p. 460), when talking about his *'Relative State' Formulation of Quantum Mechanics*, which has later become the many-worlds interpretation of quantum mechanics (Goldstein 2012):

> We wish to make quantitative statements about the relative frequencies of the different possible results of observation—which are recorded in the memory—for a typical observer state; but to accomplish this we must have a method for selecting a typical element from a superposition of orthogonal states. […] The situation here is fully analogous to that of classical Statistical Mechanics, where one puts a measure on trajectories of systems in the phase space by placing a measure on the phase space itself, and then making assertions (such as ergodicity, quasi-ergodicity, etc.) which hold for "almost all" trajectories. (Everett 1957, p. 460)

This passage is important for two reasons. First, Everett uses the word "typical" in passing. Given the uses of similar notions before 1957, Everett seems to address a community that is familiar with the main idea behind typicality, especially those who have worked on ergodicity. Second, with Everett's mentioning of the word "typical" he started to unify previous notions of *almost all*, *almost everywhere*, and *almost always true*. This may be a bit of an over-interpretation, but Everett mentions the

different uses of *almost all* in Statistical Mechanics, and he wants to use these ideas in quantum mechanics too.

I think it is not farfetched to set 1877 and 1957 as two historical milestones for the history of typicality, according to current historical knowledge. 1877 is important because Boltzmann explicitly mentions for the first time in a published paper that the Second Law of Thermodynamics when reduced to Statistical Mechanics is not a universally valid law and that "one can only prove that the number of initial states leading to a uniform state is infinitely larger than that of initial states leading to a non-uniform state after a given long time." 1957 is important because it is apparently the first time that typicality was used in a physical context.

Going back further in time, 1713 is another milestone for the history of typicality, as Jakob Bernoulli's work *Ars Conjectandi* was published in this year, in which he proved the Law of Large Numbers that he based on the notion of *moral certainty*, which is another concept close to typicality. The periods 1713–1877 and 1877–1957 pose their own challenges with respect to the historical development of typicality: Who worked on similar concepts between 1713–1877, and did they have a particular influence *on* Boltzmann and the physicists working in Statistical Mechanics at the time? Who worked on similar concepts between 1877–1957, and were they influenced *by* Boltzmann and the physicists working in Statistical Mechanics at the time? These questions are largely unanswered up to now, although excellent historical work has been done in the history of probability and measure theory that deals with formalizing the idea of *almost everywhere* in the period after 1877 (see, for instance, von Plato 1994; Hald 1998).

In this paper, however, I want to investigate the *pre-history* of typicality, namely the period before 1713. This early period is important not only for historical reasons:

1. The period before 1713 developed essential ideas for subsequent periods.
2. We notice a long tradition, in which philosophers and mathematicians tried to get a grip on uncertainty.
3. We can understand the essential idea of typicality by investigating its historical precursor *moral certainty*.
4. There has been a recent interest to develop in detail a theory of probability that is grounded on typicality (Dürr, Froemel, and Kolb, 2017; Maudlin, 2020; Hubert, 2021; Allori, 2022).[4] We can appreciate this research project by studying how Jakob Bernoulli based his Law of Large Numbers on the notion of moral certainty, which was a standard notion for a couple of centuries before.

In the following, I will examine in chronological order in which historical situations people came up with ideas similar to typicality. My endpoint will be Bernoulli's *Ars Conjectandi* published posthumously in 1713.

[4] The fundamental idea to relate typicality with probability has been expressed in Dürr, Goldstein, and Zanghì (1992); Goldstein (2001); Dürr and Teufel (2009); Goldstein (2012).

2 Aristotle: Scientific Demonstrations Versus Dialectial Deductions

Aristotle distinguished two basic kinds of explanations. In the *Posterior Analytics*, he defines a scientific explanation as a demonstration, a deductive proof, from first principles:

> By a demonstration I mean a scientific deduction; and by scientific I mean a deduction by possessing which we understand something. If to understand something is what we have posited it to be, then demonstrative understanding in particular must proceed from items which are true and primitive and immediate and more familiar than and prior to and explanatory of the conclusions. (Aristotle, Posterior Analytics, 71b15-25, translated in Aristotle 1994, pp. 2–3)

The first principles a scientific demonstration relies on are true propositions that have a distinguished epistemic status of being better known than the propositions that logically follow. They are found by induction, which is successful because we have intuition (famously stated at the end of the *Posterior Analytics* and also mentioned in the *Nicomachean Ethics*, 1139b19-39, 1140b30-1141a8; Bronstein 2012, 2016 for a first-class recent commentary, as well as Gerson 2009, Chap. 4, and Salmieri et al. 2014). Scientific demonstrations pose a subclass of deductions. If one can logically deduce from first principles a proposition about some state of affairs, one has a scientific explanation and scientific knowledge of these state of affairs, if the deduction elucidates the causes of the explanandum. Standard examples of first principles are universal claims. We can, for example, deduce from *All ravens are birds* and from *All birds have wings* that *All ravens have wings*. These kinds of syllogisms are the archetype of scientific explanations for Aristotle.

Aristotle is aware that scientific demonstrations are a very special group of arguments that require a particularly high standard of accuracy that we can only demand as an ideal[5] in the sciences; therefore, for intellectual training, casual encounters, and philosophical inquiry, we need to lower this high bar (*Topics*, Book 1, Chaps. 1–3). In these areas, one needs to replace scientific demonstrations with *dialectial deductions* (Greek: *enthumêma*, another subtype of deductions). Dialectical deductions have premises that are common beliefs (or reputable or probable opinions). As Aristotle says, "The common beliefs are the things believed by everyone or by most people or by the wise (and among the wise by all or by most or by those most known and commonly recognized)" (Aristotle 1995, *Topics*, 100b20–b25). The conclusions of dialectical deductions do not need to be true but at least sufficiently convincing.

In the ethical or political realm, Aristotle argues for a precision of arguments that lies between scientific demonstrations and those dialectical deductions:

[5] Pasnau (2013) convincingly argues that epistemology in its entire history from (at least) Aristotle until very recently has been mostly focusing on epistemic ideals. Hubert and Malfatti (2022) endeavor to revive this tradition for a modern theory of understanding (for a meta-epistemological debate, see Carr 2022; Thorstad 2023).

> Therefore in discussing subjects and arguing from evidence, conditioned in this way, we must be satisfied with a broad outline of the truth; that is, in arguing about what is for the most part so from premisses which are for the most part true we must be content to draw conclusions that are similarly qualified. […] for it is a mark of the trained mind never to expect more precision in the treatment of any subject than the nature of that subject permits; for demanding logical demonstrations from a teacher of rhetoric is clearly about as reasonable as accepting mere plausibility from a mathematician. (Aristotle, Nicomachean Ethics, 1094b15-25, translated in Aristotle 2004, p. 5)

The premises we use in ethical or political explanations are not universal truths but, at best, true *in most cases*. Therefore, the conclusions we reach from these premises are only true *in most cases* as well. Asking for universal truths in ethics and politics is not only impossible for all practical purposes, but it would also miss the point of what these disciplines are about in the first place, for "[i]t is a mark of the trained mind" to realize what kinds of explanations to expect in a particular context. G. W. Leibniz echoed Aristotle when he writes in 1670, "Only that degree of *certainty* is to be had which a given matter admits" (translated in Leibniz 1989, p. 122).

There seems to be a shift in the scope of the premises in dialectical deductions as introduced in the *Topics* and in the *Nicomachean Ethics*. In the *Topics*, the acceptance of the premises is purely epistemic depending on how many and which people support them, and these premises justify the persuasiveness of the conclusion (given the validity of the deduction). This appears to be close to a Bayesian approach, but rather one basing the degree of belief of a proposition on the beliefs of a collective not of a single agent.

In the *Nicomachean Ethics*, on the other hand, the premises are said to be true "for the most part", which can be interpreted in (at least) two ways: either in an epistemic (collective Bayesian) way, such that they are regarded as true by most people, or (as I think is more plausible) in an ontological–statistical way, such that the premises are true in most instances or in most situations. It is this interpretation that captures the central idea of typicality, that something is valid only *in most cases* and not universally valid. But before typicality has been developed in the 20th century, the notion of *moral certainty* elaborated on Aristotle's idea and was a familiar term among philosophers thereafter.

3 Gerson: The Inventor of Moral Certainty

Scholastic philosophy combined Christian theology with Ancient Greek philosophy; especially Aristotle's philosophy played a pivotal role for metaphysics and epistemology. A major project in scholastic philosophy was moral philosophy with the aim of finding rational grounds for how to act morally. The standard approach was to study particular real cases and to extract from them general principles for moral behavior (this method is called *casuistry*, from Latin *casus* meaning "case") that are the guiding principles for other (future) occasions. Today, casuistry is still a popular method in ethics, especially in business ethics, bioethics, and ethics of AI.

A reasonable standard method in casuistry is to find out what the best reasons are for moral behavior. During the Renaissance era, this method was challenged by two skeptical schools (see Schüssler 2009, for a detailed analysis, which I heavily rely upon in this section): (i) Neo-Pyrrhonism and (ii) Probabilism (*doctrina probabilatis*).

Neo-Pyrrhonism questioned the reliability of weighing different reasons for action against each other aiming for a suspension of judgement (Floridi 2002). The Probabilists were less skeptical and argued that it is possible to act morally despite not knowing the best reasons, as long as sufficient rationality standards were followed.

Neo-Pyrrhonism and Probabilism noticed that we act and need to act under uncertainty without knowing the best reasons. A problem was that people became too anxious in their actions under these circumstances—this type of exaggerated irrational anxiety was termed *scruples* (Latin: *scrupuli*) at this time. The Parisian Christian theologian Jean Gerson (1363–1429) tried to solve this dilemma: although it is unrealistic to have infallible knowledge in our decision-making processes, we can aim for sufficient knowledge for our behavior so that we do not need to suffer from scruples, that is, we can have a clear conscience (Gerson, of course, addressed the Medieval Christian community). Building on Aristotle's insight that "Only that degree of *certainty* is to be had which a given matter admits" (translated in Leibniz 1989, p. 122), Gerson coined the term *moral certainty* (Latin: *certitudo moralis*) for the appropriate level of certainty for moral actions that is not deemed sinful and that is not justified to be followed by scruples:

> There is, however, a moral certainty, which in our purpose is required, and which suffices. And this we have, when, in our recollection and examen of conscience, we find we have done that, which both our own discretion and the good counsel of others suggested, and have for some time been wont commonly so to do. But if our own judgment should not accuse us of mortal sin, then there is no new peril in going to holy communion, […] (Gerson 1883, pp. 57)

There are essentially three paths to reach moral certainty (Gerson 1883, p. 40): (i) through other people, (ii) through the study of scripture, and (iii) by using your own faculty of reason. Gerson's justification of moral certainty is similar to Aristotle's justification of common beliefs as the premises in a dialectical deduction: "The common beliefs are the things believed by everyone or by most people or by the wise (and among the wise by all or by most or by those most known and commonly recognized)" (Aristotle 1995, Topics, 100b20–b25). Aristotle also emphasized the reliance on other people, but, for obvious reasons, he did not add scripture as a means for justification. Reason, for Aristotle, seem to rather play a role for making deductions or for justifying first principles by intuition.

Gerson continues the above passage and adds important details about moral certainty:

> […] even though, as it may often happen, some slight doubts may come into our mind. These doubts we ought to repel, and we ought to force ourselves to act contrary to them. I call that a slight doubt, when a person judges of a thing, rather that it is just and good, than that it is evil; yet some reasons or thoughts occur to the mind, leading to some hesitation, but still the first judgment appears far the most certain. Now if both sides seem equally probable, we

> ought to stop till we get more ground for decision one side or other, either by the help of our own reason, or by consultation with others, or by a divine inspiration obtained through prayer. For unless in this mode a person obtain security in himself, he will always judge that he has made a bad confession, and will never feel easy or at peace, and this can never be good. (Gerson 1883, pp. 57–58)

Moral certainty still leaves room for *slight doubt*, because there always remains a degree of uncertainty. Still, a good Christian can act upon moral certainty and should not suffer from anxiety that this action is sinful, because everything that could be done and considered has been taken care of. If this person does not understand this, the consequence would be that one remains unhappy and burdened by fear of having done a sinful act, even when one had done everything humanly possible to act morally. Like the Pyrrhonic and Neo-Pyrrhonic Skeptics, Gerson aimed at calmness of the mind (Greek: *ataraxia*). Whereas the Pyrrhonic Skeptics thought to reach this state by suspension of judgement, as for any argument there can be found a counterargument, Gerson suggested to reasonably lower the bar for certainty that is suitable for the average Christian to act morally without suffering from scruples. Political and ethical experts, on the other hand, can reach a higher level of certainty closer to Aristotle's proposed level in the *Nicomachean Ethics* (Schüssler 2009, p. 453).

After Gerson invented *moral certainty* it has become a standard concept in scholastic philosophy.[6] It seems to have been so prevalent and useful that even stark critics of scholasticism, like René Descartes (1596–1650), used it without any scruples.

4 Descartes: Moral Certainty Versus Absolute/Metaphysical Certainty

Descartes used the concept of moral certainty implicitly in his *Discourse on the Method* (1637) and explicitly in *The Principles of Philosophy* (1644) – a detailed analysis of these works with respect to moral certainty can be found in Ariew (2011) and Samjetsabam (2022). In the spirit of Gerson, Descartes wrote in Part 3 in the *Discourse on the Method*:

> Similarly, since in everyday life we must often act without delay, it is a most certain truth that when it is not in our power to discern the truest opinions, we must follow the most probable. Even when no opinions appear more probable than any others, we must still adopt some; and having done so we must then regard them not as doubtful, from a practical point of view, but as most true and certain, on the grounds that the reason which made us adopt them is itself true and certain. (Descartes 1985, p. 123)

[6] In the 18th century, the concept of moral certainty entered the Anglo-Saxon legal system, and it merged into the concept of *reasonable doubt* by 1824 (Waldman 1959). Other concepts that were used before moral certainty in the judicial system were *satisfied conscience* and *satisfied understanding* (Shapiro 2012, p. 20).

Descartes gives a here a manual to adjust our degree of certainty. The best would be, of course, to find out the truth. Since we need to make quick decisions in our practical life, we often do not have the time to thoroughly examine our thoughts and ideas (and even if we had we may not figure out the truth). In this case, we need to "follow the most probable." But this situation of having a most probable opinion is not guaranteed, and still we need to choose one of the available options. If we did so with good reason, we can be morally certain about the correctness of the opinion (even if Descartes does not explicitly mention moral certainty here). Descartes point is that our acts would be at least epistemically justified when we are morally certain of them. This seems to be a slightly different position from Gerson's who seems to say something stronger, namely, that morally certain acts are morally justified so that we would not suffer scruples. It seems plausible that Descartes thinks along these lines too, but he does not say so explicitly.

For his theoretical project, on the other hand, Descartes wants to treat any proposition that can be doubted however slightly as absolutely false, as he says so in Part 4 of the *Discourse*[7]:

> For a long time I had observed, as noted above [in Part 3], that in practical life it is sometimes necessary to act upon opinions which one knows to be quite uncertain just as if they were indubitable. But since I now wished to devote myself solely to the search for truth, I thought it necessary to do the very opposite and reject as if absolutely false everything in which I could imagine the least doubt, in order to see if I was left believing anything that was entirely indubitable. (Descartes 1985, pp. 126–127)

Descartes introduces here his *methodological skepticism*, in which he merely *pretends* that certain opinions are false if he can doubt them in the slightest. His goal is to find those opinions that he cannot doubt in the slightest in order to establish a firm foundation for his epistemology and natural philosophy. Descartes is here about to use moral certainty within natural philosophy and so outside the realm of theology and moral philosophy (Wootton 2015, Chap. 11, Sect. 3, mentions that moral certainty was used by the first statisticians by 1662).

In a subsequent passage in Part 4 of the *Discourse on the Method*, Descartes clarifies his distinction between moral certainty (French: *assurance morale*) and metaphysical certainty (French: *certitude métaphysique*):

> Finally, if there are still people who are not sufficiently convinced of the existence of God and of their soul by the argument I have proposed, I would have them know that everything else of which they may think themselves more sure—such as their having a body, there being stars and an earth, and the like—is less certain. For although we have a moral certainty about these things, so that it seems we cannot doubt them without being extravagant, nevertheless when it is a question of metaphysical certainty, we cannot reasonably deny that there are adequate grounds for not being entirely sure about them. (Descartes 1985, pp. 129–130)

[7] Descartes' strong position leads to the following problem (I thank Charles Sebens for pointing this out). If Descartes treats the proposition p as false, he would need to treat the proposition $\neg p$ as false, too, as both can be doubted. It would seem to be more reasonable to interpret Descartes' position as withholding belief in p if p can be doubted.

Domain	Method	Certainty
- existence of God - existence of the human mind - mathematics	- recognition of innate ideas - radical skepticism - mathematical demonstration	absolute or metaphysical certainty
- natural philosophy in particular: - existence of the outside world - existence of the human body	Descartes' scientific arguments and results	moral certainty

Fig. 1 Descartes' distinction between absolute and moral certainty. In the *Discourse* and in the *Principles*, Descartes directly argues for the metaphysical or absolute certainty of the existence of God and the existence of the human mind (the human soul). About the existence of the outside world and the human body, Descartes first showed that we can be morally certain; in a second step, he lifts the moral certainty about their existence and even their features to the level of absolute certainty

We have metaphysical certainty of a certain idea or proposition, when we have no rational ground for doubts. Descartes claims that he proved the existence of God and of the human soul to be metaphysically certain. On the other hand, morally certain claims are less certain, but they are sufficiently certain that they can be only doubted on "extravagant" grounds. The existence of the outside world (that humans have a body, that the earth and the stars exist) is morally certain, because we can only doubt them by an extravagant thought experiments, such as a dream (see Fig. 1).

Descartes' use of moral certainty is different from Gerson's in two important respects. First, Descartes applies moral certainty beyond moral philosophy. He is primarily interested in existence claims about the natural world. Although these existence claims are indeed relevant for our practical actions, they are rather the preconditions for dealing with the world. Second, Descartes sets a higher bar for moral certainty than Gerson. For all practical purposes, morally certain existence claims about the world can be in principle doubted, but these doubts do not justify us in not believing these claims. Gerson, on the other hand, focuses on morally right behavior which does not require so high a degree of certainty as our belief in an outside world.

At the end of his *Principles of Philosophy*, Descartes examines how certain his philosophical edifice is to be the true story of the world. In §205 of Part IV, he concludes that his "explanations appear at least to be morally certain," and he defines moral certainty in the following way:

> moral certainty [Latin: *certa moraliter*] is certainty which is sufficient to regulate our behavior, or which measures up to the certainty we have on matters relating to the conduct of life, which we never normally doubt, though we know that it is possible, absolutely speaking, that they may be false. (translated in Descartes 1985, p. 289)

Unlike Aristotle, Descartes focuses his distinction on the states of belief of a *single* agent. There are many things in our daily life we have good reasons to take for granted, because they are useful or beneficial for us, although we are not absolutely certain about them (see also Simmons 2001). Even if you have never been to Kuala Lumpur,

you are morally certain that this city exists, because of the pictures you have seen and the testimonies of other people, and you would buy a ticket if you wish to visit this city lacking absolute certainty of its existence. Descartes gave a similar example with the city of Rome, and then poses an analogy that his theory is morally certain: Imagine you have an encrypted message and you stumble on a key that is able to decode this message into some sensible text, a string of words that is grammatically correct and meaningful. Then you can be morally certain, certain beyond any reasonable doubt or certain for all practical purposes, that this was the correct key, although it is logically and physically possible that the writer of the message indeed used a different key that would result in a different message. By analogy, Descartes argues, that we can be morally certain about the truth of the axioms that he introduces in the *Principles of Philosophy* because they coherently decode and explain the natural world. Different axioms or principles would be logically possible but highly implausible or unlikely, given the success of his system.

In this context, moral certainty is related to the *inference to the best explanation* (Lipton 2004). The best explanation that a certain key will decrypt an encrypted message into a sensible text is that this key is the right key. Given the accuracy and success of Descartes' system, it best explains the behavior of the world. It seems that if X is the best explanation for Y, we can be more certain about the truth of X than the alternative explanations X'. As Lipton (2000) says, "Inference to the Best Explanation must thus be glossed by the more accurate but less memorable phrase, 'inference to the best of the available competing explanations, when the best one is sufficiently good'." If X is *by far* the best explanation of Y then we may be morally certain about the truth of X.[8] Moral certainty is also used in other situations, like in moral behavior, where it does not make sense to talk about a best *explanation*, but rather about the right moral behavior or the best moral behavior. We may call this situation *inference to the best moral behavior*.

In the following paragraph (§206), Descartes argues that certain aspects of his theory of the world are even *absolutely certain* (Latin: *certa absolute*):

> Absolute certainty arises when we believe that it is wholly impossible that something should be otherwise than we judge it to be. This certainty is based on a metaphysical foundation, namely that God is supremely good and in no way a deceiver, and hence that the faculty which he gave us for distinguishing truth from falsehood cannot leas us into error, so long as we are suing it properly and are thereby perceiving something distinctly. Mathematical demonstrations have this kind of certainty, as does the knowledge that material things exist; and the same goes for all evident reasoning about material things. And perhaps even these results of mine will be allowed into the class of absolute certainties, if people consider how they have been deduced in an unbroken chain from first and simple principles of human knowledge. (translated inch31Descartes:1985aa, p. 290)

While Descartes referred to this degree of certainty in the *Discourse* as *metaphysical certainty*, he changed the terminology in the *Principles* to *absolute certainty*. Apart from the existence of God, the existence of the human soul, and mathematical results, Descartes wants to argue that even some features of his natural philosophy

[8] I thank Charles Sebens for pointing out the subtleties regarding the relationship between moral certainty and the inference to the best explanation.

are justified to be true to a higher degree than moral certainty.[9] He gives two reasons. First, since God is not a deceiver, He created us so that we are able to have epistemic access to the true structure of the world. Second, Descartes deduced his results from true first principles.

5 Locke: Probability Versus Real Certainty

The period after Descartes, especially in 17th century England, was marked by meticulous refinements of theories explaining and guiding an agent's degree of certainty given a specific domain and specific arguments within the domain (Shapiro 1985, Chap. 2). Scholars, at this time, largely agreed that demonstrations, like mathematical proofs, lead to the highest level of certainty humans can attain, and they also largely agreed that one cannot demand this high epistemic standard for other disciplines, like the natural sciences, history, law, or even religion. They disagreed, however, on three things: (i) the domain for reaching moral certainty, (ii) the level of certainty for moral certainty, (iii) levels of certainty below moral certainty.

As different scholars made different proposals for hierarchies of knowledge, John Locke (1632–1704) wrote *An Essay Concerning Human Understanding* in 1689 to synthesize this complex debate, which had a huge influence later on. Although much of Locke scholarship went into analyzing his theory of knowledge, it seems that Locke's aim was to elucidate how one can properly make rational *probable* judgements (Wolterstorff 1996; Boespflug 2023). For our purposes, I want to focus on two aspects of Locke's *Essay*: (i) he subsumed absolute and moral certainty within his notion of *real certainty* (Enquiry, Book IV, Chap. IV), (ii) for probable reasoning, he establishes, what Wolterstorff (1996, p. 79) calls *The Principle of Proportionality* (Enquiry, Book IV, Chap. XV & XVI), which is the historical ancestor of David Lewis's *Principal Principle* (Lewis 1981).

Let us first discuss Locke's notion of real certainty. For Locke, the highest form of certainty is attained through knowledge, where he defines knowledge as *the perception of the agreement and disagreement of our own ideas*. Ideas are the contents of our mind, and we have the capacity to compare these ideas. For example, we have an idea of the number 2 and an idea of the number 3. By comparing these ideas of these numbers, we find out that 2 is smaller than 3; therefore, we know that 2 is smaller than 3. Nevertheless, not all types of knowledge lead to the same degree of certainty: thus, Locke introduces a hierarchy of three types of knowledge (see Fig. 2):

1. Intuitive knowledge (Book IV, Chap. II, Sect. 1),
2. Demonstrative knowledge (Book IV, Chap. II, Sects. 2–12),
3. Sensitive Knowledge (Book IV, Chap. II, Sect. 14).

[9] Descartes also strives for the absolute certainty of the outside world in his *Meditations*, which I have left out in this section for lack of space.

Domain	Method	Certainty
relations between simple ideas	direct intuition	intuitive knowledge
- mathematics - moral philosophy - possibly other areas	- reasoning - proofs	- demonstrative knowledge - real certainty
- existence of the external world - existence of particular external objects	- unclear - God connecting simple ideas with the world	sensitive knowledge
- natural philosophy - 17th century science - other areas	- conformity with one's own experience - testimony	- belief, opinion - degrees of assent

Fig. 2 Locke's Hierarchies of Knowledge and Probability (Locke 1689/1997, Book IV, Chap. II)

Intuitive knowledge is reached when the mind directly or immediately recognizes the relation between simple ideas, for example, that white is not black. One level down the ladder is demonstrative knowledge, which can be reached once the mind has found a proof to connect two ideas upon reasoning. Mathematics is the prime example for demonstrative knowledge. Another level down the ladder, we find sensitive knowledge. This is the knowledge that there exists an external world with particular external objects beyond our mere ideas in our minds. Locke's arguments for this type of knowledge in Book IV, Chap. II, Sect. 14 and Chap. IV, Sects. 2–5 is not clear. His main argument, however, sounds Cartesian, because simple ideas match the external world by the acts of God.[10] Another step down the ladder, Locke locates belief, opinion, and different degrees of assent. Among other domains, natural philosophy and 17th century science are notable cases here (see also Boespflug 2023, Chap. 6), and testimony from other people become crucial, too, as another method for reaching one's degree of assent.

Building on his justification for the reality of simple ideas in Book IV, Chap. IV, Sects. 2–5, Locke argues that it follows that mathematical knowledge (through proofs) is *real knowledge* (Sect. 6). This argument prepares Locke to break with the tradition of moral certainty—he does not even mention this concept at all in his *Essay*. Remember that moral certainty was originally used as a degree of certainty for moral claims that one can be certain of for all practical purposes. But for Locke, the domain of morality allows for a higher form of certainty, because it is on a par with mathematics, since moral claims can be *proven*. A correct moral proof will lead to *real certainty*, which is the same certainty as attained through mathematical demonstrations (see Fig. 2)[11]:

[10] "*First*, the first are simple ideas, which since the mind, as has been showed, can by no means make to itself, must necessarily be the product of things operating on the mind in a natural way, and producing therein those perceptions which by the wisdom and will of our Maker thy are ordained and adapted to." (Book IV, Chap. IV, Sect. 4).

[11] Locke used *real certainty* only in this context and did not develop this notion further elsewhere.

> And hence it follows that *moral knowledge* is as *capable of real certainty* as mathematics. For certainty being but the perception of the agreement or disagreement of our ideas, and demonstration nothing but the perception of such agreement, by the intervention of other ideas or mediums; our *moral ideas*, as well as mathematical, being *archetypes* themselves, and so adequate and complete ideas; all the agreement or disagreement which we shall find in them will produce real knowledge, as well as in mathematical figures. (Locke 1689/1997, Book IV, Chap. IV, Sect. 7)

Locke sounds overtly optimistic about our ability for moral knowledge. In contrast to his predecessors, Locke thought in this passage that it is indeed possible to demonstrate moral propositions similar to mathematical propositions. Later, however, he realized that this is not possible (see Boespflug 2023, Chap. 7, for details).

Let us now turn to Locke's theory of probable reasoning (last row in Fig. 2) which will culminate in the *Principle of Proportionality*. First, Locke defines probability in contrast to demonstration:

> As demonstration is the showing the agreement, or disagreement of two ideas, by the intervention of one or more proofs, which have a constant, immutable, and visible connexion one with another: so *probability* is nothing but the appearance of such an agreement, or disagreement, by the intervention of proofs, whose connexion is not constant and immutable, or at least is not perceived to be so, but is, or appears for the most part to be so, and is enough to induce the mind to *judge* the proposition to be true, or false, rather than the contrary. (1689/1997, Book IV, Chap. XV, Sect. 1)

Whereas demonstrations are logical deductions that generate truths, probability signifies a scheme of reasoning that does not rely on logical entailments but rather on entailments that are valid "for the most part." Locke realizes that these kinds of looser arguments, which are prevalent in our everyday life, are still strong enough to be believed to a certain degree, as "[Man] would be often utterly in the dark, and in most of the Actions of his Life, perfectly at a stand, had he nothing to guide him in the absence of clear and certain Knowledge." (Book IV, Chap. XIV Sect. 1).

Since Locke aims to give us a hierarchy of probability, he has to tell us along which criteria we can evaluate the probability of a proposition:

> *Probability* then, being to supply the defect of our knowledge, and to guide us where that fails, is always conversant about Propositions, whereof we have no certainty, but only some inducements to receive them. The grounds of it are, in short, these *two* following:
>
> *First*, the conformity of any thing with our own knowledge, observation, and experience. *Secondly*, the testimony of others, vouching their observation and experience. (Locke 1689/1997, Book IV, Chap. XV, Sect. 4)

Locke presents two criteria for evaluating the probability of a proposition: (i) how much the arguments put forward agree (or conform) with our previous knowledge, observation, and experience, (ii) the trustworthiness of the testimony of others.[12] When we evaluate a proposition along these two axes, we can assign it a degree of

[12] Locke warns us to take *the opinions of others* as a criterion for probability, though (Book IV, Chap. XV, Sect. 6).

probability in the following order: (i) demonstration, (ii) improbability, (iii) unlikeliness, (iv) impossibility (Book IV, Chap. XV, Sect. 2). Having done this, the question is now how to adjust one's degree of assent according to this categorization. This is where Locke introduces his *Principle of Proportionality*:

> *the mind if it will proceed rationally ought to examine all the grounds of probability*, and see how they make more or less, for or against any probable proposition, before it assents to or dissents from it, and upon a due balancing the whole, reject or receive it, with a more or less firm assent, proportionably to the preponderancy of the greater grounds of probability on one side or the other. (Locke 1689/1997, Book IV, Chap. XV, Sect. 5)

Roughly, the *Principle of Proportionality* says that one should adjust one's assent (or degree of certainty) in accordance to the probability of the proposition.[13] For the degree of assent, Locke proposes this hierarchy: (i) full assurance and confidence, (ii) conjecture, (iii) doubt, and (iv) distrust.

With the invention of probability calculus, Locke's *Principle of Proportionality* would get more fine-grained steps for both probability and assent. Jakob Bernoulli will prove a crucial theorem in probability theory, the Law of Large Numbers, which combines combines probability, assent, and frequencies.

6 Jakob Bernoulli: Moral Certainty and the Law of Large Numbers

In this work, Bernoulli introduced and rigorously proved the first limit theorem in mathematics, which Siméon Denis Poisson later called the *Law of Large Numbers* and of which Bernoulli himself said later on, "I esteem this discovery more than if I had given the quadrature of the circle itself, which even if it were found very great, would be of little use" (quoted in Sylla 2016, p. 50).

Bernoulli's goal was to find out the best way to assign probabilities to certain events. One can determine the probability of a coin toss by examining the physical qualities and the tossing mechanism of the coin toss. If the coin is symmetrical and the tossing mechanism is not prepared in a special way, the coin has probability ½ to land heads or tails; this method is in Bernoulli's terminology a priori, because it is about dissecting the causes of the phenomenon (Bernoulli 2006, p. 327).

Bernoulli suggested that it had to be possible to assign the right probabilities also from empirical data, even in cases where no a priori method is possible. This is the a posteriori method of assigning probabilities:

> What cannot be ascertained a priori, may at least be found out a posteriori from the results many times observed in similar situations, since it should be presumed that something can happen or not happen in the future in as many cases as it was observed to happen or not to happen in similar circumstances in the past (Bernoulli 2006, p. 327).

[13] This is very close to the *Principal Principle*. A detailed analysis of these two principles would go beyond the scope of this paper.

Making this a posteriori strategy mathematically rigorous was the aim of the Law of Large Numbers. Consider Bernoulli's example of a large urn consisting of $1/3$ black and $2/3$ white balls (mentioned in a letter to Leibniz, quoted in Bernoulli 2006, p. 40, also in the *Ars Conjectandi* on p. 328)—in this case, $1/3$ and $2/3$ would be the a priori probabilities. The Law of Large Numbers in Bernoulli's interpretation says that, under these circumstances, you may reach moral certainty after a sufficiently long trial that the true proportion does not deviate too much from $2/3$ (more precisely, that it is within $[2/3 - \epsilon, 2/3 + \epsilon]$ for any given ϵ).

To answer the above question in full mathematical detail, Bernoulli needs a formal definition of moral certainty:

> Something is *morally certain* if its probability comes so close to complete certainty that the difference cannot be perceived. By contrast, something is *morally impossible* if it has only as much certainty as the amount by which moral certainty falls short of complete certainty. Thus if we take something that possesses $999/1000$ of certainty to be morally certain, then something that has only $1/1000$ of certainty will be morally impossible. (Bernoulli 2006, p. 316)

Two things are remarkable in this quote. First, Bernoulli introduces precise numbers when to reach moral certainty. Second, these numbers are not universally valid in all circumstances. Moral certainty may have a different mathematical number in different settings (this is also an important feature of the formalization of typicality, see Maudlin 2020). Bernoulli is explicit on that later on:

> It would be useful accordingly, if definite limits for moral certainty were established by the authority of the magistracy. For instance, it might be determined whether $99/100$ of certainty suffices or whether $999/1000$ is required. Then a judge would not be able to favor one side, but would have a reference point to keep constantly in mind in pronouncing a judgment. (Bernoulli 2006, p. 321)

There is no unanimously agreed number for what exactly counts as morally certain: in one case $99/100$ is sufficient, in another $999/1000$ may be required. Bernoulli managed to give a mathematical formalization of moral certainty and in tandem a mathematical theorem, which states under which circumstances one can expect certain empirical results with moral certainty. For Bernoulli, probabilities are, on the one hand, fractions of certainty, degrees of belief in modern parlance (Sylla 2016, p. 62); on the other hand, Bernoulli connected these a posteriori probabilities with frequencies for how often one would pick a black or white ball from an urn, for example.

In a modern approach to ground probability on typicality, Bernoulli's picture is turned upside-down: the frequencies are identified with probability, while typicality replaces moral certainty (Dürr et al. 2017; Maudlin 2020; Hubert 2021). Typicality arises from counting physical degrees of freedom and is therefore by definition not reflecting degrees of belief, but it is *formally* astoundingly similar to Bernoulli's notion of moral certainty. To bridge the gap between counting physical degrees of freedom and certainty, Wilhelm (2022) invokes a new principle of rationality, the *Typical Principle*, to connect typicality claims with belief about these claims. If a proposition makes a typicality claim, an agent should believe this claim (as long as the agent has no further information that would undermine this claim). One cannot be

absolutely certain about typicality claims because it is possible that they are violated in a particular case; instead, an agent may be morally certain of them.

7 Conclusion

Although moral certainty had a long tradition in philosophy, it is no longer used. With recent works in statistical mechanics and Bohmian mechanics, we have with typicality in fact a notion that is similar to moral certainty. Typicality formalizes the idea of *almost all* by counting physical degrees of freedom by means of a measure, whereas moral certainty is an epistemic notion that we can be *almost certain* about the truth of a proposition. Therefore, we may regard moral certainty as the epistemic counterpart of typicality. It requires, however, a deeper analysis to compare both concepts in detail, but my aim here was a historical one: the history of moral certainty can be regarded as the pre-history of typicality.

Acknowledgements I started to write this paper in 2019 when I was a Postdoctoral Instructor at Caltech, and I wish to thank the Division of the Humanities and Social Sciences for the support and resources I have received at the time that made this research possible in the first place. I wish to thank the participants of the *Caltech Probability Seminar* for helpful discussions. I thank Frederick Eberhardt, Christopher Hitchcock, and Charles Sebens for their insightful comments on a previous draft of the paper. I also thank Ariane Schneck for her helpful comments on the section on Descartes, and I thank Mark Boespflug for his helpful comments on the section on Locke. I wish to thank Barry Loewer, who was the first to tell me about the relation between moral certainty and typicality. I also thank Tim Maudlin for encouraging me to delve into the history of philosophy and for the many hours we spoke about typicality. Lastly, I thank an anonymous referee for helpful comments, especially on clarifying the recent history of typicality. I am very grateful for having been one of Detlef Dürr's students; his teaching and magnanimity have always inspired me.

References

V. Allori, The paradox of deterministic probabilities. Inquiry (2022). https://doi.org/10.1080/0020174X.2022.2065530

R. Ariew, The new matter theory and its epistemology: descartes (and late scholastics) on hypotheses and moral certainty, in *Vanishing Matter and the Laws of Motion: Descartes and Beyond, Routledge Studies in Seventeenth Century Philosophy, Chapter 2*. ed. by P. Anstey, D. Jalobeanu (Routledge, London, 2011), pp.31–46

Aristotle. *Posterior Analytics*, 2nd ed. (Clarendon Press, Oxford, 1994). Translated with a commentary by Jonathan Barnes

Aristotle. *Selections*. Indianapolis (Hackett Publishing, IN, 1995). Translated, with Introduction, Notes, and Glossary by Terence Irwin and Gail Fine

Aristotle. *The Nicomachean Ethics*. London: Penguin, *Translated by J* (A. K, Thomson, 2004)

J. Bernoulli, *The Art of Conjecturing* (The Johns Hopkins University Press, Baltimore, 2006). Translated with an introduction and notes by Edith Dudley Sylla

M. Boespflug, *Locke's Twilight of Probability: An Epistemology of Rational Assent* (Routledge, New York, 2023)

J. Bricmont, *Making Sense of Statistical Mechanics* (Springer International Publishing, Switzerland, 2022)
D. Bronstein, The origin and aim of "Posterior Analytics" II.19. Phronesis **57**(1), 29–62 (2012)
D. Bronstein, *Aristotle on Knowledge and Learning: The Posterior Analytics* (Oxford University Press, Oxford, 2016)
S.G. Brush, *The Kind of Motion We Call Heat: A History of the Kinetic Theory of Gases in the 19th Century*, volume 2—Statistical Physics and Irreversible Processes (Amsterdam, North-Holland, 1976)
J. Carr, Why ideal epistemology? Mind **131**, 1131–1162 (2022)
O. Darrigol, *Atoms, Mechanics, and Probability: Ludwig Boltzmann's Statistico-Mechanical Writings—An Exegesis* (Oxford University Press, Oxford, 2018)
O. Darrigol, J. Renn, The emergence of statistical mechanics, in *The Oxford Handbook of History of Physics, Chapter 25*. ed. by J.Z. Buchwald, R. Fox (Oxford University Press, Oxford, 2013), pp.765–88
R. Descartes, *The Philosophical Writings of Descartes*, vol. 1 (Cambridge University Press, Cambridge, UK, 1985). Translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch
D. Dürr, *Bohmsche Mechanik als Grundlage der Quantenmechanik* (Springer, Berlin, 2001)
D. Dürr, S. Teufel, *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory* (Springer, Berlin, 2009)
D. Dürr, S. Goldstein, N. Zanghì, Quantum equilibrium and the origin of absolute uncertainty. Journal of Statistical Physics **67**(5), 843–907 (1992)
D. Dürr, A. Froemel, M. Kolb, *Einführung in die Wahrscheinlichkeitstheorie als Theorie der Typizität* (Springer, Heidelberg, 2017)
H. Everett, "Relative state" formulation of quantum mechanics. Reviews of Modern Physics **29**(3), 454–62 (1957)
L. Floridi, *Sextus Empiricus: The Transmission and Recovery of Pyrrhonism* (Oxford University Press, Oxford, 2002)
R. Frigg, Typicality and the approach to equilibrium in Boltzmannian statistical mechanics. Philosophy of Science **76**(5), 997–1008 (2009)
R. Frigg, Why typicality does not explain the approach to equilibrium, in *Probabilities, Causes and Propensities in Physics, Chapter 4*. ed. by M. Suárez (Springer, Heidelberg, 2011), pp.77–93
R. Frigg, C. Werndl, Demystifying typicality. Philosophy of Science **79**(5), 917–29 (2012)
J. Gerson, *The Snares of Devil* (Thomas Richardson, London, 1883). https://archive.org/details/SnaresOfTheDevil/mode/2up. Translated by Beta
L.P. Gerson, *Ancient Epistemology* (Cambridge University Press, Cambridge, UK, 2009)
S. Goldstein, Boltzmann's approach to statistical mechanics, in *Chance in Physics: Foundations and Perspectives*. ed. by J. Bricmont, D. Dürr, M. Galavotti, G. Ghirardi, F. Petruccione, N. Zanghì (Springer, Heidelberg, 2001), pp.39–54
S. Goldstein, Typicality and notions of probability in physics, in *Probability in Physics, Chapter 4*. ed. by Y. Ben-Menahem, M. Hemmo (Springer, Heidelberg, 2012), pp.59–71
S. Goldstein, L.J. Lebowitz, R. Tumulka, N. Zanghì, Long-time behavior of macroscopic quantum systems. European Phys. J. H **35**(2), 173–200 (2010)
S. Goldstein, J.L. Lebowitz, R. Tumulka, N. Zanghì, Gibbs and Boltzmann entropy in classical and quantum mechanics, in *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, ed. by V. Allori, Chapter 14 (World Scientific, 2020), pp. 519–581
A. Hald, *A History of Mathematical Statistics from 1750 to 1930* (Wiley, 1998)
M. Hubert, Reviving frequentism. Synthese **199**, 5255–5284 (2021)
M. Hubert, F. Malfatti, Towards ideal understanding. *Ergo* (2022). Forthcoming
D. Lazarovici, P. Reichert, Typicality, irreversibility and the status of macroscopic laws. Erkenntnis **80**(4), 689–716 (2015)
J.L. Lebowitz, Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy. Physica A **194**, 1–27 (1993)
J.L. Lebowitz, Time's arrow and Boltzmann's entropy. Scholarpedia **3**(4), 3348 (2008)

G.W. Leibniz, *Philosophical Papers and Letters*, 2nd ed. (Kluwer, Dordrecht, 1989). Translated and edited, with an introduction by Leroy E. Loemker
D. Lewis, A subjectivist's guide to objective chance, in *IFS: Conditionals*. ed. by W.L. Harper, R. Stalnaker, G. Pearce. Belief, Decision, Chance and Time. (Springer, Netherlands, 1981), pp.267–97
P. Lipton, Inference to the best explanation, in *A Companion to the Philosophy of Science*. ed. by W.H. Newton-Smith (Blackwell, Oxford, 2000), pp.184–193
P. Lipton, *Inference to the Best Explanation* (Routledge, New York, 2004)
J. Locke, *An Essay Concerning Human Understanding*, ed by R. Woolhouse (Penguin, London, 1689/1997)
T. Maudlin, The grammar of typicality, in *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, ed. by V. Allori, Chapter 7 (World Scientific, 2020), pp. 231–51
W.C. Myrvold, Explaining thermodynamics: what remains to be done? in *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, ed. by V. Allori, chapter 4 (World Scientific, 2020), pp. 113–143
A. Oldofredi, D. Lazarovici, D.-A. Deckert, M. Esfeld, From the universe to subsystems: why quantum mechanics appears more stochastic than classical mechanics. Fluctuations and Noise Letters **15**, 1–16 (2016)
R. Pasnau, Epistemology idealized. Mind **122**, 987–1021 (2013)
K. Robertson, Search of the holy grail: how to reduce the second law of thermodynamics. British J. Philosop. Sci. (2021). Forthcoming
G. Salmieri, D. Bronstein, D. Charles, J.G. Lennox, Episteme, demonstration, and explanation: a fresh look at aristotle's Posterior Analytics. Metascience **23**(1), 1–35 (2014)
M. Samjetsabam, Moral certainty of faculty of reason in Descartes' discourse. Tattva J. Philos. **13**(2), 1–18 (2022). https://doi.org/10.12726/tjp.26.1
R. Schüssler, Jean Gerson, Moral Certainty and the Renaissance of Ancient Scepticism. Renaissance Stud. **23**(4), 445–462 (2009). ISSN 02691213, 14774658. http://www.jstor.org/stable/24419382
B.J. Shapiro, *Probability and Certainty in Seventeenth Century England* (Princeton University Press, Princeton, NJ, 1985)
B.J. Shapiro, Beyond reasonable doubt: The evolution of a concept, in *Fictions of Knowledge: Fact, Evidence, Doubt*, ed. by Y. Batsaki, S. Mukherji, J.-M. Schramm (Palgrave Macmillan UK, London, 2012), pp. 19–39. ISBN 978-0-230-35461-6. https://doi.org/10.1057/9780230354616_2. https://doi.org/10.1057/9780230354616_2
A. Simmons, Sensible ends: latent teleology in descartes' account of sensation. Journal of the History of Philosophy **39**(1), 49–75 (2001)
E.D. Sylla, Probability in the 17th and 18th century continental Europe from the perspective of Jacob Bernoulli's art of conjecturing, in *The Oxford Handbook of Probability and Philosophy, Chapter 3*. ed. by A. Hájek, C. Hitchcock (Oxford University Press, Oxford, 2016), pp.50–68
D. Thorstad, Why bounded rationality (in epistemology)? *Philosophy and Phenomenological Research* (2023). Forthcoming
S.B. Volchan, Probability as typicality. Stud. Hist. Philosop. Mod. Phys. **38**(4), 801–14 (2007)
J. von Plato, *Creating Modern Probability: Its Mathematics, Physics and Philosophy in Historical Perspective* (Cambridge University Press, Cambridge, UK, 1994)
G. Wagner, Typicality and minutes rectis laws: From physics to sociology. Journal for General Philosophy of Science (2020). https://doi.org/10.1007/s10838-020-09505-7
T. Waldman, Origins of the legal doctrine of reasonable doubt. J. Hist. Ideas **20**(3), 299–316 (1959). ISSN 00225037, 10863222. http://www.jstor.org/stable/2708111
I. Wilhelm, The typical principle. British J. Philosop. Sci. (2022). https://doi.org/10.1086/723240. https://doi.org/10.1086/723240
N. Wolterstorff, *John Locke and the Ethics of Belief*. Cambridge Studies in Religion and Critical Thought (Cambridge University Press, Cambridge, UK, 1996)

D. Wootton, *The Invention of Science: A New History of the Scientific Revolution* (HarperCollins Publishers, Madrid, 2015)