

POLISH ACADEMY OF SCIENCES
INSTITUTE OF PHYSICS

*Established in 1920 by
the Polish Physical Society*



ACTA PHYSICA POLONICA

- General Physics
- Atomic and Molecular Physics
- Condensed Matter
- Optics and Quantum Optics
- Quantum Information
- Biophysics
- Applied Physics

Special issue in honor of
scientific achievements of
Professor Iwo Białynicki-Birula
on his 90th birthday

Guest editor
Tomasz Sowiński



RECOGNIZED BY THE EUROPEAN
PHYSICAL SOCIETY

Volume 143 — Number 6, WARSAW, JUNE 2023

Editor-in-Chief:

Jan Mostowski

Associate Editors:

Anna Ciechan	Łukasz Cywiński
Elżbieta Guzewicz	Anna Niedźwiecka
Jerzy Pełka	Maciej Sawicki
Henryk G. Teisseyre	Andrzej Wawro

Editorial Committee:

Jerzy Kijowski	Maciej Kolwas
Jacek Kossut	Leszek Sirko
Andrzej Sobolewski	Henryk Szymczak

Editorial Council:

Jacek K. Furdyna	Tadeusz Luty
Józef Szudy	Jakub Zakrzewski
Ryszard Horodecki	Karol I. Wysokiński

Managing Editor:

Joanna Pietraszewicz

Executive Editors:

Katarzyna Dug	Marcin Ł. Staszewski
---------------	----------------------

Address of the Publisher:
Instytut Fizyki PAN
al. Lotników 32/46
02-668 Warszawa, Poland
e-mail: appol@ifpan.edu.pl

Printed in Poland:
Drukarnia HAJSTRA Sp. z o.o.
M. Langiewiczza 28
05-825 Grodzisk Mazowiecki
drukarnia-hajstra.com
e-mail: hajstra.biuro@gmail.com

POLISH ACADEMY OF SCIENCES
INSTITUTE OF PHYSICS

*Established in 1920 by
the Polish Physical Society*



ACTA PHYSICA POLONICA

- General Physics
- Atomic and Molecular Physics
- Condensed Matter
- Optics and Quantum Optics
- Quantum Information
- Biophysics
- Applied Physics

Special issue in honor of
scientific achievements of
Professor Iwo Białynicki-Birula
on his 90th birthday

Guest editor
Tomasz Sowiński



RECOGNIZED BY THE EUROPEAN
PHYSICAL SOCIETY

Volume 143 — Number 6, WARSAW, JUNE 2023

This issue is dedicated to
Professor Iwo Białynicki-Birula on
the occasion of his 90th birthday



Editor of the Special Issue
Tomasz Sowiński

WARSAW

POLISH ACADEMY OF SCIENCES
INSTITUTE OF PHYSICS

Preface

In 1952, an unknown student from the Mechanical Technical High School in Rzeszów, a provincial city, took first place (along with two other participants) in the First Physics Olympiad — a prestigious contest organized by the Polish Physical Society. Apart from the great splendor that befell this young student, the victory allowed him to avoid mandatory (due to the regulations of that time) employment in one of the local factories and opened the route for him to study at the University of Warsaw. There, he quickly came under the guidance of outstanding theoreticians led by Professor Leopold Infeld. This is how the successful scientific career of Professor Iwo Białynicki-Birula began, as well as that of all his students and collaborators, who for decades have been continuously drawing on his incredible intuition, brilliance, and kindness. Just as physics and the world have been changing over the past 70 years, the memories of his students who came under his mentorship are surely diverse. However, what definitely unites them all is the belief in the incredibly profound physical intuition of Our Professor, which he constantly expresses in one of his most beloved sayings: *Przyroda jest łaskawa* (eng. *Nature is kind*). Indeed, striking is his readiness to undertake risky and sometimes slightly controversial research directions. Who else, if not the Professor, would have dared to question the validity of the widely used Feynman's proof [1] or to construct a consistent formulation of a nonlinear correction to, by definition linear, quantum mechanics [2, 3]? Only the Professor and his wife, Professor Zofia Białynicka-Birula, could envision and later prove that photons can undergo splitting in an external magnetic field [4] or that the uncertainty principle, similar to that of massive particles, can also be formulated for quanta of light [5, 6]. Once, for purely bureaucratic reasons, I asked the Professor what I should put in the “scientific interests of the supervisor” section of a certain form. Without hesitation and with full conviction he answered: *Theoretical Physics*. Although quantum electrodynamics is his greatest passion, his horizons extend to all corners of contemporary physics, where he always finds interesting questions that are still awaiting answers.

On the occasion of Professor Iwo Białynicki-Birula's 90th birthday, I invite everyone to read this special issue of *Acta Physica Polonica A*, in which his students, collaborators, and friends publish scientific papers from their respective fields of expertise. The richness of the topics covered and references to the Professor's scientific activities once again demonstrate how versatile and respected a scientist he is.

Happy Birthday, Master!

Tomasz Sowiński

Guest Editor

- [1] I. Białynicki-Birula, *Phys. Rev.* **155**, 1414 (1967)
- [2] I. Białynicki-Birula, J. Mycielski, *Bull. Acad. Pol. Sci. Cl. III* **23**, 461 (1975)
- [3] I. Białynicki-Birula, J. Mycielski, *Ann. Phys.* **100**, 62 (1976)
- [4] Z. Białynicka-Birula, I. Białynicki-Birula, *Phys. Rev. D* **2**, 2341 (1970)
- [5] I. Białynicki-Birula, Z. Białynicka-Birula, *Phys. Rev. Lett.* **108**, 140401 (2012)
- [6] I. Białynicki-Birula, Z. Białynicka-Birula, *Phys. Rev. A* **86**, 022118 (2012)

Contributions:

1. J.H. Eberly, *Thinking BIG*
Doi: [10.12693/APhysPolA.143.S9](https://doi.org/10.12693/APhysPolA.143.S9)
2. S. Evans, J. Rafelski, *Improving Euler–Heisenberg–Schwinger Effective Action with Dressed Photons*
Doi: [10.12693/APhysPolA.143.S13](https://doi.org/10.12693/APhysPolA.143.S13)
3. A. Bechler, F. Cajiao Vélez, K. Krajewska, J.Z. Kamiński, *Vortex Structures and Momentum Sharing in Dynamic Sauter–Schwinger Process*
Doi: [10.12693/APhysPolA.143.S18](https://doi.org/10.12693/APhysPolA.143.S18)
4. M.G. Raymer, P. Polakos, *States, Modes, Fields, and Photons in Quantum Optics*
Doi: [10.12693/APhysPolA.143.S28](https://doi.org/10.12693/APhysPolA.143.S28)
5. P. Stammer, M. Lewenstein, *Quantum Optics as Applied Quantum Electrodynamics Is Back in Town*
Doi: [10.12693/APhysPolA.143.S42](https://doi.org/10.12693/APhysPolA.143.S42)
6. A. Friedrich, D. Moll, M. Freyberger, L. Plimak, W.P. Schleich, *The Wave Functional of the Vacuum in a Resonator*
Doi: [10.12693/APhysPolA.143.S52](https://doi.org/10.12693/APhysPolA.143.S52)
7. P.W. Milonni, *Free Energy of Coupled Oscillators: Lamb Shifts and van der Waals Interactions*
Doi: [10.12693/APhysPolA.143.S78](https://doi.org/10.12693/APhysPolA.143.S78)
8. W. Price, M. Formanek, J. Rafelski, *Born–Infeld Nonlinear Electromagnetism in Relativistic Heavy Ion Collisions*
Doi: [10.12693/APhysPolA.143.S87](https://doi.org/10.12693/APhysPolA.143.S87)
9. T. Linowski, Ł. Rudnicki, *Classicality of Bogoliubov Transformations and the Dynamical Casimir Effect Through the Reduced State of the Field*
Doi: [10.12693/APhysPolA.143.S95](https://doi.org/10.12693/APhysPolA.143.S95)
10. J. Przeszowski, *Smeared Field Description of Free Electromagnetic Field*
Doi: [10.12693/APhysPolA.143.S107](https://doi.org/10.12693/APhysPolA.143.S107)
11. F. Gleisberg, W.P. Schleich, *Factorization with a Logarithmic Energy Spectrum of a Central Potential*
Doi: [10.12693/APhysPolA.143.S112](https://doi.org/10.12693/APhysPolA.143.S112)
12. J. Kijowski, *Arrival Time in Quantum Mechanics (Demonstrated in Geometrical Order)*
Doi: [10.12693/APhysPolA.143.S123](https://doi.org/10.12693/APhysPolA.143.S123)
13. F. Gampel, M. Gajda, *On Repeated Measurements of a Quantum Particle in a Harmonic Potential*
Doi: [10.12693/APhysPolA.143.S131](https://doi.org/10.12693/APhysPolA.143.S131)
14. J. Mostowski, J. Pietraszewicz, *Electron Localization in Rydberg States*
Doi: [10.12693/APhysPolA.143.S140](https://doi.org/10.12693/APhysPolA.143.S140)

15. D. Wójcik, K. Życzkowski, *Fractality of Certain Quantum States*
Doi: [10.12693/APhysPolA.143.S146](https://doi.org/10.12693/APhysPolA.143.S146)
16. M. Czachor, *Contra Bellum: Bell's Theorem as a Confusion of Languages*
Doi: [10.12693/APhysPolA.143.S158](https://doi.org/10.12693/APhysPolA.143.S158)
17. M.B. Kruk, T. Vibel, J. Arlt, P. Kulik, K. Pawłowski, K. Rzażewski, *Fock State Sampling Method — Characteristic Temperature of Maximal Fluctuations for Interacting Bosons in Box Potentials*
Doi: [10.12693/APhysPolA.143.S171](https://doi.org/10.12693/APhysPolA.143.S171)
18. J. Zakrzewski, *Dynamical Quantum Phase Transitions from Quantum Optics Perspective*
Doi: [10.12693/APhysPolA.143.S179](https://doi.org/10.12693/APhysPolA.143.S179)
19. Ł. Turski, *Vlasov–Klimontovich Equation in Action*
Doi: [10.12693/APhysPolA.143.S183](https://doi.org/10.12693/APhysPolA.143.S183)
20. Van Cao Long, Dung Nguyen Trong, *The Influence of Factors on the Characteristic Quantities of Structure, Phase Transition, and Crystallization Progress of Bulk Au Materials Considered by Molecular Dynamics Simulation*
Doi: [10.12693/APhysPolA.143.S191](https://doi.org/10.12693/APhysPolA.143.S191)

DEDICATED TO PROFESSOR IWO BIALYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— ◇ —

Thinking BIG

J.H. EBERLY*

*Center for Coherence and Quantum Optics, and Department of Physics and Astronomy,
University of Rochester, Rochester, New York 14627, USA*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S9](https://doi.org/10.12693/APhysPolA.143.S9)*e-mail: eberly@pas.rochester.edu

Here we develop an informal speculation, which is focused on the existence of fundamental scales for measures of physical interest. In the present case, the scale of interest is the one for energy density, or pressure, for which a fundamental scale is not commonly known. Currently, however, the results being achieved in connection with high, and particularly extremely high, values of energy density are of interest. These speculative remarks are submitted as a contribution to the celebration of a birthday anniversary of Iwo Bialynicki-Birula and address the possibility that a search for the missing energy density scale can conceal and/or reveal something of fundamental interest. This may be particularly true, thinking of a search for such a scale, because a parallel example, a search of historical significance for a new physical scale, can be cited. The present author has enjoyed knowing Professor Iwo Bialynicki-Birula, and thanks him for the continuing pleasure of communication and consultation, as well as for a period of random occasions of table tennis warfare, and a close cooperation shared thirty years after that, which is recalled here.

topics: fundamental physical scales, unit of largeness, unit of smallness, serendipity

1. Introduction

The cooperation mentioned in the Abstract was enlivened by shared co-workers, notably including a senior theorist working behind the gates of Livermore National Laboratory in California, mostly closed to Polish scientists at the time, and an enthusiastic Ph.D. student who moved between Warsaw and Rochester. One memorable challenge was to fully understand the role that we guessed would be played by our discovery [1, 2] of Lagrange points for atomic electrons that are irradiated by microwave fields. These fields were originally considered circularly polarized and later linearly polarized, able to trap an atomic electron in very large quantum-atomic orbits with stability previously unsuspected, and conceptually a match for planetary orbits controlled by Lagrange-point force balance.

Later experimental work [3] not only confirmed the theoretical field-created orbits [4, 5], but was able to produce quantum wave packets for electrons

in long-time stably transported and non-spreading orbits, engaging principal quantum numbers larger than $n = 600$, thus greatly exceeding previous large-size records for stable non-spreading atomic orbits [6].

With such a dramatically new situation under experimental control, one is attracted to begin *Thinking Big* in a fundamental but natural way. Taking the very large quasi-Rydberg orbits recorded in Houston [3] as a starting example, one could think to ask, how large is LARGE? This question would have a ready answer if a fundamental scale of largeness were known. This introduces our central question, namely what is the *fundamental* scale for large size? Does such a fundamental limit exist to be consulted? What do large physical size scales imply? Do they indirectly suggest, or even define, an entire domain of ultimate “largeness” in physics?

We will consider the possibility of ultimate largeness. We believe that search for an answer could begin with commonly understood facts such as

the following. Within the most recent half century, cosmology has provided an unsuspected frontier for physics, and cosmological largeness may be a prominent characteristic. One knows that blackbody radiation is a phenomenon that is truly cosmic in scope, and relatively new. Since its discovery it has been carefully observed and analyzed. Much more recently than the discovery of cosmic blackbody radiation, both dark matter and dark energy have been accepted as newly-emerging phenomena with cosmological scope. They are still without universally agreed details of origin or ultimate consequence. Approaches to understanding them have been proposed and are being explored in a variety of ways. There is no fundamental scale yet associated with them. Here we speculate that a scale associated with their poorly defined large extent is attractive to consider. This could even be paired with direct attention to largeness in a domain where largeness itself can be recognized as under the current study. This is widely understood as the so-named high energy density and pressure (HEDP) domain of exceptionally high energy density and/or pressure. There is already wide international cooperation engaged in aggressive attack on examples of HEDP physics [7].

On elementary dimensional grounds, pressure is the same as energy density and is well-suited as an experimental measure. Large values of pressure have been obtained in several ways. One way has become sufficiently developed to be recognized with the award of the Physics Nobel Prize. Gerard Mourou and Donna Strickland won the Nobel Prize for physics in 2018 with the invention of CPA lasing (chirped-pulse amplification of lasing) in the Laboratory for Laser Energetics (LLE) at the University of Rochester. This now allows laboratory delivery of high values of tightly focused electromagnetic energy and has created a growing awareness of the HEDP regime. It is now accepted as a regime for experimental entry, by terrestrial laboratories [7], into studies of the highest energy densities and pressures on Earth.

Coincidence should not be overlooked. It is an obvious fact that *large* is the generic opposite to *small* and is understood as a conventional marker term, both scientifically and conversationally, for relative size among any array of physical objects. More intriguing is this question, namely is it possible that wide-ranging HEDP studies, perhaps reaching toward extreme cosmic-scaled largeness, can lead to disruptions of established physics at truly fundamental levels? Consequences of a scale for largeness, for great physical size, suggest attention to the consequences that followed attention to its opposite earlier counterpart, as follows. One knows that there was a centuries-long focus on smallness just in the casual sense, i.e., the term “atom” was widely familiar and used conversationally to mean an object so small that its smallness was incomparable. The modern epoch for the first meaningful use

of the word atom was the 1800’s, when the atom acquired a specific scientific meaning that accompanied the striking scientific advances occurring in chemistry. This happened by identifying as well as naming different types or kinds of atoms as actual objects. These were thought and taught as unbreakable and so were able to combine in fixed proportions to make different composite compounds (i.e., molecules). For example, salt is a common compound made of sodium and calcium atoms, but the atoms were still objects unquantifiably small. This use of “atom” for smallness did not yet mean that a reliable scale for smallness existed.

A backward look can remind us how a scale emerges. Measurement comes first. Fundamental scales recognize measurable quantities having limits. These scales serve to compare the values that are obtained as a result of measurement. The speed of light c is the fundamental value against which all other speeds can be judged by comparison, and the Compton wavelength provides the fundamental quantum value for particles by which its observed quantum momentum can be judged. No scale of fundamental origin is commonly accepted now as associated with energy density or pressure, especially high pressure. In regard to this, there might be some relevance in the way science did obtain a smallness measure. By inventing the “history” of quantitative smallness one can examine the question of whether we are presently entering a zone of experimentation that could unexpectedly, serendipitously quantify “fundamental largeness” for the first time. It is fair to say that HEDP work is now within an experimental regime that is “scalelessly high”. This recognizes a goal to be approached, and concedes that no naturally “fundamental” scale presently exists against which to compare high HEDP pressures.

Discoveries of natural scales usually occur accidentally, i.e., without deliberate intent. As mentioned, questions that in retrospect engaged the nature of “fundamental smallness” were being asked over many years (centuries) up to about 1890. In those times, not every alchemist or chemist or physicist was completely convinced that an “atom” was an actual thing that existed even to be detected. Thus “atom” served as a natural but not precisely defined limit for the obvious concept of “smallness”. There was no fundamental size that could be identified as the characteristic size or “typical” size for an atom, although different chemical elements were gradually and widely conceded to be made of different atoms and to have different small sizes. An interesting early example of an approach to the measurement of atomic (or molecular) size was a reported observation of the size of molecules at Clapham Pond in London in the 1750’s by the perennially curious and carefully observant Benjamin Franklin — the first American scientist. This is a topic, which Franklin is known to have speculated on, because he noted and reported the amazingly large area over which a small quantity of oil could spread freely

and very thinly on the surface of the pond water and still remain an intact film. Later, Lord Rayleigh improved Franklin's observation. Quantitative estimates based on knowledge of oil volume and pond area then led Rayleigh to a value consistent with an oil-molecule size of about 1 nanometer — 100 thousand times smaller than a human hair is wide. This set an amazing new record-low value for direct measurement of any small physical size. But this gave rise to no fundamental scale for comparison, but the story was continuing in England.

Soon afterward an astonishingly smaller value of particle size was not directly measured but was convincingly implied by the use of an entirely different kind of observation. This was reported in 1913 by the doctoral students Johannes Geiger and Ernest Marsden working for Ernest Rutherford in Manchester. Their laboratory experiments revealed the existence of a relatively and enormously very massive and tiny nucleus (so-named by Rutherford) in an empty space within gold atoms. It was about 5 additional orders of magnitude smaller than Franklin's and Rayleigh's work could provide, but still not accompanied by a fundamental scale unit.

What happened next? Actually not next, but what had happened a bit earlier? In the half-century before Rayleigh's simple pond-side experiment and 100 years after Franklin's observations, systematic measurements of another type with an entirely different motivation in mind, having nothing to do with particle sizes, were being carefully made by large numbers of scientists working worldwide. They were rapidly developing the new field of atomic spectroscopy, making innovations and then reporting results that we associate with such names as Fraunhofer, Bunsen, Kirchhoff, Ritz, Rydberg, Angstrom, Balmer and others. Values for different frequencies of light emitted by atoms were steadily accumulated, with mostly unsuccessful attempts to correlate them. A combination of squares of small integers extracted from the frequencies and published by Balmer in 1885 was just one of many inexplicable correlations of frequency data. It paid for all that experimental effort about 3 decades later in 1913. As every physicist knows, Balmer's numerical formula from 1885 provided the amazingly accurate numerical confirmation of Niels Bohr's new theory of hydrogen in 1913, which was based on Rutherford's "planetary" view of atomic electrons but included angular momentum quantization. From Bohr's theory, a fundamental measure of atomic smallness finally emerged as the "Bohr radius" $a_0 = \hbar^2/(e^2m)$. It serendipitously found both the explanation for the vast array of previously uncorrelated wavelengths collected by spectroscopists, as well as a new use for Max Planck's constant h . Thus, Bohr identified in fundamental terms what "atomic size" really means for smallness. This was a clear breakthrough. After that, atoms could be seen as just well-coordinated assemblies of electrons, all attracted by Coulomb's law to Rutherford's central

nucleus. However, the real mystery of atoms took 15 more years for it to be fully resolved. Entirely unexpectedly, the final resolution for atoms was not about atoms, but it used atoms to explore the first consequences of a completely new and very unexpected wave, which turned out to be in complete control of the electrons. That new wave is currently taught to physics students and called quantum mechanics.

Now we can reflect again on the HEDP physics. A natural question is whether its scale can be extracted and comprehended in a similar way. Maybe so, with the right orientation. It also has a key missing factor needed for describing poorly understood and emerging physical phenomena. The factor that is missing is the unrecognized fact that cosmologic distance is practically and fundamentally scale-less today. Currently, it has an out-of-scale largeness that is (inversely) similar to the out-of-scale smallness of atoms in 1900.

It is intriguing to push even further than the existing facts justify. One can imagine, without yet having the necessary ideas in their proper order, that LLE in Rochester and its cooperating partner laboratories in the world [7] are already taking data that will be serendipitously relevant to a Balmer-type first correlation of data, possibly to be made in 2030–2060 (recall the 3 decades from 1885 to 1913) and predict a new phenomenon of fundamental importance. Pushing harder, can one imagine that the central 2060 phenomenon underlying the newly scaled largeness will be "dark energy", playing a role similar to "atom" as the key phenomenon that led toward the a_0 smallness scale in 1915? Some similarities suggest a positive answer. Many physicists (weren't then — aren't now), finally and fully convinced that existence of atoms (then) and dark energy (now) is being treated appropriately and fully accepted. Can HEDP experiments turn out to be an opening wedge of innovative studies of dark matter, and use the results to reveal the existence of a fundamentally based "largeness" scale for dark matter? While wildly speculative, this would be an analog of previous worldwide cooperative atomic spectroscopy and the Rutherford invention of an atomic nucleus. The experiment in Rutherford's laboratory gave his visiting scientist Bohr something to go home to Denmark with and think about. Then Bohr theory [8] permanently unified our picture of atoms, showing where the scale of atomic smallness comes from. Who will supply the work, take the data, that sets a fundamental scale for cosmo-galactic largeness?

Acknowledgments

We acknowledge the financial support provided by National Science Foundation Grants PHY-1501589 and PHY-1539859 (INSPIRE).

References

- [1] I. Białynicki-Birula, M. Kalinski, J.H. Eberly, *Phys. Rev. Lett.* **73**, 1777 (1994).
- [2] I. Białynicki-Birula, M. Kalinski, J.H. Eberly, *Phys. Rev. Lett.* **75**, 973 (1995).
- [3] B. Wyker, S. Ye, F.B. Dunning, S. Yoshida, C.O. Reinhold, J. Burgdörfer, *Phys. Rev. Lett.* **108**, 043001 (2012).
- [4] M. Kalinski, J.H.Eberly, *Phys. Rev. A* **52**, 4285 (1995).
- [5] M. Kalinski, J.H. Eberly, *Optics Express* **1**, 216 (1997).
- [6] M. Kalinski, J.H. Eberly, I. Białynicki-Birula, *Phys. Rev. A* **52**, 2460 (1995).
- [7] There are numerous locations where experimental studies associated with high energy density and pressure (HEDP) are currently active and can be found internationally widely distributed. The following list of laboratories and countries with active programs is only suggestive, not complete: China, England, France, Germany, Italy, India, Japan, Korea, Poland, Russia, Spain, and USA. As an example, see [High-Energy-Density Physics \(HEDP\) Theory Group](#), University of Rochester.
- [8] N. Bohr, *Philos. Mag.* **26**, 1 (1913).

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

Improving Euler–Heisenberg–Schwinger Effective Action with Dressed Photons

S. EVANS^{a,b,*} AND J. RAFELSKI^a^a*Department of Physics, The University of Arizona, Tucson, AZ 85721, USA*^b*Helmholtz-Zentrum Dresden-Rossendorf, Bautzner Landstraße 400, 01328 Dresden, Germany*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S13](https://doi.org/10.12693/APhysPolA.143.S13)*e-mail: evanss@arizona.edu

We implement a longstanding proposal by Weisskopf to apply virtual polarization corrections to the in/out external fields in the study of the Euler–Heisenberg–Schwinger effective action. Our approach requires distinguishing the electromagnetic and polarization fields based on mathematical tools developed by Białynicki-Birula, originally for the Born–Infeld action. Our solution is expressed as a differential equation where the one-loop effective action serves as input. As a first result of our approach, we recover the higher order one-cut reducible loop diagrams discovered by Gies and Karbstein.

topics: Euler–Heisenberg–Schwinger (EHS), quantum electrodynamics (QED), non-perturbative vacuum structure, resummation methods

1. Introduction

Victor Weisskopf in 1936 [1, 2] suggested and attempted to improve the derivation of the Heisenberg–Euler effective action [3]; for further insights, see later work by Schwinger [4] and the review by Dunne [5]. Weisskopf considered that the polarization of the vacuum should be “fortwährend” (everlasting), and thus photons should contain the polarization effects already present in a self-consistent manner. In present-day language, the class of diagrams he envisaged requires the summation of one-cut reducible loop diagrams, i.e., photons dressed by one-loop Euler–Heisenberg–Schwinger (EHS) action. In this work, we present a path to the solution of this problem and give examples using constant homogeneous electromagnetic (EM) fields.

At first, the reducible loop diagram contributions to quantum electrodynamic (QED) effective action were assumed to vanish in the infrared, i.e., constant field limit. Ritus [6] claimed that as the photon momentum $k \rightarrow 0$, the pertinent two-loop diagrams vanish in view of the current $\propto k^2$. However, Gies and Karbstein [7] discovered that the pole of the virtual photon propagator ($\propto 1/k^2$) perfectly cancels the vanishing current in the quasi-constant EM field limit. This study of the nonvanishing two-loop

reducible diagram corrections to EHS effective action was extended via further perturbative summation to higher order loops [8–10], to scalar [11] and spinor propagators [12], and to a more general class of field configurations [13].

In this work, we demonstrate the connection between the Weisskopf conjecture and these reducible loop diagrams discovered in the present-day field-theoretical context. We implement a classical polarization approach for summing the virtual photon excitations in the infrared limit. By dressing the external field with polarization corrections at the start of the derivation of EHS action, we recover the two-loop result of Gies and Karbstein [7].

A key input into our nonperturbative solution is a class of Legendre transforms of nonlinear EM actions formulated by Białynicki-Birula [14], allowing to transform the nonlinear EHS action — a function of EM fields $\mathcal{L}_1(\mathcal{E}, \mathcal{B})$ — into an expression employing the superposable fields \mathcal{D}, \mathcal{H} . In this step, we can insert polarization corrections to dress the external fields. Lastly, we inverse the Legendre transform to return to an effective action formulation in terms of EM fields.

In Sect. 2, we develop an approach for implementing Weisskopf’s proposal to improve the EHS result, based on polarization corrections to the external fields. We implement the corrections in Sect. 3,

using the Legendre transformed EHS action, and apply our theoretical result to the case of pure electric fields. In Sect. 4, we recover the two-loop effective action of Gies and Karbstein. Extension to higher order loop contributions is straightforward, as we show with the three-loop action as an example. We believe that our approach can be applied to extend any one-loop effective action in the same everlasting manner, including the case of special interest, the strongly-interacting vacuum structure.

2. Implementing Weisskopf

2.1. Nonlinear EM action overview

We consider a general expression for EM effective action in the infrared external field limit (photon momentum $k \rightarrow 0$)

$$\mathcal{L}_{M+1}(\mathcal{E}, \mathcal{B}) = \frac{\mathcal{E}^2 - \mathcal{B}^2}{2} + \mathcal{L}_1(\mathcal{E}, \mathcal{B}), \quad (1)$$

where subscript M+1 denotes the Maxwell plus one-loop EHS contributions to the action. The EM fields \mathcal{E}, \mathcal{B} are generated by the 4-potential A^μ governing the Lorentz force as $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ and are related to the superposable fields \mathcal{D}, \mathcal{H} governing Maxwell equations with sources as

$$\begin{aligned} \mathcal{D}(\mathcal{E}, \mathcal{B}) &= \frac{\partial \mathcal{L}_{M+1}}{\partial \mathcal{E}} = \mathcal{E} + \frac{\partial \mathcal{L}_1}{\partial \mathcal{E}}, \\ \mathcal{H}(\mathcal{E}, \mathcal{B}) &= -\frac{\partial \mathcal{L}_{M+1}}{\partial \mathcal{B}} = \mathcal{B} - \frac{\partial \mathcal{L}_1}{\partial \mathcal{B}}. \end{aligned} \quad (2)$$

The nonlinear response of the vacuum thus distinguishes \mathcal{E}, \mathcal{B} from these superposable fields

$$\begin{aligned} \mathcal{E} &\equiv \mathcal{D}(\mathcal{E}, \mathcal{B}) - \mathcal{P}(\mathcal{E}, \mathcal{B}), \\ \mathcal{B} &\equiv \mathcal{H}(\mathcal{E}, \mathcal{B}) + \mathcal{M}(\mathcal{E}, \mathcal{B}), \end{aligned} \quad (3)$$

where the polarization fields \mathcal{P}, \mathcal{M} render the EM fields \mathcal{E}, \mathcal{B} non-superposable. This distinction will be necessary in order to implement Weisskopf's proposal to dress the externally applied EM fields.

All the relevant expressions for effective action in terms of EM and superposable fields are shown in Table I. The auxiliary quantity U is obtained from \mathcal{L} by Legendre transform, as we will describe below.

2.2. Reconciling EM fields with the everlasting vacuum

In Fig. 1, we show how Weisskopf's extension of QED-EHS action works in the context of in/out states: in panel (a), a photon scatters off a finite-sized polarizable material medium. The asymptotic in/out states, i.e., the EM fields before and after the interaction (black), are equivalent to the superposable fields ($\mathcal{E} = \mathcal{D}, \mathcal{B} = \mathcal{H}$). The screening by the medium (red) occurs inside the material target, with nonzero polarization fields \mathcal{P}, \mathcal{M} .

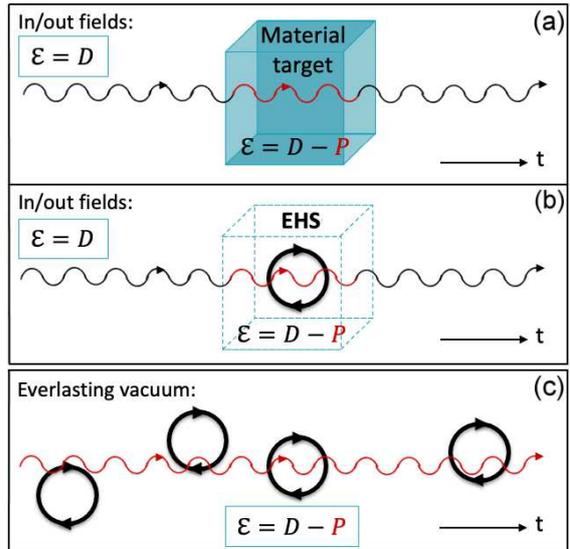


Fig. 1. EM fields interacting with (a) a finite-sized material medium, (b) prior treatment of the perturbative QED-EHS vacuum in the image of a scattering problem, (c) nonperturbative vacuum existing at all times.

TABLE I

EHS action (first two rows) and the higher order one-cut reducible loop action (last two rows); M+W refers to Maxwell+Weisskopf action, with Maxwell being the $(\mathcal{E}^2 - \mathcal{B}^2)/2$ and $(\mathcal{D}^2 - \mathcal{H}^2)/2$ contributions.

	Lagrange form	Auxiliary form
EHS	$\mathcal{L}_1(\mathcal{E}, \mathcal{B})$	$U_1(\mathcal{D}, \mathcal{H})$
Maxwell+EHS	$\mathcal{L}_{M+1}(\mathcal{E}, \mathcal{B})$	$U_{M+1}(\mathcal{D}, \mathcal{H})$
Dressed photons	$\mathcal{L}_W(\mathcal{E}, \mathcal{B})$	$U_W(\mathcal{D}, \mathcal{H})$
Maxwell+Dressed photons	$\mathcal{L}_{M+W}(\mathcal{E}, \mathcal{B})$	$U_{M+W}(\mathcal{D}, \mathcal{H})$

Following Weisskopf's insight that the external fields in EHS effective action see only one electron loop, we illustrate a perturbative EHS analog to the material target scattering (Fig. 1a). The EHS analog (Fig. 1b) comprises, in place of a material target, the quantum vacuum structure spanning a bounded spacetime domain sufficiently small that each photon in the external field sees only a single electron loop. Outside of this bounded region, no virtual electron excitations are considered, thus the asymptotic in/out external fields are approximated as $\mathcal{E} = \mathcal{D}, \mathcal{B} = \mathcal{H}$, i.e., without polarization effects.

This perturbative approach is amended in Fig. 1c. Since the vacuum structure exists at all times rather than in a bounded spacetime domain, we cannot distinguish the asymptotic in/out fields from the fields interacting with the virtual electron pairs. The polarization effects contained in fields \mathcal{P}, \mathcal{M} are

TABLE II

Legendre transforms and derivative expressions relating electromagnetic and superposable fields, after [14].

Legendre transform	Electric field	Magnetic field
$\mathcal{L}(\mathcal{E}, \mathcal{B}) = \mathcal{E} \cdot \mathcal{D} - \mathcal{B} \cdot \mathcal{H} - U$	$\mathcal{D} = \partial \mathcal{L} / \partial \mathcal{E}$	$\mathcal{H} = -\partial \mathcal{L} / \partial \mathcal{B}$
$U(\mathcal{D}, \mathcal{H}) = \mathcal{E} \cdot \mathcal{D} - \mathcal{B} \cdot \mathcal{H} - \mathcal{L}$	$\mathcal{E} = \partial U / \partial \mathcal{D}$	$\mathcal{B} = -\partial U / \partial \mathcal{H}$

always present, and thus $\mathcal{E} = \mathcal{D} - \mathcal{P}$ and $\mathcal{B} = \mathcal{H} + \mathcal{M}$ throughout Fig. 1c. These are the dressed fields to be implemented in the EHS action.

3. Derivation of effective action loop summation via everlasting vacuum properties

3.1. Legendre transform

We now show how to implement polarization field \mathcal{P}, \mathcal{M} corrections into the externally applied fields of EHS action. This cannot be done for the EHS action $\mathcal{L}_1(\mathcal{E}, \mathcal{B})$ directly due to the EM fields (see (3)) being non-superposable. Thus the first step is to transform the $\mathcal{L}_{M+1}(\mathcal{E}, \mathcal{B})$ into an auxiliary form written in terms of superposable fields $U_1(\mathcal{D}, \mathcal{H})$, based on the Legendre transforms seen in Table II.

Carrying out the Legendre transform of the EM action (1)

$$U_{M+1}(\mathcal{D}, \mathcal{H}) = \mathcal{E}(\mathcal{D}, \mathcal{H}) \cdot \mathcal{D} - \mathcal{B}(\mathcal{D}, \mathcal{H}) \cdot \mathcal{H} - \mathcal{L}_{M+1}(\mathcal{E}(\mathcal{D}, \mathcal{H}), \mathcal{B}(\mathcal{D}, \mathcal{H})), \quad (4)$$

where the EM fields

$$\begin{aligned} \mathcal{E}(\mathcal{D}, \mathcal{H}) &= \frac{\partial U_{M+1}}{\partial \mathcal{D}}, \\ \mathcal{B}(\mathcal{D}, \mathcal{H}) &= -\frac{\partial U_{M+1}}{\partial \mathcal{H}}. \end{aligned} \quad (5)$$

Separating the nonlinear contribution we define

$$U_{M+1}(\mathcal{D}, \mathcal{H}) \equiv \frac{\mathcal{D}^2 - \mathcal{H}^2}{2} + U_1(\mathcal{D}, \mathcal{H}), \quad (6)$$

distinguishing the contribution to the action, in terms of \mathcal{D}, \mathcal{H} , arising from the virtual electron interaction. Note that $U_1(\mathcal{D}, \mathcal{H})$ and $\mathcal{L}_1(\mathcal{E}, \mathcal{B})$ are not the same expressions, since the superposable fields take on a different functional dependence than non-superposable EM fields. Determining U_1 requires solving an implicit differential equation as defined in (4) and (5). An analytic solution is available for the special case of the Born–Infeld action [14–16].

3.2. Polarization corrections

Only in this auxiliary form of EHS effective action, using superposable \mathcal{D}, \mathcal{H} fields, can the asymptotic in/out fields be corrected to account for

everlasting polarization fields. Where \mathcal{D}, \mathcal{H} appear in the nonlinear part of EM action $U_1(\mathcal{D}, \mathcal{H})$ in (6), we take

$$\mathcal{D} \rightarrow \mathcal{D} - \mathcal{P}(\mathcal{D}, \mathcal{H}) = \mathcal{E}(\mathcal{D}, \mathcal{H}), \quad (7)$$

and similarly for the magnetic field

$$\mathcal{H} \rightarrow \mathcal{H} + \mathcal{M}(\mathcal{D}, \mathcal{H}) = \mathcal{B}(\mathcal{D}, \mathcal{H}), \quad (8)$$

thereby dressing the asymptotically defined EM field that any single electron loop is exposed to. The polarization fields \mathcal{P}, \mathcal{M} introduce the one-cut reducible loop sum $U_W(\mathcal{D}, \mathcal{H})$, defined as

$$U_W(\mathcal{D}, \mathcal{H}) \equiv U_1(\mathcal{D} - \mathcal{P}, \mathcal{H} + \mathcal{M}). \quad (9)$$

Including the Maxwell term and plugging in (6) and (3), we obtain

$$U_{M+W}(\mathcal{D}, \mathcal{H}) \equiv \frac{\mathcal{D}^2 - \mathcal{H}^2}{2} - \frac{\mathcal{E}^2(\mathcal{D}, \mathcal{H}) - \mathcal{B}^2(\mathcal{D}, \mathcal{H})}{2} + U_{M+1}(\mathcal{E}(\mathcal{D}, \mathcal{H}), \mathcal{B}(\mathcal{D}, \mathcal{H})), \quad (10)$$

where $U_{M+1}(\mathcal{E}, \mathcal{B})$ follows from (4), with the replacements $\mathcal{D} \rightarrow \mathcal{E}(\mathcal{D}, \mathcal{H})$ and $\mathcal{B} \rightarrow \mathcal{B}(\mathcal{D}, \mathcal{H})$.

3.3. Inverse Legendre transform

As a final step, we inverse Legendre transform (10) to return to the effective action formulation as a function of EM fields \mathcal{E}, \mathcal{B} . Using the transform from Table II,

$$\begin{aligned} \mathcal{L}_{M+W}(\mathcal{E}, \mathcal{B}) &\equiv \mathcal{E} \cdot \mathcal{D}(\mathcal{E}, \mathcal{B}) - \mathcal{B} \cdot \mathcal{H}(\mathcal{E}, \mathcal{B}) \\ &- U_{M+W}(\mathcal{D}(\mathcal{E}, \mathcal{B}), \mathcal{H}(\mathcal{E}, \mathcal{B})) = \\ &\mathcal{E} \cdot \mathcal{D}(\mathcal{E}, \mathcal{B}) - \mathcal{B} \cdot \mathcal{H}(\mathcal{E}, \mathcal{B}) + \frac{\mathcal{E}^2 - \mathcal{B}^2}{2} \\ &- \frac{\mathcal{D}^2(\mathcal{E}, \mathcal{B}) - \mathcal{H}^2(\mathcal{E}, \mathcal{B})}{2} - U_{M+1}(\mathcal{E}, \mathcal{B}), \end{aligned} \quad (11)$$

where now the derivative identities

$$\begin{aligned} \mathcal{D}(\mathcal{E}, \mathcal{B}) &= \frac{\partial \mathcal{L}_{M+W}(\mathcal{E}, \mathcal{B})}{\partial \mathcal{E}}, \\ \mathcal{H}(\mathcal{E}, \mathcal{B}) &= -\frac{\partial \mathcal{L}_{M+W}(\mathcal{E}, \mathcal{B})}{\partial \mathcal{B}}. \end{aligned} \quad (12)$$

Separating the Maxwell contribution from the nonlinear vacuum contribution, we define

$$\mathcal{L}_W(\mathcal{E}, \mathcal{B}) \equiv \mathcal{L}_{M+W}(\mathcal{E}, \mathcal{B}) - \frac{\mathcal{E}^2 - \mathcal{B}^2}{2}. \quad (13)$$

Combining (11)–(13), we now have at our disposal a differential equation requiring input EHS, which, when solved, creates the effective action for the summed reducible loop diagrams.

3.4. Summary and generalized form

To summarize, we build upon the one-loop effective action $\mathcal{L}_{M+1}(\mathcal{E}, \mathcal{B})$ in (1) by applying:

- *Legendre transform*

$$U_{M+1}(\mathcal{D}, \mathcal{H}) = \frac{\mathcal{D}^2 - \mathcal{H}^2}{2} + U_1(\mathcal{D}, \mathcal{H}), \quad (14)$$

- *Polarization corrections*

$$U_{M+W}(\mathcal{D}, \mathcal{H}) = \frac{\mathcal{D}^2 - \mathcal{H}^2}{2} + U_1(\mathcal{D} - \mathcal{P}, \mathcal{H} + \mathcal{M}), \quad (15)$$

- *Inverse Legendre transform*

$$\mathcal{L}_{M+W}(\mathcal{E}, \mathcal{B}) = \frac{\mathcal{E}^2 - \mathcal{B}^2}{2} + \mathcal{L}_W(\mathcal{E}, \mathcal{B}). \quad (16)$$

4. Perturbative series for $\alpha = 1/137$

As an illustrative example, we consider the pure electric field case to study the two-loop action of Gies and Karbstein [7]. Taking $\mathcal{B} \rightarrow 0$, (11) becomes then

$$\mathcal{L}_{M+W}(\mathcal{E}) = \mathcal{E} \cdot \mathcal{D}(\mathcal{E}) + \frac{\mathcal{E}^2}{2} - \frac{\mathcal{D}^2(\mathcal{E})}{2} - U_{M+1}(\mathcal{E}). \quad (17)$$

We evaluate (17) by applying a perturbative loop expansion.

We first write the EHS Lagrangian dependence in (17) explicitly using the Legendre transform (4)

$$\begin{aligned} \mathcal{L}_{M+W}(\mathcal{E}) &= \mathcal{E} \cdot \mathcal{D}(\mathcal{E}) - \frac{\mathcal{D}^2}{2} - \frac{\partial U_{M+1}(\mathcal{E})}{\partial \mathcal{E}} \cdot \mathcal{E} \\ &+ \frac{1}{2} \left(\frac{\partial U_{M+1}(\mathcal{E})}{\partial \mathcal{E}} \right)^2 + \mathcal{L}_1 \left(\frac{\partial U_{M+1}(\mathcal{E})}{\partial \mathcal{E}} \right) + \frac{\mathcal{E}^2}{2}. \end{aligned} \quad (18)$$

We take the case of small polarization corrections to the externally applied EM field

$$\frac{|\mathcal{D} - \mathcal{E}|}{|\mathcal{E}|} \ll 1. \quad (19)$$

Under condition (19), the leading one-loop EHS contribution dominates the higher order loop effects. The perturbative summation of reducible diagrams to ℓ -loop order can be written as

$$\lim_{\frac{|\mathcal{D} - \mathcal{E}|}{|\mathcal{E}|} \ll 1} \mathcal{L}_{M+W}(\mathcal{E}) \equiv \mathcal{L}_{M+1}(\mathcal{E}) + \sum_{\ell=2}^{\infty} \mathcal{L}_{\ell}(\mathcal{E}), \quad (20)$$

where the one-loop EHS contribution is included in $\mathcal{L}_{M+1}(\mathcal{E})$, followed by summation over the two-loop and higher orders.

To determine the form of loop corrections $\mathcal{L}_{\ell}(\mathcal{E})$ in (20), we take the small polarization limit of the auxiliary function U_{M+1} defined in (4) and differentiate with respect to \mathcal{E} to obtain

$$\lim_{\frac{|\mathcal{D} - \mathcal{E}|}{|\mathcal{E}|} \ll 1} \frac{\partial U_{M+1}(\mathcal{E})}{\partial \mathcal{E}} = \mathcal{E} - \frac{\partial \mathcal{L}_1(\mathcal{E})}{\partial \mathcal{E}}. \quad (21)$$

Similarly for the superposable field \mathcal{D} ,

$$\lim_{\frac{|\mathcal{D} - \mathcal{E}|}{|\mathcal{E}|} \ll 1} \mathcal{D}(\mathcal{E}) = \mathcal{E} + \frac{\partial \mathcal{L}_1(\mathcal{E})}{\partial \mathcal{E}}. \quad (22)$$

Plugging (21) and (22) into (18),

$$\lim_{\frac{|\mathcal{D} - \mathcal{E}|}{|\mathcal{E}|} \ll 1} \mathcal{L}_{M+W}(\mathcal{E}) = \frac{\mathcal{E}^2}{2} + \mathcal{L}_1 \left(\mathcal{E} - \frac{\partial \mathcal{L}_1(\mathcal{E})}{\partial \mathcal{E}} \right). \quad (23)$$

Note that (23) shows the iterative structure of the effective action describing the higher order loop summation. Expanding in powers of \mathcal{L}_1

$$\mathcal{L}_2(\mathcal{E}) = - \left(\frac{\partial \mathcal{L}_1(\mathcal{E})}{\partial \mathcal{E}} \right)^2, \quad (24)$$

the two-loop of Gies and Karbstein (see (32) of [7]). The original result in [7] contains both \mathcal{E} and \mathcal{B} contributions, expressed using derivatives with respect to the EM field tensor $\mathcal{L}_2 = \frac{1}{2}(\partial \mathcal{L}_1 / \partial F^{\mu\nu})^2 = (\partial \mathcal{L}_1 / \partial \mathcal{B})^2 - (\partial \mathcal{L}_1 / \partial \mathcal{E})^2$, which reduces to (24) in the pure \mathcal{E} limit.

To obtain the three-loop contribution, we iterate the two-loop (24) into (23) as a complement to \mathcal{L}_1 appearing in the polarization correction $\partial \mathcal{L}_1(\mathcal{E}) / \partial \mathcal{E}$. Expanding again in powers of \mathcal{L}_1 , this time to third order,

$$\mathcal{L}_3(\mathcal{E}) = \frac{5}{2} \frac{\partial^2 \mathcal{L}_1(\mathcal{E})}{\partial \mathcal{E}^2} \left(\frac{\partial \mathcal{L}_1(\mathcal{E})}{\partial \mathcal{E}} \right)^2. \quad (25)$$

This perturbative higher order loop summation procedure can be carried out ad infinitum as in [9], with the replacement $\mathcal{B} \rightarrow -i\mathcal{E}$ to recast Karbstein's original summation for \mathcal{B} fields in terms of \mathcal{E} fields.

5. Conclusions

We have implemented Weisskopf's proposal [1, 2] to dress the external EM fields in EHS effective action with polarization corrections. This shows that the one-cut reducible QED loop diagram summation of Gies and Karbstein [7–9] was indeed foretold in the work of Weisskopf. We developed a generalized approach to summing such diagrams, which can be applied to any nonlinear EM theory, with a one-loop effective action as input and in principle carried to higher order in coupling constant as we have demonstrated evaluating the next-to-next order correction.

It is important to note that we include only the one-cut reducible loop diagram contributions to effective action. A full summation includes: higher order cut reducible diagrams and internal photon line (irreducible) loops producing, e.g., anomalous magnetic moment and field-dependent mass. Irreducible contributions to the action in constant fields are well-known to two-loop order [6], and a subset of such diagrams comprising vertex corrections enclosing a single external line — to all orders [17].

Rather than the conventional in/out method for computing effective action, which treats the structured vacuum as bounded in spacetime akin to a finite-sized material target, our approach takes into account an everlasting vacuum structure spanning all spacetime. Finally, we remark that this work complements in the “opposite” direction the insight by Białynicki-Birula, Rudnicki, and Wienczek [18] that the finite time duration of external fields regularizes the essential singularity seen in (the imaginary part of) the one-loop EHS result in the limit of weak electrical fields.

The analytical properties of the one-loop action resurface in the higher loops as a striking interplay between real and imaginary (containing the essential singularity) parts of the effective action, and between reducible and irreducible diagram contributions. Strong field asymptotics need further exploration as they are highly nontrivial, depending on which EM field invariant dominates the external EM fields ($(\mathcal{E}^2 - \mathcal{B}^2)/2$ versus the pseudoscalar $\mathcal{E} \cdot \mathcal{B}$).

To conclude, we have improved the formulation of effective action in the presence of an everlasting vacuum structure. Our result connects Weiskopf’s conjectured extension of EHS effective action to Gies and Karbstein’s discovered higher order reducible loop diagrams.

Acknowledgments

This work is dedicated to Professor Iwo Białynicki-Birula on the occasion of his 90th birthday.

References

- [1] V. Weisskopf, *Über die Elektrodynamik des Vakuums auf Grund der Quantentheorie des Elektrons*, Mathematisk-Fysiske Meddelelser XIV, Danske Videnskabernes Selskab, No. 6, 1936.
- [2] A.I. Miller, in: *Early quantum electrodynamics: A Source book*, Cambridge Univ. Press, 1994.
- [3] W. Heisenberg, H. Euler, *Z. Phys.* **98**, 714 (1936).
- [4] J.S. Schwinger, *Phys. Rev.* **82**, 664 (1951).
- [5] G.V. Dunne, in: *From Fields to Strings*, Vol. 1, Ed. M. Shifman et al., World Scientific, Singapore 2005 p. 445.
- [6] V.I. Ritus, *Sov. Phys. JETP* **42**, 774 (1975).
- [7] H. Gies, F. Karbstein, *J. High Energy Phys.* **1703**, 108 (2017).
- [8] F. Karbstein, *J. High Energy Phys.* **1710**, 075 (2017).
- [9] F. Karbstein, *Phys. Rev. Lett.* **122**, 211602 (2019).
- [10] F. Karbstein, *J. High Energy Phys.* **01**, 057 (2022).
- [11] J.P. Edwards, C. Schubert, *Nucl. Phys. B* **923**, 339 (2017).
- [12] N. Ahmadinia, F. Bastianelli, O. Corradini, J. P. Edwards, C. Schubert, *Nucl. Phys. B* **924**, 377 (2017).
- [13] N. Ahmadinia, J.P. Edwards, A. Ilderton, *J. High Energy Phys.* **05**, 038 (2019).
- [14] I. Białynicki-Birula, in: *Quantum Theory Of Particles and Fields*, Eds. B. Jancewicz, J. Lukierski, World Scientific, Philadelphia 1983, p. 262.
- [15] M. Born, L. Infeld, *Proc. Roy. Soc. Lond. A* **144**, 425 (1934).
- [16] W. Price, M. Formanek, J. Rafelski, “Born–Infeld nonlinear electromagnetism in relativistic heavy ion collisions”, in preparation.
- [17] W. Dittrich, *J. Phys. A* **11**, 1191 (1978).
- [18] I. Białynicki-Birula, L. Rudnicki, A. Wienczek, [arXiv:1108.2615](https://arxiv.org/abs/1108.2615), 2013.

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

Vortex Structures and Momentum Sharing in Dynamic Sauter–Schwinger Process

A. BECHLER*, F. CAJIAO VÉLEZ, K. KRAJEWSKA AND J.Z. KAMIŃSKI

Institute of Theoretical Physics, Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S18](https://doi.org/10.12693/APhysPolA.143.S18)*e-mail: adam.bechler@fuw.edu.pl

Vortex pattern formation in the electron–positron pair creation from a vacuum by a time-dependent electric field of linear polarization is analyzed. It is demonstrated that in such a scenario the momentum distributions of the created particles exhibit vortex–antivortex pairs. Their sensitivity to the laser field parameters, such as field frequency and intensity, is also studied. Specifically, it is shown that with increasing field frequency across a threshold, additional vortex–antivortex pairs appear. Their location in the momentum space is consistent with the general threshold behavior of the probability distributions of the created electrons (positrons). Namely, while for small field frequencies the particles tend to be created along the field polarization direction, for large enough frequencies, they are predominantly generated in the perpendicular direction. Such a change in the longitudinal and transverse momentum sharing of the created particles occurs across a threshold.

topics: electron–positron pair creation, vortices, threshold effects, Dirac–Heisenberg–Wigner formalism

1. Introduction

The nonlinear response of the quantum vacuum to macroscopic electromagnetic fields, leading to the creation of electron–positron (e^-e^+) pairs, has been predicted by Sauter [1], Heisenberg and Euler [2], and Schwinger [3]. Since then, various authors have made significant contributions to our current understanding of this process, which we will refer to as the Sauter–Schwinger process. Specifically, Białynicki-Birula, Górnicki and Rafelski have established a new framework for treating the quantum vacuum in electromagnetic fields [4] (see also [5, 6] and the Ph.D. thesis of Ł. Rudnicki [7]). This is by means of what they called the Dirac–Heisenberg–Wigner (DHW) function, which describes the e^-e^+ densities in phase space. Later on, the method was largely explored for the case of spatially homogeneous electric fields (see, e.g., [8–16]). For instance, the quantum kinetic approach was recovered in that case [8] and various analytical results for exactly solvable fields were derived [8–10]. Recently, the spontaneous formation of time-crystal structures in the e^-e^+

pair creation was discovered by Białynicki-Birula and Białynicka-Birula in [16]. Other applications of the DHW formalism in the context of pair creation concern the case of parallel spatially homogeneous electric and magnetic fields [17], the standing electric wave [18–21], and inhomogeneous electric and magnetic fields in one spatial direction [22–24]. The latter limitation follows entirely from the performing capabilities of current computers, as the DHW method is very general and can be used in arbitrary dimensions. It is also important to emphasize that the DHW method is not limited to describing pair creation from a vacuum. For instance, it was argued that DHW is very useful for practical plasma applications such as studies of Langmuir waves in high-density plasma [25].

Another area of research to which Professor I. Białynicki-Birula contributed largely is related to quantum vortices. It follows from the hydrodynamical formulation of quantum mechanics that a probability fluid can inherently possess vortices [26]. They are defined as phase singularities of the wave function and their strength is measured in terms

of the topological charge [26–28]. As was discussed in [27, 28], vortices form isolated lines that either emerge from a single point forming a closed loop, or can be created as a pair of lines with opposite topological charges. These mechanisms of creation and subsequent annihilation of vortex–antivortex pairs were confirmed recently in a series of papers focused on vortex structures in strong-field ionization [29–33]. Specifically, it was demonstrated that vortex structures are very sensitive to the laser field parameters, so they can be easily steered by the field. While the aforementioned papers deal with quantum vortices in nonrelativistic quantum mechanics, their notion can also be extended to relativistic quantum theory, as proposed by Białynicki-Birula and Białynicka-Birula in [34]. See also, the construction of knotted vortex states, or hopfion-like states in relativistic quantum mechanics [35] by the same authors.

Note that the creation and ionization of an electron–positron pair are formally similar since they are both threshold-related phenomena that can be driven by external dynamically changing fields. For this reason, one might expect similar effects to be exhibited in both processes. Keeping this in mind, in the current paper we investigate whether vortex structures similar to [29–33] can be observed in the probability amplitude of e^-e^+ pair creation in the presence of a linearly polarized time-dependent electric field. Our emphasis will be on the threshold behavior of those patterns, which can be studied, for instance, by changing the frequency of the driving field. As we will show, this is in agreement with the longitudinal and transverse

momentum sharing of the created particles across the threshold, which has been studied in [36]. At this point, we would like to mention that other structures, known as spiral vortex patterns, were found in strong-field ionization [37, 38] and later in pair production [14, 15] for certain combinations of circularly polarized electric field pulses. However, as demonstrated in [31], in the case of ionization such spirals in the momentum distributions of photoelectrons do not necessarily carry a nonzero topological charge, which distinguishes them from vortices analyzed in [29–33]. The same is expected to hold for pair creation.

Our paper is organized as follows. Based on the original derivation presented in [4], we introduce the DHW formalism in Sect. 2. The bispinorial decomposition of the DHW-function for a spatially homogeneous electric field is presented in Sect. 3, and the final equations for a linear polarization are given in Sect. 4. Section 5 is devoted to the vortex patterns in the creation of e^-e^+ pairs and their sensitivity to external field parameters, especially when passing across a threshold. Another threshold-related effect is discussed in Sect. 6, where we demonstrate how particle momentum is redistributed across the threshold of pair creation. Our final remarks are given in Sect. 7.

2. The DHW-function for fermion field

The DHW-function for the fermion field is defined as [4]

$$W_{\alpha\beta}(\mathbf{x}, \mathbf{p}, t) = -\frac{1}{2} \int d^3s e^{-i\mathbf{p}\cdot\mathbf{s}} \langle 0 | \mathcal{U}(\mathbf{s}, \mathbf{x}, t) \left[\Psi_\alpha(\mathbf{x} + \mathbf{s}/2, t), \Psi_\beta^\dagger(\mathbf{x} - \mathbf{s}/2, t) \right] | 0 \rangle, \quad (1)$$

where the factor $\mathcal{U}(\mathbf{s}, \mathbf{x}, t)$ contains line integral of the vector potential in temporal gauge $A^0 = 0$,

$$\mathcal{U}(\mathbf{s}, \mathbf{x}, t) = \exp \left[-ie \int_{-1/2}^{1/2} d\xi \mathbf{s} \cdot \mathbf{A}(\mathbf{x} + \xi\mathbf{s}, t) \right], \quad (2)$$

and assures gauge invariance of the DHW-function, whereas Ψ_α, Ψ_β are the fermion field operators in the Heisenberg picture. We use here the version of the DHW-function with the vacuum expectation value [8]; in general, however, any pure or mixed state can be used [4]. The DHW-function is a 4×4 Hermitian matrix and as such can be decomposed in terms of 16 Hermitian matrices Γ_a with real coefficients depending generally on \mathbf{x}, \mathbf{p} and time t . Matrices Γ_a ($a = 0, 1, 2, \dots, 15$) can be constructed as Kronecker products of two sets of Pauli matrices (including the identity matrix), (I_2, ρ_j) and (I_2, σ_j) [4]. The correspondence is as follows

$$\begin{aligned} \Gamma_0 &= I_4, & \Gamma_j &= \rho_j \otimes I_2, \\ \Gamma_{j+3} &= I_2 \otimes \sigma_j, & \Gamma_{j+6} &= \rho_1 \otimes \sigma_j, \\ \Gamma_{j+9} &= \rho_2 \otimes \sigma_j, & \Gamma_{j+12} &= \rho_3 \otimes \sigma_j, \end{aligned} \quad (3)$$

where index $j = 1, 2, 3$. In terms of standard γ -matrices,

$$\begin{aligned} \Gamma_0 &= I_4, & \Gamma_1 &= \gamma_5, & \Gamma_2 &= -i\gamma^0\gamma_5, \\ \Gamma_3 &= \gamma^0, & \Gamma_{j+3} &= \Sigma^j, & \Gamma_{j+6} &= \alpha^j, \\ \Gamma_{j+9} &= -i\gamma^j, & \Gamma_{j+12} &= \gamma^0\Sigma^j, \end{aligned} \quad (4)$$

where $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$, and $\Sigma^j = \gamma^5\alpha^j$ are the 4×4 spin matrices. With the use of (4) expansion of the DHW-function can be written in the form [4]

$$\begin{aligned} W(\mathbf{x}, \mathbf{p}, t) &= \frac{1}{4} (f_0 + \gamma_5 f_1 - i\gamma^0 \gamma_5 f_2 + \gamma^0 f_3 \\ &+ \boldsymbol{\Sigma} \cdot \mathbf{g}_0 + \boldsymbol{\alpha} \cdot \mathbf{g}_1 - i\boldsymbol{\gamma} \cdot \mathbf{g}_2 + \gamma^0 \boldsymbol{\Sigma} \cdot \mathbf{g}_3). \end{aligned} \quad (5)$$

The dimensionless expansion coefficients are the same as f_0, f_1, f_2, f_3 and $\mathbf{g}_0, \mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3$ used in [4].

The equations fulfilled by the expansion coefficients can be found by calculating their time derivatives using the Dirac equation for the fermion field operators. In deriving these equations one usually adopts the Hartree-, or mean electromagnetic field-approximation, neglecting its quantum fluctuations [4, 8]. This is equivalent to the replacements $\langle 0 | \hat{F}^{\mu\nu}(\mathbf{x}, t) \mathcal{U}(\mathbf{s}, \mathbf{x}, t) [\Psi(\mathbf{x}_1, t), \Psi^\dagger(\mathbf{x}_2, t)] | 0 \rangle \rightarrow F^{\mu\nu}(\mathbf{x}, t) \langle 0 | \mathcal{U}(\mathbf{s}, \mathbf{x}, t) [\Psi(\mathbf{x}_1, t), \Psi^\dagger(\mathbf{x}_2, t)] | 0 \rangle$, i.e., the operator of the quantum electromagnetic field is replaced by classical C-number field. The application of the Dirac equation for the fermion field operators in the general case of the space- and time-dependent electromagnetic field results in a complicated system of 16 integro-differential equations for the expansion coefficients of the DHW-function. These equations significantly simplify in the case of a spatially homogeneous electric field, which is the subject of main interest in the present paper. The initial conditions are determined by the free vacuum value of the DHW-function. It follows then from (1) with zero electromagnetic field and free Dirac field operators that only the coefficients f_3 and \mathbf{g}_1 survive, and their vacuum values are

$$f_3^{\text{vac}} = -\frac{2mc^2}{E_{\mathbf{p}}}, \quad \mathbf{g}_1^{\text{vac}} = -\frac{2c\mathbf{p}}{E_{\mathbf{p}}}, \quad (6)$$

where $E_{\mathbf{p}} = \sqrt{c^2\mathbf{p}^2 + m^2c^4}$ is the free particle energy. In the case of a spatially homogeneous electric field, only the coefficients \mathbf{g}_0 and \mathbf{g}_2 couple to the vacuum values (6), so that it is sufficient to consider the 10 equations for f_3 , \mathbf{g}_0 , \mathbf{g}_1 , \mathbf{g}_2 . They have the form [9, 10]

$$\left(\partial_t + e\mathcal{E}(t) \cdot \nabla_{\mathbf{p}} \right) W(\mathbf{p}, t) = \frac{c}{\hbar} M(\mathbf{p}) W(\mathbf{p}, t), \quad (7)$$

where W denotes the 10-dimensional vector

$$W = [f_3, \mathbf{g}_0, \mathbf{g}_1, \mathbf{g}_2], \quad (8)$$

and the 10×10 matrix M has the following block structure

$$M(\mathbf{p}) = \begin{bmatrix} 0 & \mathbf{0}^T & \mathbf{0}^T & 2\mathbf{p}^T \\ \mathbf{0} & \mathbb{O}_3 & 2\mathbf{p} \times & \mathbb{O}_3 \\ \mathbf{0} & 2\mathbf{p} \times & \mathbb{O}_3 & -2mcI_3 \\ -2\mathbf{p} & \mathbb{O}_3 & 2mcI_3 & \mathbb{O}_3 \end{bmatrix}, \quad (9)$$

where $\mathbf{0}$ and \mathbf{p} are the 3-dimensional null and momentum column vectors, \mathbb{O}_3 — 3×3 null matrix, and I_3 is the 3-dimensional identity matrix. The notation $\mathbf{p} \times$ means that when acting on the 3-dimensional vector to the right, it gives its vector product with \mathbf{p} . Explicitly,

$$\mathbf{p} \times = \begin{bmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{bmatrix}. \quad (10)$$

In closing this section, we note that the physical interpretation of the DHW-functions can be found in [4]. In particular, the phase space energy density is given by [4, 11],

$$\varepsilon(t, \mathbf{r}, \mathbf{p}) = c\mathbf{p} \cdot \mathbf{g}_1(t, \mathbf{r}, \mathbf{p}) + mc^2 f_3(t, \mathbf{r}, \mathbf{p}). \quad (11)$$

The one particle distribution function, which will be used in Sect. 5 for numerical analysis of momentum distributions, is defined as [11]

$$f(t, \mathbf{r}, \mathbf{p}) = \frac{\varepsilon(t, \mathbf{r}, \mathbf{p}) - \varepsilon_{\text{vac}}}{2E_{\mathbf{p}}} = \frac{\varepsilon(t, \mathbf{r}, \mathbf{p})}{2E_{\mathbf{p}}} + 1, \quad (12)$$

where ε_{vac} was expressed by vacuum DHW-functions (6). It is also worth noting that the DHW formalism is very general, as it allows one to account for an arbitrary electromagnetic field. However, for a spatially homogeneous electric field, other approaches can be conveniently applied; one of which is developed next.

3. Bispinorial representation of the DHW-functions for spatially homogeneous electric field

We consider the Dirac equation in the spatially homogeneous electric field $\mathcal{E}(t) = -\partial_t \mathcal{A}(t)$, with the vector potential vanishing both for $t \rightarrow -\infty$ and $t \rightarrow \infty$. Due to the translational invariance of the problem, the spatial dependence of the wave function is of the plane wave type,

$$\Psi(t, \mathbf{x}) = \exp\left(\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}\right) \Phi_{\mathbf{p}r}(t), \quad (13)$$

where the time-dependent bispinor $\Phi_{\mathbf{p}r}(t)$ is labeled by the asymptotic momentum \mathbf{p} and the spin index r . It fulfills the equation

$$i\hbar \partial_t \Phi_{\mathbf{p}r}(t) = H_{\text{D}}(t) \Phi_{\mathbf{p}r}(t), \quad (14)$$

where the time-dependent Hamiltonian reads

$$H_{\text{D}}(t) = c\boldsymbol{\alpha} \cdot (\mathbf{p} - e\mathcal{A}(t)) + \gamma^0 mc^2. \quad (15)$$

To make contact with the DHW-functions, we construct 16 expressions bilinear in the bispinor $\Phi_{\mathbf{p}r}(t)$

$$S_a(\mathbf{p}, t) = \sum_r \Phi_{\mathbf{p}r}^\dagger(t) \Gamma_a \Phi_{\mathbf{p}r}(t). \quad (16)$$

Using the Dirac equation (14) and its Hermitian conjugate, one finds equations fulfilled by the functions S_a ,

$$\partial_t S_a = \frac{i}{\hbar} \sum_r \Phi_{\mathbf{p}r}^\dagger [H_{\text{D}}(t), \Gamma_a] \Phi_{\mathbf{p}r}. \quad (17)$$

The Dirac Hamiltonian $H_{\text{D}}(t)$ can be written in terms of the Γ -matrices as

$$H_{\text{D}}(t) = c\Gamma_{j+6}(p^j - e\mathcal{A}^j(t)) + mc^2\Gamma_3, \quad (18)$$

where the summation convention for the Cartesian index j is used. The Γ matrices fulfill commutation relations

$$[\Gamma_a, \Gamma_b] = i \sum_{c=0}^{15} f_{ab}^c \Gamma_c, \quad (19)$$

where f_{ab}^c are the real structure constants of the algebra of Γ matrices. Substituting (18) into (17) and using (19) gives

$$\partial_t S_a = -\frac{c}{\hbar} (p^j - e\mathcal{A}^j) \sum_{b=0}^{15} f_{j+6,a}^b S_b - \frac{mc^2}{\hbar} \sum_{b=0}^{15} f_{3a}^b S_b. \quad (20)$$

The nonvanishing structure constants are (indices i, j, k take the values 1, 2, 3)

$$\begin{aligned}
 f_{ij}^k &= 2\epsilon_{ijk}, \\
 f_{1,i+9}^{i+12} &= 2, \quad f_{1,i+12}^{i+9} = -2, \quad f_{2,i+6}^{i+12} = -2, \\
 f_{2,i+12}^{i+6} &= 2, \quad f_{3,i+6}^{i+9} = 2, \quad f_{3,i+9}^{i+6} = -2, \\
 f_{i+3,j+3}^{k+3} &= f_{i+6,j+6}^{k+3} = f_{i+9,j+9}^{k+3} = f_{i+12,j+12}^{k+3} = 2\epsilon_{ijk}, \\
 f_{i+3,j+6}^{k+6} &= f_{i+3,j+9}^{k+9} = f_{i+3,j+12}^{k+12} = 2\epsilon_{ijk}, \\
 f_{i+9,j+12}^1 &= 2\delta_{ij}, \quad f_{i+6,j+12}^2 = -2\delta_{ij}, \quad f_{i+6,j+9}^3 = 2\delta_{ij},
 \end{aligned} \tag{21}$$

plus the structure constants obtained from the antisymmetry relation $f_{c,b}^a = -f_{b,c}^a$. It is now straightforward, though a little tedious, to derive 16 equations fulfilled by the functions S_a , i.e.,

$$\begin{aligned}
 \partial_t S_0 &= 0, \\
 \partial_t S_1 &= -2\frac{mc^2}{\hbar} S_2, \\
 \partial_t S_2 &= -2\frac{c}{\hbar}(p^j - e\mathcal{A}^j)S_{j+12} + 2\frac{mc^2}{\hbar} S_1, \\
 \partial_t S_3 &= 2\frac{c}{\hbar}(p^j - e\mathcal{A}^j)S_{j+9}, \\
 \partial_t S_{k+3} &= 2\frac{c}{\hbar}\epsilon_{kjl}(p^j - e\mathcal{A}^j)S_{l+6}, \\
 \partial_t S_{k+6} &= 2\frac{c}{\hbar}\epsilon_{kjl}(p^j - e\mathcal{A}^j)S_{l+3} - 2\frac{mc^2}{\hbar} S_{k+9}, \\
 \partial_t S_{k+9} &= -2\frac{c}{\hbar}(p^k - e\mathcal{A}^k)S_3 + 2\frac{mc^2}{\hbar} S_{k+6}, \\
 \partial_t S_{k+12} &= 2\frac{c}{\hbar}(p^k - e\mathcal{A}^k)S_2.
 \end{aligned} \tag{22}$$

Note that the equations containing the 6 functions $S_0, S_1, S_2, S_{13}, S_{14}, S_{15}$ do not couple to remaining ten equations for $S_3, S_4, S_5, S_6, S_7, S_8, S_9, S_{10}, S_{11}, S_{12}$. Denoting

$$\begin{aligned}
 S_3 &= \mathbf{h}_3, & (S_4, S_5, S_6) &= \mathbf{h}_0, \\
 (S_7, S_8, S_9) &= \mathbf{h}_1, & (S_{10}, S_{11}, S_{12}) &= \mathbf{h}_2,
 \end{aligned} \tag{23}$$

we see that (22) for ten-dimensional vector $V = [h_3, \mathbf{h}_0, \mathbf{h}_1, \mathbf{h}_2]$ can be written in the matrix form as

$$\partial_t V = \frac{c}{\hbar} M(\mathbf{p}(t)) V, \tag{24}$$

where

$$\mathbf{p}(t) = \mathbf{p} - e\mathcal{A}(t), \tag{25}$$

and the matrix M is given by (9). The same system of ordinary differential equations follows from (7) after applying the method of characteristics to first-order partial differential equations [4, 8, 10, 11]. Therefore the two vectors W and V obey the same system of ordinary differential equations. In order to identify fully V and W one needs to show that they fulfill also the same initial conditions, which for W are given by (6) and $\mathbf{g}_0^{\text{vac}} = 0 = \mathbf{g}_2^{\text{vac}}$.

The Dirac wave function pertaining to the pair creation process should fulfill the Feynman boundary conditions: (i) for $t \rightarrow -\infty$ it contains only solutions of the free Dirac equation with negative energy, (ii) for $t \rightarrow \infty$ it is a combination of positive and negative energy parts with a negative energy contribution equal to the wave function of the created positron. An extensive discussion of the boundary conditions fulfilled by solutions of the Dirac equation in a classical electromagnetic field can be found in [39]. It can also be shown that the Feynman boundary conditions are ‘‘forced’’ by LSZ-reduction formulae for the S -matrix element of pair creation. For $t \rightarrow -\infty$, we have therefore

$$\Phi_{\mathbf{p}s}(t) = \exp\left(\frac{i}{\hbar} E_{\mathbf{p}} t\right) w_{-\mathbf{p}s}^{(-)}. \tag{26}$$

Substituting (26) into (16) with the bispinors normalized to unity, one can show that the coefficients (23) fulfill the following initial conditions for $t \rightarrow -\infty$

$$\mathbf{h}_0^0 = 0 = \mathbf{h}_2^0, \quad h_3^0 = -\frac{2mc^2}{E_{\mathbf{p}}}, \quad \mathbf{h}_1^0 = -\frac{2c\mathbf{p}}{E_{\mathbf{p}}}, \tag{27}$$

corresponding exactly to vacuum initial conditions for the vector W .

The bispinorial approach to the dynamic Sauter-Schwinger pair production by a spatially homogeneous electric field has been developed in this section. Importantly, the approach has been proven to be equivalent to the DHW formalism described in Sect. 2. And, like the DHW method, it has an advantage over other approaches. Specifically, it allows the treatment of an arbitrarily polarized time-dependent electric field. Having said that, we turn to the case of linear polarization, for which other well-established theories exist and can be tested against (see, for instance, [40, 41] and references therein).

4. Linearly polarized field and analogy with two level atom

In general, the vector W (or, equivalently V) can be expressed as a combination of ten orthonormal basis vectors \mathbb{E}_a

$$W = -2 \sum_{a=1}^{10} u_a \mathbb{E}_a. \tag{28}$$

With the choice of $-\frac{1}{2}W^{\text{vac}}$ as one of the basis elements, one can show that for a linearly polarized field $\mathcal{A}(t) = \mathcal{A}(t)\mathbf{n}$, three vectors

$$\mathbb{E}_1 = \frac{c}{E_{\mathbf{p}}\epsilon_{\perp}} \begin{bmatrix} -mc^2(\mathbf{n} \cdot \mathbf{p}) \\ \mathbf{0} \\ \frac{E_{\mathbf{p}}^2 \mathbf{n}}{c} - c(\mathbf{n} \cdot \mathbf{p})\mathbf{p} \\ \mathbf{0} \end{bmatrix}, \quad \mathbb{E}_2 = \frac{c}{\epsilon_{\perp}} \begin{bmatrix} 0 \\ \mathbf{p} \times \mathbf{n} \\ \mathbf{0} \\ mc\mathbf{n} \end{bmatrix}, \quad \mathbb{E}_3 = \frac{1}{E_{\mathbf{p}}} \begin{bmatrix} mc^2 \\ \mathbf{0} \\ c\mathbf{p} \\ \mathbf{0} \end{bmatrix}, \tag{29}$$

form a set closed under the action of $\mathbf{n} \cdot \nabla_{\mathbf{p}}$ and M in (7), where $\epsilon_{\perp} = \sqrt{c^2 \mathbf{p}_{\perp}^2 + m^2 c^4}$. Note that $\mathbb{E}_3 = -\frac{1}{2} W^{\text{vac}}$. Choosing \mathbf{n} in the z -direction ($\mathbf{n} = \mathbf{e}_3$) we have

$$\begin{aligned} \frac{\partial}{\partial p_3} \mathbb{E}_1 &= -\frac{c \epsilon_{\perp}}{E_{\mathbf{p}}^2} \mathbb{E}_3, & \frac{\partial}{\partial p_3} \mathbb{E}_2 &= 0, & \frac{\partial}{\partial p_3} \mathbb{E}_3 &= \frac{c \epsilon_{\perp}}{E_{\mathbf{p}}^2} \mathbb{E}_1, \\ M \mathbb{E}_1 &= \frac{2E_{\mathbf{p}}}{c} \mathbb{E}_3, & M \mathbb{E}_2 &= -\frac{2E_{\mathbf{p}}}{c} \mathbb{E}_1, & M \mathbb{E}_3 &= 0. \end{aligned} \quad (30)$$

Solution of (7) can be expressed as

$$W(\mathbf{p}, t) = -2 \sum_{a=1}^3 u_a(\mathbf{p}, t) \mathbb{E}_a(\mathbf{p}(t)). \quad (31)$$

Substituting (31) to (24) and denoting $[u_1, u_2, u_3] = \mathbf{u}$, we obtain the precession-type equation for \mathbf{u}

$$\partial_t \mathbf{u} = \mathbf{a} \times \mathbf{u}, \quad (32)$$

with the vector \mathbf{a} given by

$$\mathbf{a} = [0, -2\Omega_{\mathbf{p}}(t), 2\omega_{\mathbf{p}}(t)], \quad (33)$$

where

$$\begin{aligned} \omega_{\mathbf{p}}(t) &= \frac{E_{\mathbf{p}}(t)}{\hbar} = \frac{1}{\hbar} \sqrt{c^2 \mathbf{p}_{\perp}^2 + c^2 [p_3 - eA(t)]^2 + m^2 c^4}, \\ \Omega_{\mathbf{p}}(t) &= \frac{c e \epsilon_{\perp} \mathcal{E}(t)}{2E_{\mathbf{p}}^2}, \end{aligned} \quad (34)$$

and where the temporal dependence of the electric field is given by $\mathcal{E}(t) = -\dot{A}(t)$. Note that the initial condition for \mathbf{u} has the form $\mathbf{u}^{\text{vac}} = [0, 0, 1]$.

Three equations resulting from (32) can be reduced to a system of two equations by expressing \mathbf{u} in the form of a spinorial decomposition, analogous to that used in [9, 10], i.e.,

$$\mathbf{u} = \chi^{\dagger} \boldsymbol{\sigma} \chi, \quad (35)$$

where χ is the two-component spinor and $\boldsymbol{\sigma}$ are the Pauli matrices. Substitution of (35) to (32) leads to the equation for χ which has the same structure as the Schrödinger equation describing the time evolution of a two-level atom. This equation has been derived in the context of pair-creation by a different method earlier (see, e.g., [9, 40] and references therein),

$$i \partial_t \begin{bmatrix} c_{\mathbf{p}}^{(1)}(t) \\ c_{\mathbf{p}}^{(2)}(t) \end{bmatrix} = \begin{bmatrix} \omega_{\mathbf{p}}(t) & i\Omega_{\mathbf{p}}(t) \\ -i\Omega_{\mathbf{p}}(t) & -\omega_{\mathbf{p}}(t) \end{bmatrix} \begin{bmatrix} c_{\mathbf{p}}^{(1)}(t) \\ c_{\mathbf{p}}^{(2)}(t) \end{bmatrix}, \quad (36)$$

where $c_{\mathbf{p}}^{(1)}(t)$ and $c_{\mathbf{p}}^{(2)}(t)$ are, respectively, upper and lower components of χ . Initial conditions read

$$c_{\mathbf{p}}^{(1)}|_{t \rightarrow -\infty} = 1, \quad c_{\mathbf{p}}^{(2)}|_{t \rightarrow -\infty} = 0. \quad (37)$$

The third component of \mathbf{u} is equal to $|c_{\mathbf{p}}^{(1)}|^2 - |c_{\mathbf{p}}^{(2)}|^2$ and for a two-level atom corresponds to ‘‘population inversion’’ (with opposite sign). Before the action of the electric field $u_3 = 1$, which corresponds to the vacuum state with no pairs. During the action of the electric field, e^+e^- pairs are created so that $|c_{\mathbf{p}}^{(1)}|^2 < 1$ and $|c_{\mathbf{p}}^{(2)}|^2 > 0$ with $|c_{\mathbf{p}}^{(1)}(t)|^2 + |c_{\mathbf{p}}^{(2)}(t)|^2 = 1$. Hence, $|c_{\mathbf{p}}^{(2)}|^2$ for $t \rightarrow \infty$

can be interpreted as the momentum distribution of the created fermionic pairs, $f(\mathbf{p})$ (see (12)). Explicitly,

$$f(\mathbf{p}) = 1 - u_3 = 2|c_{\mathbf{p}}^{(2)}|^2. \quad (38)$$

Let us note in closing this section that the system of equations similar to (38) can be derived for bosons by applying other, than those based on the Wigner formalism, methods of QED (see, e.g., [41] and references therein). However, in this case, the time-evolution is pseudounitary.

From now on, we use units where $\hbar = 1$. Moreover, m and e will refer to the electron rest mass and charge, respectively.

5. Threshold effects and vortices

In our further investigations we choose an electric field $\mathcal{E}(t)$ such that

$$\mathcal{E}(t) = \begin{cases} \mathcal{E}_0 \sin^4\left(\frac{1}{2N}\phi\right) \cos(\phi), & \phi \in [0, 2\pi N], \\ 0, & \phi \notin [0, 2\pi N], \end{cases} \quad (39)$$

where $\phi = \omega t$ and $N = 3$. The integer N determines the number of cycles within the electric field pulse, and for $N \geq 3$ the condition

$$\int_{-\infty}^{\infty} dt \mathcal{E}(t) = 0 \quad (40)$$

is satisfied. Due to this property, the vector potential function,

$$\mathcal{A}(t) = - \int_{-\infty}^t d\tau \mathcal{E}(\tau), \quad (41)$$

can be chosen such that it vanishes both in the remote past and in the far future

$$\lim_{t \rightarrow \pm\infty} \mathcal{A}(t) = 0. \quad (42)$$

The shapes of both functions for $\mathcal{E}_0 = 0.1\mathcal{E}_S$ and $\omega = mc^2$ are presented in Fig. 1, where $\mathcal{E}_S = m^2 c^3 / |e|$ is the Sauter–Schwinger electric field strength [1, 3, 42]. For the electron momentum vector, we will separate its parallel and perpendicular components as measured with respect to the direction of the electric field oscillations \mathbf{e}_3 such that $\mathbf{p} = p_{\perp} \mathbf{e}_{\perp} + p_{\parallel} \mathbf{e}_3$, where \mathbf{e}_{\perp} is the unit vector perpendicular to \mathbf{e}_3 .

As was mentioned in Sect. 1, the process of creating electron–positron pairs in QED has many analogies with the ionization of atoms, in which the role of the time-dependent electric field is played by a strong laser pulse in the dipole approximation. In this case, the concept of photons is commonly used as quanta of energy absorbed or emitted by the system. One can then talk about multiphoton ionization and the energy threshold for that process. Moreover, such a threshold is dynamically increased as the electric field becomes stronger, which leads

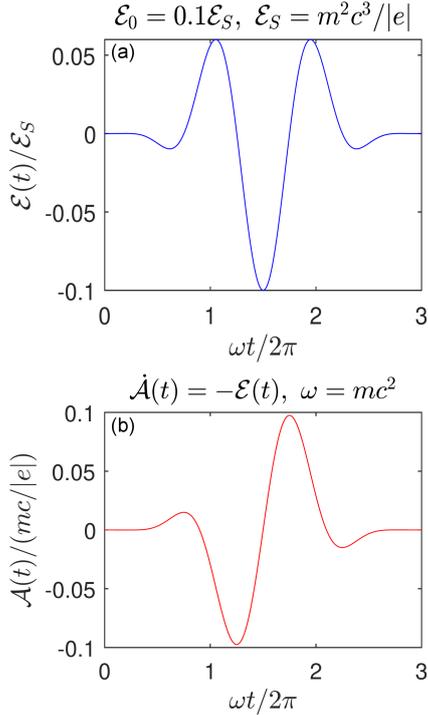


Fig. 1. (a) Time-dependent electric field strength $\mathcal{E}(t)$ for $\mathcal{E}_0 = 0.1\mathcal{E}_S$ and $N = 3$, as defined by (39), and (b) the corresponding vector potential function $\mathcal{A}(t)$. Contrary to the electric field, the amplitude of the vector potential depends on the frequency ω .

to the so-called threshold effects and channel closing in ionization [43]. It turns out that in the case of the dynamic Sauter–Schwinger process this heuristic picture can also be applied in order to describe qualitative changes in the momentum distributions of the created particles (for instance, as in the coherent energy combs studied in [40, 44]). This can be done even for the very short pulses considered here.

Another interesting effect, which appears as a result of the interaction of the time-dependent electric field with the QED vacuum, is the creation or annihilation of vortex lines in the electron momentum distributions. The properties of vortex lines and their entanglement were thoroughly analyzed in [27, 28]. In both photoionization and photodetachment, the creation and annihilation of vortex lines were studied for linearly [30] and circularly [29, 31–33] polarized fields. It was shown how the time-reversal symmetry of the laser pulse leads to the annihilation of vortex–antivortex pairs and the creation of spirals in the momentum distributions [33]. Note that such spirals have been predicted theoretically in [37] and confirmed experimentally in [38]. Moreover, the application of the DHW-function formalism allowed one to show that similar spiral structures also appear in the pair creation by a train of two circularly polarized electric field pulses of an opposite helicity [14, 15].

Therefore, the question arises: Can vortices be expected in the momentum distributions of the created pairs for linearly polarized electric field pulses?

To address this question, in Fig. 2 we present the momentum distributions of electrons created by a linearly polarized electric field pulse of different frequencies, which were selected close to the “two-photon” threshold of pair creation. For $\omega = 0.99 mc^2$ (i.e., just before opening the “two-photon” channel), we observe two singular points for which the amplitude $c_{\mathbf{p}}^{(2)}$ vanishes, and the phase $\arg[c_{\mathbf{p}}^{(2)}]$ cannot be uniquely defined. Because of the axial symmetry of the problem, it can be concluded that these two points belong to the same vortex line. In the current case, the latter is represented by a circle in three-dimensional momentum space. In fact, one can even define the orientation of this closed line by exploiting analogies with a circuit along which an electric current flows and generates, according to Amperé’s law, the vortex-type magnetic field. To this end, let us define the “magnetic field” $\mathbf{B}(\mathbf{p})$ such that

$$\mathbf{B}(\mathbf{p}) = \nabla_{\mathbf{p}} \left(\arg [c_{\mathbf{p}}^{(2)}] \right). \quad (43)$$

Its circulation around the singular point is $\pm 2\pi$, hence the “electric current” becomes $I = \pm 2\pi$ if we put the magnetic permeability $\mu_0 = 1$. In particular, for $\omega = 0.99 mc^2$ (Fig. 2a), we have ‘−’ for $(p_{\parallel} = 0, p_{\perp} > 0)$ and the current flows behind the plane, whereas for $(p_{\parallel} = 0, p_{\perp} < 0)$ we have ‘+’ and the current flows towards the reader. Thus the orientation of the vortex line can be uniquely attributed to the direction of the “electric current”. As the frequency ω increases, we observe the appearance of a new vortex line. The case of $\omega = mc^2$ (which is the threshold frequency for the two-photon pair creation) corresponds to a transition in which, for $\mathbf{p} = \mathbf{0}$, the radius of the new vortex circle is close to zero (Fig. 2b and e). After exceeding this value (the case of $\omega = 1.01 mc^2$ in Fig. 2c and f), the second circular vortex line appears, with an orientation opposite to the previous one (panels (d–f)). While increasing the frequency ω , the radii of both circular vortex lines also grow. This, in turn, results in the merging of the two well-defined lobes of high probability into a single structure, the maximum of which is found at zero momentum. This situation is discussed in Sect. 6.

In Fig. 3 we demonstrate the same phenomenon, but for a larger amplitude of the electric field. The only significant difference is that now the threshold frequency for the two-photon pair creation is shifted upwards and its value is between $1.1 mc^2$ (one vortex line) and $1.2 mc^2$ (two vortex lines). A plausible interpretation of this fact can be based on analogies with multiphoton ionization, in which, for a larger intensity of the electromagnetic field, the so-called ponderomotive shift of the threshold energy is observed [43]. Similar effects, but in the context of photodetachment by circularly polarized laser pulses, were discussed in [32].

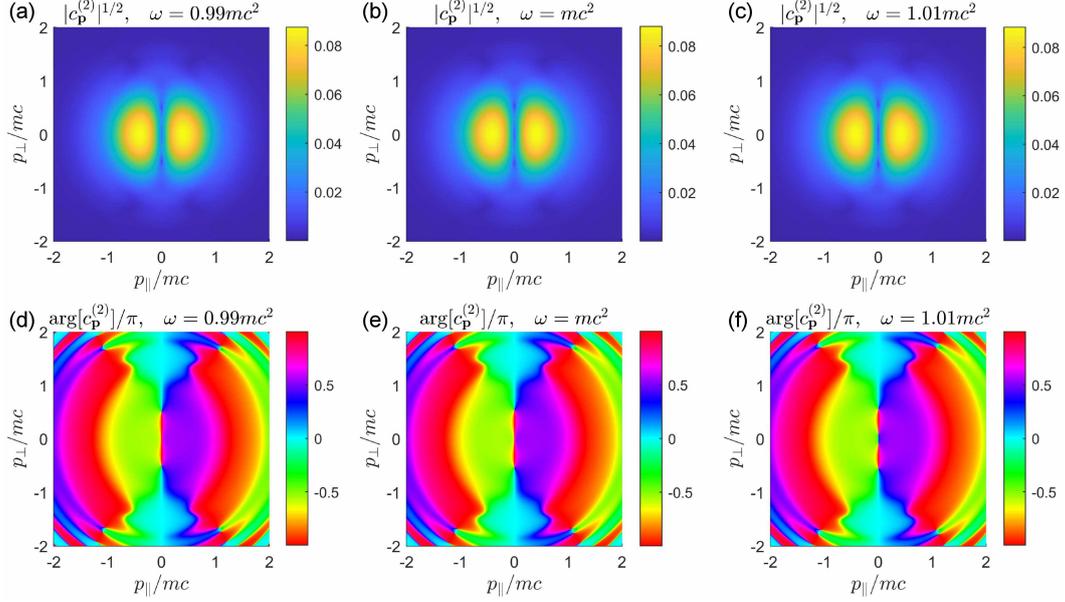


Fig. 2. Momentum distributions of electrons created from the QED vacuum by the electric field illustrated in Fig. 1. In panels (a–c), the distributions $|c_{\mathbf{p}}^{(2)}|^{1/2}$ (the power 1/2 is chosen for visual purposes) are presented for three chosen frequencies ω (equivalent to photon energies). In panels (d–f), the corresponding phases of $c_{\mathbf{p}}^{(2)}$ are demonstrated.

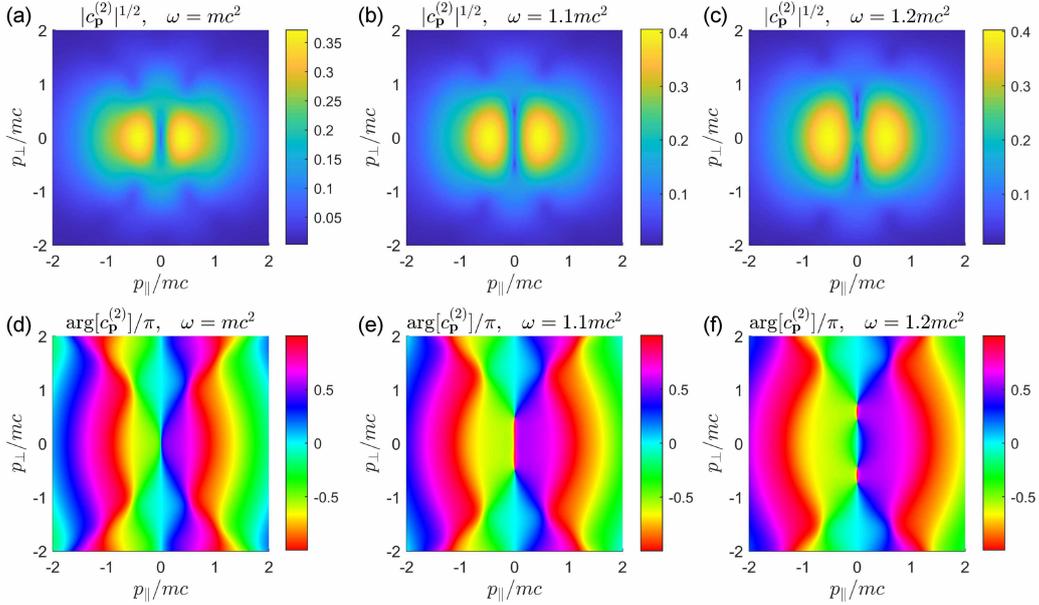


Fig. 3. The same description as in Fig. 2, but for larger electric field amplitude $\mathcal{E}_0 = 0.5\mathcal{E}_S$ and larger frequencies.

6. Longitudinal and transverse momentum sharing

The Schwinger formula for the probability rate of pair production per unit volume in the case of a constant (or slowly-changing-in-time) electric field can be derived using the tunneling formalism [45]. According to this formula, an increment of the perpendicular momentum of the particles,

$|p_{\perp}|$, is accompanied by a rapidly vanishing creation rate. However, for rapidly changing fields, tunneling theory is no longer applicable. This is supported by the analysis presented above because the momentum distributions for pair production are not elongated in the direction of the electric field. As it has been shown in [36], for sufficiently high frequencies ω , particles prefer to be created in the direction perpendicular to the electric field. This counterintuitive

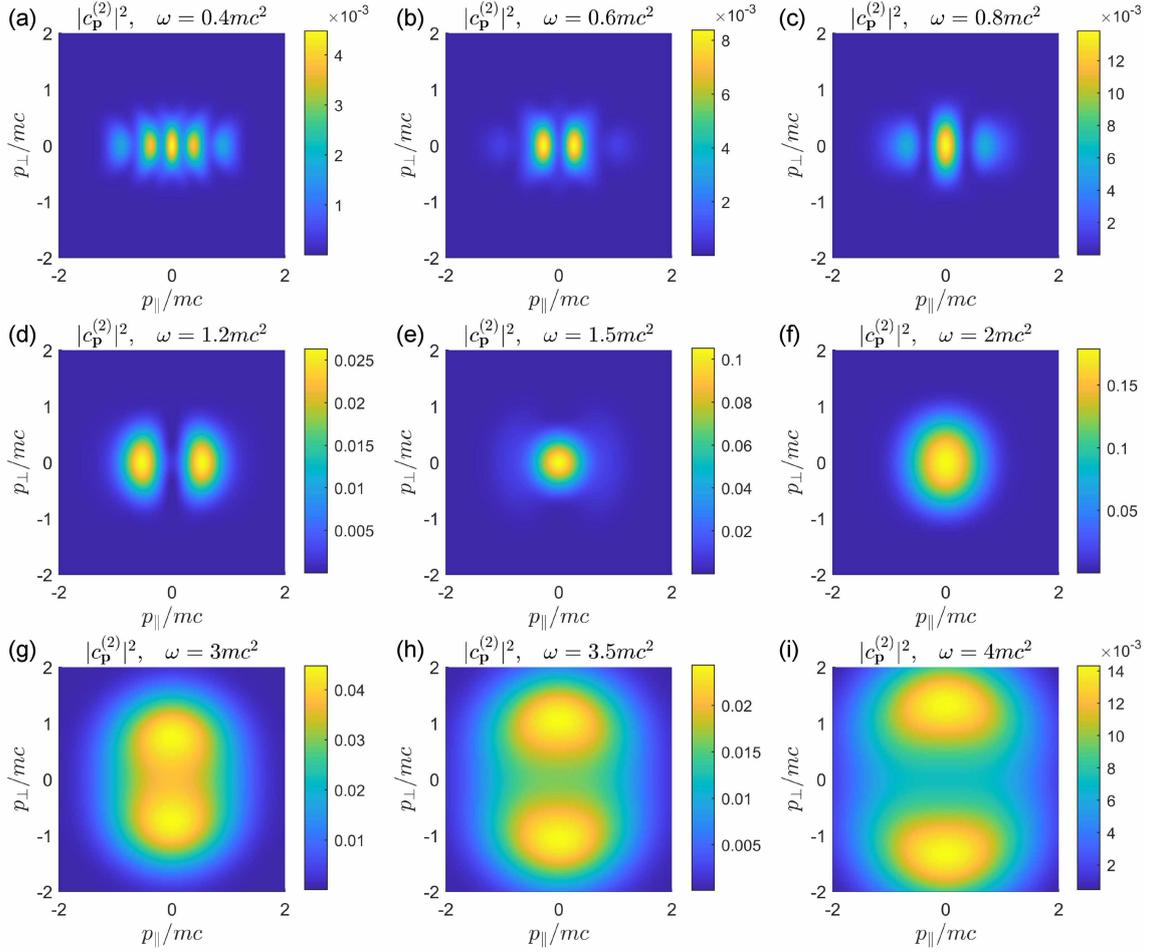


Fig. 4. Momentum distributions of electrons created by an oscillating electric field for different frequencies and for electric field amplitude $\mathcal{E}_0 = 0.5\mathcal{E}_S$. Starting from the one-photon threshold frequency, at roughly $\omega = 2mc^2$, a qualitative change in the shapes of high-probability structures is observed.

phenomenon is illustrated in Fig. 4. For low frequencies $\omega \leq 0.8mc^2$, the distributions are concentrated around the axis of vanishing transverse momentum. However, as the frequency increases, the distributions begin to concentrate around zero momentum. This happens until the one-photon threshold is reached. A further increase of frequency causes the position of the high-probability regions in the distribution to migrate towards the direction perpendicular to the electric field (i.e., towards larger p_{\perp}). Furthermore, at $\omega = 4mc^2$, the high-probability zone in the three-dimensional space takes the shape of a torus centered around $\mathbf{p}_{\perp} = \mathbf{0}$. This means that under such conditions the particles prefer to be ejected in the direction perpendicular to the electric field vector. As shown in [36], the distribution for pair creation, when integrated over particles momenta, starts to saturate (or even decreases) with increasing frequency, leading to the seemingly unexpected stabilization phenomenon. In fact, the stabilization effects appear to be quite common in the strong-field QED, as discussed for instance in [46–50].

In summary, although we have concentrated our discussion on the fermionic distribution function $f(\mathbf{p})$ and the phase of the momentum amplitude $c_{\mathbf{p}}^{(2)}$, the other components of the DHW-functions (8) can be also determined by applying (35), (31) and (29). This topic is, however, beyond the scope of the present work and is going to be considered in due course.

7. Conclusions

In this paper, we formulated the bispinorial approach to the production of e^-e^+ pair in spatially homogeneous electric fields. The method turned out to be equivalent to the DHW formalism, which was introduced in [4]. We have shown that the production of Sauter–Schwinger pair with linearly polarized time-dependent electric field is formally reduced to solving a two-level model, in compliance with [40] (see also references therein). The advantage of this approach is that one gains access to the probability amplitudes and, therefore, to their

phases. The latter allows to uniquely identify vortices and antivortices in the momentum distributions of created pairs. As it has been demonstrated in our paper, for a linearly polarized pulsed electric field, they appear in pairs. We also analyzed the vortex patterns while increasing the field frequency across a two-photon threshold. While we have observed a new vortex–antivortex pair, the general features of the momentum distributions also change across the threshold. Specifically, we have seen that below the one-photon threshold, particles are created most efficiently along the polarization direction of the electric field, whereas above the threshold — in the perpendicular direction. This shows the different characteristics of the Sauter–Schwinger process while passing from low- to high-frequency regimes of electric-field–vacuum interactions.

Acknowledgments

The work of (F.C.V., K.K., and J.Z.K.) was supported by the National Science Centre (Poland) under Grant No. 2018/31/B/ST2/01251.

References

- [1] F. Sauter, *Z. Phys.* **69**, 742 (1931).
- [2] W. Heisenberg, H. Euler, *Z. Phys.* **98**, 714 (1936).
- [3] J.S. Schwinger, *Phys. Rev.* **82**, 664 (1951).
- [4] I. Białynicki-Birula, P. Górnicki, J. Rafelski, *Phys. Rev. D* **44**, 1825 (1991).
- [5] G.R. Shin, I. Białynicki-Birula, J. Rafelski, *Phys. Rev. A* **46**, 645 (1992).
- [6] I. Białynicki-Birula, E.D. Davis, J. Rafelski, *Phys. Lett. B* **311**, 329 (1993).
- [7] Ł. Rudnicki, *Ph.D. Thesis.*, Center for Theoretical Physics, PAS, 2021.
- [8] F. Hebenstreit, R. Alkofer, and H. Gies, *Phys. Rev. D* **82**, 105026 (2010).
- [9] I. Białynicki-Birula, Ł. Rudnicki, *Phys. Rev. D* **83**, 065020 (2011).
- [10] I. Białynicki-Birula, Ł. Rudnicki, A. Wienczek, *arXiv:1108.2615v3*, 2013.
- [11] A. Blinne, H. Gies, *Phys. Rev. D* **89**, 085001 (2014).
- [12] C. Kohlfürst, *Ph.D. Thesis*, *arXiv:1512.06082*, 2015.
- [13] A. Blinne, E. Strobel, *Phys. Rev. D* **93**, 025014 (2016).
- [14] Z.L. Li, Y.J. Li, B.S. Xie, *Phys. Rev. D* **96**, 076010 (2017).
- [15] Z.L. Li, B.S. Xie, Y.J. Li, *J. Phys. B* **52**, 025601 (2019).
- [16] I. Białynicki-Birula, Z. Białynicka-Birula, *Phys. Rev. A* **104**, 022203 (2021).
- [17] X. Sheng, R. Fang, Q. Wang, D.H. Rischke, *Phys. Rev. D* **99**, 056004 (2019).
- [18] F. Hebenstreit, R. Alkofer, H. Gies, *Phys. Rev. Lett.* **107**, 180403 (2011).
- [19] D. Berényi, S. Varró, V.V. Skokov, P. Lévai, *Phys. Lett. B* **749**, 210 (2015).
- [20] H. Al-Naseri, J. Zamanian, G. Brodin, *Phys. Rev. E* **104**, 015207 (2021).
- [21] M. Mohamedsedik, L.-J. Li, B.S. Xie, *Phys. Rev. D* **104**, 016009 (2021).
- [22] C. Kohlfürst, R. Alkofer, *Phys. Lett. B* **756**, 371 (2016).
- [23] C. Kohlfürst, *Eur. Phys. J. Plus* **133**, 191 (2018).
- [24] C. Kohlfürst, *Phys. Rev. D* **101**, 096003 (2020).
- [25] G. Brodin, J. Zamanian, *Rev. Mod. Plasma Phys.* **6**, 4 (2022).
- [26] I. Białynicki-Birula, M. Cieplak, J. Kamiński, *Theory of Quanta*, Oxford University Press, New York 1992.
- [27] I. Białynicki-Birula, Z. Białynicka-Birula, C. Śliwa, *Phys. Rev. A* **61**, 032110 (2000).
- [28] A.J. Taylor, *Analysis of Quantised Vortex Tangle*, Springer, Cham 2017.
- [29] L. Geng, F. Cajiao Vélez, J.Z. Kamiński, L.-Y. Peng, K. Krajewska, *Phys. Rev. A* **102**, 043117 (2020).
- [30] F. Cajiao Vélez, L. Geng, J.Z. Kamiński, L.-Y. Peng, K. Krajewska, *Phys. Rev. A* **102**, 043102 (2020).
- [31] L. Geng, F. Cajiao Vélez, J.Z. Kamiński, L.-Y. Peng, K. Krajewska, *Phys. Rev. A* **104**, 033111 (2021).
- [32] F. Cajiao Vélez, *Phys. Rev. A* **104**, 043116 (2021).
- [33] M.M. Majczak, F. Cajiao Vélez, J.Z. Kamiński, K. Krajewska, *Opt. Express* **30**, 43330 (2022).
- [34] I. Białynicki-Birula, Z. Białynicka-Birula, *Phys. Rev. Lett.* **118**, 114801 (2017).
- [35] I. Białynicki-Birula, Z. Białynicka-Birula, *Phys. Rev. A* **100**, 012108 (2019).
- [36] K. Krajewska, J.Z. Kamiński, *Phys. Rev. A* **100**, 012104 (2019).
- [37] J.M.N. Djiokap, S.X. Hu, L.B. Madsen, N.L. Manakov, A.V. Meremianin, A.F. Starace, *Phys. Rev. Lett.* **115**, 113004 (2015).
- [38] D. Pengel, S. Kerbstadt, D. Johannmeyer, L. Englert, T. Bayer, M. Wollenhaupt, *Phys. Rev. Lett.* **118**, 053003 (2017).
- [39] I. Białynicki-Birula, Z. Białynicka-Birula, *Quantum Electrodynamics*, Ch. 5, Pergamon Press Ltd., Polish Scientific Publishers 1975.

- [40] J.Z. Kamiński, M. Twardy, K. Krajewska, *Phys. Rev. D* **98**, 056009 (2018).
- [41] K. Krajewska, J.Z. Kamiński, *Phys. Rev. A* **100**, 062116 (2019).
- [42] F. Sauter, *Z. Phys.* **73**, 547 (1932).
- [43] K. Krajewska, I.I. Fabrikant, A.F. Starace, *Phys. Rev. A* **74**, 053407 (2006).
- [44] E. Akkermans, G.V. Dunne, *Phys. Rev. Lett.* **108**, 030401 (2012).
- [45] L. Labun, J. Rafelski, *Phys. Rev. D* **79**, 057901 (2009).
- [46] J.Z. Kamiński, K. Krajewska, F. Ehlotzky, *Phys. Rev. A* **74**, 033402 (2006).
- [47] L. Labun, J. Rafelski, *Phys. Rev. D* **98**, 016006 (2018).
- [48] S. Evans, J. Rafelski, *Eur. Phys. J. A* **57**, 341 (2021).
- [49] A.I. Titov, U. Hernandez Acosta, B. Kämpfer, *Phys. Rev. A* **104**, 062811 (2021).
- [50] U. Hernandez Acosta, A.I. Titov, B. Kämpfer, *New J. Phys.* **23**, 095008 (2021).

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

States, Modes, Fields, and Photons in Quantum Optics

M.G. RAYMER^{a,*} AND P. POLAKOS^b^a*Oregon Center for Optical, Molecular and Quantum Science, Department of Physics, University of Oregon, 1585 East 13th Avenue Eugene, OR 97403, USA*^b*Cisco Systems, One Penn Plaza, 9th Floor New York, NY 10119, USA*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S28](https://doi.org/10.12693/APhysPolA.143.S28)*e-mail: raymer@uoregon.edu

The quantum nature of light enables potentially revolutionary communication technologies. A key to advancing this area of research is a clear understanding of the concepts of states, modes, fields, and photons. The concept of field modes carries over from classical optics, while the concept of state has to be considered carefully when treating light quantum mechanically. The term “photon” is an overloaded identifier in the sense that it is often used to refer to either a quantum particle or the state of a field. This overloading, often used without placing it in context, has the potential to obfuscate the physical processes that describe the reality we measure. We review the uses and relationships between these concepts using modern quantum optics theory, including the concept of a photon wave function, the modern history of which was moved forward in a groundbreaking paper in this journal by Iwo Białynicki-Birula, to whom this article is dedicated.

topics: quantum optics, nonclassical light, photon wave function, tutorial

1. Introduction

When beginning the study of quantum optics, it is natural to ask, “What is a photon?” But perhaps a better question is, “What is a quantum field?” Given that quantum theory is agnostic to the names we give to the mathematical elements of the theory, when does it matter how we name and interpret them? Properly conceptualizing and naming the elements of theory helps when trying to build intuition about a problem without the benefit of having a complete mathematical solution at hand. This contribution to the Special Issue dedicated to Professor Iwo Białynicki-Birula reviews in a tutorial manner the role of states, modes, fields, and photons in quantum optics, recognizing his important contributions to the subject. We hope to enlighten researchers who are perhaps new to the field, such as those working in the classical networks arena and now starting to consider the potentially useful applications of quantum networks. We review the concept of a photon wave function, the modern history of which begins more-or-less with a paper in this journal by Białynicki-Birula [1] and a contemporaneous paper by John Sipe [2].

States, modes, and fields are concepts that apply to both classical and quantum domains. The paper reviews in a pedagogical style how these concepts arise and are defined within the two domains, describes how quantization of electromagnetic (EM) field excitations introduces new (and measurable) behaviors, and clarifies the connections between the two domains.

In the arena of applications, we note that in any quantum optical computing or communication system, it is required to control the states of light that interact to carry out a quantum information processing (QIP) task. If control is imprecise, “errors” can occur. In fact, such errors are the main barrier to developing scalable QIP [3]. While for single-particle qubits (e.g. the spin of an electron), the concept of state is clear and routine, for optical qubits, it is not the case due to the multimode nature of the electromagnetic field, and it is worthwhile discussing some of the subtleties that arise in this case. Much has been written on the general subject, such as [4–8] and studies cited therein, but we aim here to cover topics that are not widely emphasized, in particular, the quantization of the EM field in terms of temporal (wave-packet) modes.

2. Classical fields, modes, and states

In the classical physics description, light is a transverse wave of the EM field. The Maxwell equations provide a means to calculate the energy per unit volume stored in the EM field, which may vary continuously. They also provide a wave equation that allows us to calculate the temporal and spatial evolution of the EM field, transporting energy, and momentum. A monochromatic plane wave in free space is identified by specifying values for four distinct attributes (degrees of freedom), any of which can be used to encode information: polarization and three spatial propagation constants k_x, k_y, k_z . In a beamlike geometry, it is often useful to restate these degrees of freedom as polarization, two spatial degrees of freedom describing the transverse beam profile, and frequency, $\omega = c\sqrt{k_x^2 + k_y^2 + k_z^2}$, where c is the speed of light. In either case, the four degrees of freedom define a “mode” of the EM field. A mode can be thought of as a “container” into which differing amounts of energy and momentum can be deposited and carried along by the wave.

Definition. A *classical electromagnetic field* is a physical entity of infinite spatial extent that can transport energy and momentum in the form of wave-like excitations.

Definition. A *mode* $\mathbf{u}_j(\mathbf{r})$ of a classical field is a particular form of a field, which satisfies the Maxwell equations, and a set of which can serve as a mode basis. A common example is given by plane-wave modes propagating in vacuum,

$$\mathbf{u}_j(\mathbf{r}) = \mathbf{e}_j \exp [i(k_{xj}x + k_{yj}y + k_{zj}z)], \quad (1)$$

where j is the mode index (label) and \mathbf{e}_j are the polarization vectors. In a classical description, the energy content of any mode may assume a continuous spectrum of values, proportional to the square of the field amplitude.

It is understood that modes of different frequencies can be added or superposed linearly with differing complex amplitudes a_j to form the (real-valued) electric field, expressed mathematically as $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}^+(\mathbf{r}, t) + \text{cc}$, where cc means complex conjugate and the “positive-frequency part” of the complex field is represented by

$$\mathbf{E}^+(\mathbf{r}, t) = \sum_j^\infty \mathcal{E} a_j e^{-i\omega_j t} \mathbf{u}_j(\mathbf{r}), \quad (2)$$

with a similar expression for the magnetic field. Boldface italic font represents vector quantities and \mathcal{E} is a scalar factor. The amount of energy “occupying” a given mode is proportional to $|\mathcal{E} a_j|^2$. The mode’s shape and propagation direction are contained in the forms of $\mathbf{u}_j(\mathbf{r})$, which form a mutually orthogonal and complete set of functions. As seen by the time evolution $\exp(-i\omega_j t)$, the field in

each mode undergoes single-frequency oscillations and can be described as a simple harmonic oscillator.

Each mode can be viewed as a separate subsystem, the totality of which forms the overall field. When we discuss states of the overall field, in general, we have to specify composite states involving the states of more than one mode. Such composite states can imply correlations between measurement outcomes on different modes.

While (2) is written as a discrete sum of modes, as appropriate in a closed cavity, in unbounded free space the expression for the classical field becomes an integral over a continuum of frequencies. For a beam or pulse propagating in more-or-less a single direction, it is convenient to express it as the integral

$$\mathbf{E}^+(\mathbf{r}, t) = \sum_{\sigma=1,2} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \mathcal{E} a^{(\sigma)}(\omega) e^{-i\omega t} \mathbf{u}^{(\sigma)}(\omega, \mathbf{r}), \quad (3)$$

where σ labels one of two polarization helicities in the case of circular polarization. In the simplest cases, such as in a waveguide or a well-collimated beam in free space, the spatial mode can be separated into transverse and longitudinal parts,

$$\mathbf{u}^{(\sigma)}(\omega, \mathbf{r}) = \mathbf{e}^{(\sigma)} \mathbf{w}^{(\sigma)}(\omega, x, y) \exp(i k_z^{(\sigma)}(\omega) z), \quad (4)$$

where $\mathbf{e}^{(\sigma)}$ is a polarization vector, $\mathbf{w}^{(\sigma)}$ is the transverse part of the mode function, and the function $k_z^{(\sigma)}(\omega)$ describes a dispersion relation (relationship between propagation constant and frequency).

In these equations, $\mathbf{E}(\mathbf{r}, t)$ represents the vector-valued amplitude of the field. Of course, the physical field itself is distinct from its representation; the symbol $\mathbf{E}(\mathbf{r}, t)$ is not the field itself, it is rather a *description* of the field, and many physicists take the field to be an actual element of physical reality. We follow this way of thinking in this paper.

How do we describe the *state* of the field? A general definition of state can be said to be a description of everything that is known about the condition of a physical system at a certain time. In the simplest classical picture, we can have complete knowledge of the field, and the mathematical forms of $\mathbf{E}(\mathbf{r}, t)$ and the magnetic field $\mathbf{B}(\mathbf{r}, t)$ give a full description of its state, that is, a specific configuration of the classical system. Often in optics problems the electric field dominates interactions with matter, such as detectors, and for freely propagating field the electric field is often sufficient for a complete description; so here we focus on describing the state of the electric field.

In general, however, we may possess only partial knowledge, in which case we describe the field’s state using statistical means. For example, thermal light emitted by a blackbody is described as having field amplitudes that are random variables (or random processes) with zero mean value and Gaussian probability density. In principle, one could know

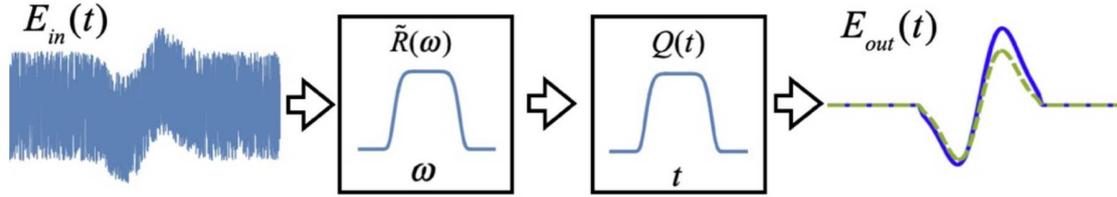


Fig. 1. A coherent signal pulse embedded in a noise background is filtered by passing through a sequence of a spectral filter $\tilde{R}(\omega)$ and a time gate $Q(t)$, resulting in a nearly coherent (single-temporal-mode) field. The envelope waveform (i.e., the field with the carrier wave removed) at the output illustrates both the ideal signal pulse (dashed curve) and the realistic simulated pulse (solid). Reproduced from [11].

the field values (that is, they have definite values at each instant), but in practice, we do not. We call such a state a *statistical state*. Let us summarize.

Definition. A *classical state* is a description of the condition of a system, either representing completely possessed information or a statistical description representing partially possessed information.

In classical physics, a specific state is specified by a point in phase space and, along with the dynamical equations of motion, determines how the phase-space point evolves in time. On the other hand, a statistical state is described by a probability density function (pdf) giving joint probabilities of all possible values of system variables at all combinations of space-time points (for review, see Mandel and Wolf, and Goodman [9, 10]),

$$P_E(\mathbf{E}(\mathbf{r}_1, t_1), \mathbf{E}(\mathbf{r}_2, t_2), \mathbf{E}(\mathbf{r}_3, t_3), \dots). \quad (5)$$

With the classical mode decomposition, we can replace this pdf with a pdf for all the complex mode amplitudes,

$$P_a(a_1, a_2, a_3, \dots). \quad (6)$$

Any expectation values of quantities involving the field can, in principle, be calculated using either of these pdfs.

For a thermal-like classical state of a single mode, the pdf for the complex zero-mean random variable a is

$$P_a(a) = \frac{1}{2\pi\sigma^2} e^{-|a|^2/(2\sigma^2)}, \quad (7)$$

where $2\sigma^2 = \langle |a|^2 \rangle$ is the variance of a . The corresponding pdf for the energy (proportional to $W = |a|^2$) in the mode is

$$P_W(W) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-W^2/(2\sigma^2)}. \quad (8)$$

When one speaks of “mode” in optics, it is often assumed to be the spatial mode (as in a laser cavity). As in (3), one can always decompose the field in terms of products of spatial modes $\mathbf{u}^{(\sigma)}(\omega, \mathbf{r})$ and a multiplicative temporal factor $e^{-i\omega t}$. But in practice, such a monochromatic field would require an infinite time duration to be fully defined or measured. How can we realistically prepare and measure

a single mode in the laboratory? A simple example is shown in Fig. 1: open a small hole in a blackbody cavity for a time T , then spatially filter the emerging light with a distant, small pinhole, then pass it through a spectral filter with small transmitting bandwidth $\Delta\nu$ such that $\Delta\nu T \ll 1$, as described theoretically in [11]. The probability for energy content, in this case, is given by (8).

Such a time-frequency filtering process selects one “time-frequency mode,” also called a temporal-spectral mode, or temporal mode (TM) for short^{†1} [5, 12]. In the following, we will put an emphasis on temporal modes because they provide a mode basis for efficiently describing optical wave packets that are localized in space and time. Encoding and receiving information in such wave-packet form generally requires synchronizing a transmitter with the receiver.

Another important benefit of temporal modes is that they form a discrete set rather than a continuous set, as is the case for monochromatic modes. The discreteness makes it easier to distinguish one mode from another during a detection process. Their discreteness arises from the fact that, by definition, they are confined to a particular space-time region; that is, the boundary condition is that they go to zero at infinity in space and in time. This result is mathematically analogous to the quantization of spatial modes in a cavity or an optical fiber.

To construct TMs, consider a transformation from the monochromatic modes of (3) to the non-monochromatic temporal modes. We can choose any complete orthonormal set of spectral amplitude functions $\{f_j(\omega)\}$ that go to zero at frequencies far from a chosen central frequency ω_0 . By definition, they satisfy

$$\begin{aligned} \int \frac{d\omega}{2\pi} f_i^*(\omega) f_j(\omega) &= \delta_{ij}, \\ \sum_j f_j^*(\omega') f_j(\omega) &= 2\pi \delta(\omega' - \omega), \end{aligned} \quad (9)$$

^{†1}Note: Please do not confuse the abbreviation TM used here with the common terminology TM-mode, meaning transverse magnetic spatial mode.

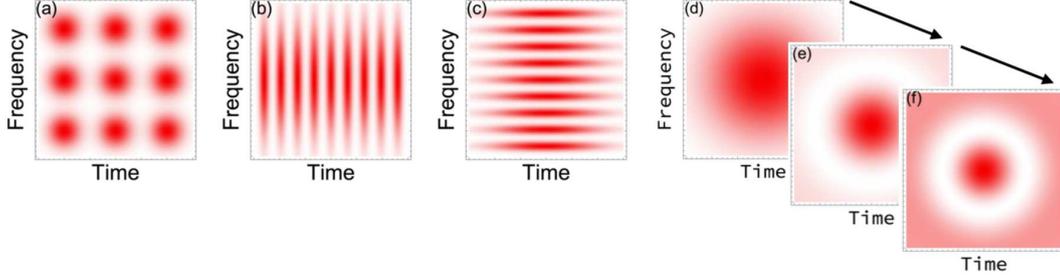


Fig. 2. Examples of sets of temporal modes, plotted as densities in a certain area of time-frequency phase (phasor) space. For a fixed scaling of the time and frequency axes, the modes may be equally broad in both variables, as in panel (a), or they may be narrow in time, as in panel (b), or narrow in frequency, as in panel (c). Such temporal modes may be Gaussian in form and are approximately orthogonal if their separations in time and frequency are large enough. As in panels (d), (e), and (f), an alternative covering of the phase-space area can be accomplished using mode functions that cover the whole region and can be made strictly orthogonal in terms of coherent overlap integration as in (9).

where δ_{ij} and $\delta(\omega)$ are the Kronecker and Dirac deltas, respectively. We can use these functions to define a set of “temporal-mode” amplitudes,

$$A_j = \int \frac{d\omega}{2\pi} f_j^*(\omega) a(\omega), \quad (10)$$

where $f_j(\omega)$ is the spectral amplitude that defines such a mode labeled by j . Hereafter we drop the polarization label σ for notational simplicity. The inverse relation is

$$a(\omega) = \sum_j f_j(\omega) A_j. \quad (11)$$

We see that the continuous (uncountably infinite) set of amplitudes $a(\omega)$ has been converted into a discrete (countably infinite) set A_j .

In terms of TMs, the field in (3) is expressed as

$$\begin{aligned} \mathbf{E}^{(+)}(\mathbf{r}, t) = & \sum_j A_j \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \mathcal{E} f_j(\omega) e^{-i\omega t} \mathbf{e} w(\omega, x, y) e^{ik_z(\omega)z} \simeq \\ & \mathcal{E} \mathbf{e} w(x, y) \sum_j A_j u_j(z, t), \end{aligned} \quad (12)$$

where, for simplicity, we assumed a common polarization vector \mathbf{e} and made the approximation that the transverse mode function $w(\omega, x, y) \simeq w(x, y)$ is independent of frequency, which is valid for reasonably narrow-band fields. Then the propagating temporal modes are defined as

$$u_j(z, t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} f_j(\omega) e^{-i\omega t} e^{ik_z(\omega)z}. \quad (13)$$

At position $z = 0$, the temporal mode equals the Fourier transform of the spectral amplitude

$$u_j(0, t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} f_j(\omega) e^{-i\omega t} = \tilde{f}_j(t). \quad (14)$$

Several example sets of temporal modes are shown in Fig. 2.

In our time-frequency filtering example in Fig. 1, the strongly filtered field supports essentially only one TM,

$$\mathbf{E}^{(+)}(\mathbf{r}, t) \simeq \mathcal{E} A_{j=0} \mathbf{e} w(x, y) u_{j=0}(z, t). \quad (15)$$

Thus, the space and time behavior of the field is determined by the spatial-temporal mode $w(x, y) u_{j=0}(z, t)$, while the (classical) state is determined by the value of (or the statistical properties of) the mode amplitude $A_{j=0}$. The description in terms of temporal modes carries over directly to the quantum treatment of light.

Note that the “incoherent” time-frequency filtering method described in Fig. 1 is necessarily inefficient, in the sense that to achieve a nearly single-temporal-mode field, the filtering needs to be so strong as to pass almost no light. Superior efficiency, approaching 100%, can be achieved using “coherent” filtering with a scheme called a quantum pulse gate, as reviewed in [12] and utilized for noise filtering in [11].

3. Quantum fields, modes, states, and photons

The fact that the classical EM field can be decomposed into a set of classical oscillators inspires us to seek a representation of the EM field as a collection of (bosonic) quantum harmonic oscillators. The creation and annihilation operators are labeled by the continuous frequency variable ω and are defined to obey the commutation relations [13, 14]

$$[\hat{a}(\omega), \hat{a}^\dagger(\omega')] = 2\pi \delta(\omega - \omega'). \quad (16)$$

The non-commutativity of operators for a given frequency embodies the essential difference between quantum and classical theory (and the nature of the physical systems being described). The energy eigenstates of each quantum oscillator are quantized, i.e., they occur at specific, identically-spaced values, $m\hbar\omega$ above the ground state with m

a nonnegative integer. In the quantum theory, the mathematical representation of the field is given by the (Hilbert-space) operator (operators being indicated by carets),

$$\mathbf{E}^{(+)}(\mathbf{r}, t) = \sum_{\sigma=1,2,-\infty} \int \frac{d\omega}{2\pi} \mathcal{E}(\omega) \hat{a}^{(\sigma)}(\omega) e^{-i\omega t} \mathbf{u}^{(\sigma)}(\omega, \mathbf{r}). \quad (17)$$

The form of the commutator (16), along with energy quantization, requires the scale factor to be frequency dependent, $\mathcal{E}(\omega) = \sqrt{\hbar\omega/(2\varepsilon_0 cn)}$, where ε_0 and c are the vacuum electric permittivity and speed of light, respectively, and n is the medium's refractive index at the frequency of interest.

For simplicity, we neglect modal dispersion as occurs in waveguide geometry, discussed in detail in [15]. This simplification allows us to drop any mode labels that refer to which waveguide mode is being considered.

The mode functions $\mathbf{u}^{(\sigma)}(\omega, \mathbf{r})$ are the same as in the classical theory and thus satisfy Maxwell's equations. As in the classical theory, a transformation from monochromatic modes to temporal (wave-packet) modes can be made using (13) for reasonably narrow-band fields. Then the field operator is

$$\mathbf{E}^{(+)}(\mathbf{r}, t) = \mathcal{E} e \mathbf{w}(x, y) \sum_j \hat{A}_j u_j(z, t), \quad (18)$$

where the scale factor for a center frequency ω_0 is $\mathcal{E} = \sqrt{\hbar\omega_0/(2\varepsilon_0 cn)}$ and \hat{A}_j are annihilation operators for the state of light "occupying" the temporal mode $u_j(z, t)$. They are given by

$$\hat{A}_j = \int \frac{d\omega}{2\pi} f_j^*(\omega) \hat{a}(\omega), \quad (19)$$

and by (16) it is easily shown that they satisfy the discrete, rather than continuous, commutation relation

$$[\hat{A}_i, \hat{A}_j^\dagger] = \delta_{ij}. \quad (20)$$

In free space (that is, with no interactions), the time evolution, expressed in the Heisenberg picture, is fully contained in the wave-packet propagation of the modes in (13)^{†2}.

Now we may ask, "What is the quantum field?" It is not $\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t)$, which is an operator that represents mathematically the annihilation of energy excitations in the field. The quantum field itself (from one meaningful point of view) is a physical entity that is "out there" and is capable of carrying energy and momentum from one place to another.

How do we specify states of the quantum field?

Definition. A *quantum state* is a mathematical form used to determine the probabilities for particular outcomes of any possible measurement on a system, either as a pure state (representing maximal possessed information) or a mixed state (representing partially possessed information).

The most general case is the mixed state, expressed mathematically as a density operator,

$$\hat{\rho} = \sum_j P_j |\Psi_j\rangle \langle \Psi_j|, \quad (21)$$

where P_j is the (classical) probability that the system is in the pure state $|\Psi_j\rangle$. We say such a state is a statistical mixture of pure states. In the ideal limit, if all P_j are known to be zero except a single one, say $P_0 = 1$, then we can describe the state simply by specifying the form of $|\Psi_0\rangle$, which in some cases is described by a wavefunction of some systems variable or in other cases as a vector in an abstract linear vector space.

A starting point for the quantum state description of the field is the vacuum state, $|\text{vac}\rangle$, which carries no energy or momentum (at least not in a way that is detectable via absorption by an atom or a photodetector). The simplest non-vacuum state of the field is the single-photon state, and its generalization, the n -photon (Fock) state, written in the temporal-mode formalism for mode $u_j(z, t)$ as

$$\begin{aligned} |1\rangle_j &= \hat{A}_j^\dagger |\text{vac}\rangle = \int \frac{d\omega}{2\pi} f_j(\omega) \hat{a}^\dagger(\omega) |\text{vac}\rangle, \\ |n\rangle_j &= \frac{1}{\sqrt{n!}} (\hat{A}_j^\dagger)^n |\text{vac}\rangle, \end{aligned} \quad (22)$$

where \hat{A}_j^\dagger is the creation operator for a given TM and n is the photon occupation number of a given TM of the field.

It is interesting that although an n -photon state of a particular temporal mode does contain a specific sharp number of photons, it does not contain a sharp value of energy, because the mode itself is constructed as a sum of modes having differing frequencies and each frequency mode has an unspecified number of photons (although they must sum to n). If the light is well collimated, quantum measurement of its energy content can be carried out using a spectrometer — disperse it on a diffraction grating followed by a dense array of photon-counting detectors. If the number of detectors is much greater than the number of photons n , and the detectors are 100% efficient, then exactly n detectors will register a detection event ("click"). Each detector will correspond to a resolved channel l with energy $\hbar\omega_l$. For $n = 1$, the probability for a given detector to click is given by $|f_j(\omega)|^2$. In general, the pattern of detectors that click will indicate the total energy observed for that measurement trial.

To sum up, in the quantum theory of a collection of oscillators of different frequencies (in the temporal-mode formalism), what gets quantized is

^{†2}By the way, the assumption of narrow-band wave packets is not an essential step, but when this approximation is removed, it turns out that the resulting wave packets are not strictly orthogonal in space. This complication arises only for wave packets whose duration is less than around 10 fs, which is not usually the case in optical communications applications. See [16].

not total energy but total excitation. You can have zero, one, or two excitations but not half an excitation.

A general *pure state* of the field in a given temporal mode is expressed as

$$|\Psi\rangle_j = \sum_{n=0}^{\infty} c_n |n\rangle_j = \sum_{n=0}^{\infty} \frac{c_n}{\sqrt{n!}} \hat{A}_j^{\dagger n} |\text{vac}\rangle, \quad (23)$$

where $|c_n|^2$ is the probability of observing n clicks in a detector array, as just described. The coherent state, with $c_n = \exp(-|\alpha|^2/2)\alpha^n/\sqrt{n!}$ for some complex amplitude α , is the state from an ideal laser emitting a pulse in the temporal mode $u_j(z, t)$. Then the probabilities are given by the Poisson distribution $|c_n|^2 = \exp(-|\alpha|^2)|\alpha|^{2n}/n!$.

A *mixed state* of the EM field can be represented by a density operator for a given temporal mode,

$$\hat{\rho} = \sum_j p_j \frac{1}{\sqrt{n!}} \hat{A}_j^{\dagger n} |\text{vac}\rangle \langle \text{vac}| \frac{1}{\sqrt{n!}} \hat{A}_j^n, \quad (24)$$

where p_j is the probability to find n photons in mode j if the photon number is measured. Interestingly, there is a different kind of mixed state, which has a definite number of photons (field excitations), but they are spread incoherently across several modes, for example, a single-photon mixed state is

$$\hat{\rho} = \sum_j P_j \hat{A}_j^{\dagger} |\text{vac}\rangle \langle \text{vac}| \hat{A}_j, \quad (25)$$

where P_j is the probability that the photon will be found in mode j if detected in a mode-selective manner.

4. Particles or fields?

Now we can try to clarify “What is a photon?” It is preferable not to think of the photon as a thing or a physical entity, but rather as simply one of the names we use to specify states of the field.

So, when we say, “The atom emitted a photon,” what we actually mean is “The atom lost energy, creating a single-photon state of the field.” It is almost always safe (prudent) to replace “photon” with “single-photon state of the field.” If we wish to have a more physically suggestive way to define a photon, we can say it is a single “excitation” of the quantum field.

Nevertheless we note that a single-photon state can be thought of in two equivalent ways: as the state of a photon as a distinct entity, or as the state of a field. Consider such a state occupying a particular wave-packet mode $\mathbf{u}_1(\mathbf{r}, t)$ that is concentrated in a finite volume and traveling through space. Horizontal and vertical polarization states $|H\rangle, |V\rangle$ define a basis for describing the state of the photon. Say, the photon’s state is diagonal $|D\rangle = (|H\rangle + |V\rangle)/\sqrt{2}$. Alternatively, we can write this photon state in terms of the state of the field, using modes that specify the field’s spatial and polarization aspects denoted

$$\mathbf{u}_H(\mathbf{r}) = \mathbf{e}_H u_1(\mathbf{r}, t), \quad \mathbf{u}_V(\mathbf{r}) = \mathbf{e}_V u_1(\mathbf{r}, t), \quad (26)$$

where $\mathbf{e}_H, \mathbf{e}_V$ are polarization vectors. Then we can express the single-photon diagonal-polarization state in terms of the occupation numbers of the $\mathbf{u}_H(\mathbf{r}), \mathbf{u}_V(\mathbf{r})$ modes,

$$\frac{|1\rangle_H |0\rangle_V + |0\rangle_H |1\rangle_V}{\sqrt{2}}. \quad (27)$$

Furthermore, we can choose a different mode basis to represent the same state. If we choose modes that define the field as diagonal and anti-diagonal polarized, $\mathbf{u}_D(\mathbf{r}) = (\mathbf{u}_H(\mathbf{r}) + \mathbf{u}_V(\mathbf{r}))/\sqrt{2}$, $\mathbf{u}_A(\mathbf{r}) = (\mathbf{u}_H(\mathbf{r}) - \mathbf{u}_V(\mathbf{r}))/\sqrt{2}$, then the same state is represented as $|1\rangle_D |0\rangle_A$. Thus, a transformation of the “mode basis” from $\mathbf{u}_H(\mathbf{r}), \mathbf{u}_V(\mathbf{r})$ to $\mathbf{u}_D(\mathbf{r}), \mathbf{u}_A(\mathbf{r})$ corresponds to a change of the “state basis” from $|H\rangle, |V\rangle$ to $|D\rangle, |A\rangle$.

The quantum theory of light can be constructed from either of two distinct starting points, as was made clear by Dirac in his book [17], where he wrote, “The dynamical system consisting of an assembly of similar bosons is equivalent to the dynamical system consisting of a set of oscillators — the two systems are just the same system looked at from two different points of view.”

Given the equivalence of the two points of view, why do many quantum physicists find it more fruitful to consider the field (that is, the collection of oscillators) as the fundamental physical entity, as we have done above? Steven Weinberg, a Nobel-winning quantum theorist, said, “Thus, the inhabitants of the universe were conceived to be a set of fields — an electron field, a proton field, an electromagnetic field — and particles were reduced to mere epiphenomena.” The reasons for this choice are many and have been summarized in language suitable for the general reader in several accounts, including Hobson [18], Wilczek [19], and Raymer [20]. Here we offer a summary of such arguments.

1. Quantum fields respect relativity.

Frank Wilczek, also a Nobel-winning quantum theorist, writes, “The concept of locality, in the crude form that one can predict the behavior of nearby objects without reference to distant ones, is basic to scientific practice.” Quantum field theory successfully describes all known phenomena (that it has been applied to) without invoking action at a distance, which would violate Einstein’s relativity [19].

2. Quantum fields evince identical particles.

Wilczek also writes, “Undoubtedly the single most profound fact about Nature that quantum field theory uniquely explains is the existence of different, yet indistinguishable, copies of elementary particles.” It is known that the world is made of a limited number of particle types, and that any two members of the same type are identical. For example, any two

electrons are identical, that is, interchangeable. “We understand this as a consequence of the fact that both are excitations of the same underlying ‘ur-stuff’, the electron field. The electron field is thus the primary reality,” Wilczek says. And the same holds true for the electromagnetic field and its photon excitations [19].

3. Quantum fields naturally account for changing numbers of particles.

Quantum field theory not only accounts for the creation and destruction of photons when atoms emit or absorb light. It also accounts for processes such as creation and destruction of electrons and positrons. Wilczek writes, “In this picture it is only the fields, and not the individual objects they create and destroy, that are permanent” [19].

4. Quantum fields give a clearer picture of wave-particle duality.

The electron–matter field is not an electron. Rather, an electron is an individual excitation of the electron–matter field, just as a photon is an individual excitation of the EM field. The quantum fields themselves behave in a wave-like manner and represent possible measurements to determine where the electron (or photon) is located. Therefore, it is not surprising that if one mistakenly believes that an electron, for example, is a particle, meaningless questions can arise. For example, the question “Which path did the electron take on its way to a detector?” has no meaning. On the other hand, a quantum field permeates all of space; therefore, it exists within both paths. So, the proper statement is not that an electron sometimes behaves like a wave and sometimes like a particle. One should rather say that the quantum field always behaves like a quantum field with its wave-like behaviors, and the electron is a manifestation of that field. It is best to replace the mysterious concept of “wave-particle duality” with the less mysterious concept of “quantum field-quantum particle duality” (adapted from [20]).

5. Superposition, separability, and entanglement

How is entanglement different from superposition? For a single quantum entity, $|\psi\rangle_A + |\phi\rangle_A$ is a superposition state, wherein the “+” symbol represents the *superposition of possibilities* and can be read as “in superposition with.” Here normalization factors have been dropped. For a pair of entities, $|\psi\rangle_A |\phi\rangle_B + |\phi\rangle_A |\psi\rangle_B$ is also a superposition state, but now involving a larger state space.

A nonseparable state of two entities, A and B , is one that *cannot* be written as a product state, $|\phi\rangle_A |\psi\rangle_B$. (For simplicity, we consider only pure states).

It is worth unpacking what is meant by “quantum entity.” In quantum photonics, the entities are the field modes, and if they are correlated or entangled, we need to specify a composite (joint) state in a higher-dimension state (Hilbert) space. Whether or not we call a nonseparable state entangled depends on the situation and, to some extent, the semantic preference of the user. According to some users, it is fair to call all nonseparable quantum states entangled. That would include the state of the two electrons in a helium atom. Other users will insist on reserving the use of the word “entangled” for cases in which the two entities are not interacting (unlike two electrons in a helium atom) and can be measured independently in separated regions of space. The justification for this stricter definition is that in quantum information science, entanglement is regarded as a resource for accomplishing tasks such as the teleportation of a state across some distance. In this context, a powerful known fact is that the amount of such entanglement cannot be increased by any quantum operations in which each entity is transformed only locally, even in the presence of classical (non-quantum) communication between parties at the two locations. If one wishes to have distinct names to classify entanglement, one could say that nonseparability with spatial separation is called “useful entanglement.”

For a two-mode nonseparable photonic state, can we transform to a new mode basis for which the state is separable? From the above discussion, we see the answer is “No” if the entities are spatially separated, and we restrict the basis changes to involve only the local states of each entity. But if we carry out arbitrary global mode-basis change involving all relevant modes of the combined system, the answer is “Yes, for all typically realizable states.” For example, note that we can write the above-mentioned single-photon state (now with normalization) as

$$\frac{1}{\sqrt{2}}(|1\rangle_A |0\rangle_B + |0\rangle_A |1\rangle_B) = \frac{1}{\sqrt{2}}(\hat{A}^\dagger + \hat{B}^\dagger)|\text{vac}\rangle, \quad (28)$$

where $\hat{A}^\dagger, \hat{B}^\dagger$ are creation operators for two spatially orthogonal modes, as represented by (19). Consider the two-mode transformation of the spectral amplitudes,

$$\begin{pmatrix} f_A(\omega) \\ f_B(\omega) \end{pmatrix} \rightarrow \begin{pmatrix} f_C(\omega) \\ f_D(\omega) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} f_A(\omega) \\ f_B(\omega) \end{pmatrix}, \quad (29)$$

which is equivalent to the transformation of the mode functions on a 50/50 beam splitter,

$$\begin{pmatrix} u_A(z, t) \\ u_B(z, t) \end{pmatrix} \rightarrow \begin{pmatrix} u_C(z, t) \\ u_D(z, t) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u_A(z, t) \\ u_B(z, t) \end{pmatrix}, \quad (30)$$

such that the operators transform as

$$\begin{pmatrix} \hat{A}^\dagger \\ \hat{B}^\dagger \end{pmatrix} \rightarrow \begin{pmatrix} \hat{C}^\dagger \\ \hat{D}^\dagger \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \hat{A}^\dagger \\ \hat{B}^\dagger \end{pmatrix}. \quad (31)$$

The state in (28) can then be written as

$$\frac{1}{\sqrt{2}} (\hat{A}^\dagger + \hat{B}^\dagger) |\text{vac}\rangle = \hat{C}^\dagger |\text{vac}\rangle = |1\rangle_C |0\rangle_D. \quad (32)$$

This is just a formal way to say that sending a single-photon state into a 50/50 beam splitter yields a mode-entangled state of the two emerging fields. The beam-splitter transformation is reversible, so we can send the mode-entangled state into the ports of a beam splitter and end up with a single-photon state in mode C (or in mode D if we adjust the phase). Thus, the state is nonseparable (entangled) in one mode basis but not in the other mode basis. Physically, undoing the entanglement (nonseparability) requires bringing the modes together to perform a “global” transformation on them. As said above, the disentanglement cannot be accomplished by any “local” transformations involving modes A and B separately.

Sperling et al. [7] pointed out that there exist entangled (nonseparable) states that cannot be disentangled (made separable) by any unitary mode transformation, global or not. At present, such states are not readily created using common techniques and have not played a role in quantum information science.

A very interesting point is that entanglement can exist with only one “particle” present. If we consider two EM field modes, A and B , that occupy separate regions of space (e.g., two well-separated wave-packet modes, $u_A(\mathbf{r}, t)$ and $u_B(\mathbf{r}, t)$), we could prepare each mode so it contains either zero or one photon’s worth of quantum-field excitation, but there is only one photon in total. Label the state of the EM field in each mode by either (1) if it has one photon’s worth of excitation, or (0) if it has none. The state $|1\rangle_A |0\rangle_B + |0\rangle_A |1\rangle_B$ refers to an entangled state of the fields in two modes. It may be helpful to recognize that modes represent different degrees of freedom of the EM field.

The state is not an entanglement of particle states. If you “believe” in photons as particles, the state would be a mere superposition, $|A\rangle + |B\rangle$, not entangled in that context. The fact that two distinct EM modes can have entanglement even when there is only one photon shared between them suggests again that EM fields, not particles, are truly the physical entities [6].

Note that there is an important difference between the joint state of the A , B modes being considered here and the state of the H , V modes in (27). The latter represents an excitation that is necessarily confined to a single spatial region, defined by the common spatial mode. Therefore, such a state is not considered to be the same kind of quantum resource as the entangled state of two

spatially separated modes. To make the point more mathematical, note that if we were to specify the latter state fully, we should indicate the state of all the relevant degrees of freedom of the field, including spatial location and polarization. Therefore, the state being considered in the present example is actually, assuming a common (H) polarization,

$$|1\rangle_{A,H} |0\rangle_{A,V} |0\rangle_{B,H} |0\rangle_{B,V} + |0\rangle_{A,H} |0\rangle_{A,V} |1\rangle_{B,H} |0\rangle_{B,V}. \quad (33)$$

6. Mode errors and state errors

We turn to the question of what types of errors can occur when quantum information is being encoded in the states of light, such as in a quantum communication system.

We saw in the examples just given that a mode transformation can be thought of as a change of state, which is a Schrödinger-picture way of thinking. Or we can think, in a Heisenberg-picture way, that the global state has not changed, but only the state basis has been altered. Both ways of thinking are valid if we keep our pictures clear. Nevertheless, the distinction between modes and states can have practical consequences, or at least can allow one to categorize “errors” that might occur in a quantum information science (QIS) scheme such as a quantum network.

Let us say we created a single-photon state as a superposition of two time-bin states with complex amplitudes α and β ,

$$|\psi\rangle = (\alpha \hat{A}_1^\dagger + \beta \hat{A}_2^\dagger) |\text{vac}\rangle = \alpha |1\rangle_{u1} |0\rangle_{u2} + \beta |0\rangle_{u1} |1\rangle_{u2}, \quad (34)$$

where $u1$ and $u2$ refer to temporal modes $u_1(z, t)$ and $u_2(z, t)$ defined by (13) and taken here to be separated time-bin modes that are orthogonal to a good approximation. We could encode a qubit using the two states

$$|\psi\rangle_1 = |1\rangle_{u1} |0\rangle_{u2}, \quad |\psi\rangle_2 = |0\rangle_{u1} |1\rangle_{u2}, \quad (35)$$

or we could choose a transformed (“rotated”) basis in which to encode the qubit,

$$|\psi\rangle_+ = \frac{|1\rangle_{u1} |0\rangle_{u2} + |0\rangle_{u1} |1\rangle_{u2}}{\sqrt{2}},$$

$$|\psi\rangle_- = \frac{|1\rangle_{u1} |0\rangle_{u2} - |0\rangle_{u1} |1\rangle_{u2}}{\sqrt{2}}. \quad (36)$$

These four states are illustrated in Fig. 3.

These two choices of bases are called mutually unbiased bases (MUBs) [21]. If a qubit is created in any of the two states in one MUB, the probabilities for detecting the qubit in either of the states of the other MUB are equal and thus unbiased. Such pairs of MUBs play important roles in quantum key distribution (QKD). For a fun simulation of QKD, see [22].

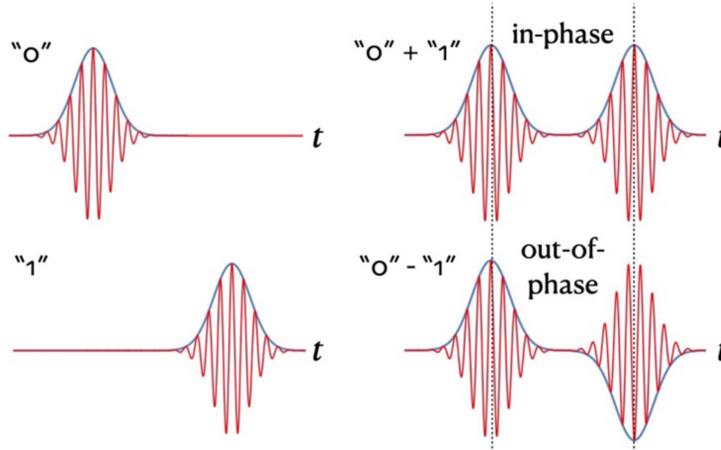


Fig. 3. A qubit can be encoded in two possible bases. On the left are two time-bin states that are nearly orthogonal, and on the right are two superposition states, which are nearly orthogonal by virtue of a relative phase shift of the components and can equally well be used as a basis.

Given that we launch one of these four states into a QIP system such as a quantum network, two kinds of errors can occur — state errors and mode errors. Recall that we are thinking about states of fields, not of particles.

A state error occurs if coefficients of the state expansion (e.g., in (36)) are modified while the forms of the modes (in this example, time-bin or temporal modes) remain unchanged. For example, photons (field excitations) could be lost from the modes of interest by scattering or absorption, never to be recovered. Or photons could be added by leakage of light from other modes into the modes of interest. Another kind of state error is non-deterministic (random) dephasing between components of the state, for example,

$$|\psi\rangle_+ \rightarrow \frac{|1\rangle_{u1}|0\rangle_{u2} + e^{i\phi(t)}|0\rangle_{u1}|1\rangle_{u2}}{\sqrt{2}}, \quad (37)$$

where $\phi(t)$ is an unknown, uncontrollable phase. Such a change would drive the pure state into a mixed state, from which the original pure state cannot be recovered.

In contrast, a *mode error* occurs if the forms of modes become altered by some physical process while the state remains unchanged. Linear dispersion in a fiber, which is deterministic, can often be reversed or otherwise compensated by a physical process such as a prism pair. Alternatively, one can simply redefine the modes of interest to be those that exist following the predictable effects of dispersion. Notice that dispersion is already accounted for in the definition of the temporal modes $u_j(z, t)$ in (13) through the dispersion relation $k_z(\omega)$. Another kind of deterministic mode change occurs simply by time delay, either of the two-pulse wave packet as a whole, or a change of the time delay between the time bins being used to encode the qubit. As with dispersion, such changes can be compensated or accounted for theoretically. These examples

point out the importance of knowing any changes of the modes during propagation, as well as the need for synchronization in most designs for a quantum network (although important progress has been made toward using single-atom quantum memories to remove the need for synchronization [23]).

Other kinds of deterministic (unitary) mode changes can also occur, such as linear mixing of the modes of interest with other modes in processes analogous to linear beam splitting. Such a process might be reversible if all the involved modes can be controlled.

Interestingly, the state change illustrated in (37) can be interpreted instead as a mode change, in particular,

$$u_1(z, t) \rightarrow u_1(z, t), \quad u_2(z, t) \rightarrow e^{-i\phi(t)}u_2(z, t), \quad (38)$$

i.e., a random dephasing between temporal-mode components of the state. That is, in this case, the error may be viewed as either a mode or state error since the phase change can be absorbed into either the state coefficients or the mode functions.

7. Photon wave functions

Finally, we discuss the question — if we, for pedagogical reasons, want to depart from the approach taken until now in this paper, which treats fields, not particles, as the fundamental quantities, how far can we go? Let us first remark that one often hears talk about “which path” a photon might take in, for example, a double-slit experiment. Such a question presupposes that a photon is a particle-like entity that has a trajectory, or at least a set of possible trajectories. In contrast, in the field-theoretic approach, such a question never needs to be asked because the field fills all of space, so it makes no sense to ask which path it takes.

Possible answers to the question depend on how strict one wants to be in defining what a particle is. If one simply conceives of particles as discrete but otherwise abstract entities that carry energy and momentum, then there seems to be no problem defining a wave function that describes its properties and dynamics. However, if one insists that a particle be an entity that can be localized to a *point* in space, then complications arise.

The question, therefore, leads us to the intriguing possibility of defining a wave function for a photon — a topic that was explored in the early days of quantum theory and more recently and fruitfully by Iwo Białynicki-Birula, whom this paper is meant to honor [1]. For a wave function for a photon in coordinate space, there should be a corresponding Schrödinger equation that it satisfies. But what is this wave function, and what is its Schrödinger equation? The simple answer is that the photon wave function is a wave packet that satisfies Maxwell's equations. Therefore, the single-photon wave function could be given by a monochromatic mode such as $\mathbf{u}^{(\sigma)}(\omega, \mathbf{r})$ in (3) or superpositions forming a temporal mode such as $u_j(z, t)$ in (13), and its Schrödinger equation has the form of Maxwell's equations. Let us explore this statement further.

As developed in [1, 2] and reviewed in [16], a compact way to write classical Maxwell's equations is to combine the (real) electric and magnetic fields, $\mathbf{E}(\mathbf{r}, t)$, $\mathbf{B}(\mathbf{r}, t)$, into a single complex field, called the Riemann–Silberstein (RS) vector field,

$$\Psi_\sigma(\mathbf{r}, t) = \sqrt{\frac{\varepsilon_0}{2}} \left(\mathbf{E}(\mathbf{r}, t) + i\sigma c\mathbf{B}(\mathbf{r}, t) \right), \quad (39)$$

where $\sigma = \pm 1$ describes the field's helicity (circular polarization). In (39), $\Psi_\sigma(\mathbf{r}, t)$ represents two fields, one for each value of σ , and, if one wishes, it can be combined into a single two-entry entity $\{\Psi_{+1}, \Psi_{-1}\}$ (analogous to a Dirac spinor). For a single helicity, the free-space Maxwell's equations with no charges or currents present can be written as a vector cross product,

$$i \frac{\partial}{\partial t} \Psi_\sigma(\mathbf{r}, t) = c\sigma \nabla \times \Psi_\sigma(\mathbf{r}, t) \quad (\sigma = \pm 1). \quad (40)$$

We can think of this form as a Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} \Psi_\sigma(\mathbf{r}, t) = H \Psi_\sigma(\mathbf{r}, t), \quad (41)$$

where the Planck constant \hbar has been inserted on both sides and the (Maxwell) Hamiltonian operator is defined as $H = \hbar c\sigma \times$.

It is interesting that (41) is a wave equation for the electromagnetic field that is first-order in the time derivative, whereas in classical optics we usually think of a wave equation that is second order. The first-order wave equation here is simply an alternate way to write the two Maxwell's equations together, and it embodies more information than the familiar second-order wave equation^{†3}.

One can fruitfully consider that (41) is the Schrödinger equation for a single photon (whether one regards the photon as a particle or as an excitation of the quantized EM field). This statement can be “tested” by performing a so-called second quantization of the photon wave function to construct a quantum field theory that permits more than a single excitation. That is, replace the classical function with an operator, $\Psi_\sigma(\mathbf{r}, t) \rightarrow \hat{\Psi}_\sigma(\mathbf{r}, t)$,

$$\hat{\Psi}_\sigma(\mathbf{r}, t) = \sum_j \hat{b}_j^{(\sigma)} \Psi_j^{(\sigma)}(\mathbf{r}, t) + \sum_j \hat{b}_j^{(\sigma)\dagger} \Psi_j^{(\sigma)*}(\mathbf{r}, t), \quad (42)$$

where $\Psi_j^{(\sigma)}(\mathbf{r}, t)$ are (vector) mode functions, and the creation and annihilation operators satisfy $[\hat{b}_j^{(\sigma)}, \hat{b}_k^{(\sigma')\dagger}] = \delta_{jk} \delta_{\sigma\sigma'}$. Hereafter for concreteness, we consider a single helicity, $\sigma = +1$.

We identify this field operator as the complex sum of electric and magnetic field operators (indicated by carets),

$$\hat{\Psi}_{+1}(\mathbf{r}, t) = \sqrt{\frac{\varepsilon_0}{2}} \left(\hat{\mathbf{E}}(\mathbf{r}, t) + ic\hat{\mathbf{B}}(\mathbf{r}, t) \right). \quad (43)$$

Then we find the electric field part, using

$$\frac{1}{\sqrt{2\varepsilon_0}} \left(\hat{\Psi}_{+1}(\mathbf{r}, t) + \hat{\Psi}_{+1}^\dagger(\mathbf{r}, t) \right) = \hat{\mathbf{E}}(\mathbf{r}, t), \quad (44)$$

and defining positive and negative-frequency parts, $\hat{\mathbf{E}}(\mathbf{r}, t) = \mathbf{E}^{(+)}(\mathbf{r}, t) + \mathbf{E}^{(-)}(\mathbf{r}, t)$, where

$$\begin{aligned} \mathbf{E}^{(+)}(\mathbf{r}, t) &= \frac{1}{\sqrt{2\varepsilon_0}} \sum_j \hat{b}_j \Psi_j(\mathbf{r}, t), \\ \mathbf{E}^{(-)}(\mathbf{r}, t) &= \frac{1}{\sqrt{2\varepsilon_0}} \sum_j \hat{b}_j^\dagger \Psi_j^*(\mathbf{r}, t), \end{aligned} \quad (45)$$

and we dropped the σ label for simplicity. Comparing with (18) and identifying the operators by $\hat{b}_j = \hat{A}_j$ and photon wave functions by $(2\varepsilon_0)^{-\frac{1}{2}} \Psi(\mathbf{r}, t) = \mathcal{E}w(x, y)\mathbf{e}_j u_j(z, t)$, we see that the second quantization procedure leads directly to the spatial-temporal-mode formalism of quantum optics.

As long as we restrict our considerations to reasonably narrow-band fields in each mode, where the frequencies are near a central carrier frequency $\omega \simeq \omega_0$, and the bandwidth is much smaller than ω_0 , we have $\mathcal{E}(\omega) \simeq \mathcal{E} = \sqrt{\hbar\omega_0}/(2\varepsilon_0 cn)$. Then the single-photon wave functions corresponding to different annihilation operators are orthogonal to a good approximation, in the sense that integrating over all space yields

$$\int d^3r \Psi_j^*(\mathbf{r}, t) \cdot \Psi_k(\mathbf{r}, t) = \delta_{jk}. \quad (46)$$

A subtlety arises when considering exotic ultra-broadband photons with bandwidth comparable to (say, 50% or greater than) the carrier frequency.

^{†3}Historically, the desire for a first-order-in-time wave equation is what drove Dirac to formulate his famous relativistic Schrödinger wave equation for the electron.

Then a more careful analysis starting from (17) with the frequency dependence retained in $\mathcal{E}(\omega)$ shows that the photon wave functions corresponding to different annihilation operators cannot be strictly orthogonal in space [16]. This non-orthogonality means that, strictly speaking, a single-photon state cannot be localized to a point in space; that is, there is no local spatial probability of finding the photon (thinking of a particle) at any particular point. In general, it is more accurate to say that the modulus-squared of a photon wave function $\Psi_j(\mathbf{r}, t)$ describes the spatial distribution of probabilities to detect the photon's *energy* concentrated around different locations \mathbf{r} , rather than to find the photon (as if it were a particle) at a specified point location. The photon, viewed as a particle or as a state of the field, always remains "spread out" within a region with a minimum volume equal to a cubic wavelength.

It should be mentioned that the same complication arises in the temporal-modes formalism when considering ultra-broadband temporal modes, because it is mathematically equivalent to the second-quantized photon-wave-function formalism. In practice, such details have not (yet) been found to have significant consequences in quantum information science, where ultra-broadband photons are not typically employed.

If there are two field excitations (photons) present, the concept of a two-photon wave function becomes relevant. Its modulus-squared gives the probability for finding the energies of the two photons concentrated around locations \mathbf{r}_1 and \mathbf{r}_2 . By analogy with a two-electron wave function, such a function has been defined as (suppressing the polarization label for simplicity) [16]

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t) = \sum_{j,k} C_{j,k} \Psi_j(\mathbf{r}_1, t) \otimes \Phi_k(\mathbf{r}_2, t), \quad (47)$$

where $\Psi_j(\mathbf{r}_1, t)$ and $\Phi_k(\mathbf{r}_2, t)$ are single-photon wave functions, and the product is a vector direct product. In fact, $C_{j,k}$ can be chosen arbitrarily as long as it ensures the symmetry properties of a two-boson state, namely $\Psi(\mathbf{r}_2, \mathbf{r}_1, t) = \Psi(\mathbf{r}_1, \mathbf{r}_2, t)$. Two-photon wave functions expressed in spatial coordinates are not often used in quantum optics theory because the quantized field method is nearly always more direct and convenient, especially in scenarios where the number of photons changes in time. A caveat is that the two-photon wave function is equivalent to the so-called *two-photon detection amplitude*, which arises naturally in standard quantum optics theory when considering joint detection of two-photon states [9, 24]. Such a formalism arises naturally from the standard quantum optics theory when analyzing optical detection.

Thus, we see that a photon-as-particle viewpoint can be formulated and used if one carefully understands its limitations and its relation to standard quantum optics theory, which is based on quantizing the EM field. Further developments have included

treating the case when light interacts with matter; then, the one- or two-photon wave-function approach has to be modified, as done, for example, by Saldanha and Monken and by Keller [25, 26].

When considering fields with more than two excitations (photons), the wave-function picture quickly becomes inconvenient and overly cumbersome (it becomes a many-body quantum theory), again giving credence to the preference among theorists to stick with the quantized-field approach [16].

8. Mode interference and state interference

Electromagnetic modes satisfy Maxwell's equations, and therefore optical interference is built into the quantum theory from the start. The mode transformation in (30) is an example of interference. What is perhaps confusing is that modes interfere (classically) and quantum-state amplitudes can also interfere (quantumly). For a single-photon state, you can think of "classical" mode interference (e.g., on a beam splitter) as a quantum change of state or a change of mode basis, as in (30). For multi-photon states, or states with an indefinite number of photons, the situation is more subtle.

As mentioned earlier, a coherent state of a single monochromatic mode is expressed as

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} (\hat{a}^\dagger)^n |\text{vac}\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \quad (48)$$

and is found (upon measurement) to contain n photons with probability $\exp(-|\alpha|^2)|\alpha|^{2n}/n!$, a Poisson distribution. If the coherent-state field passes through a phase-shifting element such as a piece of transparent glass, it picks up a phase shift θ , which manifests in the coherent state as [27]

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{(\alpha e^{-i\theta})^n}{\sqrt{n!}} |n\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n e^{-in\theta}}{\sqrt{n!}} |n\rangle, \quad (49)$$

that is, the n -photon component is phase shifted n times more than the one-photon component, keeping the state still a coherent state. (This phase shift is simply the Schrödinger time-evolution factor $\exp(-in\hbar\omega t/\hbar)$, where $n\hbar\omega$ is the energy of the state component). In this case, the "quantum phases" of each state component are related simply to the "classical phase" shift θ . This way, we can say there is "One phase to rule them all," with apologies to J.R.R. Tolkien. Such is not the case for more general (non-coherent) states of the field, where in general the state components $\exp(-i\theta_n)|n\rangle$ can have arbitrary values of their quantum phases θ_n , depending on the means of their generation. This fact,

along with the facts that quantum interference of coherent states mimics perfectly classical interference of fields and that photoelectron statistics (Poisson) are the same as in a semiclassical model of detection [9, 28], are some of the reasons why coherent states are considered to be the “most classical” states possible.

A single-photon state also behaves classically in propagation in the sense that its wavefunction follows classical Maxwell’s equations, as pointed out in the previous section. The quantum nature of this state becomes apparent only when detection occurs; the photon is “found” to be localized (within a cubic wavelength) at only one detector, not simultaneously at two or more. Yet the average rate of detection events upon repeated trials yields results identical to those expected from a classical theory treatment augmented by detection statistics. For this reason, single-photon optics experiments are sometimes thought to be not “classical enough” to bring out uniquely quantum aspects of nature. Correlations of detection events are more revealing, and such experiments require the detection of at least two photons.

Paul Dirac famously tried to sum up the situation by saying, “Each photon then interferes only with itself.” [8]. Later it became clear that nature is not that simple. A state of the field in which two excitations are present, possibly in different modes, is called a “biphoton.” The well-known Hong–Ou–Mandel (HOM) two-photon interference is illustrated in Fig. 4 and is understood by considering two single-photon states impinging on two separate input sides of a 50/50 beam splitter, using $\hat{C}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$,

$$\begin{aligned}
 |\Psi\rangle &= |1\rangle_A|1\rangle_B = \hat{B}^\dagger\hat{A}^\dagger|\text{vac}\rangle = \\
 &= \frac{(\hat{C}^\dagger + \hat{D}^\dagger)}{\sqrt{2}} \frac{(\hat{C}^\dagger - \hat{D}^\dagger)}{\sqrt{2}} |\text{vac}\rangle = \frac{(\hat{C}^\dagger)^2 - (\hat{D}^\dagger)^2}{2} |\text{vac}\rangle = \\
 &= \frac{|2\rangle_C|0\rangle_D - |0\rangle_C|2\rangle_D}{\sqrt{2}}. \tag{50}
 \end{aligned}$$

That is, both photons are detected in either the C mode or in the D mode, with a 50% probability; we never see coincidence events of a detector placed in the outgoing C mode with a detector in the D mode [29]. This phenomenon is a strictly quantum one and does not occur in classical electromagnetic theory. (Although there are classical mimics of this effect, the coincidence probability cannot go to zero in such examples). Therefore, for two-photon interference, we can say, “Each biphoton interferes only with itself.” [30, 31].

The general statement might be phrased best as “Upon detection, each quantum state component interferes only with itself, and only if they occupy the same mode.” One might then wonder, how to understand the common situation that occurs when two fields of different carrier frequencies

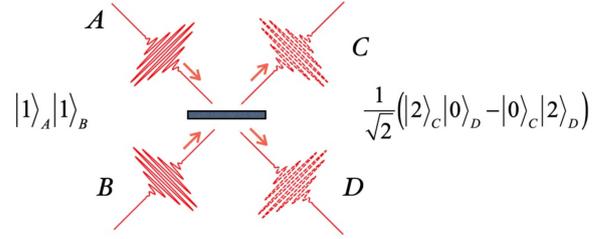


Fig. 4. A separable, nonentangled biphoton state enters the input paths of a 50/50 beam splitter. The result is an entangled biphoton state in which both photons appear in one or the other output path. Note that because single-photon states do not carry phase information *per se*, the introduction of a phase shift in either path before the beam splitter will not affect the two-photon interference outcome.

come together and create intensity “beats” at the difference frequency. It might seem to contradict the idea that modes interfere only with themselves. But it has to be noted that when time-resolved detection is performed in order to observe the beats, a mode projection takes place: the monochromatic modes of distinct frequencies are both projected onto a set of common temporal modes (fields confined within a certain short interval). These temporal modes then interfere with themselves. Such issues were understood as long ago as 1969 [32].

9. Conclusions

In summary: A classical field can be viewed as a physical entity that fills all of space and can transport energy and momentum in the form of continuous wave-like excitations. Similarly, a quantum field can be viewed as a physical entity that fills all of space and can transport energy and momentum only in the form of discrete excitations; these excitations exhibit both wave-like and particle-like behaviors. A “photon” is a label we give to a state representing a single excitation of the EM field. In most cases, it is safer to say “a single-photon state of the field” rather than simply “a photon.”

A convenient way to represent the state of the field is to decompose the field into a weighted sum of mode functions. In classical theory, the weighting coefficients can take on continuous values, representing a continuum of possible energies. In quantum theory, the weighting coefficients are operators representing the quantization (discreteness) of possible energies.

If we are careful, we can describe states in the field picture or in the particle picture, a viewpoint discussed in detail by Iwo Białynicki-Birula. We like to imagine that a given mode is like a container into which we can put any field state. For the special case of the single-photon state, “mode” in the

field picture means the same as “state” in the particle picture. In that case, a single photon distributed coherently between distinct modes represents an entangled state, although such entanglement cannot be verified experimentally by simply detecting the photon in one mode or the other. The modes must interact with other physical entities, such as separate atoms.

When more than one photon is present, we recognize that different modes can be put into many possible combined field states — separable, correlated, or entangled. Biphoton is the name given to a state representing a double excitation of the EM field, whether the excitations are in the same or different modes. Modes of the field are essentially classical constructs and satisfy Maxwell’s equations. As such, they can interfere “classically.” Quantum states can interfere “quantumly,” as in the HOM effect. These two kinds of interference create a rich structure for quantum photonics.

In the context of optics and photonics, the distinction or boundary between quantum and classical is somewhat murky, although useful operational definitions have been developed. Often, we define “classical” to mean that detection statistics can be predicted correctly by a theory in which light is treated as a classical EM wave (although it may be random, stochastic), and the detectors (photo-emissive detectors, photodiode, photomultiplier) are treated by quantum theory. By this “semiclassical” definition, coherent states and any mixture of them, such as thermal (e.g., blackbody) states, are considered classical [9]. Single-photon and biphoton states are then considered quantum. Still, there are cases where coherent states can be used to implement an intrinsically quantum task, such as QKD. There, a highly attenuated laser pulse can be engineered to be in one of several possible weak coherent states, and the quantum behavior occurs upon detection. Any intermediate measurement of the state (such as by an eavesdropper) will necessarily cause a disturbance of the state and thus be detectable. The fundamentally lowest disturbance is dictated by quantum principles related to the Heisenberg uncertainty principle, and thus, even though the state of light is considered “classical,” the security of communication can be assured by the quantum physics of measurement [33]. Harnessing the various degrees of freedom of the optical field, including the temporal-spectral one, can provide novel means to encoding and manipulating quantum information, and therefore is an ongoing topic of research [12].

Acknowledgments

We thank Brian Smith for many discussions over the years about the contents of this paper. And MR thanks Iwo Białynicki-Birula for many helpful discussions about photon wave functions. This work

was supported by the Engineering Research Centers Program of the National Science Foundation under Grant #1941583 to the NSF-ERC Center for Quantum Networks.

References

- [1] I. Białynicki-Birula, *Acta Physica Polonica A* **86**, 97 (1994).
- [2] J.E. Sipe, *Phys. Rev. A* **52**, 1875 (1995).
- [3] S. Wehner, D. Elkouss, R. Hanson, *Science* **362** eaam9288, (2018).
- [4] C. Fabre, N. Treps, *Rev. Mod. Phys.* **92**, 035005 (2020).
- [5] M.G. Raymer, I.A. Walmsley, *Phys. Scr.* **95**, 064002 (2020).
- [6] S.J. Van Enk, *Phys. Rev. A* **72**, 064306 (2005).
- [7] J. Sperling, A. Perez-Leija, K. Busch, C. Silberhorn. *Phys. Rev. A* **100**, 062129 (2019).
- [8] S.M. Barnett, *Phys. Scr.* **97**, 114004 (2022).
- [9] L. Mandel, E. Wolf. *Optical Coherence and Quantum Optics*, Cambridge University Press, Cambridge 1995.
- [10] J.W. Goodman, *Statistical Optics*, John Wiley & Sons, 2015.
- [11] M.G. Raymer, K. Banaszek, *Opt. Express* **28**, 32819 (2020).
- [12] B. Brecht, D.V. Reddy, C. Silberhorn, M.G. Raymer, *Phys. Rev. X* **5**, 041017 (2015).
- [13] R. Loudon, *The Quantum Theory of Light*, Oxford University Press, Oxford 2000.
- [14] K.J. Blow, R. Loudon, S.J.D. Phoenix, T.J. Shepherd, *Phys. Rev. A* **42**, 4102 (1990).
- [15] M.G. Raymer, *J. Mod. Opt.* **67**, 196 (2020).
- [16] B.J. Smith, M.G. Raymer, *New J. Phys.* **9**, 414 (2007).
- [17] P.A.M. Dirac, *The Principles of Quantum Mechanics*, No. 27, Oxford University Press, Oxford 1981.
- [18] A. Hobson, *Am. J. Phys.* **81**, 211 (2013).
- [19] F. Wilczek, *Rev. Mod. Phys.* **71** S83 (1999).
- [20] M.G. Raymer, *Quantum Physics: What Everyone Needs to Know*, Oxford University Press, Oxford 2017.
- [21] C. Spengler, M. Huber, S. Brierley, T. Adaktylos, B.C. Hiesmayr., *Phys. Rev. A* **86**, 022311 (2012).
- [22] QuVis, [Quantum key distribution \(BB84 protocol\) using polarized photons](#).
- [23] M.K. Bhaskar, R. Riedinger, B. Machielse et al., *Nature* **580**, 60 (2020).

- [24] W.A.T. Nogueira, S.P. Walborn, S. Pádua, C.H. Monken *Phys. Rev. Lett.* **92**, 043602 (2004).
- [25] P.L. Saldanha, C.H. Monken, *New J. Phys.* **13**, 073015 (2011).
- [26] O. Keller, *Quantum Theory of Near-Field Electrodynamics*, Springer, Berlin 2011.
- [27] U. Leonhardt, *Essential Quantum Optics: From Quantum Measurements to Black Holes*, Cambridge University Press, Cambridge 2010.
- [28] B. Saleh, *Photoelectron Statistics: With Applications to Spectroscopy and Optical Communication*, Vol. 6. Springer, 2013.
- [29] L. Mandel, in: *More Things in Heaven and Earth*, Springer, New York 1999, p. 460.
- [30] Y.H. Shih, D.V. Strekalov, T.D. Pittman, in: *Causality and Locality in Modern Physics*, Vol. 97, Eds. G. Hunter, S. Jeffers, J.P. Vigiér, Springer, Dordrecht 1998.
- [31] E.J.S. Fonseca, C.H. Monken, S. Pádua *Fortschr. Phys.* **48**, 517 (2000).
- [32] M.C. Teich, *Appl. Phys. Lett.* **14**, 201 (1969).
- [33] C. Silberhorn, T.C. Ralph, N. Lütkenhaus, G. Leuchs, *Phys. Rev. Lett.* **89**, 167901 (2002).

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— ◊ —

Quantum Optics as Applied Quantum Electrodynamics is Back in Town

P. STAMMER^a AND M. LEWENSTEIN^{a,b,*}^a*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, Castelldefels (Barcelona) 08860, Spain*^b*ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S42](https://doi.org/10.12693/APhysPolA.143.S42)*e-mail: maciej.lewenstein@icfo.es

We start this short note by remembering the beginnings of the Warsaw School of Quantum Optics, evidently stimulated by Iwo Białynicki-Birula at the Warsaw University, and then Centre for Theoretical Physics of Polish Academy of Sciences and Adam Kujawski and Zofia Białynicka-Birula at the Institute of Physics of Polish Academy of Sciences. In the theoretical approaches of the Warsaw School, quantum field theory was always present, and quantum optics was considered to be applied quantum electrodynamics. All of us who grew up in this fantastic community have carried and are still carrying the gospel to others. In particular, now quantum electrodynamics began her run on the red carpet of super intense laser–matter interactions, attosecond physics, and ultrafast laser physics in general. We will elaborate on the recent progress in this direction and on the open questions for future investigations. This paper celebrates the 90th birthday of Professor Iwo Białynicki-Birula, our quantum electrodynamics guru!

topics: quantum electrodynamics, strong field physics, quantum state engineering, attoscience

1. Introduction

1.1. Memories

On the occasion like this, it is appropriate to start the paper with some personal memories, in this case by M. Lewenstein: Me and one of my best friends, Marek Kuś, were supposed to do our Diplomas at the Department of Physics of Warsaw University in the academic year 1978–1979. Like many other top theory students, our preference was Katedra Metod Matematycznych Fizyki (KMMF), led by Professor Krzysztof Maurin. I even had a favorite supervisor — Krzysztof Gawędzki. When I asked him about the possibility, he told me literally: “Mr. Maciek, quantum field theory is difficult, and renormalization group even harder,”^{†1} and

he left Poland starting his Odyssey via Harvard, Princeton, Institute of Advanced Scientific Studies (IHÉS), and École Normale Supérieure de Lyon (ENS Lyon). Still, we wanted to go to KMMF, but the Dean of the Department, Professor Jerzy Pniewski, issued a rule that there would be no diplomas in KMMF this year. We had to look for something comparably challenging, and we chose Zakład Teorii Pola i Fizyki Statystycznej of Professor Iwo Białynicki-Birula, the author of the seminal handbook of Quantum Electrodynamics [1]. It was indeed a Mekka of the Warsaw Statistical Physics with Jarosław Piasecki, Łukasz Turski, and Bogdan Cichoński, but we were interested in quantum field theory (QFT). And then came two younger and very convincing guys, Kazimierz[†] Rzażewski and Krzysztof Wódkiewicz, who said: “Let us do quantum optics (QO), which is applied quantum electrodynamics (QED).” And we both got seduced.

^{†1}(Polish translation) “Panie Maćku, kwantowa teoria pola jest trudna, a teoria renormalizacji grup jeszcze trudniejsza”

^{†2}Often called Kazik in the community

Indeed, the training of QO in Warsaw was heavily biased toward QFT. Master equation approaches were not “allowed,” one was using full Hamiltonian and Heisenberg equations. This has taught us very early that there are no Markov processes in Nature — everything must have long-time tail corrections and more...

There is another twist to this story related to strong laser field physics. On the desk of Kazik Rzażewski, I found a preprint of Luiz Davidovich that Kazik got when they shared the same bureau at the International Centre for Theoretical Physics (ITCP) with Luiz. I got absolutely fascinated by Keldysh’s theory of tunnel ionization and decided to work on it. At the beginning of 1970, Pierre Agostini in Saclay published the first result on the so-called above-threshold ionization. Zofia Białynicka-Birula published a seminal paper [2] on the subject in 1984. That was the moment when I decided to join the operation.

The situation of the super-intense laser–matter physics is well described below in Sect. 1.2. We clearly face the situation when QED is on the move again. This paper is based on the thesis proposal of Philipp Stammer, a PhD student at ICFO. So, the plan is to present the motivation to bring quantum optics as applied quantum electrodynamics back to town. This is done by introducing various future investigations, all related to QED of strong laser fields physics, so to the clear heritage of Iwo Białynicki-Birula.

1.2. Quantum optics meets strong laser field physics

For decades the interaction of intense and short laser pulses with matter has been described successfully with semi-classical methods, in which the quantum nature of the electromagnetic field was not taken into account. The characteristics of the observed features in the spectra for the processes of high harmonic generation (HHG) [3, 4] or above threshold ionization (ATI) [5, 6] were well reproduced within the semi-classical picture. Furthermore, the semi-classical approach for the process of HHG (or even fully classical [7]) provides a powerful picture by means of the so-called 3-step model to gain intuition about the electron dynamics. There, (i) an electron tunnel ionizes into the continuum through the barrier formed by the Coulomb potential of the core and the electric field (via dipole coupling), then (ii) the freed electron is driven in the presence of the electric field and can (iii) eventually recombine to the core by emitting the gained energy in terms of radiation. This description has led to fruitful analysis in terms of quantum trajectories [8–10] within the strong field approximation [11]. The progress of strong field and attosecond physics based on the semi-classical description was immense, but neglecting the quantum properties of the field did not allow the use of the language for posing specific questions about the field observables.

However, including the quantum electrodynamical characteristics of the field can lead to new observations in the radiation field inaccessible from the classical perspective, and further allows us to ask questions unamenable before, for instance to obtain insights about the quantum state of the field. In fact, recent theoretical and experimental advances have indicated that intense laser–matter interaction can exhibit non-classical features. In particular, quantum optical approaches for the process of high-order harmonic generation asked for the quantum state of the harmonic field modes [12, 13] and studied the back-action on the fundamental driving field [13, 14]. Furthermore, the experimental advances in combining strong field physics with methods known from quantum optics [15, 16] allowed conceiving new experiments in which non-classical states of light can be generated from the HHG process [13, 14, 17]. This progress has then triggered the subsequent analysis of quantum state engineering of light using intense laser–matter interaction [18–20]. Nevertheless, and despite using Hilbert space constructs for the electromagnetic field, the investigation has not yet revealed inherent quantum signatures in the emitted radiation from the HHG process itself.

Besides these achievements in the quantum optical description of intense laser-driven processes, the full quantum optical properties of the emitted radiation in the process of high harmonic generation have not been revealed yet. The radiation is obtained from classical dipole antenna-like sources and thus exhibits the same characteristics as classical coherent radiation sources. Furthermore, the quantum state of the electromagnetic field is given in terms of product coherent states, which are classical states. Those features originate from the neglected dipole moment correlations in the current theory [13, 18, 19, 21], which, if taken into account, would eventually lead to non-classical contributions in the properties of the emitted harmonic radiation. Thus, further investigation towards accessing this information with potential hidden and interesting properties seems promising for a more detailed understanding of the HHG process and for potential applications in optical technologies. Nevertheless, introducing conditioning measurements on the field after the HHG process leads to the generation of non-classical field states by means of optical Schrödinger cat states with high photon numbers [14, 17–19]. This suggests the potential applicability of these methods in modern optical quantum technologies and could provide a new photonic platform for information processing [22, 23]. In particular, since quantum information processing often requires entangled or superposition states as a resource, there is a clear need to generate such states.

The next section (Sect. 1.3) provides an introduction to the current quantum optical formulation of the process of high harmonic generation.

This serves to define the stage for introducing the current open question within the new formalism. This will then allow us to propose further investigation in this direction. In particular, it highlights the assumptions and approximations used, which are then questioned and analyzed in the proposed future analysis.

1.3. Quantum optical high harmonic generation

In the process of high harmonic generation, coherent radiation of higher-order harmonics of the driving laser frequency is generated [4, 24]. The transfer of coherence and energy from the intense laser source to the harmonic field modes (initially in the vacuum) is achieved by a highly nonlinear interaction of the driving field with the HHG medium, in which the electron is used as an intermediary between the optical modes. Until recently, this was mainly described in semi-classical terms, in which only the electronic degrees of freedom are quantized [4], although there have been early approaches to introduce a fully quantized description of the HHG process [21, 25, 26]. However, recent advances in the quantum optical analysis of HHG have established a new direction in the investigation of strong field physics. This allows us to study the quantum mechanical properties of the harmonic radiation or to take into account the back-action on the driving field [12–20, 27]. In particular, it has been shown that conditioning procedures on processes induced by intense laser–matter interaction can lead to the generation of high-photon number controllable non-classical field states in a broad spectral range [13, 14, 17–19].

What now follows is a brief introduction to the quantum optical description of the process of HHG. We will consider discrete field modes for the sake of simplicity and would like to refer the reader to the full quantum-electrodynamical description, including a continuum of field modes given in [19]. To describe the process of HHG in the single-atom picture (see [21], in which case this is legitimate), we assume that a single active electron is initially in the ground state $|g\rangle$ and is driven by a strong laser field which is described by a coherent state $|\alpha\rangle$ in the fundamental driving mode. The harmonic field modes $q \in \{2, \dots, N\}$ are initially in the vacuum $|\{0_q\}\rangle = \otimes_{q \geq 2} |0_q\rangle$. The interaction Hamiltonian describing the process in the length gauge and within the dipole approximation is given by

$$H_I(t) = -\mathbf{d}(t) \cdot \mathbf{E}_Q(t), \quad (1)$$

where the electric field operator

$$\mathbf{E}_Q(t) = -ig \sum_{q=1}^N \sqrt{q} (b_q^\dagger e^{iq\omega t} - b_q e^{-iq\omega t}) \quad (2)$$

is coupled to the time-dependent dipole moment operator

$$\mathbf{d}(t) = U_{sc}^\dagger(t, t_0) \mathbf{d} U_{sc}(t, t_0). \quad (3)$$

The dipole moment is in the interaction picture of the semi-classical frame $U_{sc}(t, t_0) = \mathcal{T} \exp[-i \int_{t_0}^t d\tau H_{sc}(\tau)]$, with respect to the Hamiltonian of the electron

$$H_{sc}(t) = H_A - \mathbf{d} \cdot \mathbf{E}_{cl}(t). \quad (4)$$

This semi-classical Hamiltonian is the same as traditionally considered in semi-classical HHG theory [4], where $H_A = \frac{1}{2}\mathbf{p}^2 + V(\mathbf{r})$ is the pure electronic Hamiltonian, and

$$\mathbf{E}_{cl}(t) = \text{Tr} \left[\mathbf{E}_Q(t) |\alpha\rangle \langle \alpha| \right] = ig \left(\alpha e^{-i\omega t} - \alpha^* e^{i\omega t} \right) \quad (5)$$

is the classical part of the driving laser field. A detailed derivation of the interaction Hamiltonian $H_I(t)$ can be found in [19]. It now remains to solve the time-dependent Schrödinger equation (TDSE) for the dynamics of the total system of electron and field. Since we are interested in the quantum optical dynamics of the field, and in particular in the process of HHG, we consider the field evolution conditioned on the electronic ground state (this is because the electron returns to the ground state in the HHG process). We thus project the TDSE on $|g\rangle$, and it remains to solve

$$i\partial_t |\Phi(t)\rangle = -\langle g| \mathbf{d}(t) \cdot \mathbf{E}_Q(t) |\Psi(t)\rangle, \quad (6)$$

where $|\Phi(t)\rangle = \langle g| \Psi(t)\rangle$ with the state of the total system $|\Psi(t)\rangle$. Taking into account that the electron is initially in the ground state, it is equivalent to solving for the operator

$$K_{\text{HHG}} = \langle g| \mathcal{T} \exp \left[i \int_{t_0}^t dt' \mathbf{d}(t') \cdot \mathbf{E}_Q(t') \right] |g\rangle, \quad (7)$$

which solely acts on the initial field state $|\Phi_i\rangle = |\alpha\rangle |\{0_q\}\rangle$. This can be solved exactly when neglecting correlations in the dipole moment of the electron [18, 21], such that we can write

$$K_{\text{HHG}} \approx \mathcal{T} \exp \left[i \int_{t_0}^t dt' \langle g| \mathbf{d}(t') |g\rangle \cdot \mathbf{E}_Q(t') \right] = \prod_{q=1}^N e^{i\varphi_q} D(\chi_q), \quad (8)$$

where the shift in each mode is given by the respective Fourier component of the time-dependent dipole moment expectation value

$$\chi_q = -ig \int_{t_0}^t dt' \langle \mathbf{d}(t') \rangle e^{iq\omega t'}. \quad (9)$$

Thus, the solution to (8) is given by a displacement operation acting on the field modes

$$|\Phi\rangle = K_{\text{HHG}} |\Phi_i\rangle = K_{\text{HHG}} |\alpha\rangle \otimes_{q \geq 2} |0_q\rangle = |\alpha + \chi_1\rangle \otimes_{q \geq 2} |\chi_q\rangle. \quad (10)$$

The harmonic modes are described by coherent states due to the fact that the source for the coherent radiation is related to the electron dipole moment expectation value $\langle \mathbf{d}(t) \rangle = \langle g| \mathbf{d}(t) |g\rangle$, which acts as a classical charge current coupled to the

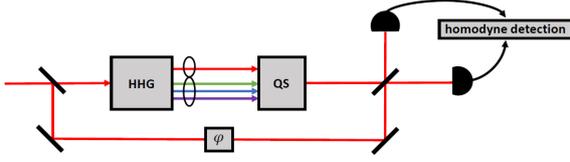


Fig. 1. Schematic illustration of the HHG conditioning experiment performed to generate optical cat states with controllable quantum features. An intense laser field drives the process of HHG, in which an entangled state of the fundamental mode and all harmonics is generated. A conditioning measurement on the harmonic field modes in the quantum spectrometer (QS) leads to a coherent state superposition in the driving field of the form (11), and is measured with a homodyne detection scheme after overlapping with a local oscillator of varying phase delay φ . The reconstructed Wigner functions of the homodyne measurement are shown in Fig. 2.

field operator. It thus only represents the coherent contribution to the harmonic radiation field, and no genuine quantum signature is found. Furthermore, the fact that the final state is a product coherent state over all modes is a consequence of the approximation of neglecting the dipole moment correlations. Otherwise, if going beyond the linear order in $\mathbf{E}_Q(t)$, the field operators for different modes would mix when evaluating the exact propagator in (7) (see Sect. 2.3). Nevertheless, a phenomenological approach to take into account the entanglement between the field modes was performed by the authors in [17, 18].

However, we can employ conditioning schemes on certain field modes, which allows for quantum state engineering of light with non-classical properties [18, 19]. In particular, it has been shown experimentally that a conditioning procedure on the process of HHG can lead to coherent state superposition states (CSS) in the driving laser mode (in the infrared (IR) regime) in close analogy to optical cat states [13, 14]. The experimental configuration is schematically shown in Fig. 1, in which the conditioning on HHG is carried out, and a homodyne detection measurement of the fundamental driving field is performed [13, 19]. To formally describe the generation of these optical CSS via a conditioning operation on the HHG state $|\Phi\rangle = |\alpha + \chi_1\rangle \otimes_{q \geq 2} |\chi_q\rangle$ from (10), M. Lewenstein has recognized that it can be obtained through the projection onto $P = \mathbb{1} - |\alpha\rangle\langle\alpha|$. This projector was phenomenologically introduced in [13] and led to the CSS state

$$|\psi\rangle = |\alpha + \chi_1\rangle - \langle\alpha|\alpha + \chi_1\rangle |\alpha\rangle. \quad (11)$$

Then P. Stammer showed in [17, 18] how this projector follows from a projective conditioning measurement on the harmonic field modes when further taking into account the correlations between the field modes, and also derived the actual measurement

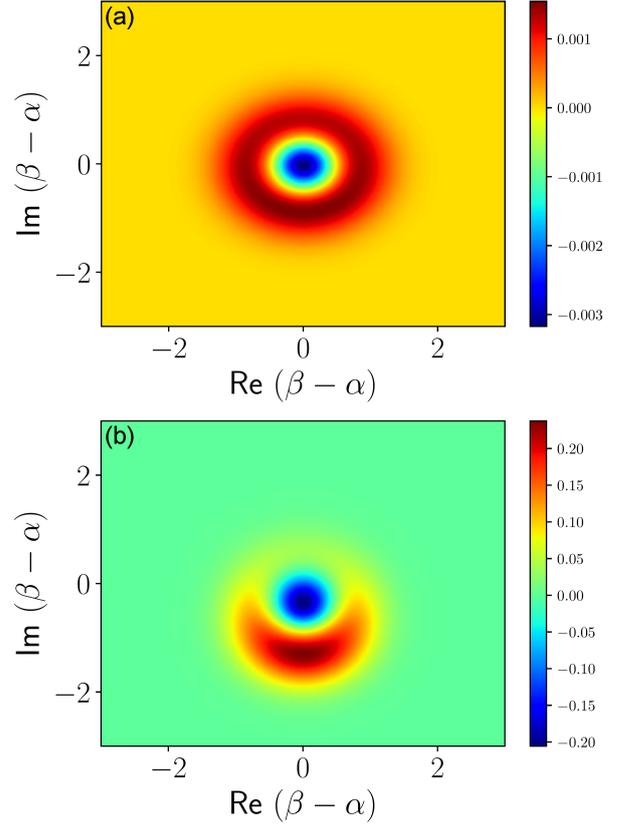


Fig. 2. Wigner function of the coherent state superposition in (11) for different displacement of (a) $\chi_1 = 0.1$, (b) $\chi_1 = 1.0$, which shows features of an optical “kitten” state and a “cat” state, respectively.

operation $M_\alpha^\chi = \mathbb{1} - \exp\left(-\sum_{q \geq 2} |\chi_q|^2\right) |\alpha\rangle\langle\alpha|$, which converges to the projector $M_\alpha^\chi \simeq P = \mathbb{1} - |\alpha\rangle\langle\alpha|$ since $\sum_{q \geq 2} |\chi_q|^2$ is on the order $\mathcal{O}(1/N)$, where N is the harmonic cutoff. The completeness relation of the associated positive operator-valued measure for the measurement operator was shown in [18] within the framework of the quantum theory of measurement. To reconstruct the quantum state of the coherent state superposition in (11), a homodyne detection measurement is performed (see Fig. 1), and the Wigner function of the state is reconstructed. The Wigner function corresponding to the CSS in (11) is shown in Fig. 2 for two different values of the displacement χ_1 . The possibility of experimentally varying the displacement χ_1 , for instance by changing the gas density in the HHG interaction region, allows for a change of the CSS from an optical “kitten” state for small displacement (displaced first Fock state) to an optical “cat” state for larger displacement, as shown in Fig. 2a and 2b, respectively. This allows us to have control over the non-classical properties of the generated CSS in order to generate high-photon number optical cat states from the infrared to the extreme ultraviolet regime [13, 17]. We note that the displacement χ_1

can not be arbitrarily large, since it would destroy the superposition in (11) due to the pre-factor in the second term, which is given by the overlap of the two states in the superposition. However, since α is the initial amplitude of the coherent state, this value has a very high photon number, and thus the optical cat and kitten states can live far away in phase space while the two states in the superposition are not too distinguishable.

2. Open questions about quantum optics of high harmonic generation

In the previous section, we have outlined the current state of the art of our efforts to have a quantum optical description of the process of high harmonic generation. However, there we made assumptions about the experimental boundary conditions and performed approximations by neglecting particular contributions. These need to be tested. Furthermore, the quantum optical description of the light-matter interaction has not yet revealed any genuine quantum mechanical feature in the HHG emission process itself. It turned out that the states of the harmonic field modes $\{q\}$ are described by product coherent states $|\chi_q\rangle$ — which are purely classical. Non-classical signatures, by means of the optical cat state, emerged through the conditioning process. However, we believe that the emitted radiation in the process of HHG contains non-classical signatures once the incoherent contributions from the dipole moment correlations are taken into account, and furthermore, that the field state will be entangled.

In the following, we will outline some open questions in the description of the process of high-order harmonic generation from a quantum optical point of view and provide a motivation why this should be a matter of interest for future investigations.

2.1. On the role of the optical phase in high harmonic generation

To describe the experimental conditions of the HHG experiment, we have assumed that the radiation field which drives the process can be described by a single-mode coherent state $|\alpha\rangle$. This would imply that the source emits continuous coherent laser light in a single-mode with a well-defined phase (coherent in the sense of having non-vanishing off-diagonal density matrix elements in the photon number basis). However, standard HHG experiments are performed using a pulsed source of radiation. On the one hand, this would automatically require a multi-mode description in the frequency domain due to the finite duration of the pulses (they are not just finite but rather super short in the regime of femtoseconds). And thus, we extended the theory to a continuum of modes given in [19]. Furthermore, assuming a pure coherent state description implies that the field has a well-defined

phase and would thus require a phase-stabilized laser system, such that the carrier wave and the envelope of the pulse have a fixed phase relation from shot to shot (carrier-envelope phase (CEP) stabilization [28]). Otherwise, for non-phase-stabilized driving lasers, where the phase varies from shot to shot, one has to average over all possible phases and take into account a proper mixed initial state

$$\rho_{|\alpha|} = \frac{1}{2\pi} \int_0^{2\pi} d\varphi |\alpha e^{i\varphi}\rangle \langle \alpha e^{i\varphi}| = e^{-|\alpha|^2} \sum_n \frac{|\alpha|^{2n}}{n!} |n\rangle \langle n|. \quad (12)$$

In particular, the experiments in [13, 14], which use the process of HHG to generate optical cat states, do not use CEP-stable driving fields. When one analyzes the process of HHG and the conditioning experiment introduced in [13], without the assumption of having a pure coherent initial state within the current quantum optical description, there arise formal difficulties and interpretational inconsistencies with the well-accepted picture of the HHG process.

The difficulty arising in the formal analysis is that the semi-classical frame from the interaction picture of the Hamiltonian $H_I(t)$ (see Sect. 1.3) is not well defined for mixed initial states. Within a fixed semi-classical frame, which is defined via the unitary transformation $D(\alpha)$, we have seen that HHG effectively leads to a shift in the field modes, i.e., $\rho_0 \rightarrow K_{\text{HHG}} \rho_0 K_{\text{HHG}}^\dagger$ (see (10)). However, for the mixed state $\rho_{|\alpha|}$, there is no well-defined semi-classical frame defined through a unique displacement operation $D(\alpha)$. This can also be seen from the fact that the classical part of the driving field vanishes

$$\mathbf{E}_{\text{cl}}(t) = \langle \mathbf{E}_Q(t) \rangle = \text{Tr} [\mathbf{E}_Q(t) \rho_{|\alpha|}] = 0, \quad (13)$$

which implies that there is a vanishing mean electric field amplitude. Hence, this conflicts with the traditionally used powerful picture of HHG in terms of the 3-step model introduced in Sect. 1.2. In this picture, the presence of a non-vanishing electric field amplitude is crucial for describing the tunnel ionization process and the electron dynamics in the continuum driven by the field. The underlying physical property, for the fact that the semi-classical frame is only uniquely defined for a pure coherent initial state $|\alpha\rangle$, is the phase of the field. A coherent state has a well-defined phase, which implies that the semi-classical frame exists via

$$\mathbf{E}_{\text{cl}}(t) = \langle \mathbf{E}_Q(t) \rangle = \text{Tr} [\mathbf{E}_Q(t) |\alpha\rangle \langle \alpha|] = \langle \alpha | \mathbf{E}_Q(t) | \alpha \rangle \propto \sin(\omega t) \quad (14)$$

and the classical picture of an electric field driving the electron process holds. However, it is now natural to ask if the process of high harmonic generation requires non-vanishing field amplitudes, as suggested by the 3-step model, and if harmonics can be generated from driving fields without optical

coherence, such as the phase randomized state in (12), which is diagonal in the photon number basis. Such a state with vanishing off-diagonal density matrix elements in the photon number basis does not exhibit optical coherence, and we thus ask if optical coherence in the driving field is a necessary requirement to generate high-order harmonics. For instance, the electric field expectation value of the mixed state (12) vanishes $\langle \mathbf{E}_Q \rangle = \text{Tr}[\mathbf{E}_Q \rho_{|\alpha\rangle}] = \mathbf{E}_{\text{cl}} = 0$, due to the totally arbitrary phase, and thus there is no well-defined semi-classical frame. This ultimately leads to the question whether processes driven by sufficiently large photon number states $|n\rangle$, which have a completely random phase due to the well-defined photon number, allow for the generation of high-order harmonics. Or, even more general, if incoherent radiation can be used to drive the parametric process of HHG as recently observed for spontaneous parametric down-conversion in [29].

In many optical experiments, the presence of optical coherence is not required to explain the measurement results, and the question of the requirement of optical coherence was first posed in [30]. It is thus natural to ask if the process of HHG requires optical coherence (in the sense of a non-diagonal density matrix in the photon number basis) or if an optical field with a vanishing mean electric field amplitude is sufficient to drive the HHG process. If this is not the case, and we can generate high-order harmonics with incoherent light, how do the harmonic radiation properties differ? And furthermore, how can the powerful picture of the 3-step model be understood for driving fields with vanishing mean-field amplitude? Those questions suggest that there is a need for further theoretical investigation about the role of the optical phase in the HHG process, and furthermore whether the conditioning experiment in [13] is sensitive to the phase of the field or not. From an experimental perspective, we are eager to observe the reconstruction of the Wigner function for CEP-stabilized driving laser fields. From the theoretical point of view, the first question necessary to answer in order to describe the experimental boundary conditions is: *What is the quantum state of an ultrashort few-cycle (CEP-stable) laser pulse?* One way to approach this question could be by following the arguments similar to [31, 32] or [33], just for pulses of radiation with and without CEP-stabilization.

2.2. Theory of quantum optical coherence of high harmonic generation

In the derivation of the field state after the process of HHG, we have thus far always neglected the correlations in the dipole moment of the electron, i.e., approximating (7) with (8). Consequently, we only considered a classical charge (by virtue of the dipole moment expectation value) coupled to the field operator. Therefore, we have only considered

the coherent contribution to the harmonic radiation field. This has the advantage of being exactly solvable. However, as commonly known [34], the incoherent contribution of the emitted radiation can exhibit non-classical signatures and can lead to interesting observations, such as photon antibunching [35]. This incoherent contribution originates from the correlations in the dipole moment. In order to access the full properties of the harmonic radiation, we should not perform the approximation of neglecting the dipole moment correlations. Including those correlations, one can obtain the complete properties of the light field in the process of HHG, which further allows one to obtain a detailed *theory of quantum optical coherence for the process of high harmonic generation*. Furthermore, including those correlations allows asking for the actual quantum state of the field after HHG, going beyond the product coherent states in (10). Taking into account terms beyond linear order in $\mathbf{E}_Q(t)$ would lead to a coupling of different field modes, and thus to entanglement and squeezing.

However, all the previous analysis was performed in the Schrödinger picture (or more precisely, in the interaction picture). However, computing the observables of the field, such as the spectra or two-time correlation functions, and eventually finding non-classical signatures, does not necessarily require the knowledge of the field state after the interaction. That's why we will switch to the Heisenberg picture, making the field operators time-dependent, which allows us to obtain two-time averages including the dipole moment correlations.

We will start with the Hamiltonian of the intense-laser-matter interaction (here in 1D for linear polarization)

$$H = \sum_q \omega_q b_q^\dagger b_q + H_A - d E_Q, \quad (15)$$

where H_A is the atomic Hamiltonian, and the electric field operator is given by $E_Q = -ig \sum_q \sqrt{q}(b_q^\dagger - b_q)$. First, we have to transform the field operator into the Heisenberg picture

$$b_q(t) = b_q e^{-i\omega_q t} + \sqrt{q} g \int_0^t dt' d(t') e^{-i\omega_q(t-t')}. \quad (16)$$

We will then compute the first-order correlation function [34]

$$G(t, t+\tau) = \langle b_q^\dagger(t) b_q(t+\tau) \rangle = q g^2 e^{i\omega_q \tau}$$

$$\times \int_0^t dt_1 e^{-i\omega_q t_1} \int_0^{t+\tau} dt_2 e^{i\omega_q t_2} \langle g | d(t_1) d(t_2) | g \rangle, \quad (17)$$

such that we can use the Wiener-Khinchin theorem [36], stating that the auto-correlation function of a stationary random process and the spectral density of this process are a Fourier-transform pair in the ensemble average, to obtain the power spectrum given by

$$S(\omega) = \frac{1}{\pi} \operatorname{Re} \left[\int_0^\infty d\tau \lim_{t \rightarrow \infty} \langle b_q^\dagger(t) b_q(t+\tau) \rangle e^{i\omega\tau} \right]. \quad (18)$$

It turns out that the power spectral density $S(\omega)$ consists of two terms, the coherent part and an incoherent contribution coming from the dipole moment correlations

$$G^{(1)}(t, t+\tau) = G_{\text{coh}}^{(1)}(t, t+\tau) + q g^2 e^{i\omega_q \tau} \times \int_0^t dt_1 e^{-i\omega_q t_1} \int_0^{t+\tau} dt_2 e^{i\omega_q t_2} \times \int dp \langle g | d(t_1) | p \rangle \langle p | d(t_2) | g \rangle, \quad (19)$$

where the coherent contribution (first term) comes from the dipole moment expectation value. In the stationary limit, this term reads

$$\lim_{t \rightarrow \infty} G_{\text{coh}}^{(1)}(t, t+\tau) = g^2 q |\langle d(\omega_q) \rangle|^2 e^{-i\omega_q \tau}, \quad (20)$$

such that the coherent contribution to the power spectrum is given by

$$S_{\text{coh}}(\omega) = g^2 q |\langle d(\omega_q) \rangle|^2 \delta(\omega - \omega_q). \quad (21)$$

It shows that the HHG spectrum consists of peaks at frequency $\omega_q = q\omega$ (when properly taking into account the finite duration of the driving pulse, the harmonic peaks will have a finite width), with the weight of each harmonic given by the Fourier transform of the time-dependent dipole moment expectation value, and it remains to compute the incoherent contribution. However, it also needs to be carefully analyzed whether the Wiener–Khinchin theorem (WKT) can be used, since it only holds for a stationary random process in the ensemble average (see discussion about time-dependent spectra in [37, 38]). One should also analyze if HHG is an ergodic process, which would then allow one to use the WKT since the ensemble and time average agree for a stationary process, and the autocorrelation function in (18) only depends on the temporal difference (stationarity in the ensemble or temporal average are not sufficient for ergodicity). Furthermore, we then want to compute the second-order correlation function

$$[H_I(t_1), H_I(t_2)] =$$

$$-g^2 \sum_{qp} \sqrt{q_p} \sum_{ijk} |i\rangle \langle j| \left(d_{ik}(t_1) d_{kj}(t_2) - d_{ik}(t_2) d_{kj}(t_1) \right) \left(b_q^\dagger b_p^\dagger e^{i\omega_q t_1} e^{i\omega_p t_2} - b_q^\dagger b_p e^{-i\omega_p t_2} e^{i\omega_q t_1} + \text{h. c.} \right) + g^2 \sum_q q \sum_{ijk} \left(d_{ik}(t_1) d_{kj}(t_2) e^{-i\omega_q(t_1-t_2)} - d_{ik}(t_2) d_{kj}(t_1) e^{i\omega_q(t_1-t_2)} \right) |i\rangle \langle j|, \quad (24)$$

where we have used a discrete basis for the atomic degree of freedom $\mathbb{1} = \sum_i |i\rangle \langle i|$, and introduced the transition dipole matrix elements $d_{ij}(t) =$

$$g^{(2)}(\tau) = \lim_{t \rightarrow \infty} \frac{\langle b_q^\dagger(t) b_q^\dagger(t+\tau) b_q(t+\tau) b_q(t) \rangle}{\langle b_q^\dagger(t) b_q(t) \rangle \langle b_q^\dagger(t+\tau) b_q(t+\tau) \rangle}, \quad (22)$$

since this would provide insights into possible antibunching signatures, i.e., $g^{(2)}(0) < g^{(2)}(\tau)$. However, we imagine that the coherent contribution dominates the incoherent contribution, and one needs to conceive clever experiments to either separate the two processes for individual harmonics or to find the conditions in which the two contributions are on the same order of magnitude. This could eventually be realized with a two-color driving field (ω and its second harmonic 2ω), which leads to the appearance of even harmonics in the spectrum. By varying the phase between the two driving fields, the amplitude of the even harmonics can be altered, such that there might be a regime in which the coherent and incoherent contributions can compete.

2.3. Entanglement and squeezing in high harmonic generation

Thus far, we found that the field state of the harmonic modes is given by product coherent states of all field modes (10). This is a consequence of the approximation performed in (8) (neglecting the dipole moment correlations), which effectively leads to a linear expression in the field operators $b_q^{(\dagger)}$. While the commutator of the exact interaction Hamiltonian $H_I(t) = -d(t)E_Q(t)$ at different times is an operator in the total Hilbert space of atom plus field,

$$[H_I(t_1), H_I(t_2)] \in \mathcal{H}_A \otimes \mathcal{H}_F. \quad (23)$$

The approximate interaction Hamiltonian $H_I^{app}(t) = -\langle d(t) \rangle E_Q(t)$ is just a complex number, i.e., $[H_I^{app}(t_1), H_I^{app}(t_2)] \in \mathbb{C}$, and thus when solving (8), the modes do not mix. Going beyond the linear term of the field operator $E_Q(t)$ would lead, for instance, to squeezing in the field modes. Furthermore, all field modes will become entangled due to the mixing of the field operators $b_q^{(\dagger)}$ of the different modes. We can thus start to evaluate the commutator of the exact interaction Hamiltonian at different times, yielding

$\langle i | d(t) | j \rangle$. Note that for the approximation of neglecting the dipole moment correlations and taking the expectation value in the electronic ground

state leads to $\sum_{ijk} d_{ik}(t_1)d_{kj}(t_2)\langle g|i\rangle\langle j|g\rangle \simeq \langle d(t_1)\rangle\langle d(t_2)\rangle$, and thus the first line in (24) vanishes (where the squeezing and mixing of modes would come from), and the second line reduces to what one would get from $[H_I^{app}(t_1), H_I^{app}(t_2)]$. However, for the exact interaction Hamiltonian $H_I(t) = -d(t)E_Q(t)$, we observe that the different field modes mix, which would lead to squeezing and entanglement. One could, for instance, already observe the first signatures of such non-classical states due to the higher-order terms of $E_Q(t)$ when taking into account up to the quadratic order in the coupling $g \propto \sqrt{\omega/V_{eff}}$ with the quantization volume V_{eff} . Thus, when solving (7) by using Baker–Campbell–Hausdorff for infinitesimal time steps, one obtains an approximate solution up to quadratic order in g when only taking into account $[H_I(t_1), H_I(t_2)] \propto g^2$, and the time-dependent transition dipole matrix elements $d_{ij}(t)$ can be computed within the strong field approximation [11].

3. Conclusions

Motivated by recent studies on the quantum optical description of the process of high harmonic generation from intense-laser-field-driven atoms, we identified current challenges and how this can lead to future investigations. With the proposed studies, we anticipate that more complete insights into the process of HHG will be obtained, and that the full characteristics of the radiation field will be found. The current quantum optical framework treats the source of the scattered field as a classical charge current, similar to a dipole antenna, and thus only the coherent contribution is obtained through the dipole moment expectation value. Thus, the radiation properties, as well as the final field state, do not indicate genuine quantum signatures in the HHG process. Only via conditioning experiments, through a post-selection procedure, we obtained non-classical signatures in the reconstructed Wigner function. It would thus be of great interest to see if, already at the level of the HHG process itself, without conditioning, non-classical observations can be obtained in the radiation properties of the scattered field. In addition to the proposed approaches present in this manuscript, there exist further efforts in this direction. For instance, there are the following options to achieve such situations:

- So far, we have considered high-order harmonics generated in atomic systems. Alternatively, one can consider HHG from solid-state targets. Even in the case of “trivial” solid-state systems, such as electrons in the Wannier–Bloch picture [39], one can obtain electron-field entanglement [40] since the electron can transition on one site in the lattice, but might recombine in another side. A similar mechanism, of semiconductors driven by

strong coherent radiation, is studied in a recent paper [41], where the potential for generating non-classical light fields is discussed.

- Another option, besides driving HHG in simple uncorrelated solid-state targets, is to look for HHG in laser-driven strongly correlated materials, such as high-temperature superconductors [42]. For a simple, yet pedagogical, model of such a mechanism, see [43, 44].
- Finally, one can use non-classical, for instance, squeezed light to drive the HHG process in atoms, which leads to its fingerprints in the field observable, such as the HHG spectra [45]. Which, however, also do not depict non-classical signatures in the harmonic radiation based on this observable.

Acknowledgments

ICFO group acknowledges support from: ERC AdG NOQIA; Ministerio de Ciencia y Innovation Agencia Estatal de Investigaciones (PGC2018-097027-B-I00/10.13039/501100011033, CEX2019-000910-S/10.13039/501100011033, Plan National FIDEUA PID2019-106901GB-I00, FPI, QUANTERA MAQS PCI2019-111828-2, QUANTERA DYNAMITE PCI2022-132919, Proyectos de I+D+I “Retos Colaboración” QUSPIN RTC2019-007196-7); MICIIN with funding from European Union NextGenerationEU(PRTR-C17.I1) and by Generalitat de Catalunya; Fundació Cellex; Fundació Mir-Puig; Generalitat de Catalunya (European Social Fund FEDER and CERCA program, AGAUR Grant No. 2021 SGR 01452, QuantumCAT U16-011424, co-funded by ERDF Operational Program of Catalonia 2014-2020); Barcelona Supercomputing Center MareNostrum (FI-2022-1-0042); EU (PASQuanS2.1, 101113690); EU Horizon 2020 FET-OPEN OPTologic (Grant No 899794); EU Horizon Europe Program (Grant Agreement 101080086 — NeQST), National Science Centre, Poland (Symfonia Grant No. 2016/20/W/ST4/00314); ICFO Internal “QuantumGaudi” project; European Union’s Horizon 2020 research and innovation program under the Marie-Skłodowska-Curie grant agreement No 101029393 (STREDCH) and No 847648 (“La Caixa” Junior Leaders fellowships ID100010434: LCF/BQ/PI19/11690013, LCF/BQ/PI20/11760031, LCF/BQ/PR20/11770012, LCF/BQ/PR21/11840013). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union, European Commission, European Climate, Infrastructure and Environment Executive Agency (CINEA), nor any other granting authority. Neither the European Union nor any granting authority can be held responsible for them.

P.S. acknowledges funding from The European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 847517.

References

- [1] I. Białynicki-Birula, Zofia Białynicka-Birula, *Quantum electrodynamics*, Vol. 70, Elsevier, 2013.
- [2] Z. Białynicka-Birula, *J. Phys. B: At. Mol. Phys.* **17**, 3091 (1984).
- [3] M. Ferray, A. L'Huillier, X.F. Li, L.A. Lompre, G. Mainfray, C. Manus, *J. Phys. B: At. Mol. Opt. Phys.* **21**, L31 (1988).
- [4] M. Lewenstein, P. Balcou, M.Y. Ivanov, A. L'huillier, P.B. Corkum, *Phys. Rev. A* **49**, 2117 (1994).
- [5] P. Agostini, F. Fabre, G. Mainfray, G. Petite, N. Ko Rahman, *Phys. Rev. Lett.* **42**, 1127 (1979).
- [6] M. Lewenstein, K.C. Kulander, K.J. Schafer, P.H. Bucksbaum, *Phys. Rev. A* **51**, 1495 (1995).
- [7] P.B. Corkum, *Phys. Rev. Lett.* **71(13)**, 1994 (1993).
- [8] P. Salières, B Carré, L. Le Déroff et al., *Science* **292**, 902 (2001).
- [9] M.Y. Ivanov, M. Spanner, O. Smirnova, *J. Mod. Opt.* **52**, 165 (2005).
- [10] O. Smirnova, M. Ivanov, [arXiv:1304.2413](https://arxiv.org/abs/1304.2413), 2013.
- [11] K. Amini, J. Biegert, F. Calegari et al., *Rep. Prog. Phys.* **82**, 116001 (2019).
- [12] A. Gorlach, O. Neufeld, N. Rivera, O. Cohen, I. Kammer, *Nat. Commun.* **11**, 1 (2020).
- [13] M. Lewenstein, M.F. Ciappina, E. Pisanty, J. Rivera-Dean, P. Stammer, T. Lamprou, P. Tzallas, *Nature Physics* **17**, 1104 (2021).
- [14] J. Rivera-Dean, T. Lamprou, E. Pisanty, P. Stammer, A.F. Ordóñez, A.S. Maxwell, M.F. Ciappina, M. Lewenstein, P. Tzallas, *Phys. Rev. A* **105**, 033714 (2022).
- [15] I.A. Gonoskov, N. Tsatrafyllis, I.K. Kominis, P. Tzallas, *Sci. Rep.* **6**, 1 (2016).
- [16] N. Tsatrafyllis, I.K. Kominis, I.A. Gonoskov, P. Tzallas, *Nat. Commun.* **8**, 1 (2017).
- [17] P. Stammer, J. Rivera-Dean, T. Lamprou, E. Pisanty, M.F. Ciappina, P. Tzallas, M. Lewenstein, *Phys. Rev. Lett.* **128**, 123603 (2022).
- [18] P. Stammer, *Phys. Rev. A* **106**, L050402 (2022).
- [19] P. Stammer, J. Rivera-Dean, A. Maxwell, T. Lamprou, A. Ordóñez, M.F. Ciappina, P. Tzallas, M. Lewenstein, *PRX Quantum* **4**, 010201 (2023).
- [20] J. Rivera-Dean, P. Stammer, A.S. Maxwell, T. Lamprou, P. Tzallas, M. Lewenstein, M.F. Ciappina, *Phys. Rev. A* **106**, 063705 (2022).
- [21] B. Sundaram, P.W. Milonni, *Phys. Rev. A* **41**, 6571 (1990).
- [22] M. Lewenstein, N. Baldelli, U. Bhattacharya et al., [arXiv:2208.14769](https://arxiv.org/abs/2208.14769), 2022.
- [23] U. Bhattacharya, T. Lamprou, A.S. Maxwell, A.F. Ordóñez, E. Pisanty, J. Rivera-Dean, P. Stammer, M.F. Ciappina, M. Lewenstein, P. Tzallas, [arXiv:2302.04692](https://arxiv.org/abs/2302.04692), 2023.
- [24] P. Salieres, A. L'huillier, P. Antoine, M. Lewenstein, [arXiv:quant-ph/9710060](https://arxiv.org/abs/quant-ph/9710060), 1997.
- [25] J.H. Eberly, M.V. Fedorov, *Phys. Rev. A* **45**, 4706 (1992).
- [26] D.J. Diestler, *Phys. Rev. A* **78**, 033814 (2008).
- [27] I.K. Kominis, G. Kolliopoulos, D. Charalambidis, P. Tzallas, *Phys. Rev. A* **89**, 063827 (2014).
- [28] F. Krausz, M. Ivanov, *Rev. Mod. Phys.* **81**, 163 (2009).
- [29] C. Li, B. Braverman, G. Kulkarni, R.W. Boyd, *Phys. Rev. A* **107**, L041701 (2023).
- [30] K. Mølmer, *Phys. Rev. A* **55**, 3195 (1997).
- [31] S.J. Van Enk, C.A. Fuchs, *Phys. Rev. Lett.* **88**, 027902 (2001).
- [32] S.J. van Enk, C.A. Fuchs, [arXiv:quant-ph/0111157](https://arxiv.org/abs/quant-ph/0111157), 2001.
- [33] S.D. Bartlett, T. Rudolph, R.W. Spekkens, *Int. J. Quant. Inf.* **4**, 17 (2006).
- [34] H. Carmichael, *Statistical Methods in Quantum Optics 1: Master Equations and Fokker-Planck Equations*, Vol. 1, Springer Science & Business Media, 1999.
- [35] H.J. Kimble, M. Dagenais, L. Mandel, *Phys. Rev. Lett.* **39**, 691 (1977).
- [36] L. Mandel, E. Wolf, *Optical Coherence and Quantum Optics*, Cambridge University press, 1995.
- [37] J.H. Eberly K. Wodkiewicz, *J. Opt. Soc. Am.* **67**, 1252 (1977).
- [38] J.H. Eberly, C.V. Kunasz, K. Wodkiewicz, *J. Phys. B: At. Mol. Phys.* **13**, 217 (1980).
- [39] E.N. Osika, A. Chacón, L. Ortmann, N. Suárez, J.A. Pérez-Hernández, B. Szafran, M.F. Ciappina, F. Sols, A.S. Landsman, M. Lewenstein, *Phys. Rev. X* **7**, 021017 (2017).

- [40] J. Rivera-Dean, P. Stammer, A.S. Maxwell, T. Lamprou, A.F. Ordóñez, E. Pisanty, P. Tzallas, M. Lewenstein, M.F. Ciappina, [arXiv:2211.00033](https://arxiv.org/abs/2211.00033), 2023.
- [41] I. Gonoskov, R. Sondenheimer, C. Hünecke, D. Kartashov, U. Peschel, S. Gräfe, [arXiv:2211.06177](https://arxiv.org/abs/2211.06177), 2022.
- [42] J. Alcalá, U. Bhattacharya, J. Biegert et al., *Proc. Natl. Acad. Sci.* **119**, e2207766119 (2022).
- [43] A. Pizzi, A. Gorlach, N. Rivera, A. Nunnenkamp, I. Kaminer, *Nat. Phys.* **19**, 551 (2023).
- [44] P. Tzallas, *Nat. Phys.* **19**, 472 (2023).
- [45] A. Gorlach, M. Even Tzur, M. Birk, M. Krüger, N. Rivera, O. Cohen, I. Kaminer, [arXiv:2211.03188](https://arxiv.org/abs/2211.03188), 2022.

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

The Wave Functional of the Vacuum in a Resonator

A. FRIEDRICH^{a,*}, D. MOLL^a, M. FREYBERGER^a,
L. PLIMAK^a AND W.P. SCHLEICH^{a,b,c}

^a*Institut für Quantenphysik and Center for Integrated Quantum Science and Technology (IQST), Universität Ulm, Albert-Einstein-Allee 11, D-89069 Ulm, Germany*

^b*Texas A&M AgriLife Research, Texas A&M University, College Station, Texas 77843-4242, USA*

^c*Hagler Institute for Advanced Study and Department of Physics and Astronomy, Institute for Quantum Science and Engineering (IQSE), Texas A&M University, College Station, Texas 77843-4242, USA*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S52](https://doi.org/10.12693/APhysPolA.143.S52)

*e-mail: alexander.friedrich@uni-ulm.de

We show that despite fundamentally different situations, the wave functional of the vacuum in a resonator is identical to that of free space. The infinite product of the Gaussian ground state wave functions defining the wave functional of the vacuum translates into an exponential of a sum rather than an integral over the squares of mode amplitudes weighted by the mode volume and power of the mode wave number. We express this sum by an integral of a bilinear form of the field containing a kernel given by a function of the square root of the negative Laplacian acting on a transverse delta function. For transverse fields, it suffices to employ the familiar delta function, which allows us to obtain explicit expressions for the kernels of the vector potential, the electric field, and the magnetic induction. We show for the example of the vector potential that different mode expansions lead to different kernels. Lastly, we show that the kernels have a close relationship with the Wightman correlation functions of the fields.

topics: wave functional, vacuum, Wightman tensor, cavity quantum electrodynamics

1. Introduction

The standard approach [1, 2] towards quantization of the electromagnetic field is straightforward: decomposition of the field into modes and quantization of the resulting harmonic oscillator amplitudes by canonical commutation relations. The wave functional of the vacuum proposed by John Archibald Wheeler [3–5] and extended [6–10] and refined by Iwo Białynicki-Birula does not rely on a mode expansion but involves the complete electromagnetic field. The essence of the wave functional is best summarized by the following quote from Białynicki-Birula’s article [7] employing the wave functional to obtain the Wigner phase space distribution of the whole electromagnetic field:

“The whole electromagnetic field is treated as one huge, infinitely dimensional harmonic oscillator. The

wave function and the corresponding Wigner function become then functionals of the field variables.”

Recent impressive progress in cavity and circuit quantum electrodynamics invites us to reconsider the wave functional of the vacuum in the case of a resonator. Indeed, so far, investigations have concentrated exclusively on free space. In the present article, we show that the expressions for the wave functional of the vacuum in the two situations are identical.

1.1. The cradle of the quantum theory of fields

The year 1925 marks not only the birth of modern quantum mechanics, but is also arguably the beginning of quantum electrodynamics (QED). Indeed, the “Drei-Männer-Arbeit” [11] not only provided the foundations of matrix mechanics, but also presented

the quantization of a free electromagnetic field for the first time. This was extended only two years later to include the interaction with quantized matter [12].

The discovery of the Lamb shift [13] and the anomalous magnetic moment [14] in 1947 demonstrated that the theory, so far plagued by infinities, contained some truth. The renormalization theory [15, 16], developed shortly after, removed these infinities and gave rise to the field of QED, a theory [17] with an unprecedented agreement with experiment.

Almost 40 years later, new experimental manifestations of QED emerged from the use of high-Q microwave cavities [18, 19] and the interaction of individual atoms with single modes of the radiation field. Whereas in the first era of cavity QED, the experiments were only in the microwave domain, the optical domain soon followed. The last 20 years have seen the development of a new rapidly moving branch of quantum optics summarized by circuit QED [20] and, recently, waveguide QED [21].

Ever since the proposal of quantized electrodynamics, there has been a constant drive toward a deeper understanding of the associated vacuum fluctuations and the measurability of the field components. For example, Lev Davidovich Landau and Rudolf Peierls [22] applied the uncertainty principle to relativistic quantum theory and concluded:

“The assumptions of wave mechanics which have been shown to be necessary in section 2 are therefore not fulfilled in the relativistic range and the application of wave mechanics methods to this range goes beyond their scope. It is therefore not surprising that the formalism leads to various infinities; it would be surprising if the formalism bore any resemblance to reality.”

Needless to say, this grim outlook was not shared by Niels Bohr, who, together with Léon Rosenfeld, immediately started to correct this article. However, it took them two years to achieve this goal for the case of free fields [23], and they stated:

“Not only is it an essential complication of the problem of field measurements that, when comparing field averages over different space-time regions, we cannot in an unambiguous way speak about a temporal sequence of the measurement processes.”

After the discovery of renormalization Bohr and Rosenfeld returned [24] to this problem and included charges. For an interesting commentary by Rosenfeld providing the historical context of both articles we refer to [25].

The analogous question of the measurability of the gravitational field, pioneered by Helmut Salecker and Eugene Paul Wigner [26], led to Wheeler’s *Geometrodynamics* [3] and the quantum fluctuations of gravity and the quantum foam. It

was in this context that he proposed to consider the wave functional [4, 5] of electromagnetism as a guide to linearized gravity. Armed with the insights from electromagnetism, he was able to derive an estimate for the fluctuations of the space-time geometry at distances of the Planck length. For a detailed discussion of the wave functional of linearized gravity, we refer to the classic paper by Karel Kuchař [27].

Similarly, but on more general grounds, Julian Schwinger investigated the effect of the so-called *fluctuating sources* (i.e., transient fields) in quantum field theories [28]. Some of these ideas [29, 30] eventually found their way into the framework, which later became effective (quantum) field theory.

Recent years have seen a renaissance of the wave functional of the vacuum. It now appears not only in the Schrödinger representation of quantum field theory [31, 32] but also in possible realizations [33, 34] of the *Gedanken Experiment* of Richard P. Feynman [35] addressing the question of measurability [36] of entanglement between two quantum systems due to gravity which has recently attracted significant attention. This field has become quite an active area of research, due to the emerging technical possibility of preparing almost macroscopic systems in motional quantum states, and also because direct tests of the quantum nature of gravity via the detection of gravitons seem highly unlikely, as suggested [37] by yet another founding father of QED, Freeman Dyson.

For this reason, we find it appropriate to revisit the wave functional of the vacuum and analyze it for the case of a resonator. This situation is not only timely, but the set of discrete modes makes the derivation much cleaner. On the other hand, the discreteness adds a different complication arising from the sum over the modal indices, confirming the well-known adage: “There ain’t no such thing as a free lunch.”

1.2. Road to the wave functional

We now summarize our path to the wave functional of the vacuum in a resonator using the example of the electric field representation. In Fig. 1, we start from the decomposition of the electric field $\mathbf{E} \equiv \mathbf{E}(t, \mathbf{r})$ (left lower corner) into a discrete set of modes \mathbf{u}_ℓ . Here, the subscript ℓ combines the polarization index as well as the indices characterizing the wave vector \mathbf{k}_ℓ enforced by the boundary conditions on the Helmholtz equation by the shape of the resonator.

The subsequent quantization of the corresponding electric field amplitudes $E_\ell \equiv \mathcal{E}_\ell p_\ell$ using the canonical commutation relations leads us to the eigenvalue equation of the electric field operator \hat{E}_ℓ in mode \mathbf{u}_ℓ . Together with the definition of the ground state $|0_\ell\rangle$ of the ℓ -th mode in terms of the annihilation operator \hat{a}_ℓ , we find the Gaussian wave function $\psi_\ell(E_\ell) \equiv \langle E_\ell | 0_\ell \rangle$ in the electric field representation.

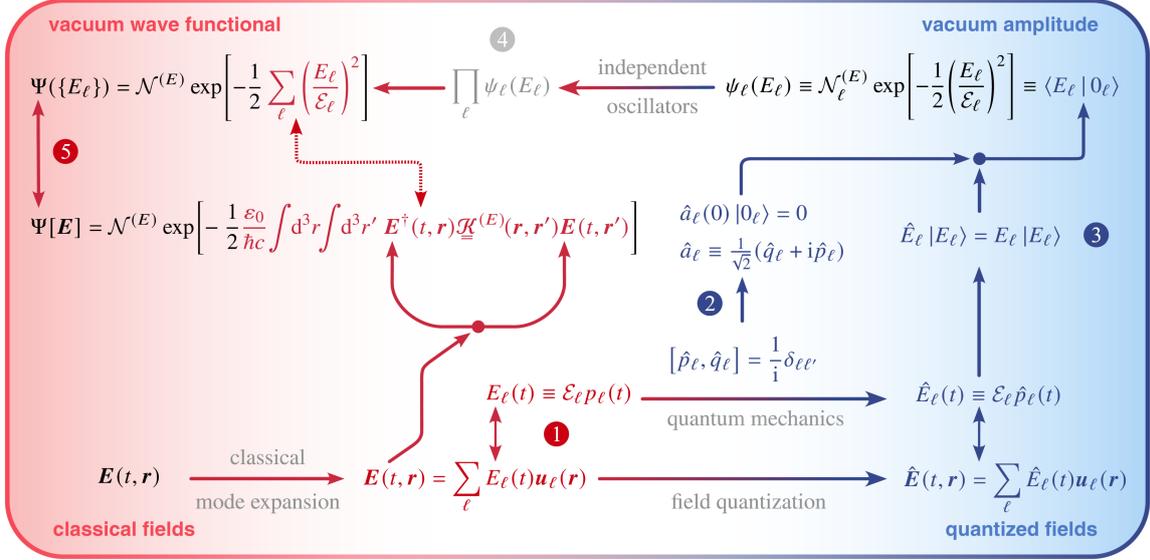


Fig. 1. Road to the wave functional $\Psi[\mathbf{E}]$ of the vacuum in a resonator. We identify five different ingredients marked by numbers: ① mode expansion of the electric field $\mathbf{E} = \mathbf{E}(t, \mathbf{r})$, ② quantization of the field E_{ℓ} in each mode \mathbf{u}_{ℓ} according to the canonical commutation relations, ③ definition of the electric field eigenstates $|E_{\ell}\rangle$ and the ground state $|0_{\ell}\rangle$ of the ℓ -th mode, ④ infinite product of all ground state wave functions $\psi_{\ell}(E_{\ell})$, and ⑤ wave functional $\Psi[\mathbf{E}]$ of the vacuum after the elimination of the mode decomposition. The last ingredient, i.e., the connection between the discrete sum over the modes and the space integrals of the bilinear form of the electric field \mathbf{E} and a kernel $\underline{\mathcal{K}}^{(E)}$, indicated by a dashed line, constitutes the topic of our article.

In the absence of matter and interactions, the modes are independent of each other and correspond to a product state with all modes in the ground state. Hence, we arrive at an *infinite product* of Gaussian wave functions. Due to the functional equation of the exponential function, this product reduces to a *single* exponential of an *infinite sum* over the squares of the scaled fields $E_{\ell}/\mathcal{E}_{\ell}$ in the modes.

The mode expansion we started with shows that this sum is identical to an appropriate space integral consisting of a bilinear form of the electric field and a kernel. In this way, we have eliminated the mode decomposition and have arrived at an expression containing the quantum mechanics of the vacuum as well as the complete electric field distribution $\mathbf{E} = \mathbf{E}(t, \mathbf{r})$ without resorting to modes.

We conclude this section by briefly addressing the differences and detours enforced by free space due to the continuous superposition of plane wave modes.

In the case of free space, the continuous superposition of the plane waves, rather than the discrete set of modes, involves an *integration* over the wave vector rather than a *summation* over mode indices ℓ . The quantization of the field is identical to that in a resonator, with the exception of the commutation relation where the Kronecker delta in ℓ and ℓ' is replaced by the Dirac delta function in the difference of the wave vectors \mathbf{k} and \mathbf{k}' .

However, the *infinite continuous* product of the ground state wave functions now requires either a discretization of the continuum of the wave vectors

or a more sophisticated technique. Once the functional equation of the exponential function has transformed the *infinite product* into an *infinite sum*, we can continue with the integral, which is a continuous superposition. In free space as well as in the resonator, we arrive at the same expression for the wave functional $\Psi[\mathbf{E}]$ of the vacuum.

1.3. In a nutshell

Before we dive into the mathematics, we motivate our results without detailed derivations and summarize them in Tables I and II. We start our discussion by recalling in Table I the essential ingredients of the expansion of a vector field into modes.

Throughout the article, we focus on an expansion of the vector potential \mathbf{A} , the electric field \mathbf{E} , and the magnetic induction \mathbf{B} into a set of discrete mode functions.

Whereas the decomposition of \mathbf{A} and \mathbf{E} involves the mode functions \mathbf{u}_{ℓ} , the one of \mathbf{B} brings in the curl of \mathbf{u}_{ℓ} due to the fact that there are no magnetic monopoles. In order to make the curl of \mathbf{u}_{ℓ} dimensionless, we have introduced the inverse of the wave number k_{ℓ} .

The field strengths A_{ℓ} , E_{ℓ} , and B_{ℓ} of \mathbf{A} , \mathbf{E} , and \mathbf{B} in the mode ℓ are determined by the products of the corresponding vacuum fields \mathcal{A}_{ℓ} , \mathcal{E}_{ℓ} , and \mathcal{B}_{ℓ} , and a dimensionless amplitude. In the case of A_{ℓ} and B_{ℓ} , this amplitude is given by q_{ℓ} , whereas for E_{ℓ} it is p_{ℓ} . They are analogs of the familiar coordinate and momentum variables of a harmonic oscillator.

TABLE I

Elements of the expansion of a vector field \mathbf{F} , such as the vector potential \mathbf{A} , the electric field \mathbf{E} , and the magnetic induction \mathbf{B} into a discrete set of mode functions \mathbf{u}_ℓ of the vector potential. Here, ℓ denotes the mode index consisting of the polarization and three integers characterizing the wave vector \mathbf{k}_ℓ determined by the boundary conditions of the Helmholtz equation imposed by the shape of the resonator. The field strengths A_ℓ , E_ℓ , and B_ℓ in the ℓ -th mode are given by the products $\mathcal{A}_\ell q_\ell$, $\mathcal{E}_\ell p_\ell$, and $\mathcal{B}_\ell q_\ell$ of the vacuum field strengths \mathcal{A}_ℓ , \mathcal{E}_ℓ , and \mathcal{B}_ℓ as well as the dimensionless quadrature variables q_ℓ and p_ℓ of a harmonic mode oscillator of frequency ω_ℓ . Here we have also included the mode expansions in terms of the eigenmodes of the individual fields defined by the solution of the Helmholtz equation for each field. These are the eigenmodes \mathbf{u}_ℓ of the vector potential \mathbf{A} , the eigenmodes $\boldsymbol{\nu}_\ell$ of the electric field, and the eigenmodes \mathbf{w}_ℓ of the magnetic induction.

Field \mathbf{F}	Eigenmodes $\{\mathbf{f}_\ell\}$	Eigenmode expansion	$\{\mathbf{u}_\ell\}$ -mode expansion	Mode field strength	Vacuum field
\mathbf{A}	$\{\mathbf{u}_\ell\}$	$\sum_\ell A_\ell \mathbf{u}_\ell$	$\sum_\ell A_\ell \mathbf{u}_\ell$	$A_\ell = \mathcal{A}_\ell q_\ell$	$\mathcal{A}_\ell = \sqrt{\frac{\hbar}{\varepsilon_0 \omega_\ell \mathcal{V}_\ell}}$
\mathbf{E}	$\{\boldsymbol{\nu}_\ell\}$	$\sum_\ell E_\ell \boldsymbol{\nu}_\ell$	$\sum_\ell E_\ell \mathbf{u}_\ell$	$E_\ell = \mathcal{E}_\ell p_\ell$	$\mathcal{E}_\ell = \mathcal{A}_\ell \omega_\ell$
\mathbf{B}	$\{\mathbf{w}_\ell\}$	$\sum_\ell B_\ell \mathbf{w}_\ell$	$\sum_\ell B_\ell k_\ell^{-1} (\nabla \times \mathbf{u}_\ell)$	$B_\ell = \mathcal{B}_\ell q_\ell$	$\mathcal{B}_\ell = \frac{1}{c} \mathcal{E}_\ell$

We recall from the Maxwell equations that in the Coulomb gauge, the electric field is determined by the time derivative of the vector potential. As a result, the vacuum electric field \mathcal{E}_ℓ differs from that of the vector potential \mathcal{A}_ℓ by the frequency ω_ℓ of the mode.

In general, the ratio of the magnetic induction to the electric field is governed by the speed of light c . This property also holds true for the corresponding vacuum fields. Thus, the ratio between the magnetic induction and the vector potential is given by the wave number k_ℓ due to the dispersion relation $k_\ell \equiv \omega_\ell/c$ of light.

This difference in the wave number dependence of the vacuum fields has important implications when we now make the transition to quantum mechanics and motivate the wave functional of the vacuum in a resonator. We summarize our path to this expression in Table II.

We start by recalling that the ground state wave function ψ_ℓ of a single mode is determined by a Gaussian. Since its argument \mathbf{f}_ℓ has to be dimensionless, it must involve the ratio of the field strength F_ℓ divided by the associated vacuum field \mathcal{F}_ℓ .

The wave function of the complete electromagnetic field describing a quantum state with every mode in the ground state is defined by the infinite product of the corresponding single mode wave functions. Due to the functional equation of the exponential function, this product of exponentials reduces to a single exponential whose argument is determined by the sum of the arguments of the individual exponentials. Hence, we arrive at a sum of the squares of the dimensionless variables \mathbf{f}_ℓ over all modes.

When we recall from Table I the definitions of these vacuum fields, we obtain for \mathbf{f}_ℓ^2 the product of the parameter $\beta^{(F)}$, determined by fundamental

constants such as the dielectric constant ε_0 , reduced Planck's constant \hbar , speed of light c , and resonator specific parameters such as the square of the field strengths F_ℓ^2 , the mode volume \mathcal{V}_ℓ and the wave number k_ℓ , or its inverse.

Since F_ℓ^2 emerges in this sum, it is tempting to replace it with an integral of a bilinear form of the complete field. Indeed, this sum over modes is reminiscent of the energy of the electromagnetic field in a resonator. However, in contrast to the present discussion, where the sums involve either the mode wave number or its inverse, the expression for the energy contains the square of it.

It is at this point that the difference in the descriptions of the electromagnetic field in terms of a continuous or a discrete superposition of modes enters the stage. This subtle point originates from the definition of the frequency of the mode.

Indeed, when we use a continuous superposition of plane waves, the wave number given by the absolute value of the wave vector is directly related to the integration variable representing the superposition. In contrast, for a discrete superposition of mode functions, the summation is over the mode indices defining the frequency, which is determined by the boundary conditions for the Helmholtz equation.

It is this distinct feature that forces us to take advantage of the concept of a fractional root of the negative Laplacian. This tool allows us to represent the kernel as an operator acting on the completeness relation, which is ultimately a Dirac delta function.

Hence, the difference between the kernels of the vector potential and the electric field or the magnetic induction manifests itself in an additional factor to the Fourier representation of what would normally be the Dirac delta function by the same power of the wave number as in the mode sum. This feature stands out most clearly in Table II.

TABLE II

Building blocks of the wave functional $\Psi[\mathbf{F}] \equiv \mathcal{N}^{(F)} \exp(-\frac{1}{2}\beta^{(F)}\mathcal{I}^{(F)}[\mathbf{G}])$ of the vacuum in a resonator for a free field $\mathbf{F} = \mathbf{F}(t, \mathbf{r})$, given either by the electric field \mathbf{E} , the magnetic induction \mathbf{B} or the vector potential \mathbf{A} emerging from the infinite product of ground state wave functions $\psi_\ell(F_\ell) \equiv \mathcal{N}_\ell^{(F)} \exp[-(F_\ell/\mathcal{F}_\ell)^2/2]$ of the ℓ -th mode. For the example of \mathbf{A} we obtain two different kernels and two different fields in the double integral. For the modes \mathbf{u}_ℓ , we find a kernel $\sim 1/r^4$ with \mathbf{A} in the integral whereas for \mathbf{w}_ℓ , we arrive at the same kernel as in \mathbf{E} and \mathbf{B} , but now $\nabla \times \mathbf{A}$ appears. Here $\mathcal{N}^{(F)} \equiv \prod_\ell \mathcal{N}_\ell^{(F)}$ denotes a normalization constant, and the bilinear form $\mathcal{I}^{(F)} \equiv \int d^3r \int d^3r' \mathbf{G}^\dagger(\mathbf{r}) \underline{\mathcal{K}}^{(F)}(\mathbf{r}, \mathbf{r}') \mathbf{G}(\mathbf{r}')$ associated with \mathbf{F} can be reduced to a scalar kernel $\mathcal{K}^{(F)} \equiv \mathcal{K}^{(F)}(\mathbf{r})$ in the mode basis $\{\mathbf{f}_\ell\}$ and is given by the Fourier integral $\mathfrak{F}\{F(k)\} \equiv (2\pi)^{-3} \int d^3k F(k) e^{i\mathbf{k}\cdot\mathbf{r}}$ extending over all space.

Field \mathbf{F}	Mode basis $\{\mathbf{f}_\ell\}$	F_ℓ/\mathcal{F}_ℓ	$(F_\ell/\mathcal{F}_\ell)^2$	$\beta^{(F)}$	$F(k_\ell)$	\mathbf{G}	Scalar kernel $\mathcal{K}^{(F)}$
\mathbf{A}	$\{\mathbf{u}_\ell\}$	$\frac{A_\ell}{\mathcal{A}_\ell}$	$\beta^{(A)} A_\ell^2 k_\ell \mathcal{V}_\ell$	$\frac{\varepsilon_0 c}{\hbar}$	k_ℓ	\mathbf{A}	$\mathfrak{F}\{k\} \sim -1/r^4$
\mathbf{E}	$\{\mathbf{v}_\ell\} = \{\mathbf{u}_\ell\}$	$\frac{E_\ell}{\mathcal{E}_\ell} = \frac{E_\ell}{\mathcal{A}_\ell \omega_\ell}$	$\beta^{(E)} E_\ell^2 k_\ell^{-1} \mathcal{V}_\ell$	$\frac{\varepsilon_0}{\hbar c}$	$1/k_\ell$	\mathbf{E}	$\mathfrak{F}\{1/k\} \sim 1/r^2$
\mathbf{B}	$\{\mathbf{w}_\ell\} = \{k_\ell^{-1} \nabla \times \mathbf{u}_\ell\}$	$\frac{B_\ell}{\mathcal{B}_\ell} = \frac{c B_\ell}{\mathcal{E}_\ell}$	$\beta^{(B)} B_\ell^2 k_\ell^{-1} \mathcal{V}_\ell$	$\frac{\varepsilon_0 c}{\hbar}$	$1/k_\ell$	\mathbf{B}	$\mathfrak{F}\{1/k\} \sim 1/r^2$
\mathbf{A}	$\{\mathbf{w}_\ell\} = \{k_\ell^{-1} \nabla \times \mathbf{u}_\ell\}$	$\frac{A_\ell^{(w)}}{\mathcal{A}_\ell}$	$\beta^{(A)} A_\ell^{(w)2} k_\ell^{-1} \mathcal{V}_\ell$	$\frac{\varepsilon_0 c}{\hbar}$	$1/k_\ell$	$\nabla \times \mathbf{A}$	$\mathfrak{F}\{1/k\} \sim 1/r^2$

1.4. Overview

Our article is organized as follows: in Sect. 2, we derive an expression for the wave function of the vacuum in a resonator in terms of a sum over modes. For this purpose, we start from the corresponding probability amplitudes of every mode being in the ground state for \mathbf{A} , \mathbf{E} , and \mathbf{B} . Since these expressions are identical in their form for the three fields of interest, we confine ourselves to a general field \mathbf{F} .

The wave function of the complete field in the vacuum is then the infinite product of all Gaussian wave functions, which translates into an exponential whose argument is a sum of all field strengths weighted by a function $F(k_\ell)$ whose form depends on the field \mathbf{F} that we consider.

We devote Sect. 3 to the elimination of the modes in the infinite product of the ground state wave functions by expressing the sum over modes by a double integral over space containing a bilinear form of the fields and a kernel. For this purpose, we replace the expansion coefficient F_ℓ by the integral over the product of the field and mode functions \mathbf{f}_ℓ and arrive, due to the appearance of the square of F_ℓ in the mode sum, at a double integral of a bilinear form of \mathbf{F} and a kernel. The kernel is then determined by the function F of the square root of the negative Laplacian acting on the completeness relation of the modes given by the transverse delta function. Since the field \mathbf{F} is already transverse, it suffices to work with the familiar delta function, which allows us to derive an explicit expression for the kernel and thus for the wave functional of the vacuum in a resonator.

This analysis demonstrates that the kernels of \mathbf{E} and \mathbf{B} are identical, but different from the one of \mathbf{A} . In Sect. 4, we show that when we use the eigenmodes of \mathbf{B} to expand \mathbf{A} , we find the same kernel as for \mathbf{E} and \mathbf{B} .

We dedicate Sect. 5 to a comparison of the resulting expressions for the wave functional of the vacuum in the different representations. Moreover, we connect our results to the literature.

In Sect. 6, we calculate the Wightman tensor of the vacuum fields and show how it is related to the kernels of \mathbf{E} , \mathbf{B} , and \mathbf{A} . Furthermore, we sketch how vacuum expectation values can be expressed in terms of the wave functional.

We conclude in Sect. 7, by summarizing our results and providing an outlook.

In order to keep our article self-contained, we have included additional material that is helpful in understanding the main sections and keeping track of factors of 2. For example, in Appendix A, we summarize the essential building blocks of the free electromagnetic field. Here we concentrate on the expansions of \mathbf{A} , \mathbf{E} , and \mathbf{B} into a complete set of discrete modes. Moreover, we define the corresponding vacuum electric fields by equating the energy in a single mode of a given frequency to that of a quantized harmonic oscillator of the same frequency.

In Appendix B, we re-derive the energy of an electromagnetic field in a resonator. This calculation also most clearly brings out the difference in the powers of the mode frequency in the energy and the infinite product of the ground state wave functions. Moreover, we verify that the mode volumes of the \mathbf{u}_ℓ -modes and the \mathbf{w}_ℓ -modes are identical.

We devote Appendix C to the derivation of the ground state wave function in three different representations, that is, in the variables of \mathbf{E} , \mathbf{B} , and \mathbf{A} . In each case, we find a Gaussian whose dimensionless argument is determined by the ratio of the variable and the vacuum field strength.

In Appendix D, we present an alternative derivation of the double integral containing the bilinear form of the field and the kernel by *reverse engineering*. In contrast to the derivation in Sect. 3, we start by already assuming that the kernel is a scalar function and given by a Fourier integral of the function F . We then reduce this double integral to a single one of the square of the root of the negative Laplacian acting on \mathbf{F} . The mode expansion of \mathbf{F} then leads us straight to the mode sum of the wave function of the vacuum. We also point out a curious analogy to the P - and R -distributions [38] of quantum optics.

Finally, in Appendix E, we provide an explicit expression for the kernel by performing the relevant integrations with the help of a convergence factor.

Finally, in Appendix F, we derive an identity for the scalar product of two curls, evaluated at different positional arguments, needed in the evaluation of the Hamiltonian density of the electromagnetic field in Appendix B.

2. Infinite product of ground state wave functions

In this section, we derive the wave function of the electromagnetic vacuum in a resonator in terms of an infinite product of the ground state wave functions. Throughout the section, we use the field \mathbf{F} , which represents either the electric field \mathbf{E} , the magnetic induction \mathbf{B} , or the vector potential \mathbf{A} , and rely on the expansion

$$\mathbf{F}(t, \mathbf{r}) \equiv \sum_{\ell} F_{\ell}(t) \mathbf{f}_{\ell}(\mathbf{r}) \quad (1)$$

of these fields into their natural modes \mathbf{f}_{ℓ} determined by the Helmholtz equation subjected to the boundary conditions of the resonator as outlined in Appendix A. For the sake of simplicity in notation, we have not attached a superscript F on the modes \mathbf{f}_{ℓ} but emphasize that they depend on the choice of \mathbf{F} .

The expansion coefficient F_{ℓ} denotes the field strength in the mode \mathbf{f}_{ℓ} . Hence, F_{ℓ} depends on the choice of the modes. Obviously, in a different mode expansion, the field strength would be different. Again, for the sake of simplicity in notation, we suppress this dependence in F_{ℓ} but keep it in mind.

2.1. Wave function of the ground state

In Appendix C, we have recalled the expressions for the ground state wave functions ψ_{ℓ} in the representations of the electric field E_{ℓ} , the magnetic

induction B_{ℓ} , or the vector potential A_{ℓ} in the ℓ -th natural mode given by $\mathbf{f}_{\ell} = \mathbf{f}_{\ell}(\mathbf{r})$. Since the not yet normalized ground state is completely symmetric in phase space, it takes the same form in each of these representations and reads

$$\psi_{\ell}(\mathfrak{f}_{\ell}) \equiv \frac{1}{\sqrt[4]{\pi}} \exp\left(-\frac{1}{2}\mathfrak{f}_{\ell}^2\right), \quad (2)$$

where the dimensionless variable

$$\mathfrak{f}_{\ell} \equiv \frac{F_{\ell}}{\mathcal{F}_{\ell}} \quad (3)$$

involves the field F_{ℓ} in the ℓ -th mode \mathbf{f}_{ℓ} , and \mathcal{F}_{ℓ} is the corresponding field strength of the vacuum. Here F_{ℓ} is either E_{ℓ} , B_{ℓ} , or A_{ℓ} .

The quantities \mathcal{F}_{ℓ} are different for the three fields. Indeed, the strength

$$\mathcal{A}_{\ell} \equiv \sqrt{\frac{\hbar}{\epsilon_0 \omega_{\ell} \mathcal{V}_{\ell}}} \quad (4)$$

of the vector potential, which involves the mode volume \mathcal{V}_{ℓ} , is defined by postulating the electromagnetic energy of the ground state of the mode to be identical to $\frac{1}{2}\hbar\omega_{\ell}$, where ω_{ℓ} denotes the frequency of the ℓ -th mode \mathbf{f}_{ℓ} .

We emphasize that also the mode volume \mathcal{V}_{ℓ} depends on the choice of modes. For this reason, it should also carry a superscript indicating the type of eigenmodes used, such as \mathbf{u}_{ℓ} for the eigenmodes of \mathbf{A} , \mathbf{v}_{ℓ} for the eigenmodes of \mathbf{E} , or \mathbf{w}_{ℓ} for the eigenmodes of \mathbf{B} . However, for the sake of simplicity in notation, we suppress it.

The strength

$$\mathcal{E}_{\ell} \equiv \mathcal{A}_{\ell} \omega_{\ell} = \sqrt{\frac{\hbar\omega_{\ell}}{\epsilon_0 \mathcal{V}_{\ell}}} \quad (5)$$

follows from the Maxwell equations, that is, from the fact that in the Coulomb gauge without currents and charges, \mathbf{E} is the time derivative of \mathbf{A} .

Moreover, for \mathbf{B} we obtain in Appendix A the expression

$$\mathcal{B}_{\ell} \equiv \frac{\mathcal{A}_{\ell} \omega_{\ell}}{c} = \frac{\mathcal{E}_{\ell}}{c} \quad (6)$$

for the field strength B_{ℓ} of \mathbf{B} . Hence, apart from a factor of c , the field strengths \mathcal{B}_{ℓ} and \mathcal{E}_{ℓ} are identical.

When we substitute the dimensionless variable \mathfrak{f}_{ℓ} given by (3) into (2), the probability amplitude $\psi_{\ell} = \psi_{\ell}(F_{\ell})$ of finding the field F_{ℓ} of the mode ℓ in the ground state of this mode reads

$$\psi_{\ell}(F_{\ell}) \equiv \mathcal{N}_{\ell}^{(F)} \exp\left[-\frac{1}{2}\left(\frac{F_{\ell}}{\mathcal{F}_{\ell}}\right)^2\right], \quad (7)$$

where the normalization constant $\mathcal{N}_{\ell}^{(F)}$ takes the form

$$\mathcal{N}_{\ell}^{(F)} \equiv \frac{1}{\sqrt[4]{\pi} \sqrt{\mathcal{F}_{\ell}}}. \quad (8)$$

Due to the presence of \mathcal{F}_{ℓ} , the normalization constant $\pi^{-1/4}$ of the Gaussian in (2) is modified to achieve the condition

$$\int_{-\infty}^{\infty} dF_{\ell} |\psi_{\ell}(F_{\ell})|^2 = 1, \quad (9)$$

dictated by the Born interpretation.

2.2. Sum over modes

Hence, the corresponding probability amplitude $\Psi(\{F_\ell\})$ for finding the field F_{ℓ_1} in the mode ℓ_1 , F_{ℓ_2} in the mode ℓ_2 , etc., in the ground state is the infinite product

$$\Psi(\{F_\ell\}) \equiv \prod_{\ell} \psi_{\ell}(F_{\ell}) \quad (10)$$

of the ground state wave functions ψ_{ℓ} of all modes.

With the expression (7) for ψ_{ℓ} and the elementary property $e^a e^b = e^{a+b}$ of the exponential function, we arrive at the formula

$$\Psi(\{F_\ell\}) = \mathcal{N}^{(F)} \exp\left(-\frac{1}{2}\beta^{(F)}\Sigma^{(F)}\right), \quad (11)$$

where we have defined

$$\mathcal{N}^{(F)} \equiv \prod_{\ell} \mathcal{N}_{\ell}^{(F)} \quad (12)$$

representing the infinite product of all normalization factors $\mathcal{N}_{\ell}^{(F)}$, and introduced the abbreviation

$$\Sigma^{(F)} \equiv \sum_{\ell} F_{\ell}^2 F(k_{\ell}) \mathcal{V}_{\ell} \quad (13)$$

for the sum over all modes. Here we have taken into account that \mathcal{F}_{ℓ} is slightly different for the three fields. For this reason, the factor $\beta^{(F)}$ containing constants of nature such as \hbar , ε_0 , and c , listed in Table II, depends on the choice of F .

Moreover, since \mathcal{A}_{ℓ} and \mathcal{E}_{ℓ} depend differently on ω_{ℓ} , as shown by (4) and (5), we have a different dependence of $\Sigma^{(F)}$ on the wave number $k_{\ell} \equiv |\mathbf{k}_{\ell}| \equiv \omega_{\ell}/c$ of the mode indicated in (13) by the contribution $F(k_{\ell})$. Indeed, for \mathbf{E} and \mathbf{B} , we find

$$F^{(E/B)}(k_{\ell}) = k_{\ell}^{-1}, \quad (14)$$

while for \mathbf{A} , we obtain

$$F^{(A)}(k_{\ell}) = k_{\ell}. \quad (15)$$

It is this difference in F that leads to different expressions for the wave functional of the vacuum in a resonator, as we shall show in the next section.

2.3. Connection to free space

We conclude this analysis of the product of all ground state wave functions with a side, but not snide, remark about the corresponding calculation in free space. Since in this case we have a *continuous* superposition of modes, we have to deal with a *continuous product* of ground state wave functions. One possibility to describe this unusual quantity, which is fundamentally different from the discrete product arising in the case of a resonator, is to employ the Volterra–Schlesinger product integral [39] used to define in QED the quantum state after a time-dependent interaction [40].

However, a much more elementary approach to overcome this complication of a continuous product is to first discretize the modes, perform the discrete product and then replace the sum over modes again with the appropriate integral. Hence, in free space, we retreat from the continuous superposition

of modes to a discrete set and then return again to the continuous one.

In contrast, in the case of a resonator, we always deal with a discrete set, and the complication of the infinite product never occurs. We note that it would be interesting to perform the calculation in free space evaluating the continuous product, for example, with the help of the Volterra–Schlesinger product integral.

3. Bilinear forms and kernels

The goal of the present section is to *construct* from the mode expansion and the mode sum $\Sigma^{(F)}$ given by (13), an equivalent expression in terms of the complete field \mathbf{F} rather than the field amplitudes F_{ℓ} . For this purpose, we note that the terms in $\Sigma^{(F)}$ are quadratic in the fields F_{ℓ} . Therefore, $\Sigma^{(F)}$ might be represented by a quadratic form of the total field \mathbf{F} . Since $\Sigma^{(F)}$ is independent of the coordinate, there must be an integration over space involved.

However, this integral cannot just contain $\mathbf{F}^2 \equiv \mathbf{F}^{\dagger}\mathbf{F} = \mathbf{F} \cdot \mathbf{F}$, since that would lead to a quantity proportional to the energy in the resonator. Indeed, as shown in Appendix B, the contribution of the electric field or the magnetic induction to the energy scale is ω_{ℓ}^2 in the field oscillator frequency. Hence, a bilinear form of \mathbf{F} and a position-dependent kernel are necessary to obtain the scaling in k_{ℓ} required by the function $F(k_{\ell})$ given by (14) and (15).

In the present section, we pursue this approach in four steps. (i) We first obtain an explicit expression for the expansion coefficients F_{ℓ} of \mathbf{F} into the natural modes \mathbf{f}_{ℓ} and establish the completeness relation of \mathbf{f}_{ℓ} . (ii) Then we cast the mode sum $\Sigma^{(F)}$ into a double integral of the fields \mathbf{F} and \mathbf{F}' together with a matrix kernel. (iii) Since \mathbf{F} is transverse, this kernel reduces to a scalar, and (iv) we finally evaluate this kernel.

3.1. Completeness relation of transverse modes

Central to the representation of $\Sigma^{(F)}$ by a double integral of a bilinear form of \mathbf{F} and a kernel is the expansion

$$\mathbf{F} = \sum_{\ell} F_{\ell} \mathbf{f}_{\ell} \quad (16)$$

of the free field \mathbf{F} into the modes \mathbf{f}_{ℓ} discussed in Appendix A.

Indeed, the strength F_{ℓ} of \mathbf{F} in the mode \mathbf{f}_{ℓ} , which appears quadratically in $\Sigma^{(F)}$, follows from (16) by multiplication of \mathbf{f}_m , integration over space, and using the orthonormality relation

$$\frac{1}{\mathcal{V}_{\ell}} \int d^3r \mathbf{f}_{\ell}^{\dagger}(\mathbf{r}) \mathbf{f}_m(\mathbf{r}) = \delta_{\ell m} \quad (17)$$

of the modes. Moreover, the integration extends over the resonator volume, unless specified otherwise.

Indeed, the models f_ℓ y form a complete and orthonormal basis of *transverse mode-space* since they are eigenfunctions of the self-adjoint Helmholtz operator applied to the field \mathbf{F} as discussed in Appendix A.

We arrive at the explicit form

$$F_\ell = \frac{1}{\mathcal{V}_\ell} \int d^3 r' f_\ell^\dagger(\mathbf{r}') \mathbf{F}(t, \mathbf{r}') \quad (18)$$

or

$$F_\ell = \frac{1}{\mathcal{V}_\ell} \int d^3 r' f_\ell'^\dagger \mathbf{F}'. \quad (19)$$

Here we have attached a prime on \mathbf{F} and f_ℓ to emphasize the fact that both depend on the integration variable \mathbf{r}' rather than \mathbf{r} .

Since the field \mathbf{F} and the modes f_ℓ are hermitian fields, we have the identity $F_\ell^\dagger = F_\ell$, and thus,

$$\frac{1}{\mathcal{V}_\ell} \int d^3 r' \mathbf{F}'^\dagger f_\ell' = \frac{1}{\mathcal{V}_\ell} \int d^3 r f_\ell^\dagger \mathbf{F}. \quad (20)$$

When we substitute (19) into the expansion (16), we find

$$\mathbf{F} = \sum_\ell \frac{1}{\mathcal{V}_\ell} \int d^3 r' (f_\ell'^\dagger \mathbf{F}') f_\ell, \quad (21)$$

which when we interchange the sum and the integral reduces to

$$\mathbf{F} = \int d^3 r' \left[\sum_\ell \frac{1}{\mathcal{V}_\ell} f_\ell(\mathbf{r}) f_\ell^\dagger(\mathbf{r}') \right] \mathbf{F}' \quad (22)$$

or

$$\mathbf{F} \equiv \int d^3 r' \underline{\mathcal{Q}}(\mathbf{r}, \mathbf{r}') \mathbf{F}', \quad (23)$$

where we have introduced the term

$$\underline{\mathcal{Q}}(\mathbf{r}, \mathbf{r}') \equiv \sum_\ell \frac{1}{\mathcal{V}_\ell} f_\ell(\mathbf{r}) f_\ell^\dagger(\mathbf{r}'). \quad (24)$$

In order to maintain the identity $\mathbf{F} = \mathbf{F}$ in (23), the kernel $\underline{\mathcal{Q}}$ has to act as a delta-function-like object with respect to the spatial coordinates. However, since our modes are in the Coulomb gauge and are thus transverse, $\underline{\mathcal{Q}}$ cannot be an ordinary delta function, but must be a transverse delta function $\underline{\delta}^\perp$. Thus (24) takes the form

$$\underline{\mathcal{Q}}(\mathbf{r}, \mathbf{r}') \equiv \underline{\delta}^\perp(\mathbf{r} - \mathbf{r}') \quad (25)$$

with the expansion (24) in terms of the modes f_ℓ .

Hence, the matrix $\underline{\mathcal{Q}}$ defines a completeness relation and represents the kernel of a projection operator $\underline{\mathcal{P}}^\perp$ onto the (function) space spanned by the transverse (generalized Fourier) modes, i.e.,

$$\underline{\mathcal{P}}^\perp(\bullet) = \int d^3 r' \underline{\mathcal{Q}}(\mathbf{r}, \mathbf{r}') \bullet, \quad (26)$$

where (\bullet) acts as a placeholder for an arbitrary vector field to be projected onto that space.

3.2. Mode sum as double integral

We are now in the position to cast the sum $\Sigma^{(F)}$ over modes defined by (13) into a double integral containing a bilinear form of \mathbf{F} and a kernel $\underline{\mathcal{K}}$. In particular, we can obtain an exact expression for $\underline{\mathcal{K}}$.

For this purpose, we substitute the expression (19) for F_ℓ combined with the symmetry relation (20) of F_ℓ into $\Sigma^{(F)}$ and find the identity

$$\Sigma^{(F)} \equiv \Sigma^{(F)}[\mathbf{F}] = \int d^3 r \int d^3 r' \mathbf{F}^\dagger \underline{\mathcal{K}}(\mathbf{r}, \mathbf{r}') \mathbf{F}' \quad (27)$$

with the kernel

$$\underline{\mathcal{K}}(\mathbf{r}, \mathbf{r}') \equiv \sum_\ell \frac{1}{\mathcal{V}_\ell} F(k_\ell) f_\ell(\mathbf{r}) f_\ell^\dagger(\mathbf{r}'). \quad (28)$$

When we compare $\underline{\mathcal{K}}$ to the completeness relation (24), we find that, apart from the appearance of $F(k_\ell)$ from Table II, which is due to the different powers of k_ℓ in the vacuum field strength \mathcal{F}_ℓ , they are identical. Therefore, we want to eliminate $F(k_\ell)$ from the sum over modes in (28) by recalling the Helmholtz equation in the form

$$(-\Delta) f_\ell = k_\ell^2 f_\ell, \quad (29)$$

which shows that f_ℓ is the eigenvector of the negative Laplacian associated with the eigenvalue k_ℓ^2 .

As a result, we find the identity

$$F(k_\ell) f_\ell = F(\sqrt{-\Delta}) f_\ell, \quad (30)$$

and the kernel $\underline{\mathcal{K}}$ given by (28) reduces to

$$\underline{\mathcal{K}}(\mathbf{r}, \mathbf{r}') = F(\sqrt{-\Delta}) \underline{\delta}^\perp(\mathbf{r} - \mathbf{r}') \quad (31)$$

or equivalently

$$\underline{\mathcal{K}}_{mn}(\mathbf{r}, \mathbf{r}') = F(\sqrt{-\Delta}) \underline{\delta}_{mn}^\perp(\mathbf{r} - \mathbf{r}') \quad (32)$$

in component notation. Here we have recalled (25).

We emphasize that in (31) and (32), the differentiation in the Laplacian could be with respect to \mathbf{r} or \mathbf{r}' . This fact follows directly from the definition, (28), of the kernel or from the argument of the transverse delta function. For this reason, we have not attached a subscript \mathbf{r} to the Laplacian.

3.3. Simplification of the kernel

Next, we recall that the tensorial version $\delta(\mathbf{r}) \underline{\mathbb{1}}_3$ of the familiar Dirac delta function $\delta(\mathbf{r})$ contains not only the transverse part $\underline{\delta}^\perp(\mathbf{r})$, but also the longitudinal part $\underline{\delta}^\parallel(\mathbf{r})$, and reads in components

$$\delta(\mathbf{r}) \delta_{mn} = \underline{\delta}_{mn}^\perp(\mathbf{r}) + \underline{\delta}_{mn}^\parallel(\mathbf{r}) \quad (33)$$

or

$$\underline{\delta}_{mn}^\perp(\mathbf{r}) = \delta(\mathbf{r}) \delta_{mn} - \underline{\delta}_{mn}^\parallel(\mathbf{r}). \quad (34)$$

The operator $F(\sqrt{-\Delta})$ acting on $\underline{\delta}^\parallel$ does not change the directionality of the longitudinal part. This property stands out most clearly in its Fourier representation

$$\underline{\delta}_{mn}^\parallel(\mathbf{r}) \equiv \int \frac{d^3 k}{(2\pi)^3} e^{i\mathbf{k}\mathbf{r}} \frac{k_m k_n}{k^2}. \quad (35)$$

Indeed, we find

$$F(\sqrt{-\Delta}) \underline{\delta}_{mn}^\parallel = \int \frac{d^3 k}{(2\pi)^3} e^{i\mathbf{k}\mathbf{r}} F(k) \frac{k_m k_n}{k^2}, \quad (36)$$

where we have used the fact that a plane wave is also an eigenfunction of the negative Laplacian in free space corresponding to the eigenvalue k^2 , i.e.,

$$(-\Delta)e^{i\mathbf{k}\mathbf{r}} = k^2 e^{i\mathbf{k}\mathbf{r}}. \quad (37)$$

As a result, the kernel $\underline{\underline{\mathcal{K}}}_{mn}$ given by (32) reads

$$\underline{\underline{\mathcal{K}}}_{mn} = \delta_{mn}\mathcal{K}^{(F)} + \underline{\underline{\mathcal{K}}}_{mn}^{\parallel}, \quad (38)$$

where the part

$$\mathcal{K}^{(F)}(\mathbf{r} - \mathbf{r}') \equiv F(\sqrt{-\Delta}) \delta(\mathbf{r} - \mathbf{r}') \quad (39)$$

of the kernel, which is diagonal, arises from the operator $F(\sqrt{-\Delta})$ acting on the familiar Dirac delta function.

On the other hand, according to (36), the expression

$$\underline{\underline{\mathcal{K}}}_{mn}^{\parallel} = F(\sqrt{-\Delta}) \underline{\underline{\delta}}_{mn}^{\parallel}(\mathbf{r} - \mathbf{r}') \quad (40)$$

is still longitudinal.

Next, we recall that in the double integral (27), the fields \mathbf{F} and \mathbf{F}' , on which the kernel acts, are already transverse, since they are expanded into the transverse modes \mathbf{f}_ℓ . Hence $\underline{\underline{\mathcal{K}}}_{mn}^{\parallel}$ does not contribute, and we arrive at the expression

$$\Sigma^{(F)} = \int d^3r \int d^3r' \mathcal{K}^{(F)} \mathbf{F}^\dagger \mathbf{F}', \quad (41)$$

where we have made use of the fact that $\mathcal{K}^{(F)}$ is a scalar, which can now be moved out of the matrix products.

3.4. Evaluation of the kernel

Finally, we evaluate the scalar kernel $\mathcal{K}^{(F)}$ given by (39). Here, two possibilities offer themselves:

(i) We recall the Green's function relation

$$(-\Delta)\frac{1}{r} = 4\pi\delta(r), \quad (42)$$

which leads us to the expression

$$\mathcal{K}^{(F)}(\mathbf{r}) = \frac{1}{4\pi} F(\sqrt{-\Delta}) (-\Delta)\frac{1}{|\mathbf{r}|} \quad (43)$$

for $\mathcal{K}^{(F)}$, or

(ii) we employ the Fourier representation

$$\delta(\mathbf{r}) \equiv \frac{1}{(2\pi)^3} \int d^3k e^{i\mathbf{k}\mathbf{r}} \quad (44)$$

of the Dirac delta function to evaluate $\mathcal{K}^{(F)}$.

In this article, we pursue the second approach since it is straightforward. Indeed, from (39) we immediately find with (37) the representation

$$\mathcal{K}^{(F)}(\mathbf{r}) = \frac{1}{(2\pi)^3} \int d^3k F(k) e^{i\mathbf{k}\mathbf{r}}. \quad (45)$$

In Appendix E, we evaluate this integral for the two cases $F(k) = k^{-1}$ or $F(k) = k$ corresponding to the fields \mathbf{E} and \mathbf{B} or \mathbf{A} , and we find

$$\mathcal{K}^{(E)}(\mathbf{r}) = \mathcal{K}^{(B)}(\mathbf{r}) = \frac{1}{2\pi^2} \frac{1}{|\mathbf{r}|^2} \quad (46)$$

or

$$\mathcal{K}^{(A)}(\mathbf{r}) = -\frac{1}{\pi^2} \frac{1}{|\mathbf{r}|^4}. \quad (47)$$

We note that apart from slightly different prefactors, the power laws of the two kernels in (46) and (47) are different. While $\mathcal{K}^{(E)} = \mathcal{K}^{(B)}$ decays as $\mathcal{K}^{(E/B)} \sim 1/r^2$, the one for \mathbf{A} , i.e., $\mathcal{K}^{(A)}$, decays as $\mathcal{K}^{(A)} \sim 1/r^4$. Moreover, they also differ in sign. While $\mathcal{K}^{(E/B)}$ is positive, $\mathcal{K}^{(A)}$ is negative.

At first sight, this sign change might cause a problem in the exponential. However, when we recall that the double integral with the bilinear form of \mathbf{A} and $\mathcal{K}^{(A)}$ is identical to the mode sum $\Sigma^{(A)}$ where each term is positive, we recognize that there is really no problem here.

3.5. Wave functional

We conclude by combining our results to obtain the wave functional $\Psi[\mathbf{F}]$ of the vacuum in a resonator expressed by the field \mathbf{F} . Indeed, when we use the connection (41) between the mode sum $\Sigma^{(F)}$ and the double integral, we find the expression

$$\Psi[\mathbf{F}] = \mathcal{N}^{(F)} \exp \left[-\frac{\beta^{(F)}}{2} \int d^3r \int d^3r' \mathcal{K}^{(F)} \mathbf{F} \cdot \mathbf{F}' \right], \quad (48)$$

where the kernel $\mathcal{K}^{(F)}$ involves the difference $\mathbf{r} - \mathbf{r}'$ of the two integration variables only.

We emphasize that in contrast to the infinite product $\Psi(\{F_\ell\})$, which is in terms of the set $\{F_\ell\}$ of field strengths in all modes and given by (11), we now have the complete field \mathbf{F} . Hence, the quantity Ψ defined by (48) represents a *functional* of \mathbf{F} as indicated by the square brackets in $\Psi[\mathbf{F}]$.

4. Vector potential once more

In the preceding section, we have derived the wave functional $\Psi[\mathbf{A}]$ in terms of the vector potential \mathbf{A} and have found a kernel (47), which is different from the ones of \mathbf{E} and \mathbf{B} , given by (46). However, it has been argued [4, 7, 41] that an expression for a wave functional solely in terms of \mathbf{A} is problematic since the vacuum, and hence the wave functional should be gauge invariant, and the full vector potential is not gauge invariant. This line of reasoning was first used by Wheeler [4] in his original article on the wave functional, where we find the quote:

"Often the dynamics of the electromagnetic field is discussed in terms of the vector potential \mathbf{A} , connected with \mathbf{H} , by the equation

$$\mathbf{H} = \text{curl } \mathbf{A}.$$

Then the probability amplitude is evaluated in the first instance as a functional of \mathbf{A} . Only later is it discovered, as a consequence of gauge invariance, that \mathbf{A} comes into evidence in the state functional only in the form of $\mathbf{H} = \text{curl } \mathbf{A}$."

Other authors [7, 41] argued in the same vein and thus concentrated their effort on expressions for the wave functional of the vacuum in terms of $\nabla \times \mathbf{A} \equiv \mathbf{B}$ instead of \mathbf{A} . However, we found in (48) exactly such a wave functional $\Psi[\mathbf{A}]$ and a corresponding kernel (47). Hence, we are led to the question of how to reconcile these opposing points of view.

Our answer to this question rests on the fact that the appearance of $\nabla \times \mathbf{A}$ is not a consequence of gauge invariance but a specific choice of mode expansion. Indeed, we first argue that due to the expansion in transverse modes, our expression is already gauge invariant. We then obtain an expression for the wave functional $\Psi[\mathbf{A}]$ in terms of $\nabla \times \mathbf{A}$ by use of the eigenmodes $\{\mathbf{w}_\ell\}$ of the magnetic induction \mathbf{B} without appealing to gauge invariance.

4.1. Field functionals, quantization and gauge invariance

While the gauge invariance argument seems superficially sound, it contains a very subtle flaw and is thus not applicable. Indeed, we start by noticing that electromagnetism is a gauge field theory [29], and it is thus essential to remove redundant gauge degrees of freedom during the quantization procedure. It is then, and only then, that we can identify the actual physical degrees of freedom of the theory. Any observable, such as correlation functions or the wave functional, are afterwards expressed solely in terms of the quantized physical degrees of freedom.

In contrast to earlier works of Wheeler [4] and Białyński-Birula [7, 41], we state and rely on a specific gauge choice from the beginning. Accidentally, the *gauge-fixing* of the Coulomb gauge directly isolates easy-to-interpret physical degrees of freedom in non-relativistic situations for the electromagnetic field. However, this comes at the cost of sacrificing the manifest Lorentz invariance of the theory. This procedure partitions the electromagnetic degrees of freedom into quantized (transverse) and non-quantized (longitudinal) degrees of freedom by enforcing the conditions $A_0 \equiv 0$ and $\nabla \cdot \mathbf{A} \equiv 0$ for the vector potential.

These quantized physical degrees of freedom are exactly our transverse fields \mathbf{A} , \mathbf{E} , and \mathbf{B} . Only these fields are associated with quantum states, that is the wave functions of our theory.

In Appendix C, we determine these wave functions for the ground state of the respective fields.

Since these wave functions form the starting point of our derivation, any expression we obtain from them is naturally expressed in terms of *gauge invariant quantities*, even if the *transverse part* of the vector potential, namely \mathbf{A} , appears in it. Consequently, our expression for the wave functional $\Psi[\mathbf{A}]$ of the vector potential (48), together with the associated kernel (47), is perfectly valid.

We conclude by returning to the subtle flaw in the argument of gauge invariance we have alluded to. Ultimately, a wave functional can only be defined *after quantization* of a gauge theory such as electromagnetism has already been achieved, as it is a fundamentally quantum object. More specifically, the fields \mathbf{F} appearing in it are *not classical fields* and, in general, do not even obey the *classical field equations*, but are mere c-number fields that parameterize all *quantum mechanically valid* field configurations *interfering* in an appropriate functional integral.

Simultaneously, at this point in the development of the theory, the gauge-freedom is already incorporated in the choice of the quantized degrees of freedom, since all physically relevant quantities that appear, are by construction expressed without the gauge-redundant degrees of freedom. As a consequence, we cannot argue about the gauge-invariance of a quantity like a wave functional anymore when it is expressed in these quantities. Thus ultimately, it is the simple oversight that not the vector potential but only its transverse part can appear in field functionals, which leads to the demise of any post-quantization argument relying on gauge transformations/invariance.

Finally, although we worked in Coulomb gauge throughout this article, our reasoning applies to any gauge-fixing chosen during quantization. Moreover, it translates to the wave-functional of other theories featuring gauge-invariances [42], e.g., the quantization of weak field gravity [36]. However, we note that when one is studying such cases, starting from a more modern path-integral formulation seems preferable [43] since gauge-fixings are implemented more easily via functional δ -functions inside the path integral.

With these ideas in mind, we briefly comment on possible generalizations of our calculation to relativistic situations using the standard QED approach. While we have sacrificed the manifest Lorentz covariance by our choice of the Coulomb gauge, this was simply due to our interest in the cavity QED situation of the quantization in a resonator. If wanted, retaining Lorentz covariance and determining relativistically invariant analogs of the expressions (48) for the wave functionals is possible by resorting to the Gupta–Bleuler [45, 45] method or the more general approach of BRST quantization [46, 47]. For a modern discussion contrasting these approaches as applied to electromagnetism in ξ -gauge, a generalization of Lorenz gauge, we refer to [47].

4.2. Wave functional in eigenmodes of magnetic induction

In order to reexpress the wave functional $\Psi[\mathbf{A}]$, as suggested by Wheeler and Białyński-Birula, in terms of $\nabla \times \mathbf{A}$, we expand \mathbf{A} into the eigenmodes

TABLE III

Wave functional $\Psi[\mathbf{F}]$ of the vacuum in a resonator for the three fields $\mathbf{F} = \mathbf{E}, \mathbf{B}$, or \mathbf{A} and their corresponding kernel $\underline{\mathcal{K}} \equiv \underline{\mathcal{K}}(\mathbf{r}, \mathbf{r}')$ when expressed in the mode basis $\{\mathbf{f}_\ell\} = \{\mathbf{u}_\ell\}, \{\boldsymbol{\nu}_\ell\}$, or $\{\mathbf{w}_\ell\}$. Here, the prime indicates the field at the integration variable \mathbf{r}' rather than \mathbf{r} .

Field \mathbf{F}	Mode basis $\{\mathbf{f}_\ell\}$	Mode basis kernel $\underline{\mathcal{K}}(\mathbf{r}, \mathbf{r}')$	Wave functional $\Psi[\mathbf{F}]$ in field basis
\mathbf{E}	$\{\boldsymbol{\nu}_\ell\} = \{\mathbf{u}_\ell\}$	$\sum_\ell \frac{k_\ell^{-1}}{\mathcal{V}_\ell} \boldsymbol{\nu}_\ell \boldsymbol{\nu}'_\ell^\dagger$	$\mathcal{N}^{(E)} \exp\left(-\frac{1}{4\pi^2} \frac{\varepsilon_0}{\hbar c} \int d^3r \int d^3r' \frac{\mathbf{E} \cdot \mathbf{E}'}{ \mathbf{r} - \mathbf{r}' ^2}\right)$
\mathbf{B}	$\{\mathbf{w}_\ell\} = \{k_\ell^{-1} \nabla \times \mathbf{u}_\ell\}$	$\sum_\ell \frac{k_\ell^{-1}}{\mathcal{V}_\ell} \mathbf{w}_\ell \mathbf{w}'_\ell^\dagger$	$\mathcal{N}^{(B)} \exp\left(-\frac{1}{4\pi^2} \frac{\varepsilon_0 c}{\hbar} \int d^3r \int d^3r' \frac{\mathbf{B} \cdot \mathbf{B}'}{ \mathbf{r} - \mathbf{r}' ^2}\right)$
\mathbf{A}	$\{\mathbf{w}_\ell\} = \{k_\ell^{-1} \nabla \times \mathbf{u}_\ell\}$	$\sum_\ell \frac{k_\ell^{-1}}{\mathcal{V}_\ell} \mathbf{w}_\ell \mathbf{w}'_\ell^\dagger$	$\mathcal{N}^{(A^{(w)})} \exp\left(-\frac{1}{4\pi^2} \frac{\varepsilon_0 c}{\hbar} \int d^3r \int d^3r' \frac{(\nabla \times \mathbf{A}) \cdot (\nabla \times \mathbf{A}')}{ \mathbf{r} - \mathbf{r}' ^2}\right)$
\mathbf{A}	$\{\mathbf{u}_\ell\}$	$\sum_\ell \frac{k_\ell}{\mathcal{V}_\ell} \mathbf{u}_\ell \mathbf{u}'_\ell^\dagger$	$\mathcal{N}^{(A)} \exp\left(\frac{1}{2\pi^2} \frac{\varepsilon_0 c}{\hbar} \int d^3r \int d^3r' \frac{\mathbf{A} \cdot \mathbf{A}'}{ \mathbf{r} - \mathbf{r}' ^4}\right)$

$$\mathbf{w}_\ell \equiv k_\ell^{-1} \nabla \times \mathbf{u}_\ell \quad (49)$$

of the wave equation for \mathbf{B} , rather than the one for \mathbf{A} , i.e.,

$$\mathbf{A} \equiv \sum_\ell A_\ell^{(w)} \mathbf{w}_\ell. \quad (50)$$

Here we have attached a superscript w to the amplitude A_ℓ to reflect the fact that this expansion is in the set of modes $\{\mathbf{w}_\ell\}$.

When we now take the curl of this representation of \mathbf{A} , recall the Coulomb gauge condition, as well as the Helmholtz equation for \mathbf{u}_ℓ , we find

$$\nabla \times \mathbf{A} = \sum_\ell A_\ell^{(w)} k_\ell \mathbf{u}_\ell. \quad (51)$$

Consequently, the expansion coefficient $A_\ell^{(w)}$ in the w -representation takes the form

$$A_\ell^{(w)} = \frac{1}{k_\ell} \frac{1}{\mathcal{V}_\ell} \int d^3r \mathbf{u}_\ell^\dagger (\nabla \times \mathbf{A}). \quad (52)$$

When we compare this expression to the corresponding one for \mathbf{F}_ℓ , expressed in the natural modes \mathbf{f}_ℓ , i.e., to (19), we note an additional factor k_ℓ^{-1} , which allows us to regain the same kernel in the double integral as in \mathbf{E} and \mathbf{B} .

Since the quantization of \mathbf{A} now takes place in the \mathbf{w}_ℓ -modes, the wave function of the vacuum in the resonator reads

$$\Psi[\mathbf{A}] = \mathcal{N}^{(A)} \exp\left(-\frac{1}{2} \beta^{(A)} \Sigma^{(A^{(w)})}\right), \quad (53)$$

where now the sum

$$\Sigma^{(A^{(w)})} \equiv \sum_\ell \left(A_\ell^{(w)}\right)^2 k_\ell^{-1} \mathcal{V}_\ell \quad (54)$$

runs over the \mathbf{w}_ℓ -modes.

When we substitute the explicit form (52) of the expansion coefficients $A_\ell^{(w)}$ into the mode sum (54), we arrive at

$$\Sigma^{(A^{(w)})} = \int d^3r \int d^3r' (\nabla \times \mathbf{A})^\dagger \underline{\mathcal{K}}(\mathbf{r}, \mathbf{r}') (\nabla \times \mathbf{A}'), \quad (55)$$

where according to (15) the term $F(k_\ell)$ in the kernel $\underline{\mathcal{K}}$ defined by (28) takes the form

$$F(k_\ell) = k_\ell^{-1}, \quad (56)$$

and is thus identical to the one for \mathbf{E} and \mathbf{B} in their natural modes.

As a consequence, the kernel for the vector potential \mathbf{A} expanded into \mathbf{w}_ℓ - rather than \mathbf{u}_ℓ -modes is identical to that of \mathbf{E} and \mathbf{B} . However, now the wave functional of the vacuum in the representation of \mathbf{A} contains only \mathbf{A} in the form $\nabla \times \mathbf{A}$. In this way, $\Psi[\mathbf{A}]$ is expressed in terms of the magnetic induction, which is a gauge invariant quantity.

5. Discussion of wave functionals

We are now in a position to present the explicit expressions for the wave functionals of the vacuum in a resonator, as summarized in Table III. Moreover, we compare and contrast the corresponding expressions to the ones in the literature.

5.1. Dependence on mode expansion

The central message of Table III is that the kernel of the wave functional depends on the mode expansion of the field. At first sight, this property is surprising since the creation of the bilinear form of the complete field removes the field expansion. However, the wave functional $\Psi[\mathbf{A}]$ of the vacuum in the representation of the vector potential \mathbf{A} , summarized in the first and last row of Table III, demonstrates this feature in a striking way.

Indeed, when we use the eigenmode expansion of \mathbf{A} , given by $\{\mathbf{u}_\ell\}$, which is identical to the one of the electric field \mathbf{E} , we find a kernel that is proportional to $1/r^4$ and negative. In this case, the bilinear form involves only \mathbf{A} .

However, when we employ the eigenmode expansion of the magnetic induction \mathbf{B} , i.e., the modes $\{k_\ell^{-1}\nabla \times \mathbf{u}_\ell\}$, the kernel of \mathbf{E} , which is identical to that of \mathbf{B} , emerges and enjoys the decay $1/r^2$. In this case, the kernel is positive. However, most importantly, the bilinear form does not involve \mathbf{A} but $\nabla \times \mathbf{A} \equiv \mathbf{B}$.

This dependence of the kernel on the mode representation, and the associated form of the bilinear form, is reminiscent of the different operator orderings in quantum mechanics and the associated quasi-probability distribution functions. We recall [48] that symmetric ordering requires the use of the Wigner function, whereas anti-normal ordering leads us to the Husimi or Q -function. Normal ordering brings in the P -distribution.

Hence, the same quantum state can enjoy different phase space distribution functions depending on the choice of the operator ordering. Nevertheless, the quantum mechanical average of interest is always the same.

This analogy draws attention to the quantity so far not addressed in our article, that is, the field operators. Indeed, we have concentrated excessively on

the wave functional, which of course, could be employed to calculate expectation values of the field operators. In order to perform this evaluation in an effective way, it is necessary to have the operators to be averaged in the same modes as the wave functional. Indeed, an identical mode expansion in operators and wave functionals is necessary to express the operator in a c-number representation. This requirement is analogous to the familiar technique of one-particle quantum mechanics to perform averages using wave functions in the eigenrepresentation of the operator. In this way, we can evaluate the expectation values by functional integration as discussed in the next section.

5.2. Connection to free space

We conclude by comparing and contrasting the form of the functionals in a resonator to the ones in free space first suggested by Wheeler [4] and discussed and extended by Białynicki-Birula. Here we confine ourselves to the one involving $\nabla \times \mathbf{A}$, which according to [7, 41], reads

$$\Psi[\mathbf{A}] = \mathcal{N}^{(A)} \exp\left(-\frac{1}{4\pi^2\hbar} \sqrt{\frac{\varepsilon_0}{\mu_0}} \int d^3r \int d^3r' \frac{(\nabla \times \mathbf{A}) \cdot (\nabla' \times \mathbf{A}')}{|\mathbf{r} - \mathbf{r}'|^2}\right). \quad (57)$$

The only difference to the expression in the fourth row of Table III is in the prefactor $\beta^{(A)}$ containing fundamental constants. Whereas we always use ε_0 and c , Białynicki-Birula's expression involves the ratio $\sqrt{\varepsilon_0/\mu_0}$. Here μ_0 denotes the permeability of the vacuum.

However, the Kirchhoff identity

$$\frac{1}{\mu_0\varepsilon_0} = c^2 \quad (58)$$

immediately yields the connection formula

$$\sqrt{\frac{\varepsilon_0}{\mu_0}} = \varepsilon_0 c, \quad (59)$$

in complete agreement with our expression in Table III.

6. Wave functionals and expectation values

In the preceding sections, we have made our way to explicit expressions for the wave functional of the electromagnetic vacuum, beginning with the quantization of the electromagnetic field in a resonator. Most of the expressions we obtained coincide with the ones found previously by Wheeler [4] and Białynicki-Birula [7, 41] for free space, although now obtained for the case of a resonator. However, one expression in terms of a bilinear functional of \mathbf{A} is new to the best of our knowledge.

While these functionals are certainly interesting from a fundamental point of view, we ultimately go through the trouble of setting up field theory in order to calculate observables, i.e., scattering cross-sections, correlation functions, and their more complicated cousins. Naturally, we must thus face the question of how these calculations can be performed with the field wave functions and functionals. This problem constitutes the topic of this section, and we shall show by the example of such a calculation for a specific correlation function how this can be done.

We focus our effort on the Wightman tensor $\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}(t)$ for the field \mathbf{F} , which contains all first-order correlation functions of the vector field \mathbf{F} evaluated at two points, \mathbf{r} and \mathbf{r}' , in space. Furthermore, it is of specific interest because it can be used to easily determine the excitation probability [49] for an atom in a cavity due to the vacuum field.

6.1. A general correlation function

We begin by stating the definition [49] of the equal-time two-point Wightman tensor

$$\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}(t) \equiv \langle \mathbf{0} | \hat{\mathbf{F}}(t, \mathbf{r}) \hat{\mathbf{F}}^\dagger(t, \mathbf{r}') | \mathbf{0} \rangle \quad (60)$$

for the field \mathbf{F} , which is the expectation value of the outer product of the field operators $\hat{\mathbf{F}}(t, \mathbf{r}) \hat{\mathbf{F}}^\dagger(t, \mathbf{r}')$ at fixed time t but in different locations \mathbf{r} and \mathbf{r}' .

In fact, (60) describes the spatial correlations in the vacuum field \mathbf{F} at the respective positions \mathbf{r} and \mathbf{r}' .

For the purpose of illustrating the formalism, $\underline{\mathcal{W}}_{\mathbf{r}\mathbf{r}'}^{(F)}$ may be seen as a tensorial version of the correlation functions introduced by Glauber [38, 50] in quantum optics. For example, taking the trace of the Wightman tensor yields an intensity correlation function, which is a precursor of the (spatial) first-order coherence function $G^{(1)}(t, \mathbf{r}; t, \mathbf{r}')$.

6.2. Wightman tensor via mode decomposition

We begin by expressing the Wightman tensor in terms of the \mathbf{f}_ℓ -modes (16), which for (60) yields the decomposition

$$\underline{\mathcal{W}}_{\mathbf{r}\mathbf{r}'}^{(F)}(t) = \sum_{\ell, \ell'} \langle \mathbf{0} | \hat{F}_\ell(t) \hat{F}_{\ell'}(t) | \mathbf{0} \rangle \mathbf{f}_\ell(\mathbf{r}) \mathbf{f}_{\ell'}^\dagger(\mathbf{r}'). \quad (61)$$

Here we have used the linearity of the mode sums and acted with the vacuum directly on the operator parts of the fields. Note that in the process we used the fact that the fields are hermitian operators, i.e., $\hat{F}_{\ell'}^\dagger = \hat{F}_{\ell'}$.

6.2.1. Determination of vacuum expectation value

Proceeding from (61), our next task is to calculate the field operator expectation value with respect to the vacuum state $\mathcal{O}_{\ell\ell'}^{(F)}$, for which we introduce the abbreviation

$$\mathcal{O}_{\ell\ell'}^{(F)} \equiv \langle \mathbf{0} | \hat{F}_\ell(t) \hat{F}_{\ell'}(t) | \mathbf{0} \rangle. \quad (62)$$

Since the time argument is identical for both field operators and is immaterial for what follows, we will suppress it going forward and simply write $\hat{F}_\ell(t) \equiv \hat{F}_\ell$ from now on to compactify the notation.

In order to evaluate the expectation value (62), we recall that the non-interacting vacuum ket-state $|\mathbf{0}\rangle$ of the free (electromagnetic) field \mathbf{F} is a direct product

$$|\mathbf{0}\rangle \equiv \bigotimes_k |0_k\rangle = |0_1\rangle |0_2\rangle |0_3\rangle \dots |0_\ell\rangle \dots \quad (63)$$

of all ground states of all modes and that the operator \hat{F}_ℓ only acts on the ℓ -th mode. Other ground states $|0_k\rangle$ with $k \neq \ell$ are not affected by \hat{F}_ℓ .

Obviously, the same property holds true for the vacuum bra-vector $\langle \mathbf{0} |$, and none of the ground states $\langle 0_{k'} |$ with $k' \neq \ell'$ is affected by $\hat{F}_{\ell'}$, and they pass to the right, where they meet the ground states $|0_k\rangle$ from the ket-vacuum.

Since we can only take the scalar product between the same modes, we have to distinguish the two cases $\ell = \ell'$ and $\ell \neq \ell'$.

The first case of identical modes, i.e., $\ell \equiv \ell'$, leads us to the expression

$$\mathcal{O}_{\ell\ell}^{(F)} = \langle 0_\ell | \hat{F}_\ell^2 | 0_\ell \rangle \prod_{k \neq \ell} \langle 0_k | 0_k \rangle \quad (64)$$

or

$$\mathcal{O}_{\ell\ell}^{(F)} = \langle 0_\ell | \hat{F}_\ell^2 | 0_\ell \rangle, \quad (65)$$

where we have used the normalization condition $\langle 0_k | 0_k \rangle = 1$ of the ground state, which in the field representation reads

$$\int_{-\infty}^{\infty} dF_k \langle 0_k | F_k \rangle \langle F_k | 0 \rangle = \int_{-\infty}^{\infty} dF_k |\psi_k(F_k)|^2 \quad (66)$$

and is satisfied, since according to Appendix C we find

$$\psi_k(F_k) = \frac{1}{\sqrt[4]{\pi}} \frac{1}{\sqrt{\mathcal{F}_k}} \exp \left[-\frac{1}{2} \left(\frac{F_k}{\mathcal{F}_k} \right)^2 \right]. \quad (67)$$

Moreover, the field operator of the ℓ -th mode obeys the eigenvalue equation

$$\hat{F}_\ell |F_\ell\rangle = F_\ell |F_\ell\rangle \quad (68)$$

and as a consequence, we have the spectral representation

$$g(\hat{F}_\ell) \equiv \int_{-\infty}^{\infty} dF_\ell g(F_\ell) |F_\ell\rangle \langle F_\ell| \quad (69)$$

for integrable functions $g \equiv g(x)$.

When we introduce this spectral representation for the ℓ -th mode into (66), we obtain

$$\mathcal{O}_{\ell\ell}^{(F)} = \int_{-\infty}^{\infty} dF_\ell F_\ell^2 |\psi_\ell(F_\ell)|^2, \quad (70)$$

which with the help of the Gaussian wave function (67) reads

$$\mathcal{O}_{\ell\ell}^{(F)} = \frac{1}{2} \mathcal{F}_\ell^2. \quad (71)$$

Next, we consider the case $\ell \neq \ell'$, which yields the expression

$$\mathcal{O}_{\ell\ell'}^{(F)} = \langle 0_{\ell'} | \hat{F}_{\ell'} | 0_{\ell'} \rangle \langle 0_\ell | \hat{F}_\ell | 0_\ell \rangle \prod_{k \neq \ell, \ell'} \langle 0_k | 0_k \rangle. \quad (72)$$

We emphasize that, in contrast to (64), the mode indices ℓ and ℓ' appear now. Nevertheless, the normalization condition is again $\langle 0_k | 0_k \rangle = 1$ for each mode and reduces (72) to

$$\mathcal{O}_{\ell\ell'}^{(F)} = \langle 0_\ell | \hat{F}_\ell | 0_\ell \rangle \langle 0_{\ell'} | \hat{F}_{\ell'} | 0_{\ell'} \rangle. \quad (73)$$

When we now employ the field representation, again we find with the eigenvalue equation (68) for $\ell \neq \ell'$ the formula

$$\langle 0_{\ell'} | \hat{F}_{\ell'} | 0_{\ell'} \rangle = \int_{-\infty}^{\infty} dF_\ell F_\ell |\psi_\ell(F_\ell)|^2 = 0, \quad (74)$$

where in the last step we have used the symmetric Gaussian wave function (67) of the ground state.

When we combine the results (71) and (74), we find

$$\mathcal{O}_{\ell\ell'}^{(F)} = \frac{1}{2} \delta_{\ell\ell'} \mathcal{F}_\ell^2. \quad (75)$$

With the respective definitions of the vacuum fields \mathcal{F}_ℓ in (4) and (6), we can bring (75) into the final form

$$\mathcal{O}_{\ell\ell'}^{(F)} = \frac{\delta_{\ell\ell'}}{2} \frac{1}{\beta^{(F)} F(k_\ell)} \frac{1}{\mathcal{V}_\ell}, \quad (76)$$

which constitutes our result for the vacuum expectation value (62). This expression for the field \mathbf{F} is determined by the physical constants contained in $\beta^{(F)}$, the wave number k_ℓ together with the function F , and the mode volume \mathcal{V}_ℓ of the ℓ -th mode.

6.2.2. Wightman tensor and kernels

With the result for the vacuum expectation value, we are now in a position to determine the Wightman tensor of the field \mathbf{F} . Using the result from (76) and inserting it into (61), we arrive at

$$\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}(t) = \frac{1}{2\beta^{(F)}} \sum_{\ell} F^{-1}(k_\ell) \frac{1}{\mathcal{V}_\ell} \mathbf{f}_\ell(\mathbf{r}) \mathbf{f}_\ell^\dagger(\mathbf{r}') \quad (77)$$

for the mode-expanded version of the Wightman tensor. We observe that this expression seems reminiscent of the expression for the transverse delta function in terms of the modes (24).

Actually, with the help of the square root of the negative Laplacian, we can move the function $F^{-1}(k_\ell)$ out of the sum by reversing its action on the modes via

$$F^{-1}(k_\ell) \mathbf{f}_\ell(\mathbf{r}) = F^{-1}(\sqrt{-\Delta_{\mathbf{r}}}) \mathbf{f}_\ell(\mathbf{r}) \quad (78)$$

and using the independence of the right-hand side from the summation index ℓ . Together with the representation of the transverse delta function (24), we arrive at the expression

$$\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}(t) = \frac{1}{2\beta^{(F)}} F^{-1}(\sqrt{-\Delta_{\mathbf{r}}}) \underline{\underline{\delta}}^\perp(\mathbf{r} - \mathbf{r}') \quad (79)$$

for the Wightman tensor, which is fully consistent with the results obtained in [49] in free space for the electric or magnetic field.

When we now compare the expression (79) for the Wightman tensor $\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}$ with the one (31) of the kernel $\underline{\underline{\mathcal{K}}}_{\mathbf{r}\mathbf{r}'}^{(F)}$, we find that $F(\sqrt{-\Delta_{\mathbf{r}}})$ is either in the denominator or in the numerator. At the same time we obtain from the definitions (14) and (15) of F for \mathbf{E} , \mathbf{B} , and \mathbf{A} the relation

$$F^{(E)} = F^{(B)} = 1/F^{(A)}. \quad (80)$$

As a result, we arrive at the connection formulae

$$\begin{aligned} \underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(E)}(t) &= \frac{\hbar c}{2\varepsilon_0} \underline{\underline{\mathcal{K}}}_{\mathbf{r}\mathbf{r}'}^{(A)}(\mathbf{r} - \mathbf{r}') = \\ &= \frac{\hbar c}{2\varepsilon_0} (-\Delta_{\mathbf{r}}) \underline{\underline{\mathcal{K}}}_{\mathbf{r}\mathbf{r}'}^{(E)}(\mathbf{r} - \mathbf{r}') \end{aligned} \quad (81)$$

and

$$\begin{aligned} \underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(B)}(t) &= \frac{\hbar}{2\varepsilon_0 c} \underline{\underline{\mathcal{K}}}_{\mathbf{r}\mathbf{r}'}^{(A)}(\mathbf{r} - \mathbf{r}') = \\ &= \frac{\hbar}{2\varepsilon_0 c} (-\Delta_{\mathbf{r}}) \underline{\underline{\mathcal{K}}}_{\mathbf{r}\mathbf{r}'}^{(B)}(\mathbf{r} - \mathbf{r}') \end{aligned} \quad (82)$$

for the Wightman tensors $\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(E)}$ and $\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(B)}$.

From (81) and (82), we make the observation that the Wightman tensors $\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(E/B)}$ are intimately related to our kernels $\underline{\underline{\mathcal{K}}}_{\mathbf{r}\mathbf{r}'}^{(A/E/B)}$ — either via the application of a negative Laplacian or even directly identical to the Wightman tensor except for a dimensionful proportionality constant.

While the existence of a relation like this seems initially surprising, it is only partially so, since kernels can be seen as the field theoretical analog of covariance matrices for the Gaussian vacuum state. The Wightman tensors, in turn, collect all possible quadratic field correlation functions. Thus an intimate relationship between both quantities is to be expected.

6.3. Wightman tensor using functional integrals

While our approach to determine the explicit form of the Wightman tensor $\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}$ via mode expansion and vacuum wave functions was ultimately successful, it did not rely on wave functionals themselves. Thus, the task arises how similar questions can be framed and answered using the wave functional. We will now give a sketch using functional methods on how this might be achieved.

We start by recalling the relation [38, 50] between the functional integration measure and the field basis,

$$\int \mathcal{D}[\mathbf{F}] \equiv \prod_{\ell} \int_{-\infty}^{\infty} dF_{\ell}. \quad (83)$$

Moreover, we note that the field operators $\hat{\mathbf{F}} = \hat{\mathbf{F}}(t, \mathbf{r})$ and $\hat{\mathbf{F}}' = \hat{\mathbf{F}}(t, \mathbf{r}')$ can be expressed as functional Schrödinger integrals via

$$\hat{\mathbf{F}}(t, \mathbf{r}) = \int \mathcal{D}[\mathbf{F}] |\mathbf{F}\rangle \langle \mathbf{F}| \mathbf{F}(t, \mathbf{r}) \quad (84)$$

and

$$\hat{\mathbf{F}}'^{\dagger}(t, \mathbf{r}') = \int \mathcal{D}[\mathbf{F}'] |\mathbf{F}'\rangle \langle \mathbf{F}'| \mathbf{F}'^{\dagger}(t, \mathbf{r}'), \quad (85)$$

where $|\mathbf{F}\rangle \equiv |\{F_{\ell}\}\rangle = |F_1\rangle |F_2\rangle \dots$ and $|\mathbf{F}'\rangle \equiv |\{F'_{\ell}\}\rangle = |F'_1\rangle |F'_2\rangle \dots$ correspond to the state vectors of the field.

With these preliminaries settled, we recall the definition of the Wightman tensor (60)

$$\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}(t) = \langle \mathbf{0} | \hat{\mathbf{F}}(t, \mathbf{r}) \hat{\mathbf{F}}^{\dagger}(t, \mathbf{r}') | \mathbf{0} \rangle \quad (86)$$

and obtain, by inserting the operator expansions from (84) and (85), the double functional integral

$$\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}(t) = \int \mathcal{D}[\mathbf{F}] \int \mathcal{D}[\mathbf{F}'] \langle \mathbf{0} | \mathbf{F}\rangle \langle \mathbf{F} | \mathbf{F}'\rangle \langle \mathbf{F}' | \mathbf{0} \rangle \mathbf{F}(t, \mathbf{r}) \mathbf{F}'^{\dagger}(t, \mathbf{r}') \quad (87)$$

representation for the Wightman tensor of the field \mathbf{F} .

At first, this result appears to be too cumbersome for actual practical use. However, with the help of the functional Dirac delta function and the relation $\delta[\mathbf{F}-\mathbf{F}'] \equiv \prod_{\ell} \delta(F_{\ell}-F'_{\ell}) = \prod_{\ell} \langle \{F_{\ell}\} | \{F'_{\ell}\} \rangle = \langle \mathbf{F} | \mathbf{F}' \rangle$,

$$(88)$$

we can collapse one of the functional integrations in (87) and arrive at

$$\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}(t) = \int \mathcal{D}[\mathbf{F}] |\langle \mathbf{F} | \mathbf{0} \rangle|^2 \mathbf{F}(t, \mathbf{r}) \mathbf{F}^{\dagger}(t, \mathbf{r}'). \quad (89)$$

Note that in the process of collapsing the integration, only a relabeling due to the replacement $\mathbf{F}' \mapsto \mathbf{F}$ has taken place, while the spatial dependence on \mathbf{r}' , characteristic of a two-point correlation function in the expression, was completely retained.

At this point in the development of the functional approach, we are finally in the position to identify our wave functionals of the vacuum by

$$|\langle \mathbf{F} | \mathbf{0} \rangle|^2 \equiv \mathcal{N} \exp(-\beta^{(F)} \Sigma^{(F)}[\mathbf{F}]), \quad (90)$$

where we have made use of the mode sum in functional form (27), and defined the normalization constant $\mathcal{N} \equiv (\mathcal{N}^{(F)})^2$ of the functional.

As a consequence we are lead to a *single* functional integral representation

$$\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}(t) = \mathcal{N} \int \mathcal{D}[\mathbf{F}] e^{-\beta^{(F)} \Sigma^{(F)}[\mathbf{F}]} \mathbf{F}(t, \mathbf{r}) \mathbf{F}^{\dagger}(t, \mathbf{r}'), \quad (91)$$

for the Wightman tensor.

Moreover, the normalization constant can be expressed [32, 51] as another functional integral, namely

$$\mathcal{N}^{-1} = \mathcal{Z}^{(F)}[\mathbf{F}] \equiv \int \mathcal{D}[\mathbf{F}] \exp(-\beta_F \Sigma[\mathbf{F}]), \quad (92)$$

which we have named $\mathcal{Z}^{(F)}$ to allude to a close analogy with the partition sum in statistical physics.

In summary, we obtain the now purely functional expression

$$\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}(t) \equiv \frac{\int \mathcal{D}[\mathbf{F}] e^{-\beta_F \Sigma[\mathbf{F}]} \mathbf{F}(t, \mathbf{r}) \mathbf{F}^{\dagger}(t, \mathbf{r}')}{\mathcal{Z}^{(F)}[\mathbf{F}]} \quad (93)$$

for the Wightman function $\underline{\underline{\mathcal{W}}}_{\mathbf{r}\mathbf{r}'}^{(F)}$. This expression is the moment of a Gaussian functional integral [51] and can, in principle, be computed, similar to its distant cousin — the Gaussian integral in \mathbb{R}^n — by completing the square and calculating a (functional) determinant. However, since we are dealing with a vector field and not the usual case of a scalar field [32], things are a bit more complicated. Hence, we postpone this task together with the detailed discussion of how the partition sum (92) may be used together with functional differentiation as a generating functional to calculate more complex correlation functions.

7. Conclusions

Motivated by the thriving fields of cavity QED and circuit QED, we analyzed the wave functional of the vacuum in a resonator. We have found

expressions identical to those of free space discussed in the literature.

At first sight, this identity is surprising since the two situations differ considerably in the way the frequency of the mode enters into the mode expansion. In the continuous superposition of free space, it is the integration variable governed by the wave number. In the discrete case of the resonator, summation rather than integration extends over the mode indices, which in turn determines the mode frequency in a nontrivial way.

We were able to overcome this complication with the help of the introduction of the square root of the negative Laplacian. In this way, we could express the mode sum by the double integral of a bilinear form of the fields and of a scalar kernel given by the Fourier integral of the function reflecting the difference in the dependence of the vacuum fields on the wave number.

Moreover, our analysis emphasizes the important role of mode choice. Although modes have been eliminated in the wave functional, its form still depends on them. We have illustrated this phenomenon for the wave functional $\Psi[\mathbf{A}]$ of the vector potential \mathbf{A} , which involves either \mathbf{A} or $\nabla \times \mathbf{A}$ resulting from the \mathbf{u}_{ℓ} - or \mathbf{w}_{ℓ} -modes.

In hindsight of our calculation, one could argue that this is not as surprising as one might have thought. Especially since the wave functional for the *quantum state* of the vacuum fields is most naturally expressed in eigenmodes, as they correspond to physical degrees of freedom that *are quantized*. Once we retreat from employing explicit mode expansion, all the information that is left to fix the *quantum state* needs to be retained in the *associated kernel*.

We conclude by noting that despite the beauty of the wave functional, we are not aware of any application of evaluating, for example, the vacuum expectation values prevalent in QED. One elementary example of its usefulness could be the sum of modes appearing in the second moment of the displacement of an electron due to a vacuum electric field. This quantity determines the Lamb shift in the Welton picture [52] and leads to the Bethe logarithm.

Indeed, due to the integration of the second-order time derivative in the Lorentz equation, the displacement contains ω_{ℓ}^{-2} in the mode expansion of the electric field. Since we deal with the second moment, the electric field appears in a bilinear way, and actually, ω_{ℓ}^{-4} enters into the sum of the modes.

Moreover, the vacuum electric field is proportional to $\omega_{\ell}^{1/2}$, reducing to the bilinearity of the second moment of the displacement in the field the power to ω_{ℓ}^{-3} . When we replace the sum with an integration, the volume element contains ω_{ℓ}^2 leaving us with ω_{ℓ}^{-1} , creating, after the integration, the Bethe logarithm.

It would be interesting to see how this expression emerges from the use of the wave functional,

which would eliminate the need for performing the sum over the modes. For this purpose, we first note that the complication of the square of frequencies appearing in the mode expansion of the free field like ω_ℓ^{-2} can be removed by the use of the inverse of the negative Laplacian. Since we deal with the second moment, the electric field appears in a bilinear way, and functional integration with respect to the wave functional should yield an expression for the displacement in a straightforward way.

The result we obtained for the (electric field) Wightman tensor might be a first step in such a direction, as its elements contain all the necessary correlation functions for such a calculation. However, it is implicitly expected that it also has a singular behavior in the coincidence limit due to it being a derivative of a transverse delta function.

Unfortunately, this topic goes beyond the scope of the present article and has to be postponed to future publication.

Acknowledgments

It is a great honor and pleasure for us to dedicate our article to Professor Iwo Białynicki-Birula on the occasion of his 90th birthday. He taught us to love the wave functional of the vacuum and thereby triggered our curiosity about the corresponding quantity in a resonator, which constitutes the topic of our paper. We are enormously grateful to him for numerous stimulating and illuminating discussions about this and other problems over the last decades. Since our first joint article [53] on the quantum phase uncertainties, we have learned so much from him and are proud to be his friends. Happy Birthday Iwo, and many more healthy and happy years!

The authors are grateful to M. Keck, N. Rach, J. Seiler, E. Giese, R. Lopp and Ch. Ufrecht for many interesting and fruitful discussions. AF is grateful to R. Lopp for pointing him to [49].

AF is grateful to the Carl Zeiss Foundation (Carl-Zeiss-Stiftung) and IQST for funding in terms of the project MuMo-RmQM. The QUANTUS and INTENTAS projects are supported by the German Space Agency at the German Aerospace Center (Deutsche Raumfahrtagentur im Deutschen Zentrum für Luft- und Raumfahrt, DLR) with funds provided by the Federal Ministry for Economic Affairs and Climate Action (Bundesministerium für Wirtschaft und Klimaschutz, BMWK) due to an enactment of the German Bundestag under Grant Nos. 50WM2250D-2250E (QUANTUS+), as well as 50WM2177-2178 (INTENTAS).

Appendix A: Modes

In this appendix, we briefly summarize the key ingredients of the description of the electromagnetic

field in a resonator with discrete modes in the absence of charges and currents. We concentrate on the mode expansions and the energy of the electromagnetic field. Throughout this section and the article, we employ the Coulomb gauge. Although these expressions are well-established, we present them here for the sake of completeness.

A1: Mode functions and amplitudes

Central to our review of the electromagnetic field in a resonator are the Maxwell equations consisting of the two sets of equations

$$\nabla \cdot \mathbf{B} = 0 \quad \text{and} \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (94)$$

and

$$\nabla \cdot \mathbf{E} = 0 \quad \text{and} \quad \nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} \quad (95)$$

in the absence of currents and charges, where c denotes the speed of light.

We solve the homogenous equations by introducing the vector potential $\mathbf{A} = \mathbf{A}(t, \mathbf{r})$ in Coulomb gauge

$$\nabla \cdot \mathbf{A} = 0 \quad (96)$$

and the ansatz

$$\mathbf{E} \equiv -\frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} \equiv \nabla \times \mathbf{A}. \quad (97)$$

As a result, (95) implies the free-space wave equation

$$\square \mathbf{A}(t, \mathbf{r}) \equiv \left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right] \mathbf{A}(t, \mathbf{r}) = 0 \quad (98)$$

for the vector potential $\mathbf{A} \equiv \mathbf{A}(t, \mathbf{r})$, in the absence of currents and charges, where Δ is the three-dimensional Laplacian.

We emphasize that in the derivation of this wave equation, we have already used the Coulomb gauge condition (96) to simplify

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - \Delta \mathbf{A} = -\Delta \mathbf{A}. \quad (99)$$

Next, we make the separation ansatz

$$\mathbf{A}(t, \mathbf{r}) \equiv \mathcal{A} q(t) \mathbf{u}(\mathbf{r}) \quad (100)$$

with a real dimensionless spatial real function $\mathbf{u} = \mathbf{u}(\mathbf{r})$ and the real dimensionless time-dependent function $q = q(t)$. In order to ensure that \mathbf{A} has the appropriate units, we have introduced the constant \mathcal{A} . The vectorial nature of \mathbf{A} is contained in the function \mathbf{u} .

When we substitute the ansatz (100) into the wave equation (98), we arrive at the Helmholtz equation

$$\left[\Delta + \left(\frac{\omega}{c} \right)^2 \right] \mathbf{u}(\mathbf{r}) = 0, \quad (101)$$

and the harmonic oscillator equation

$$\ddot{q} + \omega^2 q = 0 \quad (102)$$

with frequency ω . Here, dots denote differentiation with respect to time.

We emphasize that the solutions of the Helmholtz equation (101) become unique once we specify a proper boundary condition. For example, we could choose

$$\mathbf{n}(\mathbf{r}) \times \mathbf{A}(t, \mathbf{r}) \equiv \mathbf{0} \quad (103)$$

for all points $\mathbf{r} \in \partial\mathcal{V}$ making up the cavity walls, which corresponds to a perfectly conducting cavity surface $\partial\mathcal{V}$ with normal vector $\mathbf{n}(\mathbf{r})$.

When we apply the Coulomb gauge condition (96) to the separation ansatz (100), we obtain the transversality constraint

$$\nabla \cdot \mathbf{u}(\mathbf{r}) = 0. \quad (104)$$

While we work in the classical theory this constraint is no issue, but as Paul Dirac first noticed [54], it can come to haunt us when we quantize electromagnetism [55–57] or any other gauge field [29].

The general solution of the harmonic oscillator equation (102) reads

$$q(t) = q_0 \cos(\omega t) + \frac{\dot{q}_0}{\omega} \sin(\omega t), \quad (105)$$

where we have introduced the arbitrary initial conditions $\dot{q}_0 \equiv \dot{q}(t=0)$ and $\dot{q}_0 \equiv \dot{q}(t=0)$.

The time-derivative of q leads us to

$$\dot{q} = \omega p \quad (106)$$

with

$$p \equiv p(t) = -\left(q_0 \sin(\omega t) - \frac{p_0}{\omega} \cos(\omega t)\right). \quad (107)$$

The boundary conditions imposed by the resonator enforce a discrete set of mode functions \mathbf{u} of the vector potential enumerated by a set of three indices [58] determining an effective wave vector. Moreover, due to the Coulomb gauge (96) and (104), we find two polarization directions for \mathbf{u} .

For the sake of implementing a concise notation, we abbreviate these indices consisting of wave vector *and* polarization indices by a single quantity ℓ , and use the set $\{\mathbf{u}_\ell\}$ for the eigenmodes of the vector potential.

A2: Vector potential

As a result of the linearity of the wave equation (98), the vector potential \mathbf{A} in the resonator is the superposition

$$\mathbf{A}(t, \mathbf{r}) = \sum_{\ell} A_{\ell}(t) \mathbf{u}_{\ell}(\mathbf{r}) \quad (108)$$

of all modes $\{\mathbf{u}_{\ell}\}$ which are the eigen-(mode) function of \mathbf{A} . Here we have introduced the abbreviation

$$A_{\ell}(t) \equiv \mathcal{A}_{\ell} q_{\ell}(t) \quad (109)$$

for the vector potential contribution originating from the mode \mathbf{u}_{ℓ} .

The mode functions $\{\mathbf{u}_{\ell}\}$ of the vector potential form an orthonormal basis of transverse vector fields inside the resonator with the orthogonality relation

$$\frac{1}{\mathcal{V}_{\ell}} \int d^3r \mathbf{u}_{\ell}^{\dagger}(\mathbf{r}) \mathbf{u}_m(\mathbf{r}) = \delta_{\ell m}, \quad (110)$$

where \mathcal{V}_{ℓ} denotes the mode volume.

A more general definition for the mode volume is for example given by

$$\bar{\mathcal{V}}_{\ell} \equiv \frac{\int d^3r |\mathbf{u}_{\ell}(\mathbf{r})|^2}{|\mathbf{u}_{\ell}(\mathbf{r}_c)|^2}, \quad (111)$$

where \mathbf{r}_c is a point of special interest of a given resonator.

For example, in a box resonator with perfectly reflecting and conducting surfaces exhibiting sinusoidal modes, one typically [48] picks \mathbf{r}_c as the point of maximal mode amplitude. Alternatively, in the presence of an atomic dipole at a fixed location inside the cavity, one can also use its position. Such choices can be directly linked to single-atom cavity QED analogs of the Purcell effect [59], i.e., the enhancement (or suppression) of the spontaneous emission rate of the dipole in a resonant cavity environment. For recent generalizations to more complicated systems and open cavities, we refer to [60, 61].

A3: Electric field

Since there are no charges and currents present, the electric field (97) in Coulomb gauge takes the explicit form

$$\mathbf{E}(t, \mathbf{r}) = -\sum_{\ell} \mathcal{A}_{\ell} \dot{q}_{\ell}(t) \mathbf{u}_{\ell}(\mathbf{r}), \quad (112)$$

where we made use of the mode expansion of the vector potential (108).

With the general solution (105) of the harmonic oscillator equation (102) and the connection (106) between \dot{q}_{ℓ} and p_{ℓ} , we find

$$\mathbf{E}(t, \mathbf{r}) = \sum_{\ell} \mathcal{E}_{\ell} p_{\ell}(t) \mathbf{u}_{\ell}(\mathbf{r}), \quad (113)$$

where we have introduced the relation

$$\mathcal{E}_{\ell} \equiv \mathcal{A}_{\ell} \omega_{\ell}. \quad (114)$$

Hence, the contribution of each mode to the total electric field is determined by the amplitude

$$E_{\ell}(t) \equiv \mathcal{E}_{\ell} p_{\ell}(t) \quad (115)$$

in the mode expansion

$$\mathbf{E}(t, \mathbf{r}) = \sum_{\ell} E_{\ell}(t) \mathbf{u}_{\ell}(\mathbf{r}). \quad (116)$$

A comparison of this expression to the expansion of the electric field

$$\mathbf{E}(t, \mathbf{r}) = \sum_{\ell} E_{\ell}(t) \mathbf{v}_{\ell}(\mathbf{r}) \quad (117)$$

in its eigenmodes $\{\mathbf{v}_{\ell}\}$, reveals that \mathbf{A} and the \mathbf{E} share the same set of eigenmodes $\{\mathbf{u}_{\ell}\}$. Consequently, the set $\{\mathbf{u}_{\ell}\}$ of modes of the vector potential can be mapped *one-to-one* to the set $\{\mathbf{v}_{\ell}\}$ of eigenmodes of the electric field. We emphasize that this property is only true in the absence of currents and charges, within and on the resonator boundary, because otherwise the wave equations for both fields \mathbf{A} and \mathbf{E} might differ in their boundary conditions and thus lead to different eigenmode expansions.

A4: Magnetic induction

We conclude this discussion of the fields by presenting a similar representation for the magnetic induction \mathbf{B} in terms of the mode functions of the vector potential \mathbf{A} . However, in contrast to the electric field \mathbf{E} , linked to \mathbf{A} by differentiation in time, the field \mathbf{B} is linked to the vector potential by taking the curl, that is a coordinate derivative.

Indeed, we find from the definition $\mathbf{B} \equiv \nabla \times \mathbf{A}$ of \mathbf{B} in terms of \mathbf{A} given by (108) the expression

$$\mathbf{B}(t, \mathbf{r}) = \sum_{\ell} \mathcal{A}_{\ell} q_{\ell}(t) [\nabla \times \mathbf{u}_{\ell}(\mathbf{r})]. \quad (118)$$

In order to bring out the analogy to \mathbf{E} , we multiply and divide in the expansion the mode function by ω_{ℓ}/c , which yields

$$\mathbf{B}(t, \mathbf{r}) = \sum_{\ell} B_{\ell}(t) \frac{c}{\omega_{\ell}} [\nabla \times \mathbf{u}_{\ell}(\mathbf{r})], \quad (119)$$

where we have introduced the magnetic induction in the mode

$$B_{\ell} \equiv \mathcal{B}_{\ell} q_{\ell}(t) \quad (120)$$

with the vacuum magnetic induction

$$\mathcal{B}_{\ell} \equiv \mathcal{A}_{\ell} \omega_{\ell} / c = \mathcal{E}_{\ell} / c. \quad (121)$$

In the last step, we have recalled from (114) the definition of the vacuum electric field.

When we compare (119) to the eigenmode expansion,

$$\mathbf{B}(t, \mathbf{r}) = \sum_{\ell} B_{\ell}(t) \mathbf{w}_{\ell}(\mathbf{r}) \quad (122)$$

of \mathbf{B} , we can again find a *one-to-one* mapping between eigenmodes. However, now we have to make the matching by comparing the expressions

$$\sum_{\ell} B_{\ell}(t) \frac{c}{\omega_{\ell}} [\nabla \times \mathbf{u}_{\ell}(\mathbf{r})] \stackrel{!}{=} \sum_{\ell} B_{\ell}(t) \mathbf{w}_{\ell}(\mathbf{r}). \quad (123)$$

When we note that there can be *no reshuffling* of the sequence of mode indices since only the coefficient $B_{\ell}(t)$ contributes to the field energy, the eigenmodes of \mathbf{B} must be related to the eigenmodes of \mathbf{A} by making the identification

$$\mathbf{w}_{\ell}(\mathbf{r}) \equiv \frac{c}{\omega_{\ell}} [\nabla \times \mathbf{u}_{\ell}(\mathbf{r})]. \quad (124)$$

However, when we recall that (eigen)-modes are determined by the boundary conditions resulting from (124), this is not surprising. The magnetic induction has to fulfill different boundary conditions to be consistent with Maxwell's equations on the resonator surface. We emphasize again that our elementary treatment is valid only in the absence of currents and charges within and on the resonator surface. Otherwise, significant changes can arise. For more details, we refer, for example, to the classic text [62] on nano-photonics, or more recent work referenced therein.

A5: Determination of the vacuum field amplitude

In order to define the quantity \mathcal{A}_{ℓ} , we recall from Appendix B that the energy

$$H(t) = \frac{\varepsilon_0}{2} \int d^3r \left[\mathbf{E}(t, \mathbf{r})^2 + (c\mathbf{B}(t, \mathbf{r}))^2 \right] \quad (125)$$

of the electromagnetic field in the resonator takes the form

$$H = \sum_{\ell} \varepsilon_0 \mathcal{A}_{\ell}^2 \omega_{\ell}^2 \frac{\mathcal{V}_{\ell}}{2} [p_{\ell}^2(t) + q_{\ell}^2(t)], \quad (126)$$

where we have used the expansions (112) and (119) for \mathbf{E} and \mathbf{B} .

When we compare (126) to the representation

$$H = \sum_{\ell} \frac{\hbar \omega_{\ell}}{2} [p_{\ell}^2(t) + q_{\ell}^2(t)] \quad (127)$$

of the total energy as a sum of all modes, where each mode contains the energy $\hbar \omega_{\ell}$, we obtain the explicit expression

$$\mathcal{A}_{\ell} \equiv \sqrt{\frac{\hbar}{\varepsilon_0 \omega_{\ell} \mathcal{V}_{\ell}}} \quad (128)$$

for the amplitude \mathcal{A}_{ℓ} of the vector potential due to a single mode.

Due to the connection (114) between \mathcal{E}_{ℓ} and \mathcal{A}_{ℓ} , we find the corresponding relation

$$\mathcal{E}_{\ell} \equiv \sqrt{\frac{\hbar \omega_{\ell}}{\varepsilon_0 \mathcal{V}_{\ell}}} \quad (129)$$

for the electric field. In the quantized theory, discussed in Appendix C, \mathcal{E}_{ℓ} will become the amplitude of the vacuum field.

In Table I, we summarize key features of the mode expansions based on the eigenmodes or the \mathbf{u}_{ℓ} -modes, such as the strength of the fields and the vacuum field amplitude in each mode. Here, we emphasize the different power laws of the mode frequency ω_{ℓ} in \mathcal{A}_{ℓ} , \mathcal{E}_{ℓ} , and \mathcal{B}_{ℓ} .

A6: Natural modes

In this appendix, we have expanded the three fields \mathbf{A} , \mathbf{E} , and \mathbf{B} into the modes \mathbf{u}_{ℓ} of \mathbf{A} . However, since we focus on a situation with no charges and currents, we can also express \mathbf{E} and \mathbf{B} in their natural modes, \mathbf{v}_{ℓ} and \mathbf{w}_{ℓ} . Indeed, \mathbf{E} and \mathbf{B} also satisfy the homogeneous wave equations, i.e.,

$$\square \mathbf{E}(t, \mathbf{r}) = \left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right] \mathbf{E}(t, \mathbf{r}) = 0 \quad (130)$$

and

$$\square \mathbf{B}(t, \mathbf{r}) = \left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right] \mathbf{B}(t, \mathbf{r}) = 0 \quad (131)$$

following from the Maxwell equations, (94) and (95), in the absence of currents and charges.

Needless to say, \mathbf{E} and \mathbf{B} have to obey boundary conditions imposed by the resonator, leading us to the natural modes $\mathbf{f}_{\ell} = \mathbf{f}_{\ell}(\mathbf{r})$ defined by the Helmholtz equation

$$(\Delta + k_{\ell}^2) \mathbf{f}_{\ell} = 0, \quad (132)$$

and the boundary conditions with $k_{\ell} = \omega_{\ell}/c$.

For the sake of simplicity, we have not included in the modes \mathbf{f}_{ℓ} a superscript A , E , or B to express the fact that they depend on the choice of the field.

Indeed, for \mathbf{A} and \mathbf{E} , the natural modes are obviously \mathbf{u}_ℓ , i.e.,

$$\mathbf{u}_\ell \equiv \mathbf{f}_\ell^{(A)} = \mathbf{f}_\ell^{(E)} \equiv \mathbf{v}_\ell, \quad (133)$$

but for \mathbf{B} , we find

$$\mathbf{f}_\ell^{(B)} \equiv k_\ell^{-1} (\nabla \times \mathbf{u}_\ell) \equiv \mathbf{w}_\ell. \quad (134)$$

The introduction of natural modes allows us to represent the mode expansions of all three fields by the single expression

$$\mathbf{F} = \sum_\ell F_\ell \mathbf{f}_\ell, \quad (135)$$

where \mathbf{F} denotes either \mathbf{A} , \mathbf{E} , or \mathbf{B} , and the modes \mathbf{f}_ℓ depend on the choice of \mathbf{F} .

Appendix B: Field energy in a resonator

In order to bring out most clearly the similarities and differences between the total energy H of the radiation field and the mode sum $\Sigma^{(F)}$ defining the wave functional of the vacuum and, in particular, the difference in the powers of the frequency of the mode in H and $\Sigma^{(F)}$, we re-derive in this appendix the energy

$$H = \frac{\varepsilon_0}{2} \int d^3r \left[\mathbf{E}^2 + (c\mathbf{B})^2 \right] \quad (136)$$

of the electromagnetic field in a resonator in two slightly different ways: (i) first, we calculate in typical textbook fashion the electric and magnetic contribution to the field energy, (ii) then, we use the previously defined eigenmodes of the field \mathbf{B} to find the magnetic contribution to the field energy.

B1: Textbook quantum optics approach

We begin with the textbook treatment, following along the lines of [48]. The contribution

$$H^{(E)} \equiv \frac{\varepsilon_0}{2} \int d^3r \mathbf{E}^2 \quad (137)$$

to H due to the electric field

$$\mathbf{E} = \sum_\ell E_\ell \mathbf{u}_\ell \quad (138)$$

leads us immediately to the expression

$$H^{(E)} = \frac{\varepsilon_0}{2} \sum_{\ell, \ell'} E_\ell E_{\ell'} \int d^3r \mathbf{u}_\ell(\mathbf{r}) \cdot \mathbf{u}_{\ell'}(\mathbf{r}), \quad (139)$$

which reduces with the orthonormality relation (110) of the modes to

$$H^{(E)} = \frac{\varepsilon_0}{2} \sum_\ell E_\ell^2 \mathcal{V}_\ell. \quad (140)$$

It is slightly more complicated to calculate the term

$$H^{(B)} \equiv \frac{\varepsilon_0}{2} \int d^3r (c\mathbf{B})^2 \quad (141)$$

associated with the magnetic induction

$$\mathbf{B} = \sum_\ell B_\ell \frac{c}{\omega_\ell} [\nabla \times \mathbf{u}_\ell]. \quad (142)$$

Indeed, when we substitute the mode representation (142) into $H^{(B)}$ given by (141), we find the expression

$$H^{(B)} = \frac{\varepsilon_0 c^2}{2} \sum_{\ell, \ell'} B_\ell B_{\ell'} \frac{c^2}{\omega_\ell \omega_{\ell'}} \mathcal{J}_{\ell\ell'}, \quad (143)$$

where we have introduced the abbreviation

$$\mathcal{J}_{\ell\ell'} \equiv \int d^3r [\nabla \times \mathbf{u}_\ell(\mathbf{r})] \cdot [\nabla \times \mathbf{u}_{\ell'}(\mathbf{r})]. \quad (144)$$

With the help of the identity proven in Appendix F, the integrand in (144) can be rewritten as

$$\begin{aligned} [\nabla \times \mathbf{u}_\ell] \cdot [\nabla \times \mathbf{u}_{\ell'}] &= \nabla \cdot [\mathbf{u}_{\ell'} \times (\nabla \times \mathbf{u}_\ell)] \\ &+ \mathbf{u}_{\ell'} \cdot [\nabla \times (\nabla \times \mathbf{u}_\ell)], \end{aligned} \quad (145)$$

where the first term on the right-hand side is a complete divergence. Hence, the application of the Gauss theorem converts the volume integral $\mathcal{J}_{\ell\ell'}$ into a surface integral, which vanishes due to the mode functions respecting the boundary conditions of the resonator.

The remaining term

$$\nabla \times (\nabla \times \mathbf{u}_\ell) = \nabla(\nabla \cdot \mathbf{u}_\ell) - \Delta \mathbf{u}_\ell \quad (146)$$

in (145) reduces with the Coulomb gauge condition (104) and the Helmholtz wave equation (101) to

$$[\nabla \times (\nabla \times \mathbf{u}_\ell)] = \left(\frac{\omega_\ell}{c}\right)^2 \mathbf{u}_\ell. \quad (147)$$

Hence, the integral $\mathcal{J}_{\ell\ell'}$ given by (144), yields

$$\mathcal{J}_{\ell\ell'} = \left(\frac{\omega_\ell}{c}\right)^2 \int d^3r \mathbf{u}_\ell(\mathbf{r}) \cdot \mathbf{u}_{\ell'}(\mathbf{r}) = \left(\frac{\omega_\ell}{c}\right)^2 \mathcal{V}_\ell \delta_{\ell\ell'}, \quad (148)$$

where in the last step we have used the orthonormality relation (110) of the mode functions.

Consequently, we arrive at the expression

$$H^{(B)} = \frac{\varepsilon_0}{2} \sum_\ell c^2 B_\ell^2 \mathcal{V}_\ell \quad (149)$$

for the magnetic field energy (143).

We conclude by combining the formulae for the electric $H^{(E)}$ and magnetic part $H^{(B)}$ given by (140) and (149), and arrive at the representation

$$H = \frac{\varepsilon_0}{2} \sum_\ell \mathcal{A}_\ell^2 \omega_\ell^2 \mathcal{V}_\ell (p_\ell^2 + q_\ell^2) \quad (150)$$

of the energy in terms of modes. Here, we have recalled the definitions (115) and (120) of E_ℓ and B_ℓ , respectively, together with the connections (114) and (121).

B2: Magnetic field energy via eigenmodes

When we recall our discussion of the respective eigenmodes of \mathbf{E} and \mathbf{B} , and their relation to the eigenmodes of \mathbf{A} , one might think that we could have avoided the cumbersome calculation of the

scalar product of the curls of the modes entirely. However, this suspicion is not quite true, and to show why, we perform the relevant calculation in this section.

When we expand the magnetic induction in its eigenmodes $\{\mathbf{w}_\ell\}$, we directly obtain for the magnetic field energy (141) the expression

$$H^{(B)} = \frac{\varepsilon_0}{2} \sum_{\ell, \ell'} c^2 B_\ell B_{\ell'} \int d^3r \mathbf{w}_\ell^\dagger(\mathbf{r}) \mathbf{w}_{\ell'}(\mathbf{r}). \quad (151)$$

Next, we make use of the orthonormality of the eigenmodes \mathbf{w}_ℓ , i.e.,

$$\frac{1}{\tilde{\mathcal{V}}_\ell} \int d^3r \mathbf{w}_\ell^\dagger(\mathbf{r}) \mathbf{w}_{\ell'}(\mathbf{r}) = \delta_{\ell\ell'}, \quad (152)$$

which leads us to the preliminary result

$$H^{(B)} = \frac{\varepsilon_0}{2} \sum_\ell c^2 B_\ell^2 \tilde{\mathcal{V}}_\ell. \quad (153)$$

We emphasize that, instead of the mode volume \mathcal{V}_ℓ of the vector potential modes \mathbf{u}_ℓ , the mode volume $\tilde{\mathcal{V}}_\ell$ corresponding to the eigenmodes \mathbf{w}_ℓ of \mathbf{B} has appeared. Hence, if one wants to express the total field energy H solely in terms of one mode volume, a connection between \mathcal{V}_ℓ and $\tilde{\mathcal{V}}_\ell$ is needed.

However, the only link available between the eigenmodes \mathbf{w}_ℓ and \mathbf{u}_ℓ is (124), i.e.,

$$\mathbf{w}_\ell(\mathbf{r}) = \frac{c}{\omega_\ell} [\nabla \times \mathbf{u}_\ell(\mathbf{r})]. \quad (154)$$

When we take the scalar product of this equation with itself and integrate over the resonator volume, we obtain the relation

$$\tilde{\mathcal{V}}_\ell = \int d^3r |\mathbf{w}_\ell(\mathbf{r})|^2 = \frac{c^2}{\omega_\ell^2} \int d^3r [\nabla \times \mathbf{u}_\ell(\mathbf{r})] \cdot [\nabla \times \mathbf{u}_\ell(\mathbf{r})]. \quad (155)$$

The integrand on the right-hand side of this equation is an old acquaintance of ours — (144) evaluated at $\ell = \ell'$.

Hence, even in the approach with the eigenmodes ultimately no true simplification is gained, but it is just a slightly different detour. As a consequence, we again need to apply (145) to (147) to simplify the scalar product of the two curls, and we obtain

$$\tilde{\mathcal{V}}_\ell = \int d^3r |\mathbf{w}_\ell(\mathbf{r})|^2 = \left(\frac{\omega_\ell}{c}\right)^2 \left(\frac{c}{\omega_\ell}\right)^2 \int d^3r |\mathbf{u}_\ell(\mathbf{r})|^2 = \mathcal{V}_\ell, \quad (156)$$

where we have made use of (110) defining the mode volume of the vector potential modes \mathbf{u}_ℓ .

As a consequence of the identity $\mathcal{V}_\ell = \tilde{\mathcal{V}}_\ell$, we also arrive at the expression

$$H^{(B)} = \frac{\varepsilon_0}{2} \sum_\ell c^2 B_\ell^2 \tilde{\mathcal{V}}_\ell = \frac{\varepsilon_0}{2} \sum_\ell c^2 B_\ell^2 \mathcal{V}_\ell \quad (157)$$

for the field energy $H^{(B)}$ due to the magnetic induction.

As an afterthought, we note that one could have naively imagined that the mode volumes might be defined independently such that they differ by a numeric factor — maybe via choosing different reference points in their respective definition of the mode volume. However, then the expression for the Hamiltonian (150) would be rescaled in the mode oscillator coordinate q_ℓ corresponding to the magnetic field by the factor of $\tilde{\mathcal{V}}_\ell/\mathcal{V}_\ell$. In turn, this feature would lead to problems in the Hamilton equations of motion since the symmetry between q_ℓ and p_ℓ would be broken, leading to a rescaled Poisson bracket. This would directly impact quantization by also rescaling the commutator $[\hat{q}_\ell, \hat{p}_\ell] = i$ by the factor $\tilde{\mathcal{V}}_\ell/\mathcal{V}_\ell$, which is undesirable. Nevertheless, we note that the simple argument we have formulated here might not be as clear-cut when complicated boundary conditions enter, or open resonators in the presence of currents and charges are considered.

Appendix C: Wave function representations of the ground state

In this appendix we derive the wave function ψ_ℓ of the ground state of the electromagnetic field in the modes \mathbf{u}_ℓ , \mathbf{v}_ℓ , or \mathbf{w}_ℓ specified by the mode index ℓ and the field. Indeed, for the vector potential \mathbf{A} and the electric field \mathbf{E} , the eigenmodes are \mathbf{u}_ℓ . However, for the magnetic induction \mathbf{B} they are $\mathbf{w}_\ell \equiv k_\ell^{-1} \nabla \times \mathbf{u}_\ell$.

Although the material in this appendix is partially contained in standard textbooks on quantum optics [48], we find it useful to include it in our article to gain a complete understanding of the origin and form of the dimensionless arguments of the Gaussian ground state wave function in the different representations. We first address in detail the case of \mathbf{E} , and then we will turn briefly to the analogous calculations for \mathbf{B} and \mathbf{A} .

C1: Electric field representation

We start from the mode decomposition

$$\mathbf{E}(t, \mathbf{r}) = \sum_\ell \mathcal{E}_\ell p_\ell(t) \mathbf{u}_\ell(\mathbf{r}) \quad (158)$$

of the electric field and make a transition to quantum mechanics, namely to the electric field operator $\hat{\mathbf{E}}$, by promoting the dimensionless amplitude functions q_ℓ and p_ℓ of the harmonic field oscillator of the ℓ -th mode defined by the mode function \mathbf{u}_ℓ to operators $q_\ell \mapsto \hat{q}_\ell$ and $p_\ell \mapsto \hat{p}_\ell$, and demanding the canonical commutation relations

$$[\hat{p}_\ell, \hat{q}_{\ell'}] = \frac{1}{i} \delta_{\ell\ell'}. \quad (159)$$

Hence, $\hat{\mathbf{E}}$ takes the form

$$\hat{\mathbf{E}}(t, \mathbf{r}) = \sum_\ell \hat{E}_\ell(t) \mathbf{u}_\ell(\mathbf{r}) \quad (160)$$

with

$$\hat{E}_\ell(t) \equiv \mathcal{E}_\ell \hat{p}_\ell(t), \quad (161)$$

which forces us to introduce a quantum state space for each mode.

A representative state could be, for example, the eigenstate $|E_\ell\rangle$ defined by the eigenvalue equation

$$\hat{E}_\ell |E_\ell\rangle \equiv E_\ell |E_\ell\rangle \quad (162)$$

for the electric field operator, where $E_\ell \equiv \mathcal{E}_\ell q_\ell$ corresponds to the eigenvalue. Thus, $|E_\ell\rangle$ describes a state, where the electric field in the ℓ -th mode assumes the well-defined value E_ℓ .

The ground state $|0_\ell\rangle$ of the ℓ -th field oscillator is determined by the condition

$$\hat{a}_\ell |0_\ell\rangle = 0, \quad (163)$$

where the linear combination

$$\hat{a}_\ell \equiv \frac{1}{\sqrt{2}}(\hat{q}_\ell + i\hat{p}_\ell) \quad (164)$$

of \hat{q}_ℓ and \hat{p}_ℓ represents the annihilation operator \hat{a}_ℓ .

When we now substitute the expression for \hat{a}_ℓ given by (164) into the definition (163) of the ground state and multiply by the bra-vector $\langle E_\ell|$, we arrive at the equation

$$\langle E_\ell| \hat{q}_\ell + i\hat{p}_\ell |0_\ell\rangle = 0 \quad (165)$$

determining the ground state wave function

$$\psi_\ell(E_\ell) \equiv \langle E_\ell | 0_\ell \rangle. \quad (166)$$

in the electric field representation, which corresponds to the first-order differential equation

$$\left[-\frac{1}{i} \frac{d}{d(E_\ell/\mathcal{E}_\ell)} + i(E_\ell/\mathcal{E}_\ell) \right] \psi_\ell(E_\ell) = 0 \quad (166)$$

Here we have used the fact that, according to (162), $|E_\ell\rangle$ is an eigenstate of \hat{E}_ℓ , and therefore of \hat{p}_ℓ , leading us to the identifications

$$\hat{p}_\ell \mapsto p_\ell \quad \text{and} \quad \hat{q}_\ell \mapsto -\frac{1}{i} \frac{d}{dp_\ell} \quad (168)$$

to satisfy the canonical commutation relation (159). Moreover, in (167), we have expressed the derivative with respect to p_ℓ by $\mathcal{E}_\ell p_\ell \equiv E_\ell$.

Hence, we arrive at the Gaussian wave function

$$\psi_\ell(E_\ell) = \mathcal{N}_\ell^{(E)} \exp \left[-\frac{1}{2} \left(\frac{E_\ell}{\mathcal{E}_\ell} \right)^2 \right], \quad (169)$$

where the normalization constant

$$\mathcal{N}_\ell^{(E)} \equiv \frac{1}{\sqrt[4]{\pi} \sqrt{\mathcal{E}_\ell}} \quad (170)$$

follows from the condition

$$\int_{-\infty}^{\infty} dE_\ell |\psi_\ell(E_\ell)|^2 = 1, \quad (171)$$

imposed by the Born interpretation.

C2: Magnetic induction representation

Next, we turn to the magnetic induction \mathbf{B} , where the corresponding operator reads

$$\hat{\mathbf{B}}(t, \mathbf{r}) = \sum_\ell \hat{\mathbf{B}}_\ell(t) \mathbf{w}_\ell(\mathbf{r}) \quad (172)$$

with

$$\hat{\mathbf{B}}_\ell(t) \equiv \mathcal{B}_\ell \hat{q}_\ell(t). \quad (173)$$

This decomposition leads us to the eigenvalue equation

$$\hat{\mathbf{B}}_\ell |B_\ell\rangle = B_\ell |B_\ell\rangle \quad (174)$$

for the state $|B_\ell\rangle$ of a well-defined value B_ℓ of the magnetic induction \mathbf{B} in the ℓ -th mode $\mathbf{w}_\ell(\mathbf{r}) \equiv k_\ell^{-1}(\nabla \times \mathbf{u}_\ell)$. Here, similarly to the electric field case, the expression

$$B_\ell \equiv \mathcal{B}_\ell q_\ell \quad (175)$$

denotes the eigenvalue.

Indeed, in this representation, we have to make the identification

$$\hat{p}_\ell \mapsto \frac{1}{i} \frac{d}{dq_\ell} \quad \text{and} \quad \hat{q}_\ell \mapsto q_\ell \quad (176)$$

leading us directly to the differential equation

$$\frac{d}{d(B_\ell/\mathcal{B}_\ell)} \psi_\ell(B_\ell) = -(B_\ell/\mathcal{B}_\ell) \psi_\ell(B_\ell) \quad (177)$$

for the wave function

$$\psi_\ell(B_\ell) \equiv \langle B_\ell | 0_\ell \rangle \quad (178)$$

of the ground state of the ℓ -th mode in the magnetic induction representation.

The differential equation (177) also admits a solution in the form of a Gaussian

$$\psi_\ell(B_\ell) \equiv \mathcal{N}_\ell^{(B)} \exp \left[-\frac{1}{2} \left(\frac{B_\ell}{\mathcal{B}_\ell} \right)^2 \right] \quad (179)$$

with the normalization constant

$$\mathcal{N}_\ell^{(B)} \equiv \frac{1}{\sqrt[4]{\pi} \sqrt{\mathcal{B}_\ell}}. \quad (180)$$

The only difference from the electric field representation discussed in the preceding section is the fact that the eigenstates $|B_\ell\rangle$ are now, apart from the vacuum fields \mathcal{B}_ℓ , eigenstates of \hat{q}_ℓ rather than of \hat{p}_ℓ .

C3: Vector potential representation

We conclude by briefly discussing the vector potential representation

$$\psi_\ell(A_\ell) \equiv \langle A_\ell | 0_\ell \rangle \quad (181)$$

of the ground state wave function in the ℓ -th mode resulting from the operator

$$\hat{A}(t, \mathbf{r}) \equiv \sum_\ell \hat{A}_\ell(t) \hat{u}_\ell(\mathbf{r}) \quad (182)$$

of the vector potential with

$$\hat{A}_\ell(t) \equiv \mathcal{A}_\ell \hat{q}_\ell(t). \quad (183)$$

Since the operator \hat{A}_ℓ like $\hat{\mathbf{B}}_\ell$ is also proportional to \hat{q}_ℓ , we immediately find

$$\psi_\ell(A_\ell) = \mathcal{N}_\ell^{(A)} \exp \left[-\frac{1}{2} \left(\frac{A_\ell}{\mathcal{A}_\ell} \right)^2 \right] \quad (184)$$

with the normalization constant

$$\mathcal{N}_\ell^{(A)} \equiv \frac{1}{\sqrt[4]{\pi} \sqrt{\mathcal{A}_\ell}} \quad (185)$$

in complete analogy to the distributions (169) and (179) in the electric field and magnetic induction variables E_ℓ and B_ℓ , respectively.

Appendix D: Reduction scheme for the kernel

In the main body of this article, we have derived an exact expression for the kernel $\underline{\mathcal{K}}$ of the wave functional of the vacuum in a resonator represented by the field \mathbf{F} in terms of the natural modes \mathbf{f}_ℓ . This kernel is a matrix according to (31), this kernel is a matrix, defined by the action of the function F containing the root of the negative Laplacian on the transverse delta function. Since the fields in the double integral are transverse, we can replace it with the familiar Dirac delta function of free space. As a result, the kernel reduces to a scalar $\mathcal{K}^{(F)}$.

In this appendix, we rederive the expression for the scalar kernel from a slightly different perspective. From the outset, we assume the kernel to be a scalar in the form of a Fourier representation of a root of the negative Laplacian. We first obtain an exact expression for the double integral containing the bilinear form of a field \mathbf{F} and the scalar kernel $\mathcal{K}^{(F)}$ expressed as a single integral of the square of F containing the fourth root of the negative Laplacian acting on \mathbf{F} . We then evaluate this integral for a given mode representation and match the result with the formula for the mode sum.

This procedure yields the individual scalar kernels. We conclude by comparing and contrasting this approach to the diagonal and non-diagonal representation of the density operator in terms of coherent states, and given by the P - and R -distribution [38], respectively.

D1: A general identity for Fourier transformable kernels

We now verify the identity

$$\tilde{\mathcal{I}}^{(F)} \equiv \int d^3r \int d^3r' \mathbf{F} \cdot \mathbf{F}' \mathcal{K}^{(F)}(\mathbf{r} - \mathbf{r}') = \int d^3r |F \sqrt[4]{-\Delta} \mathbf{F}|^2 \quad (186)$$

for a vector field $\mathbf{F} = \mathbf{F}(t, \mathbf{r})$, where the kernel

$$\mathcal{K}^{(F)}(\mathbf{r}) \equiv \frac{1}{(2\pi)^3} \int d^3k F(k) e^{i\mathbf{k} \cdot \mathbf{r}} \quad (187)$$

appears in the double integral with the difference $\mathbf{r} - \mathbf{r}'$ of the integration variables \mathbf{r} and \mathbf{r}' . Here F is not a generic scalar function but the function $F(k) = k$ or $F(k) = 1/k$ appearing in the mode sum $\Sigma^{(F)}$ defined by (13), and given for \mathbf{E} and \mathbf{B} by (14), and for \mathbf{A} by (15).

Central to the relation (186) is the eigenvalue equation (37) of $e^{i\mathbf{k} \cdot \mathbf{r}}$ leading us immediately to the representation

$$\mathcal{K}^{(F)}(\mathbf{r}) = F \left(\sqrt{-\Delta} \right) \delta(\mathbf{r}), \quad (188)$$

where we have recalled the Fourier representation (44) of the Dirac delta function.

When we substitute (188) into the left-hand side of (186), we arrive at the expression

$$\tilde{\mathcal{I}}^{(F)} \equiv \int d^3r \int d^3r' \mathbf{F} \cdot \left[F \left(\sqrt[4]{-\Delta} \right) \times F \left(\sqrt[4]{-\Delta} \right) \delta(\mathbf{r} - \mathbf{r}') \right] \mathbf{F}'. \quad (189)$$

Here we have used the relation

$$F \left(\sqrt{-\Delta} \right) \delta(\mathbf{r}) = F \sqrt[4]{-\Delta} F \sqrt[4]{-\Delta} \delta(\mathbf{r} - \mathbf{r}'), \quad (190)$$

which is only true for $F(k)=k$ and $F(k)=1/k$ and follows from the fact that the delta function is in the difference of the integration variables, i.e., $\mathbf{r} - \mathbf{r}'$.

When we recall that the field \mathbf{F} vanishes outside of the resonator, we can integrate both integrals by part. As a result, we arrive at the representation

$$\tilde{\mathcal{I}}^{(F)} = \int d^3r \int d^3r' \delta(\mathbf{r} - \mathbf{r}') \times \left[F \left(\sqrt[4]{-\Delta} \right) \mathbf{F} \right] \left[F \left(\sqrt[4]{-\Delta} \right) \mathbf{F}' \right] \quad (191)$$

of the integral $\tilde{\mathcal{I}}^{(F)}$. The Dirac delta function allows us to reduce the double integral into a single one leading us to the identity (186).

D2: Evaluation of the integral

Next, we evaluate the integral on the right-hand side of the identity (186) using the expansion

$$\mathbf{F} = \sum_{\ell} F_{\ell} \mathbf{f}_{\ell} \quad (192)$$

of \mathbf{F} into the natural modes \mathbf{f}_{ℓ} , and find

$$\int d^3r \left| F \left(\sqrt[4]{-\Delta} \right) \mathbf{F} \right|^2 = \sum_{\ell} F_{\ell}^2 F(k_{\ell}) \mathcal{V}_{\ell} = \Sigma^{(F)}, \quad (193)$$

where we have used the identity (29) for the action of the fourth root of the negative Laplacian on \mathbf{f}_{ℓ} , and the orthonormality relation (17). In the last step in (193), we used the identities $F(k) = 1/k$ and $F(k) = k$ and have recalled the definition (13) of the mode sum $\Sigma^{(F)}$.

Together with the identity (186), we finally arrive at the relation

$$\Sigma^{(F)} = \int d^3r \int d^3r' \mathbf{F} \cdot \mathbf{F}' \mathcal{K}^{(F)}(\mathbf{r} - \mathbf{r}') \quad (194)$$

with the kernels

$$\mathcal{K}^{(A)}(\mathbf{r}) \equiv \frac{1}{(2\pi)^3} \int d^3k k e^{i\mathbf{k} \cdot \mathbf{r}} \quad (195)$$

and

$$\mathcal{K}^{(E/B)}(\mathbf{r}) \equiv \frac{1}{(2\pi)^3} \int d^3k \frac{1}{k} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad (196)$$

in complete agreement with the derivation in Sect. 3.

D3: A curious analogy

This approach is reminiscent of the representation [48] of the density operator $\hat{\rho}$ in terms of coherent states. Multiplying the completeness relation of the coherent states from the left and from the right onto the density operator $\hat{\rho}$, we obtain the non-diagonal representation

$$\hat{\rho} = \frac{1}{\pi^2} \int d^2\alpha \int d^2\beta |\alpha\rangle \langle\alpha|\hat{\rho}|\beta\rangle \langle\beta|. \quad (197)$$

When we compare this expression to the corresponding one of the double integral $\tilde{\mathcal{I}}^{(F)}$, given by (186), we note three similarities: (i) the two different coherent states $|\alpha\rangle$ and $|\beta\rangle$ play the role of the fields \mathbf{F} and \mathbf{F}' , (ii) the matrix element $\langle\alpha|\hat{\rho}|\beta\rangle$ corresponds to the kernel, and (iii) the two integration over the coherent states translate into a double integral over coordinates.

Needless to say, there are also fundamental differences between the two expressions. For example, coherent states live in the state space and describe the *quantum mechanics* of a *single mode*. In contrast, the bilinear form involves the *classical total fields*. Nevertheless, in both cases, states and fields are associated with vector spaces and therefore take advantage of similar mathematical tools.

Roy Glauber and George Sudarshan, independently, introduced the diagonal representation

$$\hat{\rho} = \frac{1}{\pi} \int d^2\alpha P(\alpha) |\alpha\rangle \langle\alpha| \quad (198)$$

of the density operator $\hat{\rho}$ involving the P -distribution.

In our problem, this concept corresponds to the right-hand side of (186) which, according to (193), is identical to the mode sum $\Sigma^{(F)}$, which only contains the squares of the field strength and is therefore diagonal. This transition from non-diagonal to diagonal representation is made possible by derivatives acting on delta functions. Indeed, the P -distribution of a coherent state is already a Dirac delta function, and non-classical states are more singular [48].

Appendix E: Explicit expressions for kernels

In this appendix, we derive an explicit expression for the kernel

$$\mathcal{K}^{(j)}(\mathbf{r}) \equiv \frac{1}{(2\pi)^3} \int d^3k k^j e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (199)$$

and consider especially the two cases $j = 1$ and $j = -1$ corresponding to $\mathcal{K}^{(A)}$ and $\mathcal{K}^{(E/B)}$.

We note that while we formally calculate the integral for all integer values of j in this section, the resulting expressions and integrals are obviously problematic from the simple viewpoint of Riemann or Lebesgue integration of functions, since they are either singular at the origin or at infinity depending on the j value. Methods to deal with such singular integrals have been developed in the theory of

generalized functions [63–65] in terms of Hadamard finite part regularization. This is the framework in which the following calculation should be understood.

In the case of an integral with a singularity at the origin, the standard Hadamard regularization [63] can be directly applied. In the case of a singularity at infinity, tools with similar scope were developed in [66]. For an example of the necessary procedures, we refer to [64], where the regularization of $1/r^j$ is discussed in detail. In our calculation, we implicitly assume that such a regularization is performed and the kernel expressions are understood in this way. After the dust settles, the resulting kernel may be made sense of as a pseudo-function/generalized function induced by the meromorphic continuation of the remaining finite part, with the singular parts removed.

We begin the formal integration by choosing spherical coordinates $k \equiv |\mathbf{k}|$, ϑ and φ , noting that the integrand does not depend on φ . Thus we immediately arrive at the two-dimensional integral

$$\mathcal{K}^{(j)} = \frac{1}{(2\pi)^2} \int_0^\infty dk k^{j+2} \int_0^\pi d\vartheta \sin(\vartheta) e^{ikr \cos(\vartheta)}, \quad (200)$$

which after integration over ϑ yields the expression

$$\mathcal{K}^{(j)} = \frac{1}{(2\pi)^2} \frac{1}{r} \int_0^\infty dk k^{j+1} (e^{ikr} - e^{-ikr}). \quad (201)$$

Next we eliminate the power k^{j+1} by $j+1$ by differentiating the radial wave $\exp(\pm ikr)$ with respect to r in total $j+1$ -times and find

$$\begin{aligned} \mathcal{K}^{(j)} &= -\frac{1}{(2\pi)^2} \frac{1}{r} \frac{1}{i^j} \\ &\times \frac{\partial^{j+1}}{\partial r^{j+1}} \left[\int_0^\infty dk e^{ikr} + (-1)^j \int_0^\infty dk e^{-ikr} \right]. \end{aligned} \quad (202)$$

In order to evaluate the two remaining integrals, we introduce the convergence factor $\exp(-\epsilon k)$ to calculate the resulting integral, and then let $\epsilon > 0$ approach zero afterward. With the help of the relation

$$\int_0^\infty dk e^{-(\epsilon \mp ir)k} = \frac{1}{\epsilon \mp ir}, \quad (203)$$

we finally obtain

$$\mathcal{K}^{(j)}(\mathbf{r}) = \text{Pf} \frac{1}{2\pi r} \frac{(-1)}{i^j} \frac{\partial^{j+1}}{\partial r^{j+1}} d_\epsilon^{(j)}(r). \quad (204)$$

Here we have introduced the abbreviation

$$d_\epsilon^{(j)}(r) \equiv \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + r^2} \frac{1 + (-1)^j}{2} + \frac{i}{\pi} \frac{r}{\epsilon^2 + r^2} \frac{1 - (-1)^j}{2} \quad (205)$$

and added the pseudo-function [63, 64] operator Pf to remind us that the kernel is a pseudo-functions/generalized function resulting from implicitly performing Hadamard finite part regularization on the integral leading to it, if necessary.

With the representation

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + r^2} = \delta(r) \quad (206)$$

of the Dirac delta function and the identity

$$\lim_{\epsilon \rightarrow 0} \frac{r}{\epsilon^2 + r^2} = \mathcal{P}\left(\frac{1}{r}\right), \quad (207)$$

where \mathcal{P} denotes the Cauchy principal part, we obtain the expression

$$d^{(j)}(r) = \frac{1 + (-1)^j}{2} \delta(r) + \frac{i}{\pi} \frac{1 - (-1)^j}{2} \mathcal{P}\left(\frac{1}{r}\right). \quad (208)$$

Hence, for even values of j only the delta function contributes to

$$d_\epsilon^{(j)} \equiv \lim_{\epsilon \rightarrow 0} d_\epsilon^{(j)}, \quad (209)$$

whereas for odd ones only the contribution due to the derivatives of the Cauchy principal part appears, leading us to the expressions

$$\mathcal{H}^{(2n)} = \text{Pf} \frac{(-1)^{n+1}}{2\pi r} \frac{\partial^{2n+1}}{\partial r^{2n+1}} \delta(r) \quad (210)$$

and

$$\mathcal{H}^{(2n+1)} = \text{Pf} \frac{(-1)^{n+1}}{2\pi^2 r} \frac{\partial^{2(n+1)}}{\partial r^{2(n+1)}} \mathcal{P}\left(\frac{1}{r}\right). \quad (211)$$

Both kernel expressions should be understood as pseudo-functions including an implicit regularization lending the needed context [63] in which, e.g., the derivatives of the Cauchy principal part are to be interpreted. As is often done in physics, we will from now on suppress the pseudo-function operator again for brevity in the notation, assuming the resulting kernels and objects involving them are understood implicitly in that sense from now on.

With these considerations, after performing the derivatives for $j = -1$ that is $n = -1$, we find the kernel

$$\mathcal{H}^{(-1)} = \mathcal{H}^{(E/B)} = \frac{1}{2\pi^2 r^2}, \quad (212)$$

whereas for $j = +1$, that is $n = 0$, we arrive at the kernel

$$\mathcal{H}^{(1)} = \mathcal{H}^{(A)} = \frac{(-1)}{\pi^2 r^4}. \quad (213)$$

This expression for $\mathcal{H}^{(1)}$ also follows in a straight-forward way when we note from the definition (199) of $\mathcal{H}^{(j)}$ the connection

$$(-\Delta) \mathcal{H}^{(-1)} = \mathcal{H}^{(1)} \quad (214)$$

between $\mathcal{H}^{(-1)}$ and $\mathcal{H}^{(1)}$, i.e., between $\mathcal{H}^{(E/B)}$ and $\mathcal{H}^{(A)}$.

Indeed, by direct differentiation of (212), we obtain

$$\mathcal{H}^{(A)} = (-\Delta) \mathcal{H}^{(E/B)} = -\frac{1}{2\pi^2} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) \frac{1}{r^2} = -\frac{1}{\pi^2 r^4}, \quad (215)$$

in complete agreement with (213).

Appendix F:

Scalar product of two mode functions

The scalar product of the curls of two-mode functions is crucial for calculating the contribution $H^{(B)}$ of the magnetic induction to the total energy H of the electromagnetic field in a resonator performed in Appendix B.

In (145) we applied an identity for the scalar product of the curl of two vector fields, which we derive here. We start with a more general identity for the three vector fields $\mathbf{f} = \mathbf{f}(\mathbf{r}, \mathbf{r}')$ and $\mathbf{g} = \mathbf{g}(\mathbf{r}, \mathbf{r}')$ and $\mathbf{h} = \mathbf{h}(\mathbf{r}, \mathbf{r}')$.

When we take the divergence of the cross product between \mathbf{f} and \mathbf{h} , we obtain

$$\begin{aligned} \nabla_{\mathbf{r}} \cdot [\mathbf{f}(\mathbf{r}, \mathbf{r}') \times \mathbf{h}(\mathbf{r}, \mathbf{r}')] &= \mathbf{h} \cdot (\nabla_{\mathbf{r}} \times \mathbf{f}) \\ &\quad - \mathbf{f} \cdot (\nabla_{\mathbf{r}} \times \mathbf{h}), \end{aligned} \quad (216)$$

where from now on we suppress the functional dependencies of the fields for brevity.

Replacing $\mathbf{h} \mapsto \nabla_{\mathbf{r}'} \times \mathbf{g}$ yields

$$\begin{aligned} (\nabla_{\mathbf{r}} \times \mathbf{f}) \cdot (\nabla_{\mathbf{r}'} \times \mathbf{g}) &= \nabla_{\mathbf{r}} \cdot [\mathbf{f} \times (\nabla_{\mathbf{r}'} \times \mathbf{g})] \\ &\quad + \mathbf{f} \cdot [\nabla_{\mathbf{r}} \times (\nabla_{\mathbf{r}'} \times \mathbf{g})]. \end{aligned} \quad (217)$$

Using the definition of the cross product in terms of the Levi-Civita symbol, i.e., $\mathbf{a} \times \mathbf{b} = \mathbf{e}_j \epsilon_{jkl} a_k b_l$ with summation over double-indices implied, the terms on the right-hand side of the previous equation can be transformed into

$$\mathbf{f} \times (\nabla_{\mathbf{r}'} \times \mathbf{g}) = \nabla_{\mathbf{r}'} (\mathbf{f} \cdot \mathbf{g}) - (\mathbf{f} \cdot \nabla_{\mathbf{r}'}) \mathbf{g} \quad (218)$$

and

$$\nabla_{\mathbf{r}} \times (\nabla_{\mathbf{r}'} \times \mathbf{g}) = \nabla_{\mathbf{r}'} (\nabla_{\mathbf{r}} \cdot \mathbf{g}) - (\nabla_{\mathbf{r}} \cdot \nabla_{\mathbf{r}'}) \mathbf{g}. \quad (219)$$

Reinsertion of these identities into (217) leads to the desired identity

$$\begin{aligned} (\nabla_{\mathbf{r}} \times \mathbf{f}) \cdot (\nabla_{\mathbf{r}'} \times \mathbf{g}) &= \nabla_{\mathbf{r}} \cdot [\nabla_{\mathbf{r}'} (\mathbf{f} \cdot \mathbf{g})] \\ &\quad - \nabla_{\mathbf{r}} \cdot [(\mathbf{f} \cdot \nabla_{\mathbf{r}'}) \mathbf{g}] + (\mathbf{f} \cdot \nabla_{\mathbf{r}'}) (\nabla_{\mathbf{r}} \cdot \mathbf{g}) \\ &\quad - \mathbf{f} \cdot [(\nabla_{\mathbf{r}} \cdot \nabla_{\mathbf{r}'}) \mathbf{g}] \end{aligned} \quad (220)$$

for the scalar product of two curls with differentiation with respect to different arguments \mathbf{r} and \mathbf{r}' .

Alternatively, starting from (217) and using the case of $\mathbf{r} \equiv \mathbf{r}'$ and the definition of the vector Laplacian we obtain the identity

$$\begin{aligned} (\nabla_{\mathbf{r}} \times \mathbf{f}) \cdot (\nabla_{\mathbf{r}} \times \mathbf{g}) &= \nabla_{\mathbf{r}} \cdot [\mathbf{f} \times (\nabla_{\mathbf{r}} \times \mathbf{g})] \\ &\quad + \mathbf{f} \cdot [\nabla_{\mathbf{r}} (\nabla_{\mathbf{r}} \cdot \mathbf{g}) - \Delta_{\mathbf{r}} \mathbf{g}], \end{aligned} \quad (221)$$

used in (145).

References

- [1] E. Fermi, *Rev. Mod. Phys.* **4**, 87 (1932).
- [2] W.E. Lamb, *Appl. Phys. B* **60**, 77 (1995).
- [3] J.A. Wheeler, *Ann. Phys.* **2**, 604 (1957).
- [4] J.A. Wheeler, *Geometrodynamics*, Academic Press, London 1962.
- [5] C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation*, W.H. Freeman and Company, San Francisco 1973.
- [6] Z. Białynicka-Birula, I. Białynicki-Birula, *J. Opt. Soc. Am. B* **4**, 1621 (1987).
- [7] I. Białynicki-Birula, *Opt. Commun.* **179**, 237 (2000).
- [8] I. Białynicki-Birula, *Prog. Opt.* **36**, 245 (1996).
- [9] I. Białynicki-Birula, in: *Decoherence and Entropy in Complex Systems*, Springer, Berlin 2003, p. 287.
- [10] I. Białynicki-Birula, Z. Białynicka-Birula, [arXiv:2303.12183](https://arxiv.org/abs/2303.12183), 2023.
- [11] M. Born, W. Heisenberg, P. Jordan, *Z. Phys.* **35**, 557 (1926).
- [12] P.A.M. Dirac, *Proc. R. Soc. London A* **114**, 243 (1927).
- [13] W.E. Lamb, R.C. Retherford, *Phys. Rev.* **72**, 241 (1947).
- [14] H.M. Foley, P. Kusch, *Phys. Rev.* **73**, 412 (1948).
- [15] J. Schwinger, *Selected Papers on Quantum Electrodynamics*, Dover Publications, Dover 1958.
- [16] S.S. Schweber, *QED and the Men Who Made It*, Princeton University Press, Princeton (NJ) 1994.
- [17] I. Białynicki-Birula, Z. Białynicka-Birula, *Quantum Electrodynamics*, Pergamon Press, New York 1975.
- [18] S. Haroche, *Rev. Mod. Phys.* **85**, 1083 (2013).
- [19] H. Walther, B.T.H. Varcoe, B.-G. Englert, Th. Becker, *Rep. Prog. Phys.* **69**, 1325 (2006).
- [20] A. Blais, A.L. Grimsmo, S.M. Girvin, A. Wallraff, *Rev. Mod. Phys.* **93**, 025005 (2021).
- [21] A.S. Sheremet, M.I. Petrov, I.V. Iorsh, A.V. Poshakinskiy, A.N. Poddubny, *Rev. Mod. Phys.* **95**, 015002 (2023).
- [22] L. Landau, R. Peierls, *Z. Phys.* **69**, 56 (1931).
- [23] N. Bohr, L. Rosenfeld, *Kgl. Danske Vidensk. Selskab. Math.-Fys. Medd* **12**, 3 (1933).
- [24] N. Bohr, L. Rosenfeld, *Phys. Rev.* **78**, 794 (1950).
- [25] J.A. Wheeler, W.H. Zurek, *Quantum Theory and Measurement*, Princeton University Press, 2016.
- [26] H. Salecker, E.P. Wigner, *Phys. Rev.* **109**, 571 (1958).
- [27] K. Kuchař, *J. Math. Phys.* **11**, 3322 (1970).
- [28] J. Schwinger, *Phys. Rev.* **158**, 1391 (1967).
- [29] S. Weinberg, *The Quantum Theory of Fields - Volume 1: Foundations*, Cambridge University Press, Cambridge 1995.
- [30] T. Padmanabhan, *Quantum Field Theory*, Springer, Cham 2016.
- [31] R.W. Jackiw, “Schrödinger Picture Analysis of Boson and Fermion Quantum Field Theories”, Tech. Rep., MIT. Cent. Theor. Phys., Cambridge (MA) 1987.
- [32] B. Hatfield, *Quantum Field Theory of Point Particles And Strings*, CRC Press, Boca Raton 2018.
- [33] S. Bose, A. Mazumdar, G.W. Morley, H. Ulbricht, M. Toroš, M. Paternostro, A.A. Geraci, P.F. Barker, M.S. Kim, G. Milburn, *Phys. Rev. Lett.* **119**, 240401 (2017).
- [34] C. Marletto, V. Vedral, *Phys. Rev. Lett.* **119**, 240402 (2017).
- [35] D. Rickles, C.M. DeWitt, “The Role of Gravitation in Physics: Report from the 1957 Chapel Hill Conference”, Max-Planck Gesellschaft zur Förderung der Wissenschaften, Berlin 2011.
- [36] L.-Q. Chen, F. Giacomini, C. Rovelli, *Quantum* **7**, 958 (2023).
- [37] F. Dyson, *Int. J. Mod. Phys. A* **28**, 1330041 (2013).
- [38] R.J. Glauber, *Phys. Rev.* **131**, 2766 (1963).
- [39] G. Rasch, H. Boerner, *J. Reine Angew. Math.* **171**, 65 (1934).
- [40] H. Salecker, *Z. Naturforsch. A* **5**, 480 (1950).
- [41] I. Białynicki-Birula, *Acta Phys. Pol. B* **29**, 3569 (1998).
- [42] C. Anastopoulos, B.-L. Hu, *Entropy* **24**, 490 (2022).
- [43] S. Weinberg, *The Quantum Theory of Fields: Modern Applications*, Vol. 2, Cambridge University Press, Cambridge 1996.
- [44] S.N. Gupta, *Proc. Phys. Soc. A* **63**, 681 (1950).
- [45] K. Bleuler, *Helv. Phys. Acta* **23**, 567 (1950).
- [46] A. Fuster, M. Henneaux, A. Maas, *Int. J. Geom. Meth. Mod. Phys.* **2**, 939 (2005).
- [47] F. Falceto, *Ann. Henri Poincaré*, 1 (2022).
- [48] W.P. Schleich, *Quantum Optics in Phase Space*, Wiley-VCH, Weinheim 2001.

- [49] R. Lopp, E. Martín-Martínez, *Phys. Rev. A* **103**, 013703 (2021).
- [50] R.J. Glauber, *Quantum Theory of Optical Coherence: Selected Papers and Lectures*, Wiley, Weinheim 2007.
- [51] M.E. Peskin, D.V. Schroeder, *An Introduction To Quantum Field Theory*, Taylor & Francis, Andover 2018.
- [52] T.A. Welton, *Phys. Rev.* **74**, 1157 (1948).
- [53] I. Białynicki-Birula, M. Freyberger, W. Schleich, *Phys. Scr.* **T48**, 113 (1993).
- [54] P.A.M. Dirac, *Lectures on Quantum Mechanics: Quantization with Constraints*, Dover Publications, New York 2013.
- [55] R.G. Woolley, *Phys. Rev. Res.* **2**, 013206 (2020).
- [56] A. Stokes, A. Nazir, *Phys. Rev. A* **104**, 032227 (2021).
- [57] A. Stokes, A. Nazir, *Rev. Mod. Phys.* **94**, 045003 (2022).
- [58] K. Kakazu, Y.S. Kim, [arXiv:9511012](https://arxiv.org/abs/9511012), 1995.
- [59] E.M. Purcell, H.C. Torrey, R.V. Pound, *Phys. Rev.* **69**, 37 (1946).
- [60] E.A. Muljarov, W. Langbein, *Phys. Rev. B* **94**, 235438 (2016).
- [61] J. Ren, S. Franke, S. Hughes, *Phys. Rev. X* **11**, 041020 (2021).
- [62] J.D. Joannopoulos, S.G. Johnson, J.N. Winn, R.D. Meade, *Photonic Crystals*, Princeton University Press, Princeton, 2008.
- [63] A.H. Zemanian, *Distribution Theory and Transform Analysis: An Introduction to Generalized Functions, with Applications*, Dover Publications, New York 1987.
- [64] R. Estrada, R.P. Kanwal, *J. Math. Anal. Appl.* **141**, 195 (1989).
- [65] E.A. Galapon, *J. Math. Phys.* **57**, 033502 (2016).
- [66] D.S. Jones, *Math. Methods Appl. Sci.* **19**, 1017 (1996).

DEDICATED TO PROFESSOR IWO BIAŁYŃICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

Free Energy of Coupled Oscillators: Lamb Shifts and van der Waals Interactions

P.W. MILONNI*

Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S78](https://doi.org/10.12693/APhysPolA.143.S78)*e-mail: peter_milonni@comcast.net

The Helmholtz free energy of oscillators in thermal equilibrium with electromagnetic radiation is obtained from the Pauli–Hellmann–Feynman theorem and applied to some aspects of Lamb shifts and van der Waals interactions.

topics: Lamb shift, van der Waals interactions, Pauli–Hellmann–Feynman theorem

1. Introduction

Lamb shifts and van der Waals interactions may be attributed to the coupling of atoms to the zero-point electromagnetic field. These effects are modified at finite temperatures and depend on the mode structure of the field. Analyses of these effects have involved different formalisms and physical interpretations, all based in one way or another on quantum fluctuations of electromagnetic fields and their sources, and many invoking in particular the zero-point energy of the field or its finite-temperature generalization. Here we take an approach based on the Pauli–Hellmann–Feynman (PHF) theorem. We begin in Sect. 2 with brief, heuristic derivations of the (nonrelativistic) Lamb shift and the van der Waals interaction based on changes in zero-point field energy. In Sect. 3, we use the PHF theorem to derive an exact expression for the Helmholtz free energy of a system coupled to a heat bath, including many-body interactions. This is then applied in Sects. 3–6 to some aspects of Lamb shifts and van der Waals interactions, and in particular to the form of the van der Waals interaction when there is strong coupling to a single field mode. The physical interpretation of these results is briefly discussed in Sect. 7.

2. Scatterings: Lamb shift and van der Waals interaction at zero temperature

Sixty years ago, in a talk at the Relativity Conference in Warsaw, Richard Feynman [1] returned to an interpretation of the hydrogen Lamb shift he had suggested earlier [2, 3]. The argument, briefly, is as follows. In a box of volume V containing N identical atoms per unit volume, the zero-point energy of a field mode of frequency ω is $\frac{1}{2}\hbar\omega/n(\omega)$, where $n(\omega)$ is the refractive index. The change in the total zero-point energy due to the presence of the atoms is therefore

$$\begin{aligned}\Delta E &= 2V \int \frac{d^3k}{(2\pi)^3} \frac{\hbar\omega}{2} \left(\frac{1}{n} - 1 \right) \\ &\cong -\frac{\hbar}{\pi c^3} \int_0^\infty d\omega \omega^3 \alpha_0(\omega)\end{aligned}\quad (1)$$

in the case of a single atom ($NV=1$) with polarizability $\alpha_0(\omega)$, $n(\omega)\cong 1 + 2\pi N\alpha_0(\omega)$. If we use the Kramers–Heisenberg formula for $\alpha_0(\omega)$, subtract out the free-electron energy given by ΔE with $\alpha_0(\omega)=-e^2/(m\omega^2)$, and introduce a high-frequency cutoff mc^2/\hbar , we obtain, without any need for mass renormalization, exactly the “Bethe log” expression for the (nonrelativistic) Lamb shift [4, 5]. This is

discussed in a bit more detail in Sect. 4. A more explicit analysis based on Feynman's idea was done by E.A. Power [3].

The formula (1) can be expressed in terms of the forward scattering amplitude $f(\omega) = \alpha_0(\omega)\omega^2/c^2$

$$\Delta E = -2\pi\hbar c^2 \int \frac{d^3k}{(2\pi)^3} \frac{f(\omega)}{\omega}, \quad (2)$$

which is essentially Feynman's formula^{†1} [2]. It is equivalent to Bethe's, but it involves a scattering amplitude for a *real* scattering process, whereas Bethe's formula involves the single closed-loop diagram for the emission and absorption of *virtual* photons. Feynman remarked that the formula (2) is simple but "very peculiar. The reason it's peculiar is that the forward scatterings are real processes. At last I had discovered a formula I had always wanted, which is a formula for energy differences (which are defined in terms of virtual fields) in terms of actual measurable quantities..." [1].

A more direct calculation for an atom at position \mathbf{r}_A leads to

$$\begin{aligned} \Delta E(\mathbf{r}_A) = & -\frac{1}{2} \sum_{\mathbf{k}\lambda} \sqrt{\frac{2\pi\hbar\omega}{V}} \left| \hat{\mathbf{e}}_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{r}_A} \right|^2 \alpha_0(\omega) = \\ & -\frac{\hbar}{8\pi^2} \sum_{\lambda=1}^2 \int d^3k \omega \left| \hat{\mathbf{e}}_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{r}_A} \right|^2 \alpha_0(\omega), \end{aligned} \quad (3)$$

which of course is equivalent to (2). Here $\hat{\mathbf{e}}_{\mathbf{k}\lambda}$ is a linear polarization unit vector ($\mathbf{k} \cdot \hat{\mathbf{e}}_{\mathbf{k}\lambda} = 0$, $\lambda = 1, 2$). Now, suppose there is an identical atom B at a position \mathbf{r}_B , both atoms in their ground states. The effect on atom A is to replace $\hat{\mathbf{e}}_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{r}_A}$ by

$$\begin{aligned} \hat{\mathbf{e}}_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{r}_A} + \alpha_0(\omega) e^{i\mathbf{k}\cdot\mathbf{r}_B} k^3 e^{i\mathbf{k}\cdot\mathbf{r}} \left[\frac{\hat{\mathbf{e}}_{\mathbf{k}\lambda} - (\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}}}{kr} \right. \\ \left. + \left(\frac{i}{k^2 r^2} - \frac{1}{k^3 r^3} \right) (\hat{\mathbf{e}}_{\mathbf{k}\lambda} - 3(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}}) \right], \end{aligned} \quad (4)$$

where $\mathbf{r} = \mathbf{r}_A - \mathbf{r}_B$, $\hat{\mathbf{r}} = \mathbf{r}/r$, and $\mathbf{k} = k\hat{\mathbf{k}}$. The second term may be thought of as the field at A from the dipole moment induced in B by the vacuum field incident on B, i.e., it may be attributed to the scattering of the vacuum field by atom B. When we use this expression in place of $\hat{\mathbf{e}}_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{r}_A}$ in (3) and retain only terms up to second order in $\alpha_0(\omega)$, we obtain, in addition to the r -independent Lamb shift of atom A, an r -dependent energy

$$\begin{aligned} \Delta E(r) = & -\frac{\hbar c}{\pi r^2} \int_0^\infty du u^4 \alpha_0^2(icu) \\ & \times \left(1 + \frac{2}{ur} + \frac{5}{u^2 r^2} + \frac{6}{u^3 r^3} + \frac{3}{u^4 r^4} \right) e^{-2ur}, \end{aligned} \quad (5)$$

^{†1}Feynman added to the energy (2) a contribution from vacuum polarization, expressed similarly in terms of electron and positron forward scattering amplitudes.

a well-known expression for the van der Waals interaction of two molecules in a vacuum, neither of which has a permanent dipole moment. (We have used the analyticity of $\alpha_0(\omega)$ in the first quadrant of the complex frequency plane to analytically continue the integral along the positive real axis to an integral along the positive imaginary axis.) In the limit of very large separations, this gives the Casimir-Polder result $\Delta E(r) = -23\hbar c \alpha_0^2(0)/(4\pi r^7)$ for the retarded van der Waals interaction, whereas at small separations, it gives the London result in which $\Delta E(r) \propto 1/r^6$. Like the Lamb shift, the van der Waals interaction can be expressed in terms of a *real* scattering process and a forward scattering amplitude. The zero-point field is Rayleigh-scattered by each atom according to the expression (4), and the scattered field modifies the zero-point field at the other atom from its free-space form, resulting, in effect, in an r -dependent Lamb shift. This is the van der Waals interaction energy. The extension to many-atom systems, multiple scattering, and finite temperatures is perhaps most easily done with a simple extension of the Pauli-Hellmann-Feynman theorem (Sect. 3).

3. Free energy of atoms in thermal equilibrium with radiation

3.1. Pauli-Hellmann-Feynman theorem for free energy

Consider a Hamiltonian of the general form $H = H_0 + \lambda H_1$, where H_0 is the unperturbed Hamiltonian, and the interaction Hamiltonian is parametrized by a coupling constant λ . The eigenvalues $E(\lambda)$ and eigenvectors $|\psi(\lambda)\rangle$ of H will, of course, depend on λ . According to the Pauli-Hellmann-Feynman theorem [6-9],

$$\frac{dE}{d\lambda} = \left\langle \psi(\lambda) \left| \frac{dH}{d\lambda} \right| \psi(\lambda) \right\rangle. \quad (6)$$

In its integral form, the PHF theorem gives the change $E(1) - E(0)$ in the energy of the system in the form of the coupling-constant integration algorithm

$$E(1) - E(0) = \int_0^1 \frac{d\lambda}{\lambda} \langle \psi(\lambda) | \lambda H_1 | \psi(\lambda) \rangle, \quad (7)$$

the difference between the energy with ($\lambda = 1$) and without ($\lambda = 0$) the interaction H_1 [10]. For a nice discussion of the history of this "theorem", see [9].

In the case of thermal equilibrium, there is an expression analogous to (6), now involving the average $\langle \dots \rangle$ over the canonical ensemble, for the Helmholtz free energy $F(\lambda, T)$ [11, 12]

$$\frac{dF}{d\lambda} = \left\langle \frac{dH}{d\lambda} \right\rangle, \quad (8)$$

which follows simply from the definition

$$F(\lambda, T) = -k_B T \ln \left(\text{Tr} \left[e^{-H(\lambda)/(k_B T)} \right] \right). \quad (9)$$

Integration of (8) gives the change in the free energy in a form similar to the zero-temperature expression (7)

$$\Delta F = F(1, T) - F(0, T) = \int_0^1 \frac{d\lambda}{\lambda} \langle \lambda H_1 \rangle. \quad (10)$$

For additional information, see, for instance, [11] and references therein. For more discussion of the PHF theorem for finite temperatures, see [12].

3.2. Coupling of induced dipoles and thermal radiation

We now consider a collection of \mathcal{N} atoms coupled to a heat bath, specifically an electromagnetic field in thermal equilibrium at temperature T . The atoms are assumed to remain in their ground states with high probability. We assume there are no permanent dipole moments, only electric dipole moments induced by the field. The interaction Hamiltonian in the electric dipole approximation is

$$H_{\text{int}} = -\frac{1}{2} \sum_{n=1}^{\mathcal{N}} \sum_{i=1}^3 \left[p_i(\mathbf{r}_n, t) E_i(\mathbf{r}_n, t) + E_i(\mathbf{r}_n, t) p_i(\mathbf{r}_n, t) \right], \quad (11)$$

where $E_i(\mathbf{r}_n, t)$ is the i -th component of the electric field operator for the thermal field at the position \mathbf{r}_n of the atom with dipole moment $\mathbf{p}(\mathbf{r}_n, t)$. Effects of fields from the atoms themselves are subsumed in the polarizability, as done below. The coupling constant for the application of the PHF theorem is the electron charge e . We write $E_i(\mathbf{r}_n, t)$ in terms of positive- and negative-frequency components as

$$E_i(\mathbf{r}_n, t) = \int_0^{\infty} d\omega \left[E_i^{(+)}(\mathbf{r}_n, \omega) e^{-i\omega t} + E_i^{(-)}(\mathbf{r}_n, \omega) e^{i\omega t} \right], \quad (12)$$

and likewise,

$$p_i(\mathbf{r}_n, t) = \int_0^{\infty} d\omega \left[p_i^{(+)}(\mathbf{r}_n, \omega) e^{-i\omega t} + p_i^{(-)}(\mathbf{r}_n, \omega) e^{i\omega t} \right], \quad (13)$$

with

$$p_i^{(+)}(\mathbf{r}_n, \omega) = \alpha_0(\omega + i0^+) E_i(\mathbf{r}_n, \omega) \quad (14)$$

in the case of a single atom^{†2}. The polarizability $\alpha_0(\omega + i0^+)$ is given by the Kramers–Heisenberg formula

$$\alpha_0(\omega + i0^+) = \frac{2}{3\hbar} \sum_s \frac{\omega_{sg} |\mathbf{d}_{sg}|^2}{\omega_{sg}^2 - (\omega + i0^+)^2}, \quad (15)$$

where ω_{sg} (> 0) is the frequency for the transition between the ground state g and the excited state s and \mathbf{d}_{sg} is the corresponding electric dipole matrix element. For \mathcal{N} atoms, the dipole moment induced in every atom is

$$p_i^{(+)}(\mathbf{r}_n, \omega) = \alpha_0(\omega + i0^+) E_i^{(+)}(\mathbf{r}_n, \omega) + \alpha_0(\omega + i0^+) \times \sum_{m=1}^{\mathcal{N}} \sum_{j=1}^3 G_{ij}(\mathbf{r}_n, \mathbf{r}_m, \omega) p_j^{(-)}(\mathbf{r}_m, \omega). \quad (16)$$

The dyadic Green function $G(\mathbf{r}_n, \mathbf{r}_m, \omega)$ is defined by (66) in Appendix. In matrix form,

$$p^{(+)}(\omega) = \alpha_0(\omega + i0^+) E^{(+)}(\omega) + \alpha_0(\omega + i0^+) \times G(\omega) p^{(+)}(\omega) \quad (17)$$

or

$$p^{(+)}(\omega) = \frac{\alpha_0(\omega + i0^+) E^{(+)}(\omega)}{1 - \alpha_0(\omega + i0^+) G(\omega)} \equiv \alpha(\omega) E^{(+)}(\omega), \quad (18)$$

where $\alpha(\omega)$ and $G(\omega)$ are $3\mathcal{N} \times 3\mathcal{N}$ matrices and $p^{(+)}(\omega)$ is a $3\mathcal{N}$ -dimensional vector.

For thermal radiation the different frequency components of $\mathbf{E}(\mathbf{r}, t)$ are uncorrelated. As reviewed in Appendix,

$$\begin{cases} \langle E_i^{(+)}(\mathbf{r}_n, \omega) E_j^{(-)}(\mathbf{r}_m, \omega') \rangle = \frac{\hbar}{\pi} [q(\omega) + 1] G_{ij}^I(\mathbf{r}_n, \mathbf{r}_m, \omega) \delta(\omega - \omega'), \\ \langle E_i^{(-)}(\mathbf{r}_n, \omega) E_j^{(+)}(\mathbf{r}_m, \omega') \rangle = \frac{\hbar}{\pi} q(\omega) G_{ij}^I(\mathbf{r}_n, \mathbf{r}_m, \omega) \delta(\omega - \omega'), \end{cases} \quad (19)$$

where $G_{ij}^I(\mathbf{r}_n, \mathbf{r}_m, \omega)$ is the imaginary part of $G_{ij}(\mathbf{r}_n, \mathbf{r}_m, \omega)$ ($= G_{ji}(\mathbf{r}_m, \mathbf{r}_n, \omega)$) and $q(\omega) = [e^{\hbar\omega/(k_B T)} - 1]^{-1}$. Thus,

$$\begin{aligned} \langle H_{\text{int}} \rangle &= -\frac{\hbar}{\pi} \sum_{n,m=1}^{\mathcal{N}} \sum_{i,j=1}^3 \int_0^{\infty} d\omega \alpha_{ij}(\mathbf{r}_n, \mathbf{r}_m, \omega) [2q(\omega) + 1] G_{ji}^I(\mathbf{r}_m, \mathbf{r}_n, \omega) = \\ &= -\frac{\hbar}{\pi} \text{Im Tr} \left\{ \int_0^{\infty} d\omega \alpha(\omega) [2q(\omega) + 1] G(\omega) \right\} = -\frac{\hbar}{\pi} \text{Im Tr} \left\{ \int_0^{\infty} d\omega \frac{\alpha_0(\omega + i0^+) G(\omega)}{1 - \alpha_0(\omega + i0^+) G(\omega)} \coth \left(\frac{\hbar\omega}{2k_B T} \right) \right\}. \end{aligned} \quad (20)$$

^{†2}We do not include any line broadening effects that would give an imaginary part to the polarizability. In particular, in our formulation, radiation reaction is accounted for in (16)

but not in (15). But the polarizability must not have any poles in the upper half of the complex frequency plane, whence we add $i0^+$ to ω in the Kramers–Heisenberg formula.

3.3. Free energy

As noted above, the coupling constant for the application of the PHF theorem may be taken to be the electron charge e . Since $\alpha_0(\omega+i0^+)$ is proportional to e^2 , it follows from (20) that

$$\begin{aligned} \Delta F &= -\frac{\hbar}{\pi} \text{ImTr} \left\{ \int_0^\infty d\omega \coth \left(\frac{\hbar\omega}{2k_B T} \right) \right. \\ &\quad \times \int_0^1 \frac{d\lambda}{\lambda} \frac{\lambda^2 \alpha_0(\omega+i0^+) G(\omega)}{1 - \lambda^2 \alpha_0(\omega+i0^+) G(\omega)} \left. \right\} = \\ &= \frac{\hbar}{2\pi} \text{ImTr} \left\{ \int_0^\infty d\omega \coth \left(\frac{\hbar\omega}{2k_B T} \right) \right. \\ &\quad \times \log [1 - \alpha_0(\omega+i0^+) G(\omega)] \left. \right\}. \end{aligned} \quad (21)$$

Using the identity $\text{Tr}[\log(1-X)] = \log[\det(1-X)]$, we can write this as

$$\begin{aligned} \Delta F(T) &= -\frac{\hbar}{2\pi} \text{Im} \left\{ \int_0^\infty d\omega \coth \left(\frac{\hbar\omega}{2k_B T} \right) \right. \\ &\quad \times \log \det \left[\frac{\alpha(\omega)}{\alpha_0(\omega+i0^+)} \right] \left. \right\}, \end{aligned} \quad (22)$$

which has the form of the multi-particle generalization of the ‘‘remarkable theorem’’ of Ford, Lewis, and O’Connell [13] when we identify $\alpha(\omega) = \alpha_0(\omega+i0^+)/[1 - \alpha_0(\omega+i0^+)G(\omega)]$ as their ‘‘generalized susceptibility.’’ This formula gives the Helmholtz free energy of the *interacting* system of oscillators, in this case atoms and the electromagnetic field, in terms of the polarizability of the atoms alone. A different derivation is given in the original paper of Ford et al. [13]. Another derivation, based essentially on the PHF theorem but not in the form of the coupling-constant integration algorithm used here, is given in [11].

4. Lamb shifts

Retaining only the term linear in $\alpha_0(\omega+i0^+)$, (21) gives, for a single atom at any point \mathbf{r} in free space,

$$\begin{aligned} \Delta F(T) &= -\frac{\hbar}{2\pi} \text{Im} \left\{ \sum_{i=1}^3 \int_0^\infty d\omega \coth \left(\frac{\hbar\omega}{2k_B T} \right) \right. \\ &\quad \left. \alpha_0(\omega+i0^+) G_{ii}(\mathbf{r}, \mathbf{r}, \omega) \right\} = \\ &= -\frac{\hbar}{\pi c^3} \int_0^\infty d\omega \omega^3 \coth \left(\frac{\hbar\omega}{2k_B T} \right) \alpha_0(\omega+i0^+), \end{aligned} \quad (23)$$

since $\text{Im}[\text{lim}_{\mathbf{r} \rightarrow \mathbf{r}'} G(\mathbf{r}, \mathbf{r}', \omega)] = \frac{2\omega^2 k}{c^2} = \frac{2\omega^3}{c^3}$ as follows from (66) in Appendix. For $T=0$ this reproduces (1). Subtracting the free-electron ($\omega_{sg} \rightarrow 0$) contribution and introducing a high-frequency cutoff Ω , we replace (23) with the ‘‘observable’’ shift $\Delta F(0)_{\text{obs}}$, i.e., the difference in the shift between

bound and unbound electrons

$$\begin{aligned} \Delta F(0)_{\text{obs}} &= -\frac{2P}{3\pi c^3} \int_0^\Omega d\omega \omega^3 \sum_s \omega_{sg} |\mathbf{d}_{sg}|^2 \\ &\quad \times \left(\frac{1}{\omega_{sg}^2 - \omega^2} - \frac{1}{-\omega^2} \right) = \\ &= -\frac{2}{3\pi c^3} \sum_s \omega_{sg}^2 |\mathbf{d}_{sg}|^2 P \int_0^\Omega \frac{d\omega \omega}{\omega_{sg}^2 - \omega^2} = \\ &= -\frac{2}{3\pi c^3} \sum_s \omega_{sg}^2 |\mathbf{d}_{sg}|^2 \int_0^\Omega \frac{d\omega}{\omega + \omega_{sg}} \end{aligned} \quad (24)$$

for $\Omega \gg |\omega_{sg}|$ for all transition frequencies ω_{sg} (P stands for ‘‘principal part’’). This, of course, is the ‘‘Bethe log’’ when we take the high-frequency cutoff Ω to be mc^2/\hbar .

For an atom in a homogeneous dielectric medium, $\text{Im}[\text{lim}_{\mathbf{r} \rightarrow \mathbf{r}'} G(\mathbf{r}, \mathbf{r}', \omega)] = 2n(\omega)\omega^3/c^3$ and

$$\Delta F(0)_{\text{diel}} = -\frac{2}{3\pi c^3} \sum_s \omega_{sg}^2 |\mathbf{d}_{sg}|^2 \int_0^\Omega \frac{d\omega n(\omega)}{\omega_{sg} + \omega}. \quad (25)$$

The difference between the Lamb shift of an atom in the dielectric and the atom in vacuum is

$$\begin{aligned} \Delta F(0)_{\text{diel}} - \Delta F(0)_{\text{vac}} &= \\ &= -\frac{2}{3\pi c^3} \sum_s \omega_{sg}^2 |\mathbf{d}_{sg}|^2 \int_0^\Omega \frac{d\omega (n(\omega) - 1)}{\omega_{sg} + \omega}. \end{aligned} \quad (26)$$

Since $n(\omega)-1$ can be expected to vary as $1/\omega^2$ as $\omega \rightarrow \infty$, we can take $\Omega \rightarrow \infty$. In any event, it appears that this modified Lamb shift would be very difficult to observe because of competing effects and shifts resulting from the interaction of the guest atom with the host atoms of the medium.

Note that (23) implies a T -dependent correction to the Lamb shift

$$\begin{aligned} \Delta F_i(T) - \Delta F_i(0) &= -\frac{4}{3\pi c^3} \sum_j |\mathbf{d}_{ij}|^2 P \\ &\quad \times \int_0^\infty \frac{d\omega \omega^3}{e^{\hbar\omega/k_B T} - 1} \frac{\omega_{ji}}{\omega_{ji}^2 - (\omega+i0^+)^2} \end{aligned} \quad (27)$$

for an atom in state i . For transition frequencies and temperatures such that $\hbar|\omega_{ji}| \ll k_B T$ [14],

$$\begin{aligned} \Delta F_i(T) - \Delta F_i(0) &\cong \frac{4}{3\pi c^3} \sum_j |\mathbf{d}_{ij}|^2 \omega_{ji} \\ &\quad \times \int_0^\infty \frac{d\omega \omega}{e^{\hbar\omega/k_B T} - 1} = \frac{\pi e^2}{3m\hbar c^3} (k_B T)^2, \end{aligned} \quad (28)$$

where we have used the Thomas–Reiche–Kuhn sum rule. This is just the average kinetic energy obtained from the equation of motion $m\ddot{\mathbf{x}} = e\mathbf{E}$ for an electron in a blackbody field at temperature T . Temperature-dependent corrections to the Lamb shift of Rydberg atoms have been measured and found to be consistent with a T^2 scaling [15].

5. Van der Waals Interactions

The polarizability $\alpha(\omega)$ is required from causality considerations to be analytic in the upper half of the complex frequency plane. From the definition of $G(\omega)$ it is clear that $\alpha_0(\omega+i0^+)G(\omega)$ is analytic in the upper half of the complex frequency plane. Assuming for now that $\log[1 - \alpha_0(\omega+i0^+)G(\omega)]$ is likewise analytic, we can analytically continue the integral in (21) and express the (free) energy for $T = 0$ as

$$\Delta F = \frac{\hbar}{2\pi} \text{Tr} \left\{ \int_0^\infty d\xi \log \left[1 - \alpha_0(i\xi) G(i\xi) \right] \right\}. \quad (29)$$

Considering only the contribution that goes as $\alpha_0^2(i\xi)$, and ignoring the self-energy terms with $\mathbf{r}_n = \mathbf{r}_m$, we obtain

$$\Delta F_2 = -\frac{\hbar}{4\pi} \sum_{n=1}^{\mathcal{N}} \sum_{m=1}^{\mathcal{N}} (1 - \delta_{mn}) \text{Tr} \left\{ \int_0^\infty d\xi \alpha_0^2(i\xi) G_{ij}(\mathbf{r}_n, \mathbf{r}_m, i\xi) G_{ji}(\mathbf{r}_m, \mathbf{r}_n, i\xi) \right\}, \quad (30)$$

which is found to be just the sum of pairwise van der Waals interaction energies given by (5). In particular, for small separations, the (nonretarded) van der Waals interaction between two ground-state atoms with polarizabilities $\alpha_1(\omega)$ and $\alpha_2(\omega)$ has the well-known form originally obtained by London

$$\begin{aligned} \Delta E &= -\frac{3\hbar}{\pi r^6} \int_0^\infty d\xi \alpha_1(i\xi) \alpha_2(i\xi) = -\frac{3\hbar}{\pi r^6} \left(\frac{2}{3\hbar} \right)^2 \sum_m \sum_n |\mathbf{d}_{1m}|^2 |\mathbf{d}_{2n}|^2 \omega_{1m} \omega_{2n} \int_0^\infty \frac{d\xi}{(\omega_{1m}^2 + \xi^2)(\omega_{2n}^2 + \xi^2)} = \\ &= -\frac{2}{3\hbar r^6} \sum_m \sum_n \frac{|\mathbf{d}_{1m}|^2 |\mathbf{d}_{2n}|^2}{\omega_{1m} + \omega_{2n}}, \end{aligned} \quad (31)$$

where $\omega_{\mu n}$ ($\mu = 1, 2$) are the transition frequencies between the ground state and the state n and $\mathbf{d}_{\mu n}$ are the corresponding transition moments. More generally (29) accounts for many-body interactions and retardation.

It may be worth noting that, since the magnitude of the static polarizability α_{st} is roughly on the order of an atomic radius, we require that $\alpha_{1\text{st}} \alpha_{2\text{st}} / r^6 = \alpha_{1\text{st}} \alpha_{2\text{st}} G_{ij}(\mathbf{r}_1, \mathbf{r}_2, \omega) G_{ji}(\mathbf{r}_2, \mathbf{r}_1, \omega) < 1$ for small $r = |\mathbf{r}_1 - \mathbf{r}_2|$; otherwise overlap of the atomic wavefunctions must be considered, which we have not done. This condition can also be understood from the requirement that the Hamiltonian must be bounded from below [16].

Renne [17] obtained a formula similar to (29) based on the zero-point energy of coupled harmonic oscillators, each having a frequency ω_0 . Consider (16) without the first term on the right-hand side and without allowing for the coupling of each oscillator to its own field

$$\begin{aligned} p_i^{(+)}(\mathbf{r}_n, \omega) &= \alpha_0(\omega+i0^+) \\ &\times \sum_{m \neq n}^{\mathcal{N}} G_{ij}(\mathbf{r}_n, \mathbf{r}_m, \omega) p_j^{(+)}(\mathbf{r}_m, \omega), \end{aligned} \quad (32)$$

or, in matrix form,

$$\left[1 + \alpha_0(\omega+i0^+) \mathcal{T}(\omega) \right] p^{(+)}(\omega) = 0, \quad (33)$$

where $\mathcal{T}_{ij}(\mathbf{r}_n, \mathbf{r}_m, \omega) = -(1 - \delta_{mn}) G_{ij}(\mathbf{r}_n, \mathbf{r}_m, \omega)$. The condition for a non-trivial solution of this set of $3\mathcal{N}$ equations is that the “normal-mode” frequencies ω must satisfy

$$f(\omega) = \det \left[1 + \alpha_0(\omega+i0^+) \mathcal{T}(\omega) \right] = 0. \quad (34)$$

Solutions of this equation in which all values of ω are real can be obtained in the nonretarded regime. In this case, Renne has used the argument theorem to obtain the sum of the zeros ω_s of $f(\omega)$, and he identifies $\sum_s \frac{1}{2} \hbar \omega_s$ as the zero-point energy of the system of oscillators coupled to each other by their electrostatic dipole interactions. The difference ΔE between this zero-point energy and the zero-point energy $\frac{3}{2} \mathcal{N} \hbar \omega_0$ of the uncoupled oscillators is shown to be

$$\begin{aligned} \Delta E &= \frac{\hbar}{2\pi} \int_0^\infty d\xi \log \det \left[1 + \alpha_0(i\xi) \mathcal{T} \right] = \\ &= \frac{\hbar}{2\pi} \text{Tr} \left\{ \int_0^\infty d\xi \log \left[1 + \alpha_0(i\xi) \mathcal{T} \right] \right\}, \end{aligned} \quad (35)$$

which is very similar to (29) except that self-interactions are excluded. Renne proceeds to generalize this expression to allow for retardation, and his result is equivalent, except for the Lamb shifts, to (21) with $T = 0$. (The temperature dependence of van der Waals interactions has been studied by several authors — see, for instance, [16] and references therein.)

6. Strong coupling of molecules to a single cavity mode

There has recently been much interest in modifications of molecular interactions when there is a strong coupling of the molecules to a single cavity mode. Haugland et al. [18], for instance, have shown in nonperturbative numerical studies that the distance dependence of van der Waals interactions is significantly affected by such coupling [18, 19]. They

also present an illustrative perturbation-theoretic approach based on a Hamiltonian that includes the short-distance intermolecular dipole-dipole interaction

$$V_{AB} = -\frac{1}{r^3} \left[\mathbf{d}_A \cdot \mathbf{d}_B - 3(\mathbf{d}_A \cdot \hat{\mathbf{r}})(\mathbf{d}_B \cdot \hat{\mathbf{r}}) \right] \quad (36)$$

between molecules A and B, which are assumed to have no permanent dipole moments. The alteration of the van der Waals interaction occurs as a result of the additional coupling of the molecules to the vacuum single-mode field. This follows from the PHF theorem, as we now show with a model of \mathcal{N} two-state atoms interacting with a vacuum single-mode cavity field of frequency ω and polarization $\hat{\mathbf{e}}$, and with each other via

$$\begin{aligned} V = & - \sum_{n=1}^{\mathcal{N}} \sum_{m=1}^{\mathcal{N}} \frac{\mathbf{d}_n \cdot \mathbf{d}_m - 3(\mathbf{d}_n \cdot \hat{\mathbf{r}}_{nm})(\mathbf{d}_m \cdot \hat{\mathbf{r}}_{nm})}{r_{nm}^3} \\ & \times \left[(\sigma_n + \sigma_n^\dagger)(\sigma_m + \sigma_m^\dagger) \right] = \\ & - \sum_{n=1}^{\mathcal{N}} \sum_{m=1}^{\mathcal{N}} V_{nm} \left[(\sigma_n + \sigma_n^\dagger)(\sigma_m + \sigma_m^\dagger) \right], \quad (37) \end{aligned}$$

where $r_{nm} = |\mathbf{r}_n - \mathbf{r}_m|$ and σ_n and σ_n^\dagger are respectively the two-state lowering and raising operators. The transition frequencies and dipole matrix elements of the atoms are denoted by ω_n and \mathbf{d}_n . The Hamiltonian for the interaction of the atoms with the cavity field in the electric dipole approximation has the form

$$H_c = - \sum_{n=1}^{\mathcal{N}} C_n (a + a^\dagger)(\sigma_n + \sigma_n^\dagger), \quad (38)$$

where

$$C_n = A_n (\mathbf{d}_n \cdot \hat{\mathbf{e}}) \sqrt{\hbar\omega} \quad (39)$$

and A_n is a mode function that depends on the position \mathbf{r}_n of atom n in the cavity. The complete Hamiltonian is

$$H = \sum_{n=1}^{\mathcal{N}} \hbar\omega_n \sigma_n^\dagger \sigma_n + \hbar\omega a^\dagger a + H_c + V. \quad (40)$$

We proceed as in Sect. 3.2. The solution of the Heisenberg equation of motion for $\sigma_n(t)$, omitting the freely evolving part that plays no role in what follows, is

$$\begin{aligned} \sigma_n(t) = & \frac{i}{\hbar} C_n \int_{-\infty}^t dt' [a(t') + a^\dagger(t')] e^{i\omega_n(t'-t)} \\ & + \frac{i}{\hbar} \sum_m V_{nm} \int_{-\infty}^t dt' [\sigma_m(t') + \sigma_m^\dagger(t')] e^{i\omega_n(t'-t)}. \quad (41) \end{aligned}$$

Since ground-state atoms can be treated effectively as harmonic oscillators for our purposes, we have assumed the commutation relation $[\sigma_\mu(t), \sigma_\nu^\dagger(t)] = \delta_{\mu\nu}$. Now to the lowest order in the coupling constants,

$$\begin{aligned} a(t') & \cong a(t) e^{-i\omega(t'-t)}, \\ \sigma_m(t') & \cong \sigma_m(t) e^{-i\omega_m(t'-t)}. \quad (42) \end{aligned}$$

It then follows from (41) and some simple algebra that

$$\sigma_{xn}(t) \cong \mathcal{E}_n(t) + \sum_{m=1}^{\mathcal{N}} V_{nm} \sigma_{xm}(t), \quad (43)$$

where we have defined

$$\mathcal{E}_n(t) = \frac{2C_n}{\hbar} \frac{\omega_n}{\omega_n^2 - \omega^2} [a(t) + a^\dagger(t)], \quad (44)$$

$$V_{nm} = \frac{2V_{nm}}{\hbar} \frac{\omega_n}{\omega_n^2 - \omega_m^2} \quad (45)$$

and $\sigma_{xn} = \sigma_n + \sigma_n^\dagger$.

From the Heisenberg equation of motion for $a(t)$,

$$\begin{aligned} a(t) = & a_0(t) + \frac{iC_1}{\hbar} \int_{-\infty}^t dt' [\sigma_1(t') + \sigma_1^\dagger(t')] e^{i\omega(t'-t)} \\ & + \frac{iC_2}{\hbar} \int_{-\infty}^t dt' [\sigma_2(t') + \sigma_2^\dagger(t')] e^{i\omega(t'-t)} \cong \\ & a_0(t) + \frac{C_1}{\hbar} \left(\frac{\sigma_1(t)}{\omega - \omega_1} + \frac{\sigma_1^\dagger(t)}{\omega + \omega_1} \right) \\ & + \frac{C_2}{\hbar} \left(\frac{\sigma_2(t)}{\omega - \omega_2} + \frac{\sigma_2^\dagger(t)}{\omega + \omega_2} \right) \quad (46) \end{aligned}$$

in the approximation $\sigma_\mu(t') \cong \sigma_\mu(t) e^{-i\omega_\mu(t'-t)}$, with $a_0(t)$ the freely evolving annihilation operator for the single-mode cavity field. Likewise

$$\begin{aligned} \sigma_1(t) & \cong \sigma_{10}(t) + \frac{C_1}{\hbar} \left(\frac{a_0(t)}{\omega_1 - \omega} + \frac{a_0^\dagger(t)}{\omega_1 + \omega} \right) \\ & - \frac{V_{12}}{\hbar} \left(\frac{\sigma_2(t)}{\omega_1 - \omega_2} + \frac{\sigma_2^\dagger(t)}{\omega_1 + \omega_2} \right), \quad (47) \end{aligned}$$

and

$$\begin{aligned} \sigma_2(t) & \cong \sigma_{20}(t) + \frac{C_2}{\hbar} \left(\frac{a_0(t)}{\omega_2 - \omega} + \frac{a_0^\dagger(t)}{\omega_2 + \omega} \right) \\ & - \frac{V_{12}}{\hbar} \left(\frac{\sigma_1(t)}{\omega_2 - \omega_1} + \frac{\sigma_1^\dagger(t)}{\omega_1 + \omega_2} \right), \quad (48) \end{aligned}$$

with $\sigma_{\mu 0}(t)$ the freely evolving lowering operator for atom μ .

For the application of the PHF theorem, we require the expectation value of H_c for the state $|\psi\rangle$ in which the atoms are in their ground states and the cavity field is in its vacuum state. Considering only atom 1, for instance,

$$\langle H_c^{(1)} \rangle = -C_1 \langle \sigma_1^\dagger a + a^\dagger \sigma_1 \rangle - C_1 \langle a \sigma_1 + \sigma_1^\dagger a^\dagger \rangle, \quad (49)$$

where we have used the fact that the atom and field operators commute, as assumed when writing the Heisenberg equations. Consider first the first term on the right-hand side of (49). Since $a_0(t)|\psi\rangle = 0$, the only nonvanishing part of this term would have to come from the last two terms on the right-hand side of the expression (46). But these do not contribute to $-C_1 \langle \sigma_1^\dagger a + a^\dagger \sigma_1 \rangle$ any terms involving $\langle \sigma_\mu(t) \sigma_\mu^\dagger(t) \rangle = 1$ for ground-state atoms, only terms such as $\langle \sigma_1(t) \sigma_1(t) \rangle = \langle \sigma_1(t) \sigma_2^\dagger(t) \rangle = 0$. Thus, the first term on the right-hand side of (49) vanishes within

the approximations we have made, and so

$$\langle H_c^{(1)} \rangle = -C_1 \langle a \sigma_1(t) + \sigma_1^\dagger(t) a^\dagger(t) \rangle \cong -2C_1 \langle a(t) \sigma_1(t) \rangle. \quad (50)$$

We will make the simplifying assumption in this illustrative model that $\omega \gg \omega_1, \omega_2$, as would be the case, for instance, for infrared transitions in an optical cavity. Then

$$\sigma_2(t) \cong \sigma_{20}(t) - \frac{C_2}{\hbar \omega} (a_0(t) - a_0^\dagger(t)) - \frac{V_{12}}{\hbar} \left(\frac{\sigma_1(t)}{\omega_2 - \omega_1} + \frac{\sigma_1^\dagger(t)}{\omega_2 + \omega_1} \right), \quad (51)$$

and, from the expression (41),

$$\sigma_1(t) \cong \sigma_{10}(t) + \frac{C_1}{\hbar} \left(\frac{a_0(t)}{\omega_1 - \omega} + \frac{a_0^\dagger(t)}{\omega_1 + \omega} \right) - \frac{C_2 V_{12}}{\hbar \omega} \frac{2\omega_2}{\omega_1^2 - \omega_2^2} (a_0(t) - a_0^\dagger(t)) \quad (52)$$

to first order in V_{12} . Thus,

$$\langle a(t) \sigma_1(t) \rangle \cong \langle a_0(t) \sigma_1(t) \rangle = \frac{C_1}{\hbar} \frac{1}{\omega_1 + \omega} + \frac{2C_2 V_{12}}{\hbar \omega} \frac{\omega_2}{\omega_1^2 - \omega_2^2}, \quad (53)$$

since $\sigma_{10}(t)|\psi\rangle = 0$, $\langle a_0(t) a_0(t) \rangle = \langle a_0^\dagger(t) a_0(t) \rangle = 0$, and $\langle a_0(t) a_0^\dagger(t) \rangle = 1$, and

$$\langle H_c^{(1)} \rangle \cong -\frac{2C_1^2/\hbar}{\omega_1 + \omega} - \frac{2C_1 C_2 V_{12}}{\hbar \omega} \frac{\omega_2}{\omega_1^2 - \omega_2^2}. \quad (54)$$

The same approach for atom 2 gives

$$\langle H_c^{(2)} \rangle \cong -\frac{2C_2^2/\hbar}{\omega_2 + \omega} - \frac{2C_1 C_2 V_{12}}{\hbar \omega} \frac{\omega_1}{\omega_2^2 - \omega_1^2}. \quad (55)$$

For the expectation value of $H_c = H_c^{(1)} + H_c^{(2)}$, we therefore obtain

$$\langle H_c \rangle \cong -\frac{2C_1^2/\hbar}{\omega_1 + \omega} - \frac{2C_2^2/\hbar}{\omega_2 + \omega} - \frac{2C_1 C_2 V_{12}/(\hbar \omega)}{\omega_1 + \omega_2}. \quad (56)$$

Since C_1^2 and C_2^2 are proportional to e^2 , and $C_1 C_2 V_{12}$ is proportional to e^4 , the PHF theorem introduces factors

$$\int_0^1 \frac{d\lambda}{\lambda} \lambda^2 = \frac{1}{2} \quad \text{and} \quad \int_0^1 \frac{d\lambda}{\lambda} \lambda^4 = \frac{1}{4} \quad (57)$$

for the first two terms and the last term, respectively, on the r.h.s. of (56), so that the change in the atom-field system due to their interaction is

$$\Delta E = -A_1^2 (\mathbf{d}_1 \cdot \hat{\mathbf{e}})^2 \frac{\omega}{\omega + \omega_1} - A_2^2 (\mathbf{d}_2 \cdot \hat{\mathbf{e}})^2 \frac{\omega}{\omega + \omega_2} - \frac{1}{2} \frac{A_1 A_2}{r^3} \frac{(\mathbf{d}_1 \cdot \hat{\mathbf{e}})(\mathbf{d}_2 \cdot \hat{\mathbf{e}}) [\mathbf{d}_1 \cdot \mathbf{d}_2 - 3(\mathbf{d}_1 \cdot \hat{\mathbf{e}})(\mathbf{d}_2 \cdot \hat{\mathbf{e}})]}{\omega_1 + \omega_2} \quad (58)$$

when we use the definition (39).

The two-state model simplifies some algebra in our Heisenberg-picture calculation, but the result (58) is easily generalized to include contributions from all the allowed transitions from the ground states of the two atoms. The first term on the r.h.s. of (58), for instance, generalizes to

$$\Delta E_1 = -A_1^2 \sum_s |(\mathbf{d}_1 \cdot \hat{\mathbf{e}})_{sg}|^2 \frac{\omega}{\omega + \omega_{sg}} \quad (59)$$

in the notation of (15). After accounting for additional self-energy terms, we obtain the Lamb shift due to the coupling of atom 1 to the single-mode field. But of greater interest here is the interatomic interaction term in (58). For strong coupling to a single-mode field, the nonretarded van der Waals interaction varies as $1/r^3$ rather than $1/r^6$ [18, 19]. When generalized to include all allowed transitions from the ground states, we obtain, except for the factor $A_1 A_2$, the result of the perturbation-theoretic analysis of [19]. This factor has an interesting implication for the physical interpretation of the $1/r^3$ interaction, as discussed in the following section^{†3}.

^{†3}Note that when we include a term of second order in V_{12} , we obtain another contribution to ΔE that corresponds to the familiar $1/r^6$ van der Waals interaction.

7. Conclusions

Zero-temperature Lamb shifts and van der Waals interactions have clear physical interpretations in terms of fluctuating zero-point fields. In particular, for the van der Waals interaction between two atoms in free space, each atom is “driven” by the zero-point field at its location, and the fluctuations of the zero-point fields at the two locations are correlated. The correlation decreases rapidly with the distance r between the two locations, giving the r^{-6} dependence of the nonretarded van der Waals interaction.

In the case of strong coupling of the atoms to a single cavity mode, unlike the case in which the atoms are coupled to the infinite set of modes of free space, there is no decrease in electric field correlations with r , and for small r , the van der Waals interaction varies as r^{-3} rather than r^{-6} . Such r^{-3} behavior is also found in a different scenario, when each atom experiences an *externally applied* single-mode field [20]. In this case, the interpretation of the r^{-3} behavior is obvious — the two atoms have correlated induced dipole moments and experience, for small r , just the r^{-3} electrostatic dipole–dipole interaction. But in the more subtle r^{-3} behavior resulting from the coupling of the atoms to a *zero-point*, vacuum cavity mode [18, 19], each atom has

a dipole moment induced by the zero-point field whose fluctuations are correlated for effectively all values of r . We note that the presence of the factor $A_1 A_2$ in the energy (58) implies that there is no van der Waals interaction if one of the atoms finds itself at a node of the cavity field, i.e., if either A_1 or A_2 vanishes.

The derivations of the zero-temperature Lamb shifts and van der Waals interactions using the PHF theorem make it clear that these effects are attributable to the fluctuations of the zero-point electromagnetic field. They can also be said to be attributable to changes in zero-point energy, as in Feynman's argument for the Lamb shift in Sect. 2. But the fluctuation perspective seems to offer a more physical picture of interacting dipoles as opposed to just energy "bookkeeping". Moreover, Lamb shifts and van der Waals interactions can be understood from the perspective of the quantum fluctuations not of zero-point fields but of the "source" fields generated by the dipoles themselves. The same is true for Casimir's famous attraction between conducting plates [21].

Acknowledgments

It is a pleasure to submit this paper in recognition of the creative and influential contributions of Professor Iwo Białynicki-Birula. I thank P.R. Berman and G.W. Ford for helping me to better understand the "remarkable formula" and van der Waals interactions, and G.J. Maclay for informative discussions relating to Lamb shifts.

Appendix: Electric field correlations and dyadic Green function

The positive-frequency part of the electric field operator for a vacuum or thermal field can be expressed as

$$\mathbf{E}^{(+)}(\mathbf{r}, t) = i \sum_{\mathbf{k}\lambda} \sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}}{V}} a_{\mathbf{k}\lambda} e^{-i\omega_{\mathbf{k}}t} e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k}\lambda}, \quad (60)$$

where, as usual, $a_{\mathbf{k}\lambda}$ is the photon annihilation operator for the plane-wave mode with the wave vector \mathbf{k} and the polarization index λ . For thermal radiation, $\langle a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}'\lambda'} \rangle = q(\omega) \delta_{\mathbf{k}\mathbf{k}'}^3 \delta_{\lambda\lambda'}$ and $\langle a_{\mathbf{k}\lambda} a_{\mathbf{k}'\lambda'}^\dagger \rangle = [q(\omega) + 1] \delta_{\mathbf{k}\mathbf{k}'}^3 \delta_{\lambda\lambda'}$, where $q(\omega) = [e^{\hbar\omega/k_B T} - 1]^{-1}$, and it follows after taking $\sum_{\mathbf{k}\lambda}(\dots) \rightarrow V/(2\pi)^3 \sum_{\lambda} \int d^3k(\dots)$ in the familiar fashion that

$$\begin{cases} \langle E_i^+(\mathbf{r}_n, t) E_j^{(-)}(\mathbf{r}_m, t') \rangle = \frac{\hbar}{\pi c^3} \int_0^\infty d\omega \omega^3 [q(\omega) + 1] F_{ij}(\frac{\omega r}{c}) e^{i\omega(t'-t)}, \\ \langle E_i^-(\mathbf{r}_n, t) E_j^{(+)}(\mathbf{r}_m, t') \rangle = \frac{\hbar}{\pi c^3} \int_0^\infty d\omega \omega^3 q(\omega) F_{ij}(\frac{\omega r}{c}) e^{-i\omega(t'-t)}, \end{cases} \quad (61)$$

$$F_{ij}(x) \equiv (\delta_{ij} - \hat{\mathbf{r}}_i \hat{\mathbf{r}}_j) \frac{\sin(x)}{x} + (\delta_{ij} - 3\hat{\mathbf{r}}_i \hat{\mathbf{r}}_j) \left(\frac{\cos(x)}{x^2} - \frac{\sin(x)}{x^3} \right), \quad (62)$$

where $r = |\mathbf{r}_n - \mathbf{r}_m|$. Thus, for thermal radiation,

$$\begin{cases} \langle E_i^+(\mathbf{r}_n, \omega) E_j^{(-)}(\mathbf{r}_m, \omega') \rangle = \frac{\hbar}{\pi c^3} \omega^3 [q(\omega) + 1] F_{ij}(\frac{\omega r}{c}) \delta(\omega - \omega'), \\ \langle E_i^-(\mathbf{r}_n, \omega) E_j^{(+)}(\mathbf{r}_m, \omega') \rangle = \frac{\hbar}{\pi c^3} \omega^3 q(\omega) F_{ij}(\frac{\omega r}{c}) \delta(\omega - \omega'). \end{cases} \quad (63)$$

The electric field $\mathbf{E}_S(\mathbf{r}_n, t)$ at a point \mathbf{r}_n from an electric dipole source at \mathbf{r}_m is

$$\begin{aligned} E_{Si}(\mathbf{r}_n, t) = & -\frac{1}{c^2 r} (\delta_{ij} - \hat{\mathbf{r}}_i \hat{\mathbf{r}}_j) \ddot{p}_j \left(t - \frac{r}{c} \right) \\ & - (\delta_{ij} - 3\hat{\mathbf{r}}_i \hat{\mathbf{r}}_j) \left[\frac{1}{cr^2} \dot{p}_j \left(t - \frac{r}{c} \right) + \frac{1}{r^3} p_j \left(t - \frac{r}{c} \right) \right]. \end{aligned} \quad (64)$$

We therefore identify

$$E_{Si}^{(+)}(\mathbf{r}_n, \omega) = G_{ij}(\mathbf{r}_n, \mathbf{r}_m, \omega) p_j^{(+)}(\mathbf{r}_m, \omega), \quad (65)$$

$$\begin{aligned} G_{ij}(\mathbf{r}_n, \mathbf{r}_m, \omega) = & k \frac{\omega^2}{c^2} \left[(\delta_{ij} - \hat{\mathbf{r}}_i \hat{\mathbf{r}}_j) \frac{1}{kr} \right. \\ & \left. + (\delta_{ij} - 3\hat{\mathbf{r}}_i \hat{\mathbf{r}}_j) \left(\frac{i}{k^2 r^2} - \frac{1}{k^3 r^3} \right) \right] e^{ikr}, \end{aligned} \quad (66)$$

with $k = n\omega/c$, and (19) then follows from (63).

References

- [1] R.P. Feynman, *Acta Phys. Pol.* **24**, 841 (1963).
- [2] R.P. Feynman, in: *The Quantum Theory of Fields*, Ed. R. Stoops, Interscience Publishers, New York 1961, p. 61.
- [3] E.A. Power, *Am. J. Phys.* **34**, 516 (1966).
- [4] H.A. Bethe, *Phys. Rev.* **72**, 339 (1947).
- [5] G.J. Maclay, *Physics* **2020**, 105 (2020).
- [6] W. Pauli, in: *Handbuch der Physik*, Vol. 24, 1933, p. 162.
- [7] H. Hellmann, *Einführung in die Quantenchemie*, Deuticke, Leipzig 1937, p. 285.
- [8] R.P. Feynman, *Phys. Rev.* **56**, 340 (1939).

- [9] J.I. Musher, *Am. J. Phys.* **34**, 267 (1966).
- [10] K. Sawada, K.A. Brueckner, N. Fukuda, R. Brout, *Phys. Rev.* **106**, 372 (1957).
- [11] P.R. Berman, G.W. Ford, P.W. Milonni, *J. Chem. Phys.* **141**, 16105 (2014).
- [12] M. Pons, B. Juliá-Díaz, A. Polls, A. Rios, I. Vidaña, *Am. J. Phys.* **88**, 503 (2020).
- [13] G.W. Ford, J.T. Lewis, R.F. O'Connell, *Phys. Rev. Lett.* **55**, 2273 (1985).
- [14] J.W. Farley, W.H. Wing, *Phys. Rev. A* **23**, 2397 (1981).
- [15] L. Hollberg, J.L. Hall, *Phys. Rev. Lett.* **53**, 230 (1984).
- [16] P.R. Berman, G.W. Ford, P.W. Milonni, *Phys. Rev. A* **89**, 022127 (2014).
- [17] M.J. Renne, *Physica* **53**, 193 (1971).
- [18] T.S. Haugland, C. Schäfer, E. Ronca et al., *J. Chem. Phys.* **154**, 094113 (2021).
- [19] J.P. Philbin, T.S. Haugland, T.K. Ghosh et al., [arXiv:2209.07956](https://arxiv.org/abs/2209.07956), 2022.
- [20] P.W. Milonni, A. Smith, *Phys. Rev. A* **53**, 3484 (1996).
- [21] R.L. Jaffe, *Phys. Rev. D* **72**, 021301(R) (2005).

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

Born–Infeld Nonlinear Electromagnetism in Relativistic Heavy Ion Collisions

W. PRICE^a, M. FORMANEK^b AND J. RAFELSKI^a^a*Department of Physics, The University of Arizona, Tucson, AZ, 85721, USA*^b*ELI Beamlines facility, The Extreme Light Infrastructure ERIC, 252 41 Dolní Břežany, Czech Republic*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S87](https://doi.org/10.12693/APhysPolA.143.S87)*e-mail: wprice@arizona.edu

We study the effect of the limiting field strength of Born–Infeld electromagnetism on the dynamics of charged particle scattering. We formulate the Born–Infeld limiting field in an invariant manner, showing that it is the electric field-dominated eigenvalue “ a ” of the field tensor $F^{\mu\nu}$ which is limited rather than the individual field vectors. Heavy ion collisions in particular provide uniquely large values of the field invariants that appear in the Born–Infeld action, amplifying nonlinear effects. Thus “ a ” is the dominant input into the force between heavy ions that we use to compute the scattering angle as a function of the impact parameter. We evaluate the Born–Infeld effects, showing relevance at small impact parameters and exhibiting their dependence on the value of the limiting field strength.

topics: Born–Infeld electromagnetism, heavy-ion collisions, strong fields

1. Introduction

In recent years, relativistic heavy ion collisions have been explored as a testing ground for strong electromagnetic (EM) field effects [1–6]. In peripheral collisions with large impact parameters, electromagnetic forces dominate the scattering processes, and strong field effects such as light–light scattering and spontaneous pair production can be observed.

In this paper, we suggest that heavy ion collisions also can serve as a potential means of exploring classical, nonlinear electromagnetic effects. Specifically, we will study the nonlinear Born–Infeld (BI) theory of electromagnetism and its effect on the scattering of relativistic heavy ions. The nonlinearity in BI theory is dependent on the EM field tensor invariants, \mathcal{S} and \mathcal{P} (defined in Appendix A), which are quite large in relativistic heavy ion collisions.

Nonlinear EM theories inherently contain a characteristic electric field strength scale, E_0 . In the BI theory, this field scale acts as an upper limit to the electric field in the rest frame of a particle [7, 8].

This feature yields a finite electromagnetic mass of the particle, without the divergences that appear in Maxwell’s theory, and a vanishing electromagnetic self-stress on the particle [9]. Additionally, the BI theory is the only nonlinear theory of EM in which the wave velocity does not depend on its polarization [10], and the waves can be linearly superposed with a constant background field [11].

We begin by reviewing the Lagrangian and field equations of BI theory in Sect. 2. We determine the form of the Lagrangian in terms of the EM field tensor eigenvalues a and b (see Appendix A). This allows us to formulate the BI limiting field condition in an invariant manner rather than the usual picture, where E_0 is the limiting electric field in the rest frame of the particle. Our formulation shows that the electric-like eigenvalue a is a true bounded quantity.

In Sect. 3, we formulate the equations of motion of colliding ions. In order to yield a soluble problem, we approximate the ions as test particles undergoing Lorentz force motion. We then determine the EM fields of each ion separately by first solving

the BI field equations in the rest frame and then Lorentz boosting to the center-of-momentum frame. This approach allows us to sidestep the difficulties that arise when one considers the force acting on a BI particle [12–15], which we will turn our attention to in future work.

We present numerical results for ion motion in Sect. 4. In particular, we look at the scattering angle dependence on the impact parameter for a range of E_0 values. We see that the BI limiting field suppresses the scattering angle when the impact parameter is small, allowing the ions to approach each other more closely than in the framework of Maxwell electrodynamics. For larger values of E_0 , this requires smaller impact parameters.

2. Born–Infeld electromagnetism

2.1. Review of well-known relations

In this section, we review the equations of BI theory and formulate the limiting field condition in Lorentz invariant form. Definitions of all relevant mathematical quantities and notation can be found in Appendix A. The free Lagrangian of the Born–Infeld theory is, in a flat spacetime with the metric $g_{\mu\nu}$, given by

$$\mathcal{L} = \epsilon_0 E_0^2 \left[1 - \sqrt{-\det \left(g_{\mu\nu} + \frac{c F_{\mu\nu}}{E_0} \right)} \right], \quad (1)$$

where $F_{\mu\nu}$ is the EM field tensor. Computing the determinant yields the equivalent form

$$\mathcal{L} = \epsilon_0 E_0^2 \left(1 - \sqrt{1 + \frac{2\mathcal{S}c^2}{E_0^2} - \frac{\mathcal{P}^2 c^4}{E_0^4}} \right), \quad (2)$$

where \mathcal{S} and \mathcal{P} are the EM field tensor invariants, see (54) and (55) in Appendix A. In the limit of large E_0 , (2) becomes the Maxwell Lagrangian,

$$\mathcal{L} \rightarrow \frac{\mathcal{S}}{\mu_0}, \quad \text{as } E_0 \rightarrow \infty. \quad (3)$$

Upon coupling (2) to the current density j^μ and performing the variation with respect to the four-potential A^μ , one arrives at the BI field equations

$$\partial_\mu \mathcal{H}^{\mu\nu} = j^\nu, \quad (4)$$

$$\partial_\mu \tilde{F}^{\mu\nu} = 0, \quad (5)$$

where the displacement field tensor $\mathcal{H}^{\mu\nu}$ is defined to be

$$\mathcal{H}^{\mu\nu} \equiv 2 \frac{\delta \mathcal{L}}{\delta F_{\mu\nu}} = \frac{1}{\mu_0} \frac{F^{\mu\nu} - \frac{\mathcal{P}c^2}{E_0^2} \tilde{F}^{\mu\nu}}{\sqrt{1 + \frac{2\mathcal{S}c^2}{E_0^2} - \frac{\mathcal{P}^2 c^4}{E_0^4}}} \quad (6)$$

and $\tilde{F}^{\mu\nu}$ denotes the dual EM tensor defined in (49) in Appendix A. The BI equations then take a form identical to the Maxwell equations in a dielectric medium. The nonlinear BI effects can then be interpreted as the effect of the dielectric medium on the field close to its source.

One can invert (6) and solve for $F^{\mu\nu}$ (see [7]), yielding

$$F^{\mu\nu} = \mu_0 \frac{\mathcal{H}^{\mu\nu} + \frac{\mathcal{Q}\mu_0^2 c^2}{E_0^2} \tilde{\mathcal{H}}^{\mu\nu}}{\sqrt{1 - \frac{2\mathcal{R}\mu_0^2 c^2}{E_0^2} - \frac{\mathcal{Q}^2 \mu_0^4 c^4}{4E_0^4}}}, \quad (7)$$

where \mathcal{R} and \mathcal{Q} are invariants of the displacement field tensor defined in (56) and (57) in Appendix A.

2.2. Eigenvalue formulation of BI theory

We can also formulate the BI theory in terms of the EM field tensor eigenvalues, $\pm a$ and $\pm ib$, often used in the context of the Euler–Heisenberg–Schwinger effective action of quantum electrodynamics (QED) [16] (see (63) in Appendix A). This approach has the benefit of making the limiting field condition explicit in the action. The eigenvalues are related to the invariants \mathcal{S} and \mathcal{P} through

$$a = \sqrt{-\mathcal{S} + \sqrt{\mathcal{S}^2 + \mathcal{P}^2}}, \quad (8)$$

$$b = \sqrt{\mathcal{S} + \sqrt{\mathcal{S}^2 + \mathcal{P}^2}}, \quad (9)$$

where a is the electrically-dominated eigenvalue and b is the magnetically-dominated eigenvalue, such that for zero magnetic field $a = E/c$ and for zero electric field $b = B$. Since the expressions under all four square roots are strictly positive, one can easily check that the following identities are true

$$a^2 - b^2 = -2\mathcal{S}, \quad (10)$$

$$a^2 b^2 = \mathcal{P}^2, \quad (11)$$

as well as $a^2 > 0$ and $b^2 > 0$. This allows us to re-write the determinant in (2) as the product of eigenvalues,

$$\mathcal{L} = \epsilon_0 E_0^2 \left(1 - \sqrt{\prod_{k=1}^4 \Lambda_k} \right), \quad (12)$$

where the eigenvalues Λ_k of $\delta_\nu^\mu + F_\nu^\mu c/E_0$ are defined in Appendix A, below (64). Now (12) can be simplified to

$$\mathcal{L} = \epsilon_0 E_0^2 \left(1 - \sqrt{\left(1 - c^2 \frac{a^2}{E_0^2}\right) \left(1 + c^2 \frac{b^2}{E_0^2}\right)} \right). \quad (13)$$

In fact, (13) can also be obtained by directly substituting (10) and (11) directly into (2).

The expression under the square root in (13) must be positive in order for the Lagrangian to be real and the corresponding field equations to be real. Imposing this constraint yields an upper limit for the field strength,

$$a < \frac{E_0}{c}, \quad (14)$$

while there is no limit on the value of b . Thus, it is the electric field-like invariant eigenvalue a that is limited in general, not just the electric field.

We now study the limiting field behavior of two different configurations of invariants. When $\mathcal{P} = 0$, we have

$$a = \sqrt{-\mathcal{S} + |\mathcal{S}|}. \quad (15)$$

For an electrically-dominated system, $\mathcal{S} < 0$ and $a = \sqrt{-2\mathcal{S}}$. In this case, \mathcal{S} is limited. For a magnetically-dominated system with $\mathcal{S} > 0$, we have $a = 0$ and the fields are therefore not limited.

When $\mathcal{S} = 0$, we have

$$a = \sqrt{|\mathcal{P}|}, \quad (16)$$

and therefore \mathcal{P} is limited.

2.3. Value of the limiting field constant E_0

As we will argue below, experimental data in relativistic heavy ion collisions or other strong field environments could be used to determine the BI limiting field constant E_0 . Born and Infeld originally calculated the value of E_0 on the assumption that the experimentally measured mass of the electron is entirely electromagnetic, yielding the limiting field value

$$E_0 = 1.18 \times 10^{20} \frac{\text{V}}{\text{m}}. \quad (17)$$

However, we now know that the electron mass is made up in part of non-electromagnetic components. Therefore, the Born and Infeld value of E_0 cannot be exact. Instead, we must determine E_0 by studying the dynamics of charged particles experimentally. There were previously several studies of possible bounds on the BI limiting field constant with conflicting results [17–21]. We hope that further studies of the effects of BI theory on particle dynamics in strong-field environments, such as in the present paper, as well as relatively recent papers on the BI effects in laser-plasma acceleration [22, 23] can lay the groundwork for an experimental study of the value of E_0 .

It is also interesting to compare the nonlinear classical theory with the nonlinearity inherent in QED. We can compare the classical BI limiting field (at least the value calculated by Born and Infeld) to the field strength that appears in the Euler–Heisenberg–Schwinger effective Lagrangian of QED [16, 24, 25]

$$E_{\text{EHS}} = \frac{m_e^2 c^3}{e \hbar} = 0.0112 E_0. \quad (18)$$

Thus E_{EHS} represents the scale at which quantum nonlinear effects set in, while E_0 corresponds to classical nonlinear effects. We see here that the quantum nonlinear scale is approximately two orders of magnitude smaller than the classical nonlinear scale.

Additionally, we can compare E_0 to the limiting acceleration value that appears in the Eliezer–Ford–O’Connell (EFO) radiation reaction (RR) force model [26]. The EFO radiation reaction force can be written as

$$\mathcal{F}_{\text{EFO}}^\mu = \tau_0 P_\nu^\mu \frac{d}{d\tau} (e F^{\nu\alpha} u_\alpha), \quad (19)$$

which yields the following equation of motion

$$\left(g_{\mu\nu} - \frac{e\tau_0}{m} P_\mu^\alpha F_{\alpha\nu} \right) \dot{u}^\nu = \frac{e}{m} (F_{\mu\nu} + \tau_0 \dot{F}_{\mu\nu}) u^\nu, \quad (20)$$

where the dot refers to the proper time derivative and τ_0 is the characteristic RR time scale (for electrons $\tau_0 = 6.26 \times 10^{-24}$ s). The orthogonal projection tensor is defined as

$$P_{\mu\nu} = g_{\mu\nu} - \frac{u_\mu u_\nu}{c^2}. \quad (21)$$

To solve the acceleration in (20), one must invert the tensor $g_{\mu\nu} - \frac{e\tau_0}{m} P_\mu^\alpha F_{\alpha\nu}$ by taking its determinant, which takes a similar form to the BI Lagrangian (2)

$$-\det \left(g_{\mu\nu} - \frac{e\tau_0}{m} P_\mu^\alpha F_{\alpha\nu} \right) = 1 + \frac{e^2 \tau_0^2}{m^2} \left(2\mathcal{S} + \frac{u F F u}{c^2} \right). \quad (22)$$

As we have shown in [26], in certain field configurations, (20) leads to an upper limit on the acceleration analogous to the BI limiting field

$$a_{\text{EFO}} = \frac{c}{\tau_0} = \frac{3}{2} \frac{4\pi\epsilon_0 m c^2}{e^2} = 4.80 \times 10^{31} \frac{\text{m}}{\text{s}^2}, \quad (23)$$

where the value is given for electrons. The electric field corresponding to this limiting acceleration is

$$E_{\text{EFO}} = \frac{E_{\text{EHS}}}{\alpha} = 1.53 E_0, \quad (24)$$

where $\alpha \approx 1/137$ is the fine structure constant. We see that the classical limiting field appearing in the radiation reaction force is of the same order of magnitude as the BI limiting field E_0 , which is obtained by requiring all of the electron mass to be of electromagnetic origin.

3. Heavy ion scattering

3.1. Force on a BI particle

We will consider the scattering of two identical heavy ions, both of charge Ze and mass m . We will approximate ions as test particles. In this case, the force on each charge is given by the Lorentz force. For ions located at positions \mathbf{x}_1 and \mathbf{x}_2 and with fields $F_1^{\mu\nu}$ and $F_2^{\mu\nu}$, our system of equations is

$$m \dot{u}_1^\mu = Ze F_2^{\mu\nu} u_{1\nu}, \quad (25)$$

$$m \dot{u}_2^\mu = Ze F_1^{\mu\nu} u_{2\nu}. \quad (26)$$

For symmetric collisions, the ions will have the same proper time, τ . In the test particle approximation, the field of one ion acts on the other ion. The BI effects in our equations of motion come from $F^{\mu\nu}$, which is a solution of the BI field equations and therefore is restricted by the limiting field condition $a < E_0/c$. We neglect backreaction from the self-field of each ion.

Our test particle approach allows us to formulate a tractable problem, since the closed-form expression for the force between two relativistic charges in

BI theory is not known. The Lorentz force, which we use here, is the leading order force on the BI charge for small acceleration [27]. The first-order BI corrections to the Coulomb force between two static charges have been calculated in [28].

Our first step, then, is to solve the BI equations for a single ion undergoing relativistic motion. An analogous problem for Maxwell's equations is solved by the Lienard–Wiechert fields [29]. However, in the nonlinear BI theory, the exact solution to this problem is not yet known.

3.2. Fields of a relativistic BI particle

To obtain an approximate solution for the field of the ion, we will assume that its acceleration is relatively small so that its rest frame is approximately inertial. In the rest frame, assuming that the ion is located at a position \mathbf{x}'_0 with charge Ze , the BI equations simplify to

$$\nabla \cdot \mathbf{D}'(\mathbf{x}') = Ze \delta^3(\mathbf{X}'), \quad \mathbf{X}' \equiv \mathbf{x}' - \mathbf{x}'_0 \quad (27)$$

while the magnetic field vanishes

$$\mathbf{H}'(\mathbf{x}') = 0. \quad (28)$$

Here (27) is solved by the Coulomb field

$$\mathbf{D}'(\mathbf{x}') = \frac{Ze}{4\pi} \frac{\mathbf{X}'}{|\mathbf{X}'|^3}. \quad (29)$$

We can then Lorentz boost the rest frame fields, (28) and (29), to the center-of-momentum (CM) frame where the ion is moving with velocity \mathbf{v} with negligible acceleration. The \mathbf{D} and \mathbf{H} fields transform identically to \mathbf{E} and \mathbf{B} , whose transformation formulae are given in [29]. The details of the Lorentz transformation are presented in Appendix B. There, the displacement fields in the CM frame are given in Appendix B as (72) and (73), therefore

$$\mathbf{D}(t, \mathbf{x}) = \frac{Ze}{4\pi} \frac{\gamma}{R^3} \mathbf{X}, \quad (30)$$

$$\mathbf{H}(t, \mathbf{x}) = \frac{Ze}{4\pi} \frac{\gamma}{R^3} \mathbf{v} \times \mathbf{X}, \quad (31)$$

where

$$R = \sqrt{\mathbf{X}_\perp^2 + \gamma^2 \mathbf{X}_\parallel^2}. \quad (32)$$

We then compute the invariants \mathcal{R} and \mathcal{Q} , using (30) and (31), which leads to

$$\mathcal{R} = -\frac{1}{2} \left(\frac{Zec}{4\pi R^2} \right)^2, \quad (33)$$

$$\mathcal{Q} = 0. \quad (34)$$

Now (33) is obtained by noting that

$$\left(\frac{\mathbf{v}}{c} \times \mathbf{X} \right)^2 - \mathbf{X}^2 = \frac{R^2}{\gamma^2}. \quad (35)$$

We can now compute the \mathbf{E} and \mathbf{B} fields corresponding to (30) and (31). In three-vector form, (7) reads

$$\mathbf{E} = \frac{1}{\epsilon_0} \frac{\mathbf{D}}{\sqrt{1 - \frac{2\mathcal{R}\mu_0^2 c^2}{E_0^2}}}, \quad (36)$$

$$\mathbf{B} = \mu_0 \frac{\mathbf{H}}{\sqrt{1 - \frac{2\mathcal{R}\mu_0^2 c^2}{E_0^2}}}, \quad (37)$$

where we have substituted $\mathcal{Q} = 0$. Upon combining (36) and (37) with the expressions for \mathbf{D} , \mathbf{H} , and \mathcal{R} , we find

$$\mathbf{E} = \frac{Ze\gamma}{4\pi\epsilon_0 R^3} \frac{\mathbf{X}}{\sqrt{1 + \left(\frac{Ze}{4\pi\epsilon_0 R^2} / E_0 \right)^2}}, \quad (38)$$

$$\mathbf{B} = \frac{Ze\gamma}{4\pi c^2 \epsilon_0 R^3} \frac{\mathbf{v} \times \mathbf{X}}{\sqrt{1 + \left(\frac{Ze}{4\pi\epsilon_0 R^2} / E_0 \right)^2}}. \quad (39)$$

The behavior of these fields in the ultrarelativistic limit is studied in [30].

As a final step, we will check that our fields obey the limiting field constraint from (14). For this, \mathcal{S} and \mathcal{P} need to be computed. By the well-known duality symmetry of BI theory [8],

$$\mathcal{P} = \mu_0^2 \mathcal{Q}, \quad (40)$$

and therefore \mathcal{P} vanishes. Computing \mathcal{S} using (7), we find

$$\mathcal{S} = \mu_0^2 \frac{\mathcal{R}}{1 - \frac{2\mathcal{R}\mu_0^2 c^2}{E_0^2} - \frac{\mathcal{Q}^2 \mu_0^4 c^4}{E_0^4}}, \quad (41)$$

which simplifies to

$$\mathcal{S} = -\frac{1}{2} \frac{\left(\frac{Ze}{4\pi\epsilon_0 R^2} \right)^2}{1 + \frac{1}{E_0^2} \left(\frac{Ze}{4\pi\epsilon_0 R^2} \right)^2}. \quad (42)$$

In this case $\mathcal{P} = 0$, so from (58) given in the Appendix A, we have that $a = \sqrt{-2\mathcal{S}}$ and therefore

$$a = \frac{\frac{Ze}{4\pi c \epsilon_0 R^2}}{\sqrt{1 + \frac{1}{E_0^2} \left(\frac{Ze}{4\pi\epsilon_0 R^2} \right)^2}}. \quad (43)$$

In the limit $R \rightarrow 0$,

$$a \rightarrow \frac{E_0}{c} \quad (44)$$

and we see that our fields obey the invariant limiting field condition.

4. Numerical results

Our heavy ion collisions are described by the system of equations (25) and (26), which depend on the electric (38) and magnetic (39) fields. We integrate the equations using the fifth-order Runge–Kutta–Dormand–Prince method [31] and use the constraint $u^2 = c^2$ as an estimation of the numerical error. Ions are assumed to start from $\tau = -\infty$ with incoming velocities $\pm v_0$ as well as transverse separation in the y -direction by an impact parameter ρ . The initial longitudinal separation is taken to be large enough such that the results are converged with respect to its small variations. We use ρ for the impact parameter rather than the usual b because the latter is reserved for the EM tensor eigenvalue in this paper.

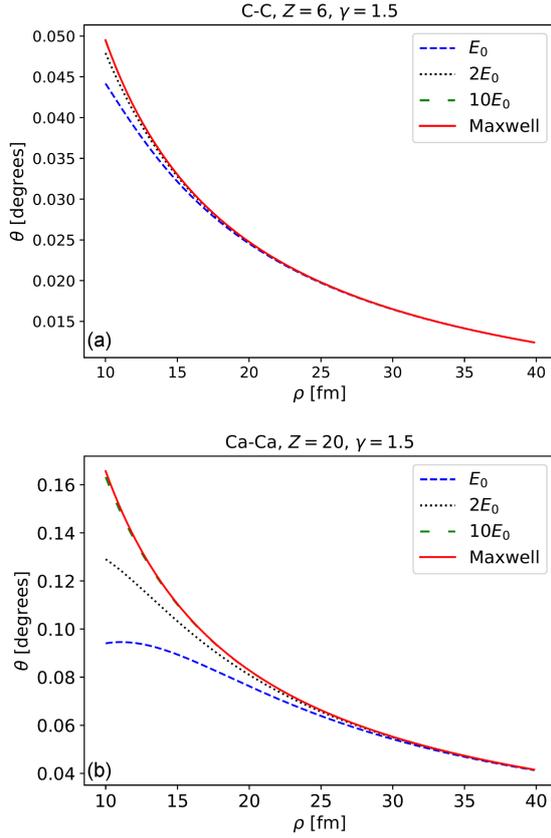


Fig. 1. Scattering angle as a function of impact parameter ρ for collisions C–C (panel (a)) and Ca–Ca (panel (b)), with $\gamma = 1.5$.

We present the results for four pairs of colliding ions: Au–Au, U–U, Ca–Ca, and C–C to study a range of charge-to-mass ratios. We compute the scattering angle θ by comparing the incoming and outgoing momentum vectors for the ions.

In Fig. 1, we plot the relationship between θ and ρ for C–C (panel (a)) and Ca–Ca (panel (b)), at $\gamma = 1.5$. In Fig. 2, we present the scattering angle for Au–Au at $\gamma = 1.5$ (panel (a)) and U–U at $\gamma = 1.2$ (panel (b)). The scattering angle is computed for the Born and Infeld value of E_0 given in (17), as well as for $2E_0$, $10E_0$, and the Maxwell theory corresponding to $E_0 \rightarrow \infty$. The results are shown for $\rho > 10$ fm to ensure that the collisions remain peripheral and the point particle approximation remains valid.

Our results show that BI effects are most significant at smaller impact parameters, where the field seen by each ion is bounded. This is in contrast to Maxwell’s theory, where the force approaches infinity for smaller and smaller impact parameters. The limiting BI field significantly suppresses the scattering angle at lower impact parameters, with the scattering angle decreasing more so as E_0 decreases. At larger impact parameters, BI effects are relevant for smaller values of E_0 and become negligible at $10E_0$ and above.

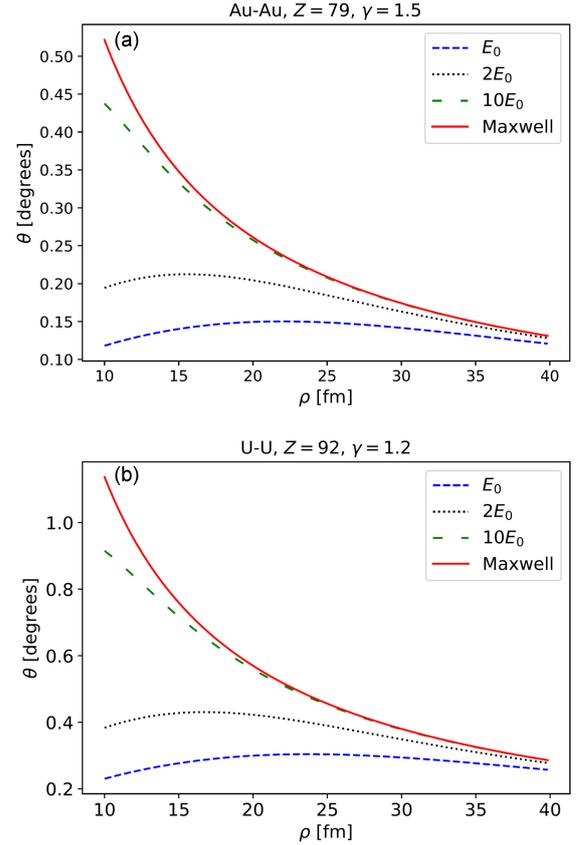


Fig. 2. Scattering angle θ as a function of impact parameter ρ for Au–Au collisions with $\gamma = 1.5$ (a) and U–U collisions with $\gamma = 1.2$ (b).

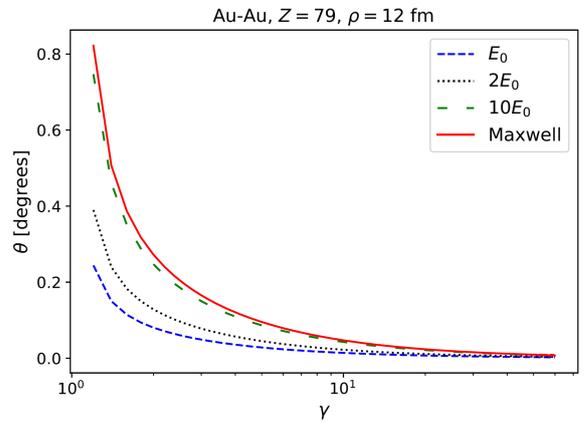


Fig. 3. Scattering angle as a function of incoming γ at a fixed impact parameter, $\rho = 12$ fm.

BI effects are additionally amplified for large Z . The C–C results in Fig. 1b show a relatively small BI contribution to the scattering, even at $1E_0$. In contrast, the Au–Au collisions in Fig. 2a with a charge approximately 13 times larger show a significant BI effect on the scattering angle.

Figure 3 plots the scattering angle as a function of the incoming γ at a fixed impact parameter $\rho = 12$ fm. The scattering angle decreases as γ ,

and thus the inertia of the ions, γm , increases. Due to this, BI effects are amplified at lower γ , where the scattering angle is in general larger for all values of E_0 . The U–U plot shown in Fig. 2b demonstrates the effect of large $Z = 92$ as well as a smaller $\gamma = 1.2$, maximizing BI effects.

5. Conclusions

In this work, we have studied the limiting field feature of BI theory and established that the electric-like eigenvalue of the EM field tensor a is limited in all inertial frames by the electric field strength constant E_0 . In the rest frame of the particle, this becomes the limit on the electrostatic field, yielding finite electric field energy as usual.

We then applied BI theory to study the dynamics of relativistic heavy ion collisions. We showed that the limiting field can have a significant impact on the scattering angle at a low impact parameter when the ions pass through the region of the field greatly altered by nonlinear effects. This causes a reduction in force and thus scattering angle at low impact parameter compared to Maxwell’s theory. The difference between the BI and Maxwell’s theory is most relevant for smaller values of the limiting field E_0 and diminishes as this limiting value is increased.

Looking at the fixed scattering angle in Fig. 2, say 0.4° , we see that the BI force predicts a smaller corresponding impact parameter by approximately 25% for $10E_0$. The BI impact parameter is $\rho_{\text{BI}} \approx 12$ fm, while the Maxwell impact parameter is $\rho_{\text{Max}} \approx 15$ fm. This impact parameter shift is even larger for smaller values of the limiting field constant. A shift in perceived impact parameter such as we see here would have a significant effect on all impact parameter-sensitive calculations using peripheral heavy ion collision data, such as nuclear size [32, 33].

Moreover, self-field effects will likely be relevant at these small impact parameters. In BI theory this includes two phenomena, i.e., the radiation reaction and the nonlinear superposition of the self-field and the external field. In future work, we will turn our attention to formulating a framework for BI particle motion that includes these effects in a self-consistent manner.

Appendix A: Mathematical identities and notation

The conventions and notation used in this paper are as follows. We use a flat spacetime metric with a negative signature

$$g_{\mu\nu} = \text{diag}(1, -1, -1, -1). \quad (45)$$

The electromagnetic field tensor is given in terms of the EM four-potential $A^\mu = (\phi/c, \mathbf{A})$ as

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \quad (46)$$

and can be written in terms of electric fields \mathbf{E} and magnetic fields \mathbf{B} in cartesian coordinates as the following 4×4 matrix

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}. \quad (47)$$

We also use the displacement field tensor $\mathcal{H}^{\mu\nu}$, which can be written in terms of the displacement fields \mathbf{D} and \mathbf{H} as

$$\mathcal{H}^{\mu\nu} = \begin{pmatrix} 0 & -cD_x & -cD_y & -cD_z \\ cD_x & 0 & -H_z & H_y \\ cD_y & H_z & 0 & -H_x \\ cD_z & -H_y & H_x & 0 \end{pmatrix}. \quad (48)$$

We can also construct the corresponding dual tensors

$$\tilde{F}_{\alpha\beta} = \frac{1}{2}\epsilon_{\alpha\beta\mu\nu}F^{\mu\nu}, \quad (49)$$

$$\tilde{\mathcal{H}}_{\alpha\beta} = \frac{1}{2}\epsilon_{\alpha\beta\mu\nu}\tilde{\mathcal{H}}^{\mu\nu}, \quad (50)$$

where the totally antisymmetric Levi–Civita symbol is defined such that

$$\epsilon_{0123} = -\epsilon^{0123} = 1. \quad (51)$$

The dual tensors written in matrix form are

$$\tilde{F}_{\mu\nu} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & -E_z/c & E_y/c \\ B_y & E_z/c & 0 & -E_x/c \\ B_z & -E_y/c & E_x/c & 0 \end{pmatrix}, \quad (52)$$

$$\tilde{\mathcal{H}}_{\mu\nu} = \begin{pmatrix} 0 & -H_x & -H_y & -H_z \\ H_x & 0 & -D_z c & D_y c \\ H_y & D_z c & 0 & -D_x c \\ H_z & -D_y c & D_x c & 0 \end{pmatrix}. \quad (53)$$

From the field tensors, we can construct the following four field invariants

$$\mathcal{S} = \frac{1}{4}F^{\mu\nu}F_{\mu\nu} = \frac{1}{2}(B^2 - E^2/c^2), \quad (54)$$

$$\mathcal{P} = \frac{1}{4}F^{\mu\nu}\tilde{F}_{\mu\nu} = \mathbf{B} \cdot \mathbf{E}/c, \quad (55)$$

$$\mathcal{R} = \frac{1}{4}\mathcal{H}^{\mu\nu}\mathcal{H}_{\mu\nu} = \frac{1}{2}(H^2 - c^2D^2), \quad (56)$$

$$\mathcal{Q} = \frac{1}{4}\mathcal{H}^{\mu\nu}\tilde{\mathcal{H}}_{\mu\nu} = \mathbf{H} \cdot c\mathbf{D}. \quad (57)$$

Evaluating the BI action requires a computation of the determinant of the tensor $g_{\mu\nu} + F_{\mu\nu}c/E_0$. The four eigenvalues of F_ν^μ are $\Lambda_k \equiv \{\pm a, \pm ib\}$, defined by [16]

$$a = \sqrt{-\mathcal{S} + \sqrt{\mathcal{S}^2 + \mathcal{P}^2}}, \quad (58)$$

$$b = \sqrt{\mathcal{S} + \sqrt{\mathcal{S}^2 + \mathcal{P}^2}}. \quad (59)$$

The parameter a is electrically dominated such that for zero magnetic field

$$a = E/c, \quad (60)$$

and the parameter b is magnetically dominated such that for zero electric field

$$b = B. \quad (61)$$

For E/c and B of equal magnitude, we have $\mathcal{S} = 0$ and

$$a = b = |\mathcal{P}|. \quad (62)$$

In general, the determinant of $F_{\mu\nu}$ can be written as the product of eigenvalues

$$\begin{aligned} -\det(F_{\mu\nu}) &= -\det(g_{\mu\alpha}F_{\nu}^{\alpha}) = \\ &= -\det(g_{\mu\alpha})\det(F_{\nu}^{\alpha}) = \det(F_{\nu}^{\alpha}) = \\ &= \prod_{k=1}^4 \Lambda_k = -a^2b^2 = -\mathcal{P}^2, \end{aligned} \quad (63)$$

because of the identity $\det(AB) = \det(A)\det(B)$ and in the flat spacetime background $\det(g_{\mu\nu}) = -1$. The last equality follows either from direct computation with the matrix form (47) or manipulation of (58) and (59). Following the same steps, we find the determinant of $g_{\mu\nu} + F_{\mu\nu}c/E_0$ to be

$$\begin{aligned} -\det\left(g_{\mu\nu} + \frac{F_{\mu\nu}c}{E_0}\right) &= \\ \det\left(\delta_{\nu}^{\mu} + \frac{F_{\nu}^{\mu}c}{E_0}\right) &= \prod_{k=1}^4 \left(1 + \Lambda_k \frac{c}{E_0}\right) = \\ \left(1 - a \frac{c}{E_0}\right) \left(1 + a \frac{c}{E_0}\right) &\left(1 - ib \frac{c}{E_0}\right) \left(1 + ib \frac{c}{E_0}\right) = \\ \left(1 - \frac{c^2a^2}{E_0^2}\right) \left(1 + \frac{c^2b^2}{E_0^2}\right) &= 1 + \frac{2\mathcal{S}c^2}{E_0^2} - \frac{\mathcal{P}^2c^4}{E_0^4}, \end{aligned} \quad (64)$$

which is used in the main text of this paper, where $\Lambda_k \equiv 1 + \Lambda_k c/E_0$ are eigenvalues of $\delta_{\nu}^{\mu} + F_{\nu}^{\mu}c/E_0$.

Appendix B: Fields of a uniformly moving charge

The displacement fields of the ion in the CM frame are given by the Lorentz transformation of the rest frame fields (29) and (28). The transformation equations are [29]

$$\mathbf{D}_{\parallel} = \mathbf{D}'_{\parallel}, \quad (65)$$

$$\mathbf{H}_{\parallel} = \mathbf{H}'_{\parallel}, \quad (66)$$

$$\mathbf{D}_{\perp} = \gamma \left(\mathbf{D}'_{\perp} - \frac{\mathbf{v}}{c^2} \times \mathbf{H}' \right), \quad (67)$$

$$\mathbf{H}_{\perp} = \gamma \left(\mathbf{H}'_{\perp} + \mathbf{v} \times \mathbf{D}' \right). \quad (68)$$

Unprimed fields represent the fields in the CM frame, while primes refer to quantities in the rest frame. We apply the Lorentz boost along the $-\mathbf{X}_{\parallel}$ direction so that the charge is seen moving along the

$+\mathbf{X}_{\parallel}$ direction. Therefore, \mathbf{X}_{\perp} and \mathbf{X}_{\parallel} represent the transverse and parallel components of \mathbf{X} with respect to the ion velocity, \mathbf{v} .

The coordinate transformation is

$$X^0 = \gamma \mathbf{v} \cdot \mathbf{X}'/c, \quad (69)$$

$$\mathbf{X}_{\parallel} = \gamma \mathbf{X}'_{\parallel}, \quad (70)$$

$$\mathbf{X}_{\perp} = \mathbf{X}'_{\perp}. \quad (71)$$

The terms t' that usually appear in the Lorentz transformation are absent above because the relative position vector $\mathbf{X}' = \mathbf{x}' - \mathbf{x}'_0$ is taken to be the difference of two vectors at the equal time in the rest frame. In (69), X^0 is the difference in time between the observation point \mathbf{x} and the source point \mathbf{x}_0 in the CM frame.

In the CM frame, we then find our boosted fields to be

$$\mathbf{D}(t, \mathbf{x}) = \frac{Ze}{4\pi R^3} \frac{\gamma}{R^3} \mathbf{X}, \quad (72)$$

$$\mathbf{H}(t, \mathbf{x}) = \frac{Ze}{4\pi R^3} \frac{\gamma}{R^3} \mathbf{v} \times \mathbf{X}, \quad (73)$$

where

$$R = \sqrt{\mathbf{X}_{\perp}^2 + \gamma^2 \mathbf{X}_{\parallel}^2}. \quad (74)$$

Note that all components of the fields (72) and (73) obtain the factor γ under the transformation. The parallel components acquire the factor γ from the coordinate transformation, while γ in the transverse components comes from the field transformation.

To validate our results for the fields \mathbf{D} and \mathbf{H} of the charge in uniform motion, obtained by the Lorentz transformation, we show that (72) and (73) are consistent with the Lienard–Wiechert solution of Maxwell’s equations. Starting with the Lienard–Wiechert solution, it can be shown that the electric field of a charge in uniform motion can be written as (see [34], page 105–106, Example 4.5)

$$\mathbf{E} = \frac{Ze}{4\pi\epsilon_0\gamma^2} \frac{\mathbf{X}}{|\mathbf{X}|^3 \left(1 - \frac{v^2}{c^2} \sin^2(\psi)\right)^{\frac{3}{2}}}, \quad (75)$$

where ψ is defined as the angle between the position vector \mathbf{X} and the direction of motion,

$$|\mathbf{X}_{\perp}| = |\mathbf{X}| \sin(\psi). \quad (76)$$

The denominator in (75) can be re-written as

$$\begin{aligned} |\mathbf{X}|^3 \left(1 - \frac{v^2}{c^2} \sin^2(\psi)\right)^{\frac{3}{2}} &= |\mathbf{X}|^3 \frac{\left(|\mathbf{X}|^2 - \frac{v^2}{c^2} \mathbf{X}_{\perp}^2\right)^{\frac{3}{2}}}{|\mathbf{X}|^3} = \\ \left(\mathbf{X}_{\parallel}^2 + \mathbf{X}_{\perp}^2/\gamma^2\right)^{\frac{3}{2}} &= \frac{R^3}{\gamma^3}. \end{aligned} \quad (77)$$

Then (75) becomes

$$\mathbf{E} = \frac{Ze}{4\pi\epsilon_0} \frac{\gamma}{R^3} \mathbf{X}, \quad (78)$$

which is consistent with (30). The magnetic field is obtained by $\mathbf{B} = \mathbf{v} \times \mathbf{E}/c^2$, which is consistent with (31).

References

- [1] V. Voronyuk, V.D. Toneev, W. Cassing, E.L. Bratkovskaya, V.P. Konchakovski, S.A. Voloshin, *Phys. Rev. C* **83**, 054911 (2011).
- [2] L. Oliva, P. Moreau, V. Voronyuk, E. Bratkovskaya, *Phys. Rev. C* **101**, 014917 (2020).
- [3] K. Mazurek, M. Khusek-Gawenda, A. Szczurek, *Eur. Phys. J. A* **58**, 245 (2022).
- [4] M. Khusek-Gawenda, P. Lebiedowicz, A. Szczurek, *Phys. Rev. C* **93**, 044907 (2016).
- [5] M. Aaboud, G. Aad, B. Abbott et al. (ATLAS Collaboration), *Nat. Phys.* **13**, 852 (2017).
- [6] G. Baur, K. Hencken, A. Aste, D. Trautmann, S.R. Klein, *Nucl. Phys. A* **729**, 787 (2003).
- [7] M. Born, L. Infeld, *Proc. Roy. Soc. Lond. A* **144**, 425 (1934).
- [8] I. Bialynicki-Birula, in: *Quantum Theory of Particles and Fields*, Ed. B. Jancewicz, J. Lukierski, World Scientific, Singapore 1983.
- [9] J. Rafelski, L.P. Fulcher, W. Greiner, *Nuovo Cim. B* **7**, 137 (1972).
- [10] G. Boillat, *J. Math. Phys.* **11**, 941 (1970).
- [11] Z. Bialynicka-Birula, *Bull. Acad. Pol. Sci. Phys. Astron.* **27**, 41 (1979).
- [12] M.K.H. Kiessling, *J. Statist. Phys.* **116**, 1057 (2004).
- [13] D. Chruscinski, *Phys. Lett. A* **240**, 8 (1998).
- [14] E. Feenberg, *Phys. Rev.* **47**, 148 (1935).
- [15] P.A.M. Dirac, *Proc. Roy. Soc. Lond. A Math. Phys. Sci.* **257**, 32 (1960).
- [16] J.S. Schwinger, *Phys. Rev.* **82**, 664 (1951).
- [17] J. Rafelski, G. Soff, W. Greiner, *Phys. Rev. A* **7**, 903 (1973).
- [18] J. Rafelski, L.P. Fulcher, W. Greiner, *Phys. Rev. Lett.* **27**, 958 (1971).
- [19] M.K.H. Kiessling, *J. Statist. Phys.* **116**, 1123 (2004).
- [20] H. Carley, M.K.H. Kiessling, *Phys. Rev. Lett.* **96**, 030402 (2006).
- [21] J.M. Davila, C. Schubert, M.A. Trejo, *Int. J. Mod. Phys. A* **29**, 1450174 (2014).
- [22] T. Dereli, R.W. Tucker, *Europhys. Lett.* **89**, 20009 (2010).
- [23] D.A. Burton, R.M.G.M. Trines, T.J. Walton, H. Wen, *J. Phys. A* **44**, 095501 (2011).
- [24] W. Heisenberg, H. Euler, *Z. Phys.* **98**, 714 (1936).
- [25] S. Evans, J. Rafelski, “Improving Euler–Heisenberg–Schwinger effective action with dressed photons” (in preparation).
- [26] W. Price, M. Formanek, J. Rafelski, *Phys. Rev. D* **105**, 016024 (2022).
- [27] M.H.L. Pryce, *Proc. Roy. Soc. Lond. A* **155**, 597 (1936).
- [28] R. Ferraro, M.E. Lipchak, *Phys. Rev. E* **77**, 046601 (2008).
- [29] J.D. Jackson, *Classical Electrodynamics*, 3rd ed., John Wiley & Sons, Hoboken (NJ) 1999.
- [30] J. Bicak, D. Kubiznak, *Phys. Lett. B* **636**, 142 (2006).
- [31] J.R. Dormand, P.J. Prince, *J. Computat. Appl. Math.* **6**, 19 (1980).
- [32] J. Zhao, J. Chen, X.G. Huang, Y.G. Ma, [arXiv:2211.03968](https://arxiv.org/abs/2211.03968), 2022.
- [33] ALICE Collaboration, [arXiv:2209.04250](https://arxiv.org/abs/2209.04250), 2022.
- [34] F. Melia, *Electrodynamics*, The University of Chicago Press, Chicago 2001.

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

Classicality of the Bogoliubov Transformations and the Dynamical Casimir Effect Through the Reduced State of the Field

T. LINOWSKI^{a,*} AND Ł. RUDNICKI^{a,b}^a*International Centre for Theory of Quantum Technologies, University of Gdańsk, Jana Bążyńskiego 1A, 80-309 Gdańsk, Poland*^b*Center for Theoretical Physics, Polish Academy of Sciences, al. Lotników 32/46, 02-668 Warszawa, Poland*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S95](https://doi.org/10.12693/APhysPolA.143.S95)*e-mail: t.linowski95@gmail.com

We use the reduced state of the field formalism (*Entropy* **21**, 705 (2019)) to derive conditions under which a Bogoliubov transformation can be considered semi-classical. We apply this result to the dynamical Casimir effect in a moving medium (*Phys. Rev. A* **78**, 042109 (2008)), discussing its classical and quantum features.

topics: Bogoliubov transformations, dynamical Casimir effect, quantum-to-classical transitions, mesoscopic degrees of freedom

1. Introduction

Arguably one of the most surprising predictions of quantum field theory is the Casimir effect, a physical force arising solely from the presence of quantum fluctuations in the vacuum [1–3]. Since its original formulation in 1948 [4], the phenomenon has garnered a lot of interest, in particular giving rise to many alternative formulations and generalizations. One such generalization, dubbed the dynamical Casimir effect, predicts the spontaneous production of particles in a medium following from non-trivial time dependence of either its boundary or its material coefficients [5–8].

In 2008, Professor Iwo Białynicki-Birula working together with Professor Zofia Białynicka-Birula[†] established a third mechanism generating the dynamical Casimir effect — oscillatory motion of

a medium [9–11]. In fact, this mechanism is more general and applies to all kinds of motion, as long as its speed varies in time, and one carefully picks the “incoming” and “outgoing” annihilation and creation operators (see an example of a uniformly accelerated medium [12]). A loosely related phenomenon occurs around large rotating and/or gravitating bodies [13].

The dynamical Casimir effect is obtained by performing a Bogoliubov transformation, i.e., a linear transformation of the creation and annihilation operators of the quantum field preserving canonical commutation relations [14]. If the Casimir effects are among the most interesting phenomena in quantum theory, Bogoliubov transformations are among its most reliable tools. Originally used to describe superconductivity [15, 16], today they are widely used in many branches of quantum physics, from optics and theories of magnetism to field theory in a curved spacetime (Unruh effect, Hawking radiation) [14, 17–19].

While the most prominent applications of the Bogoliubov transformations suggest the latter to be inherently quantum, we observe that from the formal point of view, Bogoliubov transformations are

[†]This is a good opportunity to acknowledge the fact that 59 papers out of a total of 206 so far published by Professor Białynicki-Birula, as well as the comprehensive textbook on quantum electrodynamics [9], have been written in this admirable collaboration which started as early as 1957 [10].

essentially equivalent to a change of basis of the Hilbert space. For this reason, one may expect that at least some Bogoliubov transformations could have classical analogs, similar to local unitary rotations of the Hilbert space, which do not entangle the system. If so, this could shed new conceptual light on the phenomena described by them.

In this paper, we derive an exact set of conditions under which Bogoliubov transformations can be considered semi-classical. By semi-classical (further also referred to as just “classical”), we understand models which can be described by certain kinetic equations for reduced single-particle states and their displacements — so-called *reduced state of the field* formalism [20]. This framework has recently been proven to be an effective tool in probing the classicality of quantum Gaussian evolution [21].

In the case of isolated systems, the transformations allowed by our conditions turn out to have a simple interpretation in terms of passive operations, which correspond to classical devices such as beam splitters. In the case of open systems, the conditions are less restrictive, which we interpret as some of the total dynamics’ “quantumness” being encoded into the environment. Our findings allow us to conduct an in-depth discussion of the classicality of the dynamical Casimir effect derived in [11]. We find that, while the overall phenomenon is quantum in nature, the individual photons experience each other as semi-classical dissipative effects.

This paper is organized as follows. In Sect. 2, we introduce the dynamical Casimir effect in moving media. In Sect. 3, we briefly summarize the most important properties of our main tool — the reduced state of the field (RSF). In Sect. 4, we derive our main results, namely classicality conditions for Bogoliubov transformations. In Sect. 5, we build upon these findings to assess the classicality of the dynamical Casimir effect. We conclude in Sect. 6.

2. Dynamical Casimir effect in a moving medium

The electromagnetic field is fully described by the set of four three-component vectors, \mathbf{D} and \mathbf{E} , describing the electric field, along with \mathbf{B} and \mathbf{H} , describing the magnetic field, which altogether fulfill the Maxwell equations in vacuum [22, 23]

$$\begin{aligned} \partial_t \mathbf{D}(\mathbf{r}, t) &= \nabla \times \mathbf{H}(\mathbf{r}, t), \\ \nabla \cdot \mathbf{D}(\mathbf{r}, t) &= 0, \\ -\partial_t \mathbf{B}(\mathbf{r}, t) &= \nabla \times \mathbf{E}(\mathbf{r}, t), \\ \nabla \cdot \mathbf{B}(\mathbf{r}, t) &= 0. \end{aligned} \tag{1}$$

In the Heisenberg picture, the operators associated with these fields fulfill exactly the same set of equations.

Assuming the field propagates through a homogeneous, isotropic medium moving with a velocity \mathbf{v} and characterized by constant material coefficients μ, ϵ , the field vectors are related by the Minkowski constitutive relations [24]

$$\begin{aligned} \mathbf{D} + \frac{\mathbf{v}}{c^2} \times \mathbf{H} &= \epsilon (\mathbf{E} + \mathbf{v} \times \mathbf{B}), \\ \mathbf{B} - \frac{\mathbf{v}}{c^2} \times \mathbf{E} &= \mu (\mathbf{H} - \mathbf{v} \times \mathbf{D}), \end{aligned} \tag{2}$$

where c is the speed of light.

In the convenient Riemann–Silberstein approach (see a review [25]), the electromagnetic field is combined into two vectors

$$\begin{aligned} \mathbf{F} &:= \frac{1}{\sqrt{2}\epsilon} \mathbf{D} + \frac{i}{\sqrt{2}\mu} \mathbf{B}, \\ \mathbf{G} &:= \frac{1}{\sqrt{2}\mu} \mathbf{E} + \frac{i}{\sqrt{2}\epsilon} \mathbf{H}. \end{aligned} \tag{3}$$

The advantage of this approach can already be seen in the considered problem, as the constitutive relations (2) can always be solved for \mathbf{G} , yielding

$$\mathbf{G} = \frac{c}{n} \left[\mathbf{F} + \frac{n^2-1}{c^2 n^2 - v^2} \mathbf{v} \times (\mathbf{v} \times \mathbf{F} + i c n \mathbf{F}) \right], \tag{4}$$

where $n := c\sqrt{\epsilon\mu} \geq 1$ is the refractive index of the medium. Then, assuming position-independent velocity, $\mathbf{v}(\mathbf{r}, t) = c\boldsymbol{\beta}(t)$, the vacuum Maxwell equations (1) reduce to just one equation

$$\begin{aligned} \partial_t \mathbf{F} &= -i c \delta(t) \left(\boldsymbol{\beta}(t) \cdot \nabla \right) \mathbf{F} + \frac{c}{n} \alpha(t) \nabla \times \mathbf{F} \\ &\quad - \frac{c}{n} \delta(t) \boldsymbol{\beta}(t) \times \nabla (\boldsymbol{\beta}(t) \cdot \mathbf{F}), \end{aligned} \tag{5}$$

where

$$\delta(t) := \frac{n^2-1}{n^2-\beta^2(t)}, \quad \alpha(t) := 1 - \delta(t)\beta^2(t). \tag{6}$$

Under a further assumption that the velocity has a constant direction \mathbf{m} , and with the help of the Fourier decomposition

$$\begin{aligned} \mathbf{F}(\mathbf{r}, t) &= \int \frac{d^3 \mathbf{k}}{\sqrt{(2\pi)^3}} e^{i\mathbf{k} \cdot \mathbf{r} - i\phi(\mathbf{k}, t)} \\ &\quad \times \left[\mathbf{e}(\mathbf{k}) f_+(\mathbf{k}, t) + \mathbf{e}^*(\mathbf{k}) f_-(\mathbf{k}, t) \right], \end{aligned} \tag{7}$$

where \mathbf{e} are elliptic polarization vectors [11], the Maxwell equations lead to a pair of ordinary differential equations for the functions f_{\pm}

$$\begin{aligned} \partial_t \hat{f}_{\pm}(\mathbf{k}, t) &= \\ &\quad \mp i\omega(\mathbf{k}) \left[\eta_+(\mathbf{k}, t) \hat{f}_{\pm}(\mathbf{k}, t) - \eta_-(\mathbf{k}, t) \hat{f}_{\mp}(\mathbf{k}, t) \right], \end{aligned} \tag{8}$$

with

$$\begin{aligned} \eta_{\pm}(\mathbf{k}, t) &:= \frac{1}{2} \left[\frac{\alpha(t)}{\sigma^2(\mathbf{k})} \pm \sigma^2(\mathbf{k}) \Delta(\mathbf{k}, t) \right], \\ \Delta(\mathbf{k}, t) &:= 1 - \delta(t)\beta^2(t) \cos^2(\theta(\mathbf{k})). \end{aligned} \tag{9}$$

The parameter θ denotes the angle between the wave vector \mathbf{k} and the velocity direction \mathbf{m} , while σ is a *free* real parameter defining the polarization geometry. Last but not least, the phase

$$\phi(\mathbf{k}, t) := \omega(\mathbf{k}) \cos(\theta(\mathbf{k})) \int_0^t d\tau \delta(\tau) \beta(\tau) \quad (10)$$

has been extracted to achieve a simplification of the resulting equations ($\beta = |\beta|$).

To obtain the dynamical Casimir effect, it is assumed that the medium is moving with a time-dependent velocity from time $t = 0$ up to $t = T$ [11, 12]. If the medium just before and after was “still” (characterized by $\beta(t) = \text{const}$), the corresponding operators \hat{f}_\pm , after a suitable choice of σ [11], can be interpreted in terms of the creation and annihilation operators of photons with right helicity

$$\hat{f}_+(\mathbf{k}, t) = \begin{cases} \sqrt{\hbar\omega(\mathbf{k})} \hat{a}_{R,\text{in}}(\mathbf{k}) e^{-i\omega(\mathbf{k})t}, & t < 0 \\ \sqrt{\hbar\omega(\mathbf{k})} \hat{a}_{R,\text{out}}(\mathbf{k}) e^{-i\omega(\mathbf{k})(t-T)}, & t > T \end{cases}, \quad (11)$$

and left helicity

$$\hat{f}_-(\mathbf{k}, t) = \begin{cases} \sqrt{\hbar\omega(\mathbf{k})} \hat{a}_{L,\text{in}}^\dagger(-\mathbf{k}) e^{i\omega(\mathbf{k})t}, & t < 0 \\ \sqrt{\hbar\omega(\mathbf{k})} \hat{a}_{L,\text{out}}^\dagger(-\mathbf{k}) e^{i\omega(\mathbf{k})(t-T)}, & t > T \end{cases}. \quad (12)$$

Here, $\hat{a}_{L/R,\text{in/out}}$ and their Hermitian conjugates fulfill all the expected properties of the standard annihilation and creation operators. Note that such interpretation is not possible during the acceleration period $t \in [0, T]$ itself, due to the impossibility of separation into positive and negative frequency parts.

The final operators are given by the initial ones via the relation [11]

$$\begin{aligned} \hat{a}_{R,\text{out}}(\mathbf{k}) &= e^{-i\phi} \left[f_{R+} \hat{a}_{R,\text{in}}(\mathbf{k}) + f_{R-} \hat{a}_{L,\text{in}}^\dagger(-\mathbf{k}) \right], \\ \hat{a}_{L,\text{out}}^\dagger(-\mathbf{k}) &= e^{-i\phi} \left[f_{L+} \hat{a}_{R,\text{in}}(\mathbf{k}) + f_{L-} \hat{a}_{L,\text{in}}^\dagger(-\mathbf{k}) \right], \end{aligned} \quad (13)$$

where $\phi \equiv \phi(\mathbf{k}, T)$, while $f_{L\pm} \equiv f_{L\pm}(\mathbf{k}, T)$, $f_{R\pm} \equiv f_{R\pm}(\mathbf{k}, T)$ are solutions to the differential equations (8) subject to initial conditions

$$\begin{aligned} f_{R+}(\mathbf{k}, 0) &= f_{L-}(\mathbf{k}, 0) = 1, \\ f_{R-}(\mathbf{k}, 0) &= f_{L+}(\mathbf{k}, 0) = 0. \end{aligned} \quad (14)$$

It is worth adding that, due to the canonical commutation relations for the outgoing photons (13)

$$\begin{aligned} [\hat{a}_{R,\text{out}}(\mathbf{k}), \hat{a}_{R,\text{out}}^\dagger(\mathbf{k})] &= 1, \\ [\hat{a}_{R,\text{out}}(\mathbf{k}), \hat{a}_{R,\text{out}}(\mathbf{k})] &= 0, \end{aligned} \quad (15)$$

we have

$$|f_{R+}|^2 = |f_{R-}|^2 + 1, \quad (16)$$

with an analogous relation for f_{L+} and f_{L-} .

Let us remark that in the original work [11], the functions f were denoted as $f_\pm^1 \equiv f_{R\pm}$ and $f_\pm^2 \equiv f_{L\pm}$. Here, we change the notation to make the connection to photon helicity more immediate, as well as to avoid confusing the indices with exponentiation. We stress, however, that despite corresponding to different photon helicities, the two pairs of functions are interrelated via the initial conditions and have to be considered together.

The Casimir effect is finally obtained by considering the system initially in the vacuum and computing the photon number densities after the motion

$$\begin{aligned} \langle \hat{n}_R(T) \rangle &= \langle 0 | \hat{a}_{R,\text{out}}^\dagger(\mathbf{k}) \hat{a}_{R,\text{out}}(\mathbf{k}) | 0 \rangle = \\ &= |f_{R-}(\mathbf{k}, T)|^2 \delta(0), \\ \langle \hat{n}_L(T) \rangle &= \langle 0 | \hat{a}_{L,\text{out}}^\dagger(\mathbf{k}) \hat{a}_{L,\text{out}}(\mathbf{k}) | 0 \rangle = \\ &= |f_{L+}(\mathbf{k}, T)|^2 \delta(0), \end{aligned} \quad (17)$$

where $\delta(0)$ is the Dirac delta singularity. Note that, due to the symmetry of the evolution equations governing the left and right helicity functions, the two densities are, in fact, equal

$$\langle \hat{n}_R(T) \rangle = \langle \hat{n}_L(T) \rangle \equiv \langle \hat{n}(T) \rangle. \quad (18)$$

As was verified in [11, 12], at least for some \mathbf{k} , this number is a growing function of T . Therefore, the motion of the medium results in potentially unbounded particle production in the vacuum and, hence, the prediction of the dynamical Casimir effect.

Transformation (13) at the heart of the discussed phenomenon is an example of a *Bogoliubov transformation* [15, 16], namely a linear transformation $\{\hat{a}_n, \hat{a}_n^\dagger\} \rightarrow \{\hat{a}'_n, \hat{a}'_n^\dagger\}$ of the creation and annihilation operators preserving the canonical commutation relations [14]. As the main result of this paper, we will derive the precise conditions under which such transformations can be considered semiclassical, with special emphasis put on the classicality of the dynamical Casimir effect in a moving medium.

3. Reduced state of the field

To assess the (semi)classicality of Bogoliubov transformations, we first need to define a sensible criterion for what is classical. To this end, we will employ the mesoscopic formalism of the reduced state of the field (RSF) [20], which was already used for similar purposes before [21]. Since the framework itself is not the main focus of our study, here we provide only basic information about it. For more details, see the introduction of the formalism by Robert Alicki in [20], its semi-classical interpretation in [21], and its application to thermodynamics in [26].

We consider an N -mode, continuous variable Hilbert space described by a set of N annihilation and creation operators $\hat{a}_k, \hat{a}_k^\dagger$ fulfilling the

canonical commutation relations

$$[\hat{a}_k, \hat{a}_{k'}^\dagger] = \delta_{kk'}, \quad [\hat{a}_k, \hat{a}_{k'}] = [\hat{a}_k^\dagger, \hat{a}_{k'}^\dagger] = 0. \quad (19)$$

As always, an arbitrary n -particle state in the many-body Hilbert space can be constructed by acting on the vacuum state with n appropriate creation operators. Since, in principle, the number of particles in a given mode can be arbitrary, the N -mode Hilbert space is infinitely dimensional, and so is the density operator $\hat{\rho}$ constituting the full quantum description of the system.

In some cases, however, the full quantum formalism is not necessary and can be replaced by a simpler, mesoscopic framework. For example, Gaussian states and dynamics can be efficiently studied in the symplectic picture [27–29]. Similarly, to describe macroscopic fields and associated evolution, a formalism called *reduced state of the field* (RSF) has been recently developed [20].

In the RSF framework, instead of by the density operator, the system is described by the pair $(r, |\alpha\rangle)$. Here,

$$r := \sum_{k,k'=1}^N \text{Tr} [\hat{\rho} \hat{a}_{k'}^\dagger \hat{a}_k] |k\rangle \langle k'| \quad (20)$$

is the *single-particle density matrix*, while the *averaged field* equals

$$|\alpha\rangle := \sum_{k=1}^N \text{Tr} [\hat{\rho} \hat{a}_k] |k\rangle. \quad (21)$$

The single-particle density matrix contains the simplest non-local information about the system. Additionally, its diagonal elements equal the mean particle numbers $r_{kk} = \langle \hat{a}_k^\dagger \hat{a}_k \rangle$. Consequently, the matrix is normalized to the mean total particle number. Note that, by construction, the single-particle density matrix is non-negative. The averaged field, on the other hand, contains additional local information.

Much like the previously mentioned symplectic picture requires observables and transformations that are Gaussian, the RSF formalism employs observables that are either *additive* [20]

$$\hat{O} = \sum_{k,k'=1}^N o_{kk'} \hat{a}_k^\dagger \hat{a}_{k'}, \quad (22)$$

or *linear*

$$\hat{\sigma} = \sum_{k=1}^N \left(\sigma_k^* \hat{a}_k + \sigma_k \hat{a}_k^\dagger \right). \quad (23)$$

In the case of macroscopic fields, which are usually modeled as non-interacting fields with dynamics governed by equations linear in creation and annihilation operators, the most relevant observables are of this form. For example, the Hamiltonian is additive, while the position and momentum operators are linear.

Defining the *reduced observables* corresponding to (22) and (23) as

$$o = \sum_{k,k'=1}^N o_{kk'} |k\rangle \langle k'|, \quad |\sigma\rangle = \sum_{k=1}^N \sigma_k |k\rangle, \quad (24)$$

we can indeed see that the associated expectation values can be rewritten in the RSF formalism as [21]

$$\text{Tr} [\hat{\rho} \hat{O}] = \text{tr}(r o), \quad \text{Tr} [\hat{\rho} \hat{\sigma}] = \langle \sigma | \alpha \rangle + \langle \alpha | \sigma \rangle. \quad (25)$$

The RSF framework comes equipped with dedicated entropy measures and evolution equations, both derived from the standard quantum description. In the case of entropy, we have the reduced von Neumann and Wehrl entropies [20, 21]

$$s_v(r, |\alpha\rangle) := \text{tr} [(r_\alpha + \mathbf{1}_N) \ln (r_\alpha + \mathbf{1}_N) - r_\alpha \ln (r_\alpha)],$$

$$s_w(r, |\alpha\rangle) := \text{tr} [\ln (r_\alpha + \mathbf{1}_N)] + N, \quad (26)$$

where $r_\alpha := r - |\alpha\rangle \langle \alpha|$ and $\mathbf{1}_N$ denotes the identity matrix in dimension N . The reduced entropies arise from applying the maximum entropy principle to the standard von Neumann and Wehrl entropies, respectively [30, 31].

Finally, RSF evolves according to the *reduced kinetic equations* [20, 21]

$$\frac{dr}{dt} = -\frac{i}{\hbar} [h, r] + |\zeta\rangle \langle \alpha| + |\alpha\rangle \langle \zeta|$$

$$+ \frac{1}{2} \{ \gamma_\uparrow - \gamma_\downarrow, r \} + \gamma_\uparrow + \sum_j \eta_j (u_j r u_j^\dagger - r),$$

$$\frac{d|\alpha\rangle}{dt} = -\frac{i}{\hbar} h |\alpha\rangle + \frac{1}{2} (\gamma_\uparrow - \gamma_\downarrow) |\alpha\rangle + |\zeta\rangle$$

$$+ \sum_j \eta_j (u_j - 1) |\alpha\rangle, \quad (27)$$

which are derived from the Gorini–Kossakowski–Lindblad–Sudarshan (GKLS) equation [32, 33] under the assumption that the considered quantum field can be treated as a set of individual particles subject to spontaneous decay and production, as well as interaction with coherent classical sources and random scattering by the environment. The operators entering (27) represent

- The Hamiltonian

$$h := \hbar \sum_{k=1}^N \omega_k |k\rangle \langle k|, \quad \omega_k \geq 0; \quad (28)$$

- Coherent sources

$$|\zeta\rangle := \sum_{k=1}^N \zeta_k |k\rangle; \quad (29)$$

- Particle creation rates

$$\gamma_\uparrow = \sum_{k,k'=1}^N \gamma_\uparrow^{kk'} |k\rangle \langle k'|, \quad \gamma_\uparrow \geq 0, \quad (30)$$

and analogously particle annihilation rates γ_\downarrow ;

- Unitary interactions with rates $\eta_j \geq 0$ ($\sum_j \eta_j = 1$)

$$u_j = \sum_{k,k'=1}^N u_j^{kk'} |k\rangle \langle k'|, \quad u_j^\dagger u_j = u_j u_j^\dagger = \mathbf{1}_N. \quad (31)$$

For a large number of non-commuting unitaries, the last term in either of the reduced kinetic equations represents random scattering.

Note that, while not explicitly stated in the original work [20], it is clear from the derivation of the reduced kinetic equations that all the quantities entering it may be time-dependent, provided they fulfill the respective constraints (imposed by complete positivity) during every instant of the evolution.

Although RSF was originally designed to capture the quantum features of macroscopic fields, it has been recently shown to have a convincing interpretation as a semi-classical description of bosonic many-body systems [21]. For example, it was proved that the RSF formalism contains no information about distillable entanglement in the system, and that both of the reduced entropies are akin to Wehrl's semi-classical entropy [31], typically considered as such due to its close association with the phase-space.

Because, by construction, the reduced kinetic equations (27) preserve the RSF formalism's semi-classical set of degrees of freedom, any time evolution model of the density operator, which can be rewritten as reduced kinetic equations, must be necessarily semi-classical itself. Based on this principle, in [21], quantum Gaussian evolution of light was found to be classical if and only if it consisted strictly of so-called passive optical transformations, e.g., beam splitting and phase shifting. Contrary to their remaining active counterparts, such as quantum squeezing, passive transformations can be understood operationally by treating light as a classical wave. In this paper, we adopt a similar methodology for Bogoliubov transformations — if they preserve the set of the degrees of freedom contained within RSF, we will regard them as semi-classical, and if not, we will regard them as inherently quantum.

4. Classicality of Bogoliubov transformations

We are now equipped with the tools necessary to assess the classicality of Bogoliubov transformations. We will consider two distinct cases, i.e., Bogoliubov transformations in isolated (closed) systems and in open systems. The main results of this section are presented in Propositions 1–3, with proofs in Appendix A.

4.1. Isolated system

In the case of an isolated system, the most general transformation of the density operator is unitary

$$\hat{\rho}' = \hat{U} \hat{\rho} \hat{U}^\dagger. \quad (32)$$

For the transformation to be of the Bogoliubov-type, \hat{U} must be such that, for some complex matrix \mathcal{X} ,

$$\hat{A}'_n := \hat{U}^\dagger \hat{A}_n \hat{U} = \sum_{m=1}^{2N} \mathcal{X}_{nm} \hat{A}_m, \quad (33)$$

with

$$\hat{A}_n := \begin{cases} \hat{a}_n, & n \in \{1, \dots, N\}, \\ \hat{a}_n^\dagger, & n \in \{N+1, \dots, 2N\}. \end{cases} \quad (34)$$

To preserve the canonical commutation relations, the matrix \mathcal{X} has to fulfill the so-called symplectic property [34, 35]

$$\mathcal{X} \mathcal{S} \mathcal{X}^\dagger = \mathcal{S}, \quad (35)$$

where $\mathcal{S} = \text{diag}[\mathbf{1}_N, -\mathbf{1}_N]$. As a consequence of the symplectic property,

$$\mathcal{X} = \begin{bmatrix} \mathcal{X}_\uparrow & \mathcal{X}_\downarrow \\ \mathcal{X}_\downarrow^* & \mathcal{X}_\uparrow^* \end{bmatrix}, \quad (36)$$

where \mathcal{X}_\uparrow are of size $N \times N$.

Calculating the change in RSF implied by a generic Bogoliubov transformation and forcing the result to be fully contained within the formalism, we obtain the classicality conditions for the closed system Bogoliubov transformations. Furthermore, if the unitary transformation in (32) depends smoothly on time, then so does the matrix \mathcal{X} , turning the discrete Bogoliubov transformation into a continuous *Bogoliubov evolution*. In such a case, the density operator can be differentiated with respect to time, and the resulting evolution equation can be compared with the reduced kinetic equations.

Proceeding in this way, we obtain our first major result.

Proposition 1. Isolated system Bogoliubov transformations (as described above) are compatible with the RSF formalism and are thus classical with respect to it if and only if

$$0 = \mathcal{X}_\downarrow. \quad (37)$$

Additionally, if the transformation depends smoothly on time, the corresponding reduced kinetic equations (27) exist and are governed by

$$h = \frac{i\hbar}{2} \left(\frac{d\mathcal{X}_\uparrow}{dt} \mathcal{X}_\uparrow^{-1} - \mathcal{X}_\uparrow^{-\dagger} \frac{d\mathcal{X}_\uparrow^\dagger}{dt} \right), \quad (38)$$

with the remaining terms vanishing.

Proof. See Appendix A. \square

The obtained classicality condition is easy to interpret. Substituting (37) into the symplectic condition (35), we immediately find that \mathcal{X} is also unitary in addition to being symplectic, which means that it is passive. Thus, in a complete analogy to quantum Gaussian evolution [21], Bogoliubov transformations in isolated systems are semi-classical only if they correspond to passive transformations.

Let us also remark that while the absence of the dissipative terms in the obtained reduced kinetic equations was to be expected in an isolated system, the lack of coherent classical sources was not. Indeed, it is easy to see that this lack is not

a fundamental property of the Bogoliubov evolution, but rather a consequence of the Bogoliubov transformations (33) being defined, for simplicity, without constant terms (independent of the creation and annihilation operators).

4.2. Open system

In the more general case of an open system, the total density operator of the system and environment (also called bath) is as well transformed according to (32). However, we are only interested in the state of the system, given by a partial trace over the degrees of freedom of the environment:

$$\hat{\rho}_S = \text{Tr}_E [\hat{\rho}]. \quad (39)$$

The Bogoliubov transformation itself (33) remains the same. Still, assuming the system and the environment span N_S and N_E modes respectively, it is convenient to additionally split the matrices entering the block decomposition (36) into

$$\mathcal{X}_\uparrow = \begin{bmatrix} \mathcal{X}_{\uparrow S} & \mathcal{X}_{\uparrow C} \\ \mathcal{X}_{\uparrow C'} & \mathcal{X}_{\uparrow E} \end{bmatrix}, \quad \mathcal{X}_\downarrow = \begin{bmatrix} \mathcal{X}_{\downarrow S} & \mathcal{X}_{\downarrow C} \\ \mathcal{X}_{\downarrow C'} & \mathcal{X}_{\downarrow E} \end{bmatrix}, \quad (40)$$

where $\mathcal{X}_{\uparrow S}$ is an $N_S \times N_S$ matrix associated with the system, $\mathcal{X}_{\uparrow E}$ is an $N_E \times N_E$ matrix associated with the environment, and $\mathcal{X}_{\uparrow C}$, $\mathcal{X}_{\uparrow C'}$ are appropriately-sized matrices associated with both. Note that the case of the closed system can be retrieved easily by setting $N_E = 0$ (which, in particular, implies $\mathcal{X}_\uparrow = \mathcal{X}_{\uparrow S}$) and dropping the then-redundant lower indices S .

For a generic initial state of the bath-system ensemble, the dynamics of the latter cannot be separated from the dynamics of the former, making it impossible to even compare with the RSF formalism. Nonetheless, even in this completely general setting, we were able to derive necessary conditions for classicality of Bogoliubov transformations.

Proposition 2. Open system Bogoliubov transformations (as described above) can be compatible with the RSF formalism and thus be classical with respect to it only if

$$0 = \mathcal{X}_{\downarrow S}. \quad (41)$$

Proof. See Appendix A. \square

Unlike the condition (37) for the closed system, the classicality condition for the open system is difficult to interpret. However, comparing it with its closed system counterpart, we can at least see that the latter is much more restrictive: it requires the whole matrix \mathcal{X}_\downarrow to vanish, while the former requires only its system part $\mathcal{X}_{\downarrow S}$ to vanish. Therefore, depending on how we define the degrees of freedom of the system, we may find the same total dynamics to be either classical or quantum from the point of view of the system. This will indeed be the case in

the next section, where we will find that the dynamical Casimir effect falls exactly into this category.

Still, any such interpretation has to be made with care, since it must be stressed that the condition (41) is not equivalent to classicality, but only necessary for it. In stark contrast to the closed system, in the case of an open system, whether or not a given Bogoliubov transformation is classical from the point of view of RSF depends not only on the matrix \mathcal{X} defining it, but also on the total initial state of the system-environment ensemble. It is possible that, for particularly strongly correlated total initial states, the only semi-classical Bogoliubov transformations are those that induce completely separate dynamics for the system and environment, essentially defying the notion of an open system.

To make stronger statements, we are therefore forced to make some restrictions. Firstly, we assume that the initial total state is separable with respect to the bipartition between the system and the bath. This is a typical assumption in the theory of quantum open systems. In particular, the GKLS equation cannot be derived without it [36]. Since, in particular, the reduced kinetic equations governing the time evolution in the RSF formalism are derived from a GKLS equation, it is only natural to also make this assumption in the present case.

Secondly, we assume that the bath is initially in the vacuum state. Note that while this assumption is a very strong one, it is fulfilled by many well-studied and useful models, such as quantum-limited amplification, quantum-limited attenuation, and phase conjugation channels, utilized, e.g., in studies of Gaussianity, entropy, and entanglement [37–39]. More importantly for us, as we will discuss in the next section, it is also satisfied by the dynamical Casimir effect.

Under the above assumptions, we obtain our final main result for Bogoliubov transformations.

Proposition 3. The classicality condition (41) is both necessary and sufficient for open system Bogoliubov transformations with the environment initially in the vacuum state. Additionally, if such transformations depend smoothly on time, the corresponding reduced kinetic equations exist provided

$$\mathcal{W} \geq 0, \quad \mathcal{W} - \mathcal{Y}_r \geq 0 \quad (42)$$

and are governed by

$$h = -\frac{\hbar \mathcal{Y}_i}{2}, \quad \gamma_\downarrow = \mathcal{W}, \quad \gamma_\uparrow = \mathcal{W} - \mathcal{Y}_r, \quad (43)$$

with the remaining terms vanishing. Here,

$$\begin{aligned} \mathcal{Y}_i &:= -i (\mathcal{Y} - \mathcal{Y}^\dagger), \\ \mathcal{Y} &:= \frac{d\mathcal{X}_{\uparrow S}}{dt} \mathcal{X}_{\uparrow S}^{-1}, \\ \mathcal{Y}_r &:= \mathcal{Y} + \mathcal{Y}^\dagger, \\ \mathcal{D} &:= \mathcal{X}_{\downarrow C} \mathcal{X}_{\downarrow C}^\dagger, \\ \mathcal{W} &:= \frac{d\mathcal{D}}{dt} - \mathcal{Y}\mathcal{D} - \mathcal{D}\mathcal{Y}^\dagger. \end{aligned} \quad (44)$$

Proof. See Appendix A. \square

Interestingly, the obtained Bogoliubov reduced kinetic equations do not depend on any components of the matrix \mathcal{X} labeled by the subscripts C' , despite depending on the components labeled by C . At first, this may appear surprising, since a priori both are equally responsible for describing the correlations between the system and the environment. The asymmetry is resolved by interpreting the C components as encoding the influence of the environment on the system, and the C' components as encoding the influence of the system on the environment. The lack of the C' components in the description of the system then becomes expected. As an additional argument for this view, we observe that if we exchanged the roles of the system and the environment, the equations would depend on the C' components, with the C components missing.

Proposition 3 will be our main tool in the study of the classicality of the dynamical Casimir effect. Before we do it, however, let us illustrate our results so far with a short but instructive example — the Gaussian amplification process.

Example (Gaussian amplification process). In the Gaussian amplification process, an arbitrary initial state of the N -mode system

$$\hat{\rho}(t_0) = \int \frac{d^{2N}\mathbf{z}_0}{\pi^N} P_0(\mathbf{z}_0) |\mathbf{z}_0\rangle \langle \mathbf{z}_0| \quad (45)$$

is driven by a heat bath into the state [40]

$$\hat{\rho}(t) = \int \frac{d^{2N}\mathbf{z}_0}{\pi^N} P_0(\mathbf{z}_0) \bigotimes_{j=1}^N \int \frac{d^2z_j}{\pi} \rho_j(t) |z_j\rangle \langle z_j|,$$

$$\rho_j(t) := \frac{1}{n_j(t)} \exp\left(-\frac{|z_j - z_{0j} e^{\kappa_j t}|^2}{n_j(t)}\right). \quad (46)$$

Here, the integration is over the real and imaginary parts of the complex vectors \mathbf{z}_0 , \mathbf{z} ; $P_0(\mathbf{z}_0)$ denotes the Glauber–Sudarshan P representation [41, 42] of the initial state; $|z_j\rangle$ are coherent states; κ_j is the amplification rate of the j -th mode; and

$$n_j(t) := (1 + m_j) (e^{2\kappa_j t} - 1), \quad (47)$$

where m_j is the mean number of photons in the j -th mode of the bath, assumed to be effectively constant throughout the whole process (this is true as long as the bath is much bigger than the system).

The corresponding RSF can be easily calculated

$$r_{kk'}(t) = \int \frac{d^{2N}\mathbf{z}_0}{\pi^N} P_0(\mathbf{z}_0) \prod_{j=1}^N \int \frac{d^2z_j}{\pi} \rho_j(t) z_k z_{k'}^*,$$

$$\alpha_k(t) = \int \frac{d^{2N}\mathbf{z}_0}{\pi^N} P_0(\mathbf{z}_0) \prod_{j=1}^N \int \frac{d^2z_j}{\pi} \rho_j(t) z_k. \quad (48)$$

The integrals over z_j can be performed using the standard result [43]

$$\int \frac{d^{2N}\mathbf{z}}{\pi^N} e^{-z^\dagger \mu z + s^\dagger z + z^\dagger s} = \frac{e^{s^\dagger \mu^{-1} s}}{\det[\mu]}, \quad (49)$$

where μ denotes an invertible matrix and \mathbf{s} is a vector of size N . In our case,

$$\mu^{-1} = n(t) := \sum_{j=1}^N n_j(t) |j\rangle \langle j|,$$

$$\mathbf{s} = n^{-1}(t) |\mathbf{z}_0(t)\rangle,$$

$$|\mathbf{z}_0(t)\rangle := \sum_{j=1}^N z_{0j} e^{\kappa_j t} |j\rangle. \quad (50)$$

This yields

$$r(t) = n(t) + \langle |\mathbf{z}_0(t)\rangle \langle \mathbf{z}_0(t)| \rangle_0,$$

$$|\alpha(t)\rangle = \langle |\mathbf{z}_0(t)\rangle \rangle_0, \quad (51)$$

where $\langle \cdot \rangle_0 := (\pi^{-N}) \int d^{2N}\mathbf{z}_0 P_0(\mathbf{z}_0) (\cdot)$. The formulae (51) induce the following differential evolution equations

$$\frac{dr}{dt} = \frac{1}{2} \left\{ 2\kappa(\mathbf{1} + m) - 2\kappa m, r \right\} + 2\kappa(\mathbf{1} + m),$$

$$\frac{d|\alpha\rangle}{dt} = \frac{1}{2} \left(2\kappa(\mathbf{1} + m) - 2\kappa m \right) |\alpha\rangle, \quad (52)$$

where $m := \sum_{j=1}^N m_j |j\rangle \langle j|$ and $\kappa := \sum_{j=1}^N \kappa_j |j\rangle \langle j|$. Clearly, the equations have the form of reduced kinetic equations (27) with $\gamma_\uparrow = 2\kappa(\mathbf{1} + m)$, $\gamma_\downarrow = 2\kappa m$ and $h = |\zeta\rangle = \mu(d\mathbf{u}) = 0$.

According to Proposition 2, any open system Bogoliubov evolution that can be represented by reduced kinetic equations has to necessarily fulfill the classicality condition (41). To see that this is indeed the case in the Gaussian amplification process, we observe that it is generated by a Bogoliubov transformation of the form [37]

$$\mathcal{X}_\uparrow = \cosh(\kappa t) \begin{bmatrix} \mathbf{1}_N & 0 \\ 0 & \mathbf{1}_N \end{bmatrix},$$

$$\mathcal{X}_\downarrow = \sinh(\kappa t) \begin{bmatrix} 0 & \mathbf{1}_N \\ \mathbf{1}_N & 0 \end{bmatrix}. \quad (53)$$

Clearly, $\mathcal{X}_{\downarrow S}$, being the upper left-hand side block component of \mathcal{X}_\downarrow , vanishes, as required by the aforementioned condition.

The fact that we found the Gaussian amplification process to be semi-classical is not surprising — intuitively, Gaussian amplification can be interpreted as pumping particles into the system until it reaches essentially macroscopic size. The process is well known for turning quantum phenomena into more classical ones. For example, it was previously shown that the Glauber–Sudarshan P distribution of an infinitely amplified state approaches the semi-classical Husimi Q distribution [44, 45]. Similarly, the von Neumann entropy of the maximally amplified state approaches the semi-classical Wehrl entropy [30, 38]. More recently, it has been shown that the amplified Pegg–Barnett phase formalism approaches the Paul phase formalism [46].

5. Classicality of the dynamical Casimir effect

Armed with the classicality conditions (37) and (41), we are now ready to come back to the dynamical Casimir effect. We begin by observing that while the phenomenon spans an infinite number of modes of photons with both helicities, its defining Bogoliubov transformation (13) couples them in pairs only. Any mode \mathbf{k} of the right helicity photons is coupled only to itself and the mode $-\mathbf{k}$ of the left helicity photons. For this reason, we can restrict our analysis to two modes with no loss in generality.

Written in terms of the matrix \mathcal{X} , the Bogoliubov transformation (13) reads

$$\mathcal{X} = \begin{bmatrix} e^{-i\phi} f_{R+} & 0 & 0 & e^{-i\phi} f_{R-} \\ 0 & e^{i\phi} f_{L-}^* & e^{i\phi} f_{L+}^* & 0 \\ 0 & e^{i\phi} f_{R-}^* & e^{i\phi} f_{R+}^* & 0 \\ e^{-i\phi} f_{L+} & 0 & 0 & e^{-i\phi} f_{L-} \end{bmatrix}. \quad (54)$$

The interpretation of classicality depends on what we consider to be the system.

In the most natural view, the system spans photons with both left and right helicity. Hence, we have a closed, two-mode system. Comparing (54) with (36), we easily find the classicality criterion (37) to read explicitly

$$f_{R-}(\mathbf{k}, T) = 0 = f_{L+}(\mathbf{k}, T). \quad (55)$$

Looking at (17), we can immediately see that this implies no Casimir effect, i.e., the photon production in the vacuum is zero. Thus, according to the RSF formalism, any dynamical Casimir effect is necessarily non-classical, as expected.

To see the physical reason for this, we go back to the differential equations (8), along with the initial conditions (14). It is easy to see that (8) can be fulfilled if and only if $\eta_-(\mathbf{k}, t) = 0$. This is equivalent to $\sigma(\mathbf{k}) = [\alpha/\Delta(\mathbf{k})]^{1/4}$, where, due to the time-independence of σ , α and Δ have to be time-independent too, implying constant velocity. The equations for the remaining functions can then be easily solved, yielding [11]

$$f_{R+}(\mathbf{k}, t) = f_{L-}^*(\mathbf{k}, t) = e^{-i\tilde{\omega}(\mathbf{k})t}, \quad (56)$$

where $\tilde{\omega} = \omega\sqrt{\alpha\Delta}$. Substituting this into (13), we find that the final creation and annihilation operators simplify to just

$$\begin{aligned} \hat{a}_{R,\text{out}}(\mathbf{k}) &= e^{-i[\phi(\mathbf{k}, T) + \tilde{\omega}(\mathbf{k})T]} \hat{a}_{R,\text{in}}(\mathbf{k}), \\ \hat{a}_{L,\text{out}}^\dagger(-\mathbf{k}) &= e^{-i[\phi(\mathbf{k}, T) - \tilde{\omega}(\mathbf{k})T]} \hat{a}_{L,\text{in}}^\dagger(-\mathbf{k}), \end{aligned} \quad (57)$$

i.e., they are multiplied by a phase. Obviously, this phase is irrelevant to the expectation values of the corresponding number operators on the vacuum, which is why the dynamical Casimir effect cannot take place for constant velocities.

However, there is another point of view. Nothing stops us from interpreting exclusively the left helicity photons as the system, and the right helicity photons as the environment. Then, we are dealing with an open one-mode system subject to influence from a one-mode environment. By comparing (54) with (36), (40), we immediately find that now, the classicality condition (41) always holds, regardless of the form of the functions $f_{R\pm}$, $f_{L\pm}$. Crucially, because the mode associated with the right helicity photons is initially in the vacuum state, then, due to Proposition 3, this classicality condition is both necessary and sufficient. Does this mean that the Casimir effect is, in the end, classical? Or maybe it means that the RSF formalism is not a valid tool for probing classicality after all?

In our opinion, neither. Consider, for example, the maximally entangled two-qubit Bell state [47, 48]

$$|\Phi_+\rangle := \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \quad (58)$$

If, in an analogy to the Casimir effect, we consider only the first qubit as the system, we will find it to be in the maximally mixed state

$$\hat{\rho}_S = \text{Tr}_{\text{2nd qubit}} |\Phi_+\rangle \langle \Phi_+| = \frac{1}{2} \hat{\mathbb{1}}_2, \quad (59)$$

which can certainly be considered classical. Of course, this does not mean that the Bell state that we started with was classical. Instead, its ‘‘quantumness’’ was contained in the correlations between the two qubits, rather than any of the two qubits themselves.

In the case of the Casimir effect and the Bogoliubov transformations in general, it is even more apparent what happens with the quantumness. Consider the matrix element $\mathcal{X}_{\downarrow 12} = \mathcal{X}_{\downarrow C} = f_{R-}(\mathbf{k}, T)$, which in our case, encodes the correlations between photons with left and right helicities. For a generic initial state, these correlations are potentially quantum. Thus, if a closed system is to be considered classical, they must necessarily vanish $\mathcal{X}_{\downarrow 12} = \mathcal{X}_{\downarrow C} = 0$, as they constitute an integral part of the system. However, in the case of an open system, the discussed correlations are no longer part of the system, and instead enter it only at the level of the environmental effects, most easily seen through the evolution (44). Therefore, even if they have a strictly quantum origin, the system experiences them only as dissipation, which in this case happens to have a semi-classical interpretation in terms of particle annihilation and creation rates.

Alternatively, we can think of the Casimir process as consisting of two parts. The first, captured by the matrix \mathcal{X}_\uparrow , describes the morphing of photons with left helicity into those with right helicity and vice versa. The second, captured by the matrix \mathcal{X}_\downarrow , describes the creation of photons with both helicities. The former, being semi-classical, is unconstrained by the RSF formalism. The latter, however, being more quantum in nature, is forbidden by RSF, unless the quantumness can be encoded into the environment, as discussed previously.

Finally, let us observe that even though the Bogoliubov transformation (13) is technically of the discrete type, as the creation and annihilation operators are formally ill-defined during the acceleration period $t \in [0, T]$, the functions $f_{R\pm}$ and $f_{L\pm}$ defining the transformation are well defined at all times. Adding that to the fact that the final moment of acceleration T is completely arbitrary, we can consider (13) as defining a smooth Bogoliubov evolution in the parameter T .

Since, as explained previously, the initial total state fulfills the requirements of Proposition 3, the Bogoliubov evolution at hand must have a representation in terms of the reduced kinetic equations (27) with (43) at the input. Indeed, making use of the latter equation, we find

$$\begin{aligned} h &= \hbar\omega \left(\eta_+ + \eta_- \operatorname{Re} \left[\frac{f_{R-}}{f_{R+}} \right] + \delta\beta \cos(\theta) \right), \\ \gamma_\uparrow &= 2\omega\eta_- \frac{|f_{R-}|^2}{|f_{R+}|^2} \operatorname{Im} \left[\frac{f_{R+}}{f_{R-}} \right], \\ \gamma_\downarrow &= 0. \end{aligned} \quad (60)$$

For more details regarding the derivation of these identities, see Appendix B. Here, we focus on their physical significance.

To start with, we note that, as expected, the Hamiltonian for the photons is proportional to their frequency. Furthermore, the particle annihilation rate is zero, which intuitively corresponds to the fact that the dynamical Casimir effect results only in the spontaneous creation of particles, not their disappearance. Finally, once again abusing the differential equations (8), we can easily calculate that the time derivative of the total photon density (18) equals

$$\frac{d\langle \hat{n} \rangle}{dT} = 2\omega\eta_- |f_{R-}|^2 \operatorname{Im} \left[\frac{f_{R+}}{f_{R-}} \right], \quad (61)$$

which, using (16) and (60), can be rewritten as simply

$$\frac{d\langle \hat{n} \rangle}{dT} = \gamma_\uparrow (\langle \hat{n} \rangle + 1). \quad (62)$$

This result has three worthwhile implications.

Firstly, it has a sound physical interpretation: the time derivative of the total photon density in the dynamical Casimir effect turns out to be simply proportional to the current photon density times the current particle creation rate. Secondly, it tells us that the non-negativity of γ_\uparrow , which is required for the result to be physical, is equivalent to the non-negativity of photon number growth. In particular, because of the initial condition (14), a valid matrix γ_\uparrow by its very construction prevents negative photon numbers. Finally, because of the $\langle \hat{n} \rangle$ -independent term on the r.h.s., our final result (62) proves that the dynamical Casimir effect occurs for any non-zero γ_\uparrow , which can be traced to any non-constant velocity of the medium ($\gamma_\uparrow = 0$ holds only for $\eta_- = 0$, which holds only for $\beta = \text{const}$).

6. Conclusions

In this paper, we employed the recent mesoscopic formalism of the reduced state of the field to derive the exact conditions under which Bogoliubov transformations in either isolated or open systems should be considered semi-classical. Applying our result to the case of the dynamical Casimir effect in the medium moving with a varying speed, we found that, while the photons with left and right helicity see each other as semi-classical objects, the Casimir effect itself is genuinely quantum, as expected. Let us stress that the analysis is made possible because for each wave vector, we can consider two polarization degrees of freedom. Therefore, it is essential that the described phenomenon is “based on full Maxwell equations in three dimensions” as pointed out at the end of the Conclusions section in [11].

Acknowledgments

We dedicate this work to Iwo Białyński-Birula on the occasion of his 90th birthday.

We acknowledge support from the Foundation for Polish Science (International Research Agenda Programme project, International Centre for Theory of Quantum Technologies, Grant No. 2018/MAB/5, co-financed by the European Union within the Smart Growth Operational program).

Appendix A

In this appendix, we prove our main results regarding the classicality of Bogoliubov transformations, i.e., Propositions 1–3.

To this end, in addition to RSF, we will employ two auxiliary mesoscopic fields. The first, defined originally in [21], is the *conjugate RSF*

$$\begin{aligned} c &:= \sum_{k,k'=1}^N \operatorname{Tr} [\hat{\rho} \hat{a}_{k'} \hat{a}_k] |k\rangle \langle k'|, \\ |\alpha^*\rangle &:= \sum_{k=1}^N \operatorname{Tr} [\hat{\rho} \hat{a}_k^\dagger] |k\rangle. \end{aligned} \quad (63)$$

The second is the *generalized RSF*

$$\begin{aligned} g &:= \sum_{k,k'=1}^{2N} \operatorname{Tr} [\hat{\rho} \hat{A}_{k'}^\dagger \hat{A}_k] |k\rangle \langle k'|, \\ |\mathcal{A}\rangle &:= \sum_{k=1}^{2N} \operatorname{Tr} [\hat{\rho} \hat{A}_k] |k\rangle. \end{aligned} \quad (64)$$

It is easy to see that the three reduced fields are related to each other as follows

$$\begin{aligned} g &= \begin{bmatrix} r & c \\ c^* & r^T + \mathbb{1}_N \end{bmatrix}, \\ |\mathcal{A}\rangle &= |\alpha\rangle \oplus |\alpha^*\rangle. \end{aligned} \quad (65)$$

We add that, by definition, $r = r^\dagger$, $c = c^T$, and $|\alpha\rangle^* = |\alpha^*\rangle$.

Proof of Proposition 1

We start with Proposition 1. It is easy to see that due to (64), (32), (33), under a generic Bogoliubov transformation, the generalized RSF $(g, |\mathcal{A}\rangle)$ transforms as

$$g' = \mathcal{X}g\mathcal{X}^\dagger, \quad |\mathcal{A}'\rangle = \mathcal{X}|\mathcal{A}\rangle. \quad (66)$$

Then, (65) and (36) imply

$$r' := \mathcal{X}_\uparrow r \mathcal{X}_\uparrow^\dagger + \mathcal{X}_\uparrow c \mathcal{X}_\uparrow^\dagger + \mathcal{X}_\downarrow c^\dagger \mathcal{X}_\downarrow^\dagger + \mathcal{X}_\downarrow (r_S^T + \mathbb{1}_N) \mathcal{X}_\downarrow^\dagger, \\ |\alpha'\rangle = \mathcal{X}_\uparrow |\alpha\rangle + \mathcal{X}_\downarrow |\alpha^*\rangle. \quad (67)$$

Clearly, this couples RSF to the conjugate field, meaning that it does not preserve the set of the associated degrees of freedom. For an arbitrary initial state, the coupling vanishes only if (37) is fulfilled, which is what we wanted to show.

Assuming the time-dependent case with the classicality condition (37) fulfilled, (67) reduces to

$$r(t) = \mathcal{X}_\uparrow(t)r(t_0)\mathcal{X}_\uparrow^\dagger(t), \\ |\alpha(t)\rangle = \mathcal{X}_\uparrow(t)|\alpha(t_0)\rangle. \quad (68)$$

These equations are reversible, i.e.,

$$r(t_0) = \mathcal{X}_\uparrow^{-1}(t)r(t)\mathcal{X}_\uparrow^{-\dagger}(t), \\ |\alpha(t_0)\rangle = \mathcal{X}_\uparrow^{-1}(t)|\alpha(t)\rangle. \quad (69)$$

Taking the time derivative of (68) and making use of (69), we obtain the reduced kinetic equations (27) with (38) at the input. This concludes the proof.

Proof of Proposition 2

To prove Proposition 2, we observe that the reduced fields of the total state of the system and the environment have the structure

$$r = \begin{bmatrix} r_S & r_C \\ r_C^\dagger & r_E \end{bmatrix}, \quad c = \begin{bmatrix} c_S & c_C \\ c_C^T & c_E \end{bmatrix},$$

$$|\alpha\rangle = |\alpha^*\rangle^* = |\alpha_S\rangle \oplus |\alpha_E\rangle, \quad (70)$$

where $(r_S, |\alpha_S\rangle)$, $(c_S, |\alpha_S^*\rangle)$ are the reduced fields of the system; $(r_E, |\alpha_E\rangle)$, $(c_E, |\alpha_E^*\rangle)$ are the reduced fields of the environment; and r_C, c_C contain the system-bath correlations. This fact follows directly from the definitions of the fields. For example,

$$(r_S)_{kk'} := \text{Tr} [\text{Tr}_E(\hat{\rho})\hat{a}_k^\dagger\hat{a}_k] = \text{Tr} [\hat{\rho}\hat{a}_k^\dagger\hat{a}_k] := r_{kk'}. \quad (71)$$

The remaining relations are proved in a similar fashion.

For a generic initial total state, the dynamics are quite complex. Making use of the block-form decompositions (70) and (40) in (68), we obtain a rather lengthy expression for the transformed RSF of the system, which can be written as

$$r'_S = F_{\uparrow\uparrow}(r) + F_{\downarrow\uparrow}(c^*) + F_{\uparrow\downarrow}(c) + F_{\downarrow\downarrow}(r^T + 1), \\ |\alpha'_S\rangle = \mathcal{X}_{\uparrow S} |\alpha_S\rangle + \mathcal{X}_{\uparrow C} |\alpha_E\rangle + \mathcal{X}_{\downarrow S} |\alpha_S^*\rangle + \mathcal{X}_{\downarrow C} |\alpha_E^*\rangle, \quad (72)$$

where

$$F_{ab}(x) := \mathcal{X}_{aS}x_S\mathcal{X}_{bS}^\dagger + \mathcal{X}_{aS}x_C\mathcal{X}_{bC}^\dagger + \mathcal{X}_{aC}x_C^\dagger\mathcal{X}_{bS}^\dagger \\ + \mathcal{X}_{aC}x_E\mathcal{X}_{bC}^\dagger. \quad (73)$$

Similarly to the case with the closed system transformation, (72) may preserve the set of the degrees of freedom associated with the RSF formalism in the system only if it does not depend on the conjugate field of the system, $(c_S, |\alpha_S^*\rangle)$. Close inspection of (72) reveals that this is possible only if (41) is fulfilled, which is what we wanted to prove.

Let us stress, however, that this condition is merely necessary for the RSF degrees of freedom to be preserved. Depending on the state of the bath, the remaining fields r_C, c_C, r_E , and c_E will, in general, cause the system to go beyond the RSF framework. In the most radical case, the equations may preserve the formalism's set of degrees of freedom only if all terms dependent on these additional fields vanish, reducing the system-environment ensemble to two separate closed systems.

Proof of Proposition 3

Finally, to prove Proposition 3, we note that, as is easy to calculate from their definitions, the initial reduced fields with the environment initially in the vacuum state fulfill

$$r_C = r_E = c_C = c_E = 0, \quad |\alpha_E\rangle = |\alpha_E^*\rangle = 0. \quad (74)$$

Plugging this into (72), we find that it simplifies to

$$r'_S = \mathcal{X}_{\uparrow S} r_S \mathcal{X}_{\uparrow S}^\dagger + \mathcal{X}_{\downarrow C} \mathcal{X}_{\downarrow C}^\dagger, \\ |\alpha'_S\rangle = \mathcal{X}_{\uparrow S} |\alpha_S\rangle, \quad (75)$$

where we assumed the classicality condition (41). Clearly, the final field depends only on the initial RSF, preserving the associated degrees of freedom. Therefore, in this case, the condition (43) is not only necessary, but also sufficient for classicality.

It remains to show that if the transformation depends smoothly on time, the corresponding reduced kinetic equations are given by (43). In the time-dependent case, (75) becomes

$$r_S(t) = \mathcal{X}_{\uparrow S}(t)r_S(0)\mathcal{X}_{\uparrow S}^\dagger(t) + \mathcal{X}_{\downarrow C}(t)\mathcal{X}_{\downarrow C}^\dagger(t), \\ |\alpha_S(t)\rangle = \mathcal{X}_{\uparrow S}(t)|\alpha_S(0)\rangle. \quad (76)$$

These relations are reversible

$$r_S(0) = \mathcal{X}_{\uparrow S}^{-1}(t) \left[r_S(t) - \mathcal{X}_{\downarrow C}(t)\mathcal{X}_{\downarrow C}^\dagger(t) \right] \mathcal{X}_{\uparrow S}^{-\dagger}(t), \\ |\alpha_S(0)\rangle = \mathcal{X}_{\uparrow S}^{-1}(t)|\alpha_S(t)\rangle. \quad (77)$$

Differentiating (76) with respect to time, making use of (77), and rearranging the terms, we arrive at the differential evolution equations

$$\begin{aligned} \frac{dr}{dt} &= \frac{1}{2}[\mathcal{Y}_i, r] + \frac{1}{2}\{\mathcal{Y}_r, r\} + \mathcal{W}, \\ \frac{d|\alpha\rangle}{dt} &= \frac{1}{2}\mathcal{Y}_i|\alpha\rangle + \frac{1}{2}\mathcal{Y}_r|\alpha\rangle, \end{aligned} \quad (78)$$

where the matrices \mathcal{Y}_r , \mathcal{Y}_i , \mathcal{W} are as defined in (44). Clearly, the derived equations have the form of the reduced kinetic equations characterized by (43). Thus, they describe valid dynamics provided the γ_{\uparrow} matrices are non-negative, as required by (42). This concludes the proof.

Appendix B: Proof of (60)

In this appendix, we derive the explicit forms of the operators (60) governing the reduced kinetic equations for the dynamical Casimir effect.

By comparing (54) with (36) and (40), we immediately identify

$$\mathcal{X}_{\uparrow S} = e^{-i\phi}f_{R+}, \quad \mathcal{X}_{\downarrow C} = e^{-i\phi}f_{R-}. \quad (79)$$

Plugging this into (44) and then (43), on the way utilizing the differential equations (8), we obtain, after a lengthy but straightforward calculation,

$$\begin{aligned} h &= \hbar\omega \left(\eta_+ + \eta_- \operatorname{Re} \left[\frac{f_{R-}}{f_{R+}} \right] \right) + \hbar \frac{d\phi}{dt}, \\ \gamma_{\uparrow} &= 2\omega\eta_- |f_{R-}|^2 \left(\operatorname{Im} \left[\frac{f_{R+}}{f_{R-}} \right] + \operatorname{Im} \left[\frac{f_{R-}}{f_{R+}} \right] \right), \\ \gamma_{\downarrow} &= 2\omega\eta_- \left[|f_{R-}|^2 \operatorname{Im} \left[\frac{f_{R+}}{f_{R-}} \right] + (|f_{R-}|^2 + 1) \operatorname{Im} \left[\frac{f_{R-}}{f_{R+}} \right] \right]. \end{aligned} \quad (80)$$

It remains to show that these formulas reduce to (60).

In the case of the Hamiltonian, all we need to do is to differentiate (10) with respect to time. Due to the Leibniz integral rule,

$$\frac{d\phi(\mathbf{k}, t)}{dt} = \omega(\mathbf{k})\delta(t)\beta(t) \cos\theta(\mathbf{k}), \quad (81)$$

from which we immediately see that the first lines of (80) and (60) coincide.

As for γ_{\uparrow} , we observe that for any complex number w

$$\operatorname{Im}[w^{-1}] = -\frac{\operatorname{Im}[w]}{|w|^2}. \quad (82)$$

Taking $w = f_{R+}/f_{R-}$, we get

$$\gamma_{\uparrow} = 2\omega\eta_- |f_{R-}|^2 \left(1 - \frac{|f_{R-}|^2}{|f_{R+}|^2} \right) \operatorname{Im} \left[\frac{f_{R+}}{f_{R-}} \right]. \quad (83)$$

Using (16) and simplifying, we quickly find that the second lines of (80) and (60) also coincide.

Finally, we have to show that $\gamma_{\downarrow} = 0$. Once again utilizing the relation (16), we obtain

$$\gamma_{\downarrow} = 2\omega\eta_- |f_{R-}|^2 \left(\operatorname{Im} \left[\frac{f_{R+}}{f_{R-}} \right] + \frac{|f_{R+}|^2}{|f_{R-}|^2} \operatorname{Im} \left[\frac{f_{R-}}{f_{R+}} \right] \right). \quad (84)$$

It is easy to see that the bracketed term vanishes upon the use of (82).

References

- [1] S.K. Lamoreaux, *Annu. Rev. Nucl. Part. Sci.* **62**, 37 (2012).
- [2] G.L. Klimchitskaya, U. Mohideen, V.M. Mostepanenko, *Rev. Mod. Phys.* **81**, 1827 (2009).
- [3] K.A. Milton, *J. Phys. A: Math. Gen.* **37**, R209 (2004).
- [4] H.B.G. Casimir, D. Polder, *Phys. Rev.* **73**, 360 (1948).
- [5] E. Yablonovitch, *Phys. Rev. Lett.* **62**, 1742 (1989).
- [6] J. Schwinger, *Proc. Natl. Acad. Sci.* **89**, 4091 (1992).
- [7] V.V. Dodonov, *Phys. Scr.* **82**, 038105 (2010).
- [8] V. Dodonov, *Physics* **2**, 67 (2020).
- [9] I. Białyński-Birula, Z. Białyńska-Birula, *Quantum Electrodynamics*, Elsevier, 1975.
- [10] I. Białyński-Birula, Z. Białyńska-Birula, *Bull. Acad. Pol. Sci. Cl. III* **5**, 1119 (1957).
- [11] I. Białyński-Birula, Z. Białyńska-Birula, *Phys. Rev. A* **78**, 042109 (2008).
- [12] Ł. Rudnicki, I. Białyński-Birula, *Opt. Comm.* **283**, 644 (2010).
- [13] I. Białyński-Birula, Z. Białyńska-Birula, *Phys. Rev. A* **77**, 052103 (2008).
- [14] L. Parker, D. Toms, *Quantum Field Theory in Curved Spacetime: Quantized Fields and Gravity*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, 2009.
- [15] N.N. Bogoljubov, *Il Nuovo Cimento (1955-1965)* **7**, 794 (1958).
- [16] J.G. Valatin, *Il Nuovo Cimento (1955-1965)* **7**, 794 (1958).
- [17] G.S. Agarwal, *Quantum Optics*, Cambridge University Press, 2012.
- [18] W.G. Unruh, *Phys. Rev. D* **14**, 870 (1976).
- [19] S.W. Hawking, *Nature* **248**, 30 (1974).
- [20] R. Alicki, *Entropy* **21**, 705 (2019).
- [21] T. Linowski, Ł. Rudnicki, *Phys. Rev. A* **106**, 062204 (2022).
- [22] P. Lambropoulos, D. Petrosyan, *Fundamentals of Quantum Optics and Quantum Information*, Springer, 2007.

- [23] L. Mandel, E. Wolf, *Optical Coherence and Quantum Optics*, Cambridge University Press, 1995.
- [24] H. Minkowski, *Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl.* **2**, 53 (1908).
- [25] I. Białynicki-Birula, Z. Białynicka-Birula, *J. Phys. A: Math. Theor.* **46**, 053001 (2013).
- [26] S. Cusumano, Ł. Rudnicki, *Entropy* **23**, 1198 (2021).
- [27] G. Adesso, A. Serafini, F. Illuminati, *Phys. Rev. Lett.* **92**, 087901 (2004).
- [28] S. Mancini, H.M. Wiseman, *Phys. Rev. A* **75**, 012330 (2007).
- [29] K. Koga, N. Yamamoto, *Phys. Rev. A* **85**, 022103 (2012).
- [30] A. Wehrl, *Rev. Mod. Phys.* **50**, 221 (1978).
- [31] A. Wehrl, *Rep. Math. Phys.* **16**, 353 (1979).
- [32] V. Gorini, A. Kossakowski, E.C.G. Sudarshan, *J. Math. Phys.* **17**, 821 (1976).
- [33] G. Lindblad, *Comm. Math. Phys.* **48**, 119 (1976).
- [34] J. Dereziński, *J. Math. Phys.* **58**, 121101 (2017).
- [35] P.T. Nam, M. Napiórkowski, J.P. Solovej, *J. Funct. Anal.* **270**, 4340 (2016).
- [36] A. Rivas, S.F. Huelga, *Open Quantum Systems. An introduction*, Springer, Berlin 2012.
- [37] J.S. Ivan, K.K. Sabapathy, R. Simon, *Phys. Rev. A* **84**, 042311 (2011).
- [38] G. De Palma, *Lett. Math. Phys.* **108**, 97 (2017).
- [39] G. De Palma, *J. Math. Phys.* **60**, 112201 (2019).
- [40] D.M. Davidovic, D. Lalovic, *J. Phys. A* **29**, 3787 (1996).
- [41] R.J. Glauber, *Phys. Rev.* **131**, 2766 (1963).
- [42] E.C.G. Sudarshan, *Phys. Rev. Lett.* **10**, 277 (1963).
- [43] A. Altland, B.D. Simons, *Condensed Matter Field Theory*, 2nd ed, Cambridge University Press, 2010.
- [44] W. Schleich, A. Bandilla, H. Paul, *Phys. Rev. A* **45**, 6652 (1992).
- [45] K. Husimi, *Proc. Phys. Math. Soc. Jpn* **22**, 264 (1940).
- [46] T. Linowski, K. Schlichtholz, Ł. Rudnicki, *Phys. Rev. A* **107**, 033707 (2023).
- [47] J.S. Bell, *Physics* **1**, 195 (1964).
- [48] M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information*, 10th Anniversary Edition, Cambridge University Press, 2011.

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— ◇ —

Smearred Field Description of Free Electromagnetic Field

J.A. PRZESZOWSKI*

Faculty of Physics, University of Białystok, Ciołkowskiego 4L, 15-424 Białystok, Poland

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S107](https://doi.org/10.12693/APhysPolA.143.S107)*e-mail: j.przeszowski@uwb.edu.pl

Free electromagnetic fields, satisfying Maxwell's equations with no charges and electric currents, can be described by complex vector fields. In the standard formulation with fields sharply dependent on position and time, one obtains integrals that are mathematically ill-defined. This happens for the massless Pauli–Jordan function, which is used to describe the time evolution of fields and appears in the Poisson brackets for classical fields. This difficulty can be solved by introducing smeared fields as linear functionals with test functions. In this way, the massless Pauli–Jordan function becomes a tempered distribution, allowing a mathematically rigorous analysis.

topics: Riemann–Silberstein vectors, Poisson brackets, Pauli–Jordan function, tempered distributions

1. Introduction

The electric and magnetic fields in empty space can be expressed in terms of a pair of complex vector fields $\mathbf{F}(x) = \mathbf{E}(x) - i c \mathbf{B}(x)$ and $\mathbf{F}^*(x) = \mathbf{E}(x) + i c \mathbf{B}(x)$, where $x = (t, \mathbf{x})$ and $c = (\mu_0 \epsilon_0)^{-1/2}$ [1]. This can be extended to an arbitrary homogeneous and isotropic dielectric [2]. These fields are called the Riemann–Silberstein (RS) vectors [3] and can be analysed in various aspects in both classical and quantum physics. In [3] Białynicki-Birula claims that it is a complex vector-function of space and time coordinates that adequately describes the quantum state of a single photon. It is also argued that it can be practical for describing the quantum states of excitation of a free electromagnetic field, the electromagnetic field acting on a medium, the vacuum excitation of virtual electron–positron pairs, and for comparing the photon with other quantum particles that have their wave functions. Also, the Schrödinger equation for a photon and the Heisenberg uncertainty relations can be formulated in terms of the RS vectors [4]. More mathematical aspects of this formalism are presented in [5] and [6]. An overview of many features of classical and quantum electromagnetic fields described by RS vectors can be found in [7].

In this article, we will discuss other aspects of classical fields in empty space, with particular reference to objects defined by momentum integrals that do not converge, as this can lead to self-inconsistency.

2. Poisson brackets and temporal evolution

In a free space with no charges and currents, the RS vectors allow us to express Maxwell's equations in a compact form

$$\partial_t \mathbf{F}(x) = i c \nabla \times \mathbf{F}(x), \quad \nabla \cdot \mathbf{F}(x) = 0,$$

$$\partial_t \mathbf{F}^*(x) = -i c \nabla \times \mathbf{F}^*(x), \quad \nabla \cdot \mathbf{F}^*(x) = 0. \quad (1)$$

The electromagnetic energy density can be written as a simple expression if we scale the RS vectors by a constant factor

$$\mathcal{H}(x) = \frac{\epsilon_0}{2} \mathbf{E}^2(x) + \frac{1}{2\mu_0} \mathbf{B}^2(x) = \mathbf{V}^*(x) \cdot \mathbf{V}(x),$$

$$\mathbf{V}(x) = \sqrt{\frac{\epsilon_0}{2}} \mathbf{F}^*(x),$$

$$\mathbf{V}^*(x) = \sqrt{\frac{\epsilon_0}{2}} \mathbf{F}(x). \quad (2)$$

This leads to the Hamiltonian that generates the equations of motion for $\mathbf{V}(x)$ and $\mathbf{V}^*(x)$ by means of the Poisson brackets (PB) for fields at equal time $t = t_0$,

$$\begin{aligned} \{V_i(t_0, \mathbf{x}), V_j^*(t_0, \mathbf{y})\}_{\text{PB}} &= i c \epsilon_{ikj} \partial_k^x \delta^3(\mathbf{x} - \mathbf{y}), \\ \{V_i^*(t_0, \mathbf{x}), V_j(t_0, \mathbf{y})\}_{\text{PB}} &= -i c \epsilon_{ikj} \partial_k^x \delta^3(\mathbf{x} - \mathbf{y}), \\ \{V_i(t, \mathbf{x}), V_j(t_0, \mathbf{y})\}_{\text{PB}} &= \{V_i^*(t, \mathbf{x}), V_j^*(t_0, \mathbf{y})\}_{\text{PB}} = 0, \end{aligned} \quad (3)$$

where Einstein's summation convention for repeated indices is applied. Moreover, the equations of motion can be solved by means of the integral [8]

$$\begin{aligned} \mathbf{V}(t, \mathbf{x}) &= \int_{\mathbb{R}^3} d^3 \mathbf{y} \left[\frac{\partial}{\partial t} D(t-t_0, \mathbf{x}-\mathbf{y}) \mathbf{V}(t_0, \mathbf{y}) \right. \\ &\quad \left. - i c D(t-t_0, \mathbf{x}-\mathbf{y}) \nabla \times \mathbf{V}(t_0, \mathbf{y}) \right], \end{aligned} \quad (4)$$

where the function $D(t, \mathbf{x})$ should satisfy the following properties

$$\begin{aligned} (\partial_t^2 - c^2 \nabla^2) D(x) &= 0, \quad D(0, \mathbf{x}) = 0, \\ \partial_t D(t, \mathbf{x})|_{t=0} &= \delta^3(\mathbf{x}). \end{aligned} \quad (5)$$

The general solution (4) can be used to transform the Poisson brackets for equal time (3) into the generalized Poisson brackets for arbitrary instants [8]

$$\begin{aligned} \{V_i(t, \mathbf{x}), V_j^*(t_0, \mathbf{y})\}_{\text{PB}} &= \\ &= (c^2 \delta_{ik} \nabla^2 - c^2 \partial_i \partial_k + i c \epsilon_{ikj} \partial_k^x \partial_t) D(t-t_0, \mathbf{x}-\mathbf{y}), \\ \{V_i(t, \mathbf{x}), V_j(t_0, \mathbf{y})\}_{\text{PB}} &= 0 \end{aligned} \quad (6)$$

The standard definition of $D(t, \mathbf{x})$, called the massless Pauli–Jordan function [9], is given by the momentum integral

$$D(t, \mathbf{x}) = i \int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{2\omega (2\pi)^3} (e^{-i\mathbf{k}\cdot\mathbf{x}} - e^{i\mathbf{k}\cdot\mathbf{x}}), \quad (7)$$

($\mathbf{k} \cdot \mathbf{x} = \omega t - \mathbf{k} \cdot \mathbf{x}$; $\omega = c|\mathbf{k}| = ck$), which is clearly ill-defined. However, it is generally argued that this divergent integral defines a distribution that can be written explicitly using the Dirac delta distribution accordingly

$$D(t, \mathbf{x}) = \frac{\text{sgn}(t)}{2\pi c} \delta(c^2 t^2 - \mathbf{x}^2). \quad (8)$$

Formally, the standard definition (7) is the inverse Fourier transform of a function that is not integrable but is only locally integrable in \mathbb{R}^3 . For this type of function, the corresponding Fourier transform can be consistently defined and calculated using tempered distributions, as recently proposed in [10]. One might expect that the appearance of such divergent momentum integrals as in formula (7) is a consequence of the use of fields with a sharp dependence on the position vector. Therefore, we can try a different approach, using smeared fields to check whether the new results are consistent with the previous ones.

3. Smeared Riemann–Silberstein vectors

Smeared fields are linear functionals for test functions, which vanish rapidly at infinity, so one may define smeared fields as integrals with test functions of the Schwartz class $\mathcal{S}(\mathbb{R}^3) \in f : \mathbb{R}^3 \rightarrow \mathbb{R}$ [11]

$$\begin{aligned} V_i[t, f] &:= \int_{\mathbb{R}^3} d^3 \mathbf{x} V_i(t, \mathbf{x}) f(\mathbf{x}), \\ V_i^*[t, f] &:= (V_i[t, f])^* = \int_{\mathbb{R}^3} d^3 \mathbf{x} V_i^*(t, \mathbf{x}) f(\mathbf{x}). \end{aligned} \quad (9)$$

Thus, all these integrals converge, and integration by parts can be easily performed without the boundary term at infinity. This smearing can easily be applied to relations for fields with a sharp dependence on the position vector. From Maxwell's equations (1) one obtains the relations for smeared RS vectors

$$\begin{aligned} \partial_t V_i[t, f] &= -i V_k[t, u_{ik} f], \quad V_i[t, \partial_i f] = 0, \\ \partial_t V_i^*[t, f] &= i V_k^*[t, u_{ik} f], \quad V_i^*[t, \partial_i f] = 0, \end{aligned} \quad (10)$$

where $u_{ik} := c \epsilon_{ijk} \partial_j$. Next, smearing of non-vanishing Poisson brackets for equal time (3) gives Poisson brackets for smeared RS vectors at equal time $t = t_0$

$$\begin{aligned} \{V_i[t_0, f], V_j^*[t_0, g]\}_{\text{PB}} &= i(u_{ij} g, f), \\ \{V_i^*[t_0, g], V_j[t_0, f]\}_{\text{PB}} &= -i(g, u_{ij} f), \end{aligned} \quad (11)$$

where the inner product in the space of the Schwartz test functions $f, g \in \mathcal{S}(\mathbb{R}^3)$ is [11]

$$(g, f) = \int_{\mathbb{R}^3} d^3 x g(x) f(x). \quad (12)$$

The equations of motion for smeared fields can be easily diagonalized as

$$\begin{aligned} \partial_t a_i^{(+)}[t, f] &= i a_i^{(+)}[t, \omega f], \\ \partial_t a_i^{(-)}[t, f] &= -i a_i^{(-)}[t, \omega f], \end{aligned} \quad (13)$$

with auxiliary smeared fields defined as

$$\begin{aligned} a_i^{(+)}[t, f] &:= V_i[t, \omega f] + V_j[t, u_{ij} f], \\ a_i^{(-)}[t, f] &:= V_i[t, \omega f] - V_j[t, u_{ij} f], \end{aligned} \quad (14)$$

where a new test function $\omega^a f : \mathbb{R}^3 \rightarrow \mathbb{R}$, is given by means of the inverse Fourier transform ($a \in \mathbb{N}$)

$$(\omega^a f)(x) := \int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} (c|\mathbf{k}|)^a \mathcal{F}\{f\}(k). \quad (15)$$

The equations of motion for auxiliary smeared fields can be easily solved as

$$\begin{aligned} a_i^{(+)}[t, f] &:= a_i^{(+)}[t_0, e^{i\omega(t-t_0)} f], \\ a_i^{(-)}[t, f] &:= a_i^{(-)}[t_0, e^{-i\omega(t-t_0)} f], \end{aligned} \quad (16)$$

where one has

$$(e^{\pm i\omega t} f)(\mathbf{x}) = \sum_{n=0}^{\infty} \frac{(\pm i t)^n}{n!} (\omega^n f)(\mathbf{x}) = \int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{(2\pi)^3} e^{\pm i\omega(\mathbf{k})t} \mathcal{F}\{f\}(\mathbf{k}). \quad (17)$$

These solutions enable us to write the smeared RS vector at an arbitrary instant t as

$$V_i[t, f] := \frac{1}{2} \left(a_i^{(+)}[t, \omega^{-1} f] + a_i^{(-)}[t, \omega^{-1} f] \right) = V_i[t_0, \cos(\omega\tau) f] + i V_k[t_0, u_{ik} \sin(\omega\tau) \omega^{-1} f] = \int_{\mathbb{R}^3} d^3 \mathbf{y} (\cos(\omega\tau) f)(\mathbf{y}) V_i(t_0, \mathbf{y}) - i \int_{\mathbb{R}^3} d^3 \mathbf{y} (\sin(\omega\tau) \omega^{-1} f)(\mathbf{y}) U_i(t_0, \mathbf{y}), \quad (18)$$

where $\tau = t - t_0$, $U_i(t_0, \mathbf{x}) = u_{ik} V_k(t_0, \mathbf{x})$, and we extend the previous definition (15) to the case $a = -1$. When we integrate both sides of (4) with the test function $f(\mathbf{x}) \in \mathcal{S}(\mathbb{R}^3)$, assuming local integrability of $D(\tau, \mathbf{x})$, we obtain the tempered distribution $\mathcal{S}'(\mathbb{R}^3)$. Next, by switching the order of the integrals, we get a functional, which can be compared with (18)

$$V_i[t, f] = \int_{\mathbb{R}^3} d^3 \mathbf{y} \frac{\partial}{\partial t} \left[\int_{\mathbb{R}^3} d^3 \mathbf{x} f(\mathbf{x}) D(\tau, \mathbf{x} - \mathbf{y}) \right] V_i(t_0, \mathbf{y}) - i \int_{\mathbb{R}^3} d^3 \mathbf{y} \left[\int_{\mathbb{R}^3} d^3 \mathbf{x} f(\mathbf{x}) D(\tau, \mathbf{x} - \mathbf{y}) \right] U_i(t_0, \mathbf{y}). \quad (19)$$

Hence, we conclude that the massless Pauli–Jordan function $D(\tau, \mathbf{x} - \mathbf{y})$ satisfies the integral equation

$$\int_{\mathbb{R}^3} d^3 \mathbf{x} f(\mathbf{x}) D(\tau, \mathbf{x} - \mathbf{y}) = (\sin(\omega\tau) \omega^{-1} f)(\mathbf{y}). \quad (20)$$

If we choose $t_0 = 0$ and $\mathbf{y} = 0$, then the integral equation (20) becomes the definition of the tempered distribution $D[t, f]$ as a linear functional

$$D[t, f] := \int_{\mathbb{R}^3} d^3 \mathbf{x} f(\mathbf{x}) D(t, \mathbf{x}) = (\sin(\omega t) \omega^{-1} f)(\mathbf{0}), \quad (21)$$

which can be taken as a starting point for further analysis.

4. Analysis of the Pauli–Jordan functional $D[t, f]$

Our analysis of $D[t, f]$ will use the calculation method proposed in [10], thus we start with

$$D[t, f] = (\sin(\omega t) \omega^{-1} f)(\mathbf{0}) = \int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{\sin(ckt)}{ck} \mathcal{F}\{f\}(\mathbf{k}) = \int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{\sin(ckt)}{ck} \left[\int_{\mathbb{R}^3} d^3 \mathbf{x} f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} \right]. \quad (22)$$

The next step requires switching the order of the integrals and if we do this directly in its present form, we get a divergent momentum integral, so such a final step would be mathematically incorrect,

$$D[t, f] = \int_{\mathbb{R}^3} d^3 \mathbf{x} f(\mathbf{x}) \left[\int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{\sin(ckt)}{ck} e^{-i\mathbf{k}\cdot\mathbf{x}} \right]. \quad (23)$$

Note, however, that the divergent integral in square bracket in (23) is $D(t, \mathbf{x})$, which was defined earlier by (7). Since here it appears as the result of erroneous mathematical operations, hence we can conclude that the standard definition of (7) is flawed or at best symbolic. For the smeared vector RS, we can avoid this pitfall, but we must carefully follow the steps below.

Firstly, the Fourier transform of the test function in $\mathcal{S}(\mathbb{R}^3)$ allows integration by parts without a boundary term, so we can perform the following transformation of integrals

$$D[t, f] = \int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{\sin(ckt)}{ck} \left[\int_{\mathbb{R}^3} d^3 \mathbf{x} f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} \right] = \int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{\sin(ckt)}{ck} \frac{1}{k^2} \left[\int_{\mathbb{R}^3} d^3 \mathbf{x} (-\Delta) f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} \right] = \int_{\mathbb{R}^3} d^3 \mathbf{x} (-\Delta) f(\mathbf{x}) \int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{\sin(ckt)}{ck^3} e^{-i\mathbf{k}\cdot\mathbf{x}}. \quad (24)$$

This yields a convergent momentum integral, which can be calculated analytically using formula (3.741.3) in [12],

$$\int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{\sin(ckt)}{ck^3} e^{i\mathbf{k}\cdot\mathbf{x}} = \frac{\text{sgn}(t)}{2\pi^2 c r} \int_{\mathbb{R}^3} dk \frac{\sin(c k |t|)}{k^2} \sin(k r) = \frac{\text{sgn}(t)}{4\pi c} \left(1 + \frac{c|t| - r}{r} \Theta(r - c|t|) \right), \quad (25)$$

where $r = |\mathbf{x}|$ and $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. This leads to the final stage of the calculation, where we have to perform integration by parts for the convergent integral in \mathbb{R}^3 . Omitting details, which will be presented elsewhere, we give the final result

$$D[t, f] = \frac{t}{4\pi} \int_{\Omega_3} d\omega_{\mathbf{x}} f(c|t|\hat{\mathbf{x}}), \quad (26)$$

where $d\omega_{\mathbf{x}}$ is the hypersurface element on the unit sphere Ω_3 embedded in \mathbb{R}^3 , with its surface area $|\Omega_3| = 4\pi$ and $\hat{\mathbf{x}} = \mathbf{x}/r$ being the versor of the position vector. This formula is the main result of this work and can be used to study various properties of the massless Pauli–Jordan function $D(t, \mathbf{x})$ for arbitrary time t .

First, one finds $D[0, f] = 0 \implies D(0, \mathbf{x}) = 0$ and

$$\begin{aligned} \frac{d}{dt}D[t, f] \Big|_{t=0} &= \frac{1}{4\pi} \int_{\Omega_3} d\omega_{\mathbf{x}} f(\mathbf{0}) = f(\mathbf{0}) = \\ \int_{\mathbb{R}^3} d^3\mathbf{x} f(\mathbf{x}) \delta^3(\mathbf{x}) &\implies \frac{\partial}{\partial t}D(t, \mathbf{x}) \Big|_{t=0} = \delta^3(\mathbf{x}). \end{aligned} \quad (27)$$

The other equations for $D[t, f]$ require more complicated calculations, but the use of equations presented in the Appendix can be quite helpful. First, using (39) in the Appendix, we obtain

$$\begin{aligned} \frac{\partial}{\partial t}D[t, f] &= D[t, \partial_r(r f)] \implies \\ \implies \left(t \frac{\partial}{\partial t} + r \frac{\partial}{\partial t} + 2 \right) D(t, \mathbf{x}) &= 0, \end{aligned} \quad (28)$$

which is the manifestation of the covariance at an infinitesimal dilation transformation. Then, (38) and (40) in the Appendix allow us to find

$$\begin{aligned} \frac{1}{c} \frac{d}{dt}D[t, x^i f] &= ct D[t, \partial_i f] \implies \\ \implies \left(\frac{x^i}{c} \frac{\partial}{\partial t} + ct \frac{\partial}{\partial x^i} \right) D(t, \mathbf{x}) &= 0, \end{aligned} \quad (29)$$

which is the manifestation of the invariance at an infinitesimal Lorentz boost transformation. Finally (41) and (42) in the Appendix lead to the d’Alambert equation of motion

$$\begin{aligned} \frac{d^2}{dt^2}D[t, f] &= D[t, \nabla^2 f] \implies \\ \implies \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) D(t, \mathbf{x}) &= 0. \end{aligned} \quad (30)$$

All the above implications are valid in the sense of the distributions $\mathcal{S}'(\mathbb{R}^3)$.

Finally, we can give the explicit form of the distribution $D(t, \mathbf{x})$ starting from the functional (26), for which the Dirac delta distribution can be introduced according to the equations

$$\begin{aligned} f(|a|) &= \int_{\mathbb{R}_+} dr \delta(r - |a|) f(r) = \\ 2 \int_{\mathbb{R}_+} dr r \delta(r^2 - a^2) f(r), \end{aligned} \quad (31)$$

which are valid for $a \neq 0$. Thus for $t \neq 0$, we find two equivalent functionals

$$\begin{aligned} \int_{\Omega_3} d\omega_{\mathbf{x}} f(c|t|\hat{\mathbf{x}}) &= \int_{\Omega_3} d\omega_{\mathbf{x}} \int_{\mathbb{R}_+} dr \delta(r - c|t|) f(r\hat{\mathbf{x}}) = \\ \int_{\mathbb{R}^3} d^3\mathbf{x} \frac{\delta(r - c|t|)}{r^2} f(\mathbf{x}), \end{aligned} \quad (32)$$

$$\begin{aligned} \int_{\Omega_3} d\omega_{\mathbf{x}} f(c|t|\hat{\mathbf{x}}) &= 2 \int_{\Omega_3} d\omega_{\mathbf{x}} \int_{\mathbb{R}_+} dr \delta(r^2 - c^2 t^2) f(\hat{\mathbf{x}}) = \\ 2 \int_{\mathbb{R}^3} d^3\mathbf{x} \frac{\delta(r^2 - c^2 t^2)}{r} f(\mathbf{x}). \end{aligned} \quad (33)$$

This leads to two equivalent expressions for

$$D(t, \mathbf{x}) = \frac{t}{4\pi} \frac{\delta(r - c|t|)}{r^2} = \frac{t}{2\pi} \frac{\delta(r^2 - c^2 t^2)}{r}, \quad (34)$$

and we can check that they satisfy the differential equations (28), (29), and (30), in the sense of distributions $\mathcal{S}'(\mathbb{R}^3)$. We must be aware that as long as $t \neq 0$ these distributions are well defined, but for $t = 0$ they would contain either $\delta(r)$ or $\delta(r^2)$ that are not well-defined distributions on \mathbb{R}_+ . This caveat applies equally to formula (7), which agrees with the second expression in (34). Unfortunately, this caveat is usually omitted or even unknown, and therefore there are attempts to calculate $\partial_t D(t, \mathbf{x})$ exactly at $t = 0$, as in [13], which cannot lead to the correct result. Moreover, the Poisson brackets for the sharp RS vectors at different instants of time, given by (6), do not have a simple limit for equal times if we use (7) for the distribution, which implies the appearance of inconsistency. On the contrary, if we take the Poisson bracket for the smeared RS vectors, then from (11) and (18) we obtain the relation that is smooth at the limit $\tau \rightarrow 0$, i.e.,

$$\begin{aligned} \left\{ V_i^*[t_0, g], V_j[t, f] \right\}_{\text{PB}} &= -i(g, u_{ij} \cos(\omega\tau) f) + (g, u_{ik} u_{kj} \sin(\omega\tau) \omega^{-1} f) = \\ -i \int_{\mathbb{R}^3} d^3\mathbf{x} g(\mathbf{x}) u_{ij} (\cos(\omega\tau) f)(\mathbf{x}) &+ \int_{\mathbb{R}^3} d^3\mathbf{x} u_{ik} u_{kj} (\sin(\omega\tau) \omega^{-1} f)(\mathbf{x}). \end{aligned} \quad (35)$$

5. Conclusions

The smeared RS vectors correctly describe classical free electromagnetic fields, with no ill-defined mathematical expressions at any stage of the

calculations; instead, tempered distributions appear naturally. We have explicitly calculated the distributions appearing in Poisson brackets and in the time evolution formula. Such analysis can be extended to both quantum electromagnetic field

theory and massive fields. In particular, for massive fields we can determine the Pauli–Jordan function as a tempered distribution using an improved scheme to the one presented in [13]. While the massless part carries the most singular contribution, the remaining part can be calculated quite easily.

Acknowledgments

I would like to express my deep gratitude and respect to Professor Iwo Białynicki-Birula for the help and support he gave me during the preparation of my PhD thesis, suggesting its topic and being its kind and helpful supervisor.

Appendix: Some useful integral equations

By performing direct integration over the unit sphere embedded in \mathbb{R}^3 , the following integral relations can be proved

$$\int_{\Omega_3} d\omega_{\mathbf{x}} \partial_i f(\mathbf{x}) = \frac{1}{r^2} \int_{\Omega_3} d\omega_{\mathbf{x}} \frac{\partial}{\partial r} \left[r x^i f(r\hat{\mathbf{x}}) \right], \quad (36)$$

$$\int_{\Omega_3} d\omega_{\mathbf{x}} \nabla^2 f(r\mathbf{x}) = \frac{1}{r} \int_{\Omega_3} d\omega_{\mathbf{x}} \frac{\partial^2}{\partial r^2} \left[r f(r\hat{\mathbf{x}}) \right]. \quad (37)$$

Then, (36) applied to (26) gives $D[t, \partial_i f]$

$$D[t, \partial_i f] = \frac{1}{4\pi c^2 t} \int_{\Omega_3} d\omega_{\mathbf{x}} \left[\frac{\partial}{\partial r} (r x^i f(r\hat{\mathbf{x}})) \right]_{r=c|t|}. \quad (38)$$

The temporal derivative of (26) is

$$\frac{d}{dt} D[t, f] = \frac{1}{4\pi} \int_{\Omega_3} d\omega_{\mathbf{x}} \left[\frac{\partial}{\partial r} (r f(r\hat{\mathbf{x}})) \right]_{r=c|t|}, \quad (39)$$

which for a test function $x^i f(\mathbf{x}) \in \mathcal{S}(\mathbb{R}^3)$ takes the form

$$\frac{d}{dt} D[t, x^i f] = \frac{1}{4\pi} \int_{\Omega_3} d\omega_{\mathbf{x}} \left[\frac{\partial}{\partial r} (r x^i f(r\hat{\mathbf{x}})) \right]_{r=c|t|}. \quad (40)$$

The second order temporal derivative of (26) is

$$\frac{\partial^2}{\partial t^2} D[t, f] = \frac{c \operatorname{sgn}(t)}{4\pi} \int_{\Omega_3} d\omega_{\mathbf{x}} \left[\frac{\partial^2}{\partial r^2} (r f(r\hat{\mathbf{x}})) \right]_{r=c|t|}. \quad (41)$$

If one inserts (37) into (26), then one finds

$$D[t, \nabla^2 f] = \frac{t}{4\pi c|t|} \int_{\Omega_3} d\omega_{\mathbf{x}} \left[\frac{\partial^2}{\partial r^2} (r f(r\hat{\mathbf{x}})) \right]_{r=c|t|}. \quad (42)$$

References

- [1] L. Silberstein, *Ann. Phys.* **327**, 579 (1907).
- [2] L. Silberstein, *Ann. Phys.* **329**, 783 (1907).
- [3] I. Białynicki-Birula, *Photon wave function, Progress in Optics*, Vol. 36, Ed. E. Wolf, Elsevier, Amsterdam 1996.
- [4] I. Białynicki-Birula, Z. Białynicka-Birula, *Phys. Rev. Lett.* **108**, 140401 (2012).
- [5] A. Aste, *J. Geom. Symmetry Phys.* **28**, 47 (2012).
- [6] S.A. KHan, *Phys. Scripta* **71**, 440 (2005).
- [7] I. Białynicki-Birula, Z. Białynicka-Birula, *J. Phys. A: Math. Theor.* **46**, 053001 (2013), Corrigendum *J. Phys. A: Math. Theor.* **46**, 159501 (2013).
- [8] I. Białynicki-Birula, Z. Białynicka-Birula, *Quantum Electrodynamics*, Pergamon, Oxford 1975.
- [9] P. Jordan, W. Pauli, *Z. Physik.* **47**, 151 (1928).
- [10] J.A. Przeszowski, E. Dzimida-Chmielewska, J.L. Cieśliński, *Symmetry* **14**, 241 (2022).
- [11] I.M. Gel'fand, G.E. Shilov, *Generalized Functions: Properties and Operations*, Vol. 1, original by E. Saletan (1958), translated from Russian, Academic Press, New York 1964 p. 192.
- [12] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series, and Products* 6th Ed., Academic Press, San Diego 2000.
- [13] Daqing Liu, Furui Chen, Shuyue Chen, Ning Ma, *Eur. J. Phys.* **41**, 035406 (2020).

DEDICATED TO PROFESSOR IWO BIAŁYŃICKI-BIRULA ON HIS 90TH BIRTHDAY

— ◊ —

Factorization with a Logarithmic Energy Spectrum of a Central Potential

F. GLEISBERG^a AND W.P. SCHLEICH^{a,b,*}

^a*Institut für Quantenphysik and Center for Integrated Quantum Science and Technology (IQST), Universität Ulm, D-89069 Ulm, Germany*

^b*Hagler Institute for Advanced Study at Texas A & M University, Texas A & M AgriLife Research, Institute for Quantum Studies and Engineering (IQSE) and Department of Physics and Astronomy, Texas A & M University, College Station, Texas 77843-4242, USA*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S112](https://doi.org/10.12693/APhysPolA.143.S112)*e-mail: wolfgang.schleich@uni-ulm.de

We propose a method to factor numbers based on two interacting bosonic atoms in a central potential, where the single-particle spectrum depends logarithmically on the radial quantum numbers of states corresponding to zero angular momentum. Bosons initially prepared in the ground state are excited by a sinusoidally time-dependent interaction into a state characterized by quantum numbers, which represent the factors of the number encoded in the frequency of the perturbation. We also discuss the complete single-particle spectrum as well as the limitations of our method caused by decoherence.

topics: factorization, logarithmic energy spectrum, central potential, cold atoms

1. Introduction

It is well-known that the decomposition of a positive composite integer into a product of prime factors is a difficult problem in number theory since it requires *non-polynomial* time on a classical computer, which makes it attractive for cryptological applications [1]. Indeed, decoding a message encoded by the famous Rivest–Shamir–Adleman (RSA) protocol [2] requires the decomposition of a large semiprime, i.e., an integer composed by two primes, in a reasonable time. Such a decomposition can be easily prevented by choosing larger and larger semiprimes. The topic of prime factorization is intimately connected to Peter Shor because on an ideal quantum computer Shor’s factorizing algorithm [3] takes only *polynomial time* and is therefore expected to break the RSA scheme in the future.

1.1. Factorization based on a central potential with logarithmic spectrum

As an alternative method, we have studied [4–6] the factorization of integers using bosonic atoms in one- and two-dimensional potentials, both with a logarithmic energy spectrum. Our present

theoretical study represents an extension of these thoughts and is motivated by two features: (i) it is possible [7] to create and control almost any kind of potential for the center-of-mass motion of the atom using adiabatic potentials, and (ii) bosons in a spherically symmetric *parabolic* potential as well as in a spherical box provide textbook examples for the thermodynamics of the Bose–Einstein condensation [8, 9]. For this reason, in this article, we numerically construct a *central* potential with a logarithmic energy spectrum and propose a factorization algorithm.

The two bosons originally trapped in the ground state of this potential are excited by a periodic perturbation with a frequency governed by the semiprime we want to factor. At a later time, the bosons are found with a probability of about one-half in a state where the energies of the individual bosons contain the factors of the semiprime. Thus a measurement of these energies provides us with the factors we are looking for.

Many ways to experimentally implement our scheme are offered. The most promising one takes advantage of the fact that the interaction of an atom with an electromagnetic wave, which is far detuned from the atom’s resonance, experiences [10]

a potential for its center-of-mass motion determined by the spatial dependence of the light intensity. Hence, by tailoring the intensity distribution to the predetermined shape of the potential, one creates the desired spectrum. In this way, it was possible to create [11], for example, a potential whose energy eigenvalues are given by prime numbers. Obviously, in the context of factorization, the potential with an energy spectrum given by the *logarithm* of primes is of interest and was proposed in [5].

Malcolm G. Boshier at Los Alamos National Laboratory has kindly informed us [12] that he is presently pursuing our approach to factor numbers. By shaking the one-dimensional potential associated with a logarithmic energy spectrum, he and his team could already excite individual energy states as well as their coherent superposition. Energy measurement is achieved by imaging the atoms and counting the number of nodes and anti-nodes of the energy wave functions.

We emphasize that the spherical symmetry of the unperturbed potential is crucial for the protocol proposed in this article. Among the necessities to experimentally obtain symmetry of this kind is microgravity [13]. Hence, a drop tower, for example, the one in Bremen [14], a sounding rocket in space [15], or the International Space Station [16, 17] could provide such an environment.

Moreover, central to our considerations are *s*-states, i.e., the states of vanishing angular momentum. It is worthwhile mentioning that such states have also played a major role in the studies [18–20] of the unusual dynamics of free wave packets. Indeed, they display focusing or defocusing effects even in the absence of external potentials or position-dependent phase factors, and are the result of the dependence of the Laplacian in the Schrödinger equation on the number of space dimensions. This manifestation of the dimensionality of space in quantum mechanics [21] is the analogue of the violation of Huygens’s principle in electrodynamics.

1.2. Overview

This article is organized as follows. In Sect. 2 we introduce the logarithmic energy spectrum and discuss the distribution of a given energy onto two single-particle states. Moreover, we recall a *one*-dimensional potential giving rise to such a spectrum. We then solve in Sect. 3 the Schrödinger equation in *three* dimensions and show that the *s*-states, i.e., states corresponding to the zero azimuthal quantum number, suffice to determine the central potential with a logarithmic energy spectrum. Moreover, we take into account the boundary condition at the origin and demonstrate that the single-particle *s*-states exhibit an energy spectrum identical to the one introduced in Sect. 2. In Sect. 4, we define the corresponding two-particle states using the bra–ket notation.

Section 5 constitutes the main part of our article. Here we discuss the realization of our factorizing scheme by two bosonic atoms moving in the central potential determined in Sect. 3, and being excited by a time-dependent interaction into the factor state. We derive the solution of the corresponding Schrödinger equation within the rotating wave approximation and demonstrate that after measuring the single-particle energies at random times the factor state is found with a probability of about 1/2. A brief discussion of the limitations of our method completes this section. We conclude with a short summary in Sect. 6.

Central to our proposal is the fact that the energy spectrum of our central potential does not display any accidental degeneracies. For an elementary discussion of this point, we refer to Appendix.

2. Logarithmic spectrum and potential in one dimension

In the present section, we first introduce the logarithmic energy spectrum and discuss its special role in finding factors of an integer. We then turn to the distribution of the given energy onto two subsystems. This discussion constitutes the foundation for our factorization protocol. We conclude by recalling [4, 22] the potential in one space dimension that gives rise to such a spectrum.

2.1. Central idea for factorization

Our scheme is based on the logarithmic energy spectrum

$$E_k(L) \equiv \hbar\omega_0 \ln \left(\frac{k}{L} + 1 \right) \quad (1)$$

with $k = 0, 1, 2, \dots$ and $E_0(L) = 0$. Here, the constant L plays the role of a scaling parameter and $\hbar\omega_0$ is the unit of energy.

In order to find the factors of a given semiprime $N \equiv pq$, we distribute the total energy

$$E_{\text{total}}(N; L) \equiv \hbar\omega_0 \ln \left(\frac{N}{L^2} \right) \quad (2)$$

onto *two* subsystems, each with the spectrum defined by (1) according to the relation

$$E_{\text{total}}(N; L) = \hbar\omega_0 \ln \left(\frac{p}{L} \right) + \hbar\omega_0 \ln \left(\frac{q}{L} \right) \quad (3)$$

that is

$$E_{\text{total}}(N; L) = E_{p-L}(L) + E_{q-L}(L). \quad (4)$$

Since L appears in the indices of the energies in (4), it has to be an integer. No negative indices are present in (1), therefore N cannot contain the factors $q < L$ and $p < L$.

Moreover, the factor $q = L$ or $p = L$ causes the unwanted case that the total energy given by (3) may be transferred to one of the two subsystems while the other one remains in the ground state

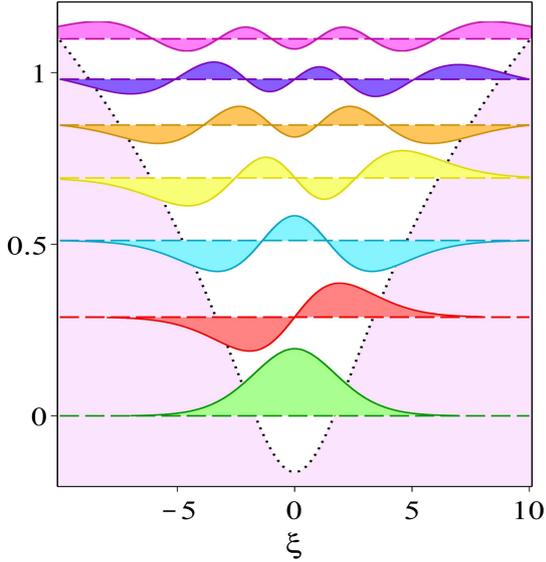


Fig. 1. The one-dimensional potential $V = V(\xi; 3)$ (dotted line) creating a logarithmic energy spectrum for the scaling parameter $L = 3$ as a function of the dimensionless coordinate $\xi \equiv \alpha x$ with $\alpha^2 \equiv \mu\omega_0/\hbar$. This potential is determined [22] numerically by an iteration algorithm based on a perturbation theory using the Hellmann–Feynman theorem, and is designed to obtain a logarithmic dependence of the energy eigenvalues $E_k(L = 3)$ on the quantum number k as given by (1). In the neighborhood of the origin, the potential is approximately harmonic whereas for large values of ξ it is logarithmic. The solid lines depict the numerically determined energy wave functions $u_k = u_k(\xi, 3)$ of the first seven states in their dependence on the dimensionless position. Both, the energies $E_k(3)$, $k = 0, 1, \dots, 6$ (dashed lines) as well as the potential $V = V(\xi; 3)$ are expressed in units of $\hbar\omega_0$.

and no factorization takes place. Hence, we have to remove the factors $2, 3, \dots, L$, which can be done by the division before we start our factorization protocols.

However, if L is chosen to be unity the trivial factorization $N = 1 \times N$ cannot be excluded. Moreover, in Sect. 3 we shall see that L has to be odd. Therefore, throughout our article, we consider the case $L \geq 3$.

The question of the uniqueness of the distribution according to (3) is easily answered because the fundamental theorem of arithmetics guarantees that the decomposition of the integer N is unique if both factors, p and q , are prime.

For our factorization protocol, the two subsystems have to be brought into a state of total energy (3), followed by a measurement of their individual energies which allows us to determine the factors p and q , as described in Sect. 5. In the remainder of our article, we shall concentrate on the factorization of semiprimes.

2.2. The inverse problem

Next, we briefly address the problem of creating such a logarithmic energy spectrum by determining the appropriate potential V in one space dimension denoted by the coordinate x . For the sake of simplicity, we assume a symmetric potential $V(x) = V(-x)$ with $-\infty < x < \infty$, where the time-independent Schrödinger equation for a particle of mass μ reads

$$\left[-\frac{\hbar^2}{2\mu} \frac{d^2}{dx^2} + V(x; L) - E_k(L) \right] u_k(x; L) = 0. \quad (5)$$

Since the eigenvalues $E_k(L)$ depend on the scaling parameter L , the potential $V = V(x; L)$ and the eigenfunctions $u_k = u_k(x; L)$ must also display the same dependence.

Under standard circumstances, the potential $V = V(x; L)$ is given and the eigenvalues must be found. However, now the energy spectrum is prescribed, and we have to determine the potential $V = V(x; L)$ from the Hellmann–Feynman theorem and the iteration algorithm described in the previous article [22]. In Fig. 1 we show the so-obtained potential $V = V(x; L = 3)$ together with the eigenfunctions $u_k = u_k(x; L = 3)$ and energy eigenvalues $E_k(L = 3)$ for $0 \leq k \leq 6$.

We conclude by noting that in [22], we constructed this potential with a logarithmic energy spectrum to obtain wave packets whose auto-correlation function yields the Dirichlet representation of the Riemann zeta function [23]. However, it has also become crucial for our factorization proposals [4–6] in one and two dimensions and plays a crucial role in the present article when we propose an algorithm for three dimensions.

3. Logarithmic energy spectrum in three dimensions

In the present section, we realize the logarithmic spectrum (1) for a particle of mass μ moving in three space dimensions in a central potential $V = V(r)$ that we shall determine. For this purpose, we start from the time-independent Schrödinger equation in three dimensions and concentrate on the radial wave functions. Due to the vanishing boundary condition at the origin, the corresponding eigenfunctions are the odd ones of the symmetric one-dimensional problem in Sect. 2. We conclude by discussing the resulting energy spectrum.

3.1. From three dimensions to one dimension

We start from the Schrödinger equation

$$\left[-\frac{\hbar^2}{2\mu} \Delta + V(r) - E \right] \psi(r, \theta, \varphi) = 0, \quad (6)$$

in spherical coordinates r , θ and φ and employ the ansatz

$$\psi_{j,\ell,m}(r, \theta, \varphi) \equiv R_{j,\ell}(r) Y_\ell^m(\theta, \varphi) \quad (7)$$

for the energy eigenfunctions $\psi_{j,\ell,m}$ that are simultaneous eigenfunctions of the Hamiltonian \hat{H} , the square of the angular momentum \hat{L}^2 and its z -component \hat{L}_z forming a complete set of commuting operators with the eigenvalues $E_{j,\ell}$, $\hbar^2 \ell(\ell+1)$ and $\hbar m$, respectively. The radial quantum number j , as well as the azimuthal quantum number ℓ , takes value $0, 1, 2, \dots$ while the magnetic quantum number m assumes $2\ell+1$ values given by $-\ell \dots \ell$. The functions $Y_\ell^m = Y_\ell^m(\theta, \varphi)$ are the spherical harmonics.

Since the solution of (6) can be found in standard textbooks on quantum mechanics, we jump directly to the radial equation

$$\left[-\frac{\hbar^2}{2\mu r} \frac{d^2}{dr^2} r + \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} + V(r) - E_{j,\ell} \right] R_{j,\ell}(r) = 0 \quad (8)$$

valid in the region $r \geq 0$ with the condition that $R_{j,\ell} = R_{j,\ell}(r)$ has to be square integrable and finite at the origin $r = 0$.

We consider s -states defined by $\ell = 0$ and set

$$R_{j,0}(r) \equiv \frac{v_{j,0}(r)}{r} \quad (9)$$

with the boundary condition

$$v_{j,0}(0) = 0 \quad (10)$$

at the origin.

3.2. Potential

For the sake of simplicity in the notation, we now suppress the index $\ell = 0$ for the time being. In Sect. 2, we obtained the potential $V = V(x, L)$ and the functions $u_k = u_k(x, L)$ for one space dimension associated with a logarithmic energy spectrum. The three-dimensional potential $V^{(3d)} = V^{(3d)}(r; L)$ as well as the eigenfunctions $v_j(r; L)$ follow by replacing the coordinate x by r in both, where now only the region $r \geq 0$ is considered that is

$$V^{(3d)}(r, L) \equiv V(x = r, L). \quad (11)$$

However, only *odd* solutions $u_k = u_k(x; L)$ of (5) with $k \equiv 2j+1$ can satisfy the boundary condition (10). Therefore, the energies $E_k(L)$ as well as the eigenfunctions $u_k(x; L)$ with *even* index k , which are present in one dimension in (5), do not appear anymore in three dimensions.

Figure 2 shows the potential $V^{(3d)} = V^{(3d)}(r; L)$ for the position vector \mathbf{r} in the x - y plane. We emphasize that states with quantum numbers $\ell > 0$ are not needed to determine V .

3.3. Energy eigenvalues and wave functions of s -states

We now show that the remaining spectrum $E_{2j+1}(L)$ does indeed have the form of (1) and therefore guarantees the validity of the results from

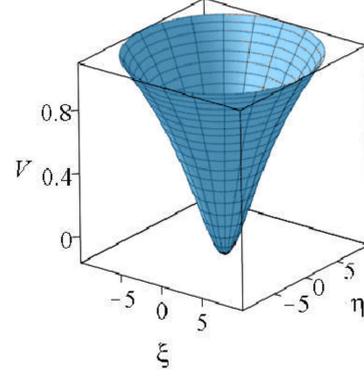


Fig. 2. Three-dimensional potential $V^{(3d)} = V^{(3d)}(r; L = 3)$ in units of $\hbar\omega_0$ creating the logarithmic energy spectrum (13) with the scaling parameter $K = 2$ as a function of the dimensionless coordinates $\xi \equiv \alpha x$ and $\eta \equiv \alpha y$, represented in the plane $z = 0$.

Sect. 2 which are essential for our factorization procedure. For this purpose, we shift the energies

$$E_{2j+1}(L) = \hbar\omega_0 \ln \left(\frac{2j+1}{L} + 1 \right) \quad (12)$$

with $j = 0, 1, 2, 3, \dots$ by a constant amount $\delta E \equiv -\hbar\omega_0 \ln(1/L + 1)$ leading us to new spectrum

$$E_j^{(3d)}(K) = \hbar\omega_0 \ln \left(\frac{j}{K} + 1 \right), \quad (13)$$

which is identical to the single-particle spectrum (1) except that L has to be replaced by the new scaling parameter

$$K \equiv \frac{L+1}{2}. \quad (14)$$

For K to be a positive integer, L has to be odd. All the statements made in Sect. 2 referring to the scaling length L remain valid here provided L is replaced by K .

The eigenfunctions $v_j = v_j(r; K)$ corresponding to $E_j^{(3d)}(K)$ read

$$v_j(r; K) \equiv u_{2j+1}(r; L). \quad (15)$$

Figure 3 shows the radial wave functions

$$R_j(r) \equiv \frac{v_j(r; K)}{r} \quad (16)$$

for indices $j = 0, \dots, 5$ together with the potential $V^{(3d)} = V^{(3d)}(r; K = 2)$ and the energy levels $E_j^{(3d)}(K = 2)$ given by (13).

4. Two-particle bosonic states

So far we have concentrated on a single particle exposed to a central potential giving rise to the logarithmic energy spectrum of (1). We now address the two-particle situation which is central to our factorization scheme.

To simplify the notation we turn to the bra-ket formalism and the time-independent single-particle Schrödinger equation for s -states takes the form ($j = 0, 1, 2, \dots$)

$$\hat{H}(K) |j\rangle = E_j^{(3d)}(K) |j\rangle, \quad (17)$$

where we have suppressed again the quantum numbers $\ell = m = 0$. The Hamiltonian $\hat{H}(K)$ is characterized by the parameter K defined by (14).

The corresponding Schrödinger equation for the two *non-interacting* bosons denoted by 1 and 2 reads

$$\left(\hat{H}_{1,2}(K) - E_{m,n}(K)\right) |m, n\rangle_B = 0, \quad (18)$$

where

$$\hat{H}_{1,2}(K) \equiv \hat{H}_1(K) + \hat{H}_2(K) \quad (19)$$

is the Hamiltonian of both bosons with total energy

$$E_{m,n}(K) \equiv E_m^{(3d)}(K) + E_n^{(3d)}(K), \quad (20)$$

in accordance with (13) and (14).

We note that bosonic two-particle states are defined by

$$|m, n\rangle_B \equiv \frac{1}{\sqrt{2}}(|m, n\rangle + |n, m\rangle), \quad (21)$$

where $|m, m\rangle_B \equiv |m, m\rangle$.

If two identical non-interacting bosons are in a state with energy

$$\hbar\omega_0 \ln\left(\frac{N}{K^2}\right) = E_{p-K}^{(3d)} + E_{q-K}^{(3d)}, \quad (22)$$

where $N \equiv pq$ is semi-prime, then according to (4) and (13) the bosons are in the factor state $|p-K, q-K\rangle_B$. A measurement of the energy of one of the bosons can only result in $\hbar\omega_0 \ln(p/K)$ or $\hbar\omega_0 \ln(q/K)$ and immediately yields the prime factors p and q , respectively.

5. Factorization algorithm

The present section contains the main results of our article. Here we propose and analyze the realization of the factorization protocol of Sect. 2 by two interacting identical bosons placed in the central potential shown in Fig. 2, with the single-particle spectrum given by (13).

Starting from the corresponding Schrödinger equation, we first derive the equations of motion for the probability amplitudes of the ground state and the relevant excited states. Here we keep all three quantum numbers and denote them by $\mathbf{k} = (j, \ell, m)$. To simplify the notation further, in the remainder of the article, we suppress the scaling parameter K as well as the subscript B and the superscript 3d.

We then derive an explicit expression for the matrix element of the Fermi point interaction [24, 25] and simplify the equations of motion for the resulting probability amplitude with the help of the rotating wave approximation (RWA) [26, 27]. The approximation reduces the equations of motion to

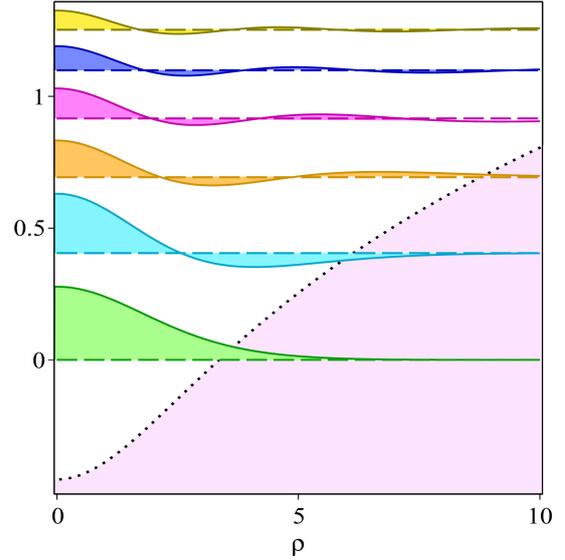


Fig. 3. Central potential $V^{(3d)} = V^{(3d)}(\rho; K = 2)$ represented by a dotted line creating the logarithmic energy spectrum $E_j^{(3d)}(K = 2)$ of (13) in units of $\hbar\omega_0$ as a function of the dimensionless radius $\rho \equiv \alpha r$ together with the corresponding radial functions $R_j = R_j(\rho)$ defined by (16) of the first six states in their dependence on ρ . We shifted the energies to ensure that the ground state has vanishing energy.

a two-level Rabi problem involving the ground state and the factor state. This insight allows us to estimate the probability of success of our factorization scheme. Moreover, we briefly discuss the limitations of our method.

5.1. Coupled set of equations

We prepare two bosons in the ground state $|\mathbf{0}, \mathbf{0}\rangle$ and expose them at $t = 0$ to the perturbation

$$\delta V(\mathbf{r}_1, \mathbf{r}_2; t) \equiv \gamma \sin(\omega_{\text{ext}} t) w(\mathbf{r}_1, \mathbf{r}_2) \quad (23)$$

where γ is a constant and the frequency ω_{ext} is chosen later in a way suitable for the factorization procedure. Moreover, the interaction term w contains the coordinates \mathbf{r}_1 and \mathbf{r}_2 of both particles.

The time evolution of the two-particle state $|\Psi(t)\rangle$ is now governed by the Schrödinger equation

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = (\hat{H}_{1,2} + \delta V(t)) |\Psi(t)\rangle \quad (24)$$

in three dimensions with the unperturbed stationary states

$$\hat{H}_{1,2} |\mathbf{k}_1, \mathbf{k}_2\rangle = E_{\mathbf{k}_1, \mathbf{k}_2} |\mathbf{k}_1, \mathbf{k}_2\rangle. \quad (25)$$

When we substitute the expansion

$$|\Psi(t)\rangle = \sum_{\mathbf{k}_1, \mathbf{k}_2} b_{\mathbf{k}_1, \mathbf{k}_2}(t) e^{-iE_{\mathbf{k}_1, \mathbf{k}_2} t/\hbar} |\mathbf{k}_1, \mathbf{k}_2\rangle \quad (26)$$

of $|\Psi(t)\rangle$ by the two-particle eigenstates $|\mathbf{k}_1, \mathbf{k}_2\rangle$ of the unperturbed Hamiltonian $\hat{H}_{1,2}$ into (24), we arrive at the coupled system

$$i\hbar \dot{b}_{\mathbf{k}_1, \mathbf{k}_2}(t) = \gamma \sin(\omega_{\text{ext}} t) \sum_{\mathbf{k}'_1, \mathbf{k}'_2} e^{i(E_{\mathbf{k}_1, \mathbf{k}_2} - E_{\mathbf{k}'_1, \mathbf{k}'_2})t/\hbar} W_{\mathbf{k}_1, \mathbf{k}_2; \mathbf{k}'_1, \mathbf{k}'_2} b_{\mathbf{k}'_1, \mathbf{k}'_2}(t) \quad (27)$$

with the initial conditions $b_{\mathbf{k}_1, \mathbf{k}_2}(0) = 1$ for $j_1 + j_2 + \ell_1 + \ell_2 = 0$, and $b_{\mathbf{k}_1, \mathbf{k}_2}(0) = 0$ otherwise, which has to be solved for the probability amplitudes $b_{\mathbf{k}_1, \mathbf{k}_2}(t)$.

We conclude by emphasizing that the eigenstates $|\mathbf{k}_1, \mathbf{k}_2\rangle$ of $\hat{H}_{1,2}$, the amplitudes $b_{\mathbf{k}_1, \mathbf{k}_2}(t)$, and the matrix elements

$$W_{\mathbf{k}_1, \mathbf{k}_2; \mathbf{k}'_1, \mathbf{k}'_2} \equiv \langle \mathbf{k}_1, \mathbf{k}_2 | w(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2) | \mathbf{k}'_1, \mathbf{k}'_2 \rangle \quad (28)$$

are ‘‘bosonic’’ ones in the sense of (21) and are built out of the eigenstates $|\mathbf{k}_1, \mathbf{k}_2\rangle$ of $\hat{H}_{1,2}$ and the spacial part w of the perturbation $\delta\hat{V}$ defined by (23). Moreover, in the summation in (26) and (27), the same states must not be counted twice.

5.2. Matrix elements for the contact interaction

Next we derive an explicit expression for the matrix element $W_{\mathbf{k}_1, \mathbf{k}_2; \mathbf{k}'_1, \mathbf{k}'_2}$ (28), assuming the contact interaction

$$w(\mathbf{r}_1, \mathbf{r}_2) \equiv \delta^{(3)}(\mathbf{r}_1 - \mathbf{r}_2). \quad (29)$$

between two particles, providing us with the selection rules of the transitions from the ground state.

Needless to say, we are well aware that we should use the *regularized* delta function [24, 25] rather than the delta function of (29) for the zero-range potential. However, in order to bring out most clearly the central points of our factorization algorithm, we resort to the elementary version of (29) of the contact potential and postpone the complete analysis to a future publication.

Due to the delta function in w , the matrix element (28) reduces in position space to

$$W_{\mathbf{k}_1, \mathbf{k}_2; \mathbf{k}'_1, \mathbf{k}'_2} \equiv \int d^3r \psi_{\mathbf{k}_1}^*(\mathbf{r}) \psi_{\mathbf{k}_2}^*(\mathbf{r}) \psi_{\mathbf{k}'_1}(\mathbf{r}) \psi_{\mathbf{k}'_2}(\mathbf{r}). \quad (30)$$

$$i\hbar \dot{b}_{0,0;0,0}(t) = \gamma \sin(\omega_{\text{ext}} t) \sum_{j_1, j_2, \ell} e^{-i(E_{j_1, \ell} + E_{j_2, \ell})t/\hbar} W_{0,0;0,0; j_1, \ell, j_2, \ell} b_{j_1, \ell; j_2, \ell}(t) \quad (34)$$

with the matrix element (31) and vanishing energy of the ground state of the two bosons.

The external frequency ω_{ext} is chosen such that the energy $\hbar\omega_{\text{ext}}$ is identical to the sum

Having in mind that we start our algorithm at time $t=0$ with the two particles in the ground state $|\mathbf{0}, \mathbf{0}\rangle$, we consider the matrix elements $W_{\mathbf{0}, \mathbf{0}; \mathbf{k}_1, \mathbf{k}_2}$ for a transition into the excited state $|\mathbf{k}_1, \mathbf{k}_2\rangle$. With the product ansatz (7) for $\psi_{\mathbf{k}}$, we therefore arrive at the expression

$$W_{\mathbf{0}, \mathbf{0}; \mathbf{k}_1, \mathbf{k}_2} = \frac{1}{4\pi} \int_0^\infty dr r^2 R_{0,0}(r)^2 R_{j_1, \ell_1}(r) \times R_{j_2, \ell_2}(r) \delta_{\ell_1, \ell_2} \delta_{m_1 + m_2, 0}. \quad (31)$$

Here we have applied the well-known orthonormality relation

$$\int d\Omega Y_{\ell_1}^{m_1*}(\theta, \varphi) Y_{\ell_2}^{m_2}(\theta, \varphi) = \delta_{\ell_1, \ell_2} \delta_{m_1, m_2} \quad (32)$$

of the spherical harmonics and the identity

$$Y_\ell^{m*}(\theta, \varphi) = Y_\ell^{-m}(\theta, \varphi) \quad (33)$$

for their complex conjugate together with $Y_0^0 \equiv 1/\sqrt{4\pi}$.

5.3. Encoding the number to be factored and rotating wave approximation

We now employ the expression (31) for the matrix element $W_{\mathbf{0}, \mathbf{0}; \mathbf{k}_1, \mathbf{k}_2}$ to simplify the system of coupled equations (27) considerably. For this purpose, we set the two magnetic quantum numbers $m_1 = m_2 = 0$ and omit them henceforth. This assumption will be justified by the calculation below. The single-particle state is now only characterized by *two* quantum numbers j and ℓ .

Thus we study the set of equations

$$E_{p-K, 0; q-k, 0} = E_{p-K, 0} + E_{q-K, 0} = \hbar\omega_0 \ln\left(\frac{N}{K^2}\right) \quad (35)$$

of the energies of the factor states, and is determined by the number $N = pq$ to be factored.

Next, we address the product

$$\mathcal{E}(t) \equiv \frac{1}{2i} \left(e^{i(E_{p-K, 0} + E_{q-K, 0})t/\hbar} - e^{-i(E_{p-K, 0} + E_{q-K, 0})t/\hbar} \right) e^{-i(E_{j_1, \ell} + E_{j_2, \ell})t/\hbar} \quad (36)$$

of time-dependent factors, which appears on the right-hand side of (34) when we decompose the sine function into the difference of two phase factors.

The essence of RWA, when applied to (34), is to retain only terms with constant coefficients on the right-hand side and to neglect all oscillating terms. Indeed, when we assume that $p \geq q$, only the term with $j_1 \equiv p-K$, $j_2 \equiv q-K$, and $\ell = 0$ survives, providing us with the contribution $(2i)^{-1}$.

The Appendix discusses the possibility of accidental degeneracy in the logarithmic single-particle spectrum $E_{j,\ell}$ given by (1), the absence of which is confirmed therein. None of the terms with $\ell \geq 1$ can therefore lead to additional constant terms in (36).

Within RWA, (34) reduces to the equation

$$i\hbar \dot{b}_{0,0}(t) = \frac{\gamma}{2i} W_{0,0;p-K,q-K} b_{p-K,q-K}(t), \quad (37)$$

where the index $\ell = 0$ is present in the matrix element, and in the probability amplitudes it is omitted here for convenience.

We derive a second equation by selecting the term with $j_1 \equiv p-K$ and $j_2 \equiv q-K$ from (27) and proceeding as before we arrive at the equation of motion

$$i\hbar \dot{b}_{p-K,q-K}(t) = -\frac{\gamma}{2i} W_{p-K,q-K;0,0} b_{0,0}(t). \quad (38)$$

of the unperturbed s -states.

5.4. Factor state and its probability

We note that (37) and (38) characterize the dynamics of a two-boson system driven by the periodic perturbation (23) with energy (35). Together with the initial conditions $b_{0,0}(0) \equiv 1$ and $b_{p-K,q-K}(0) \equiv 0$ as well as the symmetry relation

$$W_{m,n;0,0} = W_{0,0;m,n}, \quad (39)$$

the resulting probability amplitude for the ground state reads

$$b_{0,0}(t) = \cos(\Omega t), \quad (40)$$

whereas for the factor state we find

$$b_{p-K,q-K}(t) = \sin(\Omega t). \quad (41)$$

The Rabi frequency

$$\Omega \equiv \frac{\gamma}{2\hbar} W_{0,0;p-K,q-K} \quad (42)$$

is proportional to the interaction matrix element of (31).

In Sects. 2 and 4 we have shown that if the bosons are in the factor state $|p-K, q-K\rangle$ they have a two-particle energy $\hbar\omega_0 \ln(N/K^2)$ given by (22) with $N = pq$. As mentioned there, the factors p or q are determined by a measurement of the single-particle energies (3), and the factorization protocol has ended successfully.

At the time t , the system can be found with probability $|b_{p-K,q-K}(t)|^2$ in the factor state, and at the time equal to an odd multiple of $\pi/(2\Omega)$, it is there with 100% certainty. Unfortunately, the Rabi

frequency Ω is not known. Instead, we content ourselves with measuring at a time chosen randomly from a time interval $[0, T]$ much larger than π/Ω . According to (41), the probability to find the factor state is about one-half. Then the measurement of the single-particle energy gives one of the factors while the other one follows from division.

An estimate for a time of measurement by making a guess for the factors p and q and so determining the Rabi frequency (42) was presented in a previous article [4].

5.5. Limitations

In the present section, we briefly address the obstacles that prevent our protocol from factoring larger and larger semiprimes, and in particular, we derive the condition for the largest number N we can factor. Here we address especially limitations due to decoherence.

According to [26], there is a high probability for a *periodic* transition into the factor state as long as the difference between the energies of this state and of the next off-resonant state is larger than the energy $\hbar\Omega$ of the Rabi oscillation. This condition translates into the requirement

$$\hbar\omega_0 \left| \ln\left(\frac{N \pm 1}{K^2}\right) - \ln\left(\frac{N}{K^2}\right) \right| \approx \frac{\hbar\omega_0}{N} \gg \hbar\Omega. \quad (43)$$

Since the Rabi frequency Ω defined by (42) is proportional to the strength γ of the perturbation (23), the inequality (43) can easily be satisfied by choosing γ as small as needed.

Unfortunately, the second condition arises from the fact that we randomly choose the time of measuring the energies of the two bosons from the interval $[0, T]$. To find the factor state with a probability of approximately 1/2, the interval length T has to fulfill the condition $\Omega T \gg 1$.

On the other hand, the system has to be *free of decoherence* during the time interval $T < T_{\text{dec}}$ leading to the two inequalities for the Rabi frequency

$$\Omega \gg \frac{1}{T_{\text{dec}}} \quad (44)$$

and

$$\Omega \ll \frac{\omega_0}{N}. \quad (45)$$

Our aim is now to find an upper limit for the number N to be factored. In [4] and [6] for different experimental situations and models for the spatial part of the interaction, we have found an approximate N -dependence

$$W_{0,0;p-K,q-K} \propto N^{-1/2} \quad (46)$$

of the transition matrix element.

Due to (42) the same scaling holds true, of course, for the Rabi frequency Ω , and the semiprime N to be factored therefore has the upper limit

$$N < \min \left(\left[\frac{\gamma T_{\text{dec}}}{\hbar} \right]^2, \left[\frac{\hbar\omega_0}{\gamma} \right]^2 \right). \quad (47)$$

Assuming that according to (23) the interaction strength γ can be chosen at will, this relation shows that the crucial limiting factor for the magnitude of N is the decoherence time T_{dec} .

6. Conclusions

In the present article, we have proposed a method to find the factors of a semiprime N based on the quantum dynamics of two identical bosonic atoms moving in a spherically symmetric trap whose s -states exhibit a logarithmic single-particle spectrum.

In the first part of our work, we have determined the central potential, which displays such an unusual spectrum. We started by numerically calculating the one-dimensional potential from a logarithmic single-particle spectrum. Taking advantage of the close relationship between three-dimensional spherically symmetric and one-dimensional problems, the central potential was easily found. As expected, it has an energy spectrum with the logarithmic s -wave part, but with a scaling length different from the one in the one-dimensional spectrum.

In the second part, we have attacked the problem of how to force the bosons into the factor state. For this purpose, we excite them from their ground state by a periodic time-dependent contact interaction with a frequency determined by the number N to be factored. To exclude transitions between non- s states, we have discussed *in extenso* the absence of degeneracies.

Then we showed within the framework of the rotating wave approximation that the bosons perform a Rabi oscillation between the ground state and the factor state. The latter emerges with a probability of about one-half when the energies of the bosons are measured at a randomly chosen time. These energies provide us with the factors of N , and our factorization protocol has ended successfully.

Since holographic methods allow us to create almost arbitrary potentials for the center-of-mass motion of atoms and detect them by their fluorescence, an experimental implementation of our factorization scheme is within reach. Indeed, the group of Donatella Cassetari has already used this technique to experimentally realize a potential whose energy eigenvalues are given by the lowest prime numbers. Moreover, the team of Malcolm Boshier even implemented a one-dimensional potential for the logarithmic energy spectrum and observed well-defined excitations of atoms from the ground state to individual energy eigenstates of this potential. Unfortunately, the demonstration of our factorization scheme is still awaiting.

Hence, today's technology already allows us to factor small numbers using this technique. However, three phenomena make the straight-forward application to large composite integers impossible:

(i) decoherence during Rabi oscillations, (ii) scaling of separation between neighboring energy levels with an inverse of the quantum number n requiring increasing accuracy in determining the levels, and (iii) the non-vanishing time for the transition of the two atoms from the ground state to the factor states given by the inverse of the Rabi frequency, which grows with the square root of the number N to be factored.

It is interesting that the same scaling appears in the naive approach towards factoring just trying out all the primes below to the square root of N . Hence, it appears that the RSA scheme is saved by the fact that quantum transitions are not quantum jumps, but rather follow the continuous dynamics given by the Rabi oscillation as dictated by the Schrödinger equation.

Acknowledgments

His classic textbooks *Quantum Electrodynamics* and *Theory of Quanta* co-authored by his wife Zofia and by Marek Cieplak and Jerzy Kamiński, respectively, his seminal work on field quantization without mode functions based on the Wheeler functional, as well as his deep insights into the wave function of the photon, and last not least his entropic uncertainty relations represent only a few of the many outstanding scientific achievements of Professor Iwo Białynicki-Birula.

One of us (W.P.S.) met Iwo exactly 40 years ago at a NATO summer school in Boulder, Colorado, organized by Asim Barut. W.P.S. was immediately impressed by Iwo's unique ability to catch instantly the central point of an argument, his penetrating questions and his love for a good scientific fight, but also by his incredible friendliness. It was therefore only logical that 10 years later Iwo came as the first Humboldt Awardee to the still young *Institute for Quantum Physics at Ulm University*. Over the years we have had a wonderful and extremely productive collaboration, and not only did *we* become close friends but also our families.

We dedicate our article addressing the interface of quantum physics and number theory to our great teacher and fatherly friend Iwo on the occasion of his 90th birthday. We know that this field is dear to his heart, and hope that our little birthday present may find his interest and trigger a discussion leading to a deeper understanding.

Happy Birthday Iwo and many more healthy years filled with fun with physics!

We thank M.G. Boshier, M.A. Efremov and M. Freyberger for stimulating discussions on this topic and M.E.N. Tschaffon for technical assistance. W.P.S. is grateful to the Hagler Institute for Advanced Study at Texas A&M University for a Faculty Fellowship and to Texas A&M University AgriLife Research for its support. The research of the IQST is financially supported by the Ministry of Science, Research and Arts, Baden-Württemberg.

Appendix A: Absence of accidental degeneracy in the logarithmic spectrum

The energy spectrum of *any* central potential exhibits the familiar $(2\ell + 1)$ -fold *essential* degeneracy as the energy levels $E_{j,\ell}$ do not depend on the magnetic quantum number m . It has been proven long ago [28] that the only potentials that show *accidental* degeneracy are the Coulomb potential and the harmonic oscillator. This fact is a consequence of the existence of a conserved quantity [29], which does not commute with any member of the complete system of commuting operators of the problem. In the Coulomb case, this constant of the motion is the well-known Runge–Lenz vector [30, 31], whereas, for the harmonic oscillator, we shall discuss it below. Since our central potential leading to the logarithmic energy spectrum is none of the above, accidental degeneracy must be absent. We shall now study this problem in more detail.

A.1: Non-closed trajectory of a classical particle

We recall that the trajectories of a classical particle in the harmonic oscillator as well as in the Coulomb potential are closed, the latter only for negative energies. Following the textbook [32], we calculate the trajectory of a mass μ , energy E , and angular momentum J oscillating in the effective potential

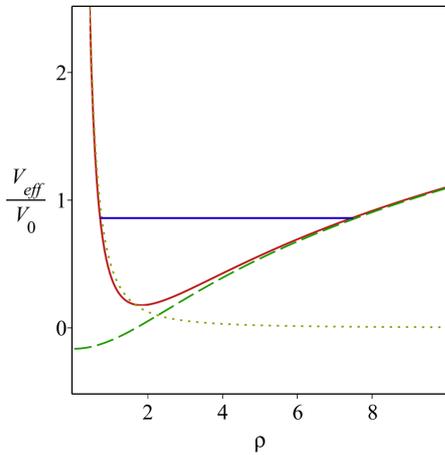


Fig. 4. Scaled effective potential V_{eff} (solid line) formed by the angular momentum barrier (dotted curve) and the potential $V^{(3d)} = V^{(3d)}(r; K = 2)$ (dashed line) which in the quantum case creates the logarithmic energy spectrum given by (13) for the scaling parameter $K = 2$ defined by (14), as a function of the dimensionless radius $\rho \equiv \alpha_{\text{cl}} r$ with $\alpha_{\text{cl}} \equiv (\mu V_0 / J^2)^{1/2}$. The horizontal line denotes the energy $E = 0.86 V_0$ of the radial coordinate $r = r(t)$ of a classical particle moving *periodically* between the left and right turning point. The angle $\theta = \theta(t)$ is *not* periodic, as is the orbit $r = r(\theta)$ shown in Fig. 5.

$$V_{\text{eff}}(r) \equiv \frac{J^2}{2\mu r^2} + V^{(3d)}(r, K) \quad (48)$$

with energy $E = 0.86 V_0$ periodically between the two turning points depicted in Fig. 4.

In Fig. 5 we display five periods of the trajectory $r = r(\theta)$. It is evident that the orbit of the particle precesses around the center of force and does not close, thus indicating the absence of accidental degeneracy.

A.2: Energy spectrum

The most direct way to check for degeneracy is to calculate the energies $E_{j,\ell}$ for the potential under consideration with radial and azimuthal quantum numbers j and ℓ , respectively. If two or more of energies with different indices are equal, degeneracy is present.

Before we get to our potential, we recall the situation for a three-dimensional isotropic harmonic oscillator where the lowest energy levels are displayed in Fig. 6. Indeed, here the energies $E_{j,\ell}$ depend on the combination of both indices j and ℓ on the principal quantum number $n = 2j + \ell$ leading to the degeneracy [33] of levels $E_n = \hbar\omega(n + 3/2)$ demonstrated by the levels with $n = 2, 3, 4$. If the x - and y -axis are oriented along the symmetry axes of the elliptic orbit of the oscillator, then the additional integral of the motion reduces [34] to the scalar function $E_x - E_y$, i.e., a difference between the energies of the motion projections onto the x - and y -axis, respectively.

For our central potential $V^{(3d)} = V^{(3d)}(r, K)$, defined by (11), leading to the logarithmic energy spectrum $E_{j,\ell}(K)$ given by (13), we numerically

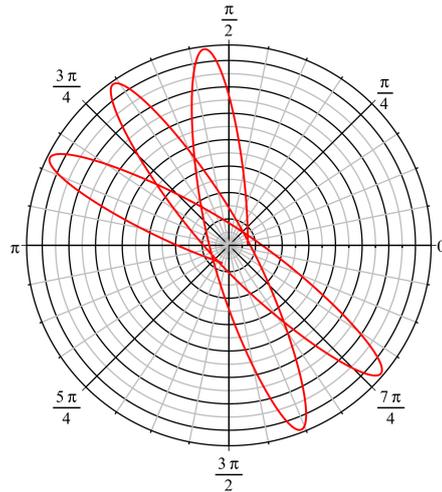


Fig. 5. Trajectory $r = r(\theta)$ of a classical particle with mass μ and energy $E = 0.86 V_0$ moving in the effective potential shown in Fig. 4. The motion starts from an inner turning point at the angle $\theta = 0$. After having covered five periods it reaches again the inner turning point, but now at the angle $\theta \approx 11\pi/8$.

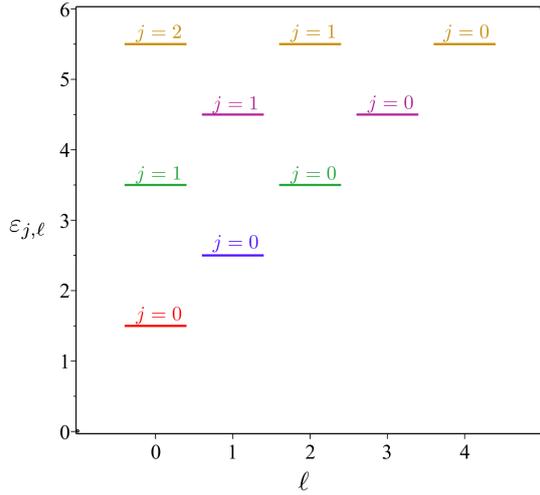


Fig. 6. Lowest dimensionless energies $\varepsilon_{j,\ell} \equiv E_{j,\ell}/(\hbar\omega)$ of an isotropic three-dimensional harmonic oscillator. The levels with the energies $E_n = \hbar\omega(n+3/2)$ show degeneracy as the principal quantum number $n = 2j + \ell$ depends on both the radial quantum number j and the azimuthal quantum number ℓ . For example, the level $n = 2$ is doubly degenerate for pairs of quantum numbers $j = 1, \ell = 0$ and $j = 0, \ell = 2$ leading to the same energy.

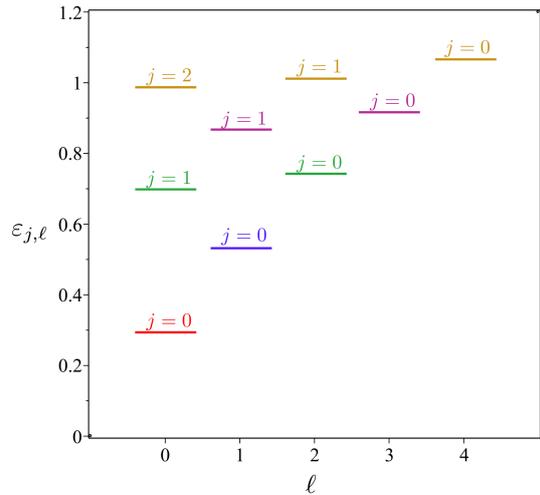


Fig. 7. Scaled energies of a particle with mass μ moving in a three-dimensional potential, leading to a spectrum where the s -state part is given by (13) and the scaling parameter $K = 2$ by (14). Each energy level is characterized by *two* quantum numbers j and ℓ . No principal quantum number can be identified and evidently, no accidental degeneracy is taking place.

solved the radial wave equation (8) and display the lowest energy levels in Fig. 7. At first sight, the scheme resembles that of the harmonic oscillator. However, closer inspection reveals that the levels, which for the harmonic oscillator were degenerate, now differ slightly. We conjecture that higher energy levels behave similarly and that no accidental

degeneracy occurs. We emphasize once more that the $(2\ell + 1)$ -fold *essential* degeneracy with respect to the magnetic quantum number m is caused by the central potential.

References

- [1] G.H. Hardy, E.M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press, Oxford 1979.
- [2] R.L. Rivest, A. Shamir, L. Adleman, *Commun. ACM* **21**, 120 (1978).
- [3] P.W. Shor, in: *Proc. 35th Annual Symposium on Foundations of Computer Science*, IEEE, 1994 p. 124.
- [4] F. Gleisberg, R. Mack, K. Vogel, W.P. Schleich, *New J. Phys.* **15**, 023037 (2013).
- [5] F. Gleisberg, F. Di Pumpo, G. Wolff, W.P. Schleich, *J. Phys. B* **51**, 035009 (2018).
- [6] F. Gleisberg, M. Volpp, W.P. Schleich, *Phys. Lett. A* **379**, 2556 (2015).
- [7] B.M. Garraway, H. Perrin, *J. Phys. B* **49**, 172001 (2016).
- [8] C.J. Pethick, H. Smith, *Bose-Einstein Condensation in Dilute Gases*, Cambridge University Press, Cambridge 2008.
- [9] S.J. Bereta, L. Madeira, V.S. Bagnato, M.A. Caracanhas, *Am. J. Phys.* **87**, 924 (2019).
- [10] R. Grimm, M. Weidemüller, Y.B. Ovchinnikov, *Adv. Atom. Mol. Opt. Phys.* **42**, 95 (2000).
- [11] D. Cassettari, G. Mussardo, A. Trombettoni, *PNAS Nexus* **2**, 1 (2022).
- [12] M.G. Boshier, private communication, October 2022; M.G. Boshier, “Sensing and Signal Processing with trapped BECs”, talk at the conference *Physics of Quantum Electronics (PQE)*, Snowbird (UT) 2023.
- [13] R.A. Carollo, D.C. Aveline, B. Rhyno, S. Vishveshwara, C. Lannert, J.D. Murphree, E.R. Elliott, J.R. Williams, R.J. Thompson, N. Lundblad, *Nature* **606**, 281 (2022).
- [14] T. van Zoest, N. Gaaloul, Y. Singh et al., *Science* **328**, 1540 (2010).
- [15] D. Becker, M.D. Lachmann, S.T. Seidel et al., *Nature* **562**, 391 (2018).
- [16] D.C. Aveline, J.R. Williams, E.R. Elliott, C. Dutenhoffer, J.R. Kellogg, J.M. Kohel, N.E. Lay, K. Oudrhiri, R.F. Shotwell, N. Yu, R.J. Thompson, *Nature* **582**, 193 (2020).

- [17] N. Gaaloul, M. Meister, R. Corgier, A. Pichery, P. Boegel, W. Herr, H. Ahlers, E. Charron, J.R. Williams, R.J. Thompson, W.P. Schleich, E.M. Rasel, N.P. Bigelow, *Nat. Commun.* **13**, 7889 (2022).
- [18] I. Białyński-Birula, M.A. Cirone, J.P. Dahl, M. Fedorov, W.P. Schleich, *Phys. Rev. Lett.* **89**, 060404 (2002).
- [19] I. Białyński-Birula, M.A. Cirone, J.P. Dahl, R.F. O’Connell, W.P. Schleich, *J. Opt. B* **4**, S393 (2002).
- [20] I. Białyński-Birula, M.A. Cirone, J.P. Dahl, T.H. Seligman, F. Straub, W.P. Schleich, *Fortschritte der Phys.* **50**, 599 (2002).
- [21] M.A. Cirone, J.P. Dahl, M. Fedorov, D. Greenberger, W.P. Schleich, *J. Phys. B* **35**, 191 (2001).
- [22] R. Mack, J.P. Dahl, H. Moya-Cessa, W.T. Strunz, R. Walser, W.P. Schleich, *Phys. Rev. A* **82**, 032119 (2010).
- [23] J. Twamley, G. Milburn, *New J. Phys.* **8**, 328 (2006).
- [24] K. Huang, C.N. Yang, *Phys. Rev.* **105**, 767 (1957).
- [25] Y.N. Demkov, V.N. Ostrovskii, *Zero-range Potentials and their Applications in Atomic Physics*, Plenum Press, New York 1988.
- [26] W.P. Schleich, *Quantum Optics in Phase Space*, VCH-Wiley, Weinheim 2001.
- [27] M.O. Scully, M.S. Zubairy, *Quantum Optics*, Cambridge University Press, Cambridge 1997.
- [28] J. Bertrand, *CR Acad. Sci* **77**, 849 (1873).
- [29] L.D. Landau, E.M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, Vol. 3, Pergamon Press, Oxford 1977.
- [30] C. Runge, W. Lenz, *Vector Analysis*, Vol. 2, Dutton, New York 1919.
- [31] W. Lenz, *Z. Phys.* **24**, 197 (1924).
- [32] H. Goldstein, C. Poole, J. Safko, *Classical Mechanics*, Addison-Wesley, New York 2002.
- [33] C. Cohen-Tannoudji, B. Diu, F. Laloë, *Quantum Mechanics*, Wiley Interscience, New York 1977.
- [34] I.B. Khriplovich, *Theoretical Kaleidoscope*, Lecture Notes in Physics, Vol. 747, Springer, Heidelberg 2008.

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

Arrival Time in Quantum Mechanics (Demonstrated in Geometrical Order)

J. KIJOWSKI*

*Center for Theoretical Physics, Polish Academy of Sciences, al. Lotników 32/46, 02-668
Warszawa, Poland*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S123](https://doi.org/10.12693/APhysPolA.143.S123)*e-mail: kijowski@cft.edu.pl

A geometric construction of the arrival time in conventional quantum mechanics is presented. It is based on a careful mathematical analysis of different quantization procedures for classical observables as functions of positions and momenta. A class of observables is selected that possess a unique (if any) quantized version. A simple criterion for the existence of such a quantized version is formulated. These mathematical results are then applied to the classical “arrival time” observable.

topics: arrival times, quantization of time, geometry of quantum mechanics, uncertainty principle

1. Introduction

At the turn of the 1960s and 1970s, many papers were published on the “time problem” in quantum mechanics (see [1]). Numerous authors have complained that in quantum mechanics only three (among four) spacetime coordinates have quantum counterparts in the form of position operators, while the fourth — time — always remains the classical parameter of evolution. This was, according to many, a flagrant violation of the relativistic invariance that should characterize any reasonable physical theory. On the other hand, a comment by Wolfgang Pauli in 1958 [2] (see also paper [3] by G.R. Allcock) clearly indicated that treating energy (the fourth component of “four-momentum”) as momentum canonically conjugate to time and requiring these quantities to satisfy the canonical commutation rules (in order to obtain the energy–time uncertainty principle as a by-product) leads to a contradiction with the positivity of the self-adjoint energy operator.

However, it is obvious that “ x ” *tout court* is not an observable. By measuring this quantity at different instants of time, we obtain different, time-dependent results. What can be measured is “ $x(t)$ ”, i.e., “the position taken by our particle at time t ”. Similarly, the arrival time “ $t(x)$ ”, i.e., “the time it takes for a particle to hit the plane $\{(x, y, z)|x = \text{const}\}$ ”, is a well-defined observable

which, at least classically, can be uniquely defined and measured. For a free particle of mass m whose initial position at time $t = 0$ is (x, y, z) , this quantity is equal to

$$t(x) = -\frac{x}{v_x} = -\frac{m x}{p}, \quad (1)$$

where v_x denotes the particle’s velocity in direction of the x -axis, whereas $p = m v_x$ is the corresponding component of the momentum vector. Indeed, solving the equation of motion

$$0 = x(t) = x + t v_x \quad (2)$$

with respect to time, we obtain (1).

According to the naive “quantization procedure”, the quantum version \hat{f} of this observable should be obtained by replacing the classical position x and the classical momentum p by the position operator \hat{x} and the momentum operator \hat{p}

$$\hat{t}(x) = -m \frac{\hat{x}}{\hat{p}}, \quad (3)$$

or

$$\hat{t}(x) = -\frac{m}{2} \left\{ \hat{x} \frac{1}{\hat{p}} + \frac{1}{\hat{p}} \hat{x} \right\}, \quad (4)$$

which looks at first glance more “hermitian”. Unfortunately, these formulas don’t make any sense. There is no self-adjoint operator that agrees with the above combinations of position and momentum operators, even if restricted to a small (but dense) subspace of quantum states.

The whole “quantization procedure”, i.e., representing classical observables by self-adjoint operators, goes wrong here. This is not surprising because from a physical point of view, classical physics is the limit of quantum physics (in situations where Planck’s constant is so relatively small to what we can measure that it can be considered equal to zero), and not vice versa. That is, the quantum theory unambiguously implies classical theory as an approximation. The universal validity of some “quantization procedure” would mean the opposite, i.e., knowing the classical theory, we would automatically know its quantum version. Such an assumption is nonsense.

However, in this article, it will be proved that there is a class of observables that admit unambiguous quantization. Using these techniques, it will be shown that there is a unique way of constructing a quantum version of the arrival time, i.e., the “arrival time operator”.

This operator was first proposed in the paper [4] (see also [5] and [6]). The present author is much indebted to I. Białynicki-Birula and S.L. Woronowicz for regular discussions concerning the fundamental structures of quantum mechanics that we had at the beginning of ’70s. These discussions were a true inspiration of the author’s analysis of the problem. The construction of the operator presented in [4] was axiomatic, based on the requirement to satisfy several physically well-founded properties. In the present paper, an entirely different construction is presented, based on a mathematical analysis of the uniqueness property of possible quantization procedures. Following the great Baruch Spinoza and his fundamental philosophical treatise *Ethica ordine geometrico demonstrata* [7], we can say that this paper describes the time of arrival “demonstrated in geometrical order”.

The construction proposed in [4] was later commented on and criticized by many authors (see, e.g., [8–11]), but none of them was able to propose another, mathematically self-consistent construction. Nevertheless, the criticism formulated by Bogdan Mielnik and Gabino Torres-Vega in [11] is well motivated from a physical point of view. It is based on the observation that the probability that a particle hits a plane,

$$P_x := \left\{ (t; x, y, z) \mid x = \text{const} \right\}, \quad (5)$$

exactly at the spacetime point $(t; y, z)$ behaves in a “strange way” as a function of x . This strange behaviour is analogous to the phenomenon known as “probability back-flow”, i.e., even if the wave function $\psi_t(x)$ contains only positive momenta, there might be regions where the probability density $|\psi_t(x)|^2$ travels in negative direction as time increases. But that’s what quantum mechanics is! The phenomenon of superposition, which does not exist in classical mechanics, leads inevitably to such behaviour of the probability density.

Recently, the “time problem” has also been intensively discussed (see, e.g., [12–14]). In the author’s opinion, these works do not add anything new to this discussion because they are either mathematically inconsistent^{†1}, or they propose new (even interesting) physical schemes, but going beyond standard quantum mechanics.

The paper is organized as follows. In Sect. 2 we show how to quantize uniquely physical observables belonging to a certain, geometrically well-defined class. Finally, in Sect. 3, we apply these techniques to the analysis of the observable (1) and discuss possible ways to quantize it.

2. Schrödinger versus Heisenberg

The equivalence between Heisenberg’s quantum mechanics and Schrödinger’s wave mechanics is obvious if we substitute the self-adjoint operators \hat{x} and $\hat{p} := i\hbar \frac{\partial}{\partial x}$, acting in the Hilbert space of square-integrable Schrödinger wave functions, for the Heisenberg q -numbers \hat{x} and \hat{p} . However, contrary to the Heisenberg’s intuition, these objects are neither finite-dimensional matrices nor continuous (bounded) operators in the infinite-dimensional Hilbert space. Consequently, even the definition of their commutator

$$[\hat{x}, \hat{p}] := \hat{x}\hat{p} - \hat{p}\hat{x}, \quad (6)$$

is, a priori, meaningless. The necessary sophisticated mathematics was then elaborated by John von Neumann and his followers. In particular, the discussion of “weak” versus “strong” commutation is necessary for the uniqueness of the above Schrödinger representation of the Heisenberg q -numbers (i.e., the uniqueness of the canonical commutation relations). Therefore, when dealing with quantum mechanics, we must remember that the algebra of unbounded operators (e.g., their product) is an extremely subtle topic and — if done without a proper mathematical background — can lead to painful paradoxes.

But physically, without going deeply into this extremely difficult mathematics, the unique representation of position and momentum as self-adjoint operators follows directly from the Schrödinger equation and from the probabilistic interpretation of the wave function. The latter obviously implies the shape of the position operator

$$(\hat{x}\psi)(\xi) := \xi\psi(\xi). \quad (7)$$

^{†1}Treating a spectral measure over a continuous spectrum as a sum over a discrete spectrum, or using “delta-normalized” wave functions, was a nice heuristic way Dirac used in 1928 to illustrate the basic concepts of quantum physics. Applying the same “techniques” in 2020 (i.e., disregarding the 90 years of progress made here in understanding the basic structures of quantum mechanics) to explain “how do we measure time in quantum theory” is unacceptable.

But, what is much less known, also the momentum operator is uniquely implied in wave mechanics, without any reference to Heisenberg's "axiomatics of q -numbers". To prove this statement, let us first consider the statistical ensemble of *classical free particles*, whose state is described by the probability density $\varphi(t; \mathbf{x}, \mathbf{p})$ in the phase space $\mathcal{P} = \{(\mathbf{x}, \mathbf{p})\}$. The corresponding densities ρ in configuration space and μ in momentum space are given as the corresponding "marginals"

$$\begin{aligned}\rho(t; \mathbf{x}) &:= \int d^3p \varphi(t; \mathbf{x}, \mathbf{p}); \\ \mu(t; \mathbf{p}) &:= \int d^3x \varphi(t; \mathbf{x}, \mathbf{p}) = \mu(\mathbf{p}),\end{aligned}\quad (8)$$

(the momentum distribution is obviously time-independent for free particles). Knowing the dynamics of the particles (their free motion)

$$\begin{aligned}\mathbf{x}(t) &= \mathbf{x}(0) + \frac{t}{m} \mathbf{p}(0), \\ \mathbf{p}(t) &= \mathbf{p}(0) = \text{const},\end{aligned}\quad (9)$$

we know the time dependence of the density

$$\varphi(t; \mathbf{x}, \mathbf{p}) = \varphi\left(0; \mathbf{x} - \frac{t}{m} \mathbf{p}, \mathbf{p}\right). \quad (10)$$

It is easy to check that the momentum probability density μ can be uniquely determined via position measurements. Indeed, we have

$$\begin{aligned}\mu(\mathbf{p}) &= \mu(0, \mathbf{p}) = \int_{\mathbb{R}^3} d\mathbf{y} \varphi(0; \mathbf{y}, \mathbf{p}) = \\ &\lim_{t \rightarrow \infty} \int_{\mathbb{R}^3} d\mathbf{y} \varphi\left(0; \mathbf{y}, \mathbf{p} - \frac{m}{t}(\mathbf{y} - \mathbf{x}_0)\right),\end{aligned}\quad (11)$$

where \mathbf{x}_0 is a fixed arbitrary point in configuration space. Using a new variable

$$\begin{aligned}\mathbf{q} &:= \mathbf{p} - \frac{m}{t}(\mathbf{y} - \mathbf{x}_0) \Leftrightarrow \mathbf{y} := \mathbf{x}_0 + \frac{t}{m}(\mathbf{p} - \mathbf{q}); \\ d^3y &= \left(\frac{t}{m}\right)^3 d^3q,\end{aligned}\quad (12)$$

we obtain

$$\begin{aligned}\mu(\mathbf{p}) &= \lim_{t \rightarrow \infty} \frac{t^3}{m^3} \int_{\mathbb{R}^3} d^3q \varphi\left(0; \mathbf{x}_0 + \frac{t}{m}(\mathbf{p} - \mathbf{q}), \mathbf{q}\right) = \\ &\lim_{t \rightarrow \infty} \frac{t^3}{m^3} \int_{\mathbb{R}^3} d^3q \varphi\left(t; \mathbf{x}_0 + \frac{t}{m} \mathbf{p}, \mathbf{q}\right) = \\ &\lim_{t \rightarrow \infty} \frac{t^3}{m^3} \rho\left(t, \mathbf{x}_0 + \frac{t}{m} \mathbf{p}\right).\end{aligned}\quad (13)$$

Thus, by measuring the probability density ρ of a particle in configuration space during its time evolution, we also obtain its probability density μ in momentum space as a result.

According to the Born probabilistic interpretation of the wave function, the quantum analog of the configuration probability density is equal to

$$\rho(t, \mathbf{x}) = \left\| \psi(t, \mathbf{x}) \right\|^2. \quad (14)$$

Taking (13) as the definition of the momentum probability density and using the free Schrödinger equation for the evolution of the wave function over time, a simple calculation is enough to prove that the above definition (14) implies the following textbook formula

$$\mu(\mathbf{p}) := \lim_{t \rightarrow \infty} \frac{t^3}{m^3} \left\| \psi\left(t, \mathbf{x}_0 + \frac{t}{m} \mathbf{p}\right) \right\|^2 = \|\tilde{\psi}(t, \mathbf{p})\|^2, \quad (15)$$

where $\tilde{\psi}$ denotes the Fourier transformation of ψ

$$\tilde{\psi}(t, \mathbf{p}) := \int_{\mathbb{R}^3} \frac{d^3x}{(2\pi\hbar)^{\frac{3}{2}}} \psi(t, \mathbf{x}) \exp\left(-\frac{i \mathbf{p} \cdot \mathbf{x}}{\hbar}\right). \quad (16)$$

Now, (15) immediately implies the form of the momentum operator in the momentum representation and, consequently, also in the position representation

$$(\hat{p}\tilde{\psi})(\mathbf{p}) := p\tilde{\psi}(\mathbf{p}) \iff (\hat{p}\psi)(\mathbf{x}) := \frac{\hbar}{i} \frac{\partial}{\partial \mathbf{x}} \psi(\mathbf{x}), \quad (17)$$

without resorting to an extremely sophisticated version of Heisenberg's axiomatics, where completely non-intuitive "strong commutation relations" between positions and momenta must be assumed a priori. To the author's knowledge, the only textbook on quantum mechanics that derives the momentum operator in this way (physically the most intuitive) and does not postulate it a priori is the excellent book by Białynicki-Cieplak-Kamiński [15].

The advantage of the geometric description of quantum physics based on Schrödinger's wave approach over Heisenberg's algebraic formulation is particularly evident when we try to "quantize" more complex observables of the form $f(x, p)$. In particular, let us consider observables that are linear with respect to momentum

$$f(\mathbf{x}, \mathbf{p}) := X^k(\mathbf{x})p_k, \quad (18)$$

where $X^k(\mathbf{x})$ is an arbitrary vector field on the configuration space. Even if the quantum operators $\hat{X}^k := X^k(\hat{x})$ and \hat{p}_k are already explicitly defined, their product depends on the order of multiplication. Unfortunately, even the symmetric order

$$\hat{f} = \frac{1}{2} \left\{ \hat{X}^k \hat{p}_k + \hat{p}_k \hat{X}^k \right\}, \quad (19)$$

although formally "hermitian", does not guarantee the self-adjointness of the resulting operator.

We are going to propose in the sequel a simple, geometric construction of the self-adjoint operator \hat{f} , together with a simple criterion for its existence. For this purpose, let us observe that the vector field \mathbf{X} generates a one-parameter group \mathcal{G}_t of local diffeomorphisms of the configuration space. These diffeomorphisms can be used to transport (drag) locally any (square-integrable) wave function. Such a transport $\hat{\mathcal{G}}_t$ is a unitary transformation (i.e., it does not "lose" any piece " $dx \|\psi\|^2$ " of the particle's probability) if and only if the transformations are *global*, i.e., the field X is complete. But a group of unitary transformations is always of the form

$$\hat{\mathcal{G}}_t = \exp\left(\frac{it}{\hbar}\hat{f}\right), \quad (20)$$

and, whence, its *self-adjoint* generator \hat{f} is uniquely defined. Mathematically, the generator of the classical transport group is called the Lie derivative with respect to the vector field \mathbf{X} and is denoted by $\mathcal{L}_{\mathbf{X}}$. We have, therefore, the following, unique formula

$$\hat{f} := \frac{\hbar}{i}\mathcal{L}_{\mathbf{X}} = \frac{\hbar}{i}\left(\frac{d}{dt}\hat{\mathcal{G}}_t\right)_{t=0}, \quad (21)$$

which is automatically self-adjoint if the diffeomorphisms \mathcal{G}_t are global. In the very special case of a constant field $\mathbf{X} = \frac{\partial}{\partial x^k}$ the Lie derivative reduces to a partial derivative and, therefore, (21) reproduces the textbook formula (17) for the momentum operator.

In this way we get a nice and practical quantization rule for the observable f together with an easy criterion for its self-adjointness, i.e., for the reasonableness of the whole procedure. Moreover, this criterion is of a topological nature, namely it imposes the existence of global solutions of the dynamical system

$$\frac{dx^k(t)}{dt} = X^k(\mathbf{x}(t)), \quad (22)$$

and has nothing to do with the algebraic complexity of the function $\mathbf{X}(\mathbf{x})$.

We will illustrate this method of “quantization” by taking as an example a 1-dimensional problem

$$f(x, p) := X(x)p. \quad (23)$$

In the simplest case, when $X(x)=1 = \text{const}$, the group \mathcal{G}_t is simply the translation group $\mathcal{G}_t(x) = x + t$. But, locally, we can always find a coordinate $s = s(x)$ such that the field $X(x)$ is constant when expressed in terms of this coordinate, i.e., that the following identity holds

$$X(x)\frac{\partial}{\partial x} = \frac{\partial}{\partial s}. \quad (24)$$

To find such a new coordinate, we therefore need to solve the following differential equation

$$\frac{ds}{dx} = \frac{1}{X(x)}. \quad (25)$$

If the solution is global, then \hat{f} is uniquely defined as the generator of the group of translations

$$\hat{f} := \frac{\hbar}{i}\frac{d}{ds}, \quad (26)$$

acting on the wave functions in the s -representation. To express this operator in the original x -representation, we must remember that the wave function *is not* a scalar but a “half density”. The correct transformation formula between the two representations is implied by the following identity

$$\int dx \|\psi(x)\|^2 = \int \frac{dx}{ds} ds \|\psi(x(s))\|^2 = \int ds \|\psi(x(s))\|^2 X(x(s)). \quad (27)$$

This means that the following transformation $U : L^2(\mathbb{R}) \mapsto L^2(\mathbb{R})$ between the two representations is unitary

$$(U\psi)(s) = \psi(x(s))\sqrt{X(x(s))}. \quad (28)$$

The existence of the unitary operator U , i.e., the global character of the group \mathcal{G}_t , is essential here. It enables us to re-calculate the Lie derivative with respect to $X(x)$ from the x -representation to the s -representation and vice versa.

Geometrically, a substantial simplification of the formulae used below is obtained if we represent the quantum state by the half-density, $\Psi := \psi\sqrt{dx}$, instead of the scalar function ψ . To transport such a quantity along the vector field \mathbf{X} , we must transport not only the scalar factor ψ , but also the half-density factor \sqrt{dx} . We — physicists, we know perfectly how to transport a density, like “ dx ”, but are not used to half-densities. For this reason, we decided to use in this paper the standard textbook notation.

In this notation we have

$$\mathcal{L}_{\mathbf{X}}\Psi = U^{-1} \circ \frac{d}{ds} \circ (U\Psi), \quad (29)$$

and, whence

$$\begin{aligned} (\mathcal{L}_{\mathbf{X}}\psi)(x) &= U^{-1}\frac{d}{ds}\left(\psi(x(s))\sqrt{X(x(s))}\right) = U^{-1}X\frac{d}{dx}\left(\psi(x)\sqrt{X(x)}\right) = \\ &= U^{-1}\left\{\left(X^{\frac{3}{2}}\psi' + \frac{1}{2}\psi X'\sqrt{X}\right)(x(s))\right\} = \left(X\frac{d}{dx} + \frac{1}{2}X'\right)\psi(x) = \frac{1}{2}\left(X\frac{d}{dx}\psi + \frac{d}{dx}(X\psi)\right)(x). \end{aligned} \quad (30)$$

This formula for the Lie derivative, together with (21), reproduces *formally* the naive quantization formula (19). Note, however, that (19) does not capture the definition of the operator \hat{f} ; indeed, this formula makes a priori no sense if the transport group \mathcal{G}_t , generated by \mathbf{X} , is not global, or in other words, if the global unitary transformation U does not exist. We conclude that

a purely algebraic approach to quantization is completely inadequate, since the problem depends entirely on the analytic and topological properties of \mathbf{X} .

Example 1. For $\mathbf{X}(\mathbf{x}) = \mathbf{x}$ the transport group generated by \mathbf{X} is the global homothety group

$$\mathcal{G}_t(\mathbf{x}) = \exp(t)\mathbf{x}, \quad (31)$$

in the 3D case, which reduces to

$$\mathcal{G}_t(x) = \exp(t)x, \quad (32)$$

in the 1D case. Hence, the operator $\hat{f} := \frac{1}{2}(\hat{x}^k \hat{p}_k + \hat{p}_k \hat{x}^k)$ and its 1D analog $\hat{f} := \frac{1}{2}(\hat{x} \hat{p} + \hat{p} \hat{x})$ are essentially self-adjoint. They are entirely described by (21) as generators of the quantum homothety group.

In the 1D case, the same result can be obtained using the variable s , which trivializes the field $X(x)$ according to (24). For this purpose we solve (25), so

$$\frac{ds}{dx} = \frac{1}{x} \Rightarrow s = \ln|x|, \quad (33)$$

and observe that the translations in the s variable are homotheties (32) in the variable x . Note that the Hilbert space $L^2(\mathbb{R})$ of the x -dependent wave functions naturally splits into the direct sum of two subspaces: $L^2(\mathbb{R}_+)$ and $L^2(\mathbb{R}_-)$, describing particles localized entirely within the positive (\mathbb{R}_+) and the negative (\mathbb{R}_-) half-axis, respectively. Each of them is isomorphic with $L^2(\mathbb{R})$ -space of s -dependent wave functions. A homothety (31) is equivalent with a simultaneous shift in the variable s (i.e., $\mathcal{G}_t(s) := s + t$) in both subspaces.

A 3D analog of the above construction is obtained if we use the spherical coordinates (r, ϑ, φ) and put $s = \ln(r)$.

Example 2. Let us consider $X(x) = x^2$. To find the transport group \mathcal{G}_t , we must solve the differential equation

$$\begin{aligned} \frac{dx(t)}{dt} &= x^2(t) \Rightarrow \frac{dx}{x^2} = dt \Rightarrow \\ &\Rightarrow t + c = -\frac{1}{x} \Rightarrow x(t) = -\frac{1}{t + c}, \end{aligned} \quad (34)$$

which describes all trajectories of the field, starting from different points. The point $\mathcal{G}_t(x_0)$ is defined by the initial condition

$$\mathcal{G}_t(x_0) = x_0 \Rightarrow c = -\frac{1}{x_0}. \quad (35)$$

This implies

$$\mathcal{G}_t(x) = -\frac{1}{t - \frac{1}{x}} = \frac{x}{1 - tx}. \quad (36)$$

This is not a global diffeomorphism and, whence, does not define a unitary transformation of wave functions. Let us first consider the case $t > 0$. We see that $\mathcal{G}_t(x)$ is defined for $x < \frac{1}{t}$ only, because it escapes to infinity as x approaches the value $x = \frac{1}{t}$. At the same time, a substantial part of the negative half-axis, namely the half-axis $]-\infty, -\frac{1}{t}[$, is not covered at all, because of the inequality $1 - tx > -tx$, which implies immediately

$$\mathcal{G}_t(x) = \frac{x}{1 - tx} > -\frac{1}{t}. \quad (37)$$

Consequently, (21) does not define any self-adjoint operator, even if purely local considerations lead to (30). This proves that the algebraically defined operator (19) is not essentially self-adjoint and, therefore, does not represent any physical observable.

Physically, the above phenomenon means that: (i) we lose a part of probability carried by the wave function $\psi(x)$ for $x > \frac{1}{t}$, and (ii) an information gap is created concerning a part of probability described by the transported wave function $\psi(x(t))$ for $x(t) < -\frac{1}{t}$.

Similarly, for $t < 0$, (i) we lose a part of probability carried by the wave function $\psi(x)$ for $x < -\frac{1}{t}$, and (ii) an information gap is created concerning a part of probability described by the transported wave function $\psi(x(t))$ for $x(t) > -\frac{1}{t}$.

Since both parts (i) information loss and (ii) information gap fit together perfectly, we can use the first one to plug the second one. Mathematically, this means that we can treat the transformation (36) as a global measurable isomorphism of the real line \mathbb{R} . Its singularity at the single point $x = \frac{1}{t}$ does not produce any problem (the transformation, even if non-continuous, is still measurable and invertible). When used to transport wave functions, it defines a continuous group of unitary transformations. Its generator (21) is therefore a self-adjoint extension of naively defined operator (19).

Physically, however, the original disease of (19) has not been cured, because the above ‘‘plugging procedure’’ is not unique. Indeed, the information loss can be plugged into the information gap with the arbitrary constant phase change ‘‘ $\exp(i\varphi)$ ’’, which leads to another self-adjoint extension. This means that in this case, the operator (19) has many inequivalent self-adjoint extensions and, physically, this formula is meaningless.

Example 3. For $X(x) = x^3$, we obtain

$$\mathcal{G}_t(x) = \frac{x}{\sqrt{1 - tx^2}}, \quad (38)$$

which, when used to transport wave functions, would imply that there is an information loss for $|x| > 1/\sqrt{t}$ and no information gap to plug it into. There is no way to repair this disease, and we conclude that (19) does not define any physical observable in this case.

To conclude these technical remarks about quantization, we stress that the canonical transformation

$$(x, p) \mapsto (p, -x), \quad (39)$$

enables us to similarly quantize functions that are linear in the position variable $f(\mathbf{x}, \mathbf{p}) := X_k(\mathbf{p}) x^k$. Indeed, the quantity \mathbf{X} defines a vector field on the space of momenta and can, therefore, be used to transport wave functions in momentum representation. Hence, the whole construction presented above applies here.

3. Arrival time

The arrival time (1) can thus be considered as a vector field on the space of momenta

$$X := X(p) \frac{\partial}{\partial p} \quad \text{where} \quad X(p) = \frac{m}{p} \quad (40)$$

(remember that the momentum canonically conjugate to p is equal to $-x$, not x). This vector field can be used to transport wave functions in the momentum representation. To find the result of such a transport let us “straighten” this field similarly as in (41), i.e., find a new variable $s = s(p)$ such that the field is a constant with respect to this variable

$$X(p) \frac{\partial}{\partial p} = \frac{\partial}{\partial s}. \quad (41)$$

To find such a new coordinate, we therefore need to solve the following differential equation

$$\frac{ds}{dp} = \frac{1}{X(p)} = \frac{p}{m} \implies s(p) = \frac{p^2}{2m} = E_{\text{kinetic}}, \quad (42)$$

(the possible additive constant is irrelevant here). Unfortunately, it is not a global coordinate on the real axis representing all possible values of momentum p . The two half-axes \mathbb{R}_+ and \mathbb{R}_- in the momentum representation are covered by two identical copies of the positive half-axis \mathbb{R}_+ in the “energy representation” (or the s -representation). The field (41) acts independently on both half-axes and is not complete on both of them. There is no way to cure this disease. The corresponding “momentum operator on a half-line” defined as $\frac{\hbar}{i} \frac{\partial}{\partial s}$ has no self-adjoint extension. In other words, the algebraic formulae (3) and (4) do not define anything, which could define a reasonable physical observable. This statement can be treated as an independent proof of W. Pauli’s statement [2] that there is no quantum observable corresponding to the classical arrival time (1).

In the paper [4], the present author have proposed to replace t with another observable, namely

$$T = -\frac{mx}{|p|} = \text{sgn}(p)t, \quad (43)$$

where the symbol “ $\text{sgn}(p)$ ” represents the “sign function”, which takes value $+1$ for $p > 0$ and -1 for $p < 0$. This observable can be called the “oriented arrival time”; it reproduces arrival time for “right movers” and “minus arrival time” for “left movers”. The corresponding vector field to quantize is now

$$X := X(p) \frac{\partial}{\partial p}, \quad \text{where} \quad X(p) = \frac{m}{|p|}. \quad (44)$$

Consequently, (44) is replaced by

$$\frac{ds}{dp} = \frac{|p|}{m}, \quad (45)$$

and the corresponding variable $s(p)$

$$s = \text{sgn}(p) E_{\text{kinetic}} = \begin{cases} \frac{p^2}{2m} & \text{for } p > 0, \\ -\frac{p^2}{2m} & \text{for } p < 0, \end{cases} \quad (46)$$

is global. Hence, the unitary operator (28) does exist. Consequently, the field $X = \frac{\partial}{\partial s}$ is perfectly complete and uniquely defines the self-adjoint operator \hat{T} as the generator (21) of the translation group in the variable s . The formula (29) enables for the transition from the “oriented energy” or s -representation to the momentum or p -representation.

According to (28), quantum states are described in s -representation by the following wave functions

$$\begin{aligned} \tilde{\phi}(s) &:= (U\tilde{\psi})(s) = \\ &\tilde{\psi} \left(\text{sgn}(s)\sqrt{2m|s|} \right) \sqrt{\frac{m}{\sqrt{2m|s|}}} = \\ &\tilde{\psi} \left(\text{sgn}(s)\sqrt{2m|s|} \right) \sqrt[4]{\frac{m}{2|s|}}, \end{aligned} \quad (47)$$

where $\tilde{\psi}(p)$ is the standard wave function in the momentum representation. Moreover, according to (46), the variable p was replaced by

$$p = \text{sgn}(s)\sqrt{2m|s|}. \quad (48)$$

The transformation $\tilde{\psi} \mapsto \tilde{\phi}$ is indeed unitary because we have

$$\begin{aligned} |\tilde{\phi}(s)|^2 ds &= |\tilde{\psi}(p)|^2 \sqrt{\frac{m}{2|s|}} ds = \\ &|\tilde{\psi}(p)|^2 \sqrt{\frac{m^2}{p^2}} ds = |\tilde{\psi}(p)|^2 dp. \end{aligned} \quad (49)$$

According to (26), the \hat{T} operator in the s -representation is defined as

$$\hat{T} := \frac{\hbar}{i} \frac{d}{ds}. \quad (50)$$

This means that the (inverse) Fourier transformation $\phi(T)$ of the function $\tilde{\phi}$,

$$\phi(t, T) := \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} \int_{\mathbb{R}^3} ds \tilde{\phi}(t, s) e^{i s T/\hbar}, \quad (51)$$

describes the spectral resolution of this operator. Physically, this means that the probability that the measurement of the observable \hat{T} gives a result $T \in [a, b] \subset \mathbb{R}$ is equal to

$$P(T \in [a, b]) = \int_a^b dT |\phi(T)|^2. \quad (52)$$

If the particle beam contains a priori only “right-movers”, without any contribution from “left-movers”, then both arrival times (oriented and non-oriented) can be identified ($T = t$). Hence, the above probability density correctly describes the arrival time and properly implements Allcock’s idea regarding time measurements in quantum mechanics. Moreover, the Schödinger evolution of the wave function is especially simple in this representation because it is given by the time translation $T \rightarrow T + t$.

Also for a beam containing “left movers” exclusively, the value of (52) has a clear physical interpretation, i.e., the probability that the measured arrival time will belong to the interval $[-b, -a]$, and so the physical arrival time coincides with $-T$.

For an arbitrary wave function, we can always decompose the quantum state ψ into the superposition $\psi = \psi_+ + \psi_-$, where ψ_+ represents the right moving component and ψ_- represents the left moving component. In the momentum representation, this decomposition is obvious

$$\tilde{\psi}_+(p) = \begin{cases} \tilde{\psi}(p) & \text{for } p > 0, \\ 0 & \text{for } p < 0, \end{cases} \quad (53)$$

$$\tilde{\psi}_-(p) = \begin{cases} 0 & \text{for } p > 0, \\ \tilde{\psi}(p) & \text{for } p < 0, \end{cases} \quad (54)$$

Consequently, we have the corresponding decomposition in s -representation

$$\tilde{\phi}(s) = \tilde{\phi}_+(s) + \tilde{\phi}_-(s). \quad (55)$$

Now, the densities $|\phi_+(T)|^2 dT$ and $|\phi_-(T)|^2 dT$ represent the probability density of arrival time for right-movers and left-movers, separately.

Under Schrödinger evolution, the component $\tilde{\phi}_+$ travels forward in time, whereas $\tilde{\phi}_-$ travels backward in time, and so do both probability densities $|\phi_+|^2$ and $|\phi_-|^2$. Unfortunately, there is no “superselection rule” between both components (right movers and right movers) of the particle beam and, whence, the total density

$$|\phi(T)|^2 = |\phi_+(T)|^2 + |\phi_-(T)|^2 + 2 \operatorname{Re} \left[\phi_+(T) \overline{\phi_-(T)} \right] \quad (56)$$

contains also the last term describing the quantum interaction between the two beams. In other words, the total probability that the particle hits the surface $x = \text{const}$ from the left and the total probability that the particle hits the surface $x = \text{const}$ from the right do not sum up to one

$$\int_{-\infty}^{+\infty} dT |\phi_+(T)|^2 + \int_{-\infty}^{+\infty} dT |\phi_-(T)|^2 \neq \int_{-\infty}^{+\infty} dT |\phi(T)|^2 = 1. \quad (57)$$

This is because, when measuring time of arrival, there are events that do not belong to either the first category (right-movers) or the second category (left-movers).

A complete 3D description of the arrival time requires also the remaining 2 coordinates, (y, z) , as independent variables of the function ψ . This way both ψ_+ and ψ_- are functions of the four variables $(x; t, y, z)$. However, the quantum interpretation applies only to the last three variables, whereas x remains a purely classical parameter numbering different 3D hypersurfaces $\{(t; x, y, z) | x = \text{const}\}$ in 4D spacetime (similarly, as t remains a purely classical parameter of the wave function $\psi = \psi(t; x, y, z)$ in position representation).

4. Conclusions

The result presented in paper [4], and then simplified slightly in [5], was obtained in an axiomatic way. The probability density (52) was derived as a unique quantity satisfying several physically motivated axioms. Such a derivation is similar to the

construction of the Newton–Wigner position operator (see [16]) in relativistic quantum mechanics, where the “up-movers” (i.e., particles) and the “down-movers” (i.e., antiparticles) were also treated separately (see also [17]) and every quantum state can be understood as a superposition of two components.

The author emphasizes that the techniques used here are based on the geometrical interpretation of the wave function as a half-density defined in the configuration space of the particle. Such an interpretation follows directly from Schrödinger’s formulation of “wave mechanics”. This formulation also contains the possibility of giving meaning to Heisenberg’s purely algebraic formulation, which, contrary to popular creeds, is not equivalent to the former. Indeed, in order to make sense of Heisenberg’s formulation, one must first answer two questions: (1) What are those “ q -numbers”? (correct answer: “non-bounded operators in a Hilbert space”), and then (2) How the commutator of non-bounded operators is defined? (correct answer: “in the so-called strong sense”). Without these two steps — highly non-intuitive from the point of view of physics — the entire Heisenberg axiomatics does not make sense, and its computational possibilities do not extend beyond the (linear!) harmonic oscillator.

As a mathematical curiosity, it is worthwhile to notice that the observables “at most linear in p ” and “at most linear in x ” span (in a certain, mathematically well-defined sense) the space of all observables $f(x, p)$. Quantization of f based on its approximation by functions belonging to those two categories for which the quantization rule is unique implies the unique quantization rule for f . It turns out that this rule coincides with the classical Weyl rule (see, e.g., [18]).

Acknowledgments

I dedicate this work to Professor Iwo Białynicki-Birula with thanks for everything I have learned from him.

This research was partially supported by Narodowe Centrum Nauki (Poland) under Grant No. 2016/21/B/ST1/00940.

References

- [1] Y. Aharonov, D. Bohm, *Phys. Rev.* **122**, 1649 (1961).
- [2] W. Pauli, *Die allgemeinen Prinzipien der Wellenmechanik* in: *Encyclopedia of Physics*, Vol. 5/1, Ed. S. Flugge, Springer, Berlin 1958, (see footnote on p. 60).
- [3] G.R. Allcock, *Ann. Phys.* **53**, 253 (1969), see also p. 286 and p. 311.
- [4] J. Kijowski, *Rep. Math. Phys.* **6**, 361 (1974).

- [5] J. Kijowski, *Phys. Rev. A* **59**, 897 (1999).
- [6] J.G. Muga, C.R. Leavens, *Phys. Rep.* **338**, 4 (2000).
- [7] B. Spinoza, *Ethica, Ordine Geometrico Demonstrata*, 1677.
- [8] N. Grot, C. Rovelli, R.S. Tate, *Phys. Rev. A* **54**, 4676 (1996).
- [9] V. Delgado, J.G. Muga, *Phys. Rev. A* **56**, 3425 (1997).
- [10] R. Giannitrapani, *Int. Journ. Theor. Phys.* **36**, 1601 (1997).
- [11] B. Mielnik, G. Torres-Vega, *Concepts Phys.* **II**, 81 (2005).
- [12] L. Maccone, K. Sacha, *Phys. Rev. Lett.* **124**, 110402 (2020).
- [13] R. Gambini, J. Pullin, *New J. Phys.* **24** 053011 (2022).
- [14] T. Jurić, H. Nikolić, *Eur. Phys. J. Plus* **137**, 631 (2022).
- [15] I. Białynicki-Birula, M. Cieplak, J. Kamiński, *Theory of Quanta*, Oxford University Press (1992); Original version in Polish: *Teoria Kwantów. Mechanika Falowa*, PWN, Warszawa (1991).
- [16] T.D. Newton, E.P. Wigner, *Rev. Mod. Phys.* **21**, 400 (1949).
- [17] J. Kijowski, G. Rudolph, *Bull. Acad. Pol. Sci. (Math. Phys. Astr.)* **24**, 1041 (1976).
- [18] H. Weyl, *Z. Phys.* **46**, 1 (1927).

On Repeated Measurements of a Quantum Particle in a Harmonic Potential

F. GAMPEL* AND M. GAJDA

Institute of Physics Polish Academy of Sciences, al. Lotnikow 32/46, 02-668 Warszawa, Poland

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S131](https://doi.org/10.12693/APhysPolA.143.S131)

*e-mail: gampel@ifpan.edu.pl

We study the evolution of a quantum particle in a harmonic potential whose position and momentum are repeatedly monitored. A back-action of measuring devices is accounted for. Our model utilizes a generalized measurement corresponding to the positive operator-valued measure. We assume that upon measurement, the particle's wavefunction is projected onto one of the possible detector states depending on the observed result. We chose these post-measurement states to be moving Gaussian wavepackets. The wavefunction quantum Monte Carlo formalism is used to simulate single quantum trajectories of the particle. We show how classical trajectories emerge in the course of observation and study in detail the dispersion of position and momentum of the particle.

topics: open quantum systems, repeated measurement, wavefunction quantum Monte Carlo

1. Introduction

Position and momentum are fundamental quantities characterizing the dynamics of a classical particle. The time-dependent position of a particle is directly related to what an observer sees while monitoring its motion. The concept is thus very intuitive. According to classical mechanics, measurements, in principle, do not affect the system, and their precision can be arbitrarily high. In contrast, quantum mechanical measurements always somehow affect the system, and moreover, the relation between a wavefunction (or density operator) describing the state of a system with what is actually being observed is not so obvious.

The first approach to resolving these issues is known as the Copenhagen interpretation [1–4], which, until today, forms the basis for the textbook version of quantum mechanics. A central role is played by the Born rule, which gives probabilities of positive answers to yes/no questions related to measurement outcomes. When a measurement is completed, an answer is obtained, and the wavefunction changes discontinuously, in accordance with the result and the von Neumann (and Lüders) postulate of wavepacket reduction [3, 5]. P. Langevin expressed this rule in the introduction to the textbook “La théorie de l’observation en mécanique quantique” [6] by F. London and E. Bauer, in the following words: “*The wave function it [the quantum theory] uses to describe the object no longer depends solely on the object, as was the case in the classical representation, but, above all, states what the*

observer knows and what, in consequence, are his possibilities for predictions about the evolution of the object. For a given object, this function, consequently, is modified in accordance to the information possessed by the observer.”

The Copenhagen interpretation gives a well-defined prescription on how to use the theory in practice. However, it is not the only existing interpretation of concepts such as wavefunction and measurements. After decades, the issue of collapse, nonlocality, and measurement still remains a subject of scientific discussion [7–12].

Iwo Białynicki-Birula and Zofia Białynicka-Birula (Z-IBB) identify in their textbook “Quantum electrodynamics” [13] the fundamental postulates of quantum theory pertaining to the relation between the density operator and measurement. The postulates are very formally combined into four axioms which can be, under some simplifications, summarized as follows: (i) the elementary questions, i.e., the yes/no questions, are represented by projectors, P ; (ii) the state of a system is represented by a non-negative, self-adjoint, and trace-one density operator ρ ; (iii) the density operator determines probabilities p of affirmative answers to elementary questions in accordance with the Born rule $p = \text{Tr}\{P\rho\}$; (iv) every dynamical variable \mathcal{A} is represented by a self-adjoint operator A , and can be assigned a spectral family of projectors, $E_\lambda^{(A)}$, symbolizing questions whether the value of a dynamical variable \mathcal{A} is not larger than λ . As mentioned by Z-IBB, “*Since the set of probabilities p is the only information in quantum theory available*

about the state of the system, from the operational point of view the concept of the state of the system should be identified with the function $p(\mathcal{P})$ defined on the set of all questions.”

The collapse postulate is missing from the Z-IBB axioms. One might possibly find it in the statement quoted above, equating the state of the system to the function $p(\mathcal{P})$. The answer to any question apparently modifies it. On the other hand, the issue might have seemed purely academic at the time, since realistic measurements in quantum mechanics were generally believed to be destructive. This excludes the possibility of repeated measurement on the same quantum system and limits the relevance of the collapse postulate. E. Schrödinger [14], one of the founding fathers of quantum mechanics, wrote: “*We never experiment with just one electron, or atom, or (small) molecule, we sometimes assume that we do; this inevitably entails ridiculous consequences. . . . In the first place, it is fair to say that we are not experimenting with single particles, any more than we can raise ichtyhsauria in the zoo.*” Nowadays, such measurements are not only theoretically considered, but also performed in labs.

Modern variations of the Copenhagen interpretation, such as QBism — quantum Bayesianism [15–18], postulate that an agent (e.g., a physicist) observing a system abruptly modifies their knowledge (the set of probabilities) once a measurement outcome becomes available. The wavefunction expresses the individual agent’s state of knowledge. “. . . *There is no real state of a physical system. What one chooses to regard as the physical system and what state one chooses to assign to it depend on the judgment of the particular physicist who questions the system and who uses quantum mechanics to calculate the probabilities of the answers,*” as stated by N. David Mermin [19]. According to Qbists, since there is no objective wavefunction of the system, there is no collapse either.

Other points of view assume that the wavefunction, the state of the system, has attributes of reality, being independent of an observer. The issue of an apparent collapse, disliked by many physicists, is resolved in various ways. Everett’s many-world interpretation (MWI) is one such approach [20]. Non-local hidden variable theories, of which Bohmian mechanics is the best-known example, are another possibility [21]. The MWI postulates that upon measurement, the system, which finds itself entangled with the measuring apparatus, does not collapse to some observed state, but rather that all components of the wavefunction associated with possible measurement outcomes continue to evolve according to the Schrödinger equation of the composite system. Because of the linearity of the Schrödinger equation, these components do not interact and form separate “branches” or “worlds.” One must accept uncountable copies of themselves and the world living different lives. Bohmian mechanics introduces additional hidden

variables — coordinates (e.g., particle positions) associated with a configuration of the system under consideration. The particles move guided by a “pilot wave”, which is equivalent to the wavefunction of orthodox quantum mechanics (QM) and evolves according to the Schrödinger equation. It is thus the hidden variables that are actually observed in a measurement. Each of the varying interpretations of QM — of which we only mentioned a few — forces us to accept some non-intuitive, seemingly problematic postulate about reality. If none of them is found satisfactory, one must accept the view that collapse — an abrupt, discontinuous change of the system, triggered by measurement — is a “real and wild” thing.

None of the interpretations presented above may be falsified on the grounds of present knowledge. The problems, at this stage of understanding, seem to be of a philosophical nature, and their experimental verification is elusive. However, the various interpretations may imply measurable effects in future experiments and lead to different generalizations of quantum theory.

The first studies of repeated measurement of continuous variables can be found in the works of Mensky [22] and Davies [23, 24]. Great experimental progress in cooling and trapping single ions opened many possibilities for repeated measurement of a single quantum system. The first spectacular example is an observation of quantum jumps, i.e., dark periods in the fluorescence spectrum of an optically driven trapped ion [25–28]. The experiments fueled the interest in the theory of repeated quantum measurements. The proper description of a system under repeated measurements calls for the inclusion of information gained — the back-action of the meters — as part of the dynamics. Different methods were developed [29–35].

The theoretical approach utilizes an open system formalism. It is based on the Gorini–Sudarshan–Kossakowski–Lindblad (GKSL) equation for the density operator [36, 37]. This allows for studying all *statistical* properties of the system. Instead of solving the GKSL equation directly, for different reasons, it may be preferable to look for single realizations of wavefunction dynamics. Obviously, such individual trajectories are stochastic in nature. Averaged over many realizations, they provide a description equivalent to the time-dependent density operator. The general theoretical framework governing wavefunction dynamics of this kind involves the introduction of the so-called stochastic Schrödinger equation (SSE) [38–42]. It should be noted that the choice of an SSE is not unique, and in general, there are many realizations (“unravelings”) corresponding to one GKSL equation. In fact, the formalism of SSE need not be invoked at all for the construction of concrete numerical schemes generating the stochastic trajectories. One notable example [43, 44] is known as the wavefunction quantum Monte Carlo (WFQMC) method. This formalism is

often used by atomic physicists since it allows them to easily generate sequences of events mimicking experiments with atoms and photons.

In this paper, we use WFQMC to analyze statistical characteristics of trajectories determined by simultaneous repeated measurement of position and momentum of a quantum particle. First, we specify our model, define jump operators, and introduce the WFQMC approach. Then we present exemplary trajectories and discuss the time dependence of dispersion of the position and momentum for different choices of detection parameters. Conclusions are presented in the final section.

2. Monte Carlo dynamics of a wavefunction

We study a phase space trajectory of a quantum particle, continuously monitored by an array of detectors. Here we use the theoretical model introduced by us in [45]. We assume that every measurement provides a value of the position and momentum of the particle at this instant. A sequence of such readouts gives a phase space trajectory. Each simultaneous measurement of position and momentum satisfies Heisenberg's uncertainty principle. We apply an open system formalism — our system is a quantum particle described by the Hamiltonian H_0 , while the detectors form a reservoir. We assume that the reservoir has no memory.

The problem of simultaneous measurement of position and momentum for the first time was considered by E. Arthurs and J.L. Kelly [46]. The recent studies of A.J. Scott and G.J. Milbourn [47] assumed a different detection model than the one studied here. They assumed the von Neumann type of coupling between a particle and a meter and used a formalism based on an Ito stochastic Schrödinger equation [31–34]. The main difference is, thus, in the form of the jump operators assumed here.

The effect of coupling the system to the reservoir of detectors is described by the “jump operators” $C_{i,j}$ specified in the following part of the paper. The general form of a completely positive and trace-preserving map which describes time-homogeneous dynamics of the density operator ρ of a system coupled to the Markovian reservoir via operators $C_{i,j}$ is given by the Gorini–Kossakowski–Sudarshan–Lindblad equation [36, 37]

$$\dot{\rho} = i [\rho, H_0] + \mathcal{L}_{relax}(\rho), \quad (1)$$

where H_0 is the self-adjoint Hamiltonian of the system, and \mathcal{L}_{relax} is a relaxation operator of the Lindblad form, accounting for an effect of the environment

$$\begin{aligned} \mathcal{L}_{relax}(\rho) = & -\frac{1}{2} \sum_{\alpha} \left(C_{i,j}^{\dagger} C_{i,j} \rho + \rho C_{i,j}^{\dagger} C_{i,j} \right) \\ & + \sum_{\alpha} C_{i,j} \rho C_{i,j}^{\dagger}. \end{aligned} \quad (2)$$

We chose $C_{i,j}$ to be proportional to projectors onto detector's states $|\alpha_{i,j}\rangle$,

$$C_{i,j} = \sqrt{\gamma} |\alpha_{i,j}\rangle \langle \alpha_{i,j}|, \quad (3)$$

where γ gives the characteristic clicking rate (probability per unit time) and $|\alpha_{i,j}\rangle$ are complex Gaussian wavepackets, which in position representations have the form

$$\langle x | \alpha_{i,j} \rangle = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-x_i)^2/4\sigma^2} e^{ik_j x}. \quad (4)$$

Spatial points x_i and momenta $\hbar k_j$ define the positions of the detectors in phase space. These locations are a matter of choice. Here we assume that they form a rectangular lattice with spacing d_x and d_p , respectively.

In what follows, we will use the index α as a shortcut notation for two indices, $\alpha \equiv (i, j)$ and $C_{\alpha} \equiv C_{i,j}$. The operators C_{α} are responsible for a reduction of the particle's wavefunction, a jump, caused by the interaction with the reservoir. Note that C_{α} projects onto non-orthogonal states, thus $C_{\alpha} C_{\beta} \neq 0$ for $\alpha \neq \beta$. Therefore, the measurement we defined does not belong to the class of a projective-valued measure (PVM). This is in accordance with the modern formulation of a measurement process, which extends the concept of measurements to account for real observations, whose results also depend on the characteristics of the measuring apparatus and procedure. For details on this positive operator-valued measure (POVM), see [24]. Projectors are substituted by an arbitrary number of positive operators — the effects' E_i — whose sum gives identity $\sum_i E_i = I$ [8, 10, 12, 38]. In the case studied here, the effects are related to a jump within a time interval dt caused by $E_{\alpha} = dt C_{\alpha}^{\dagger} C_{\alpha}$, or alternatively, a no-jump event, $E_0 = 1 - \sum_{\alpha} E_{\alpha}$. To assure that all effect operators are positive, the time step dt must be sufficiently small. We take care of this fact.

Instead of solving the GKSL equation, in the following, we use one of its possible unravelings, the quantum Monte Carlo wavefunction method [43, 44]. The idea of the approach is to generate an ensemble of individual trajectories. Each one can be viewed as a single, possible realization of the dynamics of the wavefunction. Averaging over many such trajectories yields the time dependence of the density operator in accordance with the GKSL equation $\rho(t) = |\psi\rangle\langle\psi|$. The WFQMC method simulates stochastic evolution, in which for each time step, the quantity $|\phi'(t + \delta t)\rangle$ is calculated by evolving the state for an infinitesimal time δt with the non-unitary Hamiltonian

$$H = H_0 - \frac{i}{2} \sum_{\alpha} C_{\alpha}^{\dagger} C_{\alpha}. \quad (5)$$

One of two possibilities is then selected, i.e., a jump or no-jump event. The jump to the state $|\alpha\rangle$ is selected with the probability

$$\delta p_{\alpha} = \delta t \langle \phi'(t) | C_{\alpha}^{\dagger} C_{\alpha} | \phi'(t) \rangle = \gamma \delta t |\langle \alpha | \phi'(t) \rangle|^2. \quad (6)$$

The time-step δt has to be sufficiently small to assure that $\sum_{\alpha} \delta p_{\alpha}$ is smaller than one. If the jump takes place, the particle's wavefunction changes discontinuously

$$|\phi(t + \delta t)\rangle = |\alpha\rangle. \quad (7)$$

The probability of no jump is equal to

$$P_0 = 1 - \sum_{\alpha} \delta p_{\alpha}. \quad (8)$$

If the “no-jump” event takes place, the state is essentially replaced with $|\phi'(t + \delta t)\rangle$. However, since the Hamiltonian (5) does not preserve the norm, the state is first normalized

$$|\phi(t + \delta t)\rangle = \frac{(1 - iH\delta t)|\phi(t)\rangle}{\|(1 - iH\delta t)|\phi(t)\rangle\|}. \quad (9)$$

The evolution of the wavepacket corresponds thus to a random sequence of jump and no-jump events. In our approach, every jump is interpreted as an act of measurement. Projection onto states associated with detectors is reminiscent of the reduction of a wavepacket. The non-unitary evolution accounts for the Hamiltonian dynamics of the particle as well as for interaction with the detectors. The Hermitian part H_0 is the sum of kinetic and potential energy $H_0 = \frac{1}{2m}p^2 + V(x)$. Interaction with the detectors is represented by the non-Hermitian term $\frac{1}{2}C_{\alpha}^{\dagger}C_{\alpha}$. This term causes a kind of “accumulation” of the wavefunction around the detector positions in phase space [45]. In each timestep δt , every detector contributes to the particle wavefunction, ϕ , by an amount proportional to $\propto -\frac{1}{2}\gamma\delta t\langle x|\alpha\rangle\langle\alpha|\phi\rangle$.

Our choice of jump operators C_{α} fulfills a number of basic assumptions about a sensible detector of position and momentum. First of all, a meter of position should click if the probability of finding a particle in its neighborhood is large. In our case, this probability is proportional to the squared overlap of the wavefunction with the state associated with the detector. Once the meter “fires”, the particle wavefunction should be reduced according to the information gained, so the post-measured state is localized around the position of the detector. Choosing the detector states to be Gaussians stands to reason. The width σ_x gives the precision of the measurement.

We would like our detectors to be “gentle” to the objects under measurement. By this, we do not mean a weak measurement, but we want the particle velocity, assumed to be proportional to the probability density current, to be not significantly affected due to detection. The post-measurement state of the particle should preserve some information about its pre-measurement momentum at the detection point. To this end, we equip the detectors at every spatial location with a variety of kinetic momenta by assigning to every Gaussian spatial profile plane-waves of momenta $\hbar k_n$. The momenta can take various values, as discussed above. The probability of clicking is thus maximal if both the position and momentum of the particle fit one

of the detector states. This conclusion is obvious if one considers the detector's wavefunction not in position but in momentum space. The Fourier transform of a detector state (see (4)) is

$$\langle k|\alpha_{mn}\rangle = \sqrt{\frac{2\sigma^2}{\pi}} e^{-\sigma^2(k-k_n)^2 + ix_m(k-k_n)}, \quad (10)$$

a Gaussian superposition of plane-waves of momenta centered around k_n . The detector is very sensitive to wavefunctions whose local velocity at x_m is close to k_n .

3. Statistical characterization of particle's trajectories

In our work, we study the “trajectories” of a particle resulting from the detection process, i.e., sequences of position and momentum measurements of the particle in the harmonic potential $V = \frac{1}{2}m\omega^2x^2$. We use the harmonic oscillator units, i.e., the unit of length $a_{ho} = \sqrt{\hbar/(m\omega)}$, the unit of momentum $q_0 = \hbar/a_{ho}$, the unit of time $\tau_0 = 1/\omega$, and the unit of energy $\varepsilon_0 = \hbar\omega$. From now on, all quantities are expressed in these units. The Hermitian Hamiltonian has the form

$$H_0 = \frac{1}{2}x^2 + \frac{1}{2}p^2, \quad (11)$$

placing position and momentum on equal footing. We assume that the detectors are characterized by a spatial width $\sigma = \sqrt{1/2}$. Selecting this value ensures that the detectors formally have the same width in momentum space. Moreover, as the detectors project into coherent states, the post-measurement uncertainty in position and momentum is minimal according to Heisenberg's principle. Similarly to the Hermitian part of the Hamiltonian, the coupling to detectors is symmetric, and position and momentum are on an equal footing. We choose the same numerical value for the detector spacing $d_x = d_p = d$.

In our calculations, we impose the initial wavefunction of the particle to be identical to one of the Gaussian detector states (4), centered at $(x_0 = nd, p_0 = 0)$. Here n is a natural number, chosen so that x_0 is close to 20, so for different grid densities d we get comparable initial conditions. The particle thus starts with zero velocity at some distance from the minimum of the potential. This distance defines the classical amplitude of a harmonic oscillation and comprises several other detectors ($n \gg 1$) so that subsequent motion can be monitored with sufficient resolution. In our numerical experiment, we simulate a large number of trajectories, where by trajectory we mean a time series of detection events, “clicks” of meters at phase space locations (x_i, p_i) at instants t_i . An example of a single realization of a measurement experiment is shown in Fig. 1. The particle follows a circular orbit in phase space, as would be expected for a classical particle. Some random departures from this orbit are clearly visible.

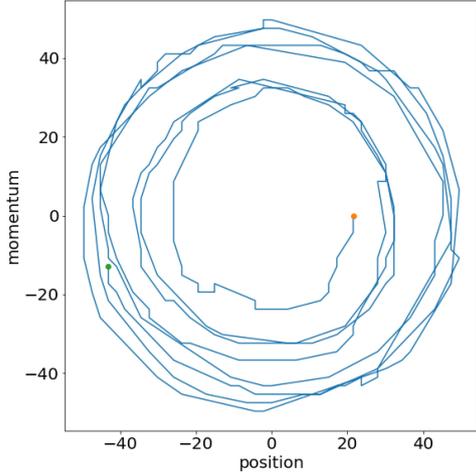


Fig. 1. Sample trajectory in phase space. The start point is marked by an orange point, while the end of the simulation is visualized as a green point. A typical trajectory in our setting is always a circular motion with a growing radius. Only isolated phase space positions of a particle are available to the observer. The line is drawn to guide the eye. Detector spacing is $d = 2.16$ and $\gamma = 1$.

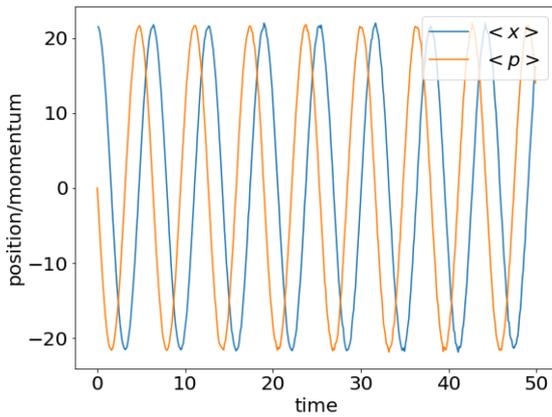


Fig. 2. Average position and momentum of a particle in the harmonic potential as a function of time. The dependence is equivalent to the classical solution of the harmonic oscillator, giving a sinusoidal motion with the respective frequency and a phase shift of $\pi/2$ between position and momentum. Detector spacing is $d = 2.16$ and $\gamma = 1$.

Moreover, the radius of the orbit grows slowly in time, i.e., the energy of the observed particle increases.

Individual trajectories, resulting from a stochastic process, differ from one another. Their statistical properties are the main objects of our interest. First, we analyze the average phase space trajectory ($\langle x(t) \rangle, \langle p(t) \rangle$). Using the WFQMC formalism, we generate 5000 trajectories for each choice of parameters. Detection events are random and discrete points in time, so to get a mean trajectory, we in-

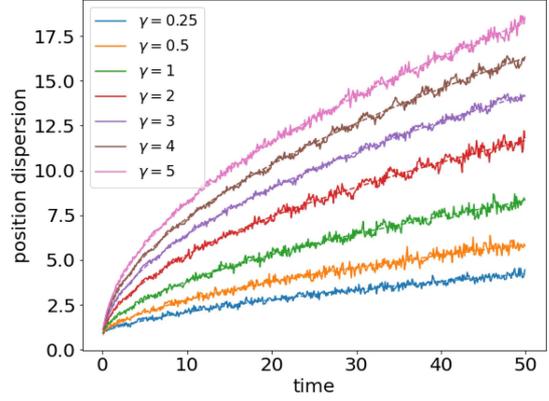


Fig. 3. Dispersion in position (and equivalently, momentum) $\sqrt{\delta_0^2(t)}$ as a function of time. The colours correspond to different values of γ for $d = 2.16$. The dispersion grows faster for larger γ . The figure also shows fits of (12) as dashed lines in the corresponding color, which are barely visible because of high agreement.

roduce coarse-grained time by dividing the timeline into small intervals, $[t, t + \delta t]$, where $\delta t = 0.1$, and calculate the mean position and momentum for all clicks from the ensemble falling into the interval. The mean trajectories, both in position and momentum space, show that, *on average*, the particle follows a classical path (cf. Fig. 2). The position as well as momentum oscillate with the harmonic oscillator frequency and are phase-shifted by $\pi/2$.

Deviations of a single realization from the average trajectory are characterized by the second moment of the click distribution, i.e., the dispersion $\delta^2 x(t) = \langle x(t)^2 \rangle - \langle x(t) \rangle^2$ and $\delta^2 p(t) = \langle p(t)^2 \rangle - \langle p(t) \rangle^2$. Because of the symmetry of the Hamiltonian (5) and (11), the dispersion in position and momentum should be equivalent to $\delta^2 x = \delta^2 p$. The simulations essentially confirm these expectations, which is why in Fig. 3, we only plot the dispersion function $\delta(t) \equiv \sqrt{\delta^2 x}$. The dispersion functions of position and momentum actually differ by a small modulation due to the $\pi/2$ phase shift of position and momentum of the particle. This will be discussed later on in this section.

The time dependence, fitted to the numerical results, is found to be

$$\delta^2(t) \approx Dt + \delta_0^2, \quad (12)$$

where D is a diffusion coefficient and δ_0^2 the initial dispersion, independent of γ ; δ_0^2 is a result of the initial wavepacket having a finite width even when identical to the detector wavefunction. In other words, because of the lack of orthogonality, immediately after localization at the detector at the position $(x_0, p_0) = (j_0 d, k_0 d)$, the particle may be captured by a different detector $(x, p) = (j d, k d)$. In our model, the probability distribution of a subsequent click of the detector $\alpha_{j,k}$, under the condition

that such a click occurs within a short time from the first one, is approximately equal to the discretized Husimi function $\mathcal{Q}(j, k)$ of the initial state [45]

$$\mathcal{Q}(j, k) = |\langle \alpha_{j,k} | \alpha_{j_0, k_0} \rangle|^2 = \frac{e^{-\frac{1}{2}d^2[(j-j_0)^2 + (k-k_0)^2]}}{\sum_{j,k} |\langle \alpha_{j,k} | \alpha_{j_0, k_0} \rangle|^2}. \quad (13)$$

According to the discussion above, the dispersion squared of the initial spatial position of the monitored particle is $\delta_0^2 = \sum_{j,k} \mathcal{Q}(j, k)(jd)^2$. If $d \ll 1$, summation can be substituted by integration, which yields $\delta_0^2 \approx 1$. In the case of the numerical results shown in Fig. 3, this condition is not satisfied ($d = 2.16$), but surprisingly we find that this continuous approximation still works quite well.

For large times, the initial dispersion can be neglected, and (12) indicates that on the top of the harmonic oscillation, the particle undergoes Brownian motion. Deviations from the mean trajectory grow as the square root of time, suggesting a diffusion process characterized by the D coefficient. Moreover, from the dimensional analysis, it seems that dynamical quantities such as $\delta^2(t)$ should depend on the dimensionless parameter γt . Indeed, detailed studies confirm this prediction (see Fig. 4a). This implies that the diffusion coefficient D grows linearly with γ , which is plausible since this implies more frequent detection of the particle. Similarly, the denser the detector grid, the more detectors monitor the particle, which in turn leads to a higher detection frequency and larger perturbations of the classical trajectory. In Fig. 4b, we show the dependence of the diffusion coefficient on the detector spacing d for a fixed value of $\gamma = 1.0$. The results clearly show that D is inversely proportional to the squared detector spacing. Our numerical experiment allows us to postulate the following dependence of the diffusion coefficient on the parameters of the observation process

$$D \approx 2\pi \frac{\gamma}{d^2}. \quad (14)$$

The analytical formula (12) shows very good agreement with numerical calculations. This formula may also be confirmed by approximate analytical considerations. The diffusion coefficient is related to the squared mean displacement of a walking particle per unit of time, i.e.,

$$D = \gamma \sum_{j,k} e^{-d^2(j^2+k^2)/2} (dj)^2. \quad (15)$$

Using the continuum approximation, $jd = x$, $kd = p$, and $\sum_{j,k} \rightarrow \frac{1}{d^2} \int dx dp$, the diffusion coefficient is equal to

$$D \approx \frac{\gamma}{d^2} \int dx dp x^2 e^{-(x^2+p^2)/2} = \frac{2\pi\gamma}{d^2}. \quad (16)$$

We thus recovered (14), which was obtained by fitting it to numerical data.

A more careful analysis indicates that in addition to the Brownian diffusion characterized by a linear growth of the dispersion $\delta^2(t)$, there are

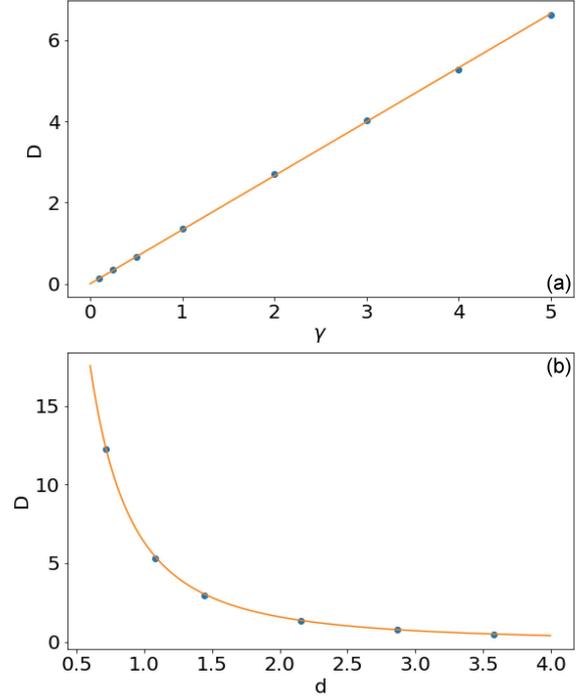


Fig. 4. Dependence of the diffusion coefficient D on parameters of detection. (a) Dependence of D on γ for $d = 2.16$. Blue points correspond to simulations while the orange curve is a fit of the linear function, $D \propto \gamma$. (b) Dependence of D on distance d between detectors for a fixed value of $\gamma = 1$. The blue points are the simulation results, while the orange line is the fit of $D \approx (\frac{5}{2})^2 \frac{1}{d^2}$ function.

small-amplitude oscillations with the frequency 2ω . These oscillations can be explained assuming small dephasings of individual trajectories $x=x_0 \cos(t + \delta\varphi)$ with respect to the average $x=x_0 \cos(t)$. Dephasing gives an oscillatory contribution to the dispersion $\langle x^2 \rangle - \langle x \rangle^2 \approx \delta\varphi^2 \sin^2(t)$. A similar oscillatory character of the dispersion of position and momentum was observed in [47], where the phase space dynamics of a continuously monitored particle in an anharmonic potential is studied. In that work, however, the dispersion is bounded, contrary to the result presented here. This is because the authors of [47] have studied the limit of very frequent and very weak measurements, whereas the present work treats a series of strong measurements at discrete points in time. Each measurement is performed at the ‘‘Heisenberg limit’’, i.e., it minimizes the uncertainty relation

$$\sigma_x \sigma_p = \frac{1}{2}. \quad (17)$$

Such a measurement necessarily introduces growing fluctuations. Our studies indicate that the dispersion of trajectories is model/system sensitive. This fact was also noticed by us in [45], where different types of diffusion were found for alternative POVMs of measurement operators.

Fluctuations in the position and momentum of a particle lead to an increase in its energy. It is because of this that when we observe the sample trajectory in phase space, it tends to be a circular motion spiraling outwards (see Fig. 1). The radius of a circle in phase space increases with time, $r(t) = \sqrt{2\langle E(t) \rangle}$. It follows directly from (12) that the average energy of the particle, $\langle E \rangle = \frac{1}{2}\langle x^2 \rangle + \frac{1}{2}\langle p^2 \rangle$, grows linearly with time,

$$\langle E(t) \rangle = \delta^2(t) + E_0 = Dt + (\delta_0^2 + E_0), \quad (18)$$

where $E_0 = \frac{1}{2}(x_0^2 + p_0^2)$ is the initial energy of a classical particle at initial position x_0 with initial momentum p_0 .

By dividing the energy scale into small intervals ΔE , we can obtain the energy distribution $p_E(t)$ of the ensemble of trajectories as a function of time. This distribution around $t = 0$, as obtained from our simulations, is shown in Fig. 5. This is a relatively narrow function centered around E_0 . Again, as in the case of position dispersion, the initial distribution of energy can be approximately obtained from analytic calculations. As previously, we use the continuum approximation: $j(j_0)d \rightarrow x(x_0)$, $k(k_0)d \rightarrow p(p_0)$, and $\mathcal{Q}_{i,j} \rightarrow \mathcal{P}(x,p) = \frac{1}{2\pi} e^{-\frac{1}{2}(x-x_0)^2} e^{-\frac{1}{2}(p-p_0)^2}$. If the particle is initially placed at phase space location (x_0, p_0) , then the initial energy distribution is

$$p_E = \int dx dp \mathcal{P}(x,p) \delta(E - \frac{1}{2}(x^2 + p^2)). \quad (19)$$

Using that $2E_0 = x_0^2 + p_0^2$, we get

$$p_E = e^{-(E+E_0)} I_0(\sqrt{2E_0}\sqrt{2E}), \quad (20)$$

where $I_0(z)$ is the modified Bessel function of the first kind.

The energy distribution, as given by (20), is plotted in Fig. 5a. Again, the continuous approximation works quite well even for the parameters that do not fully legitimate the use of the formula. We stress that to get the energy histogram, we accumulated data from the time interval $0 < t < 2\pi$, so strictly speaking, the histogram does not give the energy distribution exactly at $t = 0$, but the distribution averaged over the first period of the oscillation. For large times t , this initial energy distribution evolves into a thermal distribution

$$p_E(t) = \frac{1}{\epsilon(t)} e^{-E/\epsilon(t)}. \quad (21)$$

The width and mean of this $\epsilon(t)$ distribution depend on time. Setting $\epsilon = k_B T$ allows us to formally define a temperature for the system, identifying the repeated measurement process with the type of ‘‘heating.’’ The distribution p_E in the thermal regime is shown in Fig. 5b. The temperature of the ensemble grows with time, and for large times, it becomes $k_B T(t) = \langle E \rangle = \delta^2(t) \approx Dt$. This analytical prediction again agrees well with the numerical results.

In summary, we studied a quantum particle in an external harmonic potential that is repeatedly monitored by an array of detectors regularly dis-

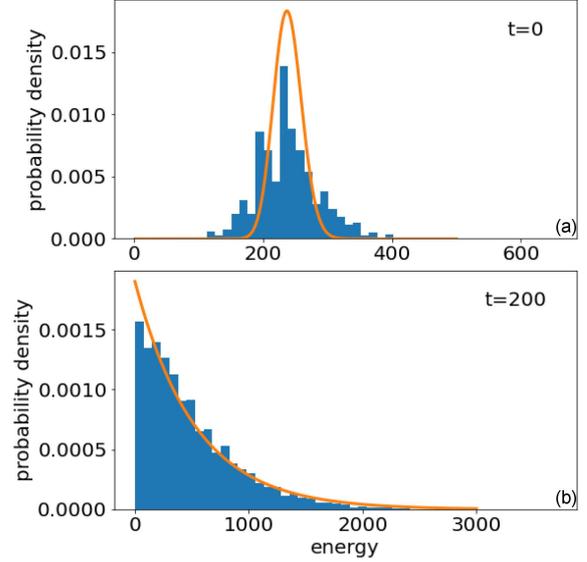


Fig. 5. Energy distribution for two different times. Panel (a) shows a distribution shortly after the beginning of the simulation, and panel (b) shows the distribution at a ‘‘late time’’ (after many oscillations). The blue histograms correspond to numerical data, while the orange curves are the function (20) (a) and an exponential (b) fit. Note that the labels $t = 0$ and $t = 200$ are approximate in the sense that in order to gather sufficient numerical data for the histogram, we consider clicks from a time interval corresponding to one full oscillation.

tributed in phase space. We employed an open system formalism, treating the detectors as an external reservoir. Coupling of the particle to the meters is given by jump operators whose action is to project the particle’s wavefunction onto coherent states characterizing the detectors. We use the wavefunction quantum Monte Carlo method to generate ensembles of time-dependent wavefunctions. We interpret every generated wavefunction as a single realization of the particle’s dynamics, which in addition to continuous evolution, experiences quantum jumps related to observations. We show that, on average, the trajectories follow the classical path. This result is similar to the one in [47], where a von Neumann type of coupling between the system — being a nonlinear oscillator — and the meters were considered. Random quantum jumps in position and momentum space introduce fluctuations on top of harmonic motion. We have shown that these fluctuations have the character of Brownian motion, as the diffusive process with the dispersion of position and momentum grows linearly with time. We numerically found the diffusion coefficient and its dependence on the detector clicking rate γ and the detector spacing d . Going back to dimensional units, we see that the diffusion coefficient D_x in position space is proportional to the Planck constant

$$D_x = 4\pi\gamma \frac{\hbar}{d_x d_p} \sigma^2, \quad (22)$$

signifying the quantum character of this process. Again, this is due to the fact that our measurements are performed at the limit set by the Heisenberg uncertainty limit (see (17)). Here, $d_x d_p$ is the action equal to the area of an elementary cell in the phase space, determined by the detector spacing.

Finally, we found that repeated observation introduces heating of the particle, the energy distribution of the trajectory ensemble at large times becomes thermal, and the effective temperature grows linearly in time.

Our studies of the system under continuous monitoring and comparison to similar studies [45, 47] show that the observed mean trajectories correspond to the classical trajectories, however, the deviation from the mean (the dispersion) significantly depends on the system studied and details of the detection process, in particular on a choice of the positive operator-valued measure.

We do not know whether the particular measurement schemes considered here can ever be realized in practice. However, the model we formulate is fully admissible in view of the present understanding of the quantum measurement theory. As such, it is legitimate to study its consequences. Paraphrasing the words of Professor Iwo Białynicki [48]: “As to the usefulness of our results, we have no opinion at all. Perhaps someone else could see whether they are good for anything.”^{†1}

Acknowledgments

The paper is dedicated to Professor Iwo Białynicki-Birula on the occasion of his 90th birthday. The lectures on quantum mechanics, given by the Professor at the Physics Department of Warsaw University in the fall semester of 1976, played a very important role in the scientific development of one of us (MG). Moreover, MG is especially grateful to the Professor for his particular care, which MG experienced during his professional life. The Professor and his wife, Zofia, were always eager to offer their friendly help at moments of important decisions.

We thank Magdalena Załuska-Kotur for illuminating us on the methods of determination of a diffusion coefficient and for pointing out to us that $\frac{5}{2} \approx \sqrt{2\pi}$.

This work was supported by the Polish National Science Centre grant No. 2019/32/Z/ST2/00016, through the project MAQS under QuantERA, which has received funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement No. 731473.

^{†1}Original quote: “As to the usefulness of my results, I have no opinion at all. Perhaps someone else could see whether they are good for anything.”

References

- [1] M. Born, *Nature* **119**, 354 (1927).
- [2] W. Heisenberg, *Die Physikalischen Prinzipien der Quantentheorie*, Ed. S. Hirzel, Leipzig 1930; W. Heisenberg, *The Physical Principles of the Quantum Theory*, English translation by C. Eckart, F.C. Hoyt, Dover Publications, 1949.
- [3] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin 1932; reprinted 1981; English translation by R.T. Beyer, *Mathematical Foundations of Quantum Mechanics*, Princeton University, Princeton (NJ) 1955.
- [4] A. Peres, *Stud. Hist. Philos. Mod. Phys.* **33**, 23 (2002).
- [5] G. Lüders, *Ann. Phys. (Leipzig)* **8**, 322 (1951).
- [6] F. London, E. Bauer, *La theorie de l’observation en mecanique quantique*, Hermann, Paris 1939; English translation in: *Quantum Theory and Measurement*, Eds. J.A. Wheeler, H. Zurek, Princeton University, Princeton (NJ) 1983, p. 217.
- [7] E.P. Wigner, *Am. J. Phys.* **31**, 6 (1963).
- [8] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, Springer, Berlin 1983.
- [9] G. Ludwig, *Foundations of Quantum Mechanics I and II*, Springer, Berlin 1983.
- [10] A.S. Holevo, *Statistical Structure of Quantum Theory, Lecture Notes in Physics Monographs 67*, Springer, Berlin 2001.
- [11] W.H. Zurek, *Rev. Mod. Phys.* **75**, 715 (2003).
- [12] H.M. Wiseman, G.J. Milburn, *Quantum Measurement and Control*, Cambridge University Press, Cambridge 2010.
- [13] I. Białynicki-Birula, Z. Białynicka-Birula, *Quantum Electrodynamics*, Elsevier, 1975.
- [14] E. Schrödinger, *Br. J. Philos. Sci.* **3**, 233 (1952).
- [15] C.M. Caves, C.A. Fuchs, R. Schack, *Phys. Rev. A* **65**, 022305 (2002).
- [16] Fuchs, C.A. Preprint, [arXiv:1003.5209](https://arxiv.org/abs/1003.5209), 2010.
- [17] C.A. Fuchs, R. Schack, *Rev. Mod. Phys.* **85**, 1693 (2013).
- [18] N. David Mermin, *Nature* **507**, 421 (2014).
- [19] N. Davis Mermin, *Phys. Today* **75**, 62 (2022).
- [20] H. Everett, *On the Foundations of Quantum Mechanics*, Mudd Manuscript Library — Remote Storage (ReCAP), PRIN 685.1957.17, 1957.

- [21] D. Bohm, *Phys. Rev.* **85**, 166 (1952); **85**, 180 (1952).
- [22] M.B. Mensky, *Phys. Rev. D* **20**, 384 (1979).
- [23] E.B. Davies, *Commun. Math. Phys.* **15**, 277 (1969).
- [24] E.B. Davies, *Quantum theory of open systems*, Academic Press London, 1976.
- [25] W. Neuhauser, M. Hohenstatt, H. Dehmelt, P. Toschek, *Phys. Rev. Lett.* **41**, 233 (1978).
- [26] W. Nagourney, J. Sandberg, H. Dehmelt, *Phys. Rev. Lett.* **56**, 2797 (1986).
- [27] T. Sauter, W. Neuhauser, R. Blatt, P.E. Toschek, *Phys. Rev. Lett.* **57**, 1696 (1986).
- [28] J.C. Bergquist, R.G. Hulet, W.M. Itano, D.J. Wineland, *Phys. Rev. Lett.* **57**, 1699 (1986).
- [29] C. Caves, *Phys. Rev. D* **33**, 1643 (1986).
- [30] C.M. Caves, G.J. Milburn, *Phys. Rev. A* **36**, 5543 (1987).
- [31] A. Barchielli, L. Lanz, G.M. Prosperi, *Nuovo Cim.* **72B**, 79 (1982).
- [32] A. Barchielli, L. Lanz, G.M. Prosperi, *Foundations of Quantum Mechanics*, Eds. S. Kamefuchi et al., Physical Society of Japan, Tokyo 1984, p. 165.
- [33] L. Diósi, *Phys. Lett. A* **129**, 419 (1988).
- [34] L. DiÁlsi, *Phys. Lett. A* **132**, 233 (1988).
- [35] N. Gisin, *Phys. Rev. Lett.* **52**, 1657 (1984).
- [36] V. Gorini, A. Kossakowski, E.C.G. Sudarshan, *CJ. Math. Phys.* **17**, 821 (1976).
- [37] G. Lindblad, *Commun. Math. Phys.* **48**, 119 (1976).
- [38] K. Jacobs, D.A. Steck, *Contemp. Physics* **47**, 279 (2006).
- [39] N. Gisin, I.C. Percival, *J. Phys. A: Math. Gen.* **25**, 5677 (1992).
- [40] H.J. Carmichael, *An Open Systems Approach to Quantum Optics*, Springer-Verlag, Berlin 1993.
- [41] P. Zoller, C.W. Gardiner, in: *Fluctuations Quantiques*, Eds. S. Reynaud, E. Giacobino, J. Zinn-Justin, North-Holland, Amsterdam 1997, p. 79.
- [42] Y. Ashida, M. Ueda, *Phys. Rev. A* **95**, 022124 (2017).
- [43] J. Dalibard, Y. Castin, K. Mølmer, *Phys. Rev. Lett.* **68**, 580 (1992).
- [44] K. Mølmer, Y. Castin, J. Dalibard, *J. Opt. Soc. Am. B* **10**, 524 (1993).
- [45] F. Gampel, M. Gajda, *Phys. Rev. A* **107**, 012420 (2023).
- [46] E. Arthurs, J.L. Kelly Jr., *Bell Syst. Tech. J.* **44**, 725 (1965).
- [47] A.J. Scott, G.J. Milburn, *Phys. Rev. A* **63**, 042101 (2001).
- [48] I. Białynicki-Birula, *Lett. Math. Phys.* **10**, 189 (1985).

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— ◊ —

Electron Localization in Rydberg States

J. MOSTOWSKI* AND J. PIETRASZEWICZ

Institute of Physics, Polish Academy of Sciences, al. Lotników 32/46, PL-02668 Warsaw, Poland

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S140](https://doi.org/10.12693/APhysPolA.143.S140)*e-mail: jan.mostowski@ifpan.edu.pl

We discuss the possibility of localizing an electron in a highly excited Rydberg state. The second-order correlation of emitted photons is the tool for the determination of electron position. This second-order correlation of emitted radiation and, therefore, the correlation of operators describing the acceleration of the electron allows for a partial localization of the electron in its orbit. The correlation function is found by approximating the transition matrix elements by their values in the classical limit. It is shown that the second-order correlation, depending on two times, is a function of the time difference and is a periodic function of this argument with the period equal to the period of the corresponding classical motion. The function has sharp maxima corresponding to large electron acceleration in the vicinity of the “perihelion.” This allows the localization of the electron in its consecutive approach to the perihelion point.

topics: Rydberg state, radiation, second order correlation, localization

1. Introduction

The measurement process, since the early days of quantum physics, has been one of the central issues in many attempts to understand the relation between the classical and quantum description of physical systems [1] (for a more recent analysis, see, e.g., [2]). The most common theory of quantum measurement [3–5] assumes that the quantum system is coupled to a meter. The interaction between them entangles the two systems. The measurement, described as a projection onto the state of the meter, provides information on the state of the system. Continuous measurements of quantum systems treated as stochastic processes were first considered in [6] in the context of photon counting. The formalism based on path integrals was initiated in [7] and further developed in [4] (see also [8]).

The classical motion of the electron bound in a Coulomb field is periodic. The wavefunction describing the bound electron in a stationary state does not show any time-dependent features. Time dependence, and hence classical features of wavefunctions, can be obtained for non-stationary states, linear combinations of energy eigenstates with different energies. Such a construction is well known in the case of a harmonic oscillator, and the most

classical states are well-known coherent states [9] (see also, e.g., [10]). The corresponding time-dependent states in the case of Rydberg states were introduced in [11] (see also [12–15]).

Another point of view was presented in [16], where it is pointed out that when a measurement breaks the time-translational symmetry of a stationary state, a periodic motion of the system is initiated. This approach was further elaborated in [17, 18].

The classical limit of quantum mechanics is still a vivid subject of investigation (see, e.g., [19]). One of the recently discussed problems in this area relates to the successive measurements of particle position and detection of the trajectory. Most of the interest has been limited to free particles, and not much has been done in the case of bound states.

Quantum description of the hydrogen atom is well known. All energies and wavefunctions of stationary states are well known. The classical limit is approached in the limit of large quantum numbers — the wavefunction should be related to classical trajectories. This relation has been discussed in many papers. Both time-dependent states, analogs of harmonic oscillator coherent states, and stationary states in the limit of high excitation were shown to exhibit classical features.

In this paper, we will present yet another aspect of the classical limit in the case of Rydberg states. Namely, we will use the radiation emitted from the highly excited state to determine the electron position as a function of time. Detection of radiation at a given time breaks the time-translational symmetry and allows observation of the time dependence of subsequent evolution. This approach provides a partial but straightforward way of estimating the elements of the time-dependent classical trajectory hidden in the stationary wavefunction.

Radiation from a quantum system, such as a hydrogen atom, is usually studied in the frequency domain. The spectrum consists of several lines. Measurement of the spectrum is not the only possibility — time dependence of radiation can be studied as well. The time dependence of the spontaneous emission from a highly excited vibrational state of a diatomic molecule was used to determine the time-dependent relative position of the constituents. This allowed us to demonstrate the time dependence of various states, such as coherent states and others, e.g., the Schrödinger cat state [20]. Let us note that in the case of Rydberg states with the principle quantum number $n \approx 100$, the characteristic frequency of radiation is $\nu \approx 10^{10}$ Hz, so the time dependence of the radiation for times smaller than $1/\nu$ is within experimental reach. Radiation observed for such small times of the order of $1/\nu$ exhibits different features as compared to the long-time measurements. This and the relation to the position measurement will be discussed below.

2. A simple case — harmonic oscillator

We will begin the discussion of electromagnetic radiation in the time domain and its relation to the measurement of the electron position with a simple example of a harmonic oscillator. The charged particle oscillates with the frequency ω along the x axis; its motion is given by $x_{cl}(t) = A \cos(\omega t)$. This electron is a source of electromagnetic radiation. We will find the x component of the electric field in the far zone along the y axis (to simplify the geometry). We have, in the dipole approximation,

$$E_x(\mathbf{R}, t) = -\frac{e}{4\pi\epsilon_0 R} \omega^2 x_{cl}(t_{ret}), \quad (1)$$

where $t_{ret} = t - |\mathbf{R}|/c$ is the retarded time, and c is the speed of light. We have skipped the R dependence of the field — it is just like in classical electrodynamics, namely $E_x \sim R^{-1}$. It follows from (1) that the electric field oscillates with the frequency ω . This classical treatment does not take into account radiation damping, thus it is valid only for a short time, shorter than the characteristic damping time.

We will now discuss an emission of radiation, taking into account the quantum nature of the oscillator. We will concentrate on the highly excited states of the oscillator and hence on the classical limit.

The position of an oscillating particle is described by the position operator x . It can be expressed in terms of the lowering and raising operators a and a^\dagger , respectively, as follows

$$x = x_0 \frac{(a + a^\dagger)}{\sqrt{2}}, \quad (2)$$

where $x_0 = \sqrt{\frac{\hbar}{M\omega}}$, \hbar is the Planck constant, and M denotes the mass of the oscillating particle. The component E_x of the electric field operator (the radiated part) in the dipole approximation is given by

$$E_x(\mathbf{R}, t) = -\frac{e}{4\pi\epsilon_0 R} \omega^2 x(t_{ret}), \quad (3)$$

just like in the classical case. This time, however, the electric field is an operator, and we will find the expectation values of this operator. We assume that at time $t = 0$, the oscillator is in the energy eigenstate $|n\rangle$ with energy $E_n = \hbar\omega n$. Thus the expectation value of the x operator, and hence of the $E_x(\mathbf{r}, t)$ operator, is equal to zero. The first-order correlation function becomes

$$\begin{aligned} \langle E_x(\mathbf{R}, t_2) E_x(\mathbf{R}, t_1) \rangle &= \frac{1}{2} \frac{e^2}{(4\pi\epsilon_0 R)^2} \omega^4 x_0^2 \\ &\times \left[n e^{i\omega(t_2-t_1)} + \left(n + \frac{1}{2} \right) e^{-i\omega(t_2-t_1)} \right]. \end{aligned} \quad (4)$$

In the case of the highly excited state, i.e., when $n \gg 1$, we can approximate $\sqrt{n(n+1)} \approx n \approx \sqrt{n(n-1)}$. Then we get

$$\begin{aligned} \langle E_x(\mathbf{R}, t_2) E_x(\mathbf{R}, t_1) \rangle &= \\ &= \frac{e^2}{(4\pi\epsilon_0 R)^2} \omega^4 x_0^2 n \cos(\omega(t_2-t_1)), \end{aligned} \quad (5)$$

just as in the classical case. The average intensity of radiation given by the first correlation function at $t_2 = t_1$ is a constant. The first correlation function for $t_2 > t_1$ gives the spectrum of radiation and, in this case, consists of one line only.

The second-order correlation function is more interesting. For $n \gg 1$, we get

$$\langle E_x^2(\mathbf{R}, t_2) E_x^2(\mathbf{R}, t_1) \rangle = n^2 \left[1 + \cos(2\omega(t_2-t_1)) \right]. \quad (6)$$

The second correlation function oscillates with the frequency 2ω . This tells us that the maxima of radiation occur every half period of the electron motion. Thus, the second correlation function can be used to determine the position of the oscillating particle in the vicinity of a turning point. The high intensity is due to the large acceleration of the oscillating charge and this takes place when the electron is close to one of the turning points. Thus, if high intensity has been detected at t_1 , then the electron will reach another turning point half the period later, and the intensity will be high once more. Thus, the time dependence of the second correlation function provides information about the motion of the electron. The information is not complete, as

the radiation does not distinguish between the two turning points. It is worth noting that the correlation function allows the detection of the particle close to the turning point in spite of the dipole approximation.

3. Classical radiation from Kepler orbit

Before we discuss radiation from the Rydberg states, we will give a classical description of motion in the Coulomb field [21]. If the motion is in the xy plane, the coordinates x and y as functions of time are given by

$$\begin{aligned} x(t) &= a [\cos(\xi(t)) - \epsilon], \\ y(t) &= a \sqrt{1 - \epsilon^2} \sin(\xi(t)), \end{aligned} \quad (7)$$

where

$$\omega t + \varphi = \xi(t) - \epsilon \sin(\xi(t)), \quad (8)$$

where φ is an arbitrary phase. The radial variable $r = \sqrt{x^2 + y^2}$ can also be expressed as a function of time

$$r = a [1 - \epsilon \cos(\xi(t))]. \quad (9)$$

The parameters a , ω , and ϵ characterize the trajectory. They can be related to energy and angular momentum in the standard way [21].

We will also need more general trajectories that differ by an orientation in the plane of the motion described by the phase χ and by the phase of the motion, φ . Thus we define

$$\begin{aligned} X(t) &= x(t) \cos(\chi) + y(t) \sin(\chi), \\ Y(t) &= -x(t) \sin(\chi) + y(t) \cos(\chi), \end{aligned} \quad (10)$$

with $\omega t + \varphi = \xi(t) - \epsilon \sin(\xi(t))$.

The classical description of the radiation of a charge moving along such an orbit is found to be in complete analogy to the harmonic oscillator case. We will use the dipole approximation since the size of the orbit is much smaller than the characteristic wavelengths of the emitted radiation. The electric field in the far zone is given by

$$\mathbf{E}(\mathbf{R}, t) = \frac{1}{R} \mathbf{n} \times [\mathbf{n} \times \mathbf{a}(t_{ret})], \quad (11)$$

where \mathbf{a} is the acceleration, and $\mathbf{n} = \mathbf{R}/|R|$. Radiation damping is neglected, as in the previous section.

Also, the Fourier decomposition of the trajectory can be found (see [21]). Here we will give the Fourier decomposition of the x variable

$$x(t) = \sum_k \exp(ik(\omega t + \varphi)) x_k, \quad (12)$$

where

$$x_k = \frac{a}{2k} [J_{k-1}(k\epsilon) - J_{k+1}(k\epsilon)], \quad k \neq 0. \quad (13)$$

A similar formula holds for $y(t)$. This will be used in the next section.

4. Classical limit of matrix elements

From now on, we will use atomic units.

Consider the quantum description of an atom in a highly excited energy eigenstate. We label the states by standard quantum numbers: n — principle quantum number, l — angular momentum quantum number, and m — magnetic quantum number. The energy E_n of this state depends on the principal quantum number n as $E_n = -1/(2n^2)$. We will be interested only in states with $m = l$, thus, we will skip the magnetic quantum number to avoid confusion. This means that the wavefunctions considered in this paper are well concentrated in the xy plane, which is perpendicular to the angular momentum. This can be seen from the explicit form of the spherical harmonics function $|Y_{l,l}(\theta, \varphi)|^2 \sim \sin^{2l}(\theta)$ that has a sharp maximum at $\theta = \frac{\pi}{2}$ for large l . We will, therefore, not consider the wavefunction dependence along the z axis.

The expectation values of the radiated field depend on the matrix elements of the position operator between the quantum states of the atom, i.e., $\langle \psi_{n,l,l} | x | \psi_{n',l',l'} \rangle$, where x is the coordinate. In spherical coordinates, $x = r \sin(\theta) \cos(\varphi)$, and a similar expression is valid for the y coordinate, $y = r \sin(\theta) \sin(\varphi)$. The wavefunctions $\psi_{n,l,l}(r, \theta, \varphi) = R_{n,l}(r) Y_{l,l}(\theta, \varphi)$ are the standard states of the hydrogen atom, with $R_{n,l}(r)$ describing the radial part of the wavefunction and $Y_{l,l}(\theta, \varphi)$ denotes the spherical harmonics. Because of selection rules, these matrix elements are different from zero only if $l' = l \pm 1$.

The radial part of the matrix element of r^k (for any k), i.e.,

$$\int_0^\infty dr r^{2+k} R_{n,l}(r) R_{n',l'}(r), \quad (14)$$

can be found explicitly in terms of special functions [22]. In fact, the classical limit of this expression, valid for $n \rightarrow \infty$, $l \rightarrow \infty$ with $l/n = \text{const}$, has been found in [23]. In this limit, (14) approaches the Fourier transform of the classical trajectory $r_{classical}^k$ for the frequency $\omega = (E_n - E_{n'})/\hbar$. The classical trajectory $r(t)$ corresponds to the average energy $E = \frac{1}{2}(E_n + E_{n'})$ and the eccentricity $\epsilon = \sqrt{1 - (l/n)^2}$. Thus, for the matrix element of r , we find for $l' = l \pm 1$ that

$$\begin{aligned} \langle n', l', l' | r | n, l, l \rangle &\approx a_0 \frac{n^2}{2(n-n')} \\ &\times \left[J_{n-n'+1}((n-n')\epsilon) - J_{n-n'-1}((n-n')\epsilon) \right], \end{aligned} \quad (15)$$

where a_0 denotes the Bohr radius, and ϵ corresponds to the eccentricity of the classical orbit with energy and angular momentum equal to the average of the energies of the initial and final state. It should be noted that (15) is analogous to (5) for the harmonic oscillator, where $\sqrt{n(n+1)}$ is replaced by n for large quantum numbers n .

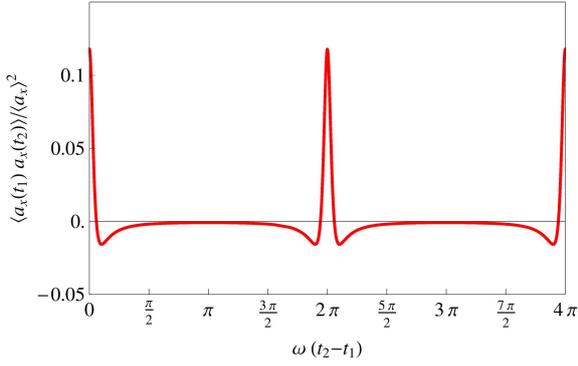


Fig. 1. The first correlation function of accelerations (23) (normalized to the average square of acceleration) as a function of $\omega(t_2 - t_1)$ for two periods ($\epsilon = 0.8$).

The transition elements for x can also be found $\langle n', l+1, l+1 | x | n, l, l \rangle + \langle n', l-1, l-1 | x | n, l, l \rangle$

$$\approx x_{n-n'}, \quad (16)$$

where $x_{n-n'}$ is given by (13). These formulas allow describing the radiation from the Rydberg states using classical approximations.

The values of the matrix elements can be modeled classically by random trajectories. Consider then the trajectories

$$X(t) = x(t) \cos(\chi) + y(t) \sin(\chi),$$

$$Y(t) = -x(t) \sin(\chi) + y(t) \cos(\chi), \quad (17)$$

with $\omega t + \phi = \xi - \epsilon \sin(\xi)$. The quantities ϕ and χ are random phases, with uniform distributions between 0 and 2π . In this case, the expectation values of the x and y operators are equal to the mean values of the classical quantities X and Y with the same values of energy and angular momentum.

5. Radiation from a Rydberg state

In this and subsequent sections, we will use atomic units in the description of a quantum state.

The electric field $\mathbf{E}(\mathbf{R}, t)$ in the far field is given by the same formula as in the classical field, with the difference that the acceleration \mathbf{a} is an operator acting on the quantum state of the system consisting of an electron and the photon vacuum. In the quantum case, also the electric field is an operator. Thus, for the radiated part of the field, we get in the dipole approximation

$$\mathbf{E}(\mathbf{R}, t) = \frac{1}{R} \mathbf{n} \times [\mathbf{n} \times \mathbf{a}(t_{ret})], \quad (18)$$

where \mathbf{n} is the unit vector in the direction of the observation point, $\mathbf{n} = \frac{\mathbf{R}}{|\mathbf{R}|}$.

In what follows, we will find the expectation values of the electric field, as well as the first and second correlation function. It should be noted that the radiation is weak, and therefore the measurement of

light intensity in the classical sense is questionable. The expectation value of the electric field squared at a given point should be understood as the photon counting rate.

We assume that at $t = 0$, the state describes the photon vacuum and the atom is in the state $\psi_{n,l,l}$. This requires matrix elements of the operators x and y and their second derivatives over time.

The first correlation function of the x component of the field radiated in the y direction is given by

$$\langle E_x(\mathbf{R}, t_2) E_x(\mathbf{R}, t_1) \rangle = \frac{\langle a_x(t_{2,ret}) a_x(t_{1,ret}) \rangle}{(4\pi\epsilon_0 R)^2}. \quad (19)$$

The expectation value of the product of accelerations will be found in the classical limit. First, we will linearize the energy in the vicinity of the initial state energy with the principal quantum number n_0 . We get

$$E_n \approx -\frac{1}{2n_0^2} + \frac{n - n_0}{n_0^3}. \quad (20)$$

This allows for the approximation of the expectation values of the acceleration operator $a(t)$ by the expectation values of the r operator

$$\begin{aligned} \langle n', l-1, l-1 | a_x(t) | n, l, l \rangle \approx \\ - (n-n')^2 \omega_0^2 \exp(-i(n-n')\omega_0 t) x_{n-n'}, \end{aligned} \quad (21)$$

with $\omega_0 = 1/n_0^3$. Thus, for the two-time correlation function of acceleration in the state $|n, l, l\rangle$, the following can be found

$$\begin{aligned} \langle a_x(t_2) a_x(t_1) \rangle = \sum_{n'l'} \langle n, l, l | x | n'l'l' \rangle \langle n'l'l' | x | n, l, l \rangle \\ \times \exp(i\omega(t_2 - t_1)(n - n')). \end{aligned} \quad (22)$$

The same can be expressed by the correlation of the classical trajectories

$$\langle a_x(t_2) a_x(t_1) \rangle \approx \int \frac{d\phi}{2\pi} \int \frac{d\chi}{2\pi} \frac{d^2 X(t_2)}{dt_2^2} \frac{d^2 X(t_1)}{dt_1^2}. \quad (23)$$

This is a good approximation for large n and l . The main point is that the matrix elements of the angular part

$$\int d\theta \sin(\theta) d\phi Y_{l,l}(\theta, \varphi) \sin(\theta) e^{i\varphi} Y_{l-1,l-1}(\theta, \varphi), \quad (24)$$

hence the matrix element of the position operator x weakly depends on l for large l . The correlation function obtained above is shown in Fig. 1.

From the above considerations, it follows that the average intensity of radiation is proportional to the correlation function at $t_1 = t_2$ and does not depend on time. The Fourier transform of the correlation function

$$\int dt \langle a_x(t) a_x(0) \rangle \exp(ik\omega t) \quad (25)$$

determines the radiation spectrum. Thus the spectrum of radiation from a Rydberg state can be approximated by the spectrum of radiation from the corresponding classical orbit.

6. Second order correlation

In this section, we will discuss the second-order correlation function of the radiation originating from a Rydberg state. This is given by

$$G(t_2, t_1) = \left\langle E_x(t_1)E_x(t_2)E_x(t_2)E_x(t_1) \right\rangle. \quad (26)$$

The state is, as before, the photon vacuum and the Rydberg state of the atom. Expressing the electric field by the acceleration of an electron in the atom, we get

$$G(t_2, t_1) = \frac{1}{R^4} \exp(-2i n\omega(t_1 - t_2)) \times \langle a_x(t_1)a_x(t_2)a_x(t_2)a_x(t_1) \rangle. \quad (27)$$

Just as before, we insert a complete set of states $|n, l, l\rangle$ between the a operators and apply the approximation of l independence of the matrix elements in the case of large l . This leads to the following representation of the correlation function

$$G(t_2, t_1) = \frac{1}{R^4} \int \frac{d\chi}{2\pi} \int \frac{d\varphi}{2\pi} \times \frac{d^2X(t_2)}{dt_2^2} \frac{d^2X(t_2)}{dt_2^2} \frac{d^2X(t_1)}{dt_1^2} \frac{d^2X(t_1)}{dt_1^2}. \quad (28)$$

Integration over the angle χ can be done explicitly, whereas integration over the angle φ has to be done numerically.

This is our final result. It gives the second correlation function of radiation emitted by the atom in a Rydberg state. The formula is approximated and valid for small time differences $t_2 - t_1$ because it does not take radiation damping into account. It is valid only in the case of Rydberg states with large n and large l , with the maximal magnetic quantum number $m = l$.

An example of the second-order correlation function is shown in Fig. 2. One can notice very strong correlation of radiation for small times — much smaller than the period of motion — and the periodic behavior of the correlation.

7. Conclusions

Electromagnetic radiation from an atom in the Rydberg state can be used to partially localize the electron on the orbit. According to the classical view of radiation, the electron moves along an elliptic orbit and emits radiation most efficiently when the acceleration is large. This happens when the electron is close to the nucleus. The quantum wavefunction $\psi(r, \theta, \varphi)$ describing the electron state does not indicate the time when the electron is close to the nucleus. Therefore the emitted radiation is

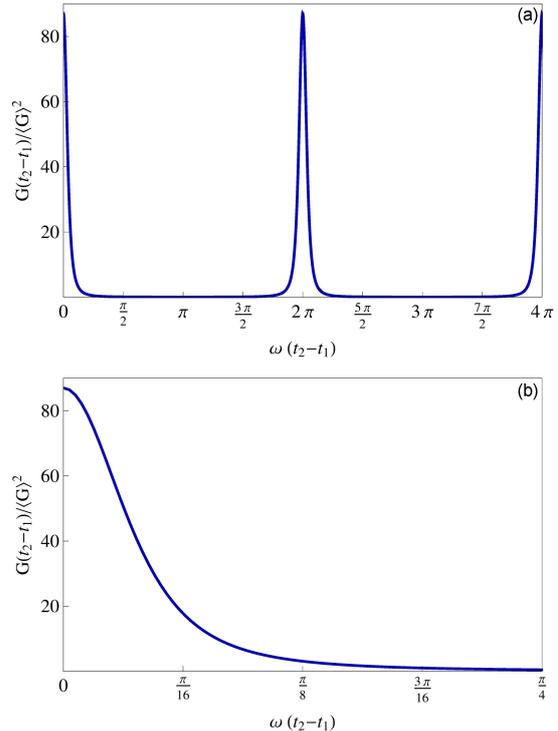


Fig. 2. The second-order correlation function of accelerations (27) (normalized to the square of the average square of acceleration) as a function of $\omega(t_2 - t_1)$ for one period ($\epsilon = 0.8$). Panel (b) shows the same for smaller values of time difference.

time-dependent, and its period reflects the period of motion. The time-averaged intensity, as well as the spectrum of radiation, is constant in time (for a relatively short time; radiation damping is not taken into account). The second correlation function, $G(t_2, t_1)$, depends on the time difference $t_2 - t_1$ and is a periodic function of time, with the frequency of the classical electron motion.

In the quantum language, the atom is in a highly excited Rydberg state with the principal quantum number n . The state is stationary, therefore, the average intensity of emitted radiation is constant in time. The spectrum is stationary since radiation damping is neglected, and consists of several narrow lines corresponding to the transition to lower energy states. The second correlation function, however, breaks the time translation symmetry, and this unravels the time evolution of radiation. Based on the measurement of radiation, we can reconstruct the motion of the electron.

The second correlation function was found in the classical approximation, however, its meaning is indeed purely quantum. The classical approximation means that transition matrix elements have been approximated by the corresponding classical expression. If exact expressions for the matrix elements had been used, the result would have been very similar. The calculations would have been numerically more complex.

We have to stress that electron localization is limited by the uncertainty principle. Thus in the case of the state with orbital quantum number l , the angle localization is possible up to $2\pi/l$. While in the case of large l considered here, this is not a strong limitation, it does play a significant role in the case of l of the order of 1, even for states with large principle quantum numbers n .

Our results show that the correlation function is strongly time-dependent. This correlation function clearly shows that if a strong and short impulse of radiation is detected, the next such pulse will come after one period of the corresponding classical motion, or in the quantum language, after time $T = 2\pi/(E_n - E_{n-1})$. This is due to the large acceleration of an electron in the vicinity of the nucleus.

The first strong pulse localizes the electron at this point and breaks the time independence of the radiation. The second pulse comes after one period. Between the strong pulses, the radiation is much weaker because of the small acceleration. Thus the observation of the time dependence of radiation allows the localization of the Rydberg electron in the vicinity of the nucleus.

This method of localizing an electron on the orbit is non-standard. The recent approach to quantum particle localization is based on successive measurements of a single particle. Measurement means entangling the particle with another system — a pointer — and then the measurement of the pointer state. In the present approach, the electromagnetic field serves as the pointer. The electron position is not measured directly — remember the dipole approximation — the electron acceleration is being measured. Obviously, the second-order correlation gives a deeper insight into the dynamics than the average values of observables. Also, it provides some insight into the measurement process in quantum mechanics, due to which the difficult process of position measurement is replaced by a standard measurement of radiation.

We have to point out that the approach described in this paper does not discuss the probabilities of single measurements, but rather it discusses averages such as a correlation function. Nevertheless, it is a possible way of detecting the motion of an electron along a trajectory.

Acknowledgments

JM dedicates this paper to Professor Iwo Białynicki-Birula, who taught me quantum mechanics and for many years guided me through quantum physics.

References

- [1] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University, Princeton (NJ) 1955.
- [2] W.H. Zurek, *Rev. Mod. Phys.* **75**, 715 (2003).
- [3] E.P. Wigner, *Am. J. Phys.* **31**, 6 (1963).
- [4] C.M. Caves, G.J. Milburn, *Phys. Rev. A* **36**, 5543 (1987).
- [5] A.J. Scott, G.J. Milburn, *Phys. Rev. A* **63**, 042101 (2001).
- [6] E.B. Davies, *Commun. Math. Phys.* **15**, 277 (1969).
- [7] M.B. Mensky, *Phys. Rev. D.* **20**, 384 (1979).
- [8] A. Barchielli, M. Gregoratti, *Phil. Trans. R. Soc. A* **370**, 5364 (2012).
- [9] R.J. Glauber, *Phys. Rev.* **131**, 2766 (1963).
- [10] J.R. Klauder, B. Skagerstam, *Coherent States*, World Scientific, Singapore 1985.
- [11] J. Mostowski, *Lett. Math Phys.* **2**, 1 (1977).
- [12] M. Nauenberg, *Phys. Rev. A* **40**, 1133 (1989).
- [13] A. Rauch, J. Parisi, *Adv. Stud. Theor. Phys.* **8**, 889 (2014).
- [14] J.C. Gay, D. Delande, A. Bommier, *Phys. Rev. A* **39**, 6587 (1989).
- [15] C.C. Gerry, *Phys. Rev. A* **33**, 6 (1986).
- [16] F. Wilczek, *Phys. Rev. Lett.* **109**, 160401 (2012).
- [17] K. Sacha, J. Zakrzewski, *Rep. Prog. Phys.* **81**, 016401 (2018).
- [18] K. Sacha, *Time Crystals, Springer Series on Atomic, Optical, and Plasma Physics*, Springer, Cham 2020, p. 114.
- [19] F. Gampel, M. Gajda, *Phys. Rev. A* **107**, 012420 (2023).
- [20] P. Kowalczyk, C. Radzewicz, J. Mostowski, I.A. Walmsley, *Phys. Rev. A* **42**, 5622 (1990).
- [21] L.D. Landau, E.M. Lifshitz, *Classical Theory of Fields*, Pergamon Press, 2000.
- [22] L.D. Landau, E.M. Lifshitz, *Quantum Mechanics*, Pergamon Press, 2000.
- [23] A. Pitak, J. Mostowski, *Eur. J. Phys.* **39**, 025402 (2018).

DEDICATED TO PROFESSOR IWO BIAŁYŃICKI-BIRULA ON HIS 90TH BIRTHDAY

— ◊ —

Fractality of Certain Quantum States

D.K. WÓJCIK^{a,b,*} AND K. ŻYCZKOWSKI^{c,d}^a*Nencki Institute of Experimental Biology of Polish Academy of Sciences, Pasteura 3, 02-093 Warsaw, Poland*^b*Institute of Applied Psychology, Jagiellonian University, Łojasiewicza 4, 30-348 Kraków, Poland*^c*Institute of Theoretical Physics, Jagiellonian University, Łojasiewicza 11, 30-348 Kraków, Poland*^d*Center for Theoretical Physics, Polish Academy of Sciences, al. Lotników 32/46, 26-680 Warszawa, Poland*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S146](https://doi.org/10.12693/APhysPolA.143.S146)*e-mail: d.wojcik@nencki.edu.pl

Fractal structures appearing in solutions of certain quantum problems are investigated. We prove the previously announced results concerning the existence and properties of fractal states for the Schrödinger equation in the infinite one-dimensional well. In particular, we show that for this problem, there exist solutions in the form of *fractal quantum carpets*: the probability density $P(x, t)$ forms a fractal surface with dimension D_{xy} , while its cross-sections $P_t(x)$ and $P_x(t)$ typically form fractal graphs with dimensions D_x and D_t respectively, where $D_{xy} = 2 + D_x/2$ and $D_t = 1 + D_x/2$ (almost everywhere).

topics: quantum carpets, quantum fractals, fractal dimension, fractal curves

1. Introduction

Fractals are sets and measures of non-integer dimension [1, 2]. They are good models of phenomena and objects in various areas of science. Their ubiquity in dynamical systems theory as attractors, repellers, and attractor boundaries is well-known [2, 3]. They are often connected with non-equilibrium problems of growth [4] and transport [5, 6]. Fractal properties of hydrodynamic modes have been shown to be connected with transport coefficients [7, 8]. Fractal dimensions are used in many nonlinear time series analysis methods [9, 10].

Fractals have also been found in quantum mechanics [11–14]. For instance, quantum models related to the problem of chaotic scattering often reveal fractal structures [15–17] relevant for quantum transport [18]. Fractal structures play a prominent role in studies of the quantum dynamics of a reduced density operator [19]. Spectroscopic characterization of the electronic wave function inside a confined structure with fractal geometry was discussed in [20]. Quantum field theories in fractal spacetimes were also analyzed [21, 22], and fractal

structures were reported in models of quantum gravity [23, 24]. Fractional calculus was found useful to describe the dynamics of quantum particles [25], while a Bohmian approach to quantum fractals was presented in [26].

It was also shown that the Schrödinger equation for the simplest non-chaotic potentials admits fractal solutions [27]. The resulting probability distributions $P(x, t)$ as functions of space and time, called *quantum carpets* [28–31], reveal fractal features [27, 32, 33]. In this paper, we have two objectives. One is to present the rigorous proofs of fractality of quantum states reported in [32]. The other is to illustrate a convenient method of calculating the dimensions of graphs of continuous functions introduced by Claude Tricot [34].

2. Methods

2.1. Box-counting dimension

In this section, we recall several equivalent definitions of box-counting dimension, state a criterion for finding the dimension of a continuous function

of one variable, and prove a connection between the dimension of the graph of a function of n variables and the dimensions of its sections. All of this is known, perhaps with the exception of Theorem 2, which might be new. We concentrate on the theory of box-counting dimension for graphs of continuous functions of one variable. More general theory and a deeper presentation can be found, for instance, in [2, 34, 35].

Let $A \subset \mathbb{R}^n$ be bounded. Consider a grid of n -dimensional boxes of side δ

$$[m_1\delta, (m_1+1)\delta] \times \cdots \times [m_n\delta, (m_n+1)\delta]. \quad (1)$$

Let $N(\delta)$ be the number of these boxes covering the set A . It is always finite because A is bounded.

Definition 1. Box-counting dimension of the set A is the limit

$$\dim_B(A) := \lim_{\delta \rightarrow 0} \frac{\ln [N(\delta)]}{\ln [1/\delta]}. \quad (1)$$

If the limit does not exist, one considers upper and lower box-counting dimensions

$$\overline{\dim}_B(A) := \limsup_{\delta \rightarrow 0} \frac{\ln [N(\delta)]}{\ln [1/\delta]}, \quad (3)$$

$$\underline{\dim}_B(A) := \liminf_{\delta \rightarrow 0} \frac{\ln [N(\delta)]}{\ln [1/\delta]}, \quad (4)$$

which always exist and satisfy

$$\overline{\dim}_B(A) \geq \underline{\dim}_B(A). \quad (5)$$

The box-counting dimension exists if the upper and lower box-counting dimensions are equal.

Several equivalent definitions are in use (see [1, 2, 34, 35] for a review). The most convenient definition to study the fractal properties of graphs of continuous functions is given in terms of δ -variations [34]. It is essentially a variant of the Bouligand definition [36]. We shall restrict our attention to dimensions of curves being subsets of a plane.

Let $K_\delta(x)$ be a closed ball $\{y \in \mathbb{R}^2 : |x - y| \leq \delta\}$.

Definition 2. *Minkowski sausage* or δ -parallel body of $A \subset \mathbb{R}^2$ is

$$A_\delta := \bigcup_{x \in A} K_\delta(x) = \{y \in \mathbb{R}^2 : \exists x \in A, |x - y| \leq \delta\}. \quad (6)$$

Thus the Minkowski sausage of A is the set of all the points located within δ of A .

Proposition 1. The box-counting dimension of a set $A \subset \mathbb{R}^2$ satisfies

$$\dim_B(A) = \lim_{\delta \rightarrow 0} \left(2 - \frac{\ln [V(A_\delta)]}{\ln [\delta]} \right), \quad (7)$$

where $V(\delta) = \text{vol}^2(A_\delta)$ is the area of the Minkowski sausage of A .

Proof. Every square from the δ -grid containing $x \in A$ is included in $K_{\sqrt{2}\delta}(x)$. On the other hand, every closed ball of radius $\sqrt{2}\delta$ can be covered by at most 16 squares from the grid. Therefore,

$$\delta^2 N(\varepsilon) \leq V(A_{\sqrt{2}\delta}) \leq 16\delta^2 N(\varepsilon). \quad (8)$$

□

Consider a continuous function on a closed interval $f : [a, b] \rightarrow \mathbb{R}$. Its graph is a curve in the plane. To find its box-counting dimension, estimate the number of boxes $N(\delta)$ intersecting the graph. Choose column $\{(x, y) : x \in [n\delta, (n+1)\delta]\}$. Since the curve is continuous, the number of the boxes in this column intersecting the graph of f is at least

$$\frac{1}{\delta} \left[\sup_{x \in [n\delta, (n+1)\delta]} f(x) - \inf_{x \in [n\delta, (n+1)\delta]} f(x) \right] \quad (9)$$

and no more than the same plus 2. If f was a record of a signal, then the difference between the maximum and minimum value of f on the given interval quantifies how the signal oscillates on this interval. That's why it is called δ -oscillation.

Definition 3. δ -oscillation of f at x is

$$\begin{aligned} \text{osc}_\delta(x)(f) &:= \sup_{|y-x| \leq \delta} f(y) - \inf_{|y-x| \leq \delta} f(y) = \\ &\sup \{|f(y) - f(z)| : y, z \in [a, b] \cap [x - \delta, x + \delta]\}. \end{aligned} \quad (10)$$

We will skip (f) if it is clear from the context which function we consider.

From (9) we obtain the following estimate on the total number of boxes covering the graph of f ,

$$\sum_{m=1}^M \frac{\text{osc}_{\delta/2}(x_m)}{\delta} \leq N(\delta) \leq 2M + \sum_{m=1}^M \frac{\text{osc}_{\delta/2}(x_m)}{\delta}, \quad (11)$$

where $x_m = a + (m - \frac{1}{2})\delta$ is the middle of the m -th column from the cover of the graph and $M = \lceil \frac{b-a}{\delta} \rceil$ is the number of columns in the cover ($\lceil x \rceil$ stands for the smallest integer greater or equal to x). Thus,

$$N(\delta) \approx M \overline{\text{osc}}_{\delta/2} / \delta. \quad (12)$$

If the graph of f has the box-counting dimension D , $N(\delta)$ scales as δ^{-D} . This implies the following scaling of the oscillations

$$\overline{\text{osc}}_{\delta/2} \approx N(\delta) \delta / M \propto \delta^{2-D}. \quad (13)$$

We have thus suggested a connection between the box-counting dimension of the graph and the scaling exponent of the average oscillation of the function f .

Definition 4. δ -variation of function f is

$$\text{Var}_\delta(f) := \int_a^b dx \text{osc}_\delta(x)(f) =: (b-a) \overline{\text{osc}}_\delta(f). \quad (14)$$

Geometrically, variation is the area of the set scanned by the graph of f moved horizontally $\pm\delta$ and truncated at $x = a$ and $x = b$, thus, it is a kind

of Minkowski sausage constructed with horizontal intervals of length 2δ . This observation leads to a convenient technique for calculating dimensions.

Theorem 1. Let $f(x)$ be a non-constant continuous function on $[a, b]$, then

$$\dim_B \text{graph} f = \lim_{\delta \rightarrow 0} \left(2 - \frac{\ln [\text{Var}_\delta(f)]}{\ln[\delta]} \right). \quad (15)$$

The proof consists of showing equivalence of $\text{Var}_\delta(f)$ with the Minkowski sausage and follows from inequality ([34], p. 130–132, 148–149)

$$\text{Var}_\delta(f) \leq V(A_\delta) \leq c \text{Var}_\delta(f), \quad (16)$$

where

$$\begin{aligned} A &= \text{graph} f, \\ c &= c_1 + c_2/s, \\ s &= \left[\sup_{x \in [a, b]} f(x) - \inf_{x \in [a, b]} f(x) \right]. \end{aligned} \quad (17)$$

This is where the assumption of non-constancy of f comes in. Derivation of (16) is not difficult but rather lengthy and will be omitted.

This theorem is the main tool to prove Theorem 3 in Sect. 3. In order to find the dimensions, we will look for estimates of δ -variation. They will usually take the following form:

Proposition 2

1. $\text{osc}_\delta(x)f(x) \leq c\delta^{2-s} \Rightarrow \dim_B \text{graph} f \leq s$.
2. $W := \int_a^b dx |f(x+\delta) - f(x-\delta)| \geq c\delta^{2-s} \Rightarrow \dim_B \text{graph} f \geq s$.

Proof.

1. $\text{Var}_\delta f = \int_a^b dx \text{osc}_\delta(x)(f) \leq (b-a)c\delta^{2-s}$.
2. $\text{osc}_{2\delta}(x)f \geq |f(x+\delta) - f(x-\delta)| \Rightarrow \text{Var}_\delta f \geq (b-a)c(\delta/2)^{2-s}$. \square

To prove the last point of Theorem 3, we need to know what is the dimension of the graph of $f: \mathbb{R}^n \rightarrow \mathbb{R}$, given all the dimensions of its one-variable restrictions.

Theorem 2. Let $f \in C^0([a_1, b_1] \times \dots \times [a_n, b_n])$. For every point $x = (x^1, \dots, x^n) \in [a_1, b_1] \times \dots \times [a_n, b_n]$ define $\tilde{x}^i := (x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^n)$. Then

$$f_i[\tilde{x}_0^i](x^i) := f(x_0^1, \dots, x_0^{i-1}, x^i, x_0^{i+1}, \dots, x_0^n) \quad (18)$$

is a restriction of f to a line parallel to i -th axis going through x_0 and $f_i[\tilde{x}_0^i] \in C^0([a_i, b_i])$.

(1) If $\forall x: \text{osc}_\delta f_i[\tilde{x}^i] \leq c_i \delta^{H_i}$, then

$$\dim_B \text{graph} f(x^1, \dots, x^n) \leq n+1 - \min\{H_1, \dots, H_n\}. \quad (19)$$

(2) If $\text{Var}_\delta f_i[\tilde{x}_0^i] \geq c_i \delta^{H_i}$ for a dense set $\tilde{x}_0^i \in A \subset \bar{A} = [a_1, b_1] \times \dots \times [a_{i-1}, b_{i-1}] \times [a_{i+1}, b_{i+1}] \times \dots \times [a_n, b_n]$, then

$$\dim_B \text{graph} f(x^1, \dots, x^n) \geq n+1 - \min\{H_1, \dots, H_n\}. \quad (20)$$

(3) If all of the above conditions are satisfied, then $\dim_B \text{graph} f(x^1, \dots, x^n) = n+1 - \min\{H_1, \dots, H_n\} =$

$$n-1 + \max\{s_1, \dots, s_n\}, \quad (21)$$

where $s_i = \sup_{\tilde{x}^i} \dim_B \text{graph} f_i[\tilde{x}^i](x^i)$.

In other words, the strongest oscillations along any direction determine the box-counting dimension of the whole $n+1$ -dimensional graph.

Proof. We will show the theorem for $n = 2$ for notational simplicity. Generalization to arbitrary n is immediate. Let $f: [a_1, b_1] \times [a_2, b_2] \rightarrow \mathbb{R}$. Divide the domain into squares $X_i \times Y_j$ of side δ . This gives rise to K columns A_{ij} of δ -grid in \mathbb{R}^3 , $1 \leq \frac{K\delta^2}{(b_1-a_1)(b_2-a_2)} \leq 2$.

(1). The number of δ -cubes having a common point with the graph of f in column A_{ij} is not greater than $\frac{1}{\delta}(\sup_{A_{ij}} f - \inf_{A_{ij}} f) + 2$. But

$$\begin{aligned} |f(x_1, y_1) - f(x_2, y_2)| &= \\ |f(x_1, y_1) - f(x_1, y_2) + f(x_1, y_2) - f(x_2, y_2)| & \\ \leq |f(x_1, y_1) - f(x_1, y_2)| + |f(x_1, y_2) - f(x_2, y_2)|. & \end{aligned} \quad (22)$$

Therefore

$$\begin{aligned} \sup_{A_{ij}} f - \inf_{A_{ij}} f &= \sup_{(x_1, y_1), (x_2, y_2) \in A_{ij}} |f(x_1, y_1) - f(x_2, y_2)| \\ &\leq \sup_{x \in X_i} \sup_{y \in Y_j} f(x, y) + \sup_{y \in Y_j} \sup_{x \in X_i} f(x, y) \\ &\leq \sup_{x \in X_i} \text{osc}_{\delta/2} f_1[x] + \sup_{y \in Y_j} \text{osc}_{\delta/2} f_2[y] \\ &\leq c \delta^{\min\{H_1, H_2\}}. \end{aligned} \quad (23)$$

Thus

$$\begin{aligned} \dim_B \text{graph} f(x^1, x^2) &\leq \lim_{\delta \rightarrow 0} \frac{\ln[Kc\delta^{\min\{H_1, H_2\}}/\delta]}{\ln[1/\delta]} \\ &\leq 3 - \min\{H_1, H_2\}. \end{aligned} \quad (24)$$

(2). Set $x \in X_i$. From (8) and (16) it follows that the number $N_i(\delta)$ of δ -cubes in columns A_{ij} covering the graph of $f_2[x](y)$ and the variation of f_2 satisfy

$$\text{Var}_\delta f_2[x] \leq c\delta^2 N_i(\delta). \quad (25)$$

Thus

$$N_i(\delta) \geq c \sup_{x \in X_i} \text{Var}_\delta f_2[x] / \delta^2 \geq c \delta^{H_2-2}. \quad (26)$$

Therefore, the number $N(\delta)$ of boxes covering the whole graph of f satisfies

$$N(\delta) \geq c \sum_{i=1}^M \sup_{x \in X_i} \text{Var}_\delta f_2[x] / \delta^2 \geq c_1 \delta^{H_2-3}. \quad (27)$$

The same can be repeated for any direction, thus

$$N(\delta) \geq c_2 \delta^{\min\{H_1, H_2\}-3}. \quad (28)$$

(3). An immediate corollary.

Generalization to arbitrary n is achieved by observing that $K\delta^n \approx \text{const}$. \square

Another definition, which has some convenient technical properties, is the Hausdorff dimension [37–39], however, it is often too difficult to calculate. For instance, as far as we know, there is still no proof that the Hausdorff dimension of the Weierstrass function is equal to its box-counting dimension. Thus in practice, one usually uses the (upper) box-counting dimension. This is also our present approach. It is often assumed that the box-counting dimension and the Hausdorff dimension are equal. A general characterization of situations when this conjecture really holds is also lacking.

2.2. Fractal functions

One of the oldest fractals is a graph of the Weierstrass function [40, 41]

$$W(x) = \sum_{n=0}^{\infty} a^n \cos(b^n x \pi), \quad (29)$$

introduced as an example of an everywhere continuous, nowhere differentiable function by Karl Weierstrass around 1872. The maximum range of parameters, for which the above series has required properties was found by Godfrey Harold Hardy in 1916 [42], who also showed that

$$\sup \{|f(x) - f(y)| : |x - y| \leq \delta\} \sim \delta^H, \quad (30)$$

where $H = \frac{\ln(1/a)}{\ln(b)}$. From this it easily follows (see below) that the box-counting dimension of the graph of the Weierstrass function $W(x)$ is

$$D_W = 2 + H = 2 + \frac{\ln(a)}{\ln(b)} = 2 - \left| \frac{\ln(a)}{\ln(b)} \right|. \quad (31)$$

Functions whose graphs have non-integer box-counting dimension are called *fractal functions*. Even though the box-counting dimension of the Weierstrass function is easy to calculate [34], the proof that its Hausdorff dimension has the same value is still lacking, as far as we know. Lower bounds on the Hausdorff dimension of the graph were found by Mauldin [43, 44]. Graphs of random Weierstrass functions were shown to have the same Hausdorff and box-counting dimensions for almost every distribution of phases [45].

3. Results

The construction of the Weierstrass function, (29), can easily be realized in quantum mechanics. Consider solutions of the Schrödinger equation

$$i\partial_t \Psi(x, t) = -\nabla^2 \Psi(x, t) \quad (32)$$

for a particle in the one-dimensional infinite potential well. The general solutions satisfying the boundary conditions $\Psi(0, t) = 0 = \Psi(\pi, t)$ have the form

$$\Psi(x, t) = \sum_{n=1}^{\infty} a_n \sin(nx) e^{-in^2 t}, \quad (33)$$

where

$$a_n = \frac{2}{\pi} \int_0^\pi dx \sin(nx) \Psi(x, 0). \quad (34)$$

Weierstrass quantum fractals are wave functions of the form

$$\Psi_M(x, t) = N_M \sum_{n=0}^M q^{n(s-2)} \sin(q^n x) e^{-iq^{2n} t}, \quad (35)$$

where $q = 2, 3, \dots$, $s \in (0, 2)$.

In the physically interesting case of finite M , the wave function Ψ_M is a solution of the Schrödinger equation. The limiting case

$$\Psi(x, t) := \lim_{M \rightarrow \infty} \Psi_M(x, t) = N \sum_{n=0}^{\infty} q^{n(s-2)} \sin(q^n x) e^{-iq^{2n} t}, \quad (36)$$

with the normalization constant $N = \sqrt{\frac{2}{\pi} \sqrt{1 - q^{2(s-2)}}$, is continuous but nowhere differentiable. It is a weak solution of the Schrödinger equation. Note that (36) converges for $(|q^{s-2}| < 1 \equiv s < 2)$. Since the probability density of wave function (36) shows fractal features for $s > 0$ (see below), the interesting range of s is $(0, 2)$.

The main results announced in [32], which we prove here, are that not only the real part of the wave function $\Psi(x, t)$, but also the physically important probability density $P(x, t) := |\Psi(x, t)|^2$ exhibit fractal nature. This is not obvious, because $|\Psi(x, t)|^2$ is the sum of squares of the real and imaginary parts having usually equal dimensions. One can easily show that the dimension of the graph of a sum of functions whose graphs have the same dimensions D can be anything[†] from 1 to D .

Our main results are given by Theorem 3.

Theorem 3. Let $P(x, t)$ denote the probability density of a Weierstrass-like wave function (36). Then

1. at the initial time $t = 0$, the probability density $P_0(x) = P(x, 0)$ forms a fractal graph in the space variable (i.e., space fractal) of dimension $D_x = \max\{s, 1\}$;
2. the dimension D_x of graph of $P_t(x) = P(x, t = \text{const})$ does not change in time;
3. for almost every x inside the well, the probability density, $P_x(t) = P(x = \text{const}, t)$, forms a fractal graph in the time variable (i.e., time fractal) of dimension $D_t(x) = D_t := 1 + s/2$;
4. for a discrete, dense set of points x_d , $P_{x_d}(t) = P(x_d, t)$ is smooth, and thus $D_t(x_d) = 1$;
5. for even q , the average velocity $\frac{d\langle x \rangle}{dt}(t)$ is fractal with the dimension of its graph equal to $D_v = \max\{(1 + s)/2, 1\}$;
6. the surface $P(x, t)$ has dimension $D_{xy} = 2 + s/2$.

[†]Let f_1 and f_2 be functions with graphs having dimensions, respectively, $1 \leq D_1 < D_2 \leq 2$. Let $g_1 = f_1 + f_2$, $g_2 = f_1 - f_2$. Then the box-counting dimension of both the graph of g_1 and g_2 is D_2 , but the dimension of the graph of their sum $g_1 + g_2 = 2f_1$ is $D_1 \in [1, D_2]$.

The physical meaning of Theorem 3 has been discussed in [32]. Here we only emphasize that to generate a fractal wave function with exact mathematically rigorous fractal features with infinite scaling properties, infinite energy is required. However, even a few terms in the series defining the function (36) can lead to physically interesting effects.

Our proof of Theorem 3 is based on the power-law behavior of the average δ -oscillation of the infinite double sum present in $P(x, t) = |\Psi(x, t)|^2$ (see (122) in Appendix). Some fundamental concepts and facts used in the proof are given in Sect. 2.1. Calculations of probability density and average velocity are provided in the Appendix. Positive real constants are denoted by c, c_1, c_2, \dots

4. Proof of Theorem 3

1. At the initial time $t = 0$, the probability density, $P_0(x) = P(x, 0)$, forms a fractal graph in the space variable (i.e., space fractal) of dimension $D_x = \max\{s, 1\}$.
2. The dimension D_x of graph of $P_t(x) = P(x, t = \text{const})$ does not change in time.

We will show that for every fixed t , the graph of the probability density $|\Psi|^2$ (see (122) in Appendix) as a function of x has the box-counting dimension s .

(a) Fix t . Let

$$P_n(x) := \sum_{k=0}^n q^{k(s-2)} \sum_{l=0}^k \sin(q^l x) \sin(q^{k-l} x) \times \cos[(q^{2l} - q^{2(k-l)})t], \quad (37)$$

$q = 2, 3, \dots$ It is a smooth function whose derivative at every point satisfies

$$|P'_n(x)| \leq 2 \sum_{k=0}^n q^{k(s-2)} \sum_{l=0}^k q^l \left| \cos(q^l x) \sin(q^{k-l} x) \right| \leq 2 \sum_{k=0}^n \frac{q^{k+1}}{q-1} q^{k(s-2)} \leq d_1(s, q) q^{n(s-1)}, \quad (38)$$

where

$$d_1(s, q) = \frac{2q^s}{(q-1)(q^{s-1}-1)}. \quad (39)$$

Let $\delta = q^{-n}$. Then

$$\text{osc}_\delta(x) P_n \leq 2\delta \sup_{x \in [0, \pi]} |P'_n(x)| \leq 2d_1(s, q) \delta^{2-s}. \quad (40)$$

On the other hand, for

$$R_n(x) := P(x) - P_n(x) = \sum_{k=n+1}^\infty q^{k(s-2)} \times \sum_{l=0}^k \sin(q^l x) \sin(q^{k-l} x) \cos[(q^{2l} - q^{2(k-l)})t], \quad (41)$$

we have

$$\text{osc}_\delta(x) R_n \leq 2 \sum_{k=n+1}^\infty q^{k(s-2)} (k+1) \leq \frac{4q^{(n+1)(s-2)} n}{(1 - q^{s-2})^2}. \quad (42)$$

Polynomial growth is slower than exponential, therefore for arbitrarily small ε there is some M such that $\forall n > M : n < (q^\varepsilon)^n$. This leads to the following estimate of the oscillation of R_n ,

$$\text{osc}_\delta(x) R_n \leq d_2(s, q) \delta^{2-s-\varepsilon}, \quad (43)$$

where

$$d_2(s, q) = \frac{4q^{s-2}}{(1 - q^{s-2})^2}. \quad (44)$$

Thus for all x and $\delta = q^{-n}$, where $\frac{\ln(n)}{n} < \varepsilon \ln(q)$, we have

$$\text{osc}_\delta(x) P \leq \text{osc}_\delta(x) P_n + \text{osc}_\delta(x) R_n \leq (2d_1 + d_2) \delta^{2-s-\varepsilon}. \quad (45)$$

From Proposition 2 it follows that $\dim_B \text{graph } P_t(x) \leq 2 - (2-s-\varepsilon) = s+\varepsilon \rightarrow_{\varepsilon \rightarrow 0} s$.

(b) Fix t . Let $f(x) = P(x, t)$. We want to show that

$$W := \int_a^b dx |f(x+\delta) - f(x-\delta)| \geq c \delta^{2-s}. \quad (46)$$

Take $a=0, b=\pi$. Notice that (we skip the normalization constant)

$$W = \int_0^\pi dx |f(x+\delta) - f(x-\delta)| = \int_0^\pi dx \left| \sum_{k=0}^\infty q^{k(s-2)} \sum_{l=0}^k \left\{ \sin[q^l(x+\delta)] \sin[q^{k-l}(x+\delta)] - \sin[q^l(x-\delta)] \sin[q^{k-l}(x-\delta)] \right\} a_{kl} \right| = \int_0^\pi dx \left| \sum_{k=0}^\infty q^{k(s-2)} \sum_{l=0}^k \left\{ \cos(q^l x) \sin(q^{k-l} x) \sin(q^l \delta) \cos(q^{k-l} \delta) \right\} a_{kl} \right|, \quad (47)$$

where $a_{kl} = \cos[(q^{2k} - q^{2l})t]$. Take $|h(x)| \leq 1$. Observe that

$$\int_a^b dx \left| \sum_i f_i(x) \right| \geq \int_a^b dx |h(x)| \left| \sum_i f_i(x) \right| \geq \left| \int_a^b dx \sum_i h(x) f_i(x) \right| \geq \left| \int_a^b dx h(x) f_k(x) \right| - \sum_{i \neq k} \left| \int_a^b dx h(x) f_i(x) \right|. \quad (48)$$

One can interchange the order of summation and integration because $f(x)$ is absolutely convergent. Let us take $\delta=q^{-N}$, $h(x) = \sin(q^m x) \cos(q^n x)$. After substitution in (47) using (48) we obtain

$$W \geq \left| \sum_{k=0}^{\infty} q^{k(s-2)} \sum_{l=0}^k \sin(q^{l-N}) \sin(q^{k-l-N}) \int_0^{\pi} dx \sin(q^l x) \cos(q^{k-l} x) \cos(q^m x) \cos(q^n x) a_{kl} \right| =$$

$$\frac{\pi}{4} q^{(m+n)(s-2)} \left| \cos(q^{m-N}) \sin(q^{n-N}) \cos[(q^{2(m+n)} - q^{2m})t] \right| =: \frac{\pi}{4} \widetilde{W}. \quad (49)$$

We will now prove that $\exists c : \widetilde{W} = q^{(m+n)(s-2)} \times |\cos(q^{m-N}) \sin(q^{n-N}) \cos[(q^{2(m+n)} - q^{2m})t]| \geq cq^{N(s-2)}$, for arbitrary real t . We will take advantage of the fact that q is an integer.

Let $N = m + n$. Then $\widetilde{W} = q^{N(s-2)} \times |\cos(q^{m-N}) \sin(q^{-m}) \cos[(q^{2N} - q^{2m})t]|$. It is enough to consider $t \in [0, \pi]$.

Let us write t/π in q

$$\frac{t}{\pi} = \frac{a_1}{q} + \frac{a_2}{q^2} + \frac{a_3}{q^3} + \dots = \sum_{k=1}^{\infty} \frac{a_k}{q^k}, \quad (50)$$

where $a_k \in \{0, 1, \dots, q-1\}$, so that t/π can be written as

$$\frac{t}{\pi} = 0.a_1 a_2 \dots a_K (a_{K+1} \dots a_{K+T}). \quad (51)$$

Therefore,

$$\cos[(q^{2N} - q^{2m})t] = \cos[\pi(q^{2N-1} a_1 + q^{2N-2} a_2 + \dots + a_{2N} + q^{-1} a_{2N+1} + \dots + q^{2m-1} a_1 + q^{2m-2} a_2 + \dots + a_{2m} + q^{-1} a_{2m+1} + \dots)] =$$

$$\cos\left[\pi\left(\frac{a_{2N+1} - a_{2m+1}}{q} + \frac{a_{2N+2} - a_{2m+2}}{q^2} + \dots\right)\right]. \quad (52)$$

If we could only choose m so that the first two terms in this series cancel out, we would have a lower estimate on the cosine, because, in this case,

$$\left| \frac{a_{2N+3} - a_{2m+3}}{q^3} + \dots \right| \leq (q-1) \left(\frac{1}{q^3} + \frac{1}{q^4} + \dots \right) = \frac{1}{q^2}. \quad (53)$$

Thus,

$$\cos[(q^{2N} - q^{2m})t] \geq \cos\left(\frac{\pi}{q^2}\right) \geq \cos\left(\frac{\pi}{4}\right). \quad (54)$$

Let A be the set of all the two-element sequences with elements from the set $\{0, 1, \dots, q-1\}$. Thus

$$A = \{\{0, 0\}, \{0, 1\}, \dots, \{0, q-1\}, \{1, 0\}, \dots, \dots, \{q-1, q-1\}\}, \quad (55)$$

and we write $A_{k,l} := \{k, l\}$, $k, l \in \{0, 1, \dots, q-1\}$. Consider all the pairs of consecutive q -digits of t/π of the form

$$\{a_{2m+1}, a_{2m+2}\}, \quad (56)$$

i.e., $\{a_1, a_2\}, \{a_3, a_4\}$ etc. Every such pair is equal to some $A_{k,l}$. Let $N_{k,l}$ be the first such m , for which

$$A_{k,l} = \{a_{2m+1}, a_{2m+2}\}. \quad (57)$$

If $A_{k,l}$ for given k, l doesn't appear in the sequence of all the pairs (56), we set $N_{k,l} = 0$. Let

$$M = \sup_{k,l} N_{k,l}. \quad (58)$$

Thus if $n > M$, the sequence $\{a_{2n+1}, a_{2n+2}\}$ has appeared at least once among the pairs $\{a_1, a_2\}, \{a_3, a_4\}, \dots, \{a_{2M+1}, a_{2M+2}\}$. Therefore, for every $N > M$ we can find such an $m \in 1, 2, \dots, M$ that

$$\left| \cos\left[\left(q^{2N} - q^{2m}\right) \frac{t}{\pi} \pi\right] \right| \geq \cos\left(\frac{\pi}{q^2}\right) \geq \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}. \quad (59)$$

Also

$$\sin(q^{-m}) \geq \sin(q^{-M}),$$

$$\cos(q^{m-N}) \geq \cos(q^{M-N}) \geq \cos(1), \quad (60)$$

which leads to

$$\widetilde{W} \geq \frac{\sqrt{2}}{2} q^{N(s-2)} \sin(q^{-M}) \cos(1) = \text{const } q^{N(s-2)}. \quad (61)$$

We have thus shown that for every t , for natural q , and for $\delta = q^{-N}$

$$W \geq \text{const } \delta^{2-s}, \quad (62)$$

therefore (Proposition 2)

$$\dim_B \text{graph } P_t(x) \geq 2 - (2-s) = s. \quad (63)$$

$$3. \text{ For almost every } x, \quad D_t(x) = \dim_B \text{graph } P_x(t) = D_t := 1 + s/2.$$

We will use the form (123) given in Appendix of the probability density. It is enough to analyze the dimension of

$$\widetilde{P}(t) := \sum_{c=1}^{\infty} q^{2c(s-2)} \sin(q^c x) \sum_{d=1}^c q^{-d(s-2)} \times \sin(q^{c-d} x) \cos[(q^{2c} - q^{2(c-d)})t]. \quad (64)$$

(a) Let

$$P_n(t) := \sum_{c=1}^n q^{2c(s-2)} \sin(q^c x) \sum_{d=1}^c q^{-d(s-2)} \times \sin(q^{c-d} x) \cos[(q^{2c} - q^{2(c-d)})t]. \quad (65)$$

Then

$$|P'_n(t)| = \left| \sum_{c=1}^n q^{2c(s-2)} \sin(q^c x) \sum_{d=1}^c q^{-d(s-2)} \sin(q^{c-d} x) (q^{2c} - q^{2(c-d)}) \sin[(q^{2c} - q^{2(c-d)})t] \right| \leq$$

$$\sum_{c=1}^n q^{2c(s-2)} \sum_{d=1}^c q^{-d(s-2)} (q^{2c} - q^{2(c-d)}) = \sum_{c=1}^n q^{2c(s-2+1)} \sum_{d=1}^c q^{-d(s-2)} (1 - q^{-2d}) =$$

$$\frac{q^{2-s}}{q^{2-s} - 1} \left[q^s \frac{q^{ns} - 1}{q^s - 1} - q^{2(s-1)} \frac{q^{2n(s-1)} - 1}{q^{2(s-1)} - 1} \right] - \frac{q^{-s}}{q^{-s} - 1} \left[q^{s-2} \frac{q^{n(s-2)} - 1}{q^{s-2} - 1} - q^{2(s-1)} \frac{q^{2n(s-1)} - 1}{q^{2(s-1)} - 1} \right]. \quad (66)$$

Therefore, for n large enough,

$$|P'_n(t)| \leq c_1 q^{n \max\{s, 2(s-1), s-2\}} = c q^{ns}. \quad (67)$$

Let $\delta = q^{-\alpha n}$. Then $q^n = \delta^{-1/\alpha}$ and

$$\text{osc}_\delta(t) P_n \leq 2c_1 \delta q^{ns} = 2c_1 \delta^{1-s/\alpha}. \quad (68)$$

Let

$$R_n(t) := \tilde{P}(t) - P_n(t). \quad (69)$$

Then

$$\text{osc}_\delta(t) R_n \leq 2 \sum_{c=n+1}^{\infty} q^{2c(s-2)} \sum_{d=1}^c q^{-d(s-2)}$$

$$\leq \frac{2q^{2-s}}{q^{2-s} - 1} \sum_{c=n+1}^{\infty} q^{2c(s-2)+c(2-s)} = c_2 \delta^{(s-2)/\alpha}. \quad (70)$$

To obtain a consistent estimate we must set

$$1 - \frac{s}{\alpha} = \frac{2}{\alpha} - \frac{s}{\alpha}, \quad (71)$$

which gives $\alpha = 2$. Thus

$$\text{osc}_\delta(t) \tilde{P} \leq (2c_1 + c_2) \delta^{1-s/2}. \quad (72)$$

(b) Now we want to show that

$$W = \int_a^b dt \left| \tilde{P}(t+\delta) - \tilde{P}(t-\delta) \right| \geq c \delta^{1-s/2}. \quad (73)$$

Set $a = 0, b = 2\pi$ for convenience. Then

$W =$

$$\int_0^{2\pi} dt \left| \sum_{c=1}^{\infty} q^{2c(s-2)} \sin(q^c x) \sum_{d=1}^c q^{-d(s-2)} \sin(q^{c-d} x) \left\{ \cos[(q^{2c} - q^{2(c-d)})(t+\delta)] - \cos[(q^{2c} - q^{2(c-d)})(t-\delta)] \right\} \right| =$$

$$\int_0^{2\pi} dt \left| \sum_{c=1}^{\infty} q^{2c(s-2)} \sin(q^c x) \sum_{d=1}^c q^{-d(s-2)} \sin(q^{c-d} x) \left\{ -2 \sin[(q^{2c} - q^{2(c-d)})t] \sin[(q^{2c} - q^{2(c-d)})\delta] \right\} \right|. \quad (74)$$

Using our standard arguments, we multiply the integrand by a suitable function smaller or equal to 1,

$$W \geq \int_0^{2\pi} dt |h(t)| \left| \tilde{P}(t+\delta) - \tilde{P}(t-\delta) \right| \geq$$

$$\left| \int_0^{2\pi} dt h(t) \left[\tilde{P}(t+\delta) - \tilde{P}(t-\delta) \right] \right|. \quad (75)$$

We choose $h(t) = \sin[(q^{2c} - q^{2(c-d)})t]$ and set $\delta = q^{-2N}$. It follows that

$$W \geq 2\pi q^{(2c-d)(s-2)}$$

$$\times \left| \sin(q^c x) \sin(q^{c-d} x) \sin(q^{2c} - q^{2(c-d)}) q^{-2N} \right|. \quad (76)$$

We now want to show that for almost all x

$$W \geq c_3 \delta^{1-s/2} = c_3 q^{N(s-2)}. \quad (77)$$

Set $2c-d = N$. Then

$$W \geq 2\pi q^{N(s-2)}$$

$$\times \left| \sin(q^c x) \sin(q^{N-c} x) \sin(q^{2(c-N)} - q^{-2c}) \right|. \quad (78)$$

Thus it is enough to bound

$$\left| \sin(q^c x) \sin(q^{N-c} x) \sin[q^{2(c-N)} - q^{-2c}] \right| \quad (79)$$

from below.

Choose rational x/π . All the rational numbers in a given basis q have finite or periodic expansion. In the first case ($x/\pi = k/q^l$), we cannot find the lower bound on (79). We cannot succeed, because at these points the function $P_x(t)$ is smooth (cf. the proof of Theorem 3.4).

The other case means that x/π can be written as

$$\frac{x}{\pi} = 0.a_1 a_2 \dots a_K (a_{K+1} \dots a_{K+T}), \quad (80)$$

where again $(a_{K+1} \dots a_{K+T})$ denotes the periodic part. Therefore, for every $N > K$, $q^n x \bmod \pi$ can take only one of T values: $q^{K+1} x \bmod \pi, \dots, q^{K+T} x \bmod \pi$. Let us take $c = 1$, $N > K$. Then $|\sin(q^c x)| = |\sin(qx)| > 0$ and is a constant. Note that $|\sin(q^{N-1} x)|$ takes one of T values, none of which is 0, therefore it is always bounded from below by

$$\inf_{l=1,2,\dots,T} |\sin(q^{K+l} x)| > 0. \quad (81)$$

Also, the last term can be bounded

$$\left| \sin(q^{-2(N-c)} - q^{-2c}) \right| \geq \sin(q^{-2} - q^{-2(N-1)}) \geq \sin(q^{-3}) \quad (82)$$

for $N \geq 3$. Thus for rational x/π with periodic expansion in q

$$W \geq 2\pi c_3 q^{N(s-2)}, \quad (83)$$

where $c_3 = |\sin(qx) \sin(\frac{x}{q^3})| \inf_{l=1,2,\dots,T} |\sin(q^{K+l} x)|$.

Consider now irrational x/π . Inequality (78) for $c = N$ takes the form

$$\begin{aligned} W &\geq 2\pi q^{N(s-2)} |\sin(q^N x) \sin(x) \sin(1 - q^{-2N})| \\ &\geq c q^{N(s-2)} |\sin(q^N x)|, \end{aligned} \quad (84)$$

for $N \geq 2$. Instead of showing it can be bounded from below, we will use it to prove that for almost every x

$$\dim_B \text{graph } P_x(t) \geq 1 + s/2. \quad (85)$$

Let

$$x_n := q^n (x/\pi) \bmod 1. \quad (86)$$

Let

$$F_N^\alpha := \left\{ x : \exists n \geq N \left(x_n \leq \frac{1}{q^{N\alpha}} \right) \vee \left(1 - x_n \leq \frac{1}{q^{N\alpha}} \right) \right\}, \quad (87)$$

where $\alpha \in [0, 1]$. Let

$$F_\infty^\alpha := \bigcap_{N=1} F_N^\alpha. \quad (88)$$

Clearly,

$$F_N^\alpha \supset F_{N+1}^\alpha \supset F_{N+2}^\alpha \dots \quad (89)$$

Since the Renyi map (86) preserves the Lebesgue measure, we have

$$\begin{aligned} \mu(F_N^\alpha) &\leq 2 \left(\frac{1}{q^{N\alpha}} + \frac{1}{q^{(N+1)\alpha}} + \frac{1}{q^{(N+2)\alpha}} + \dots \right) = \\ &= \frac{2q}{q-1} \frac{1}{q^{N\alpha}}. \end{aligned} \quad (90)$$

Therefore

$$0 \leq \mu(F_\infty^\alpha) \leq \inf_N \mu(F_N^\alpha) = 0. \quad (91)$$

It follows that for almost every $\{x_n\}$

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\ln |\sin x_n|}{n} &\geq \lim_{n \rightarrow \infty} \frac{\ln(q^{-n\alpha})}{n} \geq \\ \lim_{n \rightarrow \infty} \frac{q^{-n\alpha}}{2n} &\geq -\alpha \ln(q). \end{aligned} \quad (92)$$

Thus for every $\alpha > 0$ and for almost every $x/\pi \in [0, 1]$ we have

$$\begin{aligned} \dim_B \text{graph } P_x(t) &= \lim \left[2 - \frac{\ln(\text{Var}_\delta P_x(t))}{\ln(q^{-2N})} \right] \\ &\geq 2 + \lim \frac{\ln(\text{Var}_\delta P_x(t))}{2N \ln(q)}. \end{aligned} \quad (93)$$

But

$$\text{Var}_\delta P_x(t) \geq W, \quad (94)$$

therefore from (84)

$$\begin{aligned} \dim_B \text{graph } P_x(t) &\geq \\ &2 + \lim \frac{\ln(c) + N(s-2) \ln(q) + \ln |\sin(x_n \pi)|}{2N \ln(q)} = \\ &1 + \frac{s}{2} + \lim \frac{\ln |\sin(x_n \pi)|}{2N \ln(q)} \geq 1 + \frac{s}{2} - \alpha. \end{aligned} \quad (95)$$

But α is arbitrary, thus

$$\dim_B \text{graph } P_x(t) \geq 1 + s/2. \quad (96)$$

4. For a discrete, dense set of points x_d , $D_t(x_d) = \dim_B \text{graph } P_{x_d}(t) = 1$.

Let $x_{k,m} = \frac{m\pi}{q^k}$, where $k \in \mathbb{N}$, $m = 0, 1, \dots, q^k - 1$. The set $\{x_{k,m}\}$ is dense in $[0, 1]$. At these points, $\Psi(x_{k,m}, t)$ is a sum of a finite number of terms

$$\begin{aligned} \Psi \left(\frac{m\pi}{q^k}, t \right) &= \sqrt{\frac{2(1 - q^{-2(2-s)})}{\pi}} \sum_{n=0}^{k-1} q^{(s-2)n} \\ &\times \sin(q^{n-k} m\pi) e^{-i q^{2n} t}. \end{aligned} \quad (97)$$

Therefore,

$$\dim_B \text{graph} \left| \Psi \left(\frac{m\pi}{q^k}, t \right) \right|^2 = 1. \quad (98)$$

5. For even q the average velocity $\frac{d\langle x(t) \rangle}{dt}$ is fractal with the dimension of its graph equal to $D_v = \max\{(1+s)/2, 1\}$.

Heuristically, this is rather obvious, because

$$\begin{aligned} \frac{d\langle x(t) \rangle}{dt} &\approx \sum_{k=1}^{\infty} \frac{q^{k(s-1)}}{q^{2k}} \sin(q^{2k} t) = \\ &= \sum_{k=1}^{\infty} q^{2k(s-3)/2} \sin(q^{2k} t). \end{aligned} \quad (99)$$

Thus the average velocity is essentially a Weierstrass-like function and the dimension of its graph should be

$$2 - (3-s)/2 = (1+s)/2. \quad (100)$$

It is enough to consider

$$W(t) := \sum_{k=1}^{\infty} \frac{q^{k(s-1)}}{q^{2k} - 1} \sin[(q^{2k} - 1)t]. \quad (101)$$

(a) Let

$$W_n(t) := \sum_{k=1}^n \frac{q^{k(s-1)}}{q^{2k} - 1} \sin[(q^{2k} - 1)t]. \quad (102)$$

Set $\delta = q^{-\alpha n}$. Then

$$\begin{aligned} |W'_n(t)| &= \left| \sum_{k=1}^n q^{k(s-1)} \cos [(q^{2k}-1)t] \right| \\ &\leq \sum_{k=1}^n q^{k(s-1)} \leq c_1 \delta^{(1-s)/\alpha}, \end{aligned} \quad (103)$$

where

$$c_1 = \frac{q^{s-1}}{q^{s-1}-1}. \quad (104)$$

Therefore,

$$\text{osc}_\delta(t)W_n \leq 2c_1 \delta^{(1-s)/\alpha} \delta = 2c_1 \delta^{1+(1-s)/\alpha}. \quad (105)$$

Now, for

$$P_n(t) := W(t) - W_n(t), \quad (106)$$

we have

$$\begin{aligned} |P_n(t)| &= \left| \sum_{k=n+1}^{\infty} \frac{q^{k(s-1)}}{q^{2k}-1} \sin [(q^{2k}-1)t] \right| \\ &\leq \sum_{k=n+1}^{\infty} \frac{q^{k(s-1)}}{q^{2k}-1} \leq \sum_{k=n+1}^{\infty} \frac{2q^{k(s-1)}}{q^{2k}} = c_2 \delta^{-(s-3)/\alpha}, \end{aligned} \quad (107)$$

where

$$c_2 = \frac{2q^{s-3}}{1-q^{s-3}}. \quad (108)$$

Thus

$$\text{osc}_\delta(t)P_n \leq 2c_2 \delta^{(3-s)/\alpha}. \quad (109)$$

To obtain consistent estimates for both P_n and W_n we must set

$$1 + (1-s)/\alpha = (3-s)/\alpha, \quad (110)$$

thus $\alpha = 2$ and $\delta = q^{-2N}$. Therefore,

$$\begin{aligned} \text{osc}_\delta(t)W &\leq \text{osc}_\delta(t)W_n + \text{osc}_\delta(t)P_n \\ &\leq 2(c_1+c_2) \delta^{2-(s+1)/2}. \end{aligned} \quad (111)$$

(b) Consider

$$\begin{aligned} &\int_a^b dt |W(t+\delta) - W(t-\delta)| = \\ &\int_a^b dt \left| \sum_{k=1}^{\infty} \frac{q^{k(s-1)} \cos [(q^{2k}-1)t] \sin [(q^{2k}-1)\delta]}{q^{2k}-1} \right| \\ &\geq \left| \int_a^b dt h(t)f_N(t) \right| - \sum_{k \neq N} \left| \int_a^b dt h(t)f_k(t) \right|, \end{aligned} \quad (112)$$

where

$$f_k(t) = \frac{q^{k(s-1)}}{q^{2k}-1} \cos [(q^{2k}-1)t] \sin [(q^{2k}-1)\delta]. \quad (113)$$

Let $h(t) = \cos[(q^{2N}-1)t]$, $\delta = q^{-2N}$. Then

$$\begin{aligned} \left| \int_a^b dt h(t)f_N(t) \right| &= \frac{q^{N(s-1)}}{q^{2N}-1} \sin(1-q^{-2N}) \int_a^b dt \cos^2[(q^{2N}-1)t] \geq \frac{q^{N(s-1)}}{q^{2N}} \sin\left(\frac{\pi}{6}\right) \int_a^b dt \cos^2[(q^{2N}-1)t] \\ &\geq \frac{1}{2} q^{N(s-3)} \left[\frac{b-a}{2} + \frac{\sin[2b(q^{2N}-1)] - \sin[2a(q^{2N}-1)]}{4(q^{2N}-1)} \right] \geq \frac{1}{2} \delta^{(3-s)/2} \left[\frac{b-a}{2} - \frac{2 \cdot 2}{4q^{2N}} \right] = \\ &\frac{1}{2} \delta^{(3-s)/2} \left[\frac{1}{2}(b-a) - \delta \right] \geq \frac{1}{8} \delta^{(3-s)/2} (b-a). \end{aligned} \quad (114)$$

On the other hand,

$$\begin{aligned} \left| \int_a^b dt h(t)f_k(t) \right| &= \frac{q^{k(s-1)}}{q^{2k}-1} \sin [(q^{2k}-1)q^{-2N}] \int_a^b dt \cos [(q^{2k}-1)t] \cos [(q^{2N}-1)t] \\ &\leq \frac{q^{k(s-1)}}{q^{2k}-1} \left| \frac{\sin[b(q^{2N}-q^{2k})] - \sin[a(q^{2N}-q^{2k})]}{2(q^{2N}-q^{2k})} + \frac{\sin[b(q^{2N}+q^{2k})] - \sin[a(q^{2N}+q^{2k})]}{2(q^{2N}+q^{2k})} \right| \\ &\leq 2q^{k(s-3)} \left[\frac{1}{|q^{2N}-q^{2k}|} + \frac{1}{q^{2N}+q^{2k}} \right] \leq 2q^{k(s-3)} \left[\frac{1}{q^{2N}-q^{2(N-1)}} + \frac{1}{q^{2N}} \right] \leq 5q^{k(s-3)} \delta. \end{aligned} \quad (115)$$

Therefore,

$$W \geq \frac{1}{8} \delta^{(3-s)/2} (b-a) - \sum_k 5q^{k(s-3)} \delta \geq \frac{1}{8} \delta^{(3-s)/2} (b-a) - \frac{5q^{s-3}}{q^{s-3}-1} \delta. \quad (116)$$

But $\frac{1}{2}(3-s) < 1$, thus for large enough N (small enough δ) the first term dominates the other, therefore

$$W \geq c\delta^{2-(1+s)/2}, \quad (117)$$

with $c = (b-a)/16$, for example.

From Theorem 1 it follows that

$$D_v = \frac{1+s}{2}. \quad (118)$$

6. The surface $P(x, t)$ has dimension $D_{xy} = 2 + \frac{1}{2}s$.

Setting x or t constant, we have shown that oscillations are bounded by $c\delta^H$, where exponent H is one of $1, s, s/2$. We also showed the lower bound of variation is always $c\delta^H$, again with H being one of $1, s, s/2$. What is more, there is a dense set of points x , for which $\text{Var}_\delta P_x(t) \geq c\delta^{s/2}$. One can take, for instance, all rational x/π with periodic q -expansion. Thus from Theorem 2 we have

$$D_{xt} = 1 + \max\{D_x, D_t\} = 2 + \frac{s}{2}. \quad (119)$$

Proof. \square

5. Conclusions

In this article, we proved a theorem announced in [32] that a simple textbook problem of quantum theory — the Schrödinger equation describing a point particle in an infinite potential wall — admits continuous but nowhere differentiable solutions with fractal structure. The proposed solutions $\Psi = \Psi(x, t)$ display properties of a fractal quantum carpet, i.e., the probability density, $P(x, t) = |\Psi(x, t)|^2$, forms a fractal surface and its dimension D_{xy} is determined by the fractal dimension D_x of the cross-section $P_t(x)$.

Substitute $c = m, d = m - n$ to arrive at

$$\begin{aligned} P(x, t) &= \frac{2(1-q^{2(s-2)})}{\pi} \sum_{m=0}^{\infty} \left\{ q^{2m(s-2)} \sin^2(q^m x) + 2 \sum_{n < m} q^{(m+n)(s-2)} \sin(q^n x) \sin(q^m x) \cos[(q^{2m} - q^{2n})t] \right\} = \\ &= \frac{2(1-q^{2(s-2)})}{\pi} \left\{ \sum_{m=0}^{\infty} q^{2m(s-2)} \sin^2(q^m x) + 2 \sum_{c=1}^{\infty} \sum_{d=1}^c q^{(2c-d)(s-2)} \sin(q^c x) \sin(q^{c-d} x) \cos[(q^{2c} - q^{2(c-d)})t] \right\} = \\ &= \frac{2(1-q^{2(s-2)})}{\pi} \left\{ \sum_{m=0}^{\infty} q^{2m(s-2)} \sin^2(q^m x) + 2 \sum_{c=1}^{\infty} q^{2c(s-2)} \sin(q^c x) \sum_{d=1}^c q^{-d(s-2)} \sin(q^{c-d} x) \right. \\ &\quad \left. \times \cos \left[(q^2 - 1) q^{2(c-d)} \sum_{a=0}^{d-1} q^{2at} \right] \right\} = \frac{2(1-q^{2(s-2)})}{\pi} \sum_{m=0}^{\infty} q^{2m(s-2)} \sin^2(q^m x) + \frac{4(1-q^{2(s-2)})}{\pi} \sum_{c=1}^{\infty} q^{2c(s-2)} \sin(q^c x) \\ &\quad \times \sum_{d=1}^c q^{-d(s-2)} \sin(q^{c-d} x) \cos[(q^2 - 1)(q^{2(c-1)} + \dots + q^{2(c-d)})t] =: P_x(x) + P_{xt}(x, t). \end{aligned} \quad (123)$$

Acknowledgments

We thank Iwo Białynicki-Birula for many enlightening comments during our fruitful collaboration, which culminated in the work [32], where the results proven here were stated.

Appendix: Auxiliary calculations

A1. Probability density

Take the fractal wave function (36),

$$\Psi(x, t) = \frac{\sqrt{2(1-q^{2(s-2)})}}{\sqrt{\pi}} \sum_{n=0}^{\infty} q^{n(s-2)} \sin(q^n x) e^{-iq^{2n}t}. \quad (120)$$

Let us calculate two useful forms of the probability density $P(x, t)$

$$\begin{aligned} P(x, t) = |\Psi(x, t)|^2 &= \frac{2(1-q^{2(s-2)})}{\pi} \sum_{m, n=0}^{\infty} q^{(m+n)(s-2)} \\ &\quad \times \sin(q^n x) \sin(q^m x) e^{-i(q^{2n} - q^{2m})t}. \end{aligned} \quad (121)$$

Taking $k = m + n, l = n$ we obtain

$$\begin{aligned} P(x, t) &= \frac{2}{\pi} (1 - q^{2(s-2)}) \sum_{k=0}^{\infty} q^{k(s-2)} \\ &\quad \times \sum_{l=0}^k \sin(q^l x) \sin(q^{k-l} x) e^{-i(q^{2l} - q^{2(k-l)})t} = \\ &= \frac{2}{\pi} (1 - q^{2(s-2)}) \sum_{k=0}^{\infty} q^{k(s-2)} \\ &\quad \times \sum_{l=0}^k \sin(q^l x) \sin(q^{k-l} x) \cos[(q^{2l} - q^{2(k-l)})t]. \end{aligned} \quad (122)$$

Note that the time-independent part

$$\begin{aligned}
 P_x(x) &= \\
 &= \frac{2(1-q^{2(s-2)})}{\pi} \sum_{m=0}^{\infty} \frac{q^{2m(s-2)}(1-\cos(q^m 2x))}{2} = \\
 &= \frac{1}{\pi} - \frac{(1-q^{2(s-2)})}{\pi} \sum_{m=0}^{\infty} q^{m(2s-4)} \cos(q^m 2x), \\
 &= \frac{1}{\pi} - \frac{(1-q^{2(s-2)})}{\pi} \sum_{m=0}^{\infty} q^{m(2s-4)} \cos(q^m 2x), \quad (124)
 \end{aligned}$$

is a Weierstrass-like function with the dimension $s' = \max\{2s-2, 1\} \in [1, 2)$ (i.e., for $s \in [1, 3/2]$, $s' = 1$). From the equation (123) one immediately gets the spectrum of $P(x, t)$ — all the frequencies governing the time evolution are

$$\omega_{c,d} = (q^2 - 1)(q^{2(c-1)} + \dots + q^{2(c-d)}), \quad (125)$$

where $c = 1, 2, \dots$, $d = 1, 2, \dots, c$. Thus all the frequencies divide by $q^2 - 1$ which is also the smallest frequency, so the fundamental period of $P(x, t)$ is $2\pi/(q^2 - 1)$.

A2. Average velocity

Let us study the behavior of $\langle x \rangle$.

$$\begin{aligned}
 \langle x \rangle &= \int_0^\pi dx x |\Psi|^2 = \frac{\pi}{2} - \frac{16}{\pi} (1 - q^{2(s-2)}) \\
 &\times \sum_{k=1}^{\infty} \frac{q^{k(s-1)}}{(q^{2k} - 1)^2} \cos[(q^{2k} - 1)t]. \quad (126)
 \end{aligned}$$

The above expression is valid only for even q . For odd q we have just the first term, which is $\pi/2$.

The average $x(t)$ is of class \mathcal{C}^1 , because its derivative is given by an absolutely convergent series

$$\begin{aligned}
 \left| \frac{d\langle x \rangle}{dt} \right| &= \left| \frac{16(1-q^{2(s-2)})}{\pi} \sum_{k=1}^{\infty} \frac{q^{k(s-1)} \sin[(q^{2k} - 1)t]}{q^{2k} - 1} \right| \\
 &\leq 2c \sum_{k=1}^{\infty} \frac{q^{k(s-1)}}{q^{2k}} = 2c \frac{q^{s-3}}{1 - q^{s-3}}. \quad (127)
 \end{aligned}$$

In Sect. 4 we show that (127) is fractal, while for odd q the average velocity $|d\langle x \rangle/dt|$, of course, is not. This seemingly strange behavior is caused by the fact that

$$\int_0^\pi dx \sin(nx) \sin(mx) \quad (128)$$

is non-zero only for m, n of different parity. However, if one slightly disturbs our function, for instance, by changing an arbitrary number of terms to the next higher or lower eigenstates, the dimensions D_x and D_t will not be altered, but the average velocity will become fractal. In other words, with probability one, independently of the parity of q , the average velocity of the wave function

$$\begin{aligned}
 \Phi_0(x, t) &= M_0 \sum_{n=1}^{\infty} q^{n(s-2)} \sin[(q^n \pm 1)x] \\
 &\times e^{-i(q^n \pm 1)^2 t}. \quad (129)
 \end{aligned}$$

is fractal characterized by the same dimensions D_x and D_t as the function currently studied.

An explicit example of a similar function for odd q it is

$$\begin{aligned}
 \Phi_1(x, t) &= M_1 \left[2^{s-2} \sin(2x) e^{-i2^2 t} \right. \\
 &\left. + \sum_{n=1}^{\infty} q^{n(s-2)} \sin(q^n x) e^{-iq^{2n} t} \right]. \quad (130)
 \end{aligned}$$

One can see the only difference between this example and the original one (36) is in the *first* term. This difference accounts for the smoothness or roughness of the average velocity. It is very interesting because normally one expects that it is the *asymptotic* behavior that determines the fractal dimension. Here we have an exactly opposite case: a change in the first term (varying most slowly) of a series changes the dimension of a complicated function $\langle v \rangle$.

The average velocity of the wave packet (130) is smooth for even q and fractal for odd q . A function, which gives fractal average velocity for both even and odd q , is

$$\begin{aligned}
 \Phi_2(x, t) &= M_2 \left[2^{s-2} \sin(2x) e^{-i2^2 t} \right. \\
 &\left. + \sum_{n=0}^{\infty} q^{n(s-2)} \sin(q^n x) e^{-iq^{2n} t} \right] = \\
 &M_2 \left[2^{s-2} \sin(2x) e^{-i2^2 t} + \frac{1}{N} \Psi(x, t) \right], \quad (131)
 \end{aligned}$$

where M_2 is the normalization constant. On the other hand,

$$\Phi_3(x, t) = M_3 \sum_{n=1}^{\infty} q^{n(s-2)} \sin(q^n x) e^{-iq^{2n} t} \quad (132)$$

gives smooth average velocity for both even and odd q .

References

- [1] B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco 1982.
- [2] K.J. Falconer, *Fractal Geometry*, John Wiley & Sons, Chichester 1990.
- [3] E. Ott, *Chaos in Dynamical Systems*, Cambridge University Press, Cambridge 1993.
- [4] P. Meakin, *Fractals, Scaling and Growth Far from Equilibrium*, Cambridge University Press, Cambridge 1998.
- [5] P. Gaspard, *Chaos, Scattering and Statistical Mechanics*, Cambridge University Press, Cambridge 1998.

- [6] J.R. Dorfman, *An Introduction to Chaos in Nonequilibrium Statistical Mechanics*, Cambridge University Press, Cambridge 1999.
- [7] T. Gilbert, J.R. Dorfman, P. Gaspard, *Phys. Rev. Lett.* **85**, 1606 (2000).
- [8] P. Gaspard, I. Claus, T. Gilbert, J.R. Dorfman, *Phys. Rev. Lett.* **86**, 1506 (2001).
- [9] D. Wójcik, A. Nowak, M. Kuś, *Phys. Rev. E* **63**, 036221 (2001).
- [10] H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, 2004.
- [11] N. Laskin, *Chaos* **10**, 780 (2000).
- [12] A. Jadczyk, *Quantum Fractals: From Heisenberg's Uncertainty to Barnsley's Fractality*, World Scientific, Singapore 2014.
- [13] N. Laskin, *Fractional Quantum Mechanics*, World Scientific, Singapore 2018.
- [14] D. Bercioux, A. Iñiguez, *Nature Phys* **15**, 111 (2019).
- [15] B. Eckhardt, *Physica D* **33**, 89 (1988).
- [16] R. Ketzmerick, *Phys. Rev. B* **54**, 10841 (1996).
- [17] G. Casati, I. Guarneri, G. Maspero, *Phys. Rev. Lett.* **84**, 63 (2000).
- [18] V. Kotimäki, E. Räsänen, H. Hennig, E.J. Heller, *Phys. Rev. E* **88**, 022913 (2013).
- [19] W. Kłobus, P. Kurzyński, M. Kuś, W. Laskowski, R. Przybycień, K. Życzkowski, *Phys. Rev. E* **105**, 034201 (2022).
- [20] S.N. Kempkes, M.R. Slot, S.E. Freney, S.J.M. Zevenhuizen, D. Vanmaekelbergh, I. Swart, C. Morais Smith, *Nat. Phys.* **15**, 127 (2019).
- [21] H. Kröger, *Phys. Rep.* **323**, 181 (2000).
- [22] E. Akkermans, [arXiv:1210.6763](https://arxiv.org/abs/1210.6763), 2012.
- [23] J. Ambjørn, J. Jurkiewicz, Y. Watabiki, *Nucl. Phys. B* **454**, 313 (1995).
- [24] J. Ambjørn, J. Jurkiewicz, R. Loll, *Phys. Rev. Lett.* **95**, 171301 (2005).
- [25] V.E. Tarasov, *Fractional Dynamics. Applications of Fractional Calculus to Dynamics of Particles, Fields and Media*, Springer-Verlag, Berlin 2010.
- [26] A.S. Sanz, *J. Phys. A* **38**, 6037 (2005).
- [27] M.V. Berry, *J. Phys. A: Math. Gen.* **29**, 6617 (1996).
- [28] I. Marzoli, F. Saif, I. Białynicki-Birula, O.M. Friesch, A.E. Kaplan, W.P. Schleich, *Acta Phys. Slov.* **48**, 323 (1998).
- [29] M.J.W. Hall, M.S. Reineker, W.P. Schleich, *J. Phys. A* **32**, 8275 (1999).
- [30] M. Berry, I. Marzoli, W. Schleich, *Phys. World* **14**, 39 (2001).
- [31] P.T. Grochowski, T. Karpiuk, M. Brewczyk, K. Rzażewski, *Phys. Rev. Res.* **2**, 013119 (2020).
- [32] D. Wójcik, I. Białynicki-Birula, K. Życzkowski, *Phys. Rev. Lett.* **85**, 5022 (2000).
- [33] A.V. Gorokhov, A.V. Shaikin, in: *Proc. SPIE 5067, Saratov Fall Meeting 2002: Coherent Optics of Ordered and Random Media III*, 2003.
- [34] C. Tricot, *Curves and Fractal Dimension*, Springer-Verlag, New York 1995.
- [35] K.J. Falconer, *Techniques in Fractal Geometry*, John Wiley and Sons, New York, 1997.
- [36] G. Bouligand, *Bull. Sc. Math.* **52**, 320 (1928); *Bull. Sc. Math.* **52**, 361 (1928).
- [37] F. Hausdorff, *Math. Ann.* **79**, 157 (1919).
- [38] G.A. Edgar, *Integral, Probability, and Fractal Measures*, Springer, New York 1998.
- [39] K.J. Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, Cambridge 1985.
- [40] F. Weierstrass, in: *Mathematische Werke II*, Mayer & Muller, Berlin 1895, p. 71.
- [41] G.A. Edgar, editor, *Classics on fractals*, Addison-Wesley, Reading (MA) 1993.
- [42] G.H. Hardy, *Trans. Am. Math. Soc.* **17**, 301 (1916).
- [43] R.D. Mauldin, S.C. Williams, *Trans. Am. Math. Soc.* **298**, 793 (1986).
- [44] R.D. Mauldin, in: *Dimensions and Entropies in Chaotic Systems*, Ed. G. Mayer-Kress, Springer-Verlag, Berlin 1986.
- [45] B.R. Hunt, *Proc. Am. Math. Soc.* **126**, 791 (1998).

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY



Contra Bellum: Bell’s Theorem as a Confusion of Languages

M. CZACHOR*

Instytut Fizyki i Informatyki Stosowanej, Politechnika Gdańska, 80-233 Gdańsk, Poland

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S158](https://doi.org/10.12693/APhysPolA.143.S158)

*e-mail: marek.czachor@pg.edu.pl

Bell’s theorem is a conflict of mathematical predictions formulated within an infinite hierarchy of mathematical models. Inequalities formulated at level $k \in \mathbb{Z}$ are violated by probabilities at level $k + 1$. We are inclined to think that $k = 0$ corresponds to the classical world, while $k = 1$ — to the quantum one. However, as the $k = 0$ inequalities are violated by $k = 1$ probabilities, the same relation holds between $k = 1$ inequalities violated by $k = 2$ probabilities, $k = -1$ inequalities violated by $k = 0$ probabilities, and so forth. By accepting the logic of the Bell theorem, can we prove by induction that nothing exists?

topics: Bell’s theorem, black holes, non-Newtonian calculus, quantum cryptography

1. Introduction

Is Bell’s theorem a mathematical theorem? If we treat Bell’s theorem [1] as a theorem about the additivity of Lebesgue measures, then yes — this is a mathematical theorem. However, Bell’s theorem is more ambitious. It tells us about reality per se, the security of communication channels, the structure of space and time, and even the freedom of experimental physicists.

Although mathematical theorems cannot have counterexamples, this is not necessarily true for theorems about physical reality. The whole history of science is a series of exceptions to various well-established truths.

One such famous truth about reality was known as Euclid’s fifth axiom, which essentially states that angles in any triangle add up to 180° . It was so self-evident to 19th-century mathematicians that even Gauss himself was not eager to publish his thoughts on the subject.

Bell’s theorem is technically based on another apparently self-evident truth about additivity, namely

$$\int_{\Lambda} d\lambda (f \pm g)(\lambda) = \int_{\Lambda} d\lambda f(\lambda) \pm \int_{\Lambda} d\lambda g(\lambda). \tag{1}$$

In proofs of Bell-type inequalities, one often replaces (1) with a more elementary rule,

$$\frac{n \pm m}{N} = \frac{n}{N} \pm \frac{m}{N}. \tag{2}$$

Thus, (1) occurs in contexts of probability measures, while (2) is typical of frequentist approaches.

However, neither (1) nor (2) are universally true; (1) fails for fuzzy or fractal functions; (2) fails if n, m , represent velocities and N is the velocity of light. In the latter case, what we get is rather

$$n \oplus m = N \tanh \left(\tanh^{-1} \left(\frac{n}{N} \right) + \tanh^{-1} \left(\frac{m}{N} \right) \right) = f^{-1} \left(f(n) + f(m) \right). \tag{3}$$

Of course, nothing can prevent us from adding velocities by means of (2), but this is not what Nature does. The arithmetic of Nature is (3). A relation between \oplus and $+$ is here analogous to the one between a curvature of a general manifold and the flatness of its local chart of coordinates (charts in atlases are flat). Arithmetic in special relativity becomes as physical as geometry in general relativity.

A similar situation occurs with (1). In fuzzy and fractal applications, one often encounters [2–12]

$$\int_{\Lambda} d\lambda f(\lambda) \oplus \int_{\Lambda} d\lambda g(\lambda) = \int_{\Lambda} d\lambda (f \oplus g)(\lambda). \tag{4}$$

The exact meaning of \oplus depends on the way fuzzy sets are constructed [13] or which fractals one is dealing with [14]. A particular example of (4) occurs in Maslov's idempotent analysis [15, 16]. Here, certain optimization problems that are nonlinear in the usual framework become linear with respect to generalized arithmetic operations [17], even though the generalized arithmetic is not isomorphic to the one of \mathbb{R} [18].

The goal of the paper is to show that quantum probabilities typical of a two-particle singlet state (that is, those used by Bell in his argumentation), despite all the wisdom of theoretical physicists, *can* result from a local theory where Einstein–Podolsky–Rosen-type (EPR) elements of reality exist [19–21], with probabilities given by local realistic Clauser–Horne formulas [22], where observers have free will, and their detectors are 100% efficient. The only difference is in the form of the integral, whose linearity is with respect to \oplus , \ominus , \odot , and \oslash appropriately defined.

Violation of Bell-type inequalities is then no more paradoxical than $c+c=2c$, which could be claimed to violate the speed of light limit. Moreover, there is no problem with circumventing the Tsirelson bounds typical of Hilbert-space models of probability [23], still maintaining Bell locality, EPR elements of reality, the free will of observers, and 100% efficient detectors.

I can reassure the readers that the models we are analyzing are *not* an alternative to quantum mechanics. They do not explain why probability amplitudes interfere, but, nevertheless, shift the discussion of linearity to new, unexplored areas. The status of theorems based on algebraic properties of observables, such as the Greenberger–Horne–Zeilinger theorem [24, 25] or its single-particle analogs [26], is still open.

But what the models *do* show is that quantum correlations of an EPR type do not necessarily exclude EPR elements of reality — a conclusion with potentially dramatic implications for quantum cryptography. How serious the consequences are, remains to be investigated. Einstein's views on incompleteness of quantum mechanics receive unexpected support.

We will begin our discussion with the observation that principles of relativity are more general and ubiquitous than Einstein's relativity of uniform motion or Copernican relativity of point of observation. The most fundamental principle occurring in all natural sciences is the relativity of arithmetic [27]. It implies, in particular, principles of the relativity of calculus and the relativity of probability. Both are essential for Bell's theorem.

2. Relativity of probability

Relativity of probability occurs at several levels. The most obvious one is illustrated by the following

example. It can be regarded as a particular case of Einstein's special relativity.

Assume a source emits particles to the right or to the left, with certain probability density $\rho(v)$ of velocities. If N particles have been emitted, let N_+ denote the number of particles propagating to the right. An observer measures N_+/N and compares it with the theoretical prediction, $p_+ = \int_0^\infty dv \rho(v)$.

An observer that moves with velocity V with respect to the previous one will measure a different value of N_+/N , even though both of them analyze the same experiment with the same N .

The example is trivial, but it illustrates an important fact about probability — different observers may associate different probabilities with the same experimental situation and with the same definition of elementary events. In this concrete example, the relativity of probability, $p_+(V) = \int_0^\infty dv \rho(v+V)$, results from the relativity of motion.

As a less trivial relativistic example, consider the gravitational collapse of a star. There are two observers: Alice, who falls with the star, and Bob, who remains at rest at position r . Alice employs a broken clock that randomly fails to work (which happens with probability p_0). The motion of the clock's hand becomes a Poisson process characterized by probability $p_1 = 1 - p_0$ of a forward move.

A single run of experiment lasts a fixed amount τ of the observer's proper time. Alice measures $N = \lfloor \tau/\Delta\tau \rfloor$ bits A_1, \dots, A_N (N_1 events $A_j=1$ when the clock's hand moves; $N_0 = N - N_1$ events $A_j=0$ when it gets stuck). The average amount of proper time measured by the Alice's damaged clock is $p_1\tau$.

The experimental ratio N_1/N observed by Alice gets translated into \tilde{N}_1/\tilde{N} observed by Bob. In general, $\tilde{N}_1 \neq N_1$ and $\tilde{N} \neq N$ because the numbers of observed events differ for Alice and Bob due to the relativity of time and the presence of the horizon. The events observed by Alice after she crosses the Schwarzschild radius at her proper time τ_S will be unavailable to Bob, even though his detectors are 100% efficient.

Bob should cautiously draw conclusions about N_1 and N on the basis of \tilde{N}_1 and \tilde{N} he observes. For example, if he concludes that τ_S is greater than τ because, from his perspective, Alice cannot reach the Schwarzschild radius, this inequality can be “violated” in the world of Alice.

Bob can derive various inequalities about the data of Alice, provided he knows the map g_r that relates her N_1/N with his

$$\frac{\tilde{N}_1}{\tilde{N}} = g_r\left(\frac{N_1}{N}\right). \tag{5}$$

The exact form of g_r is irrelevant to our argument, but it could be derived on the basis of general relativity if needed.

From our perspective, it is important that g_r connects two real probabilistic processes. Both N_1/N and $g_r(N_1/N)$ are true, physically significant

probabilities. The “violation” of Bob’s world $\tau_S > \tau$ by Alice’s world $\tau_S < \tau$ is paradoxical only for those who do not understand Einstein’s theory of gravity.

3. A lemma on relativity of binary probabilities

For binary events, there exists a simple result guaranteeing that both N_1/N and $g(N_1/N)$ are probabilities.

Lemma 1: $g(p) + g(1-p) = 1$ for any $p \in [0, 1]$, if and only if

$$g(p) = \frac{1}{2} + h\left(p - \frac{1}{2}\right), \quad (6)$$

where $h(-x) = -h(x)$. Any such g has a fixed point at $p = \frac{1}{2}$.

So, any antisymmetric $h(x)$ leads to an acceptable $g(p)$. The proof can be found in [28]. For a discussion of non-binary probabilities, see [29].

As an example, consider the antisymmetric function

$$h(x) = \frac{1}{2} \sin(\pi x). \quad (7)$$

Then $g(p) = \sin^2(\frac{\pi}{2}p)$, and indeed

$$g(p) + g(1-p) = \sin^2\left(\frac{\pi}{2}p\right) + \cos^2\left(\frac{\pi}{2}p\right) = 1 \quad (8)$$

for any p . Now let $p = (\pi - \theta)/\pi$ be the probability of finding a point belonging to the overlap of two half-circles rotated by θ . Then

$$g(p) = \sin^2\left(\frac{\pi}{2} \frac{\pi - \theta}{\pi}\right) = \cos^2\left(\frac{\theta}{2}\right) \quad (9)$$

is the Malus law for spin 1/2 (or Mach–Zehnder interferometers).

Note that g is one-to-one on $[0, 1]$. Moreover, $g(0) = 0$ and $g(1) = 1$ — a property with important implications for the definition of bits: classical, quantum, and intermediate.

The readers should think of p and $\tilde{p} = g(p)$ in categories similar to those that have led us to (5). Both p and \tilde{p} can be physically meaningful. We should be as cautious as Bob in formulating statements about the level of p on the basis of the rules that apply to the level of \tilde{p} .

4. Arithmetic elements of reality

Consider some set \mathbb{X} and a bijection $f_{\mathbb{X}} : \mathbb{X} \rightarrow \mathbb{R}$. Cardinality of \mathbb{X} must be the same as the one of \mathbb{R} . The inverse map is $g_{\mathbb{X}} = f_{\mathbb{X}}^{-1}$, $g_{\mathbb{X}} : \mathbb{R} \rightarrow \mathbb{X}$. The map g from the previous section can be an example of $g_{\mathbb{R}}$ restricted to $[0, 1]$. To put it differently, the bijection $g : [0, 1] \rightarrow [0, 1]$ can be extended to a bijection $g_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}$, satisfying $g_{\mathbb{R}}(p) = \sin^2(\frac{\pi}{2}p) = g(p)$ when restricted to $p \in [0, 1]$.

We define arithmetic operations in \mathbb{X} ,

$$x \oplus_{\mathbb{X}} y = g_{\mathbb{X}}(f_{\mathbb{X}}(x) + f_{\mathbb{X}}(y)), \quad (10)$$

$$x \ominus_{\mathbb{X}} y = g_{\mathbb{X}}(f_{\mathbb{X}}(x) - f_{\mathbb{X}}(y)), \quad (11)$$

$$x \odot_{\mathbb{X}} y = g_{\mathbb{X}}(f_{\mathbb{X}}(x) \cdot f_{\mathbb{X}}(y)), \quad (12)$$

$$x \oslash_{\mathbb{X}} y = g_{\mathbb{X}}(f_{\mathbb{X}}(x)/f_{\mathbb{X}}(y)). \quad (13)$$

The arithmetic given by (10)–(13) is called projective [12, 30]. Here $f_{\mathbb{X}}$ defines an isomorphism of arithmetics. The neutral elements, $0_{\mathbb{X}} = g_{\mathbb{X}}(0)$ (projective zero in \mathbb{X}), $1_{\mathbb{X}} = g_{\mathbb{X}}(1)$ (projective one in \mathbb{X}) are to some extent analogous to qubits [29, 31].

Indeed, expressions such as $0_{\mathbb{X}} + 0_{\mathbb{Y}}$ are, in general, meaningless if $\mathbb{X} \neq \mathbb{Y}$. Just think of $\mathbb{X} = \mathbb{R}$ and $\mathbb{Y} = \mathbb{R}^2$. Even if $\mathbb{X} = \mathbb{Y}$ and $0_{\mathbb{X}} = 0_{\mathbb{Y}} = 0$, $1_{\mathbb{X}} = 1_{\mathbb{Y}} = 1$, the projective bits can be as incompatible as eigenvalues of non-commuting projectors.

However, in spite of this incompatibility, $0_{\mathbb{X}} = g_{\mathbb{X}}(0)$ and $0_{\mathbb{Y}} = g_{\mathbb{Y}}(0)$ are images of the same $0 \in \mathbb{R}$. This “ordinary zero” can play the role of an EPR-type element of reality for $0_{\mathbb{X}}$ and $0_{\mathbb{Y}}$, i.e., incompatible projective bits can be correlated by means of their elements of reality, in exact analogy to the formulas postulated by Bell in his classic analysis.

Note that (3) is an example of (10). The neutral elements are $0_{\mathbb{X}} = N \tanh(0) = 0$, $1_{\mathbb{X}} = N \tanh(1) = 0.76N$ (hence velocity $0.76c$ is the neutral element of special relativistic multiplication). The velocity of light is literally infinite, of course in the sense of $\infty_{\mathbb{X}} = N \tanh(\infty) = N$. The case $c \oplus c = c$ is an example of $\infty_{\mathbb{X}} \oplus_{\mathbb{X}} \infty_{\mathbb{X}} = \infty_{\mathbb{X}}$. Strictly speaking, a relativistic unit of velocity is not c but $c \tanh(1)$.

5. Clauser–Horne formulas for projective bits

We are interested in singlet-state probabilities,

$$\begin{aligned} P_{0_1} &= P_{1_1} = P_{0_2} = P_{1_2} = \\ \langle \psi | \hat{P}_{0_1} \otimes I | \psi \rangle &= \langle \psi | \hat{P}_{1_1} \otimes I | \psi \rangle = \\ \langle \psi | I \otimes \hat{P}_{0_2} | \psi \rangle &= \langle \psi | I \otimes \hat{P}_{1_2} | \psi \rangle = \frac{1}{2} \end{aligned} \quad (14)$$

with joint probabilities,

$$\begin{aligned} P_{0_1 0_2} &= P_{1_1 1_2} = \langle \psi | \hat{P}_{0_1} \otimes \hat{P}_{0_2} | \psi \rangle = \langle \psi | \hat{P}_{1_1} \otimes \hat{P}_{1_2} | \psi \rangle = \\ &= \frac{1}{2} \sin^2\left(\frac{\alpha - \beta}{2}\right), \end{aligned} \quad (15)$$

$$\begin{aligned} P_{0_1 1_2} &= P_{1_1 0_2} = \langle \psi | \hat{P}_{0_1} \otimes \hat{P}_{1_2} | \psi \rangle = \langle \psi | \hat{P}_{1_1} \otimes \hat{P}_{0_2} | \psi \rangle = \\ &= \frac{1}{2} \cos^2\left(\frac{\alpha - \beta}{2}\right). \end{aligned} \quad (16)$$

We will write them in a Clauser–Horne form [12]

$$P_{A_1 A_2} = \int D x \chi_{A_1}(x) \odot_{\mathbb{X}} \chi_{A_2}(x) \odot_{\mathbb{X}} \rho(x), \quad (17)$$

$$P_A = \int D x \chi_A(x) \odot_{\mathbb{X}} \rho(x) = \frac{1}{2}, \quad (18)$$

where the χ s are characteristic functions and $\rho(x) \geq 0$ is a non-negative probability density normalized to 1,

$$\int D x \rho(x) = 1. \quad (19)$$

Of course, the trick is to work with appropriate forms of the integral and employ the freedom available in possible meanings of $\odot_{\mathbb{X}}$ and $\oplus_{\mathbb{X}}$. We will assume $\mathbb{X} = \mathbb{R}$, and $g_{\mathbb{X}}(0) = 0$, $g_{\mathbb{X}}(1) = 1$. The latter two conditions imply that the values of projective bits will be given by ordinary 0 and 1.

Formulas (17) and (18) implicitly imply that measurements are modeled in the usual way by the products of $\rho(x)$ with characteristic functions,

$$\rho(x) \mapsto \chi_A(x) \odot_{\mathbb{X}} \rho(x), \quad (20)$$

$$\rho(x) \mapsto \chi_{A \cap B}(x) \odot_{\mathbb{X}} \rho(x) =$$

$$\chi_A(x) \odot_{\mathbb{X}} \chi_B(x) \odot_{\mathbb{X}} \rho(x), \quad (21)$$

and so forth. If A' denotes the set-theoretic completion of set A , then

$$\chi_A(x) \oplus_{\mathbb{X}} \chi_{A'}(x) = 1, \quad (22)$$

$$1 \ominus_{\mathbb{X}} \chi_A(x) = \chi_{A'}(x), \quad (23)$$

$$\chi_A(x) \odot_{\mathbb{X}} \chi_{A'}(x) = 0, \quad (24)$$

$$\chi_A(x) \odot_{\mathbb{X}} \chi_A(x) = \chi_A(x), \quad (25)$$

$$\chi_{A'}(x) \odot_{\mathbb{X}} \chi_{A'}(x) = \chi_{A'}(x). \quad (26)$$

The probabilities must add up to 1 in an ordinary way,

$$P_{0_1 0_2} + P_{0_1 1_2} + P_{1_1 0_2} + P_{1_1 1_2} = 1, \quad (27)$$

because this is how experimentalists will use them.

On the other hand, the integral can be additive in a more general sense of (4), similarly to fuzzy, fractal, or idempotent integrals. A dual form of normalization will be a consequence of such a generalized linearity,

$$P_{0_1 0_2} \oplus_{\mathbb{X}} P_{0_1 1_2} \oplus_{\mathbb{X}} P_{1_1 0_2} \oplus_{\mathbb{X}} P_{1_1 1_2} = 1. \quad (28)$$

Note that (27) and (28) must hold simultaneously for any $P_{A_1 A_2}$, a condition, which is not entirely trivial, but whose solution exists.

The choice of arithmetic will naturally define the integral occurring in (17)–(18). Historically the first construction of calculus based on projective arithmetic was given by Grossman and Katz in their 1972 book *Non-Newtonian Calculus* [32]^{†1}.

Bell published his paper in 1964.

^{†1}Grossman and Katz had worked on the problem since the late 1960s, but their little book, as well as its two sequels [33, 34], went practically unnoticed by both mathematicians and physicists. The main idea was rediscovered by Endre Pap and published in 1993 in a local journal of Novi Sad University [35]. Over the next two decades, the formalism developed by Pap matured into a whole new branch of applied mathematics (see [9, 10, 36]). The so-called F^α calculus on fractals [37–39] can be regarded as a special case of non-Newtonian calculus. The same can be said of Maslov's idempotent analysis [15, 16]. In 2014, the ideas of Grossman, Katz, and Pap were once again rediscovered by myself [27] and led to nontrivial applications in physics, just to mention wave equations on Koch curves (a long-standing problem of fractal analysis) [40], elements of Fourier analysis on arbitrary Cantor sets (circumventing a no-go theorem about Fourier transforms on the triadic Cantor set) [11], or the issues of dark energy and matter [41, 42]. The problem of

6. Non-Newtonian Calculus

We need an integral because Clauser–Horne formulas involve integration. In fuzzy or fractal applications, the usual strategy would be to define some measure on a fractal or fuzzy set, and only then start worrying whether the resulting integral is consistent with derivatives, typically defined by means of a completely different procedure than the one that has led to the integral. In effect, the fundamental theorem of calculus often becomes problematic [13].

The approach that starts with arithmetic is much more systematic. First, one defines a derivative by means of a formula which is a straightforward generalization of

$$\frac{dF(x)}{dx} = \lim_{\delta \rightarrow 0} \frac{F(x + \delta) - F(x)}{\delta}. \quad (29)$$

Then one demands that the integral be related to the derivative by means of the fundamental theorem of calculus. The notion of measure appears automatically at the very end, once we know how to integrate. Knowing the measure, we know how to define probability.

The non-Newtonian derivative of a function $F: \mathbb{X} \rightarrow \mathbb{Y}$ depends on the arithmetics of \mathbb{X} and \mathbb{Y} . Denoting $\delta_{\mathbb{X}} = g_{\mathbb{X}}(\delta)$, $\delta_{\mathbb{Y}} = g_{\mathbb{Y}}(\delta)$, one defines

$$\frac{DF(x)}{Dx} = \lim_{\delta \rightarrow 0} (F(x \oplus_{\mathbb{X}} \delta_{\mathbb{X}}) \ominus_{\mathbb{Y}} F(x)) \oslash_{\mathbb{Y}} \delta_{\mathbb{Y}}, \quad (30)$$

whose more practical form reads

$$\frac{DF(x)}{Dx} = g_{\mathbb{Y}} \left(\frac{d\tilde{F}(f_{\mathbb{X}}(x))}{df_{\mathbb{X}}(x)} \right). \quad (31)$$

The argument of $g_{\mathbb{Y}}$ in (31) is the (Newtonian) derivative (29) of \tilde{F} defined by the commutative diagram

$$\begin{array}{ccc} \mathbb{X} & \xrightarrow{F} & \mathbb{Y} \\ f_{\mathbb{X}} \downarrow & & \uparrow g_{\mathbb{Y}} \\ \mathbb{R} & \xrightarrow{\tilde{F}} & \mathbb{R} \end{array} \quad (32)$$

Although (31) makes non-Newtonian differentiation as simple as the Newtonian one, (30) reveals the logical structure behind the derivative. For example, it explains why we find the generalized form of additivity

$$\frac{D(F(x) \oplus_{\mathbb{Y}} G(x))}{Dx} = \frac{DF(x)}{Dx} \oplus_{\mathbb{Y}} \frac{DG(x)}{Dx} \quad (33)$$

and the generalized Leibniz rule

Bell's theorem was reformulated from a non-Newtonian perspective in a series of four papers [28, 29, 31, 43]. The version of non-Newtonian formalism introduced in [27] is based on the weakest assumptions and thus is the most general and flexible so far, at least in my opinion. A review of the formalism can be found in [12].

$$\frac{D(F(x) \odot_{\mathbb{Y}} G(x))}{Dx} = \frac{DF(x)}{Dx} \odot_{\mathbb{Y}} G(x) \oplus_{\mathbb{Y}} F(x) \odot_{\mathbb{Y}} \frac{DG(x)}{Dx}. \quad (34)$$

To define a non-Newtonian integral $\int_a^b Dx F(x)$, we demand its consistency with the derivatives (two fundamental theorems of calculus)

$$\int_a^b Dx \frac{DF(x)}{Dx} = F(b) \ominus_{\mathbb{Y}} F(a), \quad (35)$$

and

$$\frac{D}{Dx} \int_a^x Dy F(y) = F(x). \quad (36)$$

The result is

$$\int_a^b Dx F(x) = g_{\mathbb{Y}} \left(\int_{f_{\mathbb{X}}(a)}^{f_{\mathbb{X}}(b)} dr \tilde{F}(r) \right). \quad (37)$$

The argument of $g_{\mathbb{Y}}$ in (37) is the (Newtonian, hence Lebesgue, Riemann, etc.) integral of \tilde{F} defined by the commutative diagram (32). Such an integral inherits additivity,

$$\int_a^b Dx F(x) \oplus_{\mathbb{Y}} G(x) = \int_a^b Dx F(x) \oplus_{\mathbb{Y}} \int_a^b Dx G(x), \quad (38)$$

and one-homogeneity (for a constant F),

$$\int_a^b Dx F \odot_{\mathbb{Y}} G(x) = F \odot_{\mathbb{Y}} \int_a^b Dx G(x), \quad (39)$$

from the arithmetic that defines the derivative.

It should be now rather clear why non-Newtonian hidden-variable models lead to Bell-type inequalities of basically the usual form, but with the ordinary plus, minus, times, and divided replaced by \oplus , \ominus , \odot , and \oslash .

If one takes this subtlety into account, then quantum mechanical singlet-state probabilities will *not* violate the Bell inequality — not the one that can be derived for the hidden-variable model.

7. Singlet-state probabilities

It remains to construct the projective arithmetic (10)–(13) that implies (14)–(18) by means of the corresponding non-Newtonian integral (37). First, $g_{\mathbb{X}}$ will be constructed via an intermediate g , whose properties are described by the following consequence of Lemma 1.

Lemma 2: Consider four joint probabilities $p_{0_1 0_2}$, $p_{1_1 1_2}$, $p_{0_1 1_2}$, $p_{1_1 0_2}$, satisfying

$$\sum_{AB} p_{AB} = 1, \quad (40)$$

$$\sum_A p_{AA_2} = \sum_A p_{A_1 A} = \frac{1}{2}. \quad (41)$$

A sufficient condition for

$$\sum_{AB} G(p_{AB}) = 1, \quad (42)$$

for any p_{AB} satisfying (40), (41), is given by $G(p) = \frac{1}{2}g(2p)$, where g satisfies Lemma 1. Any such G has a fixed point at $p = 1/4$.

The proof can be found in [43].

Guided by Lemmas 1 and 2, we take $\mathbb{X} = \mathbb{R}$ and define (Fig. 1),

$$g_{\mathbb{X}}(x) = \frac{n}{2} + \frac{1}{2} \sin^2 \left(\pi x - \pi \frac{n}{2} \right), \quad (43)$$

$$f_{\mathbb{X}}(x) = \frac{n}{2} + \frac{1}{\pi} \arcsin \sqrt{2x - n}, \quad (44)$$

for $\frac{1}{2}n \leq x \leq \frac{1}{2}(n+1)$, $n \in \mathbb{Z}$ (for more details, see [26]). Function (43) is, up to the rescaling $g(p) \mapsto \frac{1}{2}g(2p)$ required by Lemma 2, the one we have used as the illustration of Lemma 1 for spins $1/2$, but extended from $[0, 1]$ to the whole of \mathbb{R} . Non-Newtonian integrals (17)–(18) constructed by means of (43)–(44) reconstruct singlet-state probabilities if we appropriately define the characteristic functions. For example,

$$\left(\int_{\alpha_{\mathbb{X}}}^{\beta_{\mathbb{X}}} D\lambda \right) \oslash_{\mathbb{X}} g_{\mathbb{X}}(2\pi) = \frac{1}{2} \sin^2 \left(\frac{\beta - \alpha}{2} \right), \quad (45)$$

$\alpha_{\mathbb{X}} = g_{\mathbb{X}}(\alpha)$, $\beta_{\mathbb{X}} = g_{\mathbb{X}}(\beta)$, which is the standard local hidden-variable expression postulated by Bell. It can be rewritten as a particular case of (17) if we denote

$$\rho(\lambda) = 1 \oslash_{\mathbb{X}} g_{\mathbb{X}}(2\pi) = g_{\mathbb{X}}(1/(2\pi)), \quad (46)$$

and integration is over the circle $0 \leq \lambda < g_{\mathbb{X}}(2\pi)$. The product of characteristic functions is encoded in the integration limits.

The rotational invariance of the probability is a consequence of

$$\int_{\alpha_{\mathbb{X}}}^{\beta_{\mathbb{X}}} Dx = \int_{\alpha_{\mathbb{X}} \oplus_{\mathbb{X}} \gamma_{\mathbb{X}}}^{\beta_{\mathbb{X}} \oplus_{\mathbb{X}} \gamma_{\mathbb{X}}} Dx \quad (47)$$

valid for any $\gamma_{\mathbb{X}} \in \mathbb{X}$ and any non-Newtonian integral defined by means of the arithmetic (10)–(13).

Then, what about the Bell inequality?

Of course, it is *not* violated by (45) despite the exact quantum mechanical form of the probability, and there is nothing paradoxical about this statement. Just try to derive any form of a Bell-type inequality for such a non-Newtonian local hidden-variable model. For example, following the steps of the Clauser–Horne reasoning, one arrives at the projective-arithmetic generalization of the Clauser–Horne inequality,

$$0 \leq 3 \oslash_{\mathbb{X}} P_{1_1 0_2}(\theta) \ominus_{\mathbb{X}} P_{1_1 0_2}(3\theta) \leq 1. \quad (48)$$

Inserting singlet-state probabilities into (48), one finds

$$3 \oslash_{\mathbb{X}} P_{1_1 0_2}(\theta) \ominus_{\mathbb{X}} P_{1_1 0_2}(3\theta) = 1 \quad (49)$$

for any θ , so there is no contradiction.

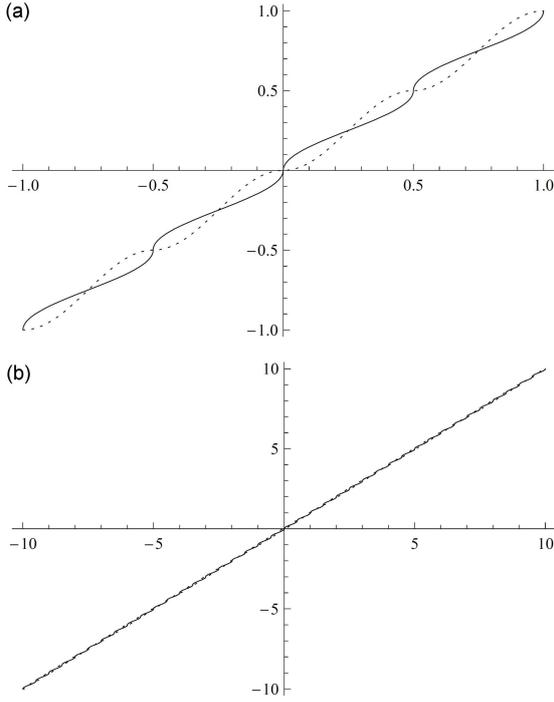


Fig. 1. One-to-one $f_x : \mathbb{R} \rightarrow \mathbb{R}$ (full line) and its inverse g_x (dotted) defined by (44) and (43), as implied by Lemma 2. Both functions have fixed points at integer multiples of $1/4$. The plots are given in two scales, (a) $-1 \leq x \leq 1$ and (b) $-10 \leq x \leq 10$, explaining the origin of the correspondence principle discussed in Sect. 10.

The inequality that will be indeed violated is

$$0 \leq 3P_{1_1 0_2}(\theta) - P_{1_1 0_2}(3\theta) \leq 1, \quad (50)$$

but it is derived under the *wrong* assumption of additivity (1), which does not hold for this concrete model of non-Newtonian integration. Standard Clauser–Horne inequality (50) cannot be proved for non-Newtonian hidden variables in question, so it is no surprise that it is not satisfied in our model.

The readers should keep in mind that although (27) and (28) are simultaneously valid, this is no longer true for arbitrary linear combinations of probabilities, in particular those occurring in (48) and (50).

We will now show that the relation between p and $\tilde{p} = g_x(p)$, which is at the core of the Bell inequality violation, is, in fact, a very special case of an infinite hierarchy of relations, based on an infinite hierarchy of arithmetics and calculi. What we intuitively regard as the “normal” or “our” arithmetic and calculus can correspond to *any* level of the hierarchy.

This will lead us to the notion of a Copernican hierarchy of models. We call them Copernican because they deprive our human point of view of the aura of uniqueness. Each level of such a hierarchy can be our level.

The standard Bell theorem describes a relation between any two neighboring levels of the hierarchy. A surprising consequence of this relation is that in the same way that Bell proved the non-existence of EPR elements of reality, it is possible to prove the non-existence of ourselves.

Well, at least the author of this paper exists as an element of reality.

8. Copernican hierarchies

Functions g that satisfy Lemma 1 form an interesting structure, closed under the composition of maps [43].

Lemma 3: Consider two functions $g_j : [0, 1] \rightarrow [0, 1]$, $j = 1, 2$, that satisfy assumptions of Lemma 1,

$$g_j(p) = \frac{1}{2} + h_j \left(p - \frac{1}{2} \right), \quad (51)$$

where $h_j(-x) = -h_j(x)$. Then $g_{12} = g_1 \circ g_2$ also satisfies Lemma 1 with $h_{12} = h_1 \circ h_2$,

$$g_{12}(p) = \frac{1}{2} + h_{12} \left(p - \frac{1}{2} \right). \quad (52)$$

Accordingly,

$$g_{12}(p) + g_{12}(1-p) = 1 \quad (53)$$

for any $p \in [0, 1]$.

Lemma 4: Let $g^k = g \circ \dots \circ g$, $g^{-k} = g^{-1} \circ \dots \circ g^{-1}$ (k times), $g^0(x) = x$. If g satisfies Lemma 1,

$$g(p) = \frac{1}{2} + h \left(p - \frac{1}{2} \right), \quad (54)$$

then g^k also satisfies Lemma 1 for any $k \in \mathbb{Z}$,

$$g^k(p) = \frac{1}{2} + h^k \left(p - \frac{1}{2} \right). \quad (55)$$

Accordingly,

$$g^k(p) + g^k(1-p) = 1 \quad (56)$$

for any $p \in [0, 1]$, and any integer k . In particular,

$$g^{-1}(p) + g^{-1}(1-p) = 1. \quad (57)$$

The proofs are straightforward [43].

As an illustration, consider again $g(p) = \sin^2\left(\frac{\pi}{2}p\right)$ and

$$g^2(p) = \sin^2 \left[\frac{\pi}{2} \sin^2 \left(\frac{\pi}{2}p \right) \right], \quad (58)$$

$$g^{-1}(p) = \frac{2}{\pi} \arcsin \sqrt{p}. \quad (59)$$

The cross-check of (56) for (58) is simple but instructive

$$\begin{aligned} g^2(p) + g^2(1-p) &= \\ \sin^2 \left[\frac{\pi}{2} \sin^2 \left(\frac{\pi}{2}p \right) \right] + \sin^2 \left[\frac{\pi}{2} \sin^2 \left(\frac{\pi}{2}(1-p) \right) \right] &= \\ \sin^2 \left[\frac{\pi}{2} \sin^2 \left(\frac{\pi}{2}p \right) \right] + \sin^2 \left[\frac{\pi}{2} \cos^2 \left(\frac{\pi}{2}p \right) \right] &= \\ \sin^2 \left[\frac{\pi}{2} \sin^2 \left(\frac{\pi}{2}p \right) \right] + \sin^2 \left[\frac{\pi}{2} \left(1 - \sin^2 \left(\frac{\pi}{2}p \right) \right) \right] &= \\ \sin^2 \left[\frac{\pi}{2} \sin^2 \left(\frac{\pi}{2}p \right) \right] + \cos^2 \left[\frac{\pi}{2} \sin^2 \left(\frac{\pi}{2}p \right) \right] &= 1. \end{aligned} \quad (60)$$

An analogous proof for (59) is left as an exercise. Figure 2 shows the result.

We are inclined to believe that “our” arithmetic corresponds to $k = 0$. So, consider any binary probabilities from level 0,

$$p_0 + p_1 = 1, \quad (61)$$

and those from level k ,

$$g^k(p_0) + g^k(p_1) = 1. \quad (62)$$

Denoting $P_0 = g^k(p_0)$, $P_1 = g^k(p_1)$, we obtain a symmetric rule,

$$P_0 + P_1 = 1, \quad (63)$$

$$g^{-k}(P_0) + g^{-k}(P_1) = 1. \quad (64)$$

In both cases, k is an arbitrary integer: positive, negative, or zero.

The question is: How do we know that it is p_A and not P_A that defines level-zero probabilities?

We can phrase the same question in arithmetic terms. To this end, assume $g(p)$ is a restriction to $[0, 1]$ of some bijection $g_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}$ that satisfies $g_{\mathbb{R}}(1) = g(1) = 1$. We can act on both sides of (62) with g^{-k} , while on both sides of (65) with g^k , obtaining

$$g^{-k}(g^k(p_0) + g^k(p_1)) = p_0 \oplus_k p_1 = 1, \quad (65)$$

and

$$g^k(g^{-k}(P_0) + g^{-k}(P_1)) = P_0 \oplus_{-k} P_1 = 1. \quad (66)$$

We have no criterion that could tell us which of the four additions — (61), (63), (65), or (66) — defines the level of description we employ in everyday life. Which of these two probabilities, and which of the several ways of adding them, is our usual way of processing experimental data?

Which of the three additions, $+$, \oplus_k , or \oplus_{-k} , is the one we have learned as kids?

Which of the three derivatives, (29), or

$$\frac{DF(x)}{Dx} = \lim_{\delta \rightarrow 0} (F(x \oplus_k \delta_k) \ominus_k F(x)) \oslash_k \delta_k, \quad (67)$$

or

$$\frac{DF(x)}{Dx} = \lim_{\delta \rightarrow 0} (F(x \oplus_{-k} \delta_{-k}) \ominus_{-k} F(x)) \oslash_{-k} \delta_{-k}, \quad (68)$$

is the one we have mastered during our undergraduate education?

Last but not least, which of the three integrals, $\int Dx F(x)$, is the one that should define a hidden-variable theory?

Numerous fundamental answers are possible.

One possibility is that Nature prefers only one $k \in \mathbb{Z}$ as the true physical arithmetic with some fixed form of g , determined by some unknown physical law. This is the situation we encounter in special relativity when we add velocities by means of $g = \tanh$. In principle, we can detect such a physical g in an experiment. In [42], it is shown that problems with dark energy may indicate that time at cosmological scales involves a nontrivial $g \sim \sinh$. If this conclusion were true, dark energy would be as unreal as the luminiferous aether.

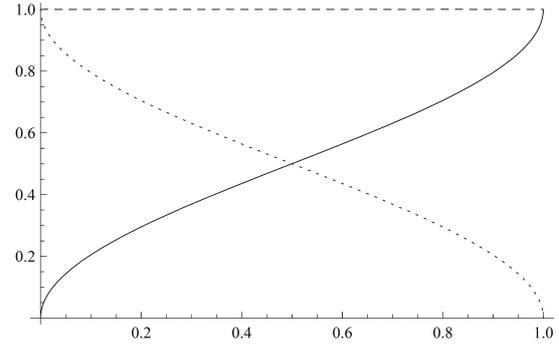


Fig. 2. The results of $g^{-1}(p) = \frac{2}{\pi} \arcsin \sqrt{p}$ (full line), $g^{-1}(1-p)$ (dotted), and their sum (dashed).

The second fundamental possibility is that all these possibilities are simultaneously true. Perhaps there is no preferred k , like there is no preferred rest frame or preferred point of observation of the universe. Only relative k might be observable. Such an option is intriguing and tempting from a theoretical perspective. It could mean, for example, that the same physical law might have its mathematical representations at any level of the hierarchy, and each of these representations might be meaningful. Violation of Bell-type inequalities would then be a conflict of predictions derived at level k but tested at level $k+1$.

A similar conflict occurs if Bob concludes that Alice will never reach the Schwarzschild radius, and yet she crosses it in a finite time.

It remains to say something about the conflicts that occur between non-neighboring levels of the hierarchy. We will see that other well-known bounds, such as the Tsirelson inequality characterizing Hilbert-space models of probability, can be easily circumvented as well.

9. Beyond Tsirelson’s bounds

The standard Clauser–Horne inequality (50)

$$0 \leq 3P_{1_1 0_2}(\theta) - P_{1_1 0_2}(3\theta) \leq 1, \quad (69)$$

is derived for joint probabilities limited by

$$0 \leq P_{1_1 0_2}(\alpha) \leq 0.5. \quad (70)$$

The absolute bounds for such a linear combination of probabilities are, therefore,

$$-0.5 \leq 3P_{1_1 0_2}(\theta) - P_{1_1 0_2}(3\theta) \leq 1.5. \quad (71)$$

Tsirelson bounds are narrower,

$$-\frac{\sqrt{2}-1}{2} \leq 3P_{1_1 0_2}(\theta) - P_{1_1 0_2}(3\theta) \leq \frac{\sqrt{2}+1}{2}. \quad (72)$$

In order to understand the influence of k on the violation of Clauser–Horne $k=0$ inequalities, we have to estimate the expression [43]

$$X(g^k, \theta) = 3g^k\left(\frac{\theta}{2\pi}\right) - g^k\left(\frac{3\theta}{2\pi}\right), \quad (73)$$

where $0 \leq \theta < \pi/3$. The singlet-state example corresponds in this range of parameters to $g^1(p) = \frac{1}{2} \sin^2(\pi p)$. For $\theta = \pi/4$, we find

$$X(g^1, \frac{\pi}{4}) = \frac{3}{2} \sin^2\left(\pi \frac{\pi/4}{2\pi}\right) - \frac{1}{2} \sin^2\left(\pi \frac{3\pi/4}{2\pi}\right) - \frac{\sqrt{2}-1}{2} = -0.20711, \quad (74)$$

that is, the maximal left Tsirelson bound. This is what is usually called the maximal (left) violation of the Clauser–Horne inequality by singlet-state quantum probabilities. For other values of k , we find

$$\begin{aligned} X(g^0, \frac{\pi}{4}) &= 0, \\ X(g^1, \frac{\pi}{4}) &= -0.20711, \\ X(g^2, \frac{\pi}{4}) &= -0.39602, \\ X(g^3, \frac{\pi}{4}) &= -0.48669, \\ X(g^4, \frac{\pi}{4}) &= -0.49978, \\ &\vdots \\ X(g^\infty, \frac{\pi}{4}) &= -0.5. \end{aligned} \quad (75)$$

Of course, as stressed before, the choice of $k=0$ as the reference level is arbitrary. Level $k=2022$ probabilities violate level $k=2021$ inequalities in exactly the same way as quantum mechanics violates the standard Clauser–Horne inequality.

More importantly, $k=2$ probabilities violate $k=1$ inequalities in the same way as $k=1$ probabilities violate $k=0$ inequalities. If $k=0$ elements of reality do not exist, then $k=1$ elements of reality do not exist either. Accepting the logic of Bell's theorem, can we prove by induction that nothing exists?

Slightly modifying the experimental configuration, one obtains the maximal right violations. In our formalism, the function to estimate is

$$Y(g^k, \theta) = 3g^k \left(\frac{3\theta}{2\pi}\right) - g^k \left(\frac{\theta}{2\pi}\right). \quad (76)$$

We find

$$\begin{aligned} Y(g^0, \frac{\pi}{4}) &= 1, \\ Y(g^1, \frac{\pi}{4}) &= \frac{1}{2}(\sqrt{2} + 1) = 1.20711, \\ Y(g^2, \frac{\pi}{4}) &= 1.39602, \\ Y(g^3, \frac{\pi}{4}) &= 1.48669, \\ Y(g^4, \frac{\pi}{4}) &= 1.49978, \\ &\vdots \\ Y(g^\infty, \frac{\pi}{4}) &= 1.5. \end{aligned} \quad (77)$$

All these models are local-realistic, observers have free will, and detectors are ideal. The only modification is in the presence of the bijection g that links arithmetics, calculi, and probabilities at various levels of the hierarchy.

Our g^k plays a role analogous to g_r that linked experimental data collected at different neighborhoods of a collapsing star.

Both examples are based on principles of relativity. We have learned to live with special relativity, general relativity, and the Copernican principle.

It is time to learn to live with the arithmetic principle of relativity.

10. Correspondence principles

Trigonometric functions $\cos_{\mathbb{X}} : \mathbb{X} \rightarrow \mathbb{X}$, $\sin_{\mathbb{X}} : \mathbb{X} \rightarrow \mathbb{X}$ are defined by

$$\cos_{\mathbb{X}}(x) = g_{\mathbb{X}}(\cos(f_{\mathbb{X}}(x))), \quad (78)$$

$$\sin_{\mathbb{X}}(x) = g_{\mathbb{X}}(\sin(f_{\mathbb{X}}(x))). \quad (79)$$

They satisfy all the usual trigonometric relations, of course with respect to appropriate arithmetic operations. They also satisfy all the usual differential relations, of course with respect to appropriate non-Newtonian derivatives. In particular, they define circles by

$$\theta \mapsto (r \odot_{\mathbb{X}} \cos_{\mathbb{X}} \theta, r \odot_{\mathbb{X}} \sin_{\mathbb{X}} \theta). \quad (80)$$

Let us now take the bijections (43) and (44), which we have used to reconstruct singlet state probabilities. Figure 3 shows seven circles defined by (80) for decreasing radii. A picture to the right is a close-up of its left neighbor. All the circles are given by the same formula (80), with the same bijection $g_{\mathbb{X}}$ — the greater the radius, the more circular the shape. Simply put, the larger the x argument, the more difficult it is to tell $g_{\mathbb{X}}(x)$ from x . However, the readers must bear in mind that all these circles are *truly* rotationally invariant! They have been generated as homogeneous spaces of the rotation group in 2D — the only nonstandard element being the choice of arithmetic.

The notion of a *hidden* or *internal* symmetry, often used in particle physics, seems especially adequate here. Each of these circles would have looked “normal” if we had reprogrammed Wolfram Mathematica to make the plots in the arithmetic $\{\mathbb{R}, \oplus_1, \ominus_1, \odot_1, \oslash_1\}$.

The limit $r \rightarrow \infty$ plays a role of a correspondence principle with the ordinary, rotational external symmetry. The obvious similarity to the classical limit of quantum mechanics is striking. Other examples of arithmetic correspondence principles can be found in [27] and [44]. An analogous correspondence principle occurs in the idempotent analysis [18].

11. Implications for cryptography

In 1862, more than a century before Bell's paper, George Boole submitted to Philosophical Transactions of the Royal Society the article *On the theory of probabilities*, where he introduced inequalities imposing constraints on our “possible experience” [45].

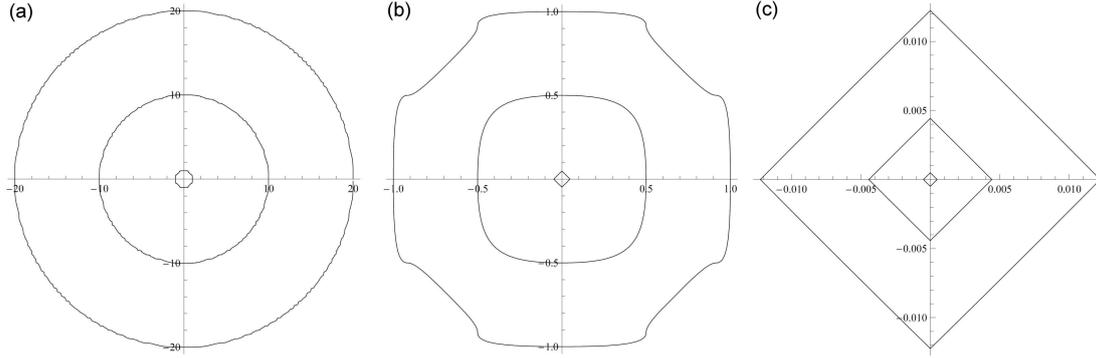


Fig. 3. Seven circles of different radii described by (80) with the arithmetic $\{\mathbb{R}, \oplus_1, \ominus_1, \odot_1, \oslash_1\}$ defined by (43) and (44). Despite appearances, all these circles are rotationally invariant. Panels to the right are the close-ups of those to the left.

Three decades after Bell’s theorem, in 1994, Itamar Pitowsky noticed that Boole’s inequalities are inequalities of a Bell-type [46].

Boole inequalities defined “possible experience” in common-sense categories appropriate for 1862. Boole’s scientific paradigm has been falsified by quantum mechanics.

If someone had asked Boole if he could give an example of a system that violates his inequalities, he probably would have answered in the negative. Treating his negative answer as the ultimate proof that Nature has to comply with Boole inequalities, we would prove that quantum mechanics is logically impossible.

Bell inequalities defined “possible experience” in common-sense categories appropriate for 1964. Grossman and Katz’s book appeared in 1972, but its implications for Bell’s theorem went unnoticed until very recently.

In light of these results, what is the actual status of all the claims about the fundamental security of quantum cryptography [47–51]? We typically base them on the belief that EPR elements of reality cannot exist. Protocols that are not based on a hidden-variable argumentation (such as the Bennett–Brassard–Mermin one [49], essentially based on rotational invariance of singlet-state probabilities) *can* be successfully attacked in non-Newtonian local hidden variable theories — non-Newtonian hidden variables are rotationally invariant because the rotation group works there by means of the *hidden* representation depicted in Fig. 3.

Furthermore, what if the Newtonian paradigm of contemporary quantum mechanics will one day be falsified by some new theory?

What if it has already been falsified? What if our enemies, whoever they may be, are well ahead of us and know systems that can mimic quantum probabilities by means of non-Newtonian hidden variables? Can they hack entangled-state quantum communication channels?

Can a no-go theorem, based on algebraic rather than probabilistic properties of quantum mechanics, cure the arithmetic loophole in quantum proofs of security?

How to guarantee that we are not in the position of German cryptographers in the 1930s, so happy with their Enigma and its security certified by appropriate theorems, while at the same time, it was systematically hacked by the Polish Cypher Bureau?

The list of open questions is longer.

12. Non-Newtonian quantum mechanics

Non-Newtonian hidden variables are not meant as an alternative to quantum mechanics.

However, non-Newtonian calculus paves the way to natural generalizations of quantum mechanics (quantum mechanics on a Cantor set is an example [42]). The resulting theory is non-Newtonian linear but Newtonian nonlinear. Such a form of nonlinear quantum mechanics [52–55] is isomorphic to the standard textbook theory, so it is free of all the difficulties that have plagued the formalisms based on nonlinear Schrödinger equations [56–59].

Yet, “mathematically isomorphic” is not synonymous with “physically equivalent”.

The following three examples illustrate the idea.

Assume $\mathbb{X} = \mathbb{Y} = \mathbb{R}$ with projective arithmetics defined by some $f_{\mathbb{X}} : \mathbb{R} \rightarrow \mathbb{R}$ and $f_{\mathbb{Y}} : \mathbb{R} \rightarrow \mathbb{R}$. Let $\psi : \mathbb{X} \rightarrow \mathbb{Y}$ be a solution of

$$H\psi(x) = \ominus_{\mathbb{Y}}\psi''(x) \oplus_{\mathbb{Y}} U(x) \odot_{\mathbb{Y}} \psi(x) = E \odot_{\mathbb{Y}} \psi(x), \tag{81}$$

where $\psi''(x)$ is the non-Newtonian second derivative. Normalization of states is assumed in the form

$$\langle \psi | \psi \rangle = \int_{(-\infty)_{\mathbb{X}}}^{\infty_{\mathbb{X}}} Dx |\psi(x)|^{2_{\mathbb{Y}}} = 1_{\mathbb{Y}} = g_{\mathbb{Y}}(1). \tag{82}$$

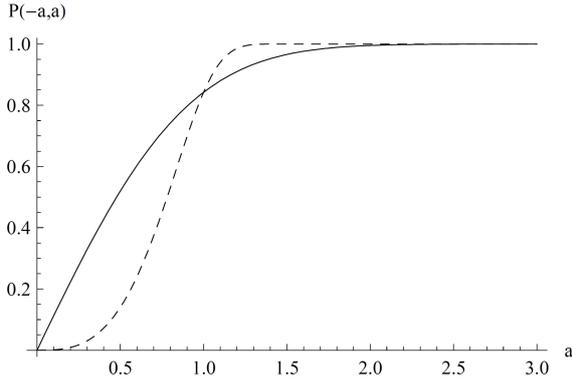


Fig. 4. Probability (88) of finding a particle in $[-a, a]$ for $0 \leq a \leq 3$, with $\tilde{\psi}(r) \sim \exp(-r^2/2)$ representing the ground state of a quantum harmonic oscillator $\tilde{U}(r) = r^2$ (in dimensionless units) for (i) the ordinary arithmetic (full line), and (ii) projective arithmetics in $\mathbb{X} = \mathbb{R} = \mathbb{Y}$ defined by $f_{\mathbb{X}}(x) = x^3$, $f_{\mathbb{Y}}(x) = x$ (dashed). The ordinary arithmetic is experimentally distinguishable from the projective one because limits of integration are $[-a^3, a^3]$ instead of the usual $[-a, a]$.

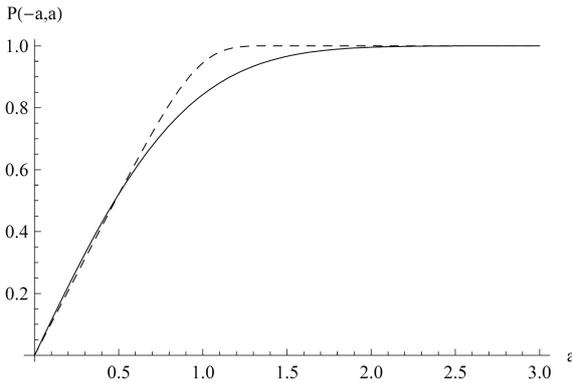


Fig. 5. The same situation as in the previous figure, but now with $f_{\mathbb{X}}(x) = x^3 = f_{\mathbb{Y}}(x)$.

(For real-valued $\psi(x)$, the modulus in

$$|\psi(x)|^{2_{\mathbb{Y}}} = \psi(x) \odot_{\mathbb{Y}} \psi(x) \quad (83)$$

is redundant, but we keep it to make the notation less awkward.) The probability of finding a particle in $[a, b] \subset \mathbb{X}$ equals

$$P(a, b) = \int_a^b D_{\mathbb{X}} |\psi(x)|^{2_{\mathbb{Y}}}. \quad (84)$$

As usual, $\psi = g_{\mathbb{Y}} \circ \tilde{\psi} \circ f_{\mathbb{X}}$, $U = g_{\mathbb{Y}} \circ \tilde{U} \circ f_{\mathbb{X}}$ (compare (32)). Let $\tilde{\psi}''(f_{\mathbb{X}}(x))$ be the Newtonian second derivative of $\tilde{\psi}$ with respect to $f_{\mathbb{X}}(x)$ so that the non-Newtonian Schrödinger equation is equivalent to the usual Newtonian equation, but with redefined parameters, i.e.,

$$f_{\mathbb{Y}}(E)\tilde{\psi}(r) = -\tilde{\psi}''(r) + \tilde{U}(r)\tilde{\psi}(r), \quad (85)$$

$$1 = \langle \tilde{\psi} | \tilde{\psi} \rangle = \int_{-\infty}^{\infty} dr |\tilde{\psi}(r)|^2. \quad (86)$$

Now let us consider $f_{\mathbb{X}}(x) = x^3$, $f_{\mathbb{Y}}(x) = x$. Then $\psi = \tilde{\psi} \circ f_{\mathbb{X}}$, $U = \tilde{U} \circ f_{\mathbb{X}}$, and the Schrödinger equation is just

$$E\tilde{\psi}(r) = -\tilde{\psi}''(r) + \tilde{U}(r)\tilde{\psi}(r), \quad (87)$$

so apparently the problem is completely equivalent to the standard one. However, due to the triviality of $f_{\mathbb{Y}}$ and the non-triviality of $f_{\mathbb{X}}$, probability (84) is now explicitly given by

$$P(a, b) = \int_{f_{\mathbb{X}}(a)}^{f_{\mathbb{X}}(b)} dr |\tilde{\psi}(r)|^2 = \int_{a^3}^{b^3} dr |\tilde{\psi}(r)|^2. \quad (88)$$

As we can see, despite the mathematical banality of the problem, the non-Diophantine arithmetic of \mathbb{X} does influence the probability of finding the particle in the interval $[a, b]$ because the integral is over $[a^3, b^3]$. Figure 4 shows the probability of finding the particle in $[-a, a]$ as a function of a .

Taking $f_{\mathbb{X}}(x) = x^3 = f_{\mathbb{Y}}(x)$, we obtain the probability depicted in Fig. 5.

As the third example consider $f_{\mathbb{X}}(x) = x$, $f_{\mathbb{Y}}(x) = x/\sqrt{|x|}$, and $g_{\mathbb{Y}}(x) = f_{\mathbb{Y}}^{-1}(x) = x^3/|x|$. Now,

$$P(a, b) = \left(\int_a^b dr |\tilde{\psi}(r)|^2 \right)^2. \quad (89)$$

The projective addition of probabilities looks here like a superposition principle from quantum mechanics,

$$P(a, c) = P(a, b) \oplus_{\mathbb{Y}} P(b, c) = \left(\sqrt{P(a, b)} + \sqrt{P(b, c)} \right)^2. \quad (90)$$

Theories based on non-Newtonian calculi involve the same physical principles, but their mathematical forms may differ from one another.

Is there any natural law that determines the form of arithmetic and calculus?

13. Towards a new paradigm

Paul Benioff, a pioneer of quantum computation, was among those physicists who believed that physics and mathematics should be logically formulated at a unified level [60, 61]. According to Benioff, physical or geometric quantities do not possess numerical values per se, but these values are introduced through “value maps”. Natural numbers are elements of any well-ordered set, and in themselves do not possess numerical values. A value map takes a number and turns it into an object with concrete numerical properties. This is somewhat similar

to the idea that “zero”, the neutral element of addition, can be, in fact, an arbitrary point $0_{\mathbb{X}} \in \mathbb{X}$, provided that \mathbb{X} can be bijectively mapped onto \mathbb{R} by means of some $f_{\mathbb{X}}$ that fulfills $f_{\mathbb{X}}(0_{\mathbb{X}}) = 0 \in \mathbb{R}$. Benioff considered only linear value maps but allowed for the possibility of value-map fields. One of his conclusions was that scalar value maps are in many respects analogous to the Higgs field [62–64].

A fundamental role of the set of bijections occurs also in Etesi’s recent reformulation of black-hole entropy [65]. The “arithmetic continuum” \mathbb{R} plays there a role of a gas subject to thermodynamic laws, while the black hole entropy is a purely set-theoretic notion, related to Gödel’s first incompleteness theorem.

The approach advocated in our work involves arithmetics and calculi constructed by means of bijections that can be regarded either as (global) nonlinear Benioff’s maps or as compositions of value maps with some bijections. More precisely, the “nonlinear” bijection $f_{\mathbb{X}} : \mathbb{X} \rightarrow \mathbb{R}$ is always linear here, but with respect to $\oplus_{\mathbb{X}}$, $\ominus_{\mathbb{X}}$, $\odot_{\mathbb{X}}$, and $\oslash_{\mathbb{X}}$. The non-Newtonian formalism in its most general form demands only the bijectivity of $f_{\mathbb{X}}$. One does not impose continuity or topological conditions on either $f_{\mathbb{X}}$ or \mathbb{X} . Note that $f_{\mathbb{X}}$ is always smooth in the topology and calculus it induces from \mathbb{R} , even if \mathbb{X} is as weird as Cantor or Sierpiński fractals. Non-Newtonian derivatives of $f_{\mathbb{X}}$ and $g_{\mathbb{X}}$ are “trivial” (equal to 1 and $1_{\mathbb{X}}$, respectively [12]) because from the point of view of the projective arithmetic in \mathbb{X} , the map $f_{\mathbb{X}}$ behaves as the identity map.

The duality between non-Newtonian linearity and Newtonian nonlinearity is one of the trademarks of the new paradigm. This is not the usual linearization of a nonlinear problem by a nonlinear change of variables. The idea can be traced back to Maslov’s superposition principle and its application to nonlinear optimization problems [15].

The resulting structure is incredibly flexible. It automatically leads to well-behaved calculi on all sets whose cardinalities are the same as the cardinality of the continuum. The resulting relativity principle (relativity of arithmetic and calculus) is much more general than the principle of general covariance.

Non-Newtonian calculus has a huge potential for the unification and systematization of various ideas scattered over mathematical and physical literature [29]. It is quite typical, however, that even if some elements of non-Newtonian thinking can be identified in those works, their arithmetic aspects are not exploited in their full generality.

For example, velocities in special relativity are added and subtracted in a projective way, but it is difficult to find a paper where repeated addition would be replaced by multiplication and its inverse — division. I have found only one place in relativistic physics where velocity $v = c \tanh(1)$, the “one” in special-relativistic projective multiplication, plays a distinguished role [66].

Kolmogorov–Nagumo averages [67, 68], the departure point of Rényi’s studies on generalized entropies [69], are exactly the averages in the sense of projective arithmetic. However, when Rényi discussed the additivity of his α -entropies, he did not think of additivity in the same sense as the one he implicitly used in Kolmogorov–Nagumo averaging. Various forms of projective arithmetic operations and derivatives have been studied in the context of generalized statistical physics and thermodynamics by Kaniadakis [70–73], but only some of the derivatives he invented were non-Newtonian, whereas the others were neither Newtonian nor non-Newtonian, a fact explaining why only the non-Newtonian ones have found applications [29]. The whole field of psychophysics is implicitly based on projective addition (see Chapter 7 in [12]). Typically, we are unaware that decibels and star magnitudes correspond to logarithmic scales because our sensory systems induce projective arithmetic in our brains, based on approximately logarithmic bijections (the Weber–Fechner law). However, although projective subtraction is here essential, the remaining three arithmetic operations are not employed. Certain elements of non-Newtonian integral calculus are present in cepstral analysis and homomorphic filtering of images [74]. Fractional derivatives can be regarded as non-Newtonian first derivatives, but only when formulated in the so-called F^{α} formalism [39]. Fuzzy calculus is non-additive but not necessarily fully non-Newtonian, and this is why the fundamental theorem of calculus does not necessarily work. Non-additive probability, somewhat similar to our non-Newtonian hidden variables (as based on non-additive Vitali and Choquet integrals), is a standard element of modern decision theory [5, 8].

Perhaps the most radical view on generalized arithmetics is due to Mark Burgin, who studied arithmetics that are *not* isomorphic to the arithmetic of natural numbers [75]. One of his goals was to replace inconsistent arithmetics (e.g., the computer arithmetic based on the notion of “machine infinity”: $\infty_M < \infty$, $\infty_M + 1 = \infty_M$) with arithmetics that are consistent but non-Diophantine.

Similarly radical is the approach of Sergeev [76], where infinities and infinitesimals are reformulated in a more intuitive and, essentially, non-Diophantine way. Here the infinity of integers is twice bigger than the infinity of natural numbers, while events of zero probability cannot happen (as opposed to the Kolmogorovian formalism based on measures) [77]. Such a new arithmetic and probabilistic paradigm often turns out to be more practical than the usual Kolmogorovian framework, just to mention Sergeev’s Infinity Calculator software.

The interference of probabilities is one of the greatest puzzles of quantum mechanics. Quantum probabilities sometimes behave as if they were negative, a situation known from projective arithmetics based on, say, $g_{\mathbb{X}}(p) = \ln(p)$. Risk aversion paradoxes in economics can be modeled by

non-additive integrals [5], but a new tendency appears where the same effects are modeled by quantum probabilities [78]. Ironically, while here we have shown that quantum probabilities can be classical but non-Newtonian, some authors are starting to replace non-additive integrals from classical economics with quantum probabilities [79].

Another aspect of interference is the superposition principle and the problem of linearity of quantum mechanics. Non-Newtonian linear Schrödinger equation can be Newtonian nonlinear. In such nonlinear quantum mechanics, the superposition principle remains the same as in the linear theory; only the meaning of “plus” and “times” is different. The same type of duality was introduced by Maslov to optimization theory [15], with the key idea that something very difficult in a nonlinear framework can become easy if we rewrite the problem in new arithmetic.

Speaking of a non-Newtonian paradigm, one typically has in mind a non-Newtonian theory of gravity (hence general relativity) or non-Newtonian mechanics (hence quantum mechanics). In the new paradigm that looms on the horizon, the term non-Newtonian may be understood in a much broader sense.

14. Conclusions

Bell overlooked the fundamental possibility of probabilistic theories based on generalized calculi and arithmetics. Both structures can be formulated at different levels, leading to a hierarchy of models related to one another by a new type of relativity principle. The violation of Bell's inequality by quantum probabilities has the same status as the paradoxes from special or general theories of relativity. They all disappear if we correctly apply appropriate relativity principles. The new framework of generalized arithmetics and calculi creates theoretical possibilities that are comparable only to those opened by the discovery of non-Euclidean geometries.

Acknowledgments

I am indebted to Kamil Nalikowski for the discussions. Calculations were carried out at the Academic Computer Center in Gdańsk. The work was supported by the CI TASK grant “Non-Newtonian calculus with interdisciplinary applications”.

References

[1] J.S. Bell, *Physics* **1**, 195 (1964).
 [2] D. Dubois, H. Prade, *Fuzzy Sets Syst.* **8**, 1 (1982).
 [3] D. Dubois, H. Prade, *Fuzzy Sets Syst.* **8**, 105 (1982).

[4] D. Dubois, H. Prade, *Fuzzy Sets Syst.* **8**, 225 (1982).
 [5] D. Schmeidler, *Econometria* **57**, 571 (1989).
 [6] R. Mesiar, *J. Math. Anal. App.* **194**, 477 (1995).
 [7] H.-J. Zimmermann, *Fuzzy Set Theory — and its Applications*, 3rd Ed., Kluwer, Boston (MA) 1996.
 [8] M. Marinacci, *Decisions Econ. Finan.* **20**, 153 (1997).
 [9] E. Pap, in: *Handbook of Measure Theory*, Vol. II, Ed. E. Pap, Elsevier, 2002, p. 1403.
 [10] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, *Aggregation Functions*, Cambridge University Press, Cambridge 2009.
 [11] D. Aerts, M. Czachor, M. Kuna, *Chaos Solitons Fractals* **91**, 461 (2016).
 [12] M. Burgin, M. Czachor, *Non-Diophantine Arithmetics in Mathematics, Physics, and Psychology*, World Scientific, Singapore 2020.
 [13] Q. Lei, Z. Xu, *Knowledge-Based Sys.* **76**, 1 (2015).
 [14] D. Aerts, M. Czachor, M. Kuna, *Rep. Math. Phys.* **81**, 357 (2018).
 [15] V.P. Maslov, *Uspekhi Mat. Nauk* **42**, 43 (1987) (in Russian).
 [16] V.N. Kolokoltsov, V.P. Maslov, *Idempotent Analysis and Applications*, Kluwer, Dordrecht 1997.
 [17] M. Gondran, M. Minoux, *Graphs, Dioids, and Semirings. New Models and Algorithms*, Springer, 2008.
 [18] G.L. Litvinov, V.P. Maslov, G.B. Shpiz, *Math. Notes* **69**, 696 (2001).
 [19] A. Einstein, B. Podolsky, N. Rosen, *Phys. Rev.* **47**, 777 (1935).
 [20] D. Bohm, *Phys. Rev.* **85**, 166 (1952).
 [21] D. Bohm, *Phys. Rev.* **85**, 180 (1952).
 [22] J.F. Clauser, M.A. Horne, *Phys. Rev. D* **10**, 526 (1974).
 [23] B.S. Cirel'son, *Lett. Math. Phys.* **4**, 93 (1980).
 [24] D.M. Greenberger, M.A. Horne, A. Shimony, A. Zeilinger, *Am. J. Phys.* **58**, 1131 (1990).
 [25] D. Bouwmeester, J.-W. Pan, M. Daniell, H. Weinfurter, A. Zeilinger, *Phys. Rev. Lett.* **82**, 1345 (1999).
 [26] M. Czachor, *Phys. Rev. A* **49**, 2231 (1994).
 [27] M. Czachor, *Quantum Stud. Math. Found.* **3**, 123 (2016).
 [28] M. Czachor, *Acta Phys. Pol. A* **139**, 70 (2021).
 [29] M. Czachor, *Entropy* **22**, 1180 (2020).

- [30] M. Burgin, [arXiv:1010.3287](https://arxiv.org/abs/1010.3287), 2010.
- [31] M. Czachor, *Found. Sci.* **25**, 971 (2020).
- [32] M. Grossman, R. Katz, *Non-Newtonian Calculus*, Lee Press, Pigeon Cove (MA) 1972.
- [33] M. Grossman, *The First Nonlinear System of Differential and Integral Calculus*, Mathco, Rockport (ME) 1979.
- [34] M. Grossman, *Bigeometric Calculus: A System with Scale-Free Derivative*, Archimedes Foundation, Rockport (ME) 1983.
- [35] E. Pap, *Zb. Rad. Prirod.-Mat. Fak. Ser. Mat.* **23**, 145 (1993).
- [36] E. Pap, *Int. J. Approx. Reasoning* **47**, 368 (2008).
- [37] A. Parvate, A.D. Gangal, *Fractals* **17**, 53 (2009).
- [38] A. Parvate, A.D. Gangal, *Fractals* **19**, 271 (2011).
- [39] A.K. Golmankhaneh, *Fractal Calculus and its Applications: F^α -calculus*, World Scientific, Singapore (2022).
- [40] M. Czachor, *Acta Phys. Pol. B* **50**, 813 (2019).
- [41] M. Czachor, *Int. J. Theor. Phys.* **56**, 1364 (2017).
- [42] M. Czachor, *Found. Sci.* **26**, 75 (2021).
- [43] M. Czachor, K. Nalikowski, *Found. Sci.* **138** (2022).
- [44] M.P. Piłat, *Found. Sci.* (2022).
- [45] G. Boole, *Phil. Trans. R. Soc.* **152**, 225 (1862).
- [46] I. Pitowsky, *British J. Phil. Sci.* **45**, 95 (1994).
- [47] C.H. Bennett, G. Brassard, in: *Proc. of the IEEE Int. Conf. on Computers, Systems, and Signal Processing, Bangalore*, 1984, p. 175.
- [48] A. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).
- [49] C.H. Bennett, G. Brassard, N.D. Mermin, *Phys. Rev. Lett.* **68**, 557 (1992).
- [50] N. Gisin, G. Ribordy, W. Tittel, H. Zbinden, *Rev. Mod. Phys.* **74**, 145 (2002).
- [51] D. Aerts, M. Czachor, M. Pawłowski, *Phys. Rev. A* **73**, 034303 (2006); Erratum: *Phys. Rev. A* **73**, 059901(E) (2006).
- [52] I. Białyński-Birula, J. Mycielski, *Ann. Phys.* **100**, 62 (1976).
- [53] R. Gähler, A.G. Klein, A. Zeilinger, *Phys. Rev. A* **23**, 1611 (1981).
- [54] S. Weinberg, *Ann. Phys.* **194**, 336 (1989).
- [55] H.D. Doebner, G.A. Goldin, *Phys. Rev. A* **54**, 3764 (1996).
- [56] N. Gisin, *Helv. Phys. Acta* **62**, 363 (1989).
- [57] J. Polchinski, *Phys. Rev. Lett.* **66**, 397 (1991).
- [58] M. Czachor, *Found. Phys. Lett.* **4**, 351 (1991).
- [59] M. Czachor, H.-D. Doebner, *Phys. Lett. A* **301**, 139 (2002).
- [60] P. Benioff, *Found. Phys.* **32**, 989 (2002).
- [61] P. Benioff, *Found. Phys.* **35**, 1825 (2005).
- [62] P. Benioff, *Quant. Stud. Math. Found.* **2**, 289 (2015).
- [63] P. Benioff, *Quant. Inf. Proc.* **15**, 1081 (2016).
- [64] P. Benioff, *Quant. Inf. Proc.* **15**, 3005 (2016).
- [65] G. Etesi, *Found. Sci.* **25**, 327 (2020).
- [66] M. Czachor, [arXiv:2209.06070](https://arxiv.org/abs/2209.06070), 2022.
- [67] A.N. Kolmogorov, *Atti. Acad. Naz. Lincei. Rend.* **12**, 388 (1930), reprinted in: *Selected Works of A.N. Kolmogorov, Vol. 1, Mathematics and Mechanics*, Ed. V.M. Tikhomirov, Kluwer, Dordrecht 1991.
- [68] M. Nagumo, *Japan J. Math.* **7**, 71 (1930), reprinted in: *Mitio Nagumo Collected Papers*, Eds. M. Yamaguti, L. Nirenberg, S. Mizohata, Y. Sibuya, Springer, Tokyo 1993.
- [69] A. Rényi, *MTA III. Oszt. Közl.* **10**, 251 (1960), reprinted in: *Selected Papers of Alfréd Rényi*, Ed. P. Turán, Akadémiai Kiadó, Budapest 1976.
- [70] G. Kaniadakis, *Physica A* **296**, 405 (2001).
- [71] G. Kaniadakis, *Phys. Lett. A* **288**, 283 (2001).
- [72] G. Kaniadakis, *Phys. Rev. E* **66**, 056125 (2002).
- [73] G. Kaniadakis, *Entropy* **15**, 3983 (2013).
- [74] A.V. Oppenheim, R.W. Schaffer, *IEEE Signal Process. Mag.* **21**, 95 (2004).
- [75] M.S. Burgin, *Uspekhi Mat. Nauk* **32**, 209 (1977) (in Russian).
- [76] Y.D. Sergeyev, *EMS Surv. Math. Sci.* **4**, 219 (2017).
- [77] C.S. Calude, M. Dumitrescu, *SN Comput. Sci.* **1**, 36 (2020).
- [78] D. Aerts, S. Sozzo, J. Tapia, in: *Quantum Interaction*, Lecture Notes in Computer Science, vol. 7620, 2012 p. 48.
- [79] J.R. Busemeyer, P.D. Bruza, *Quantum Models of Cognition and Decision*, Cambridge University Press, Cambridge 2014.

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

Fock State Sampling Method — Characteristic Temperature of Maximal Fluctuations for Interacting Bosons in Box Potentials

M.B. KRUK^{a,b}, T. VIBEL^c, J. ARLT^c,
P. KULIK^a, K. PAWŁOWSKI^{a,*} AND K. RZAŻEWSKI^a

^aCenter for Theoretical Physics, Polish Academy of Sciences, al. Lotników 32/46, 02-668 Warsaw, Poland

^bInstitute of Physics, Polish Academy of Sciences, al. Lotników 32/46, 02-668 Warsaw, Poland

^cCenter for Complex Quantum Systems, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S171](https://doi.org/10.12693/APhysPolA.143.S171)*e-mail: pawlowski@cft.edu.pl

We study the statistical properties of a gas of interacting bosons trapped in a box potential in two and three dimensions. Our primary focus is the characteristic temperature T_p , i.e., the temperature at which the fluctuations of the number of condensed atoms (or, in 2D, the number of motionless atoms) are maximal. Using the Fock state sampling method, we show that T_p increases due to interaction. In 3D, this temperature converges to the critical temperature in the thermodynamic limit. In 2D, we show the general applicability of the method by obtaining a generalized dependence of the characteristic temperature on the interaction strength. Finally, we discuss the experimental conditions necessary for the verification of our theoretical predictions.

topics: Bose–Einstein condensate, statistical ensembles, quantum gases

1. Introduction

The statistical properties of interacting ultracold gases of bosonic atoms and, in particular, of Bose–Einstein condensates (BEC) remain a considerable challenge of current interest. While the statistical properties of non-interacting gases are well described by a number of methods, a soluble model for interacting bosons exists only in one dimension. In two and three dimensions, reliable results are only available for weakly interacting gases at low temperatures within the Bogoliubov approximation. Hence, the dependence of the critical temperature on interaction remains a challenging issue.

Over the years, a large number of mutually exclusive predictions of the change of the critical temperature due to interactions were made [1–11]. Practically all of them dealt with a gas trapped in a three-dimensional cubic box potential. The conflicting results are summarized in Table I [2–14]. Note that even the sign of the correction was

uncertain initially. Later the consensus emerged that in the thermodynamic limit, the shift to the critical temperature is $\Delta T_c \approx 1.3 a \rho^{1/3}$, where ρ is the gas density, and a is the s -wave scattering length.

During the struggle to compute the shift of the critical temperature, a number of theoretical methods were used (see the review [11]). The correct result was eventually obtained by using the classical field approximation (CFA) [10]. The CFA method itself suffers from a cut-off problem, which was cleverly overcome in [10]. Recently, we proposed yet another method based on a direct quantum description of the system and the definitions of the statistical ensembles. The method, called the Fock state sampling (FSS) method [15], is presented in Sect. 2.

Most BEC experiments to date are performed with harmonic traps. However, recently Bose–Einstein condensates were created in nearly perfect box potentials [16]. Nonetheless, experimental verification of the theoretical prediction remains chal-

TABLE I

Coefficient c of the shift of the critical temperature $\Delta T_c/T_c = c a \rho^{1/3}$ obtained by various analytic and numerical methods. See also the review by J. Andersen [11]. In [14], four different many-body methods were used yielding results in the range $c \in [1, 6.7]$.

Authors	Ref.	Coefficient c
Grueter et al. (1997)	[2]	0.34 ± 0.06
Holzmann et al. (1999)	[3]	0.7
Holzmann et al. (1999)	[4]	2.3 ± 0.25
Baym et al. (1999)	[5]	2.9
Wilkens et al. (2000)	[6]	-0.93
Arnold et al. (2000)	[7]	1.71
Baym et al. (2000)	[8]	2.33
Souza Cruz et al. (2001)	[9]	3.059
Kashurnikov et al. (2001)	[10]	1.29 ± 0.05
Davis et al. (2003)	[12]	1.3 ± 0.4
Nho et al. (2004)	[13]	1.32 ± 0.14
Watabe et al. (2013)	[14]	1 to 6.7

lenging. One of the problems is due to the fact that experiments are performed with a finite number of atoms, and for such a system, there is no unique way of determining the critical temperature. Namely, for a finite-size system, the number of condensed atoms is an analytic function of temperature, and thus there is not a definite value of temperature beyond which the number of condensed atoms is strictly zero. The remedy for this difficulty, proposed in [17], is to study the temperature of maximal variance of the number of condensed atoms instead of the critical temperature. The temperature of the maximal variance tends to the critical temperature in the thermodynamic limit [17] and is well-defined for finite-size systems, which makes it applicable to gases exhibiting only quasi-condensation.

Experimentally, it is more demanding to measure the fluctuations of the condensate atom number than the mean of this number. However, the experimental difficulties were recently overcome due to a stabilization technique of the evaporation process [18], allowing for a measurement of the fluctuations [19]. Furthermore, it was shown that the canonical ensemble fails to describe the experimental situation, and one must invoke the microcanonical one [20]. These experiments directly measure the temperature of the maximal fluctuations rather than the temperature at which the condensate vanishes.

It is the purpose of this paper to discuss the interaction-induced shifts of the temperature of maximal fluctuations, which is referred to as the characteristic temperature T_p . Based on the FSS method, we provide, to our knowledge for the first time, results for a bosonic gas in a box potential in the microcanonical ensemble.

The paper is organized as follows. In Sect. 2, we briefly review the FSS method. Section 3 applies the method to a gas trapped in the three-dimensional box potential in both the canonical and microcanonical ensembles. In Sect. 4, the case of the two-dimensional box potential is discussed. Note that there is no phase transition and no critical temperature in this case, and nonetheless, the characteristic temperature can be defined. Section 5 concludes the discussion and provides an outlook on future experiments.

2. Fock state sampling method

We consider N bosonic atoms trapped in a box potential with periodic boundary conditions and interacting via short-range interaction potential. The Hamiltonian of the system is

$$\hat{H} = -\frac{\hbar^2}{2m} \int d^d r \hat{\Psi}^\dagger(\mathbf{r}) \nabla^2 \hat{\Psi}(\mathbf{r}) + \frac{g_d}{2} \int d^d r \hat{\Psi}^\dagger(\mathbf{r}) \hat{\Psi}^\dagger(\mathbf{r}) \hat{\Psi}(\mathbf{r}) \hat{\Psi}(\mathbf{r}), \quad (1)$$

where $\hat{\Psi}(\mathbf{r})$ is a bosonic annihilation operator, m is a mass, and $g_d > 0$ is a coupling constant related to short-range interactions. In Sect. 3, we consider three-dimensional systems, where the coupling constant is $g_{d=3} \equiv g_{3D} = 4\pi\hbar^2 a/m$ and a is the scattering length. In Sect. 4, devoted to two-dimensional systems, we use the notation $g_{d=2} \equiv g_{2D}$. In the case of a box potential with periodic boundary conditions, the macroscopically occupied orbital (the BEC wave function in 3D) is just a constant function, i.e., a plane wave with momentum 0.

In what follows, we focus on the fluctuations of the number of atoms in BEC at finite temperature. We use the canonical and the microcanonical ensembles, which were shown to be close to the experimental reality [19, 20].

There are several different ways to describe the statistical properties of ultra-cold Bose gases theoretically. In this paper, we sample many-body states to generate a set of copies that properly approximates the canonical ensemble of a gas. Given a sufficiently large set of copies, the expectation values are defined as the average over the set. By post-selecting the set, we obtain results in the microcanonical ensemble. To define the appropriate Metropolis algorithm [21], we need to define “the stage”, which is the set of available states, and the “Metropolis dynamics” or the specific algorithm defining the Markov chain generating the approximation to the canonical ensemble.

2.1. Setting “the stage”

All states of N particles belong to the suitable Hilbert space. A convenient parametrization is provided by the basis of single-particle states in the

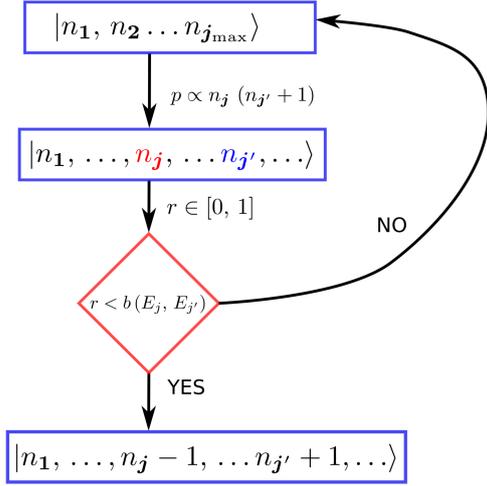


Fig. 1. Single step of the FSS method: one draws two states — one from which an atom might be taken (index j) and one in which the atom may land (index j') with probability distribution proportional to $n_j(n_{j'} + 1)$. The new state is accepted only if a random number r drawn from a uniform distribution in $[0, 1]$ is smaller than the Boltzmann factor b given in (5).

trapping potential. Since we consider box potentials with periodic boundary conditions, the basis states are just plane waves

$$f_{\mathbf{j}}(\mathbf{r}) := f_{j_x j_y j_z}(x, y, z) = \frac{1}{\sqrt{L_x L_y L_z}} e^{-i 2\pi(j_x x/L_x + j_y y/L_y + j_z z/L_z)}, \quad (2)$$

where L_x , L_y , and L_z are the lengths of the box potential and j_x , j_y , and j_z are positive and negative integers. It is worth stressing that due to translational invariance, these states remain eigenstates of the single-particle density matrix also for the interacting gas. Thus, the constant function $j_x = j_y = j_z = 0$ remains the condensate state also in the presence of interactions.

The space of all N -particle states is spanned by the Fock states

$$|\mathbf{n}\rangle := |n_1, n_2, \dots\rangle, \quad (3)$$

where n_j denotes the number of bosons in the single-particle state $f_{\mathbf{j}}(\mathbf{r})$. In the canonical ensemble, we fix the total number of atoms N and consider only the Fock states that contain N bosons

$$\sum_{\mathbf{j}} n_{\mathbf{j}} = N. \quad (4)$$

The whole Hilbert space contains all superpositions of all N -particle Fock states. The appropriate parameters are far too numerous for any efficient numerics. Instead, we restrict our set of available states just to the Fock states in (3), not accounting for their superpositions. This has two consequences. First, it neglects the phenomenon of quantum depletion. Thus, the method is expected to yield correct

results only for weak interactions. Second, it is not applicable to weakly interacting bosons confined in a harmonic trap, since, in this case, the condensate wave function is a superposition of many oscillator states.

2.2. Metropolis dynamics

The following algorithm defines our Markov chain used to generate the elements of our representation of the canonical ensemble. A single step of this algorithm is also shown in Fig. 1.

Each particle has the same probability of jumping out of a given single-particle state. The probability of jumping out is proportional to the number of particles in that state. The probability of landing in a given single-particle state is proportional to its occupation (stimulated process) plus one (to account for the spontaneous process). The acceptance criterion, usual for the Metropolis algorithm, is based on comparing a random number $0 < r < 1$ drawn from a uniform distribution versus the Boltzmann factor b of the initial and the final states

$$b(E_{\text{current}}, E_{\text{candidate}}) = e^{-\beta(E_{\text{current}} - E_{\text{candidate}})}, \quad (5)$$

where $\beta = 1/(k_B T)$, k_B is the Boltzmann constant and T is the temperature. The energy E_{current} is the expectation value of the Hamiltonian in the current Fock state. It is the sum of the kinetic energy and the interaction energy. The kinetic energy is simply

$$E_{\text{kin}} = \sum_{\mathbf{j}} n_{\mathbf{j}} e_{\mathbf{j}}, \quad (6)$$

where $e_{\mathbf{j}}$ is the energy of the \mathbf{j} -th level, i.e.,

$$e_{\mathbf{j}} := \frac{2\pi^2 \hbar^2}{m} \left[\left(\frac{j_x}{L_1} \right)^2 + \left(\frac{j_y}{L_2} \right)^2 + \left(\frac{j_z}{L_3} \right)^2 \right]. \quad (7)$$

The short-range interaction energy, which in the general case is a nontrivial quadratic form, reduces to a single sum in the case of a box potential and is averaged in a single Fock state

$$E_{\text{int}} = \left\langle \mathbf{n} \left| \frac{g_{3D}}{2} \int d^3r \hat{\Psi}^\dagger(\mathbf{r}) \hat{\Psi}^\dagger(\mathbf{r}) \hat{\Psi}(\mathbf{r}) \hat{\Psi}(\mathbf{r}) \right| \mathbf{n} \right\rangle = \frac{g_{3D}}{2} \left(2N(N-1) - \sum_{\mathbf{j}} n_{\mathbf{j}}^2 \right). \quad (8)$$

Moreover, for comparison of the Boltzmann factors, only the difference of the energies of the final and initial state enters (see (5))

$$E_{\text{current}} - E_{\text{candidate}} = e_{\mathbf{j}} - e_{\mathbf{j}'} + g_{3D} (n_{\mathbf{j}'} - n_{\mathbf{j}}), \quad (9)$$

where \mathbf{j} and \mathbf{j}' are the indices of the single-particle states from which the atom escaped and in which it lands, respectively. A single step of this algorithm is presented in Fig. 1.

The algorithm satisfies the detailed balance principle and guarantees access to all important N particle states. When the number of steps goes to infinity, the expectation value of any physical quantity

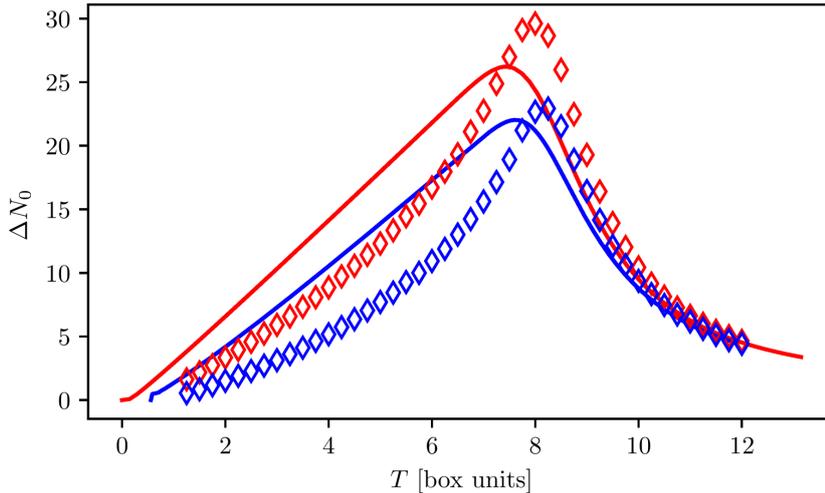


Fig. 2. Standard deviation of the number of condensed atoms confined in a 3D box potential with periodic boundary conditions for $N = 300$ atoms for a non-interacting (solid lines) and an interacting gas with a gas parameter $a\rho^{1/3} = 0.05$ (empty diamonds). Red symbols indicate canonical and blue microcanonical results. The solid lines are exact, obtained with the help of recurrence relations. The empty diamonds represent canonical and microcanonical results obtained with the help of the Fock state sampling method. Note similar shifts of the characteristic temperature in both statistical ensembles. Also, note that the maximal fluctuations are increased by interactions.

does not depend on the state used to initiate the algorithm. In practice, we perform only a finite number of steps and discard approximately $10N$ initial steps during which the quantities of interest are not only fluctuating but also drifting.

Importantly, the Fock state sampling method also offers access to the microcanonical ensemble. This is accomplished simply by reducing the number of states to those with energy in the small interval around the most probable one.

In the following, we present the results obtained with Fock state sampling method (FSSM) for a gas in a 3D (Sect. 3) and 2D (Sect. 4) box potential, including interactions between the atoms.

3. Characteristic temperature for gas in 3D box potential

We illustrate the results of our method for ultracold gases of bosonic atoms in a 3D box potential in Fig. 2. The figure shows the temperature dependence of the standard deviation of the number of atoms in a Bose–Einstein condensate ΔN_0 for both canonical and microcanonical ensembles. The results for the non-interacting gas are exact and obtained with the recurrence relations [6], while the results for the interacting gas are obtained with the FSS method described in the previous section. Note that the microcanonical ensemble yields significantly lower fluctuations. Moreover, interactions increase the peak fluctuations in both ensembles, and similarly, the temperature of the maximal fluctuations is increased.

The related shift of the critical temperature due to collisions in weakly interacting Bose gas in a 3D box potential has been the subject of a longstanding debate, as outlined in the introduction. The final result for the correction was obtained with a sophisticated numerical method [10], based on techniques developed over the past 20 years. Our method, on the contrary, although approximate, is simple to implement for as many as 10 000 atoms.

Here, we study the temperature of the maximal fluctuations T_p instead of looking at the critical temperature T_c . This characteristic temperature is well-defined for systems with a finite number of atoms. Moreover, it has been recently shown that it can be measured for a Bose gas in a harmonic trap [19, 20].

Thus, the main quantity of interest is the interaction-induced relative shift to the characteristic temperature T_p

$$\delta T_p(N, a) := \frac{T_p(N, a) - T_p(N, 0)}{T_p(N, 0)}, \quad (10)$$

where $T_p(N, a)$ is the characteristic temperature for a gas with N atoms interacting with the s -wave scattering length a . To find the dependence of δT_p on N and a , we study the system with atom numbers ranging from 100 to 10000 and interaction strengths g corresponding to gas parameters $a\rho^{1/3}$ from 0.005 to 0.02. All temperatures are given in units of $2\pi^2\hbar^2/(mk_B L^2)$.

The results are illustrated in Fig. 3, which shows the relative standard deviation of the number of condensed atoms as a function of temperature for different total numbers of atoms and various interaction strengths g . Since the main focus is the shift

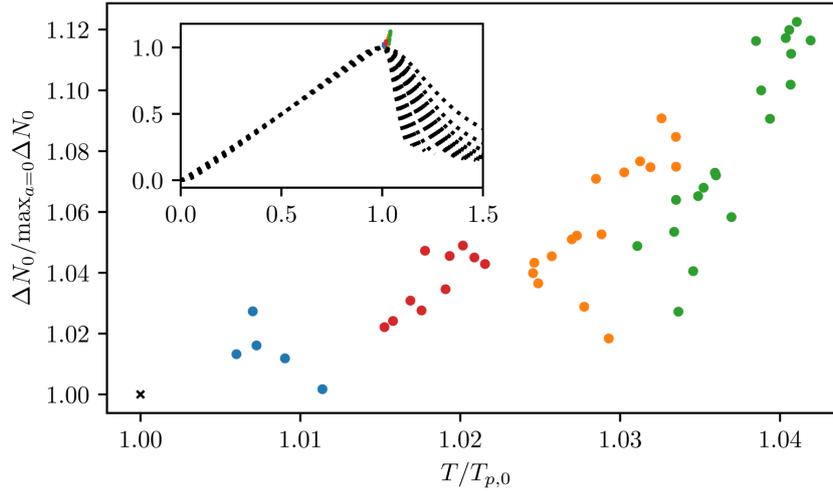


Fig. 3. Relative standard deviation of the BEC atom number for an interacting gas (colored points) for various total numbers of atoms, and interaction strengths. All points are rescaled by the maximal value of the fluctuations in the non-interacting case. Results for systems with the same gas parameter are marked with the same color, i.e., $a\rho^{1/3}$ corresponds to 0.005 (blue), 0.01 (red), 0.015 (orange), 0.02 (green). The quantity $T_{p,0}$ is the temperature of maximal fluctuations of the non-interacting gas and the symbol “x” marks the reference point — the maximal BEC fluctuations of the non-interacting gas. The atom number is in the range $N = 100, 200, \dots, 1000, 2000, \dots, 10000$. The inset shows an overview of the entire temperature range with the results for the non-interacting gas (dashed lines).

of the characteristic temperature T_p , all results for the non-interacting gas are normalized in terms of both the maximal value and its temperature. The same scaling factors are used for the results for the interacting gas, i.e., the temperature (relative standard deviation) is divided by the temperature $T_{p,0}$ (maximal relative standard deviation) of the non-interacting gas with the same number of atoms. After rescaling, one can easily follow the interaction-induced shifts. The points of a given color show the maximal variance at the characteristic temperature for various atom numbers N and scattering lengths a , but a common gas parameter $a\rho^{1/3}$.

Note that the maximal relative standard deviations are grouped into small regions for a common gas parameter. For larger gas parameters corresponding to larger interactions, the maximal relative standard deviations are larger and are reached at higher temperatures as compared to the non-interacting gas.

We fit the average shift for each characteristic temperature with a linear dependence on the gas parameter and obtain

$$\delta T_p \approx (2.039 \pm 0.014) a\rho^{\frac{1}{3}} \quad (11)$$

in the range of the number of atoms N between 4000 and 10000. Thus the scaling is similar to the one obtained for the critical temperature of an infinitely large system, while the prefactor is larger.

Note that the maximal relative standard deviations for a common gas parameter form elongated regions, indicating that the maximal variance and the characteristic temperature may have a further

dependence on the scattering length and density. The remaining spread of the points indicates the precision of our method.

Also note that interactions increase the maximal fluctuations in this case. This point has also been the subject of a long-standing controversy (see, for instance, the inset in Fig. 4 in [19]). It was recently addressed using the FSS method [15], showing that the size of the fluctuations depends on all system parameters, and thus it is not possible to generalize the effect of interactions on the magnitude of condensate fluctuations.

Importantly, the characteristic temperature discussed in this section is also well-defined for systems that do not exhibit a phase transition in a thermodynamic limit. An example of such a system is a gas in a 2D box potential, discussed in the next section.

4. Characteristic temperature in 2D box potential

It is well known that Bose–Einstein condensation appears as a phase transition for sufficiently high dimensions. In a box potential, the phase transition occurs only in three dimensions, while it is absent in lower dimensionality. Despite the fact that there is no phase transition and, therefore, no critical temperature in the case of two dimensions, the notion of the characteristic temperature T_p , marking the temperature of maximal fluctuations, is still applicable. Of course, in the absence of the phase transition, interactions still affect the fluctuations.

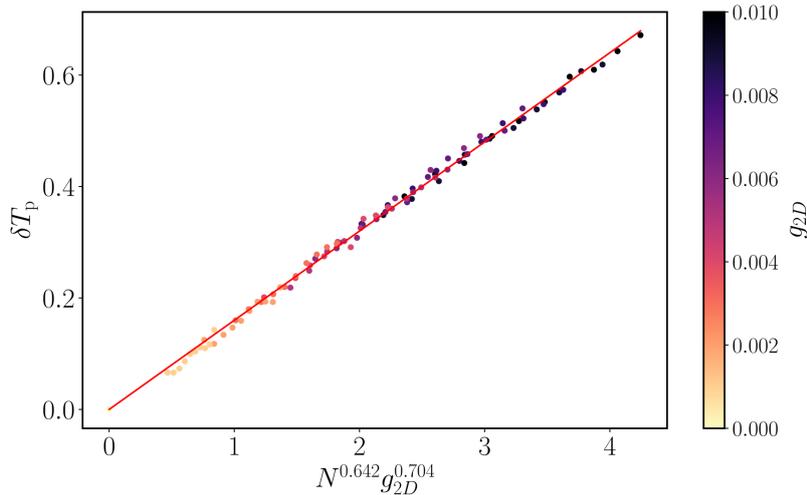


Fig. 4. Relative shift of the characteristic temperature for Bose gases in a 2D box potential with periodic boundary conditions due to interactions. The coupling constant g_{2D} was varied from 0 to 0.01 and the number of atoms was adjusted from 600 to 1500.

We illustrate this by investigating a two-dimensional box potential with periodic boundary conditions. The calculation is analogous to the 3D box potential, and the condensate wave function is still a constant one, regardless of interaction. The algorithm, after omitting all Z -dependent variables, is identical to the one introduced in the previous section. The relative shift of the characteristic temperature due to interactions was calculated for various atom numbers N and interaction strengths g_{2D} , as defined in the Hamiltonian (1). In this section, we do not refer to the gas parameter, which would be more complicated than in the 3D case.

Figure 4 shows this shift as a function of $g_{2D}^\alpha N^\beta$ with optimally chosen exponents α and β obtained from a fit to the data, yielding

$$\delta T_p^{(2D)}(N, g_{2D}) := (0.16 \pm 0.03) N^{0.642 \pm 0.015} g_{2D}^{0.704 \pm 0.006}. \quad (12)$$

The stated errors may be reduced at the expense of the numerical effort. The convergence is very slow, and errors scale with the square root of the number of Metropolis steps.

The results presented in this section illustrate the power and generality of the FSS method. The method is conceptually very simple, and its successful application merely requires a numerical effort.

5. Conclusions

In conclusion, we have investigated the fluctuations of the ideal and weakly interacting Bose-Einstein condensates trapped in box potentials with periodic boundary conditions. The temperature of maximal BEC atom number fluctuations T_p was analyzed under various conditions. The advantage of

T_p over T_c lies in the fact that it is unambiguously defined also for a finite system, and it can be studied also in systems that do not exhibit phase transition.

In our study, we used the Fock state sampling method, which turns out to be easy to use, exact for the non-interacting system (see [15]), and applicable to a wide range of problems. With this method, we found the shift of the characteristic temperature in the 3D box potential to be $\approx 2.03 a \rho^{1/3}$, where a is the scattering length, and ρ is the gas density. This is reasonably close to the expected shift of the critical temperature in this system $\approx 1.3 a \rho^{1/3}$.

We also applied our method to a two-dimensional system and obtained a generalized dependence of the characteristic temperature on the interaction strength and atom number $\approx 0.16 N^{0.642} g_{2D}^{0.704}$, showing the applicability in a system that does not exhibit a phase transition.

Experimentally, the recent realization of box potentials provides an opportunity to address the predictions presented above. In particular, combining box potentials with atomic species that allow for tunability of the interaction strength will provide access to a wide variation of gas parameters.

Box potentials are typically created using blue-detuned light to form the walls of a box, such as, e.g., a hollow beam with two narrow light sheets as end caps [22]. To flatten the bottom of the potential, gravity must be compensated using a magnetic field gradient. Alternatively, a light field with a linearly varying intensity produced by an acousto-optic deflector can be used [23]. The necessary beam shapes for box potentials can be generated using spatial light modulators, digital micromirror devices, or specialized optical elements such as axicons.

Tunability of the scattering length would be highly beneficial to isolate the effect of interactions on the characteristic temperature and the

magnitude of atom number fluctuations. This can be achieved by adjusting the magnetic field near a Feshbach resonance [24]. Since a magnetic field gradient is the most common way to cancel gravity within a box potential, this will typically necessitate independent control of the magnetic field gradient and its mean value. The field gradient will thus introduce a spatial dependence of the scattering length, and hence atomic species with broad resonances such as, e.g., the bosonic isotopes of potassium should be used.

Furthermore, it is important to distinguish between BEC and the thermal part of partially condensed atomic clouds to measure the atom number fluctuations. Thus, the bimodality of the momentum distribution is crucial for determining the BEC number and the number of thermal atoms. Fortunately, both the bimodality and an appropriate fitting function for the thermal cloud have been confirmed [16] experimentally.

The most significant outstanding challenge towards the measurement of fluctuations proposed here is the combination of box potentials with atom number stabilization. There are two primary technical sources of variations in the total number of atoms. The first one is due to the statistical nature of evaporative cooling, which relates the atom number to the temperature. This is predictable and can be accounted for in the evaluation of atom number fluctuations. The second source of atom number variation is typically due to various technical noise sources in the experiment and should be minimized since it can distort the measured atom number fluctuation when different mean values of the BEC atom number are probed. Thus, to conduct the experiment, it will be necessary to combine the box potential with atom number stabilization. However, it is not yet clear whether it is sufficient to stabilize the atom number before loading the cloud into the box potential or if methods for stabilization within the box potential must be developed.

The combination of the Fock state sampling method and current experimental developments will allow for further experiments in the near future. Especially since the FSSM can provide precise predictions for experimentally relevant atom numbers in a variety of potentials, the time has now come for a new generation of experiments on these fundamental questions.

Acknowledgments

We dedicate this paper to Professor Iwo Białynicki-Birula on the occasion of his 90th birthday. The Polish authors represent three generations of his students. We owe him a lot in terms of knowledge and style of doing research. Moreover, we pursue our research careers at the Center for Theoretical Physics PAS founded by Professor Białynicki-Birula. Working at the institute, under the influence

of Iwo Białynicki-Birula, formed us as physicists. His papers are exceptionally clearly written, and we strive to be similarly clear in our work.

We thank P. Deuar for the fruitful discussions.

M.B.K. acknowledges support from the (Polish) National Science Center Grants No. 2018/31/B/ST2/01871 and No. 2022/45/N/ST2/03511. K.P. acknowledges support from the (Polish) National Science Center Grant No. 2019/34/E/ST2/00289. K.R., P.K., and M.B.K. acknowledge support from the (Polish) National Science Center Grant No. 2021/43/B/ST2/01426. The Center for Theoretical Physics of the Polish Academy of Sciences is a member of the National Laboratory of Atomic, Molecular and Optical Physics (KL FAMO).

T.V. and J.A. acknowledge support from the Danish National Research Foundation through the Center of Excellence (Grant Agreement No. DNR156) and from the Independent Research Fund Denmark-Natural Sciences via Grant No. 8021-00233B and 0135-00205B.

References

- [1] T.D. Lee, C.N. Yang, *Phys. Rev.* **105**, 1119 (1957).
- [2] P. Grüter, D. Ceperley, F. Laloë, *Phys. Rev. Lett.* **79**, 3549 (1997).
- [3] M. Holzmann, P. Grüter, F. Laloë, *Eur. Phys. J. B* **10**, 739 (1999).
- [4] M. Holzmann, W. Krauth, *Phys. Rev. Lett.* **83**, 2687 (1999).
- [5] G. Baym, J.-P. Blaizot, M. Holzmann, F. Laloë, D. Vautherin, *Phys. Rev. Lett.* **83**, 1703 (1999).
- [6] M. Wilkens, F. Illuminati, M. Krämer, *J. Phys. B: At. Mol. Opt. Phys.* **33**, L779 (2000).
- [7] P. Arnold, B. Tomášik, *Phys. Rev. A* **62**, 063604 (2000).
- [8] G. Baym, J.-P. Blaizot, J. Zinn-Justin, *Europhys. Lett.* **49**, 150 (2000).
- [9] F.F. de Souza Cruz, M.B. Pinto, R.O. Ramos, *Phys. Rev. B* **64**, 014515 (2001).
- [10] V.A. Kashurnikov, N.V. Prokof'ev, B.V. Svistunov, *Phys. Rev. Lett.* **87**, 120402 (2001).
- [11] J.O. Andersen, *Rev. Mod. Phys.* **76**, 599 (2004).
- [12] M.J. Davis, S.A. Morgan, *Phys. Rev. A* **68**, 053615 (2003).
- [13] K. Nho, D.P. Landau, *Phys. Rev. A* **70**, 053614 (2004).
- [14] S. Watabe, Y. Ohashi, *Phys. Rev. A* **88**, 053633 (2013).

- [15] M.B. Kruk, D. Hryniuk, M. Kristensen, T. Vibel, K. Pawłowski, J. Arlt, K. Rzażewski, *SciPost Phys.* **14**, 036 (2023).
- [16] A.L. Gaunt, T.F. Schmidutz, I. Gotlibovych, R.P. Smith, Z. Hadzibabic, *Phys. Rev. Lett.* **110**, 200406 (2013).
- [17] Z. Idziaszek, K. Rzażewski, *Phys. Rev. A* **68**, 035604 (2003).
- [18] M. Gajdacz, A.J. Hilliard, M.A. Kristensen, P.L. Pedersen, C. Klempt, J.J. Arlt, J.F. Sherson, *Phys. Rev. Lett.* **117**, 073604 (2016).
- [19] M. Kristensen, M. Christensen, M. Gajdacz, M. Igllicki, K. Pawłowski, C. Klempt, J. Sherson, K. Rzażewski, A. Hilliard, J. Arlt, *Phys. Rev. Lett.* **122**, 163601 (2019).
- [20] M. Christensen, T. Vibel, A. Hilliard, M. Kruk, K. Pawłowski, D. Hryniuk, K. Rzażewski, M. Kristensen, J. Arlt, *Phys. Rev. Lett.* **126**, 153601 (2021).
- [21] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [22] N. Navon, R.P. Smith, Z. Hadzibabic, *Nat. Phys.* **17**, 1334 (2021).
- [23] K. Shibata, H. Ikeda, R. Suzuki, T. Hirano, *Phys. Rev. Res.* **2**, 013068 (2020).
- [24] C. Chin, R. Grimm, P. Julienne, E. Tiesinga, *Reviews of Modern Physics* **82**, 1225 (2010).

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

Dynamical Quantum Phase Transitions from Quantum Optics Perspective

J. ZAKRZEWSKI^{a,b,*}^a*Institut Fizyki Teoretycznej, Wydział Fizyki, Astronomii i Informatyki Stosowanej, Uniwersytet Jagielloński, Profesora Stanisława Łojasiewicza 11, PL-30-348 Kraków, Poland*^b*Mark Kac Complex Systems Research Center, Uniwersytet Jagielloński, Kraków, Poland*

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S179](https://doi.org/10.12693/APhysPolA.143.S179)*e-mail: jakub.zakrzewski@uj.edu.pl

In this work dedicated to Professor Iwo Białynicki-Birula on the occasion of his 90th birthday, I attempt to show that dynamical quantum phase transitions observed as singularities in the Loschmidt rate dynamics bear a close resemblance to the standard Rabi oscillations known from the dynamics of two-level systems. For some many-body systems, this analogy may go even further, and the behaviour observed for example transverse Ising chain can be directly mapped to such simple dynamics. A simple link between Loschmidt echo singularities and quantum scars is also suggested.

topics: phase transitions, Loschmidt echo singularities, Rabi oscillations, quantum scars

1. Introduction

The physics of complex systems may sometimes be understood (in particular limiting cases) in a simple, enlightening form. This has been often demonstrated in quantum optics, one of the many areas of Iwo Białynicki-Birula outstanding contributions. As a scientific grandson of Iwo Białynicki-Birula, I had relatively small overlap in scientific interests with him, our paths crossed for a moment in the studies of nonspreading wave-packets [1–4]. Still, however, I profited a lot from occasional conversations as well as participation, from time to time, in unusually vivid seminars with his active participation. Often his aim was to find a simple picture of the presented effects. In this contribution, I consider briefly two cases from studies of nonequilibrium dynamics of many-body systems which may be, in my opinion, understood in simple terms: dynamical quantum phase transitions (DQPT) [5, 6] and quantum many-body scars (QMBS) dynamics [7].

The simplest definition of DQPT consists of a sudden quench in which the system is prepared in the ground state $|\Psi\rangle$ of a parameter-dependent

Hamiltonian $H(\lambda = 0)$, and λ is suddenly changed to other value. It has been observed that often if a change of λ moves the Hamiltonian into a different phase, the time dynamics with $H(\lambda)$ of the now nonstationary state after quench reveals the so-called Loschmidt echo singularities. Their appearance is neither a necessary nor a sufficient condition for the phase transition between $H(0)$ and $H(\lambda)$. Still, a predominantly lack of singularities occurs if no phase transition is crossed while changing λ and vice versa.

Dynamical detection [7] of QMBS is in some sense similar. One prepares an initial nonstationary state for the many-body system described by $H(\lambda)$. When this initial state has a significant overlap with a few almost equally spaced in energy eigenstates of $H(\lambda)$, the time evolution of the observables reveals oscillations even in a weakly ergodic regime, i.e., when the dynamics of a typical generic state will lead to thermalization.

Both these phenomena, while of current interest, can be simply explained by identifying a “essential state model”, i.e., a minimal approximate level scheme allowing one to simulate the dynamics. Let us first consider DQPT in the seminal example of the transverse Ising model.

2. DQPT in transverse Ising model

The first work on DQPT [5] considers the transverse Ising model with the Hamiltonian of the form

$$H = -\frac{1}{2} \sum_i \sigma_i^z \sigma_{i+1}^z - \frac{g}{2} \sum_i \sigma_i^x, \quad (1)$$

where g is the strength of the magnetic field pointing in the Ox direction. For small g , the interactions favor the ferromagnetic (FM) orientation (along Oz), with the two degenerate (in the thermodynamic limit) ground states given for $g \rightarrow 0$ by $|\psi(\pm)^z\rangle = \prod_i |\pm\rangle_i^z$, where $|\pm\rangle_i^z$ denotes the eigenvectors of σ_i^z . The phase transition from FM to paramagnetic order occurs for $g = 1$, for large g the unique ground state is well approximated by $|\varphi\rangle = \prod_i |+\rangle_i^x$ with $|\pm\rangle_i^x$ being the eigenvectors of σ_i^x .

Let g serve as the parameter λ and let us start with the ground state of (1) for small g , say with $|\Psi\rangle = |\psi(+)^z\rangle$, and abruptly change g to a large, positive value. In the new Hamiltonian, the term proportional to g will dominate, while the first term of the interaction will be a small perturbation. The initial state can then be decomposed in the basis of eigenvectors σ_i^x as

$$|\Psi(0)\rangle = \prod_i \frac{1}{\sqrt{2}} \left(|+\rangle_i^x + |-\rangle_i^x \right). \quad (2)$$

The initial state after a quench is therefore the product of two-state combinations with coefficients of equal magnitude (the phase does not affect the result). The subsequent time evolution, still neglecting interactions in the final Hamiltonian, yields

$$|\Psi(t)\rangle = \prod_i \frac{1}{\sqrt{2}} \left(|+\rangle_i^x e^{igt/2} + |-\rangle_i^x e^{-igt/2} \right). \quad (3)$$

By the survival probability (fidelity, return amplitude, or Loschmidt echo), one calls (depending on the context) the squared overlap of initial and time evolved state, $\mathcal{L}(t) \equiv |\langle \Psi(0) | \Psi(t) \rangle|^2$. Further one may define [6] the rate function $r(t)$ via $\mathcal{L}(t) = \exp(-Lr(t))$, where L is the system size (number of degrees of freedom). Such a measure has a good thermodynamic limit. Singularities in $r(t)$ time dependence, often referred to as Loschmidt echo singularities, are the defining features of DQPT.

Let us immediately consider the example above. The squared overlap $\mathcal{L}(t)$ becomes simply $\mathcal{L}(t) = \cos^{2L}(gt/2)$, and the size-independent rate $r(t)$ reveals singularities whenever the cosine function vanishes, i.e., for $t^* = (2k+1)\pi/g$ for an integer k . This example clearly shows that Rabi-type oscillations are the real origin of rate function singularities in this case.

One can complain that the situation described above is too simplified; singularities in the form of finite cusps appear also for smaller changes of g , where the approximations made by us would not work fully. Then, however, one can use the

Jordan–Wigner transformation into a noninteracting fermion system, as in the original DQPT letter [5], and observe similar “two-level” dynamics for a given k as different k decouple.

3. Other examples

Our model, however, helps to explain also other situations. In fact, as reviewed in [6], 2-band topological noninteracting models lead to exactly the same dynamics. Again here, due to the lack of interactions, different k values can be treated independently, leading to a similar estimate of critical times at which singularities appear. Let us stress that while these singularities are essential for the phase-transition language application, they just seem to be due to the vanishing overlaps between the initial and time-evolved wavepacket.

Consider now a situation in which we make an abrupt quench within the same phase, then by definition the ground state changes slowly and continuously with the change of the parameter for a finite system. So it is quite justified to assume that the ground state at say $\lambda = 0$ expands in eigenstates $\{|\psi_k\rangle\}$ of $H(\lambda)$ as

$$|\Psi\rangle = \alpha_0 |\psi_0\rangle + \sum_k \alpha_k |\psi_k\rangle, \quad (4)$$

with $|\alpha_0| \gg |\alpha_k|$ for $k > 0$. Then the survival probability (Loschmidt echo) is dominated by the large term $|\alpha_0|^2$. The situation is more subtle in the thermodynamic limit due to the Anderson catastrophe. Still then, one may expect that many eigenstates at the final parameter value contribute to the initial wavepacket, leading to many superimposed oscillations at different frequencies. In such a situation, the rate function should not reveal strong maxima (not speaking of singularities).

Note that the situation is markedly different when the phase transition is crossed in λ because then, for the Ising system, via symmetry as described above, *two* eigenstates contribute significantly to the sum, leading to Rabi oscillations at half of their energy difference (per site).

The discussion up till now was concentrated on spin-1/2 models leading to simple Rabi oscillations. This might be a transverse Ising chain but also, e.g. a quantum dot dynamics [8]. As known from quantum optics, Rabi oscillations generalize to quantum revivals appearing when several equally spaced levels are populated [9]. Here again one may expect that between consecutive revivals, minima of the survival probability lead to maxima (and possibly cusps) of the Loschmidt rate functions. In a many-body system, an even more general situation was experimentally realized many years ago for interacting bosons in an optical lattice [10]. Initially, the bosons were kept in a shallow lattice, then abruptly the height of the lattice was increased dramatically, separating different lattice sites. Within each site,

the initial almost coherent state was a superposition of states with different site occupations, separated by a quadratic progression in the interaction strength U (within the tight binding Bose–Hubbard description), and the revivals were observed. The corresponding Loschmidt rates reveal singularities (or maxima), as discussed in detail recently [11] in the DQPT language, at times when the overlap between initial and time evolved state is minimal, i.e., roughly in the middle between two consecutive revivals.

4. Quantum scars

Recently, an interesting manifestation of ergodicity breaking as persistent oscillations for certain initial states was discovered experimentally with ultracold Rydberg atoms [7]. This feature is due to the presence of few atypical, almost equally spaced eigenstates — so-called quantum many-body scars (QMBS) [12, 13], which are embedded in the otherwise thermal spectrum of a quantum many-body system. For initial states with a high overlap with a few QMBS, one observes long-lived oscillations of observables, whereas for generic initial conditions the system quickly approaches thermal equilibrium state. The same oscillations should be present in the survival probability leading, in turn, to maxima of the Loschmidt echo rate function if the data are interpreted in that way.

QMBS borrowed their name from the single-particle quantum chaos studies, where “quantum scar” described the enhanced probability of eigenstates or wavepackets in regions of space occupied by unstable periodic orbits [14] — in close relation to the semiclassical periodic orbits quantization [15, 16]. Then also the concept of scarring by symmetries was developed in the context of hydrogen atom in magnetic field studies [17]. Similar symmetry concepts were used for the construction of nonergodic states in many-body case see, e.g. [18, 19].

Such QMBS may be easily imagined as having the origin in the approximate decoupling of a (not always apparent) single degree of freedom from other degrees of freedom. If this single degree is locally described by a harmonic oscillator (or an angular momentum), then the corresponding eigenstates are equidistant — their weak coupling to the remaining states preserves the energy structure. Now, if by accident (or cleverness) the initial state is prepared as a linear combination of those selected states (or if it has sufficiently large overlap on at least a few of them), one may naturally expect a persistent oscillation in the time dynamics. Let us mention also that quench dynamics and Rabi oscillations resulting from the excitations of two or more localized integrals of motion in the context of many-body localization have recently been studied [20, 21]. The localized, almost decoupled family of states may not be easy to identify, one may try to identify

it e.g., by adiabatic following from some analytic limit [22] or via purely numerical approaches including artificial intelligence [23].

5. Conclusions

DQPT forms a very intriguing interpretation of rapid quantum quenches. On the other hand, signatures of DQPT in the form of singularities of the Loschmidt echo rate functions appear to a large extent due to the very definition of this rate. Survival probability (Loschmidt echo) itself reveals no singularities but rather smooth oscillations (or revivals in more complicated cases).

Let us stress that the mechanism presented above considers rather simple examples. For more complicated cases one may consider the Loschmidt echo as coming back not to a single ground state, but to the degenerate manifold, if it exists [6]. After the first draft of this note was completed, a related work appeared, giving a more general picture of DQPT [24]. It has been brought also to our attention that similar to DQPT cusp structures may appear in single particle dynamics [25, 26].

Acknowledgments

This note was born as a reflection on a seminar given in Krakow by Tadeusz Domański. His great presentation skills are acknowledged. I am grateful to Dominique Delande, Mateusz Łącki, Anatoly Polkovnikov, and Tommaso Roscilde for the discussions. The work of J.Z has been realized within the Opus grant 2019/35/B/ST2/00034, financed by National Science Centre (Poland). The research has also been supported by the Strategic Programme Excellence Initiative at Jagiellonian University.

References

- [1] I. Białynicki-Birula, M. Kaliński, J.H. Eberly, *Phys. Rev. Lett.* **73**, 1777 (1994).
- [2] D. Delande, J. Zakrzewski, A. Buchleitner, *Europhys. Lett.* **32**, 107 (1995).
- [3] D. Delande, J. Zakrzewski, *Phys. Rev. A* **58**, 466 (1998).
- [4] A. Buchleitner, D. Delande, J. Zakrzewski, *Phys. Rep.* **368**, 409 (2002).
- [5] M. Heyl, A. Polkovnikov, S. Kehrein, *Phys. Rev. Lett.* **110**, 135704 (2013).
- [6] M. Heyl, *Rep. Prog. Phys.* **81**, 054001 (2018).
- [7] H. Bernien, S. Schwartz, A. Keesling et al., *Nature* **551**, 579 (2017).
- [8] K. Wrześniewski, I. Weymann, N. Sedlmayr, T. Domański, *Phys. Rev. B* **105**, 094514 (2022).

- [9] J.H. Eberly, N.B. Narozhny, J.J. Sanchez-Mondragon, *Phys. Rev. Lett.* **44**, 1323 (1980).
- [10] M. Greiner, O. Mandel, T.W. Hänsch, I. Bloch, *Nature* **419**, 51 (2002).
- [11] M. Lacki, M. Heyl, *Phys. Rev. B* **99**, 121107 (2019).
- [12] C.J. Turner, A.A. Michailidis, D.A. Abanin, M. Serbyn, Z. Papić, *Nat. Phys.* **14**, 745 (2018).
- [13] C.J. Turner, A.A. Michailidis, D.A. Abanin, M. Serbyn, Z. Papić, *Phys. Rev. B* **98**, 155134 (2018).
- [14] E.J. Heller, *Phys. Rev. Lett.* **53**, 1515 (1984).
- [15] M.C. Gutzwiller, *J. Math. Phys.* **12**, 343 (1971).
- [16] E. Bogomolny, *Physica D* **31**, 169 (1988).
- [17] D. Delande, J.C. Gay, *Phys. Rev. Lett.* **59**, 1809 (1987).
- [18] N. Shiraishi, T. Mori, *Phys. Rev. Lett.* **119**, 030601 (2017).
- [19] S. Moudgalya, S. Rachel, B.A. Bernevig, N. Regnault, *Phys. Rev. B* **98**, 235155 (2018).
- [20] L. Benini, P. Naldesi, R.A. Römer, T. Roscilde, *New J. Phys.* **23**, 023030 (2021).
- [21] L. Benini, P. Naldesi, R.A. Römer, T. Roscilde, *Ann. Phys.* **435**, 168545 (2021).
- [22] A.S. Aramthottil, U. Bhattacharya, D. González-Cuadra, M. Lewenstein, L. Barbiero, J. Zakrzewski, *Phys. Rev. B* **106**, L041101 (2022).
- [23] T. Szoldra, P. Sierant, M. Lewenstein, J. Zakrzewski, *Phys. Rev. B* **105**, 224205 (2022).
- [24] M. Van Damme, J.-Y. Desaulles, Z. Papić, J.C. Halimeh, *arXiv:2210.02453*, 2022.
- [25] J.M. Zhang, H.-T. Yang, *Europhys. Lett.* **114**, 60001 (2016).
- [26] J.M. Zhang, Y. Liu, *Eur. J. Phys.* **37**, 065406 (2016).

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— ◇ —

Vlasov–Klimontovich Equation in Action

Ł.A. TURSKI*

Center for Theoretical Physics, Polish Academy of Science, al. Lotników 32/46, 02-668 Warsaw, Poland

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S183](https://doi.org/10.12693/APhysPolA.143.S183)*e-mail: laturski@cft.edu.pl

This paper outlines a variety of possible applications of the Vlasov equation and its generalization, i.e., the Klimontovich equation, in various areas of many-body physics. In particular, these equations are shown to be used in relativistic plasma physics, the theory of semi-classical Bloch electrons, and the metriplectic description of dissipative processes.

topics: kinetic theory of relativistic plasma, semi-classical theory of Bloch electrons, metriplectic description of dissipative processes

1. Introduction

In 1938, Anatoly Vlasov wrote a seminal paper [1] in which he argued that in the description of a many-body system with long-range interparticle interactions, the conventional kinetic Boltzmann equation is inadequate and should be replaced by the continuity equation for the one-particle distribution function $f(\mathbf{r}, \mathbf{p}, t)$ in the one-particle phase space, where \mathbf{r} denotes the particle position and \mathbf{p} its momentum. The adequate equation in question is

$$\frac{\partial f(\mathbf{r}, \mathbf{p}, t)}{\partial t} + \frac{\mathbf{p}}{m} \cdot \nabla f(\mathbf{r}, \mathbf{p}, t) + \mathbf{F} \cdot \frac{\partial f(\mathbf{r}, \mathbf{p}, t)}{\partial \mathbf{p}} = 0. \quad (1)$$

Here, \mathbf{F} is the total force acting on the particle, resulting from interactions with all particles in the system. Therefore, the force \mathbf{F} depends on $f(\mathbf{r}, \mathbf{p}, t)$. When the interparticle interactions are given by the potential forces, then $\mathbf{F}(\mathbf{r}, t) = -\int d\mathbf{p}' d\mathbf{r}' \nabla V(\mathbf{r}, \mathbf{r}') f(\mathbf{r}', \mathbf{p}', t)$, and thus (1) becomes a self-consistent, nonlinear equation of one particle distribution function.

Vlasov pointed out that the collision term — the conventional right-hand side of the Boltzmann equation — is divergent for long-range Coulomb interactions between charged particles. A replacement of the collision term for a charged particle was suggested by Landau, and its formal derivation was proposed by Rosenbluth [2] in the first glory years of thermonuclear reaction physics. In fact, (1) with the Landau collision term becomes a

formidable nonlinear equation, which plays a fundamental role in plasma physics [3].

Forty years later, the posthumously published book by Vlasov [4] contained several, mostly failed, attempts to generalize his original ideas, including those for relativistic statistical mechanics. The Vlasov equation has been also used in many condensed matter applications far from classical plasma physics, for example, to describe quark–gluon plasma and some problems in heavy ion collisions [5] and in the development of late stages of phase separation in the first-order phase transformations [6].

Almost twenty years later, Braun and Hepp [7] showed that the Vlasov equation describes an asymptotically exact, equal time evolution for the N -particle Born, Bogoliubov, Green, Kirkwood, and Yvon (BBGKY) hierarchy with interactions of the form $\frac{1}{N} \sum V(\mathbf{r}, \mathbf{r}')$. The eigenfunctions of the linearized (1) appear as approximate eigenfunctions of the classical Liouville equation in the Zwanzig variational principle for fluids [8].

Independently from Vlasov, and years later, Yu.L. Klimontovich [9, 10] observed that for the N -particle system the “phase space operator”

$$\hat{f}(\mathbf{r}, \mathbf{p}, t) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \delta(\mathbf{p} - \mathbf{p}_i) \quad (2)$$

obeys the equation

$$\frac{\partial \hat{f}(\mathbf{r}, \mathbf{p}, t)}{\partial t} + \frac{\mathbf{p}}{m} \cdot \nabla \hat{f}(\mathbf{r}, \mathbf{p}, t) + \mathbf{F}\{\hat{f}\} \cdot \frac{\partial \hat{f}(\mathbf{r}, \mathbf{p}, t)}{\partial \mathbf{p}} = 0, \quad (3)$$

where

$$\mathbf{F}(\hat{f}) = \int d\mathbf{r}' d\mathbf{p}' \mathbf{F}(\mathbf{r}, \mathbf{r}') \hat{f}(\mathbf{r}' \mathbf{p}'), \quad (4)$$

here $\mathbf{F}(\mathbf{r}, \mathbf{r}')$ is the force acting between the particles. The fact that the function $\hat{f}(\mathbf{r}, \mathbf{p})$ is one of the exact solutions of (1) plays a fundamental role in all applications of the Vlasov equation and particularly in mathematically correct solutions of it [3].

For an N -particle system with the Hamiltonian $H(\mathbf{r}, \mathbf{p})$, the Hamilton equations of motion, identical to the Newton ones, can be written as

$$\begin{aligned} \dot{\mathbf{r}} &= \{\mathbf{r}, H\}, \\ \dot{\mathbf{p}} &= \{\mathbf{p}, H\}, \end{aligned} \quad (5)$$

where $\{a, b\}$ denote the Poisson brackets between the arbitrary phase space functions $a = a(\mathbf{r}, \mathbf{p})$ and $b = b(\mathbf{r}, \mathbf{p})$

$$\{a, b\} = \frac{\partial a}{\partial \mathbf{r}} \cdot \frac{\partial b}{\partial \mathbf{p}} - \frac{\partial b}{\partial \mathbf{r}} \cdot \frac{\partial a}{\partial \mathbf{p}}. \quad (6)$$

Using (5)–(6), we can derive the Poisson bracket relation between the Klimontovich distributions (2),

$$\{\hat{f}(\mathbf{r}, \mathbf{p}, t), \hat{f}(\mathbf{r}', \mathbf{p}', t)\} = [\hat{f}(\mathbf{r}, \mathbf{p}', t) - \hat{f}(\mathbf{r}', \mathbf{p}, t)] \nabla \cdot \nabla_P \delta(\mathbf{r} - \mathbf{r}') \delta(\mathbf{p} - \mathbf{p}') = \int d\mathbf{r}'' d\mathbf{p}'' C_{\mathbf{r}\mathbf{p}, \mathbf{r}'\mathbf{p}'}^{\mathbf{r}''\mathbf{p}''} \hat{f}(\mathbf{r}'', \mathbf{p}'', t), \quad (7)$$

where $\nabla \equiv \partial/\partial \mathbf{r}$ and $\nabla_P \equiv \partial/\partial \mathbf{p}$. Now, (7) shows that the algebra of the Klimontovich distribution functions forms the Lie algebra with structure coefficients $C_{\mathbf{r}\mathbf{p}, \mathbf{r}'\mathbf{p}'}^{\mathbf{r}''\mathbf{p}''}$. This algebra is, therefore, of fundamental interest in the metriplectic formulation of dissipative systems dynamics [11].

Assuming the conventional form of the Hamiltonian

$$H(\mathbf{r}, \mathbf{p}) = \sum_{i=1\dots N} \frac{\mathbf{p}^2}{2m} + \frac{1}{2} \sum_{i<j} V(|\mathbf{r}_i - \mathbf{r}_j|), \quad (8)$$

we can write the Hamiltonian H as functional of \hat{f} ,

$$\begin{aligned} H\{\hat{f}\} &= \int d\mathbf{r} d\mathbf{p} \frac{\mathbf{p}^2}{2m} \hat{f}(\mathbf{r}, \mathbf{p}) \\ &+ \frac{1}{2} \int d\mathbf{r} d\mathbf{p} \int d\mathbf{r}' d\mathbf{p}' \hat{f}(\mathbf{r}, \mathbf{p}) V(|\mathbf{r} - \mathbf{r}'|) \hat{f}(\mathbf{r}', \mathbf{p}'), \end{aligned} \quad (9)$$

and subsequently (3) can be written as

$$\frac{\partial \hat{f}(\mathbf{r}, \mathbf{p})}{\partial t} = \{\hat{f}(\mathbf{r}, \mathbf{p}), H\{\hat{f}\}\}. \quad (10)$$

Note that (2), (7), and (10) form the symplectic formulation of the many-body dynamics equivalent to the Hamiltonian formulation. The complete introduction to symplectic dynamics can be found in a classical book by Marsden and collaborators [12] and in a series of works by Morrison [13]. It is worth noticing that by replacing the Klimontovich distribution with the Wigner function [14] and its Poisson brackets through the Moyal brackets [15], we obtain the quantum version of the Vlasov–Klimontovich formulation of the many-body system.

Back in the late 1970s, Piotr Goldstein and I were working extensively on the use of the Vlasov–Klimontovich formulation to describe the properties of waves propagating in a quasi-relativistic plasma, described by an approximation, in which interactions between charged particles are described by means of the Breit–Darwin Hamiltonian containing velocity-dependent interactions [16]. Simultaneously, with Zbigniew Iwiński, a former student of Iwo Białynicki-Birula, we were analyzing the possibility of formulating a fully relativistic

form of the Vlasov equation. In a preliminary paper, we formulated such a description and derived the Poisson brackets for a relativistic generalization of the Klimontovich function [17]. Years later, Iwo Białynicki-Birula and John C. Hubbard were working on the same subject, and eventually, we published together a complete description of the gauge-independent and canonical formulation of the relativistic plasma theory [18].

2. Relativistic plasma theory

The publication [18] mentioned at the end of the previous section contained the gauge-independent formulation of the theory of relativistic plasma constituting the multicomponent particle system and the electromagnetic field. Our theoretical tool for that purpose is the symplectic (or canonical) formulation with the dynamical variables for electromagnetic field \mathbf{E}, \mathbf{B} and particle variables, namely, positions $\boldsymbol{\xi}_A$ and relativistic kinematic momenta \mathbf{P}_A . The index A labels particles belonging to a particular particle species $A \in S_a, a = 1, \dots, \mathcal{S}$, and $\mathbf{P}_A = m_a \mathbf{v}_A / \sqrt{1 - \mathbf{v}_A^2/c^2}$. These variables obey the Maxwell–Lorentz equations of the form

$$\begin{aligned} \frac{d\boldsymbol{\xi}_A}{dt} &= \mathbf{v}_A, \\ \frac{d\mathbf{P}_A}{dt} &= e_a [\mathbf{E}(\boldsymbol{\xi}_A(t), t) + \mathbf{v}_A(t) \times \mathbf{B}(\boldsymbol{\xi}_A(t), t)], \\ \frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} &= -\nabla \times \mathbf{E}(\mathbf{r}, t), \\ \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} &= \nabla \times \mathbf{B}(\mathbf{r}, t) - \sum_A e_A \mathbf{v}_A(t) \delta(\mathbf{r} - \boldsymbol{\xi}_A(t)), \\ \nabla \cdot \mathbf{B}(\mathbf{r}, t) &= 0, \\ \nabla \cdot \mathbf{E}(\mathbf{r}, t) &= \sum_A e_A \delta(\mathbf{r} - \boldsymbol{\xi}_A(t)). \end{aligned} \quad (11)$$

The relativistic invariant phase space Klimontovich function is identical to that in (2), with the relativistic kinematic momenta replacing \mathbf{p} , i.e.,

$$\hat{f}_a(\mathbf{r}, \mathbf{p}, t) = \sum_{A \in \mathcal{S}_a} \delta(\mathbf{r} - \boldsymbol{\xi}_A) \delta(\mathbf{p} - \mathbf{P}_A), \quad (12)$$

where

$$\mathbf{p} = m_a \mathbf{v} / \sqrt{1 - \mathbf{v}^2/c^2}. \quad (13)$$

The Maxwell–Lorentz equation (11) can be cast into a canonical form using the Poisson bracket relations

$$\begin{aligned} \{\xi_A^i, P_B^j\} &= \delta_{AB} \delta^{ij}, \\ \{P_A^i, P_B^j\} &= e_A \delta_{AB} \epsilon^{ijk} B^k(\boldsymbol{\xi}_A), \\ \{P_A^i, E^j(\mathbf{r})\} &= e_A \delta^{ij} \delta(\mathbf{r} - \boldsymbol{\xi}_A), \\ \{B^i(\mathbf{r}), E^j(\mathbf{r}')\} &= \epsilon^{ijk} \partial_k \delta(\mathbf{r} - \mathbf{r}'), \end{aligned} \quad (14)$$

with all other Poisson brackets vanishing. The Poisson brackets for electromagnetic fields $\{E^i, B^j\}$ are the classical form of commutators derived by

Born and Infeld [19] and discussed in greater detail in [20]. These Poisson brackets are consistent with the constraints described by the last two equations in (11).

With the above choice of canonical variables and their Poisson brackets, the full Poincaré group is realised as a subgroup of the canonical transformation group [18] and the theory of plasma becomes fully relativistic. Using (14) one can easily derive the Poisson brackets for electromagnetic fields and the phase space function (12) which we write below employing shorthand notation $\mathbf{z} = (\mathbf{r}, \mathbf{P})$, $\zeta_A = (\boldsymbol{\xi}_A, \mathbf{P}_A)$, rational system of units with $c = 1$ and following

$$\left\{ F(\{\zeta_A\}), G(\{\zeta_B\}) \right\} = \sum_{\zeta_A, \zeta_B} \frac{\partial F}{\partial \zeta_A} \left\{ \zeta_a, \zeta_B \right\} \frac{\partial G}{\partial \zeta_B}, \quad (15)$$

$$\{\hat{f}_a(\mathbf{z},) \hat{f}_b(\mathbf{z}')\} = \delta_{ab} \left[\left(\hat{f}_a(\mathbf{r}, \mathbf{P}') - \hat{f}_a(\mathbf{r}', \mathbf{P}) \right) \nabla \cdot \nabla_P + e_a \mathbf{B}(\mathbf{r}) \cdot (\nabla_P \hat{f}_b(\mathbf{z}) \times \nabla_P) \right] \delta(\mathbf{z} - \mathbf{z}'). \quad (16)$$

The remaining non-zero Poisson brackets read

$$\{\hat{f}_a(\mathbf{z}), \mathbf{E}(\mathbf{r}')\} = -e_a \nabla_P \hat{f}_a(\mathbf{z}) \delta(\mathbf{z} - \mathbf{z}'), \quad \{\hat{f}_a(\mathbf{z}), \mathbf{B}(\mathbf{r}')\} = 0. \quad (17)$$

Having the above formalism, we can express all the generators of the Poincaré group in terms of the Klimontovich function \hat{f} and fields \mathbf{E} , \mathbf{B} [18]. For example, the Hamiltonian of the system reads

$$H = \sum_a \int d\mathbf{z} \sqrt{\mathbf{p}^2 + m^2} \hat{f}_a(\mathbf{z}) + \frac{1}{2} \int d\mathbf{r} (\mathbf{E}^2 + \mathbf{B}^2), \quad (18)$$

and momentum vector

$$\boldsymbol{\Pi} = \sum_a \int d\mathbf{z} \mathbf{p} \hat{f}_a(\mathbf{z}) + \int d\mathbf{r} \mathbf{E} \times \mathbf{B}. \quad (19)$$

Note that the Hamiltonian (18) does not contain the coupling constant between the plasma and the electromagnetic field, i.e., the charge e_a . The interaction between these two is fully contained in the Poisson brackets (14), (16), (17). The Klimontovich–Vlasov formulation of relativistic plasma physics, presented in [18], therefore follows some ideas presented by Souriau and Sternberg [21, 22]. This gauge-invariant formulation of the interacting system of particles and fields can be extended for general relativity formulation [23].

The Vlasov equation and the Maxwell equation can then be written as

$$\begin{aligned} \frac{\partial \hat{f}_a}{\partial t} &= \left\{ \hat{f}_a, H \left\{ \hat{f}_a, \mathbf{E}, \mathbf{B} \right\} \right\} \equiv \\ &\quad - \left[\mathbf{v}_a \cdot \nabla + e_a (\mathbf{E} + \mathbf{v}_a \times \mathbf{B}) \cdot \nabla_P \right] \hat{f}_a, \\ \frac{\partial \mathbf{E}}{\partial t} &= \{ \mathbf{E}, H \} \equiv \nabla \times \mathbf{B} - \sum_a \int d\mathbf{p} \mathbf{v}_a \hat{f}_a, \\ \frac{\partial \mathbf{B}}{\partial t} &= \{ \mathbf{B}, H \} \equiv -\nabla \times \mathbf{E}. \end{aligned} \quad (20)$$

The relativistic statistical mechanics does not offer a mathematically rigorous formulation of the relation between the Vlasov (1) and the Klimontovich equation (3) like that in [7]. Nevertheless, there is sufficient experimental experience from hot plasma and astrophysical applications for one to make an assumption that the one-particle distribution function defined in one-particle phase space of positions and relativistic kinematical momenta — the ensemble average of the Klimontovich function — obeys identical equations as our relativistic one (20). With this assumption, we can generalize the formulation given above by including in our description direct information on destroying physical processes in plasma — direct charge particle collisions — similarly as Landau has done for the original Vlasov equation. To do this, it is convenient to follow an algebraic method of including dissipative processes in symplectic dynamics — the metriplectic method. We shall discuss this procedure in Sect. 4.

3. Semiclassical spin 1/2 Bloch electrons plasma

In the quantum theory of crystalline solids, the motion of electrons is described by means of the wave packets constructed from Bloch wave functions with periodic part $u_{\mathbf{k}}$, where \mathbf{k} labels the wave vectors for the specific band [24]. For the sake of simplicity, we consider here the solids with only one energy band $\epsilon(\mathbf{k})$. J. Zak [25] observed that the Bloch systems yield the geometric phases and that the gauge-invariant Berry curvature [26]

$$\tilde{\Omega}(\mathbf{k}) = i\langle \nabla_{\mathbf{k}} u_{\mathbf{k}} | \times | \nabla_{\mathbf{k}} u_{\mathbf{k}} \rangle \quad (21)$$

is observable and generally nonzero for crystals without inversion symmetry. Theoretical analysis [27, 28] has shown that in many important experimental applications, it is sufficient to describe the motion of electrons by a semiclassical equation of motion in which the position of the center of the localized electron wave function and wave vector (\mathbf{r}, \mathbf{k}) obey the equation of motion

$$\dot{\mathbf{r}} = \frac{\partial \epsilon(\mathbf{k})}{\hbar \partial \mathbf{k}} + \mathbf{F} \times \tilde{\Omega}/\hbar, \quad \dot{\mathbf{k}} = \mathbf{F}/\hbar, \quad (22)$$

where $\mathbf{F} = -\nabla U(\mathbf{r})$ is the net force acting on the electrons.

These equations play the role of the Hamilton equations in classical mechanics and, therefore, can be used in the formulation of the symplectic, or Lie–Poisson bracket technique [11, 12, 29], description of the semiclassical Bloch electrons. This description can subsequently be rewritten using the Vlasov–Klimontovich equation approach [30]. We begin by defining the Poisson brackets for “position” \mathbf{r} and “momenta” $\boldsymbol{\kappa} = \hbar \mathbf{k}$ (where $\Omega = \tilde{\Omega}/\hbar$) as follows

$$\begin{aligned} \{r_a, r_b\} &= \epsilon_{abc} \Omega_c, & \{r_a, \kappa_b\} &= \delta_{ab}, \\ \{\kappa_a, \kappa_b\} &= 0, \end{aligned} \quad (23)$$

a special case of non-commutative classical mechanics Poisson brackets discussed in [31]. Assuming that the Hamiltonian for semiclassical electrons can be written as $H(\mathbf{r}, \boldsymbol{\kappa}) = \epsilon(\boldsymbol{\kappa}) + U(\mathbf{r})$ equations

$$\dot{r}_a = \{r_a, H\}, \quad \dot{\kappa}_a = \{\kappa_a, H\}, \quad (24)$$

become identical to (22). That allows us to use the Klimontovich function [9, 10] with momenta \mathbf{p} replaced by $\boldsymbol{\kappa}$ to describe the semiclassical Bloch electron plasma. We can write, as in previous sections, $H\{\hat{f}\} = \int d\mathbf{1} \epsilon(\mathbf{1}) \hat{f}(\mathbf{1})$, where $\mathbf{1} = (\mathbf{r}, \boldsymbol{\kappa})$ and the mean value of a physical observable as $\langle A \rangle = \int d\mathbf{1} A(\mathbf{1}) \hat{f}(\mathbf{1})$. Using the Poisson brackets (23), one easily finds the Poisson brackets for \hat{f} function

$$\begin{aligned} \{\hat{f}(\mathbf{1}), \hat{f}(\mathbf{2})\} &= \left(\nabla \hat{f}(\mathbf{1}) \cdot \nabla_{\boldsymbol{\kappa}} - \nabla_{\boldsymbol{\kappa}} \hat{f}(\mathbf{1}) \cdot \nabla \right) \delta(\mathbf{1}-\mathbf{2}) \\ &+ \Omega \cdot \left[\nabla \hat{f}(\mathbf{1}) \times \nabla \delta(\mathbf{1}-\mathbf{2}) \right]. \end{aligned} \quad (25)$$

Now, (25) allows us to write the Vlasov–Klimontovich equation for the semiclassical electrons as

$$\frac{\partial \hat{f}(\mathbf{1})}{\partial t} = \{\hat{f}(\mathbf{1}), H\{\hat{f}\}\}. \quad (26)$$

To account for the specific properties of the semiclassical electron plasma, for example, Ohm’s law, we need to supplement the Vlasov–Klimontovich equation (26) with the proper dissipative term on its RHS, which we shall denote $W\{\hat{f}\}$. In [27, 28], the simple relaxation time approximation has been used for $W\{\hat{f}\}$. The full kinetic equation for \hat{f} does not then conserve the number of charge carriers in the system. It seems, therefore, more appropriate to replace this $W\{\hat{f}\}$ by the generalization of

the Boltzmann–Lorentz collision operator [32, 33], which offers a formulation of the collision operator for a tight-binding model. The relation between the construction of such an operator and the symplectic formulation of many-particle system dynamics will be outlined in the last section. Having done so and using the linearized version of the Chapman–Enskog approximation $\hat{f}(\mathbf{r}, \boldsymbol{\kappa}) \approx \rho(\mathbf{r}) \phi_B(\boldsymbol{\kappa})$, where $\rho(\mathbf{r})$ denotes carriers density and $\phi_B(\boldsymbol{\kappa})$ stands for equilibrium carriers distribution function at the temperature of the carrier β^{-1} defined as $m \delta_{ij} \beta^{-1} = \int d\boldsymbol{\kappa} \phi_B(\boldsymbol{\kappa}) \kappa_i \kappa_j$, and, furthermore, assuming that the Berry curvature Ω is slowly varying function of the wave vector \mathbf{k} traversing the Brillouin zone, we obtain the dispersion relation for the fluctuation of the density of carriers $\rho_{\omega, \mathbf{q}}$, which replaces the Ohm law for spinless Bloch electrons

$$\begin{aligned} \left(\tilde{\omega}(\mathbf{q}) + iz\Gamma_{\mathbf{q}} \right) \left(\tilde{\omega}(\mathbf{q}) - iS_{\mathbf{q}}\Gamma_{\mathbf{q}} \right) = \\ \left[\frac{q^2}{m\beta} - e\mathbf{q} \cdot \left(\mathbf{E}_0 \times (\mathbf{q} \cdot \bar{\Xi}) \right) + ie\mathbf{E}_0 \cdot \frac{\mathbf{q}}{m} \right], \end{aligned} \quad (27)$$

where

$$\tilde{\omega}(\mathbf{q}) = \omega - e\mathbf{q} \cdot (\bar{\Omega} \times \mathbf{E}_0) \quad (28)$$

is the frequency shift due to the anomalous Hall drift velocity $e\bar{\Omega} \times \mathbf{E}_0$. There, $\Gamma_{\mathbf{q}}$, $S_{\mathbf{q}}$, and z are the scattering amplitude, scatterers structure factor, and coordination number, respectively; $\bar{\Omega}$ and $\bar{\Xi}_{ij}$ denote, respectively, the averaged values of the Berry curvature $\bar{\Omega} = \int d\boldsymbol{\kappa} \phi_B(\boldsymbol{\kappa}) \Omega(\boldsymbol{\kappa})$ and the averaged “curvature torque” $\bar{\Xi}_{ij} = \int d\boldsymbol{\kappa} \kappa_i \Omega_j(\boldsymbol{\kappa}) \phi_B(\boldsymbol{\kappa})$.

The charge carriers in solid carry internal degrees of freedom spins. One can generalize the Vlasov–Klimontovich description of semiclassical carriers, shown above, for the case of Bloch electrons with spin $\frac{1}{2}$. We do this by describing the carriers by spinor Klimontovich distribution function $\hat{f} = \frac{1}{2} \sum_{\alpha=0}^3 f_{\alpha} \hat{\sigma}_{\alpha}$, where $\hat{\sigma}_{i=1,2,3}$ are the Pauli matrices and $\hat{\sigma}_0$ is the 2×2 unit matrix. The meaning of coefficients f_{α} stems from the meaning of the mean value of the observables $\langle A \rangle = \text{Tr}(A \hat{f})$. Here, Tr denotes the matrix trace in spinor space and phase-space integration. The coefficient f_0 is the Vlasov–Klimontovich function used for spinless carriers, and f_i are carriers spin densities $\langle S_i \rangle = \frac{1}{2} \int d\mathbf{1} \text{Tr}(\frac{\hbar}{2} \hat{\sigma}_i \sum_{\alpha} f_{\alpha} \hat{\sigma}_{\alpha}) = \frac{\hbar}{2} \int d\mathbf{1} f_i(\mathbf{1})$.

For the spinor distribution function \hat{f} , the Poisson brackets now become 4×4 functional matrices ($i, j = 1, 2, 3$)

$$\left[\hat{f}(\mathbf{1}), \hat{f}(\mathbf{2}) \right] = \begin{pmatrix} \{f_0(\mathbf{1}), f_0(\mathbf{2})\}, & \{f_0(\mathbf{1}), f_j(\mathbf{2})\} \\ \{f_i(\mathbf{1}), f_0(\mathbf{2})\}, & \{f_i(\mathbf{1}), f_j(\mathbf{2})\} \end{pmatrix}, \quad (29)$$

where $\{f_0(\mathbf{1}), f_0(\mathbf{2})\}$ is given by (25) and

$$\begin{aligned} \{f_i(\mathbf{1}), f_0(\mathbf{2})\} &= -\nabla f_i(\mathbf{1}) \cdot \nabla_{\boldsymbol{\kappa}} \delta(\mathbf{1}-\mathbf{2}) \\ \{f_i(\mathbf{1}), f_j(\mathbf{2})\} &= \epsilon_{ijk} f_k(\mathbf{1}) \delta(\mathbf{1}-\mathbf{2}). \end{aligned} \quad (30)$$

The kinetic equation for the spinor \hat{f} becomes now

$$\begin{aligned} \frac{\partial \hat{f}}{\partial t} &= [\hat{f}, H\{\hat{f}\}] + \mathcal{W}\{\hat{f}\} = \\ \text{Tr}_{\mathbf{2}}[\hat{f}(\mathbf{1}), \hat{f}(\mathbf{2})] &\frac{\delta H\{\hat{f}\}}{\delta \hat{f}(\mathbf{2})} + \mathcal{W}\{\hat{f}\}. \end{aligned} \quad (31)$$

The collision operator \mathcal{W} must preserve the length of the carrier spin \mathbf{S}^2 . The Poisson brackets for the spinor \hat{f} (see (29)) guarantee that the spin length is the Casimir of the Lie–Poisson brackets.

Using the notation from [34, 35], we write

$$\begin{aligned} \mathcal{W}\{\hat{f}\} &= \\ -\lambda \int d\mathbf{2} &\left[\delta_{ij} \phi_B(\kappa_2) \delta(\mathbf{1}-\mathbf{2}) - \hat{f}_i(\mathbf{1}) \hat{f}_j(\mathbf{2}) \right] \frac{\delta H\{\hat{f}\}}{\delta \hat{f}(\mathbf{2})}. \end{aligned} \quad (32)$$

Having formulated the kinetic equation for semiclassical spin 1/2 Bloch electrons, we can use them to analyze the properties of such a carrier plasma. For example, we can consider plasma with constant equilibrium carriers density ρ_0 , constant external electric field \mathbf{E}_0 , and diagonal pressure tensor $P_{ij}(\rho) = \int d^3\kappa f(\boldsymbol{\kappa}) \kappa_i \kappa_j$. Linearizing (31), we obtain the dispersion relation for plasma excitation

$$\tilde{\omega}(\mathbf{q})^2 = \omega_P^2 + C_{ij} q^i q^j + \frac{ie}{m} \mathbf{q} \cdot \mathbf{E}_0, \quad (33)$$

where (ω, \mathbf{q}) denotes the plasma excitations frequency and momentum; $\omega_P = \sqrt{4\pi e^2 \rho_0 / m}$ is the plasma frequency; $\tilde{\omega}(\mathbf{q})$ is given in (28); $c = (\frac{\partial P}{\partial \rho})_0^{1/2}$ is the speed of sound in carrier gas at equilibrium; and sound velocity is anisotropic $c \rightarrow C_{ij} q^i q^j$, with $C_{ij} = c^2 (\delta_{ij} - \epsilon_{ijk} E_0^l \bar{\Xi}^{kl} / c^2)$. This anisotropy is caused by the coupling of the external electric field and the Berry phase curvature torque. The plasma excitations group velocity is shifted with respect to that of usual plasma in the frame of reference drifting with the anomalous Hall velocity $\mathbf{v}_D = e \bar{\boldsymbol{\Omega}} \times \mathbf{E}_0$.

In the following sections, we shall discuss the metriplectic generalization of the Vlasov–Klimontovich formulation of both relativistic plasma and plasma of semiclassical Bloch electrons.

4. Metriplectic description

In two previous sections, we have described the use of the Lie–Poisson brackets technique, together with the use of the Vlasov–Klimontovich function, for two important physical models of plasmas: the relativistic plasma and the plasma of semiclassical Bloch, spinless and spin 1/2, electrons. The Lie–Poisson bracket technique has been applied to many other examples in non-linear physics leading to important progress in those fields [12, 36]. All these applications are examples of reversible (dissipationless) dynamical systems. Allan Kaufman and Phil Morrison suggested [37–39] that this description can

be generalized to include dissipative processes by employing a technique called now metriplectic dynamics.

This theory, described for example in [11, 13], consists of two steps.

Step one is replacing the Hamiltonian in equations of motions by the system free energy $\mathcal{F}(\psi) = \mathcal{H}(\psi) - \theta \mathcal{S}(\mathcal{C})$, where ψ stands for the system dynamical variables, \mathcal{S} is the entropy functional and \mathcal{C} denotes a set of the Casimir variables defined as quantities which Poisson brackets with all ψ vanishes identically, independently of the form of the Hamiltonian [29]. The coefficient θ depends on the type of interaction between the dynamical system and the environment. For example, it can be identified with the system temperature by assuming that the absolute minimum of $\mathcal{F}(\psi)$ is described by the equilibrium distribution function for the system given by the Hamiltonian H . In both examples discussed in previous sections, this distribution function is the proper Maxwell–Boltzmann distribution.

Step two consists in adding to the Poisson brackets in the Hamilton equations of motion (5) the symmetric–“dissipative” brackets

$$\langle \psi(\zeta_A), \psi(\zeta_B) \rangle = \langle \psi(\zeta_B), \psi(\zeta_A) \rangle = D(\psi_A, \psi_B). \quad (34)$$

The dynamical equations of motion are now written as

$$\frac{\partial \zeta_A}{\partial t} = \left[\psi_A(\zeta), \mathcal{F}(\zeta) \right] = \int D\zeta' \mathcal{L}_{AB}(\zeta, \zeta') \frac{\delta \mathcal{F}}{\delta \psi_B(\zeta')}, \quad (35)$$

where

$$\mathcal{L}_{AB}(\zeta, \zeta') = \{ \psi_A, \psi_B \} - \langle \psi(\zeta_A), \psi(\zeta_B) \rangle. \quad (36)$$

The structure of the dissipative bracket depends on the nature of the processes causing the dissipation.

In [29], the general theory of algebraic construction of $D(\psi_A, \psi_B)$ was given. For the dynamical theory equipped with Lie–Poisson brackets associated with Lie algebra with the structure constant C_{BC}^A (compare with (7)), the dissipative brackets, consistent with preservation of the Casimirs of that Lie algebra, have the form

$$D_{AB} = G^{CD} C_{CA}^M C_{DB}^N \psi_M \psi_N, \quad (37)$$

where G^{CD} is the inverse of the Cartan–Killing tensor built from the structure constant as $G_{AB} = -C_{AN}^D C_{DB}^N$.

In both applications described in previous sections, these brackets describe the direct particle–particle collisions. Thus recalling that the role of fields ψ is played by the Klimontovich function, we can rewrite (34) as

$$\begin{aligned} \langle \hat{f}(\mathbf{z}), \hat{f}(\mathbf{z}') \rangle &\equiv D(\hat{f}(\mathbf{z}), \hat{f}(\mathbf{z}')) = \\ \frac{1}{2} \int &d^6 z_1 d^6 z_2 \hat{f}(\mathbf{z}_1) \Delta(\mathbf{z}, \mathbf{z}'; \mathbf{z}_1, \mathbf{z}_2) \hat{f}(\mathbf{z}_2), \end{aligned} \quad (38)$$

where the kernel $\Delta(\mathbf{z}, \mathbf{z}'; \mathbf{z}_1, \mathbf{z}_2)$ accounts for physics of those particle–particle interactions.

In the analysis of the semiclassical Bloch electrons, we have used the collision operator \mathcal{W} . The use of a linearized version of such operator results in the dispersion relation in (27). A similar collision operator is shown in (32).

$$\begin{aligned} \Delta(\mathbf{z}, \mathbf{z}'; \mathbf{z}_1 \mathbf{z}_2) &= \int d\mathbf{k} \alpha_{\mathbf{k}}(\mathbf{z}_1, \mathbf{z}_2) \delta(\mathbf{k} \cdot (\mathbf{v}(\mathbf{p}) - \mathbf{v}(\mathbf{p}'))) \left(\mathbf{k} \cdot \left[\nabla_{P_1} \delta(\mathbf{z} - \mathbf{z}_1) - \nabla_{P_2} \delta(\mathbf{z} - \mathbf{z}_2) \right] \right) \\ &\times \left(\mathbf{k} \cdot \left[\nabla_{P_1} \delta(\mathbf{z}' - \mathbf{z}_1) - \nabla_{P_2} \delta(\mathbf{z}' - \mathbf{z}_2) \right] \right), \end{aligned} \quad (39)$$

where $\alpha_{\mathbf{k}}(\mathbf{z}_1, \mathbf{z}_2)$ describe details of particle-particle collision and as before, $\mathbf{v}(\mathbf{p}) = \mathbf{p}/\sqrt{\mathbf{p}^2 + m^2}$, [40].

The metriplectic description based on the Klimontovich function suffers from mathematical difficulties related to the problems with the operations on the singular distribution functions. In most of the applications, the distribution \hat{f} is therefore replaced by the “smooth” one-particle distribution function $f(\mathbf{r}, \mathbf{p}, t) = \langle \hat{f}(\mathbf{r}, \mathbf{p}, t) \rangle$, where $\langle \dots \rangle$ denotes initial ensemble averaging. That surely leads to the loss of information. One can attempt to restore at least part of that lost information by amending the RHS of (35) with a properly chosen Langevin “force” [40–42]

$$\frac{\partial \zeta_A}{\partial t} = \int D\zeta' \mathcal{L}_{AB}(\zeta, \zeta') \frac{\delta \mathcal{F}}{\delta \psi_B(\zeta')} + \lambda_A(\zeta) \quad (40)$$

with

$$\langle \langle \lambda_A(\zeta, t) \lambda_B(\zeta', t) \rangle \rangle = S_{AB}(\zeta, \zeta') \delta(t - t'), \quad (41)$$

where the double brackets $\langle \langle \dots \rangle \rangle$ denote averaging over the realizations of the Langevin forces λ_A . The generalized Fokker–Planck equation for the probability distribution \mathcal{P} in space of dynamical variables ψ can now be written as [40]

$$\frac{\partial \mathcal{P}}{\partial t} = \hat{L} \left(\frac{\delta}{\delta \zeta} \right) \mathcal{P}, \quad (42)$$

where

$$\begin{aligned} \mathcal{P} &= \int D\zeta D\zeta' \frac{\delta}{\delta \psi_A(\zeta)} \mathcal{L}_{AB}(\zeta, \zeta') \frac{\delta \mathcal{F}}{\delta \psi_B(\zeta')} \\ &+ \int D\zeta D\zeta' \frac{\delta}{\delta \psi_A(\zeta)} \mathcal{S}_{AB}(\zeta, \zeta') \frac{\delta}{\delta \psi_B(\zeta')}. \end{aligned} \quad (43)$$

Note that (43) has the same form for both specific examples discussed in Sect. 3. Note, therefore, that for the relativistic plasma, it is a fully relativistic Fokker–Planck equation for the dynamical variables $\psi_A(\zeta)$, which in this case are the Vlasov one-particle distribution function $f(\mathbf{r}, \mathbf{p})$, with \mathbf{p} given by (13) and electromagnetic field (\mathbf{E}, \mathbf{B}) . For semiclassical Bloch electrons, \mathbf{p} is the kinematic momentum of the carrier. Assuming that the Hamiltonian for semiclassical Bloch carriers is given as a nonrelativistic form of (18) supplemented with the

The dissipative bracket for relativistic classical plasma, discussed earlier, should yield the collision operator on the RHS of the first equation in (20) in the form of the Landau collision operator [3]. This requirement gives the kernel $\Delta(\mathbf{z}, \mathbf{z}'; \mathbf{z}_1 \mathbf{z}_2)$ in the form

Zeeman-like coupling proportional $\sum_{j=1}^3 \gamma f_j B_j(\mathbf{1})$ and neglecting the internal electric and magnetic fields generated by the motion of carriers and using the dissipative brackets (32) stemming from dissipative brackets for spins [43]

$$\prec S_i, S_j \succ = -\lambda |\mathbf{S}| \left(\delta_{ij} - \frac{S_i S_j}{\mathbf{S}^2} \right), \quad (44)$$

we can derive the conservation equation for the spin density $S_i(\mathbf{r}) = \int d\mathbf{p} f_i(\mathbf{r}, \mathbf{p})$, which is equivalent to the convective version of the Gilbert–Landau equation for sample magnetization [44]

$$\frac{\partial \mathbf{S}}{\partial t} + \nabla \cdot \mathbf{A} = \gamma \mathbf{S} \times \mathbf{B} - \lambda \mathbf{S} \times (\mathbf{S} \times \mathbf{B}), \quad (45)$$

where $A_{ij} = \int d\mathbf{p} p_i f_j(\mathbf{r}, \mathbf{p})$ is the spin current tensor. The explicit form of \mathbf{A} follows from the Chapman–Enskog approximations in solving the kinetic equation (31).

The above example of continuum equations following from metriplectic analysis of the Vlasov–Klimontovich description of the many particles system allows us to derive the hydrodynamic-like description of that system. These continuum mechanics equations can also be cast in the form of metriplectic dynamics [41], and it has been discussed in many recent publications of Massimo Materassi and Phil Morrison and their collaborators [13, 36]. Whether this technique can be useful in other applications, for example, in the theory of quark–gluon plasma, remains to be seen.

5. Conclusions

The above sections contain a discussion of the use of Vlasov–Klimontovich formulation of the classical many particle systems dynamics. Some important generalisations of that formulation, for example, quantum many-body problems or classical hydrodynamics [36] are mentioned in included references. There is essentially no applications of that formulation in equilibrium statistical mechanics in spite of the fact that the Vlasov description could easily be used within the Martin–Rose–Sigma formulation [45]. The general relativity generalisation of the kinetic theory base on the present above formulation [23] is now being prepared for publication.

Acknowledgments

This mini-review is based on collaborative work with many colleagues and students whose contributions are acknowledged in the paper bibliography.

Iwo Białynicki-Birula was the mentor for many generations of physicists, in particular, to the group which joined him at the Center for Theoretical Physics in 1979 and continues to work at the same institution until today. I personally benefited throughout all that time from intellectual interactions with Professor Iwo Białynicki-Birula and from his support and friendship in the bright and dark periods of that institution and the Country. We hope to continue to benefit from his remarkable wisdom and confidence for a better future of science and the world for many more years.

References

- [1] A.A. Vlasov, *J. Exp. Theor. Phys.* **3**, 291 (1938) (in Russian).
- [2] M.N. Rosenbluth, W.M. MacDonald, D.L. Judd *Phys. Rev.* **107**, 1 (1957).
- [3] N.G. van Kampen, B.U. Felderhof, *Theoretical Methods of Plasma Physics*, Cambridge University Press, 2009.
- [4] A.A. Vlasov, *Nonlocal Statistical Mechanics*, Nauka, Moscow 1978 (in Russian).
- [5] *Quark-Gluon Plasma: Theoretical Foundations: An Annotated Reprint Collection*, Eds. J. Kapusta, B. Müller, J. Rafelski, 2003.
- [6] G. Manzi, R. Marra, in: *Transport Phenomena and Kinetic Theory, Applications to Gases, Semiconductors, Photons, and Biological Systems*, Eds. C. Cercignani, E. Gabetta, Birkhäuser, 2007.
- [7] W. Braun, K. Hepp, *Commun. Math. Phys.* **56**, 101 (1977).
- [8] R. Zwanzig, *Phys. Rev.* **144**, 170 (1966).
- [9] Y.L. Klimontovich, *Zh. Eksp. Teor. Fiz.* **34**, 173 (1958).
- [10] Y.L. Klimontovich, *Statistical Theory of Nonequilibrium Processes in Plasma*, MIT Press, Cambridge 1984.
- [11] L.A. Turski, in: *Continuum Models and Discrete Systems*, Vol. 1, Ed. G.A. Maugin, Longman, London 1990.
- [12] J.E. Marsden, T.S. Ratiu, *Introduction to Mechanics and Symmetry*, Springer-Verlag, Berlin 2002.
- [13] P.J. Morrison *J. Phys. Conf. Ser.* **169**, 012006 (2009).
- [14] E.P. Wigner, *Phys. Rev.* **40**, 749 (1932).
- [15] J.E. Moyal, *Proc. Camb. Phil. Soc.* **45**, 99 (1949).
- [16] P. Goldstein, L.A. Turski, *Physica A* **89**, 481 (1977).
- [17] Z.R. Iwiński, L.A. Turski, *Lett. Appl. Sci. Eng.* **4**, 179 (1976).
- [18] I. Białynicki-Birula, J.C. Hubbard, L.A. Turski, *Physica A* **128**, 509 (1984).
- [19] M. Born, L. Infeld, *Proc. R. Soc. London* **150**, 141 (1935).
- [20] J.E. Marsden, A. Weinstein, *Physica D* **4**, 394, (1982).
- [21] J.M. Souriau, *Structure des Systéms Dynamique* Dunod, Paris 1970.
- [22] J.F. Carineña, H. Figueroa, P. Guha, [arXiv:1501.04917v2](https://arxiv.org/abs/1501.04917v2), 2015.
- [23] I. Białynicki-Birula, L.A. Turski, to be publish.
- [24] N.W. Ashcroft, N.D. Mermin, *Solid State Physics*, Philadelphia 1976.
- [25] J. Zak, *Phys. Rev. Lett.* **62**, 2747 (1989).
- [26] M.V. Berry, *Proc. R. Soc. A* **392**, 45 (1984).
- [27] D. Culcer, J. Sinova, N.A. Sinitsyn, T. Jungwirth, A.H. MacDonald, Q. Niu, *Phys. Rev. Lett.* **93**, 046602 (2004).
- [28] G. Sundaram, Q. Niu, *Phys. Rev. B* **59**, 14915 (2002).
- [29] L.A. Turski, *Dissipative Quantum Mechanics*, Springer Lect. Notes in Physics 4777, Eds. Z. Petru, J. Przystawa, K. Rapcewicz, New York 1996.
- [30] W.C. Kerr, M.J. Rave, L.A. Turski, *Phys. Rev. Lett.* **94** 176403 (2005).
- [31] A.E.F. Djemai, *In. J. Theor. Phys.* **43**, 299, (2004).
- [32] Z.W. Gortel, M.A. Załuska-Kotur, L.A. Turski, *Phys. Rev. B* **52**, 16916 (1995).
- [33] S. Dattagupta, L.A. Turski, *Phys. Rev. A* **32**, 1439 (1985).
- [34] S.Q. Nguyen, L.A. Turski, *Physica A* **290**, 431 (2001).
- [35] S.Q. Nguyen, L.A. Turski, *J. Phys. A* **34**, 9281 (2001).
- [36] M. Materassi, E. Tassi, *Physica D* **241**, 729 (2012).
- [37] A. Kaufman, *Phys. Lett. A* **100** (1984).
- [38] A. Kaufman, *Phys. Lett. A* **109**, 87, (1985).
- [39] P.J. Morrison, *Phys. Lett. A* **100**, 423 (1984).
- [40] A.N. Kaufman, L.A. Turski, *Phys. Lett. A* **120**, 331 (1987).
- [41] C.P. Enz, L.A. Turski, *Physica A* **96**, 369 (1979).

- [42] Ching Lok Chong, *J.Non-Newtonian Fluid Mech.* **292**, 104537 (2021).
- [43] J.A. Holyst, L.A. Turski, *Phys. Rev. A* **45**, 6180 (1992).
- [44] T.B. Boykin, *Am. J. Phys.* **69**, 793 (2001).
- [45] S.P.C. Martin, E.D. Siggia, H.A. Rose, *Phys. Rev. A* **8**, 423 (1973).

DEDICATED TO PROFESSOR IWO BIAŁYNICKI-BIRULA ON HIS 90TH BIRTHDAY

— \diamond —

Molecular Dynamics Approach to Processes in Bulk Au Materials

VAN CAO LONG* AND DUNG NGUYEN TRONG

Institute of Physics, University of Zielona Góra, Prof. Szafrana 4a, 65-516 Zielona Góra, Poland

Published: 14.06.2023

Doi: [10.12693/APhysPolA.143.S191](https://doi.org/10.12693/APhysPolA.143.S191)*e-mail: caolongvanuz@gmail.com

The paper studies the influence of several factors such as the number of atoms (N), temperature (T), and annealing time (t) on the characteristic quantities of structure, phase transition, and crystallization progress of bulk Au materials using so-called molecular dynamics simulation. The authors hope that obtained results would be very useful for near future experimental research with Au.

topics: cubic Au materials, glass temperature, number of atoms, molecular dynamics

1. Introduction

In this paper, we would like to present an example of the usefulness of scientific knowledge gained in the school of Professor Iwo Białynicki-Birula (the Master). Namely, in the middle eighties of the last century, the present author (Van Cao Long) together with other Master's outstanding descendants: Professor Jan Mostowski from the Institute of Physics, Polish Academy of Sciences, and Professor Marek Trippenbach from the University of Warsaw, we tried to construct a classical theory of ionization of many-electron atoms [1]. Our classical model of ionization is described in detail in the third chapter of M. Trippenbach's doctoral dissertation [2]. This model is practically a precursor of the simulation method — the so-called molecular dynamics, very popular nowadays in the simulation of multi-particle systems — on which the present author (Van Cao Long) is now working intensively with the current Ph.D. student Nguyen Trong Dung. During only the last two years, we published approximately 20 papers [3–22].

The quantum description of a large, multi-element physical system is usually complex and impossible to analyze by current supercomputers. Often these attempts obscure the physical picture of the phenomena under consideration. Therefore, approximate baseline simulations were used in the probabilistic treatment of the system with

the parameters included in the problem after their quantum averaging, while the tested physical quantities were already treated as random variables or classical stochastic processes, and further numerical simulations were carried out using classical equations. This approach was first used in the 1950s and then developed intensively in the 1980s. One of us (Van Cao Long) actively participated in these processes during his work in the Department of Theoretical Physics (now the Center for Theoretical Physics of the Polish Academy of Sciences) in the eighties of the last century. In particular, the research emphasized above and described in detail in M. Trippenbach's doctoral dissertation "Selected issues concerning ionization of atoms in the field of strong laser waves" [2] can be treated as a precursor of the deterministic simulation method, currently widely used in the theory of solids, chemistry, and molecular biology, namely the so-called molecular dynamics method. The model we proposed consisted in drawing the initial conditions for the position and momentum of a particular electron using the Monte Carlo method according to their certain distribution in the phase space, namely with the modified Wigner distribution. Then we simulated the dynamics of electrons classically, i.e., we solved Newton's equations. We counted the total energy of the electron at a given moment with the newly obtained distribution in the phase space. When this energy is positive, we say we have ionization. In this

way, we have studied the ionization of a multi-electron atom. This scheme is exactly the same as in the molecular dynamics (MD) method. We will take an example below concerning the study of bulk Au materials to illustrate this powerful simulation method.

2. Structure, phase transition, and crystallization progress of bulk Au materials within the molecular dynamics (MD) simulation method

It is widely known that gold (Au) is a precious metal, inert and strongly colored [23, 24], with the number of protons $Z = 79$ and the electronic structure $4f^{14}5d^{10}6s^1$. It has great applicability in scientific and technological fields due to biocompatibility, ability to perform biological/chemical analyses [25, 26]. Gold is used in biosensors [27], cancer therapy [28], as jewelry and foreign currency for exchange and economic transactions in life. To study and fabricate bulk Au materials, scientists use research methods such as experiment, theory, and simulation. One can say that the traditional division of physics and other natural sciences into “experiment” and “theory” is no longer valid. Computer simulation is the third that is complementary to these two traditional branches of science.

With the experimental method, Yuk et al. [28] successfully studied the process of crystallization of the material at the temperature of $T = 700$ K, with an appearance time of only 10 s, which is a very short time for observations [28]. The bond length of Au–Au determined by experiments is $r_{\text{Au–Au}} = 2.48, 2.88$ Å, $3.6\text{--}4.0$ Å [29], and $3.0\text{--}3.6$ Å [30], whereas by the simulation it is $r_{\text{Au–Au}} = 2.8$ Å [42] and 2.9 Å [31].

Results with such large deviations are explained by impurities. To confirm this, the authors successfully studied the influence of factors on the bond length of the CuAu alloys with $r_{\text{Au–Au}} = 3.05$ Å [10] and of the NiAu alloys with $r_{\text{Au–Au}} = 3.09\text{--}3.17$ Å when the heating rate increases from 4×10^{12} to 4×10^{14} K/s [4]. In particular, the simulation method is considered the most advanced tool today, with low cost and high efficiency. It gives access to materials with sizes less than 2 nm, placed in harsh conditions such as those in the center of the earth at a temperature of $T = 7000$ K and a pressure of $P = 360$ GPa with materials MgSiO_3 [32] and CaSiO_3 [33], when experimental methods are inaccessible.

To study the structural properties of Au cubic materials, researchers mainly use the molecular dynamics (MD) method [34, 35]. The reason is that this research method can study large-sized materials using the equations of motion according to Newtonian second laws. Research results include, among others, consideration of the radial distribution function (RDF) [36, 37]. In this formalism, the distance

to the second peak of RDF [38] defines the size of the crystalline state of the material. When the Honeycutt–Andersen (HA) structural transformation appears in these materials [39], the deformation of atoms [40, 41] can be determined and this characteristics requires the choice of properly embedded atomic interactions (EAM). Mishin et al. [42] successfully determined the structural stability and defects in the crystal lattice with Cu metal. Other authors [43, 44] studied the structure of materials in liquid metals. The obtained results show that crystallization usually occurs at a size of about 10 nm. In MD simulation studies, researchers often use commercial or free software such as LAMMPS [45]. Zhou et al. [46] successfully studied the deformation of Cu thin films with the simulation method. Here an important question appears, what factors affect the structure, phase transition, crystallization, and phase transition temperature of bulk Au materials? To answer this question, in the content of this paper we study the influence of factors such as atomic number, temperature, and annealing time on structural characteristics, phase transition, crystallinity, and phase transition temperature. The obtained results would be the basis for experimental studying the structural characteristics of Au bulk materials in the future.

3. Calculation method

According to the MD simulation scheme, initially, the Au atoms are randomly seeded into a cubic pattern of size (l), determined by

$$l = \sqrt[3]{\frac{N}{\rho}}. \quad (1)$$

For interaction potential between atoms in the classical Newtonian equations, we choose the Sutton–Chen (SC) embedded interaction potential field [47, 48] in the framework of the MD simulation method with periodic boundary conditions. The following formulas are used

$$E_{\text{tot}} = \frac{1}{2} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \Phi(r_{ij}) + F(\rho_i),$$

$$\Phi(r_{ij}) = \varepsilon \left(\frac{a}{r_{ij}} \right)^n,$$

$$F(\rho_i) = A_i E_i^0 \rho_i \ln(\rho_i),$$

$$[\rho_i(R)]^{\alpha(l)} = e^{-b^*},$$

$$b^* = \beta_i^{(l)} \left(\frac{R}{R_i^0 - 1} \right),$$

$$(\bar{\rho}_i)^2 = \sum_{l=0}^3 t_i^{(l)} (\rho_i^{(l)})^2. \quad (2)$$

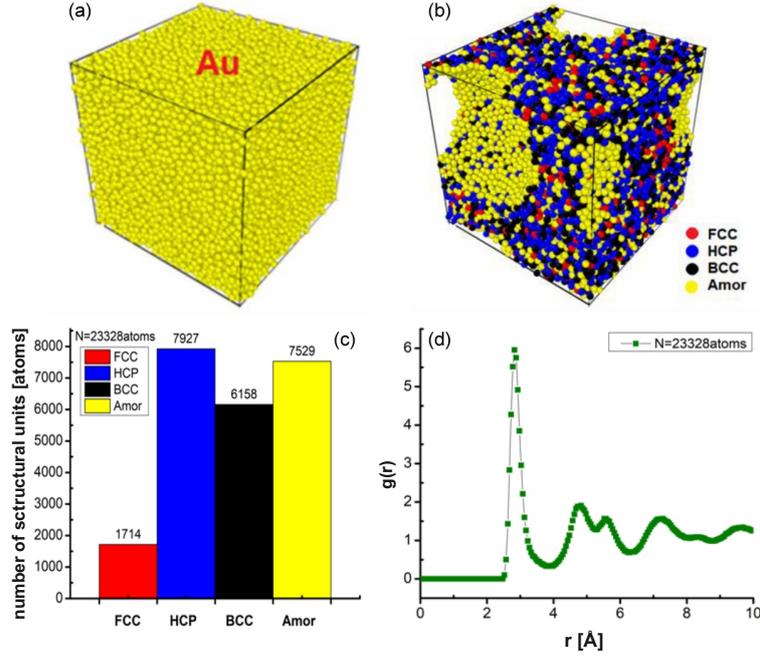


Fig. 1. Initial shape (a), structure shape (b), number of structural units (c), RDF (d) of Au_{23328} material at $T = 300$ K.

In order to increase the accuracy, we choose the interaction force field parameters as $E_i^0 = 3.93$ eV, $R_i^0 = 2.0$ Å, $\alpha = 6.34$, $A_i = 1.0$, $\beta_i^{(0)} = 5.77$, $\beta_i^{(1)} = 2.2$, $\beta_i^{(2)} = 6.0$, $\beta_i^{(4)} = 2.2$, $t_i^{(0)} = 1.0$, $t_i^{(1)} = 2.9$, $t_i^{(2)} = 1.64$, $t_i^{(4)} = 2.0$ [49]. To implement the simulation program, we use the open-source code LAMMPS [47, 48], which is completely free for the scientific research community. This source code works on the basis of the MD method combined with embedded interaction potential (EAM) to simulate the interactions between Au atoms. Initially, for the bulk Au material, we run 2×10^4 MD recovery statistic runs at a heating rate of 4×10^{11} K/s and a heating step time of 1 fs at a temperature $T = 4000$ K. Then the temperature is reduced from 4000 to 300 K to increase the crystallization state of the material. The number of atoms (N) in the considered model is taken as $N = 4000$ Au atoms (Au_{4000}), 6912 Au atoms (Au_{6912}), 10976 Au atoms (Au_{10976}), 16384 Au atom (Au_{16384}), and 23328 Au atom (Au_{23328}). After determining the Au_{23328} material with the highest crystallinity, we chose this material as a sample to continue the study of the influence of temperature. As a consequence, we successfully determine the glass temperature (T_g) as $T_g = 600$ K and continue the annealing after an annealing time $t = 400$ ps. To study the structural features, we determine the characteristics, i.e., shape, structure shape, and number of structural units. The radial distribution function (RDF) $g(r)$ [50] is given by the formula

$$g(r) = \frac{n(r)}{4\pi r^2 \rho_0 dr}, \quad (3)$$

where $g(r)$, $n(r)$ are the functions used in the calculations: the probability of finding the atom i in space, the number of atoms, respectively; ρ_0 is the atomic density, and r is the distance from the atom i to other atoms. The total energy of the system E_{tot} is given in [51]. To determine the diffusion mechanism of the atoms, mean squared distance (MSD) is applied according to the following expression [52]

$$r(t) = \sqrt{\frac{1}{N} \left[\sum_{i=1}^N |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \right]}, \quad (4)$$

where N is the number of atoms in the system, $r_i(0)$ is the initial position of the i atom, whereas $r_i(t)$ is the position of the atom i at time t . In all samples, phase transitions were studied by applying Common Neighborhood Analysis (CNA) [55, 56], according to Nosé's [53], see also Hoover [54]. OVITO software was used [57, 58] to visualize the results. Finally, to check the accuracy of the results, Dual Energy X-ray Absorbance Measurement (DXA) was used [59]. These programs are edited and used in the computer system of the Institute of Physics, University of Zielona Góra, Poland.

4. Results and discussion

4.1. Structural properties of bulk Au materials

The structural properties of bulk Au are shown in Fig. 1.

The obtained results show that the material of Au_{23328} metal has the shape of a cube, made up of Au metal atoms identified by yellow color (Fig. 1a).

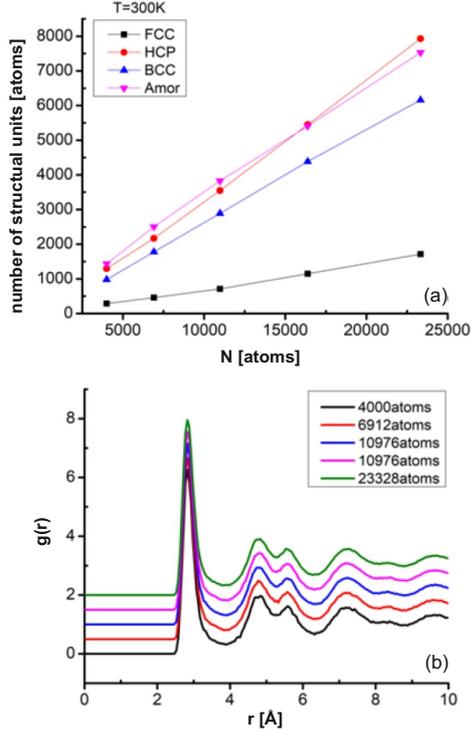


Fig. 2. Number of structural units (a), radial distribution function (b) of Au blocks with different number of atoms.

The Au_{23328} material is determined by 4 structural units: face-centered cubic (FCC) shown in red, hexagonal packing (HCP) shown in blue, body-centered cubic (BCC) shown in black, and amorphous (Amor) shown in yellow (Fig. 1b). The number of the structural units is 1714 FCC, 7927 HCP, 6158 BCC, 7529 Amor (Fig. 1c), the length of the link $r_{\text{Au-Au}} = 2.825$ Å, the height of the first peak of the radial distribution function $g(r) = 5.95$ (Fig. 1d). Our simulation results are completely consistent with the experimental results $r_{\text{Au-Au}} = 2.48, 2.88$ Å, and $3.6\text{--}4.0$ Å [29] and with the simulation results obtained before: $r_{\text{Au-Au}} = 2.8$ Å [42] and 2.9 Å [31]. In addition, the material has material size (l) corresponding to $l = 7.965$ nm and $E_{\text{tot}} = -88328$ eV. The obtained results show that the structural characteristic quantities of the Au_{23328} material model are consistent with the results obtained before [29, 31, 42] and create a premise for the process of studying the influencing factors.

4.2. Influencing factors

To study the influencing factors, we use the material models Au_{4000} , Au_{6912} , Au_{10976} , Au_{16384} , Au_{23328} . In particular, we consider the influence of size, the influence of temperature at $T = 300, 400, 500, 600, 700, 800, 900, 1000$ K and incubation time after $t = 0, 10, 50, 250, 400$ ps.

4.2.1. Effect of the number of atoms

The results in Fig. 2a show how the number of atoms affects the structural properties of the considered material. Here, the results are for the metal material Au_{4000} (its structural shape is shown in Fig. 1a). The number of structural units is represented by 1714 FCC, 7927 HCP, 6158 BCC, 7529 Amor (the same as in Fig. 1c) and $r_{\text{Au-Au}} = 2.825$ Å, $g(r) = 5.95$ (as in Fig. 1d) and $l = 7.965$ nm, $E_{\text{tot}} = -88328$ eV. When the number of atoms increases from $N = 4000$ (Au_{4000}) through $N = 6912$ atoms (Au_{6912}), 10976 atoms (Au_{10976}), 16384 atoms (Au_{16384}), up to 23328 atoms (Au_{23328}), the shape the material structure changed as follows: the number of structural units for FCC increases from 286 to 1714, for HCP it increases from 1295 to 7927, for BCC this number increases from 982 to 6158, for the Amor units it increases from 1437 to 7529 (see Fig. 2a). Now, whereas $g(r)$ is unchanged for $r < r_{\text{Au-Au}} = 2.825$ Å, it decreased from $g(r) = 6.27$ to $g(r) = 5.95$ (see Fig. 2b).

In Fig. 3a, the structure size l increases from 4.502 to 7.965 nm, and the system total energy E_{tot} decreases from -15025 to -88328 eV (Fig. 3b). The obtained results show a close relationship between the number of atoms, the size, and the total energy of the system (Fig. 3). The linear relationship

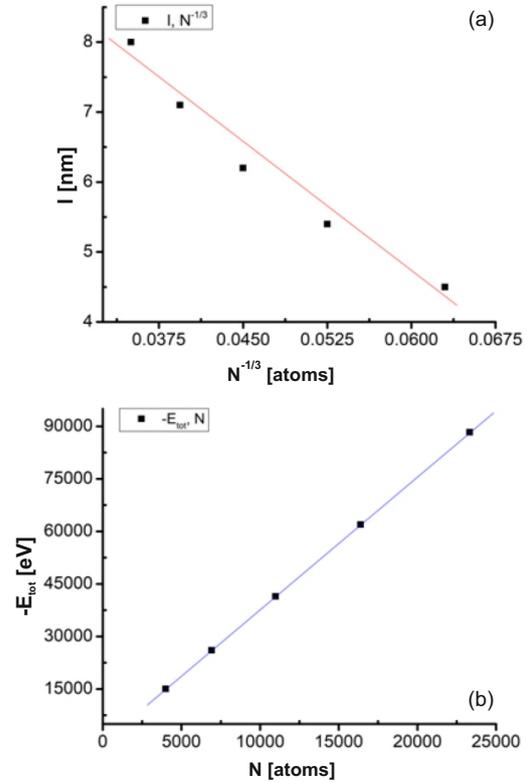


Fig. 3. Relationship between the number of atoms and size (a), between the number of atoms and total energy of the system (b) as N changes.

Structural characteristics and phase transitions of Au₂₃₃₂₈ materials at different temperatures.

TABLE I

T [K]	300	400	500	600	700	800	900	1000
r [Å]	2.825	2.825	2.825	2.825	2.825	2.825	2.825	2.825
$g(r)$	5.95	5.56	5.26	4.91	4.59	4.32	4.09	3.88
FCC	1714	1621	1536	1405	1292	1040	971	760
HCP	7927	7796	7578	7547	7386	6944	6606	6103
BCC	6158	6226	6073	5999	5863	5709	5294	4906
Amor	7529	7685	8141	8377	8787	9635	10457	11559
l [nm]	7.965	7.967	7.971	7.975	7.982	7.990	7.995	8.000
E_{tot} [eV]	-88328	-87965	-87588	-87153	-86701	-86194	-85668	-85099

between the size (l) and $N^{-1/3}$ (where N is the number of atoms) is described by the expression $l = 11.97 - 121.98 N^{-1/3}$. Similarly, the dependence of the total energy of the system (E_{tot}) and the inverse of the the number of atoms (N) is described by the linear expression $E_{\text{tot}} = -164.93 - 3.793 N^{-1}$ (Fig. 3b).

There is a clear similarity between our results and those for metals in [60, 61]. When increasing the number of atoms, the size of matter increases with the ratio $l \sim N^{-1/3}$, and the total energy of the system decreases with the ratio $E_{\text{tot}} \sim N^{-1}$. This phenomenon is due to the size effect and also to the surface effect. The obtained results can be the basis for testing methods to test future applications. To ensure the calculation speed as well as the stability of the structural features, we choose the Au₂₃₃₂₈ material model as the standard material to investigate the influencing factors in the following sections.

4.2.2. Effects of temperature

The effect of temperature on the structural characteristics and phase transition of Au₂₃₃₂₈ materials is shown in Table I. When T increases from 300 to 1000 K, the results show that E_{tot} increases from -88328 to -85099 eV, the size l increases very slightly from 7.965 to 8.000 nm, r is constant $r_{\text{Au-Au}} = 2.825$ Å, and $g(r)$ decreases from 5.95 to 3.88. Correspondingly, the number of FCC structural units decreases from 1714 to 760, of the HCP structural units it decreases from 7927 to 6103, of BCC this number decreases from 6158 to 4906, and of the Amor units it increases from 7529 to 11559 (Table I). The obtained results demonstrate that when the temperature increases, the length of the Au-Au links remains constant, while the number of FCC, HCP, and BCC structural units decrease and increases only for the Amor unit. The relationship between the sum of E_{tot} and T is shown in Fig. 4.

The obtained results indicate that as the temperature T increases, E always increases linearly and proportional to T in two ranges, i.e., 300–600 K

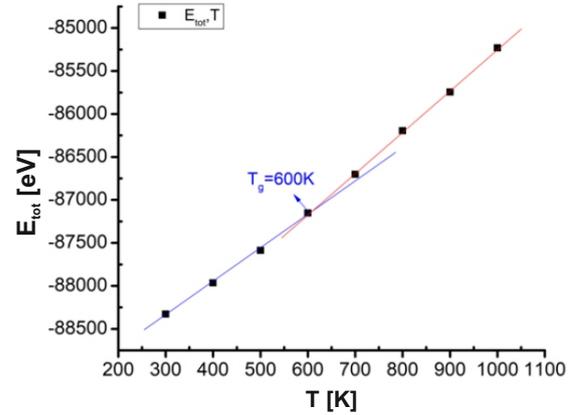


Fig. 4. Relationship between total energy of the system and temperature of the Au₂₃₃₂₈ material.

and 600–1000 K. Increasing of T leads to an increase in l , an increase in the total energy of the system E_{tot} , a decrease in the number of FCC, HCP and BCC structural units, and an increase in Amor. It follows that at $T = 600$ K the glassiness of the material Au₂₃₃₂₈ appears. The intersection between the two temperature ranges is called the glass point or glass temperature (T_g), with the value $T_g = 600$ K in Fig. 4. The cause of this phenomenon is due to the effective size, the effective surface and the fact that phase transition of materials Au₂₃₃₂₈ is of type 1. To determine the crystallization process of Au₂₃₃₂₈ materials at the glass transition temperature $T_g = 600$ K, we investigate the effect of the incubation time. The result will be presented in the following section. In another paper [62], the authors determined the glass transition temperature by taking the intersection of these lines and obtained the value of 700 K [62], which is evidently different from the value we established in the present study (600 K). Our simulations show that the simulation procedure with the use of the SC potential is stable for the number of atoms of the order of thousands. Meanwhile, the results of our analysis of structural phase transitions of bulk Au₂₃₃₂₈ materials are shown in more detail.

4.2.3. Effect of incubation time

Figure 5 presents the results, when annealing Au_{23328} material at temperatures $T = 500, 600, 700$ K is considered at different annealing times (t), i.e., t_1, t_2, t_3, t_4 and the annealing time increases. At a given temperature (T), the total energy of the system (E_{tot}) decreases, showing that the crystallization process of the material increases. At $T = 500$ K, E_{tot} decreased the least, slightly greater decrease was at $T = 700$ K, and the largest decrease was at $T = 600$ K after $t_3 = 250$ ps. The obtained results are completely consistent with the results obtained above for $T_g = 600$ K, so we confirm that this is the crystallization temperature of the material. To confirm this in a more demonstrated manner, we chose the temperature to be investigated at each different annealing time as $T = T_g = 600$ K — the results are shown in Fig. 6.

For the Au_{23328} bulk material at glass temperature $T_g = 600$ K and annealing time $t_1 = 10$ ps, with the structural form shown in Fig. 6a, the number of the structural units is 941 FCC, 6717 HCP, 5104 BCC, 10566 Amor (Fig. 6e), the length of the link is $r_{\text{Au-Au}} = 2.825$ Å, and $g(r) = 3.86$, $l = 7.644$ nm, $E_{\text{tot}} = -85988$ eV (Fig. 6i). As the annealing time increased from $t_1 = 10$ ps to $t_2 = 50$ ps, $t_3 = 250$ ps, and $t_4 = 400$ ps, the structural geometry of the material changed (Fig. 6a–d), with the number of FCC units increasing from 941 to 1481. In ad-

dition, the number of HCP units increased from 6717 to 7598, of BCC increased from 5104 to 6507, and of Amor units increased from 10566 to 7742 (Fig. 6e–h). Accordingly, for the cubic Au_{23328} , $g(r)$ has a constant value when $r < r_{\text{Au-Au}} = 2.825$ Å. Then the value of $g(r)$ increases with annealing time from 3.86 (at t_1) to 4.38 (at t_4), l decreases slightly from 7.644 to 7.643 nm, E_{tot} decreases from -85988 to -87636 eV (Fig. 6i–l).

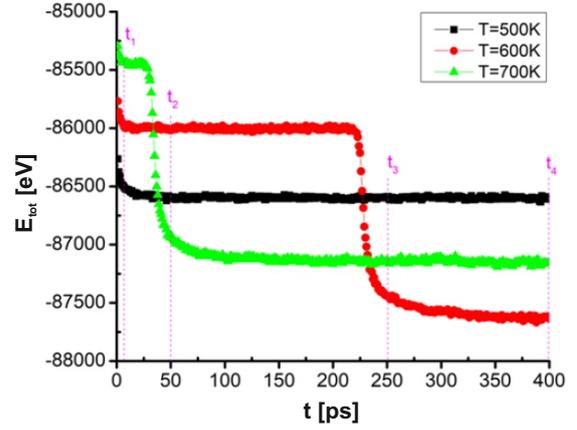


Fig. 5. Phase transition of bulk Au_{23328} at glass temperatures $T = 500, 600, 700$ K as a function of annealing time t .

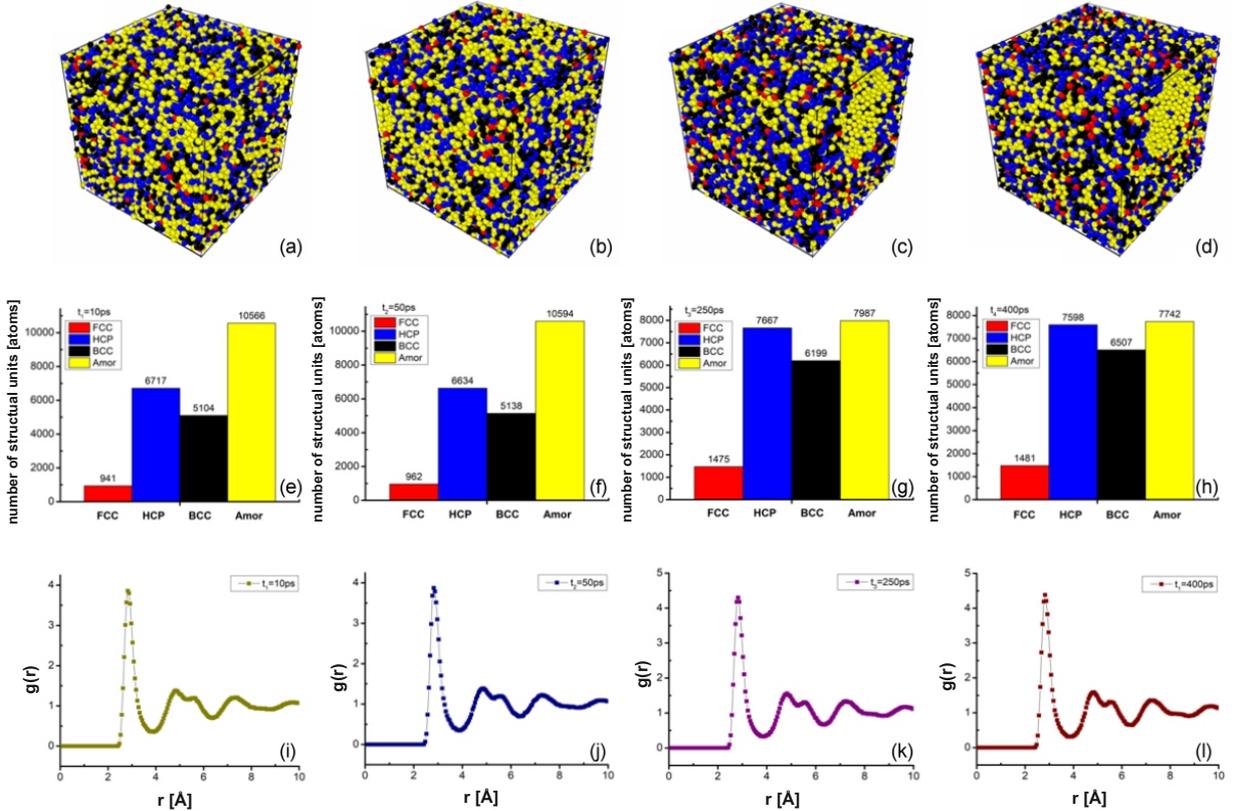


Fig. 6. The crystallization process of Au_{23328} bulk material at glass temperature $T_g = 600$ K for different annealing times (a, e, i) $t_1 = 10$ ps; (b, f, j) $t_2 = 50$ ps; (c, g, k) $t_3 = 250$ ps; (d, h, l) $t_4 = 400$ ps.

The obtained results show that when increasing the annealing time, the length of the links Au–Au has a constant value $r_{\text{Au–Au}} = 2.825 \text{ \AA}$ and this result is consistent with the results obtained previously for Au [29, 31, 42]. Also, our results showed that after annealing time at glass temperature $T_g = 600 \text{ K}$, the crystallization process increased. We hope that the recent results obtained here will serve as the basis for future experimental studies.

5. Conclusions

The results obtained in this paper show the influence of such factors as the number of atoms, temperature, and annealing time on the structural characteristic quantities of bulk Au materials. We used here the Sutton–Chen embedded interaction potential, which is the most suitable for the periodic boundary conditions and the Verlet algorithm in the MD scheme. In the equilibrium state, bulk Au materials always exist in 4 types of structures: FCC, HCP, BCC, and Amor. An increase of the number of atoms (N) leads to an increase in size (l), whereas the total energy of the system (E_{tot}) decreases, and at the same time, an increase of the annealing time (t) leads to the decrease of both l and E_{tot} . As temperature increases, both l and E_{tot} increase. In addition, by increasing t at $T = 600 \text{ K}$, the number of the structural units FCC, HCP, BCC increases, and the Amor unit decreases, which is confirmed by the fact that the glass temperature of the bulk Au material is $T_g = 600 \text{ K}$. The influence of N , T , and t on the structure, phase transition, and crystallization of Au bulk materials was considered in detail. We observed that the length of the links Au–Au has a constant value $r_{\text{Au–Au}} = 2.825 \text{ \AA}$ when the influencing factors change. It follows that the influencing factors do not change the length of links, but they change the height of the first peak of the radial distribution function. The obtained results are completely consistent with the experimental and simulation results obtained previously. Also, as N increases, the size l of the material increases proportionally to $N^{-1/3}$ and the total energy E_{tot} is proportional to N^{-1} . The results have successfully determined that the crystallization process increases at the glass temperature $T_g = 600 \text{ K}$. The question that why the crystallization temperature leads to an increase in the crystallization process has not been yet explained. This is devoted to future studies.

6. Relations with Professor Iwo Białyński-Birula — personal remarks by Cao Long Van

I was born in Vietnam, a poor country, underdeveloped due to constant wars, with no traditions in the sciences. Nevertheless, in my youth, I was well educated in mathematics, and since mathematics is



Fig. 7. A centuries-old traditional image of a teacher from the Nghệ An area of Vietnam “Ông đồ xứ Nghệ”, where Vinh University is located. Professor Iwo Białyński-Birula has many scientific descendants at this University.

the language of nature, I was able to deeply learn physics (fundamental science) with love. At the beginning of this path, I was very lucky to find my second homeland, Poland, where, through the school of life, I found exceptional teachers in my profession, in particular Professor Dr. hab. Iwo Białyński-Birula.

The history of my contact with Professor Iwo Białyński-Birula is very long, more than 50 years. During my studies at the Warsaw University in the years 1971–1976, I was one of the best students of my generation. As a consequence, from the fourth semester I was on an individual course of study under the supervision of Dr. Adam Bechler, who is now a full professor at the University of Szczecin (retired), but at that time he had just completed his doctorate under the supervision of Professor Dr. hab. Iwo Białyński-Birula in the Department of Field Theory and Statistical Physics, Institute of Theoretical Physics, University of Warsaw. Since then, my scientific and educational career has been closely related to the school of Professor I. Białyński-Birula with his outstanding students, such as Professors Kazimierz Rzażewski, Krzysztof Wódkiewicz, Jan Mostowski, and “scientific grandchildren” such as Professors Maciej Lewenstein, Marek Kuś, or Marek Trippenbach. It is enough to emphasize the fact that the physicists of three generations of this school (namely, the Master himself, his student K. Rzażewski and his “scientific grandson” M. Lewenstein) were laureates of the Polish Science Foundation Award. Over time, they became my older and younger friends[†] and work-

[†]It is a great regret that one of them, Professor K. Wódkiewicz, passed away prematurely due to a serious illness.

ing with them was a pure pleasure. It can be said that the long-term friendship and cooperation with them under the guidance of my Master (who was the supervisor of my masters and Ph.D. thesis), my mentor, an excellent scientist and teacher — Professor I. Białyński-Birula, who treated me as “the nicest student”, he provided me with basic professional knowledge necessary for my later achievements both in research and teaching (see Fig. 7). As an example of this, me and his “scientific grandson” Professor M. Trippenbach more than 20 years ago created the best School of Optics in the biggest university in Middle Vietnam, i.e., Vinh University. For this contribution, we were both honored with the title of Doctor Honoris Causa of this University. We regularly organized international conferences and workshops at Vinh University. As an example, I and another “scientific grandson” of Professor I. Białyński-Birula, i.e., Professor M. Lewenstein, organized the Workshop on Quantum Information at this University in 2016. Thus Professor I. Białyński-Birula has many Vietnamese students who played an important role in development of physics in Vietnam.

To summarize, my Master Professor I. Białyński-Birula and his outstanding former students of different generations are real masters who have vast knowledge and experience in transferring this knowledge to students. I myself had the opportunity to educate under the guidance of one of the best, then work among older and younger wonderful friends in my environment. Therefore, now, when I am close to the 70th anniversary of my life, I can say with confidence that all my dreams about my profession have come true. I am one of many successive masters who, accepting the baton from their excellent predecessors, fulfilled the mission of deepening their physical knowledge and passing this knowledge to the next generations of physicists, thus contributing to the development of this exceptionally beautiful and at the same time meritorious science. For this, I am immensely grateful to my Master Professor Iwo Białyński-Birula. On the occasion of his 90th birthday, I would like to wish him a lot of health and long life, so that we can admire the wonderful teacher for the years to come.

References

- [1] J. Mostowski, M. Trippenbach, C.L. Van, in: *Fourth Int. Conf. on mMltiphoton Processes*, Boulder (CO) 1987.
- [2] M. Trippenbach, Ph.D. Thesis, Center of Theoretical Physics, Polish Academy of Science, 1987 (unpublished).
- [3] C.-L. Van, D.-Q. Van, N.-T. Dung, *ACS Omega* **5**, 31391 (2020).
- [4] N.-T. Dung, C.-L. Van, Ş. Tǎlu, *J. Compos. Sci.* **5**, 18 (2021).
- [5] N.-Q. Hoc, N.-D. Hien, N.-T. Dung, C.-L. Van, Ş. Tǎlu, *J. Compos. Sci.* **5**, 153 (2021).
- [6] N.-T. Dung, C.-L. Van, *Adv. Civ. Eng.* **2021**, 9976633 (2021).
- [7] V.-Q. Trung, D.-B. Dai, T.-T.T. Duong et al., *Designed monomers and polymers* **24**, 274 (2021).
- [8] N.-T. Dung, C.-L. Van, *Mater. Today Commun.* **29**, 102812, (2021).
- [9] N.-T. Dung, C.-L. Van and Ş. Tǎlu, *Coatings* **11**, 1209 (2021).
- [10] T.-Q. Tuan, C.-L. Van Ş. Tǎlu, N.-T. Dung, *Appl. Sci.* **12**, 946 (2022).
- [11] C.-L. Van, S. Umut, C.-B. Mevlana, D.-T. Luong, Ş. Tǎlu, N.-T. Dung, *Coatings* **12**(3), 346 (2022).
- [12] N.-T. Dung, C.-L. Van, Ş. Tǎlu, *Appl. Sci.* **12**, 4437 (2022).
- [13] N.-T. Dung, C.-L. Van, N.-D. Phu, Ş. Tǎlu, *AIMS Mater. Sci.* **9**, 406 (2022).
- [14] V.-Q. Trung, H.-M. Hung, L.-V. Khoe et al., *ACS Omega* **7**, 19842 (2022).
- [15] N.-T. Dung, C.-L. Van, Ş. Tǎlu, *AIP Adv.* **12**, 065016, (2022).
- [16] N.-T. Dung, C.-L. Van, Ş. Tǎlu, U. Saraç, N.-D. Phu, P.-H. Kien, *J. Compos. Sci.* **6**, 234 (2022).
- [17] N.-T. Dung, C.-L. Van, Ş. Tǎlu, *Appl. Sci.* **12**, 8473 (2022).
- [18] N.-Q. Hoc, N.-T. Dung, N.-C. Cuong, B.-D. Tinh, N.-D. Hien, C.-L. Van, U. Saraç, Ş. Tǎlu, *J. Compos. Sci.* **6**, 250 (2022).
- [19] T.-Q. Tuan, N.-T. Dung, C.-L. Van, U. Saraç, Ş. Tǎlu, *J. Compos. Sci.* **6**, 278 (2022).
- [20] U. Saraç, N.-T. Dung, C.-B. Mevlana, C.-L. Van, Ş. Tǎlu, *Microsc. Res. Tech.* **85**, 3945 (2022).
- [21] N.-T. Dung, C.-L. Van, U. Saraç, D.-Q. Van, Ş. Tǎlu, *J. Compos. Sci.* **6**, 383, (2022).
- [22] N.-T. Dung, C.-L. Van, Ş. Tǎlu, *AIP Adv.* **13**, 015209 (2023).
- [23] L. Dykman, N. Khlebtsov, *Chem. Soc. Rev.* **41**, 2256 (2012).
- [24] U. Bubniene, M. Ocwieja, B. Bugeyte, Z. Adamczyk, M. Nattich-Rak, J. Voronoyic, A. Ramanaviciene, A. Ramanavicius, *Colloid Surf. A* **441**, 204 (2014).
- [25] M. Hu, J. Chen, Z.Y. Li, L. Au, G.V. Hartland, X. Li, M. Marquez, Y. Xia, *Chem. Soc. Rev.* **35**, 1084 (2006).

- [26] A. Makaraviciute, T. Ruzgas, A. Ramanavicius, A. Ramanaviciene, *Anal. Methods* **6**, 2134 (2014).
- [27] Y. Wenjie, L. Huazheng, M. Songhua, W. Devin, H. Jun, *Sustain. Mater. Technol.* **22**, e00109, (2019).
- [28] J.M. Yuk, M. Jeong, S.Y. Kim, H.K. Seo, J. Kim, J.Y. Lee, *Chem. Commun.* **49**, 11479 (2013).
- [29] H. Ohnishi, Y. Kondo, K. Takayanagi, *Nature* **395**, 780 (1998).
- [30] V. Rodrigues, D. Ugarte, *Phys. Rev. B* **63**, 073405 (2001).
- [31] Y. Takai, T. Kadwasaki, Y. Kimura, T. Ikuta, R. Shimizu, *Phys. Rev. Lett.* **87**, 106105, (2001).
- [32] T.-Q. Tuan, N.-T. Dung, *Results Phys.* **15**, 102671 (2019).
- [33] N.-T. Dung, *Mater. Sci. Eng. B* **260**, 114648 (2020).
- [34] J. L. Lewis, P. Jensen, J.-L. Barrat, *Phys. Rev. B* **56**, 2248, (1997).
- [35] Y. Li, H. Liang, X. Zhang, *Mater. Trans.* **59**, 172 (2018).
- [36] K. Weller, N. Zotov, Z.-M. Wang, L.-P.H. Jeurgens, E.J. Mittemeijer, *Non-Cryst. Solids* **427**, 104 (2015).
- [37] J. Fortner, J.S. Lannin, *Phys. Rev. B* **39**, 5527 (1989).
- [38] M.A. Tschopp, D.L. McDowell, *J. Mech. Phys. Solids* **56**, 1806 (2008).
- [39] H.-H. Kart, H. Yildirim, S.O. Kart, T. Çağ, *Mater. Chem. Phys.* **147**, 204 (2014).
- [40] F. Shimizu, S. Ogata, J. Li, *Mater. Trans.* **48**, 2923 (2007).
- [41] V. Borovikov, M.-I. Mendeleev, A.-H. King, R. LeSar, *Model. Simul. Mater. Sci. Eng.* **24**, 085017 (2016).
- [42] C. Untiedt, A. I. Yanson, R. Grande, G. Rubio-Bollinger, N. Agraït, S. Vieira, J.M. van Ruitenbeek, *Phys. Rev. B* **66**, 085418, (2002).
- [43] M.-I. Mendeleev, M.-J. Kramer, C.-A. Becker, M. Asta, *Philos. Mag.* **88**, 1723 (2008).
- [44] D.-V. Louzguine-Luzgin, A.-I. Bazlov, *Metals* **10**, 1532, (2020).
- [45] S. Plimpton, *J. Comput. Phys.* **117**, 1, (1995).
- [46] K. Zhou, T. Zhang, B. Liu, Y. Yao, *Comput. Mater. Sci.* **142**, 389 (2018).
- [47] D.-E. Spearot, M.-A. Tschopp, K.-I. Jacob, D.-L. McDowell, *Acta Mater.* **55**, 705 (2007).
- [48] M.-A. Tschopp, D.-E. Spearot, D.-L. McDowell, *Model. Simul. Mater. Sci. Eng.* **15**, 693, (2007).
- [49] S.-M. Foiles, M.-I. Baskes, M.-S. Daw, *Phys. Rev. B* **33**, 7983 (1986).
- [50] C. Song, T. Lin, P. He, Z. Jiao, J. Tao, Y. Ji, *Comput. Mater. Sci.* **83**, 35 (2014).
- [51] E.-J. Reed, L.-E. Fried, J. Joannopoulos, *Phys. Rev. Lett.* **90**, 235503, (2003).
- [52] T. Lu, G.J. Niu, Y. Xu, J. Wang, Z. An, H. Liu, H. Zhou, F. Ding, G.N. Luo, X.C. Li, *Fusion Eng. Des.* **113**, 340 (2016).
- [53] S. Nosé, *J. Chem. Phys.* **81**, 511 (1984).
- [54] W.G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).
- [55] H. Tsuzuki, P.S. Branicio, J.P. Rino, *Comput. Phys. Commun.* **177**, 518 (2007).
- [56] A. Stukowski, *Model. Simul. Mater. Sci. Eng.* **20**, 045021 (2012).
- [57] S. Ishraq, E.M. Ronald, *Acta Mater.* **57**, 4364 (2009).
- [58] B.J. Lee, J.C. Lee, Y.C. Kim, S.H. Lee, *Met. Mater. Int.* **10**, 467 (2004).
- [59] A. Stukowski, K. Albe, *Simul. Mater. Sci. Eng.* **18**, 085001 (2010).
- [60] N. Anupam, M. Nilanjan, *Sci. Rep.* **7**, 7337 (2017).
- [61] V.D. Aleksandrov, V.N. Aleksandrova, A.A. Barannikov, N.V. Dobritsa, N.E. Malinovskaya, S.A. Frolova, *Tech. Phys. Lett.* **27**, 258 (2001).
- [62] J.M. Yuk, M. Jeong, S.Y. Kim, H.K. Seo, J. Kim, J.Y. Lee, *Chem. Commun.* **49**, 3961 (2013).

SUBSCRIPTION TERMS/WARUNKI PRENUMERATY — 2023

Acta Physica Polonica is issued as two separate journals, namely *A* and *B*, in Warszawa and Kraków, respectively. Each of the periodicals is published in the indicated scope and sold by its respective Publisher (see details below). A subscription order should be submitted to the relevant Publisher.

- ◇ The subscription price of *Acta Physica Polonica A (APPA)* is **450 EUR** (surface mail delivery, priority mail on demand with extra **90 EUR** postage rate).

The Publisher of *Acta Physica Polonica A*:

The Director of Institute of Physics
Polish Academy of Sciences
al. Lotników 32/46
02-668 Warsaw, Poland
e-mail: appol@ifpan.edu.pl

bank account BGK S.A. Warszawa:
PL 35 1130 1017 0013 4373 9820 0027
(int. bank code SWIFT: GOSKPLPW)

The archival *APPA* issues (if still available) can be purchased at the Editorial Board of *APPA*.

- ◇ Cena prenumeraty rocznej *Acta Physica Polonica A* wynosi **2100 PLN** (netto).

Prenumeratę przyjmuje Wydawca:

Dyrektor Instytutu Fizyki
Polskiej Akademii Nauk
al. Lotników 32/46
02-668 Warszawa, Polska
e-mail: appol@ifpan.edu.pl

konto bankowe BGK S.A. Warszawa:
89 1130 1017 0013 4373 9820 0025

- ◇ The subscription price of *Acta Physica Polonica B* can be obtained by contacting its Publisher:

Institute of Physics
Jagiellonian University
Łojasiewicza 11
30-348 Kraków, Poland
e-mail: acta.phys.pol.b@uj.edu.pl

Special issue in honor of scientific achievements of
Profesor Iwo Białynicki-Birula on his 90th birthday

J.H. Eberly <i>Thinking BIG</i>	S9	F. Gleisberg, W.P. Schleich <i>Factorization with a Logarithmic Energy Spectrum of a Central Potential</i>	S112
S. Evans, J. Rafelski <i>Improving Euler–Heisenberg–Schwinger Effective Action with Dressed Photons</i>	S13	J. Kijowski <i>Arrival Time in Quantum Mechanics (Demonstrated in Geometrical Order)</i>	S123
A. Bechler, F. Cajiao Vélez, K. Krajewska, J.Z. Kamiński <i>Vortex Structures and Momentum Sharing in Dynamic Sauter–Schwinger Process</i>	S18	F. Gampel, M. Gajda <i>On Repeated Measurements of a Quantum Particle in a Harmonic Potential</i>	S131
M.G. Raymer, P. Polakos <i>States, Modes, Fields, and Photons in Quantum Optics</i>	S28	J. Mostowski, J. Pietraszewicz <i>Electron Localization in Rydberg States</i>	S140
P. Stammer, M. Lewenstein <i>Quantum Optics as Applied Quantum Electrodynamics is Back in Town</i>	S42	D.K. Wójcik, K. Życzkowski <i>Fractality of Certain Quantum States</i>	S146
A. Friedrich, D. Moll, M. Freyberger, L. Plimak, W.P. Schleich <i>The Wave Functional of the Vacuum in a Resonator</i>	S52	M. Czachor <i>Contra Bellum: Bell’s Theorem as a Confusion of Languages</i>	S158
P.W. Milonni <i>Free Energy of Coupled Oscillators: Lamb Shifts and van der Waals Interactions</i>	S78	M.B. Kruk, T. Vibel, J. Arlt, P. Kulik, K. Pawłowski, K. Rzażewski <i>Fock State Sampling Method — Characteristic Temperature of Maximal Fluctuations for Interacting Bosons in Box Potentials</i>	S171
W. Price, M. Formanek, J. Rafelski <i>Born–Infeld Nonlinear Electromagnetism in Relativistic Heavy Ion Collisions</i>	S87	J. Zakrzewski <i>Dynamical Quantum Phase Transitions from Quantum Optics Perspective</i>	S179
T. Linowski, Ł. Rudnicki <i>Classicality of the Bogoliubov Transformations and the Dynamical Casimir Effect Through the Reduced State of the Field</i>	S95	Ł.A. Turski <i>Vlasov–Klimontovich Equation in Action</i>	S183
J.A. Przeszowski <i>Smearred Field Description of Free Electromagnetic Field</i>	S107	Van Cao Long, Dung Nguyen Trong <i>Molecular Dynamics Approach to Processes in Bulk Au Materials</i>	S191