Lecture Notes in Physics

The Lecture Notes in Physics

The series Lecture Notes in Physics (LNP), founded in 1969, reports new developments in physics research and teaching – quickly and informally, but with a high quality and the explicit aim to summarize and communicate current knowledge in an accessible way. Books published in this series are conceived as bridging material between advanced graduate textbooks and the forefront of research to serve the following purposes:

• to be a compact and modern up-to-date source of reference on a well-defined topic;
• to serve as an accessible introduction to the field to postgraduate students and nonspecialist researchers from related areas;
• to be a source of advanced teaching material for specialized seminars, courses and schools.

Both monographs and multi-author volumes will be considered for publication. Edited volumes should, however, consist of a very limited number of contributions only. Proceedings will not be considered for LNP.

Volumes published in LNP are disseminated both in print and in electronic formats, the electronic archive is available at springerlink.com. The series content is indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks, and consortia.

Proposals should be sent to a member of the Editorial Board, or directly to the managing editor at Springer:

Dr. Christian Caron
Springer Heidelberg
Physics Editorial Department I
Tiergartenstrasse 17
69121 Heidelberg/Germany
christian.caron@springer.com

William G. Unruh Ralf Schützhold (Eds.)

Quantum Analogues: From Phase Transitions to Black Holes and Cosmology

Springer

Editors

William G. Unruh
Department of Physics & Astronomy
University of British Columbia
6224 Agricultural Road
Vancouver, B.C. V6T 1Z1, Canada
E-mail: unruh@physics.ubc.ca

Ralf Schützhold
Institut für Theoretische Physik
Technische Universität Dresden
01062 Dresden, Germany
E-mail: schuetz@theory.phy.tu-dresden.de

W.G. Unruh and R. Schützhold, *Quantum Analogues: From Phase Transitions to Black Holes and Cosmology*, Lect. Notes Phys. 718 (Springer, Berlin Heidelberg 2007), DOI 10.1007/b11804185

Library of Congress Control Number: 2007921534

ISSN 0075-8450
ISBN-10 3-540-70858-8 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-70858-2 Springer Berlin Heidelberg New York

Typesetting: by the authors and techbooks using a Springer LaTeX macro package
Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper SPIN: 11804185 54/techbooks 5 4 3 2 1 0

Acknowledgements

The present book contains a series of selected lectures[1] from the international workshop[2] on "Quantum Simulations via Analogues" which took place at the Max Planck Institute for the Physics of Complex Systems in Dresden (Germany) from July 25th till 28th in 2005. Financial support from the programme "Cosmology in the Laboratory" (COSLAB) of the European Science Foundation (ESF) as well as from the Max Planck Institute for the Physics of Complex Systems is gratefully acknowledged. Furthermore, we would like to thank Mandy Lochar (from the Max Planck Institute for the Physics of Complex Systems) as well as Sarah Mostame, Friedemann Queisser, Dr. Gernot Schaller, Markus Tiersch, Michael Uhlmann, and Dr. Yan Xu (all from the Emmy-Noether Research Group at the Institute for Theoretical Physics at the Dresden University of Technology, Germany) for their valuable help regarding the organisation of this workshop. The Emmy-Noether Research Group has been supported by the the German Research Foundation (DFG). Finally, we would like to thank the authors of the chapters for their efforts and – last but not least – Friedemann Queisser and Markus Tiersch for combining the various contributions into a book.

Vancouver, Dresden
January 2007

Ralf Schützhold
William G. Unruh

[1] The lecture given by J. Yngvason is not included here since it has already been published elsewhere and can be found in Lect. Notes Phys. **690**, 199–215 (2006).
[2] `http://www.mpipks-dresden.mpg.de/~quasim05/`

Contents

List of Contributors

Uwe R. Fischer
Eberhard-Karls-Universität
Tübingen
Institut für Theoretische Physik
Auf der Morgenstelle 14
D-72076 Tübingen, Germany
uwe.fischer@uni-tuebingen.de

Bernd Kuckert
II. Institut für
Theoretische Physik
Luruper Chaussee 149
22761 Hamburg, Germany
kuckert@mail.desy.de

U. Leonhardt
School of Physics and
Astronomy
University of St. Andrews
North Haugh
St. Andrews Fife KY16 9SS
Scotland
ulf@st-andrews.ac.uk

Stefano Liberati
International School for
Advanced Studies
Via Beirut 2-4
34014 Trieste
Italy
liberati@sissa.it

T.G. Philbin
School of Physics and
Astronomy
University of St. Andrews
North Haugh
St. Andrews Fife KY16 9SS
Scotland
tgp3@st-andrews.ac.uk

Mairi Sakellariadou
Department of Physics
King's College London
University of London
Strand WC2R 2LS
United Kingdom
Mairi.Sakellariadou@kcl.ac.uk

Ralf Schützhold
Institut für Theoretische Physik
Technische Universität Dresden
01062 Dresden, Germany
schuetz@theory.phy.tu-dresden.de

Peter Skyba
Centre of Low Temperature Physics
Institute of Experimental Physics
Slovak Academy of Sciences
Watsonova 47
04001 Košice, Slovakia
skyba@saske.sk

William G. Unruh
Department of Physics & Astronomy
University of British Columbia
6224 Agricultural Road
Vancouver, B.C. V6T 1Z1, Canada
unruh@physics.ubc.ca

Matt Visser
Victoria University
School of Mathematics, Statistics
and Computer Science
PO Box 600
Wellington, New Zealand
matt.visser@mcs.vuw.ac.nz

Giuseppe Vitiello
Dipartimento di Fisica
"E.R. Caianiello"
and
INFN Universitá di Salerno
84100 Salerno
Italia
vitiello@sa.infn.it

G. E. Volovik
Low Temperature Laboratory
Helsinki University of Technology
P.O. Box 2200
02015 HUT
Espoo, Finland
and
Landau Institute for
Theoretical Physics
Kosygina 2
119334 Moscow, Russia
volovik@boojum.hut.fi

Silke Weinfurtner
Victoria University
School of Mathematics, Statistics
and Computer Science
PO Box 600
Wellington, New Zealand
silke.weinfurtner@mcs.vuw.ac.nz

The Analogue Between Rimfall and Black Holes

W. G. Unruh

CIAR Cosmology and Gravity Program, Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, B.C. V6T 1Z1, Canada
unruh@physics.ubc.ca

Deep beneath the great encircling seas of the Discworld[1] lived a species of hyper-intelligent fish. As is known the Discworld is ruled by Magic, but the salt in the seas shorted out that magic. A small amount dissolved in the waters, and resulted in the development of the high intelligence in those fish. However, cut off from the Magic which bathed the rest of the world, their view of the world was a materialist one, one in which they tried to discover the laws of matter which governed their universe.

The deeps being dark, they experienced the world through sound, and over the centuries have discovered that the outer boundary of their world was a strange place. Known as Rimfall to those above, this was where the waters of the Discworld plunged over the edge to bathe the shell of the Great Atuin, on whose back the Discworld travelled through space. But to the fish, this Rimfall was a boundary, a horizon beyond which nothing could be heard. No fish who had ever travelled over the Rimfall had ever reported back. The shouts, or were they screams, of those intrepid explorers who had travelled too near to that boundary had suffered the most strange bass shifting, the high pitched scream rapidly descending the scales to disappear from sound. Some claimed that with the most careful measurements of the sound, one could still hear sounds, of lower and lower frequency arbitrarily far into the future, as though the sound from those explorers never ceased. However, in no case could sounds ever be heard from beyond the location of that horizon, as that peculiar surface in the Rimfall came to be known. In later years when the fish

[1] Discworld is the location of the hugely popular Discworld series of novels by Terry Pratchett (see especially "The Colour of Magic" which describes Rimfall, but obviously not this fish world view). The Discworld is a disc with continents surrounded by a great surrounding sea which falls off the edge of the disk in a waterfall called the Rimfall. ("Arrangements are made" to return the water to the sea). In Discworld, magic, rather than physics, rules and light behaves very strangely. This means that in the aboveworld physics plays only a very minor role, and is unknown except by a few insane individuals at the Unseen University.

W.G. Unruh: *The Analogue Between Rimfall and Black Holes*, Lect. Notes Phys. **718**, 1–4 (2007)
DOI 10.1007/3-540-70859-6_1

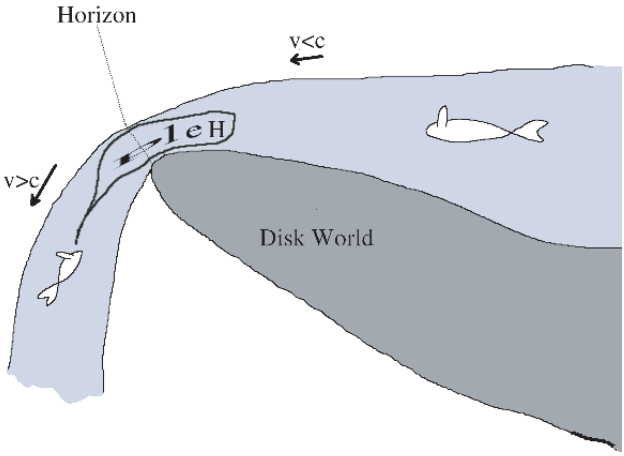

Fig. 1. Two fish from the Underwater University on Discworld experimenting at the Rimfall. The graduate student has fallen over the horizon, where the velocity of the water exceeds the velocity of sound in the water. Realising what was happening he calls out Help, but with the "p" emitted just as he crosses the horizon. The sound is stretched out, with part slowly escaping (but bass shifted) and part falling over the horizon with the student. Exactly the same happens for someone falling into a black hole, where it is light rather than sound which is trapped, and where the light emitted just before horizon crossing is exponentially red shifted as seen by the external observer

began to communicate with the world above, and in particular with visitors from far distant worlds (since their worldview differences with the aboveworld inhabitants of Discworld made communication essentially impossible), they immediately recognised that, what those visitors called black holes, were in many ways analogues in light of their experiences with Rimfall.

One of the fish in particular, playing with that strange new theory called Quantum Mechanics realised that this Rimfall horizon was not simply a one way street. Instead the theory predicted that even if the water flowing over the Rimfall was absolutely cold, that horizon would produce thermal sonic radiation, radiation with a temperature far far below any they had ever experienced, but with a non-zero temperature non-the-less. Again, in the conversations with those visitors from distant planets, they discovered that the same effect for black holes had been found by a physicist (as those materialist philosophers on the distant planets were called) and was known as the Hawking Effect. Many of the fish began to see that they could use these phenomena associated with light and gravity, as an analogue to the effect they had discovered. In fact the equations of the motion of light in the gravitation fields of the black holes were exactly the same, in both the classical and the quantum regime, as the equations which governed the sound waves travelling near the horizon of Rimfall.

Of course those visitors saw things differently. For them it was the motion of the sound waves in the fluid flowing over the waterfall which were the analogue to their equations of motion of light or of other fields near the their black holes. The bass shift of the screams of the fish falling over Rimfall were the analogue of the redshift of light for someone falling into a black hole, where again, the light would forever redshift and one would never be able to see beyond the horizon as someone or something fell into the black hole.

I first used this analogy to try to explain classical black holes to a colloquium at Oxford University in 1972. The experiences of those fish near a waterfall in which the velocity of the water became larger than the velocity of sound in the water at some point were analogous to many of the experiences of explorers near the horizon of a black hole. In 1981 [1] I realised that the equations of motion of the sound waves were identical to those near a black hole, and that the quantisation of those sound waves would produce thermal radiation in the same manner as Hawking had predicted for black holes. Since that time many other systems have also been shown to behave in the same way.

This volume represents a wide variety of views of that analogy between condensed matter systems and black hole and cosmological physics. In many cases they represent the kind of physics which those fish would have developed on the Discworld, and explicate those analogies between the gravitation physics of the visitors, and the physics of their sound world developed by the fish. Not only were the sound waves an analogy, even the ripples on the surface of the ocean also acted as an analogue to the behaviour of fields near a black hole horizon. In fact, it soon seemed like a contest to see in how many disparate physical systems one could find such analogues. This was in part a revelling in the exuberance which nature showed in the diversity with which it revealed the same physics in different contexts. Physics seems to manifest a striking unity within the diversity of physical world. However, also driving the subject was an attempt to find experimental situations in which these ideas (which were after all theoretical ideas) could be tested. While no-one believed that hearing the thermal noise from a cold fluid flowing past a sonic horizon would prove that the thermal emission from a black hole must exist, such successful experiments would greatly increase the confidence in the approximation which were being made in both the gravitational and the analogue situations. "Are quantum gravity (Planck scale) effects important in understanding either black hole evaporation or cosmological particle creation?" has its analogy in "Do atomic processes have an effect on the thermal creation of sound near a "dumb-hole' horizon?". Certainly the suggestions from the sonic case are that Planckian physics is irrelevant to black hole evaporation, and that the radiation emitted by a black hole is due to low energy processes, processes on the length scale set by the black hole, not by quantum gravity.

The experiments, if they are ever done will be hard, very hard. There is a competition between easily making a horizon, which in general requires relatively low values for the "sound waves", and obtaining sufficient energy

in those sound waves to be detectable (for example for Bose-Einstein Condensates, in which the velocity of sound is millimetres per second, making a horizon is relatively easy, but the energy emitted by a horizon correspond to only a few KHz phonons emitted per second, because of the very low velocity of the sound in such condensates).

A third area of importance of these studies is the cross fertilisation that they offer. The questions which one asks, for example of a Bose-Einstein Condensate, when one is looking at a system as an analogue of another system tend to be different from those asked by the practitioners in the field. The study of such analogies thus leads to insights both in the original area and in the area of the analogy.

Enjoy this volume, and revel in the insights into the unity of the universe which it has to offer.

Reference

1. W. G. Unruh, Phys. Rev. Lett. **46**, 1351 (1981)

Effective Horizons in the Laboratory

R. Schützhold

Institut für Theoretische Physik, Technische Universität Dresden, Germany

Abstract. The concept of a horizon known from general relativity describes the loss of causal connection and can be applied to non-gravitational scenarios such as out-of-equilibrium condensed-matter systems in the laboratory. This analogy facilitates the identification and possibly the experimental verification of exotic effects known from gravity and cosmology, such as Hawking radiation, as well as a unified description and better understanding of non-equilibrium phenomena in condensed matter systems. By means of several examples including general fluid flows, dynamical quantum phase transitions, and expanding Bose–Einstein condensates, the concepts of event, particle, and apparent horizons will be discussed together with the resulting quantum effects.

1 Introduction

1.1 Preliminaries

For many condensed-matter systems, the equilibrium state and small fluctuations around this state are quite well understood. Strictly speaking, the equilibrium state – which could be a thermal state or the ground state of the system – may describe static systems only. In the presence of an external or internal time-dependence, the situation becomes more complicated. As long as this time-dependence is sufficiently slow, the adiabatic theorem states that the actual state of the system stays very near its equilibrium state. For fast dynamics, on the other hand, this adiabaticity assumptions fails and non-equilibrium phenomena become important. For such strongly time-dependent systems, the usual split into the ground state (assuming zero temperature) and small quasi-particle excitations is no longer unique nor well-defined. One consequence of this non-uniqueness is the creation of quasi-particles, i.e., the initial ground state may evolve into an excited state during the non-adiabatic evolution.

In searching for a unified description of these non-equilibrium phenomena, one realizes that the same situation occurs for quantum fields in curved space-

R. Schützhold: *Effective Horizons in the Laboratory*, Lect. Notes Phys. **718**, 5–30 (2007)
DOI 10.1007/3-540-70859-6_2

times, see, e.g., [1]. Indeed, it turns out that many of the concepts developed in this field can be applied to condensed matter systems too. One can have the analogue of horizons in laboratory systems, which imply the same effects as in curved space-times, such as the amplification of quantum fluctuations leading to the creation of quasi-particles. Studying the analogy between gravitational and condensed-matter systems leads to a better understanding on both sides [2].

1.2 The Acoustic Analogy

A quarter of a century ago, Bill Unruh (see previous chapter and [3]) suggested an intriguing analogy between (quantised) sound waves in irrotational and inviscid fluids and quantum fields in curved space-times. His original motivation was to study analogous systems which reproduce major features of black holes and whose underlying physics is (at least in principle) understood – the black hole analogues [3]. Without rotation $\boldsymbol{\nabla} \times \boldsymbol{v} = 0$ and viscosity, the propagation of phonons is described by the wave equation

$$\left(\frac{\partial}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{v}_0\right) \frac{\varrho_0}{c_\mathrm{s}^2} \left(\frac{\partial}{\partial t} + \boldsymbol{v}_0 \cdot \boldsymbol{\nabla}\right) \phi = \boldsymbol{\nabla} \cdot (\varrho_0 \boldsymbol{\nabla} \phi) \,, \tag{1}$$

with $\boldsymbol{v}_0$ denoting the (background) velocity of the fluid, ϱ_0 its density, and c_s the speed of sound. For an irrotational flow, the velocity perturbations associated with the sound waves can be described by a potential $\delta \boldsymbol{v} = \boldsymbol{\nabla} \phi$. This complicated wave equation has exactly the same form as that for a scalar field in a curved space-time

$$\Box_\mathrm{eff} \phi = \frac{1}{\sqrt{-g_\mathrm{eff}}} \partial_\mu \left(\sqrt{-g_\mathrm{eff}}\, g_\mathrm{eff}^{\mu\nu} \partial_\nu \phi\right) = 0 \,, \tag{2}$$

provided that the effective geometry of that curved space-time is described by the Painlevé-Gullstrand-Lemaître metric [4]

$$g_\mathrm{eff}^{\mu\nu} = \frac{1}{\varrho_0 c_\mathrm{s}} \begin{pmatrix} 1 & \boldsymbol{v}_0 \\ \boldsymbol{v}_0 & \boldsymbol{v}_0 \otimes \boldsymbol{v}_0 - c_\mathrm{s}^2 \mathbf{1} \end{pmatrix} \,, \tag{3}$$

according to the background flow profile.

As a consequence, phonons in a flowing (irrotational and inviscid) fluid are completely equivalent to a (quantum) field in a curved space-time described by the above metric. Hence we may apply all the concepts known from general relativity – such as horizons [1,5,6]. However, it is worth noting that the above analogy applies to the *kinematics*, i.e., the phonon propagation – but not to the *dynamics*, i.e., the evolution of the background is different (Euler versus Einstein equations).

1.3 Generalisations

Apart from phonons in fluids such as Bose–Einstein condensates or super-fluid Helium, see, e.g., [7–9], one can introduce an effective metric for other scenarios as well: If the quasi-particle excitations possess just one non-degenerate mode, one can describe them by a single scalar field ϕ. In the case of degeneracy, this is not possible; but if we assume that the multiple degenerate modes are independent, we may pick out one mode and describe it again by a scalar field. Furthermore, if the quasi-particle excitations (and their quantum fluctuations) are small enough, we may linearise the equations of motion and hence the most general low-energy effective action for these modes ϕ reads [9, 10]

$$\mathcal{L}_{\text{eff}} = \frac{1}{2}(\partial_\mu\phi)(\partial_\nu\phi)G^{\mu\nu} + V\phi^2 + \mathcal{O}(\phi^3) + \mathcal{O}(\partial^3) \; . \tag{4}$$

The tensor $G^{\mu\nu}$ describes the background and is related to the effective metric. For Goldstone modes (such as phonons in Bose–Einstein condensates), the potential V vanishes due to $\omega(k = 0) = 0$, but in general it can be an arbitrarily space-time dependent function. The remaining higher-order terms $\mathcal{O}(\phi^3)$ and $\mathcal{O}(\partial^3)$ reflect the linearisation of the equation of motion and the low-energy expansion (into powers of ω and k).

One example for the above action are surface waves (ripplons) on thin fluid films [11] or photons in particularly designed wave-guides [12]. An example for non-scalar quasi-particles are photons in dielectric media [13]. Since the two polarisations are not independent in general media, an effective metric – which must be the same for both polarisations – can only be introduced in the presence of a strong symmetry between the two modes. For dielectric media with a constant dielectric permittivity ε but an arbitrary four-velocity u^μ, this is the case and one obtains the Gordon metric [14]

$$g^{\mu\nu}_{\text{eff}} = g^{\mu\nu}_{\text{Minkowski}} + (\varepsilon - 1)\, u^\mu\, u^\nu \;\; , \tag{5}$$

for the propagation of the spin-one photon field within that dielectric medium. This metric is slightly different from the sonic case in Eq. (3), but it may also contain a horizon with the Hawking temperature being given by the same expression as in Eq. (8) below plus relativistic corrections [13]. In addition to spin-zero (phonons in fluids) and spin-one (photons in dielectric media) fields, it is also possible to effectively model a spin-1/2 Dirac field in highly dispersive quantum optical media, which support the phenomenon of slow light [15].

1.4 Geometric Concepts

As a result, the propagation of quasi-particle excitations in a rather large class of systems displays universal behaviour in being completely equivalent to a curved space-time described by the corresponding effective metric. As far as purely kinematic aspects are concerned, we may forget the structure of the

underlying system and just work with the effective metric, which allows us to apply the geometric concepts known from general relativity. For example, we may introduce sound cones in complete analogy to light cones as null rays of the effective line element

$$ds^2_{\rm eff} = g^{\rm eff}_{\mu\nu}\, dx^\mu\, dx^\nu = 0\ . \tag{6}$$

The class of sound cones determines which points can send/receive sound waves (or other quasi-particles under consideration) to/from which other points and hence induce the causal connection. Roughly speaking, if this causal connection is lost, we have the analogue of a horizon – which clearly indicates the breakdown of adiabaticity (non-equilibrium). In order to be in equilibrium, every point must be able to exchange energy or momentum or, more generally, information with any other point. In the presence of a horizon, the quantum dynamics is far from equilibrium, which leads to effects such as the amplification of the quantum fluctuations and the creation of quasi-particles.

2 Event Horizon

2.1 Black Hole Thermodynamics

Black holes are supposed to be the final states of massive stellar objects whose internal pressure cannot compensate the gravitational attraction anymore [5, 6]. Neglecting exotic phenomena such as magnetic monopoles, these final states are largely independent of the initial object ("no-hair theorem") and can be determined in terms of three (conserved) quantities: their mass M, their angular momentum J, and their electric charge Q. (However, since black holes are extremely compact objects and hence tend to dispose of any macroscopic net charge Q very quickly via pair creation, mainly neutral black holes with $Q = 0$ will be considered in the following.)

It was Jacob Bekenstein [16] who first noticed an intriguing analogy between the laws of thermodynamics and the physics of black holes provided that one identifies the corresponding quantities according to the table provided below:

Black holes	Thermodynamics
mass M	energy E
surface area A	entropy S
surface gravity κ	temperature T
angular velocity at the horizon $\Omega_{\rm h}$	intensive quantity, e.g., pressure p
(minus) angular momentum J	extensive quantity, e.g., volume V

Indeed, by using the Einstein equations augmented with suitable energy conditions on the matter fields (such as a non-negative energy density), one can derive the four laws of black hole thermodynamics [5, 6, 16]

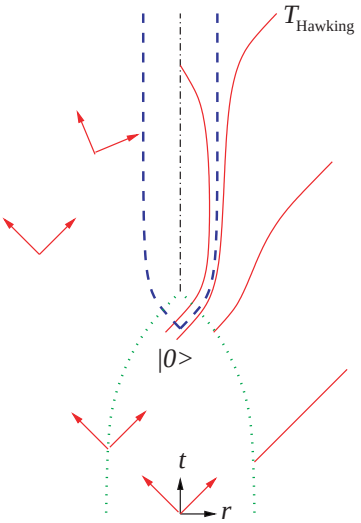

Fig. 1. Space-time diagram of the collapse to a black hole. Time increases from bottom to top and the radial coordinate from the middle to the side, where the angular dependence has been dropped assuming spherical symmetry. The *dotted green line* denotes the surface of the collapsing star and the *black dashed-dotted* line is the singularity. The *solid red lines/arrows* are light rays/cones, which are bent by the gravitational field. The *dashed blue line* is the horizon, i.e., the boundary between the last ray which is able to escape to infinity and the first ray which is trapped and hits the singularity. These two light rays are very close initially, but their fate is totally different. This shows the strong distortion of the field modes, which is responsible for the Hawking effect in the original derivation: The quantum fluctuations of the initial vacuum state $|0\rangle$ are strongly distorted such that they transform into a final state containing particles (with a thermal distribution according to the Hawking temperature T_{Hawking})

- 0^{th} law: the surface gravity κ is constant (across the horizon) in equilibrium
- 1^{st} law: the energy conservation condition reads $dM = \kappa\, dA/(8\pi) + \Omega_{\text{h}}\, dJ$
- 2^{nd} law: the total surface area A of the horizon always increases
- 3^{rd} law (weak version): one cannot reach $\kappa = 0$
  (which is only relevant for $Q \neq 0$)

2.2 Hawking Effect and Trans-Planckian Problem

However, taking this analogy seriously did impose the problem that any body with a finite temperature (and absorption cross section) should radiate – whereas nothing was supposed to be able to escape a black hole. Therefore, a major breakthrough was achieved by Stephen Hawkings seminal discovery [17] that black holes indeed emit thermal radiation due to quantum effects with the temperature being consistent with Bekensteins analogy (cf. Fig. 1)

$$T_{\rm Hawking} = \frac{1}{8\pi\, M} \frac{\hbar\, c^3}{G_{\rm N} k_{\rm B}} \propto \kappa \; . \tag{7}$$

Without taking into account the quantum effects such as Hawking radiation, there are several *Gedanken* experiments suggesting that it should be possible to violate the second law of thermodynamics – but including these effects, the analogy between black hole physics and thermodynamics provides such a consistent picture that many physicists consider this analogy as a hint of some fundamental principle in physics. For example, the black hole entropy, which is given by the surface area (instead of the volume, for example) of the horizon in Planckian units (constructed out of Newtons gravitational constant $G_{\rm N}$, Planck's constant $\hbar$, and the speed of light c) is thought of as a measure of the number of fundamental degrees of freedom. This observation is one of the main motivations behind fundamental concepts such as the holographic principle etc.

However, there is a serious flaw in this interesting picture: Tracing the particles of the Hawking radiation (e.g., photons) back in time, one has to undo the gravitational red-shift, which is exponentially strong in the vicinity of the horizon (where the escape velocity equals the speed of light, cf. Fig. 1). Therefore, the photons emitted at late times have their origin in modes with extremely short wavenumbers corresponding to very large frequencies/energies. For example, a photon given off by a solar-mass black hole after just one second originates from a mode with a frequency which is bigger than the mass of the black hole (over the speed of light squared) itself! Of course, at these energy/frequency scales, one cannot trust the theory of quantum fields (e.g., quantum electrodynamics) propagating in (classical) curved space-times anymore. For example, effects of quantum gravity are expected to become important at the Planck scale, which is around 10^{19} GeV. I.e., the derivation of the Hawking effect is based on the extrapolation of a theory (quantum fields propagating in classical curved space-times) to an energy region, where this theory is expected to break down (trans-Planckian problem). This raises the question of whether the Hawking effect depends on details of the underlying physics (e.g., quantum gravity) at Planckian energies [3,18,19].

2.3 De Laval Nozzle

In order to tackle the aforementioned trans-Planckian problem, let us consider a toy model for a black hole in the laboratory. At a first glance, a radial converging flow profile seems to be most appropriate since the insertion of $\boldsymbol{v} = -f(r)\boldsymbol{r}$ reproduces the Schwarzschild metric for an appropriate function $f(r)$. However, such a flow profile would be unstable in general [20] and generate shock waves etc. A stable realization would be a de Laval nozzle sketched in Fig. 2 with the horizon being situated at the point where the velocity of the fluid exceeds the speed of sound $v_0 = c_{\rm s}$. As it is intuitively clear, no sound can escape from the region beyond the sonic horizon. This analogy is not

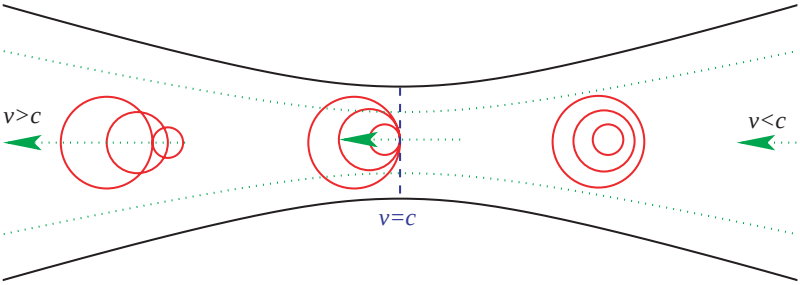

Fig. 2. Cross section of a de Laval nozzle. The *solid black lines* indicate the (frictionless) walls and the *dotted green lines* are stream-lines. The flow enters subsonically on the right-hand-side, exceeds the speed of sound at the narrowest point and exits the nozzle with a supersonic velocity. The *dashed blue boundary* is the acoustic analogue of an event horizon since no sound (*solid red circles*) from beyond this point can escape to the region on the right-hand-side

just a nice intuitive picture – there is an exact correspondence to black holes, which allows us to repeat the steps of Hawkings derivation for that system. In complete analogy to real black holes, one would predict a thermal emission of sound waves corresponding to the temperature

$$T_{\text{Hawking}} = \frac{\hbar}{2\pi\, k_{\text{B}}} \left| \frac{\partial}{\partial r} \left(v_0^{\perp} - c_{\text{s}} \right) \right| , \tag{8}$$

which is basically determined by the velocity gradient. Even though the Hawking temperature of the black hole analogues in the laboratory is typically orders of magnitude larger than for real black holes with a few solar masses (Chandrasekhar limit), it is still comparably small: Depending on the actual realizations, it varies from typical values of order nano-Kelvin for Bose–Einstein condensates [21, 22] over micro-Kelvin for superfluid Helium [23] up to fractions of a Kelvin for electromagnetic waveguides [12].

Of course, this prediction goes along with the same problems as for real black holes, since it uses the above wave equation (1) for extremely short wavelengths (e.g., below the inter-particle distance), where fluid dynamics breaks down. However, for this analogue system, the behaviour at short length scales is (in principle) understood, which allows us to actually check the validity and robustness of the prediction.

2.4 Impact of Dispersion Relation

Having established the analogy between black holes and fluids, we may now retrace the steps of Hawkings original derivation and study the impact of short-scale physics on the predicted radiation (see, e.g., [18, 19]). For many fluids, the first deviations from the macroscopic Euler equation at short distances can be described in terms of a non-trivial dispersion relation $\omega(k)$:

For small k, the dispersion relation in the local rest frame of the fluid goes as $\omega(k \downarrow 0) = c_{\rm s} k$, but for large k, there are deviations from this linear behaviour, i.e., the group and phase velocity change at small distances.

In order to study the impact of the dispersion relation on the Hawking radiation, let us make a few assumptions: Firstly, the surface gravity (and hence the Hawking temperature) should be much smaller than the cut-off frequency (which is the analogue of the Planck energy), where the dispersion relation starts to deviate from the linear behaviour present a low k; otherwise the whole concept of an effective metric breaks down. Secondly, we assume an analytic dispersion relation with a finite radius of convergence around $k = 0$. Finally, the dispersion relation $\omega(k)$ approaches the $\omega = c_{\rm s} k$-line at small k, but otherwise it is supposed to be well separated from the $\omega = c_{\rm s} k$-line and the k-axis. Using analytic continuation techniques, it can be shown that arbitrary dispersion relations $\omega(k)$ satisfying these conditions indeed reproduce the Hawking radiation – provided that the modes start off in their ground state (with respect to the the local rest frame of the fluid) at short distances [18].

Let us consider a few examples: Sound waves in Bose–Einstein condensates (see, e.g., [24]) are described by the Bogoliubov dispersion relation ($\hbar = 1$)

$$\omega^2 = c_{\rm s}^2 k^2 + \frac{k^4}{4m^2} \; , \tag{9}$$

with m being the mass of the particles forming the condensate. Since this relation satisfies all the above conditions, the result of Hawkings original derivation is correct for Bose–Einstein condensates. The origin of the acoustic analogue of Hawking radiation is sketched in Fig. 3 for this situation.

However, if we consider a second example

$$\omega^2 = \omega_0^2 \sin^2(ak) \; , \tag{10}$$

which is the dispersion relation of phonons in a lattice with the lattice spacing a, we see that this relation violates the above assumptions. Indeed, due to the additional zeros at $k \in \pi a^{-1} \mathbb{N}$, it is very easy to excite these high-wavenumber modes which then induces deviations from the thermal spectrum of the Hawking effect [18], see also Fig. 4.

As a final example, let us consider some interactions between the sound waves in the moving fluid and the walls in addition to the dispersion $\omega^2(k) = k^2 v_{\rm phase}^2(k)$ in the local rest frame of the fluid. For the frequency Ω measured in the laboratory frame (i.e., the rest frame of the walls – not of the fluid), this interaction can effectively be described by a damping term $\gamma(k)$

$$(\Omega + v_{\rm fluid} k)^2 = k^2 v_{\rm phase}^2(k) + i \Omega \gamma(k) \; , \tag{11}$$

where $\omega^2 = (\Omega + v_{\rm fluid} k)^2$ is the co-moving frequency measured in the local rest frame of the fluid. For small fluid velocities $v_{\rm fluid}$, the solutions of the above equations for Ω just describe damping of the sound waves $\Im(\Omega) > 0$ (assuming real wavenumbers $k \in \mathbb{R}$); but for $v_{\rm fluid} > v_{\rm phase}$, the imaginary

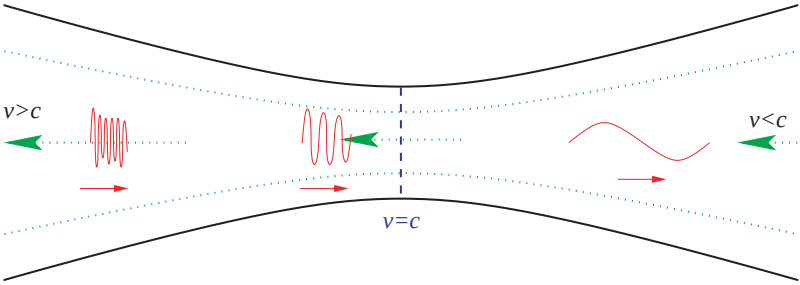

Fig. 3. Origin of Hawking radiation for Bose–Einstein condensates. The initial wave-packet on the left-hand-side has a very high wavenumber and hence a supersonic propagation velocity which enables it to overcome the drag of the flow and to propagate towards the right. During its approach to the horizon, the wave-packet (*in the middle*) is constantly stretched since the flow on its front end is slightly slower than on its back end – in complete analogy to the gravitational red-shift. This distortion squeezes the quantum fluctuations in such a way that the wave-packet, after starting off in its ground state (*left*) is no longer in its ground state after crossing the horizon (*right*) but contains particles in complete agreement with Hawkings prediction. In the presence of an additional white-hole horizon (supersonic → subsonic flow), the situation is more complicated, see chapter by U. Leonhardt and T. G. Philbin in this volume

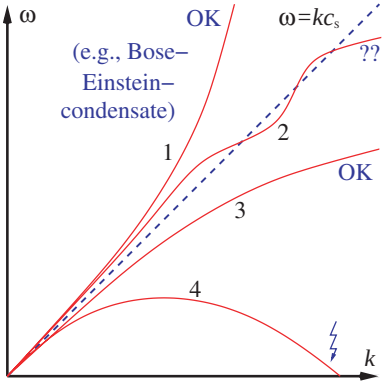

Fig. 4. Different dispersion relations: Curve 1 corresponds to a superluminal/supersonic dispersion (such as in Bose–Einstein condensates), for which Hawkings prediction is correct. For the mixed situation in curve 2, the situation is more complicated, but one would expect that it also reproduces Hawkings result. The sub-luminal/subsonic dispersion 3 also yields the same thermal spectrum, since the curve is well separated from the k-axis. The last curve 4, however, violates that assumption and would result in deviations from Hawkings prediction in general

part of Ω changes sign $\Im(\Omega) < 0$ indicating an instability[1]. Due to this instability, the short-scale modes do not stay in their ground state but get excited – which again destroys the thermal spectrum of the Hawking radiation. Translating these findings back to real black holes, the local rest frame of the fluid corresponds to the (local) freely falling frame, whereas the laboratory frame (i.e., the rest frame of the walls) is analogous to the global rest frame of the black hole. Roughly speaking, if the Planckian degrees of freedom generate interactions with respect to the global rest frame of the black hole instead of the (local) freely falling frame, one may well obtain instabilities which generate deviations from Hawkings result, for example.

In summary, one reproduces the Hawking effect for a rather large class of scenarios, but there are also counter-examples, which do not appear to be unphysical or artificial, displaying strong deviations from Hawkings result. Therefore, whether real black holes emit Hawking radiation remains an open question and could give non-trivial information about Planckian physics (e.g., whether the "space-time foam" is freely falling or at rest).

3 Cosmic Horizons

Apart from black hole horizons, the aforementioned analogy between quantum fields in curved space-times and sound waves in fluids can also be applied to cosmic horizons (occurring in a rapidly expanding or contracting universe, for example [1, 5, 6]). Even though the two scenarios seem very different at a first glance, there are fundamental similarities: The horizon causes a strong distortion of the initial quantum vacuum fluctuations and so generates potentially observable effects. Let us demonstrate such an effect for the example of cosmic inflation, which (according to our standard model of cosmology) is a very early epoch in the evolution of our universe governed by an accelerated expansion as described by the de Sitter metric ($c = 1$)

$$ds^2 = d\tau^2 - e^{2H\tau} d\boldsymbol{r}^2 \;, \tag{12}$$

where H is the Hubble constant. (We have used the Friedmann-Robertson-Walker representation.) During inflation, the most important field is supposed to be the inflaton, which is a scalar quantum field satisfying the approximate equation of motion ($c = 1$ if not otherwise indicated)

$$\left(\frac{\partial^2}{\partial \tau^2} + 3H \, \frac{\partial}{\partial \tau} - e^{-2H\tau} \boldsymbol{\nabla}^2 \right) \phi = 0 \;. \tag{13}$$

After a spatial expansion into plane waves, every $\boldsymbol{k}$-mode corresponds to a damped harmonic oscillator with a continuously decreasing potential $e^{-2H\tau} \boldsymbol{k}^2$. Consequently, the temporal evolution can roughly be split up into three stages

[1] This effect is known as the Miles instability [25], which is responsible for the generation of water waves by wind blowing of its surface.

- Oscillation $e^{-2H\tau}\boldsymbol{k}^2 \gg H^2$: Initially, the damping term $3H\partial/\partial\tau$ can be neglected with respect to the potential term and the does oscillate almost freely.
- Horizon-crossing $e^{-2H\tau}\boldsymbol{k}^2 \approx H^2$: After some period of time depending on the wavenumber $\boldsymbol{k}$, the continuously decreasing potential term becomes small enough for the damping to set in.
- Freezing $e^{-2H\tau}\boldsymbol{k}^2 \ll H^2$: Finally, the modes are over-damped like a pendulum in an extremely viscous fluid and effectively do not move anymore (i.e., they are frozen).

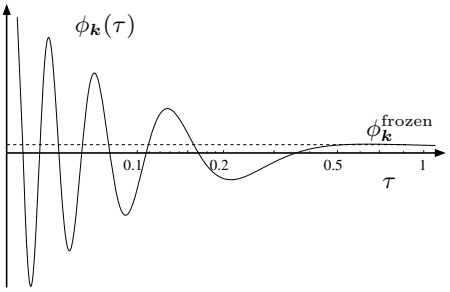

Fig. 5. Temporal evolution of a given mode $\boldsymbol{k}$ over the three stages: oscillation, horizon-crossing, and freezing

The transition from oscillation to freezing (sketched in Fig. 5) can also be understood in terms of the cosmic horizon, which represents the maximum co-moving spatial distance at which events can be in causal contact. In physical units of length, the horizon size is just given by the inverse Hubble constant $1/H$. Modes with a given wavelength λ are being continuously stretched due to the cosmic expansion and after the wavelength λ exceeds the horizon size $1/H$, causal connection is lost and hence the modes cannot oscillate anymore (freezing).

In a slowly expanding universe without horizon, the quantum vacuum fluctuations evolve adiabatically and effectively stay in their ground state – i.e., they continuously decrease. (Roughly speaking, the zero-point energy $\hbar\omega/2$ of a given mode diminishes since the wavelength λ is being continuously stretched due to the cosmic expansion.) In the presence of a cosmic horizon such as during inflation, however, the freezing of the modes prevent this continuous decrease – i.e., the initial quantum vacuum fluctuations are amplified in comparison with the ground state. In our present standard model of cosmology, this amplification mechanism provides the seeds for structure formation (such as our galaxy) since, at the end of inflation, the frozen and amplified quantum fluctuations are converted into temperature and density fluctuations, see Fig. 6.

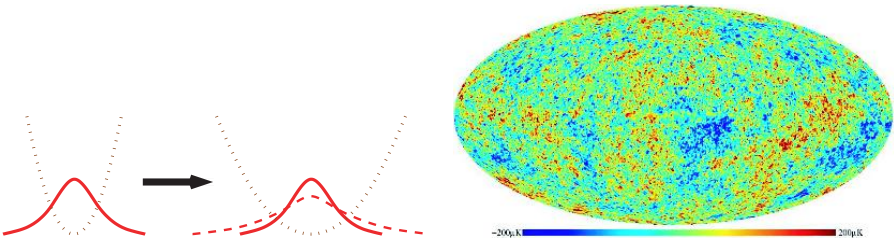

Fig. 6. Sketch of the amplification (*squeezing*) of quantum fluctuations (*left*) and picture of the anisotropies in the cosmic microwave background radiation (*right*) obtained by the recent WMAP mission (see `http://map.gsfc.nasa.gov/`). Every $\boldsymbol{k}$-mode of the $\hat{\phi}$-field corresponds to a harmonic oscillator whose potential (sketched as *brown dotted curve*) decreases/broadens $e^{-2H\tau}\boldsymbol{k}^2$. However, after the mode crossed the horizon and is frozen, the quantum state (*solid red curve*) cannot adapt to this change anymore and hence deviates from the ground state (*dashed red curve*), thereby evolving to a squeezed state. This amplification mechanism is the reason why the initial quantum fluctuations generate rather strong temperature and density fluctuations with a relative size of order 10^{-5}, which can be observed in the anisotropies of the cosmic microwave background radiation (*right*), for example

3.1 Particle Horizon

Now let us investigate how to simulate/recover these quantum effects within an expanding universe in the laboratory: Depending on the properties of the fluid, the effective metric in Eq. (3) is capable of modelling spherically symmetric as well as rotating black holes – but also the space-time of an expanding (or contracting) universe. Recalling the effective line-element in Eq. (3)

$$ds^2_{\rm eff} = \frac{\varrho_0}{c_{\rm s}} \left([c_s^2 - \boldsymbol{v}_0^2] dt^2 + 2\boldsymbol{v}_0 \cdot d\boldsymbol{r} dt - d\boldsymbol{r}^2 \right) \ , \tag{14}$$

there are basically two major possibilities for simulating an expanding universe: firstly, a decreasing sound speed for a fluid at rest, and, secondly, an expansion of the fluid itself.

Let us start with the first possibility: If the speed of sound $c_s(t)$ decreases fast enough such that the integral

$$\Delta r(t) = \int_t^\infty dt' \, c_s(t') \tag{15}$$

converges to a finite value, then a sound wave starting at a time t propagates the maximum distance $\Delta r(t)$. From this time t on, points which are further apart than $\Delta r(t)$ cannot be in causal contact (via sound waves) anymore – i.e., this distance $\Delta r(t)$ is the size of the particle horizon, see Fig. 7.

This loss of causal connection results in many interesting non-equilibrium effects – even on the classical level, where one prominent example is the

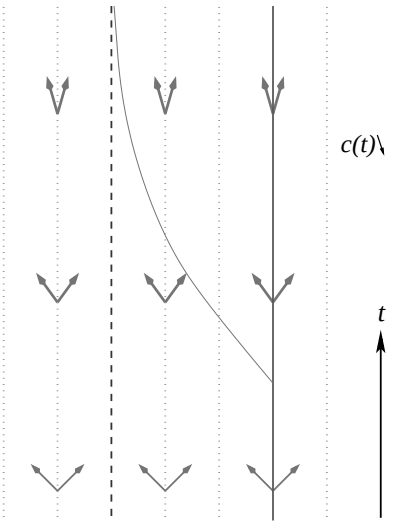

Fig. 7. Sketch of the emergence of a particle horizon. Time goes up in this picture and the speed of sound $c_s(t)$ decreases. The *vertical dotted green lines* denote the fluid at rest and the *solid brown line* is the trajectory of a chosen particle (also at rest, i.e., vertical). The *solid red arrows* are the sound cones which become constantly narrower. Hence a sound wave (*solid red curve*) emitted by the chosen particle (*solid brown line*) at a time t is bent upwards and – if $c_s(t)$ decreases fast enough – cannot propagate further than the *dashed blue line*, which then represents the associated particle horizon

Kibble-Zurek mechanism [26]. Let us consider a thermal phase transition from a symmetric phase to another phase where this symmetry is spontaneously broken. If the phase transition occurs slowly enough, the system will choose the same point in the internal symmetry-breaking manifold everywhere leading to a homogeneous state (equilibrium). However, if this transition takes place in a rapidly expanding universe with a horizon, different spatial positions at distances larger than the horizon will be out of causal contact and hence they will not choose the same point in the internal symmetry-breaking manifold in general. If this internal symmetry-breaking manifold is not simply connected, this random orientation may lead to the generation of topological defects (such as kinks, vortices, or strings, see chapter by M. Sakellariadou in this volume), cf. Fig. 8. As it becomes evident from the above considerations, this mechanism is not restricted to expanding universes but can also occur in the laboratory. In the vicinity of the transition, the response time of the system typically diverges and thus the propagation speed of the excitations vanishes. Therefore, one obtains an effective horizon which generates the same effect in a suitable transition (e.g., quench) in the laboratory [26].

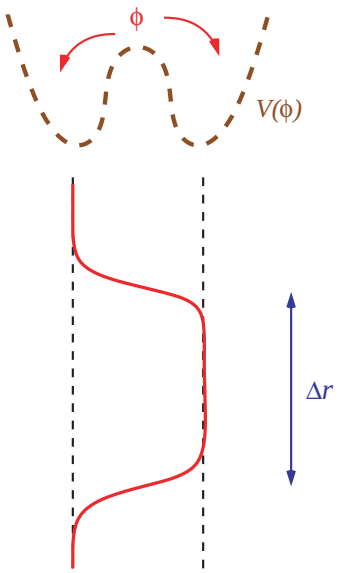

Fig. 8. Sketch of the Kibble-Zurek mechanism [26] for the simple example of a real scalar field (*solid red curve*) in one spatial dimension with a double-well potential (*dashed brown curve*). Initially, the temperature is large and the field is in the symmetric phase $\langle\phi\rangle = 0$. After cooling down, the field has to choose one of the two minima (*dashed lines*). In the presence of a real or effective horizon of size Δr, two points with a distance larger than Δr cannot exchange information about which minimum to select and hence they will choose different minima with a finite probability. This generates topological defects (here: kinks) whose typical distance is roughly determined by the horizon size Δr at the "moment of choice"

3.2 Time-dependent Phase Transitions at Zero Temperature

The Kibble-Zurek mechanism is an example for the amplification of thermal (i.e., classical) fluctuations leading to the generation of topological defects in the presence of a real or effective particle horizon. However, as we have seen at the beginning of this section, a horizon also implies the amplification of quantum fluctuations. Therefore, let us consider phase transitions at very low temperatures – where quantum fluctuations play the dominant role – instead of thermal transitions in the following: For simplicity, we assume zero temperature $T = 0$. Hence, the phase transition cannot by generated by heating or cooling, but by changing an external parameter g instead. This could be the pressure p or an external magnetic field B etc. A phase transition occurs if the structure of the ground state changes for a certain critical value of this external parameter $g = g_c$: Below this critical value $g < g_c$, the ground state $|\Psi_<(g)\rangle$ of the g-dependent Hamiltonian $\hat{H}(g)$ is different from the ground state $|\Psi_>(g)\rangle$ of $\hat{H}(g)$ above this critical value $g > g_c$, see Fig. 9. For example,

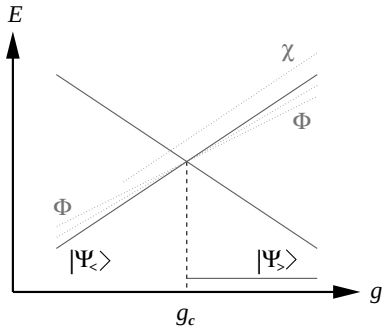

Fig. 9. Sketch of the level structure (i.e., plot of the energy E as a function of the external parameter g for different states) near a phase transition. The *solid green lines* denote the competing ground states $|\Psi_<(g)\rangle$ and $|\Psi_>(g)\rangle$, where the continuous behaviour (level crossing) of $|\Psi_>(g)\rangle$ is typical for a first-order phase transition, but the new ground state could also be non-continuous and just come into existence for $g \geq g_c$ (*lower solid green line*). However, we are mostly interested in the vicinity of the initial ground state $|\Psi_<(g)\rangle$ and its excitations (*red dotted lines*). Some of them (χ) remain stable after the transition $g > g_c$, i.e., still lie above $|\Psi_<(g)\rangle$, but others (Φ) become unstable, i.e., dive below $|\Psi_<(g)\rangle$. These quasi-particle excitations Φ, which are stable for $g < g_c$ (since there $|\Psi_<(g)\rangle$ is the true ground state) and become unstable for $g > g_c$, trigger the phase transition, i.e., the decay from the false vacuum $|\Psi_<(g)\rangle$ to the true ground state $|\Psi_>(g)\rangle$

$|\Psi_<(g)\rangle$ and $|\Psi_>(g)\rangle$ could have different global/topological properties (such as magnetisation) in the thermodynamic limit.

As in the case of thermal phase transitions, there exists a vast amount of literature regarding the equilibrium properties in the vicinity of the phase transition, for example in view of universal behaviour (e.g., scaling laws) near the critical point [27]. However, since response times typically diverge in the vicinity of the critical point, sweeping through the phase transition with a finite velocity, for example, leads to a break-down of adiabaticity and thus generates interesting dynamical (non-equilibrium) effects in complete analogy to the examples studied before. Let us consider the quantum dynamics of quasi-particle excitations Φ, which are stable for $g < g_c$ and become unstable for $g > g_c$, see Fig. 9. In order to do this quantitatively, a few assumptions are necessary: Firstly, we suppose that the modes Φ are **linear** and **independent** and obey an **analytic** dispersion relation with a **vanishing gap** – which are exactly the same conditions that lead to the action in Eq. (4). In addition, if we assume a **homogeneous** and **isotropic** medium (on large scales – which is the same assumption as in the cosmological principle), the most general low-energy effective action reads [28]

$$\mathcal{L}_{\text{eff}} = \frac{1}{2}\left(\frac{1}{\alpha[g(t)]}\,\dot{\Phi}^2 - \beta[g(t)]\,(\boldsymbol{\nabla}\Phi)^2\right) . \tag{16}$$

The two factors α and β depend on the external parameter g and therewith on time. As one would expect, this is equivalent to a scalar field in an expanding or contracting universe with the effective metric

$$ds^2_{\rm eff} = \sqrt{\alpha\beta^3}\, dt^2 - \sqrt{\beta/\alpha}\, d\boldsymbol{r}^2 \ . \tag{17}$$

Constructing the effective energy

$$\mathcal{H}_{\rm eff} = \frac{1}{2}\left(\alpha\, \Pi^2 + \beta\, (\boldsymbol{\nabla}\Phi)^2\right) \ , \tag{18}$$

we see that both parameters α and β must be positive before the phase transition (where Φ is stable), whereas at least one of the two parameters has to change its sign at the critical point (since the mode Φ becomes unstable for $g > g_c$). Consequently, the propagation speed c_s given by $c_s^2 = \alpha\beta$ vanishes at the transition (if we exclude the possibility that one of the two coefficients diverges – see next subsection).

The evanescent propagation speed $c_s(g_c) = 0$ is a consequence of the divergence of the response time (the energy gap vanishes at g_c) and results in the occurrence of a particle horizon (as discussed before). This loss of causal connection marks the breakdown of adiabaticity and implies non-equilibrium phenomena such as the amplification of quantum fluctuations explained earlier. The spectrum of these fluctuations can be calculated purely from the effective metric in Eq. (17) and hence just requires the knowledge of the parameters $\alpha(t)$ and $\beta(t)$. Note, however, that even though the *spectrum* of the amplified quantum fluctuations is largely independent of the microscopic structure (universality), the *strength* of their impact in the final state $|\Psi_>\rangle$ (such as the anisotropies of the cosmic microwave background radiation in Fig. 6) depends on the microscopic structure.

Let us consider a few simple examples: In the dilute-gas limit, the quantum phase and density fluctuations within atomic Bose–Einstein condensates are small and can be treated as linear perturbations. For wavelengths far above the healing length, the effective action of the phase fluctuations Φ reads ($\hbar = 1$)

$$\mathcal{L}_{\rm eff} = \frac{1}{2}\left(\frac{1}{g}\,\dot{\Phi}^2 - \frac{\varrho_0}{m}\,(\boldsymbol{\nabla}\Phi)^2\right) \ , \tag{19}$$

where ϱ_0 is the background density of the condensate, m the mass of the atoms, and g the time-dependent coupling strength representing the interparticle repulsion $g > 0$ or attraction $g < 0$. Obviously, a homogeneous condensate becomes unstable for attractive interactions and hence the critical point is $g_c = 0$. The speed of sound is governed by the two-particle contact repulsion, which can by modified by an external magnetic field via the mechanism based on the Feshbach resonance. If one adjusts the time-dependent external magnetic field such that $c_{\rm s}(t) \propto 1/t^2$ for late times $t \uparrow \infty$, one exactly recovers the de Sitter metric in Eq. (12). As a result the phonons (phase fluctuations) within such Bose–Einstein condensates as described by

Eq. (1) behave in the same way as the inflaton field during inflation – i.e., the modes undergo the three stages described above: oscillation, horizon crossing, and freezing. The frozen quantum fluctuations possess a scale-invariant $1/k^3$-spectrum and are amplified to a size of the order of one percent, which places an observation within reach of present experimental capabilities. Replacing the fine-tuned evolution $c_{\rm s}(t) \propto 1/t^2$ by a constant rate of change of the time-dependent external magnetic field also corresponds to an expanding universe, but with a metric different from the de Sitter metric in Eq. (12), and generates a $k^{-4/3}$-spectrum for the phase fluctuations and a $k^{4/3}$-spectrum for the density fluctuations [28]. Note that this behaviour is consistent with the amplification/suppression of quantum fluctuations by squeezing which maintains the minimal Heisenberg uncertainty of the ground state, i.e., $\Delta q_k \Delta p_k = \hbar/2$.

As a second example, we study a simple 1+1 dimensional Hopefield model of the electromagnetic field coupled to a linear medium via the magnetic component

$$\mathcal{L}_{\rm eff} = \frac{1}{2}\big(E^2 - B^2 + \dot{\Psi}^2 - \Omega^2\Psi^2 + 2gB\Psi\big) \,, \tag{20}$$

with the electromagnetic field (E, B) being governed by the potential A via $E = \partial_t A$ and $B = \partial_x A$. The field Ψ describes the (linearised and localised) dynamics of the medium (Hopefield model) with the plasma frequency Ω and g denoting the coupling (magnetic dipole moment). Integrating out (i.e., averaging over) the degrees of freedom Ψ of the medium, the low-energy effective theory for macroscopic electrodynamics in media yields the permeability $1/\mu = 1 - g^2/\Omega^2$ (which corresponds to inserting the adiabatic solution $\Psi \approx gB/\Omega^2$ back into the action). Hence there is a zero-temperature phase transition at the critical value of the coupling $g_c = \Omega$ after which the medium becomes unstable to spontaneous magnetisation and the linearised description above breaks down. If we sweep through the critical point with a finite velocity, the frozen two-point spectra behave as $k^{-2/3}$ for A, and thus $k^{4/3}$ for B (and Ψ), and finally $k^{2/3}$ for $E = \Pi$ (again respecting $\Delta q_k \Delta p_k = \hbar/2$).

3.3 Similarities to Cosmic Inflation

Interestingly, sweeping through a zero-temperature phase transition by means of a time-dependent external parameter displays various similarities to "real" cosmic inflation:

- **Release of energy**
  After crossing the critical point, the system does not jump to the new ground state $|\Psi_>\rangle$ immediately. Because its response time is very large, the system stays in the vicinity of the false vacuum $|\Psi_<\rangle$ for a while and decays down to $|\Psi_>\rangle$ after a finite period of time. At that time, there will be a finite energy difference between the states $|\Psi_<\rangle$ and $|\Psi_>\rangle$, which is released by this decay. In the absence of damping, this energy difference

will heat up the system – in close similarity to the (p)re-heating after inflation.

- **Robustness**
  The aforementioned decay typically occurs at small length scales since these are the fastest modes. Therefore, structures on large length scales may well leave their imprint in the final state, but initial small-scale perturbations/excitations will be completely washed out and do not affect the final state significantly.
- **Universality**
  As it became evident from the previous considerations, merely a few assumptions suffice to determine the relevant features of the system, i.e., no fine-tuning is necessary.
- **Amplification of quantum fluctuations**
  Finally, the emergence of an effective particle horizon implies the same amplification mechanism (oscillation → horizon crossing → freezing and squeezing) as in inflation.

The first three points are qualitative, but the last one has a quantitative nature and points to the main difference between zero-temperature phase transitions in the laboratory and cosmic inflation – the spectrum of the amplified quantum fluctuations. Unless the dynamics of $g(t)$ is fine-tuned such as $g(t) \propto 1/t^4$ for Bose–Einstein condensates, laboratory systems do not reproduce the scale-invariant $1/k^3$-spectrum of inflation in general. (E.g., a sweep with a finite velocity yields a $k^{-4/3}$-spectrum in Bose–Einstein condensates.) On the other hand, this may not be too surprising as the laboratory systems considered above break many symmetries we observe in the real universe, e.g., they possess a preferred frame and do not respect the principle of equivalence etc. This leads us to the question: can we conceive a phase transition which does not break these symmetries? If we *demand* that the effective action (at least at low energies) does not single out a locally preferred frame (remember that the two-point function $\langle\hat{\Phi}(\underline{x})\hat{\Phi}(\underline{x}')\rangle$ depends on the effective Ricci scalar etc.) and that the velocity of propagation (i.e., the speed of light) be constant (such that we can set it equal unity), there is only one possibility left

$$\mathcal{A} = \frac{1}{2}\int dt\, d^3r\, \frac{\dot{\Phi}^2 - (\boldsymbol{\nabla}\Phi)^2}{t^2} \,. \tag{21}$$

Note that, imposing the two conditions above (no locally preferred frame and constant speed of light), we get scale-invariance for free – it turns out that the effective action satisfies these requirements if and only if it is scale-invariant

$$\mathcal{A}[\lambda t, \lambda \boldsymbol{r}] = \mathcal{A}[t, \boldsymbol{r}] \,. \tag{22}$$

For a laboratory system, such an action would probably seem rather strange, but one might expect that its dynamics is dominated by (quantum) back-reaction effects, which have been omitted so far (since they are usually very

small in laboratory systems, but not necessarily in the early universe). In view of the above scale-invariance, we indeed reproduce the scale-invariant $1/k^3$-spectrum of inflation, which is also no surprise since the effective metric is just the de Sitter metric in conformal coordinates. These interesting findings entice the question/speculation of whether the epoch of cosmic inflation did not actually correspond to a real exponential expansion – but to our distorted view on such a phase transition in the very early universe instead.

3.4 Expanding Bose–Einstein Condensates

As mentioned earlier, there are basically two possibilities for simulating an expanding universe with the effective line-element in Eqs. (3) and (14)

$$ds^2_{\rm eff} = \frac{\varrho_0}{c_{\rm s}} \left([c_s^2 - \boldsymbol{v}_0^2] dt^2 + 2\boldsymbol{v}_0 \cdot d\boldsymbol{r} dt - d\boldsymbol{r}^2 \right) \ .$$

So far, we considered the fluid such as a Bose–Einstein condensate at rest with a time-dependent speed of sound $c_s(t)$. The second possibility is to let the condensate expand [29], which can be reached by changing the trapping potential (which is used to confine the dilute atomic/molecular gas), e.g., to switch it off entirely. (This is already done in time-of-flight measurements.) Assuming homogeneity and isotropy of the background fluid (which reflects the situation at the centre of the condensate quite well), we can describe the dynamics of the background by a scale-factor $b(t)$ and the velocity assumes a particularly simple form $\boldsymbol{v}_0 = \boldsymbol{r}\dot{b}/b$. In complete analogy the general relativity, we can diagonalise and thereby simplify the metric in Eq. (14) by a transformation from the laboratory coordinates $(t, \boldsymbol{r})$ to co-moving coordinates $\boldsymbol{r} = b(t)\boldsymbol{R}$

$$ds^2_{\rm eff} = \frac{\varrho_0}{c_{\rm s}} \left(c_s^2 dt^2 - b^2 d\boldsymbol{R}^2 \right) = d\tau^2 - \frac{\varrho_0 b^2}{c_{\rm s}} \, d\boldsymbol{R}^2 \ , \tag{23}$$

where $\tau = \int^t dt' \sqrt{\varrho_0(t') c_s(t')}$ denotes the effective proper co-moving time. Note that the background density ϱ_0 as well as the speed of sound c_s will be time-dependent in an expanding condensate $\varrho_0 \propto 1/b^3$ and $c_s \propto 1/b^{3/2}$ (assuming that the two-particle contact repulsion remains constant).

As an example, consider the free expansion of a three-dimensional condensate after the trap is suddenly switched off at time $t = 0$. Initially, the interaction energy is transfered to kinetic energy and the condensate accelerates ($\ddot{b} > 0$). But very quickly (after a period of time which is roughly the inverse of the initial trapping frequency) the scale factor approaches a constant velocity and we can approximate $b = \alpha t$. Insertion into Eq. (23) yields an accelerated expanding universe with an apparent horizon (see Fig. 10) at [29]

$$r_{\rm horizon} = \frac{c_s(t=0)}{\alpha^{3/2}} t^{-1/2} \ , \qquad R_{\rm horizon} = \frac{c_s(t=0)}{\alpha^{5/2}} t^{-3/2} \ , \tag{24}$$

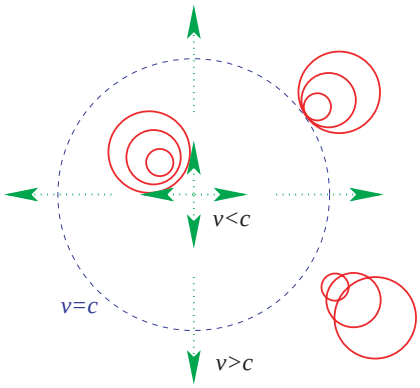

Fig. 10. Sketch of an apparent horizon (*blue dashed circle*). The *dotted green arrows* denote the streamlines of the expanding fluid (moving away from a chosen fluid particle at rest in the centre of the picture). The *solid red circles* are sound waves. The apparent horizon (*blue dashed circle*) emerges at the radius where the flow velocity v exceeds the speed of sound c_s. Obviously, no sound wave emitted beyond the horizon can reach the inner region, i.e., the chosen fluid particle at rest

expressed in terms of the laboratory and co-moving coordinates, respectively. Note that the free (linear $b = \alpha t$) expanding condensate does not exactly correspond to the de Sitter metric in Eq. (12), which would again require some fine-tuning of the temporal evolution of the trapping potential (instead of just switching it off, which is relatively easy to do experimentally).

As a consequence of the occurrence of a horizon, the initial quantum fluctuations of the sound modes will be frozen and amplified during the (free) expansion of the condensate, see Fig. 11. Again, their spectrum can be calculated using the tools known from general relativity and cosmology: the density-density correlation function $\langle \delta\hat{\varrho}(\boldsymbol{r})\delta\hat{\varrho}(\boldsymbol{r'})\rangle$ possesses a $k^{4/3}$-spectrum (as in the previous case) with its relative size being on the percent level [29]. Hence it should be feasible to observe this mechanism – which is analogous to the generation of the seeds of structure formation in the early universe – with high-precision absorption images of the expanding cloud (as it is already done in the time-of-flight measurements).

4 Summary and Outlook

The analogy between quantised sound waves in fluids and quantum fields in curved space-times – which can be extended to many other excitations such as surface waves or photons in dielectrics – facilitates an interdisciplinary know-how transfer in both directions. One the one hand, one may use the microscopic structure of the fluid as a toy model for unknown high-energy (Planckian) effects in quantum gravity, for example, and investigate the influence of

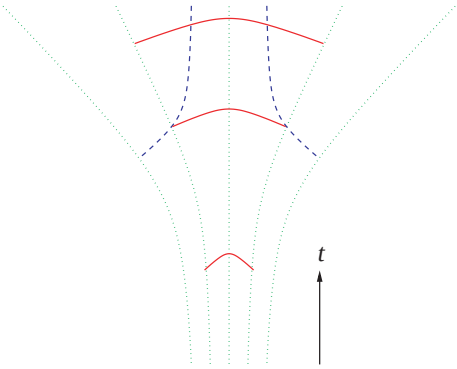

Fig. 11. Sketch of expanding condensate. Time goes up in this diagram and the *dotted green lines* denote particle trajectories of the expanding fluid. The apparent horizon (*dashed blue line*) occurs where the flow velocity v exceeds the speed of sound c_s and moves inwards since v increases but c_s decreases. The *solid red curve* denotes half a wavelength of a given phonon mode, which is being stretched by the expansion of the condensate. Initially, the wavelength is much smaller than the horizon and oscillates almost freely. Later on, the mode crosses the horizon and after that, the oscillation freezes because the two knots are separated by the horizon and hence cannot exchange energy anymore

the corresponding cut-off – see the chapter by S. Weinfurtner, S. Liberati, and M. Visser and [30], for example. Examining the derivation of the Hawking effect for various dispersion relations, one reproduces Hawking radiation for a rather large class of scenarios, but there are also counter-examples, which do not appear to be unphysical or artificial, displaying strong deviations from Hawkings result. Therefore, whether real black holes emit Hawking radiation remains an open question and could give non-trivial information about Planckian physics.

On the other hand, the emergence of an effective geometry/metric allows us to apply the vast amount of universal tools and concepts developed for general relativity (such as horizons), which provide a unified description and better understanding of (classical and quantum) non-equilibrium phenomena (e.g., freezing and amplification of quantum fluctuations) in condensed matter systems. As an example for such a universal mechanism, the Kibble-Zurek effect describes the generation of topological effects due to the amplification of classical/thermal fluctuations in non-equilibrium thermal phase transitions. This effect is one of the main themes of the COSLAB programme. It may play a role in cosmology (e.g., falsifying some models, see chapter by M. Sakellariadou) and has been observed in the laboratory. The loss of causal connection underlying the Kibble-Zurek mechanism can be understood in terms of an effective horizon – which clearly indicates the departure from equilibrium. The associated breakdown of adiabaticity leads to an amplification of thermal fluctuations (as in the Kibble-Zurek mechanism) as well as quantum fluctua-

tions (at zero temperature). The zero-temperature version of this amplification mechanism is completely analogous to the early universe and becomes particularly important for the new and rapidly developing field of quantum phase transitions (see contributions by G. E. Volovik as well as J. Yngvason[2]).

Furthermore, these analogue models might provide the exciting opportunity of measuring the analogues of these exotic effects – such as Hawking radiation or the generation of the seeds for structure formation during inflation – in actual laboratory experiments, i.e., experimental quantum simulations of black hole physics or the early universe. Even though the detection of these exotic quantum effects is partially very hard (see, e.g., [23]) and requires ultralow temperatures etc., there is no (known) principal objection against it (see chapter by P. Skyba and [22]). The analogue models range from black and/or white hole event horizons (see chapter by U. Leonhardt and T. G. Philbin) in flowing fluids and other laboratory systems (see chapter by P. Skyba) over apparent horizons in expanding Bose–Einstein condensates, for example, to particle horizons in quantum phase transitions etc.

However, one should stress that the analogy reproduces the kinematics (quantum fields in curved space-times with horizons etc.) but not the dynamics, i.e., the effective geometry/metric is not described by the Einstein equations in general. An important and strongly related problem is the correct description of the back-reaction (see chapter by U. R. Fischer) of the quantum fluctuations (e.g., phonons) onto the background (e.g., fluid flow). In gravity, the impact of the (classical or quantum) matter is usually incorporated by the (expectation value of the) energy-momentum tensor. Since this quantity can be introduced at a purely kinematic level, we may use the same construction for phonons in flowing fluids, for example – which is called the pseudo energy-momentum tensor. The relevant component of this tensor describing the energy density (which is conserved for stationary flows) may become negative as soon as the flow velocity exceeds the sound speed. These negative contributions explain the energy balance of the Hawking radiation in black hole analogues as well as super-radiant scattering. However, it turns out that the (expectation value of the) pseudo energy-momentum tensor does not determine the quantum back-reaction correctly [31].

Finally, one should not neglect to mention another possibility for the occurrence of a horizon in the laboratory – the Unruh effect (see chapter by B. Kuckert). A uniformly accelerated observer cannot see half of the (1+1-dimensional) space-time, see Fig. 12, i.e., the two Rindler wedges are completely causally disconnected by the horizon(s). In each wedge, one may introduce a set of observables corresponding to the measurements made by the observers confined to this wedge – thereby obtaining two equivalent copies of observables in one wedge. In terms of these two copies, the Minkowski vacuum is an entangled state which yields the usual phenomena (thermo-field

[2] This contribution has already been published elsewhere and can be found in Lect. Notes Phys. **690**, 199–215 (2006).

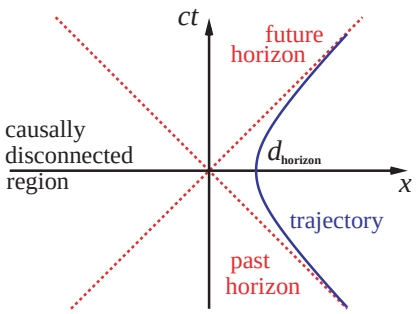

Fig. 12. Space-time diagram with a trajectory of a uniformly accelerated observer (*solid blue curve*) and the resulting particle horizons (*red dotted lines*). The observer is confined to the right Rindler wedge (region $x > |ct|$ between the two horizons) and cannot influence or be influenced by all events in the left Rindler wedge ($x < |ct|$), which is completely causally disconnected

formalism etc., see chapter by G. Vitiello) including the Unruh effect – i.e., the uniformly accelerated observer experiences the Minkowski vacuum as a thermal bath: For rather general quantum fields (Bisognano-Wichmann theorem, see chapter by B. Kuckert), it can be shown that the quantum state $\hat{\varrho}$ obtained by restricting the Minkowski vacuum to one of the Rindler wedges behaves as a mixed state $\hat{\varrho} = \exp\{-2\pi\hat{H}_\tau/\kappa\}/Z$, where $\hat{H}_\tau$ corresponds to the Hamiltonian generating the proper (co-moving wristwatch) time τ measured by the accelerated observer and κ is the analogue to the surface gravity in Sect. 2.1 and determines the acceleration.

The thermal character of this restricted state $\hat{\varrho}$ arises from the quantum correlations of the Minkowski vacuum in the two Rindler wedges, i.e., the Minkowski vacuum is a multi-mode squeezed state with respect the two equivalent copies of observables in each wedge. This is a quite general phenomenon associated with doubling the degrees of freedom (see chapter by G. Vitiello) and describes the underlying idea of the thermo-field formalism, for example. The entropy of the thermal radiation in the Unruh and the Hawking effect can be understood as an entanglement entropy: For the Unruh effect, it is caused by averaging over the quantum correlations between the two Rindler wedges. In the black hole case, each particle of the outgoing Hawking radiation has its infalling partner particle (with a negative energy with respect to spatial infinity) and the entanglement between the two generates the entropy flux of the Hawking radiation.

Instead of accelerating a detector and measuring its excitations, one could replace the accelerated observer by an accelerated scatterer. This device would scatter (virtual) particles from the thermal bath and thereby create real particles – which can be interpreted as a signature if Unruh effect (see [32] and chapter by P. Skyba).

Acknowledgements

This work was supported by the Emmy Noether Programme of the German Research Foundation (DFG) under grant No. SCHU 1557/1-1/2. The author gratefully acknowledges further support by the COSLAB Programme of the ESF, the EU-ULTI Programme, and the Pacific Institute of Theoretical Physics.

References

1. N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England 1982); S. A. Fulling, *Aspects of Quantum Field Theory in Curved Space-Time* (Cambridge University Press, Cambridge, England 1989)
2. R. Schützhold and M. Uhlmann, in *Horizon Analogues in the Laboratory*, Proceedings of the Memorial Symposium for Gerhard Soff (April 25 and 26, 2005, Frankfurt, Germany)
3. W. G. Unruh, *Experimental Black Hole Evaporation?*, Phys. Rev. Lett. **46** (1981) 1351; *Sonic Analogue of Black Holes and the Effects of High Frequencies on Black Hole Evaporation*, Phys. Rev. D **51** (1995) 2827
4. P. Painlevé, *La Mécanique classique et la théorie de la relativité*, C. R. Hebd. Seances Acad. Sci. (Paris) **173** (1921) 677; A. Gullstrand, *Allgemeine Lösung des statischen Einkörperproblems in der Einsteinschen Gravitationstheorie*, Ark. Mat. Astron. Fys. **16** (1922) 1; G. Lemaître, *L'univers en expansion*, Ann. Soc. Sci. (Bruxelles) A **53** (1933) 51
5. C. M. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freemann, San Francisco, 1973)
6. S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-time* (Cambridge University Press, Cambridge, England, 1973)
7. G. E. Volovik, *Universe in a Helium Droplet* (Oxford University Press, Oxford, 2003); *Superfluid analogies of cosmological phenomena*, Phys. Rept. **351** (2001) 195
8. M. Novello, M. Visser, and G. Volovik (editors), *Artificial Black Holes* (World Scientific, Singapore, 2002)
9. C. Barceló, S. Liberati, and M. Visser, *Analogue Gravity*, Living Rev. Rel. **8**, 12 (2005); and references therein
10. C. Barcelo, S. Liberati and M. Visser, *Analogue gravity from field theory normal modes?*, Class. Quant. Grav. **18** (2001) 3595; *Refringence, field theory, and normal modes*, *ibid* **19** (2002) 2961; *Einstein gravity as an emergent phenomenon?*, Int. J. Mod. Phys. D **10** (2001) 799
11. R. Schützhold and W. G. Unruh, *Gravity wave analogues of black holes*, Phys. Rev. D **66** (2002) 044019
12. R. Schützhold and W. G. Unruh, *Hawking radiation in an electro-magnetic wave-guide?*, Phys. Rev. Lett. **95**, 031301 (2005)
13. R. Schützhold, G. Plunien, and G. Soff, *Dielectric black hole analogs*, Phys. Rev. Lett. **88** (2002) 061101
14. W. Gordon, *Zur Lichtfortpflanzung nach der Relativitätstheorie*, Ann. Phys. (Leipzig) **72** (1923) 421

15. W. G. Unruh and R. Schützhold, *On Slow Light as a Black Hole Analogue*, Phys. Rev. D **68** (2003) 024008; see also U. Leonhardt and P. Piwnicki, *Relativistic Effects of Light in Moving Media with Extremely Low Group Velocity*, Phys. Rev. Lett. **84** (2000) 822; with comment by M. Visser, *Comment on "Relativistic Effects of Light in Moving Media with Extremely Low Group Velocity"*, ibid. **85** (2000) 5252; and reply U. Leonhardt and P. Piwnicki, *Reply to comment on "Relativistic Effects of Light in Moving Media with Extremely Low Group Velocity"*, ibid. **85** (2000) 5253
16. J. D. Bekenstein, *Black Holes and the Second Law*, Lett. Nuovo Cim. **4** (1972) 737; *Black Holes And Entropy*, Phys. Rev. D **7** (1973) 2333; *Generalized Second Law of Thermodynamics in Black Hole Physics*, ibid **9** (1974) 3292; *Statistical Black Hole Thermodynamics*, ibid **12** (1975) 3077; J. M. Bardeen, B. Carter and S. W. Hawking, *The Four Laws of Black Hole Mechanics*, Commun. Math. Phys. **31** (1973) 161
17. S. W. Hawking, *Black Hole Explosions*, Nature **248** (1974) 30; *Particle Creation by Black Holes*, Commun. Math. Phys. **43** (1975) 199
18. W. G. Unruh and R. Schützhold, *On the universality of the Hawking effect*, Phys. Rev. D **71** (2005) 024028
19. T. Jacobson, *Trans-Planckian redshifts and the substance of the space-time river*, Prog. Theor. Phys. Suppl. **136** (1999) 1; *Black Hole Evaporation and Ultrashort Distances*, Phys. Rev. D **44** (1991) 1731; ibid **48** (1993) 728; *On the Origin of the Outgoing Black Hole Modes*, ibid **53** (1996) 7082; T. Jacobson and D. Mattingly, *Hawking radiation on a falling lattice*, ibid **61** (2000) 024017; S. Corley and T. Jacobson, *Hawking Spectrum and High Frequency Dispersion*, ibid **54** (1996) **54**; *Lattice black holes*, ibid **57** (1998) 6269; *Black hole lasers*, ibid **59** (1999) 124011; S. Corley, *Particle creation via high frequency dispersion*, ibid **55** (1997) 6155; *Computing the spectrum of black hole radiation in the presence of high frequency dispersion: An analytical approach*, ibid **57** (1998) 6280
20. S. Liberati, S. Sonego and M. Visser, *Unexpectedly Large Surface Gravities for Acoustic Horizons?*, Class. Quant. Grav. **17**, 2903 (2000)
21. L. J. Garay, J. R. Anglin, J. I. Cirac, and P. Zoller, *Sonic Analog of Gravitational Black Holes in Bose-Einstein Condensates*, Phys. Rev. Lett. **85**, 4643 (2000); *Sonic black holes in dilute Bose-Einstein condensates*, Phys. Rev. A **63**, 023611 (2001); S. Giovanazzi, C. Farrell, T. Kiss, and U. Leonhardt, *Conditions for one-dimensional supersonic flow of quantum gases*, Phys. Rev. A **70**, 063602 (2004)
22. R. Schützhold, *On the detectability of quantum radiation in Bose-Einstein condensates*, Phys. Rev. Lett. **97**, 190405 (2006)
23. W. G. Unruh, *Measurability of Dumb Hole Radiation?*, in *Artificial Black Holes*, edited by M. Novello, M. Visser, and G. Volovik (World Scientific, Singapore, 2002)
24. F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, *Theory of Bose-Einstein condensation in trapped gases*, Rev. Mod. Phys. **71**, 463 (1999); A. J. Leggett, *Bose-Einstein condensation in the alkali gases: Some fundamental concepts*, ibid. **73**, 307 (2001)
25. J. W. Miles, *On the generation of surface waves by shear flows*, J. Fluid Mech. **3** (1957) 185; G. E. Vekstein, *Landau resonance mechanism for plasma and wind-generated water waves*, Am. J. Phys. **66** (1998) 886
26. T. W. B. Kibble, *Topology of Cosmic Domains And Strings*, J. Phys. A **9**, 1387 (1976); *ome Implications of A Cosmological Phase Transition*, Phys. Rept. **67**,

183 (1980); *Symmetry breaking and defects*, pp. 3-36 in *Patterns of Symmetry Breaking*, H. Arodz et al, eds. (Kluwer Academic, 2003) W. H. Zurek, *Cosmological Experiments In Superfluid Helium?*, Nature **317**, 505 (1985); *Cosmic strings in laboratory superfluids and the topological remnants of other phase transitions*, Acta Phys. Polon. B **24**, 1301 (1993); *Cosmological Experiments in Condensed Matter Systems*, Phys. Rept. **276**, 177 (1996)

27. S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, England, 1999)
28. R. Schützhold, *Dynamical zero-temperature phase transitions and cosmic inflation/deflation*, Phys. Rev. Lett. **95**, 135703 (2005); U. R. Fischer and R. Schützhold, *Quantum simulation of cosmic inflation in two-component Bose-Einstein condensates*, Phys. Rev. A **70** (2004) 063615; R. Schützhold, M. Uhlmann, Y. Xu and U. R. Fischer, *Sweeping from the superfluid to Mott phase in the Bose-Hubbard model*, Phys. Rev. Lett. **97**, 200601 (2006)
29. M. Uhlmann, Y. Xu, and R. Schützhold, *Aspects of Cosmic Inflation in Expanding Bose-Einstein Condensates*, New J. Phys. **7**, 248 (2005)
30. F. Queisser, M. Uhlmann, and R. Schützhold, *Signatures of Planck-scale interactions in the cosmic microwave background?*, `gr-qc/0601108`
31. R. Schützhold, M. Uhlmann, Y. Xu, and U. R. Fischer, *Quantum back-reaction in dilute Bose-Einstein condensates*, Phys. Rev. D **72**, 105005 (2005)
32. R. Schützhold, G. Schaller, and D. Habs, *Signatures of the Unruh effect from electrons accelerated by ultra-strong laser fields*, Phys. Rev. Lett. **97**, 121302 (2006)

Quantum Phase Transitions from Topology in Momentum Space

G. E. Volovik[1,2]

[1] Low Temperature Laboratory, Helsinki University of Technology, P.O. Box 2200, FIN-02015 HUT, Espoo, Finland
[2] Landau Institute for Theoretical Physics, Kosygina 2, 119334 Moscow, Russia

Many quantum condensed matter systems are strongly correlated and strongly interacting fermionic systems, which cannot be treated perturbatively. However, physics which emerges in the low-energy corner does not depend on the complicated details of the system and is relatively simple. It is determined by the nodes in the fermionic spectrum, which are protected by topology in momentum space (in some cases, in combination with the vacuum symmetry). Close to the nodes the behavior of the system becomes universal; and the universality classes are determined by the toplogical invariants in momentum space. When one changes the parameters of the system, the transitions are expected to occur between the vacua with the same symmetry but which belong to different universality classes. Different types of quantum phase transitions governed by topology in momentum space are discussed in this chapter. They involve Fermi surfaces, Fermi points, Fermi lines, and also the topological transitions between the fully gapped states. The consideration based on the momentum space topology of the Green's function is general and is applicable to the vacua of relativistic quantum fields. This is illustrated by the possible quantum phase transition governed by topology of nodes in the spectrum of elementary particles of Standard Model.

1 Introduction

There are two schemes for the classification of states in condensed matter physics and relativistic quantum fields: classification by symmetry (GUT scheme) and by momentum space topology (anti-GUT scheme).

For the first classification method, a given state of the system is characterized by a symmetry group H which is a subgroup of the symmetry group G of the relevant physical laws. The thermodynamic phase transition between equilibrium states is usually marked by a change of the symmetry group H.

G.E. Volovik: *Quantum Phase Transitions from Topology in Momentum Space*, Lect. Notes Phys. **718**, 31–73 (2007)
DOI 10.1007/3-540-70859-6_3

This classification reflects the phenomenon of spontaneously broken symmetry. In relativistic quantum fields the chain of successive phase transitions, in which the large symmetry group existing at high energy is reduced at low energy, is in the basis of the Grand Unification models (GUT) [1, 2]. In condensed matter the spontaneous symmetry breaking is a typical phenomenon, and the thermodynamic states are also classified in terms of the subgroup H of the relevant group G (see e.g, the classification of superfluid and superconducting states in Refs. [3, 4]). The groups G and H are also responsible for topological defects, which are determined by the nontrivial elements of the homotopy groups $\pi_n(G/H)$; cf. [5].

The second classification method reflects the opposite tendency – the anti Grand Unification (anti-GUT) – when instead of the symmetry breaking the symmetry gradually emerges at low energy. This method deals with the ground states of the system at zero temperature ($T = 0$), i.e., it is the classification of quantum vacua. The universality classes of quantum vacua are determined by momentum-space topology, which is also responsible for the type of the effective theory, emergent physical laws and symmetries at low energy. Contrary to the GUT scheme, where the symmetry of the vacuum state is primary giving rise to topology, in the anti-GUT scheme the topology in the momentum space is primary while the vacuum symmetry is the emergent phenomenon in the low energy corner.

At the moment, we live in the ultra-cold Universe. All the characteristic temperatures in our Universe are extremely small compared to the Planck energy scale $E_{\rm P}$. That is why all the massive fermions, whose natural mass must be of order $E_{\rm P}$, are frozen out due to extremely small factor $\exp(-E_{\rm P}/T)$. There is no matter in our Universe unless there are massless fermions, whose masslessness is protected with extremely high accuracy. It is the topology in the momentum space, which provides such protection.

For systems living in 3D space, there are four basic universality classes of fermionic vacua provided by topology in momentum space [6, 7]:

(i) Vacua with fully-gapped fermionic excitations, such as semiconductors and conventional superconductors.

(ii) Vacua with fermionic excitations characterized by Fermi points– points in 3D momentum space at which the energy of fermionic quasiparticle vanishes. Examples are provided by superfluid ^{3}He-A and also by the quantum vacuum of Standard Model above the electroweak transition, where all elementary particles are Weyl fermions with Fermi points in the spectrum. This universality class manifests the phenomenon of emergent relativistic quantum fields at low energy: close to the Fermi points the fermionic quasiparticles behave as massless Weyl fermions, while the collective modes of the vacuum interact with these fermions as gauge and gravitational fields.

(iii) Vacua with fermionic excitations characterized by lines in 3D momentum space or points in 2D momentum space. We call them Fermi lines, though

in general it is better to characterize zeroes by co-dimension, which is the dimension of **p**-space minus the dimension of the manifold of zeros. Lines in 3D momentum space and points in 2D momentum space have co-dimension 2: since $3-1=2-0=2$; compare this with zeroes of class (ii) which have co-dimension $3-0=3$. The Fermi lines are topologically stable only if some special symmetry is obeyed. Example is provided by the vacuum of the high T_c superconductors where the Cooper pairing into a d-wave state occurs. The nodal lines (or actually the point nodes in these effectively 2D systems) are stabilized by the combined effect of momentum-space topology and time reversal symmetry.

(iv) Vacua with fermionic excitations characterized by Fermi surfaces. The representatives of this universality class are normal metals and normal liquid ^{3}He. This universality class also manifests the phenomenon of emergent physics, though non-relativistic: at low temperature all the metals behave in a similar way, and this behavior is determined by the Landau theory of Fermi liquid – the effective theory based on the existence of Fermi surface. Fermi surface has co-dimension 1: in 3D system it is the surface (co-dimension $= 3-2=1$), in 2D system it is the line (co-dimension $= 2-1=1$), and in 1D system it is the point (co-dimension $= 1-0=1$; in one dimensional system the Landau Fermi-liquid theory does not work, but the Fermi surface survives).

The possibility of the Fermi band class (v), where the energy vanishes in the finite region of the 3D momentum space and thus zeroes have co-dimension 0, has been also discussed [8–11]. It is believed that this the so-called Fermi condensate may occur in strongly interacting electron systems $PuCoGA_5$ and $CeCoIn_5$ [12]. Topologically stable flat band may exist in the spectrum of fermion zero modes, i.e. for fermions localized in the core of the topological objects [13].

The phase transitions which follow from this classification scheme are quantum phase transitions which occur at $T=0$ [14]. It may happen that by changing some parameter q of the system we transfer the vacuum state from one universality class to another, or to the vacuum of the same universality class but different topological quantum number, without changing its symmetry group H. The point q_c, where this zero-temperature transition occurs, marks the quantum phase transition. For $T\neq 0$, the second order phase transition is absent, as the two states belong to the same symmetry class H, but the first order phase transition is not excluded. Hence, there is an isolated singular point $(q_c,0)$ in the (q,T) plane (Fig. 1(**a**)), or the end point of the first order transition (Fig. 1(**b**)).

The quantum phase transitions which occur in classes (iv) and (i) or between these classes are well known. In the class (iv) the corresponding quantum phase transition is known as Lifshitz transition [15], at which the Fermi surface changes its topology or emerges from the fully gapped state of class (i), see Sect. 2.2. The transition between the fully gapped states characterized

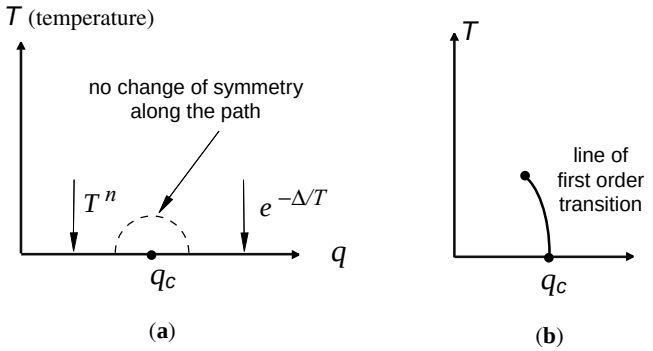

Fig. 1. Quantum phase transition between two ground states with the same symmetry but of different universality class – gapless at $q < q_c$ and fully gapped at $q > q_c$ – as isolated point (**a**) as the termination point of first order transition (**b**)

by different topological charges occurs in 2D systems exhibiting the quantum Hall and spin-Hall effect: this is the plateau-plateau transition between the states with different values of the Hall (or spin-Hall) conductance (see Sect. 5). The less known transitions involve nodes of co-dimension 3 [16–20] (Sect. 3 on Fermi points) and nodes of co-dimension 2 [21–24] (Sect. 4 on nodal lines). The quantum phase transitions involving the flat bands of class (v) are discussed in Ref. [13].

2 Fermi Surface and Lifshitz Transition

2.1 Fermi Surface as a Vortex in p-space

In ideal Fermi gases, the Fermi surface at $p = p_F = \sqrt{2\mu m}$ is the boundary in $\mathbf{p}$-space between the occupied states ($n_{\mathbf{p}} = 1$) at $p^2/2m < \mu$ and empty states ($n_{\mathbf{p}} = 0$) at $p^2/2m > \mu$. At this boundary (the surface in 3D momentum space) the energy is zero. What happens when the interaction between particles is introduced? Due to interaction the distribution function $n_{\mathbf{p}}$ of particles in the ground state is no longer exactly 1 or 0. However, it appears that the Fermi surface survives as the singularity in $n_{\mathbf{p}}$. Such stability of the Fermi surface comes from a topological property of the one-particle Green's function at imaginary frequency:

$$G^{-1} = i\omega - \frac{p^2}{2m} + \mu \, . \tag{1}$$

Let us for simplicity skip one spatial dimension p_z so that the Fermi surface becomes the line in 2D momentum space (p_x, p_y); this does not change the co-dimension of zeroes which remains $1 = 3-2 = 2-1$. The Green's function has singularities lying on a closed line $\omega = 0$, $p_x^2 + p_y^2 = p_F^2$ in the 3D momentum-frequency space (ω, p_x, p_y) (Fig. 2(**a**)). This is the line of the quantized vortex

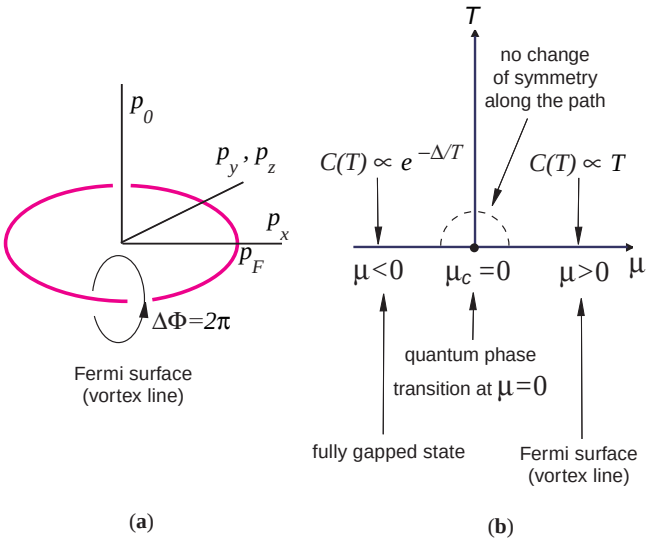

Fig. 2. (**a**) Fermi surface is a topological object in momentum space – a vortex loop. (**b**) When the chemical potential μ decreases the loop shrinks and disappears at $\mu < 0$. The point $\mu = T = 0$ marks the Lifshitz transition between the gapless ground state at $\mu > 0$ to the fully gapped vacuum at $\mu < 0$

in the momemtum space, since the phase Φ of the Green's function $G = |G|e^{i\Phi}$ changes by $2\pi N_1$ around the path embracing any element of this vortex line. In the considered case the phase winding number is $N_1 = 1$. If we add the third momentum dimension p_z the vortex line becomes the surface in the 4D momentum-frequency space (ω, p_x, p_y, p_z) – the Fermi surface – but again the phase changes by 2π along any closed loop empracing the element of the 2D surface in the 4D momentum-frequency space.

The winding number cannot change by continuous deformation of the Green's function: the momentum-space vortex is robust toward any perturbation. Thus the singularity of the Green's function on the Fermi surface is preserved, even when interaction between fermions is introduced. The invariant is the same for any space dimension, since the co-dimension remains 1.

The Green function is generally a matrix with spin indices. In addition, it may have the band indices (in the case of electrons in the periodic potential of crystals). In such a case the phase of the Green's function becomes meaningless; however, the topological property of the Green's function remains robust. The general analysis [7] demonstrates that topologically stable Fermi surfaces are described by the group Z of integers. The winding number N_1 is expressed analytically in terms of the Green's function [6]:

$$N_1 = \mathbf{tr} \oint_C \frac{dl}{2\pi i} G(\mu, \mathbf{p}) \partial_l G^{-1}(\mu, \mathbf{p}) \ . \tag{2}$$

Here the integral is taken over an arbitrary contour C around the momentum-space vortex, and **tr** is the trace over the spin, band and/or other indices.

2.2 Lifshitz Transitions

There are two scenarios of how to destroy the vortex loop in momentum space: perturbative and non-perturbative. The non-perturbative mechanism of destruction of the Fermi surface occurs for example at the superconducting transition, at which the spectrum changes drastically and the gap appears. We shall consider this later in Sect. 2.3, and now let us concentrate on the perturbative processes.

Contraction and Expansion of Vortex Loop in p-Space

The Fermi surface cannot be destroyed by small perturbations, since it is protected by topology and thus is robust to perturbations. But the Fermi surface can be removed by large perturbations in the processes which reproduces the processes occurring for the real-space counterpart of the Fermi surface – the loop of quantized vortex in superfluids and superconductors. The vortex ring can continuously shrink to a point and then disappear, or continuously expand and leave the momentum space. The first scenario occurs when one continuously changes the chemical potential from the positive to the negative value: at $\mu < 0$ there is no vortex loop in momentum space and the ground state (vacuum) is fully gapped. The point $\mu = 0$ marks the quantum phase transition – the Lifshitz transition– at which the topology of the energy spectrum changes (Fig. 2(**b**)). At this transition the symmetry of the ground state does not changes. The second scenario of the quantum phase transition to the fully gapped states occurs when the inverse mass $1/m$ in Eq. (1) crosses zero.

Similar Lifshitz transitions from the fully gapped state to the state with the Fermi surface may occur in superfluids and superconductors. This happens, for example, when the superfluid velocity crosses the Landau critical velocity [Fig. 3]. The symmetry of the order parameter does not change across such a quantum phase transition. On the other examples of the Fermi surface in superfluid/superconducting states in condensed matter and quark matter see [25]. In the non-superconduting states, the transition from the gapless to gapped state is the metal-insulator transition. The Mott transition also belongs to this class.

Reconnection of Vortex Lines in p-Space

The Lifshitz transitions involving the vortex lines in **p**-space may occur between the gapless states. They are accompanied by the change of the topology of the Fermi surface itself. The simplest example of such a phase transition discussed in terms of the vortex lines is provided by the reconnection of the

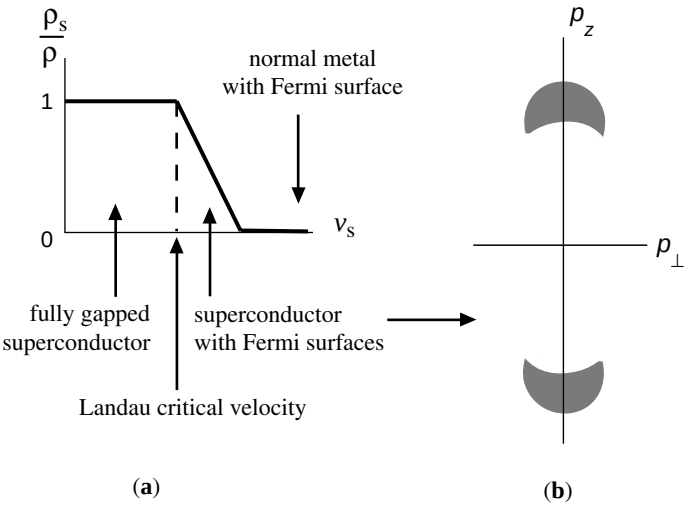

Fig. 3. Illustration of Lifshitz transition in superfluid/superconductor at Landau critical velocity. (**a**) In the presence of the superfluid motion with velocity $\mathbf{v}_s$, the spectrum of quasiparticles is Doppler shifted. In the fully gapped superconductor in Eq. (5) the spectrum becomes $E(\mathbf{p}) = \pm\sqrt{(p^2/2m - \mu)^2 + |\Delta|^2} + \mathbf{p}\cdot\mathbf{v}_s$. When the flow velocity exceeds the Landau critical velocity, $v_L \approx \Delta/p_F$ if $\Delta \ll \mu$, the positive branch crosses zero energy level. Typically this leads to instability, but in some cases, for example, in superfluid ^{3}He-B, the superfluidity is not destroyed. In this case the Landau critical velocity marks the quantum phase transition at which two Fermi surfaces with $E(\mathbf{p}) = 0$ emerge in the superfluid state (**b**). Liquid remains superfluid, but the density of the fermionic states is nonzero due to Fermi surfaces. Due to that the normal component of the liquid becomes nonzero even at $T = 0$, as a result the density of the superfluid component ρ_s (the prefactor in the superfluid current $\mathbf{j}_s = \rho_s\mathbf{v}_s$) is reduced compared with its value ρ below the threshold. See also Sect. 26.1 in Ref. [6]

vortex lines. In Fig. 4 the two-dimensional system is considered with the saddle point spectrum $E(\mathbf{p}) = p_x^2 - p_y^2 - \mu$. The reconnection quantum transition occurs at $\mu = 0$. The three-dimensional systems, in which the Fermi surface is a 2D vortex sheet in the 4D space (ω, p_x, p_y, p_z), may experience the more complicated topological transitions.

2.3 Metal-superconductor Transition

The transition to superconducting state, even if it occurs at $T = 0$, does not belong to the class of the quantum phase transitions which we discuss in this review, because it is the consequence of the spontaneously broken symmetry and does not occur perturbatively. Let us discuss this transition from the point of view of the momentum-space topology.

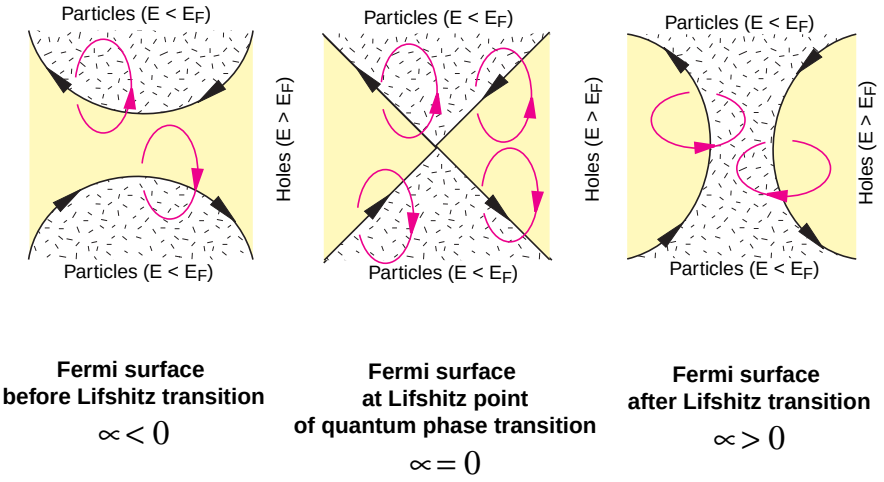

Fig. 4. Lifshitz transition with change of the Fermi surface topology as reconnection of vortex lines in momentum space. The direction of the "circulation" around the vortex lines (*grey arrows*) and "vorticity" along the vortex lines (*black arrows*) are shown

Topology of Gor'kov Function Across the Superconducting Transition

Let us first note that the breaking of $U(1)$ symmetry is not the sufficient condition for superfluidity or superconductivity. For example, the $U(1)$ symmetry of the atoms A which is the result of conservation of the number N_A of A atoms, may be violated simply due to possibility of decay of atom A to atom B. But this does not lead to superfluidity, and the Fermi surface does not disappear. For these two species of atoms the Hamiltonian is 2×2 matrix, such as

$$H = \begin{pmatrix} p^2/2m_A - \mu & \Delta \\ \Delta^* & p^2/2m_B - \mu \end{pmatrix} , \tag{3}$$

where Δ is the matrix element which mixes the atoms A and B. This mixing violates the separate $U(1)$ symmetry for each of the two gases, but the gap does not appear. Zeroes of the energy spectrum found from the nullification of the determinant of the matrix, $(p^2/2m_A - \mu)(p^2/2m_B - \mu) - |\Delta|^2 = 0$, form two Fermi surfaces if $\Delta = 0$, and these Fermi surfaces survive if $\Delta \neq 0$ but is sufficiently small. This is the consequence of topological stability of **p**-space vortices. Each Fermi surface has topological charge $N_1 = 1$, and their sum $N_1 = 2$ is robust to small perturbations.

The non-perturbative phenomenon of superfluidity in the fermionic gas occurs due to Cooper pairing of atoms (electrons), i.e. due to mixing between the particle and hole states. Such mixing requires introduction of the extended matrix Green's function even for a single fermions species. This is the Gor'kov

Green's function which is the matrix in the particle-hole space of the same fermions, i.e. we have effective doubling of the relevant fermionic degrees of freedom for the description of superconductivity. In case of s-wave pairing the Gor'kov Green's function has the following form:

$$G^{-1} = \begin{pmatrix} i\omega - p^2/2m + \mu & \Delta \\ \Delta^* & i\omega + p^2/2m - \mu \end{pmatrix} , \tag{4}$$

Now the energy spectrum

$$E^2 = (p^2/2m - \mu)^2 + |\Delta|^2 \tag{5}$$

has a gap, i.e. the Fermi surface disappears. How does this happen? At $\Delta = 0$ the matrix Green's function describes two species of fermions: particles and holes. The topological charges of the corresponding Fermi surfaces are $N_1 = 1$ for particles and $N_1 = -1$ for holes, with total topological charge $N_1 = 0$. The trivial total topological charge of the Fermi surfaces allows for their annihilation, which just occurs when the mixing matrix element $\Delta \neq 0$ and the energy spectrum becomes fully gapped. Thus the topology of the matrix Gor'kov Green's function G does not change across the superconducting transition.

Topology of Diagonal Green's Function Across the Superconducting Transition

Let us consider what happens with the conventional Green's function across the transition. This is the G_{11} element of the matrix (4):

$$G_{11} = -\frac{i\omega + p^2/2m - \mu}{\omega^2 + (p^2/2m - \mu)^2 + |\Delta|^2} . \tag{6}$$

One can see that it has the same topology in momentum space as the Green's function of normal metal in Eq. (1):

$$G_{11}(\Delta = 0) = \frac{1}{i\omega - p^2/2m + \mu} = -\frac{i\omega + p^2/2m - \mu}{\omega^2 + (p^2/2m - \mu)^2} . \tag{7}$$

Though instead of the pole in Eq. (7) for superconducting state one has zero in Eq. (6) for normal state, their topological charges in Eq. (2) are the same: both have the same vortex singularity with $N_1 = 1$. Thus the topology of the conventional Green's function G_{11} also does not change across the superconducting transition.

So the topology of each of the functions G and G_{11} does not change across the transition. This illustrates again the robustness of the topological charge. But what occurs at the transition? The Green's function G_{11} gives the proper description of the normal state, but it does not provide the complete description of the superconducting state. That is why its zeroes, though have

non-trivial topological charge, bear no information on the spectrum of excitations. On the other hand the matrix Green's function G provides the complete description of the superconducting states, but is meaningless on the normal state side of the transition. Thus the spectrum on two sides of the transition is determined by two different functions with different topological properties. This illustrates the non-perturbative nature of the superconducting transition, which crucially changes the **p**-space topology leading to the destruction of the Fermi surface without conservation of the topological charge across the transition.

Momentum Space Topology in Pseudo-Gap State

Pseudo-gap is the effect of the suppression of the density of states (DOS) at low energy [26]. Let us consider a simple model in which the pseudo-gap behavior of the normal Fermi liquid results from the superfluid/superconducting fluctuations, i.e. in this model the pseudo-gap state is the normal (non-superconducting) state with the virtual superconducting order parameter Δ fluctuating about its equilibrium zero value (see review [27] and Ref. [28]). For simplicity we discuss the extreme case of such state where Δ fluctuates being homogeneous in space. The average value of the off-diagonal element of the Gor'kov functions is zero in this state, $\langle G_{12}\rangle = 0$, and thus the $U(1)$ symmetry remains unbroken. The Green's function of this pseudo-gap state is obtained by averaging of the function G_{11} over the distribution of the uniform complex order parameter Δ:

$$G = \langle G_{11}\rangle = \int d\Delta d\Delta^* P(|\Delta|) \frac{-i\omega - \epsilon}{\omega^2 + \epsilon^2 + |\Delta|^2} \ . \tag{8}$$

Here $\epsilon(\mathbf{p}) = p^2/2m - \mu$ and $P(|\Delta|)$ is the probability of the gap $|\Delta|$. If $P(0) \neq 0$, then in the low-energy limit $\omega^2 + \epsilon^2 \ll \Delta_0^2$, where Δ_0 is the amplitude of fluctuations, one obtains

$$G = \frac{Z}{i\omega - \epsilon} \ , \quad Z \propto \frac{\omega^2 + \epsilon^2}{\Delta_0^2} \ln \frac{\Delta_0^2}{\omega^2 + \epsilon^2} \ . \tag{9}$$

The Green's function has the same topological property as conventional Green's function of metal with Fermi surface at $\epsilon(\mathbf{p}) = 0$, but the suppression of residue Z is so strong, that the pole in the Green's function is transformed to the zero of the Green's function. Because of the topological stability, the singularity of the Green's function at the Fermi surface is not destroyed: the zero is also the singularity and it has the same topological invariant in Eq. (2) as pole. So this model of the Fermi liquid represents a kind of Luttinger or marginal Fermi liquid with a very strong renormalization of the singularity at the Fermi surface.

This demonstrates that the topology of the Fermi surface is the robust property, which does not resolve between different fine structures of the Fermi liquids with different DOS.

Using the continuation of Eq. (9) to the real frequency axis ω, one obtains the density of states in this extreme model of the pseudo-gap:

$$\nu(\omega) = N_0 \int d\epsilon \ \mathrm{Im}G = \pi N_0 \int_0^{\omega} d\epsilon \ \frac{\omega + \epsilon}{\Delta_0^2} = \frac{3\pi}{2} N_0 \frac{\omega^2}{\Delta_0^2} \ , \tag{10}$$

where N_0 is the DOS of the conventional Fermi liquid, i.e. without the pseudo-gap effect. Though this state is non-superfluid and is characterized by the Fermi surface, the DOS at $\omega \ll \Delta_0$ is highly suppressed compared to N_0, i.e. the pseudo-gap effect is highly pronounced. This DOS has the same dependence on ω as that in such superconductors or superfluids in which the gap has point nodes discussed in the next Sect. 3. When the spatial and time variation of the gap fluctuations are taken into account, the pseudo-gap effect would not be so strong.

3 Fermi Points

3.1 Fermi Point as Topological Object

Chiral Fermi Points

The crucial non-perturbative reconstruction of the spectrum occurs at the superfluid transition to ^{3}He-A, where the point nodes emerge instead of the Fermi surface. Since we are only interested in effects determined by the topology and the symmetry of the fermionic Hamiltonian $H(\mathbf{p})$ or Green's function $G(\mathbf{p}, i\omega)$, we do not require a special form of the Green's function and can choose the simplest one with the required topology and symmetry. First, consider the Bogoliubov–Nambu Hamiltonian which qualitatively describes fermionic quasiparticles in the axial state of p–wave pairing. This Hamiltonian can be applied to superfluid ^{3}He-A [4] and also to the p-wave BCS state of ultracold Fermi gas:

$$\begin{aligned} H &= \begin{pmatrix} p^2/2m - \mu & c_\perp \, \mathbf{p} \cdot (\hat{\mathbf{e}}_1 + i \, \hat{\mathbf{e}}_2) \\ c_\perp \, \mathbf{p} \cdot (\hat{\mathbf{e}}_1 - i \, \hat{\mathbf{e}}_2) & -p^2/2m + \mu \end{pmatrix} \\ &= \tau_3 (p^2/2m - \mu) + c_\perp \, \mathbf{p} \cdot (\tau_1 \hat{\mathbf{e}}_1 - \tau_2 \hat{\mathbf{e}}_2), \end{aligned} \tag{11}$$

where τ_1, τ_2 and τ_3 are 2×2 Pauli matrices in Bogoliubov–Nambu particle-hole space, and we neglect the spin structure which is irrelevant for consideration. The orthonormal triad ($\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, $\hat{\mathbf{l}} \equiv \hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2$) characterizes the order parameter in the axial state of triplet superfluid. The unit vector $\hat{\mathbf{l}}$ corresponds to the direction of the orbital momentum of the Cooper pair (or the diatomic molecule in case of BEC); and $c_\perp$ is the speed of the quasiparticles if they propagate in the plane perpendicular to $\hat{\mathbf{l}}$.

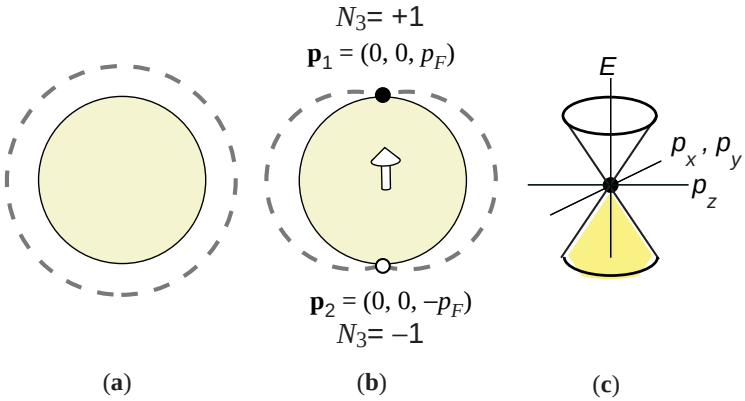

Fig. 5. Angular dependence of the superfluid/superconducting gap (*dashed lines*) at the former Fermi surface (*solid lines*). The gap Δ is (**a**) isotropic in the s-wave state and (**b**) is $\Delta(\theta) = p_F c_\perp \sin\theta$ for the p-wave state in Eq. (12), where θ is the polar angle, and arrow shows the direction of the unit vector $\hat{\mathbf{l}}$. The spectrum of quasiparticles has two nodes at the former Fermi surface: at $\theta = 0$, i.e. at $\mathbf{p}_1 = p_F\hat{\mathbf{l}}$ (*filled circle*) and at $\theta = \pi$, i.e. at $\mathbf{p}_2 = -p_F\hat{\mathbf{l}}$ (*open circle*). Their winding numbers of the Fermi points given by Eq. (13) are correspondingly $N_3 = +1$ and $N_3 = -1$. (**c**) According to Eq. (14), close to the Fermi points the quasiparticle spectrum becomes "relativistic". For the local observer, who measures the spectrum using clocks and rods made of the low-energy fermions, the Hamiltonian for fermions in the vicinity of the point with $N_3 = +1$ is equivalent to the Weyl Hamiltonian for the right handed massless fermions: $H = c\boldsymbol{\sigma}\cdot\mathbf{p}$; and the spectrum has the conical form $E = \pm cp$

The energy spectrum of these Bogoliubov–Nambu fermions is

$$E^2(\mathbf{p}) = \left(\frac{p^2}{2m} - \mu\right)^2 + c_\perp^2 \left(\mathbf{p}\times\hat{\mathbf{l}}\right)^2 . \tag{12}$$

In the BCS regime occuring for positive chemical potential $\mu > 0$, there are two Fermi points in 3D momentum space with $E(\mathbf{p}) = 0$. For the energy spectrum (12), the Fermi points are $\mathbf{p}_1 = p_F\,\hat{\mathbf{l}}$ and $\mathbf{p}_2 = -p_F\,\hat{\mathbf{l}}$, with Fermi momentum $p_F = \sqrt{2m\mu}$ [Fig. 5(**b**)].

For a general system, be it relativistic or nonrelativistic, the topological stability of the Fermi point (the node of the co-dimension 3) is guaranteed by the nontrivial homotopy group $\pi_2(GL(n,\mathbf{C})) = Z$ which describes the mapping of a sphere S^2 embracing the point node to the space of non-degenerate complex matrices [7]. This is the group of integers. The integer valued topological invariant (winding number) can be written in terms of the fermionic propagator $G(i\omega,\mathbf{p})$ as a surface integral in the 4D frequency-momentum space $p_\mu = (\omega,\mathbf{p})$: [6]

$$N_3 \equiv \frac{1}{24\pi^2}\,\epsilon_{\mu\nu\rho\sigma}\ \mathrm{tr}\oint_{\Sigma_a} dS^\sigma G\frac{\partial}{\partial p_\mu}G^{-1}\,G\frac{\partial}{\partial p_\nu}G^{-1}\,G\frac{\partial}{\partial p_\rho}G^{-1} . \tag{13}$$

Here Σ_a is a three-dimensional surface around the isolated Fermi point $p_{\mu a} = (0, \mathbf{p}_a)$ and "tr"stands for the trace over the relevant spin and/or band indices. For the case considered in Eq. (11), the Green's function is $G^{-1}(i\omega, \mathbf{p}) = i\omega - H(\mathbf{p})$; the trace is over the Bogoliubov-Nambu spin; and the two Fermi points $\mathbf{p}_1$ and $\mathbf{p}_2$ have nonzero topological charges $N_3 = +1$ and $N_3 = -1$ [Fig. 6 (*right*)].

We call such Fermi points the chiral Fermi points, because in the vicinity of these point the fermions behave as right-handed or left handed particles (see below). These nodes of co-dimension 3 are the diabolical points – the exceptional degeneracy points of the complex-valued Hamiltonian which depends on the external parameters (see Ref. [29–32]). At these points two different branches of the spectrum touch each other. Topology of these points has been discussed in Ref. [33]. In our case the relevant parameters of the Hamiltonian are the components of momentum $\mathbf{p}$, and we discuss the contact point of branches with positive and negative energies [34]. Topology of the chiral Fermi points in relation to the spectrum of elementary particles has been discussed in Ref. [35].

Emergent Relativity and Chiral Fermions

Close to any of the Fermi points the energy spectrum of fermionic quasiparticles acquires the relativistic form (this follows from the so-called Atiyah-Bott-Shapiro construction [7]). In particular, the Hamiltonian in Eq. (11) and spectrum in Eq. (12) become [6]:

$$H \rightarrow e_k^i \sigma^k (p_i - eA_i) \ , \quad E^2(\mathbf{p}) \rightarrow g^{ik}(p_i - eA_i)(p_k - eA_k) \ . \tag{14}$$

Here the analogue of the dynamic gauge field is $\mathbf{A} = p_F \hat{\mathbf{l}}$; the "electric charge" is either $e = +1$ or $e = -1$ depending on the Fermi point; the matrix e_i^k is the analogue of the dreibein with $g^{ik} = e_j^i e_j^k = \mathrm{diag}(c_\perp^2, c_\perp^2, c_\parallel^2 = p_F^2/m^2)$ playing the role of the effective dynamic metric in which fermions move along the geodesic lines. Fermions in Eq. (14) are chiral: they are right-handed if the determinant of the matrix e_j^i is positive, which occurs at $N_3 = +1$; the fermions are left-handed if the determinant of the matrix e_j^i is negative, which occurs at $N_3 = -1$. For the local observer, who measures the spectrum using clocks and rods made of the low-energy fermions, the Hamiltonian in Eq. (14) is simplified: $H = \pm c\boldsymbol{\sigma} \cdot \mathbf{p}$ [Fig. 5(**c**). Thus the chirality is the property of the behavior in the low energy corner and it is determined by the topological invariant N_3.

Majorana Fermi Point

The Hamiltonians which give rise to the chiral Fermi points with non-zero N_3 are essentially complex matrices. That is why one may expect that in

systems described by real-valued Hamiltonian matrices there are no topologically stable points of co-dimension 3. However, the general analysis in terms of K-theory [7] demonstrates that such points exist and are described by the group Z_2. Let us denote this Z_2 charge as $N_{3\mathrm{M}}$ to distinguish it from the Z charge N_3 of chiral fermions. The summation law for the charge $N_{3\mathrm{M}}$ is $1+1=0$, i.e. two such points annihilate each other. Example of topologically stable massless real fermions is provided by the Majorana fermions [7]. The summation law $1+1=0$ also means that $1=-1$, i.e. the particle is its own antiparticle. This property of the Majorana fermions follows from the topology in momentum space and does not require the relativistic invariance.

Summation Law for Majorana Fermions and Marginal Fermi Point

The summation law $1-1=0$ for chiral fermions and $1+1=0$ for Majorana fermions is illustrated using the following 4×4 Hamiltonian matrix:

$$H = c\tau_1 p_x + c\tau_2\sigma_2 p_y + c\tau_3 p_z \ . \tag{15}$$

This Hamiltonian describes either two chiral fermions or two Majorana fermions. The first description is obtained if one chooses the spin quantization axis along σ_2. Then for the direction of spin $\sigma_2=+1$ this Hamiltonian describes the right-handed fermion with spectrum $E(p)=cp$ whose Fermi point at $\mathbf{p}=0$ has topological charge $N_3=+1$. For $\sigma_2=-1$ one has the left-handed chiral fermion whose Fermi point is also at $\mathbf{p}=0$, but it has the opposite topological charge $N_3=-1$. Thus the total topological charge of the Fermi point at $\mathbf{p}=0$ is $N_3=1-1=0$.

In the other description, one takes into account that the matrix (15) is real and thus can describe the real (Majorana) fermions. In our case the original fermions are complex, and thus we have two real fermions with the spectrum $E(p)=cp$ representing the real and imaginary parts of the complex fermion. Each of the two Majorana fermions has the Fermi (Majorana) point at $\mathbf{p}=0$ where the energy of fermions is zero. Since the Hamiltonian (15) is the same for both real fermions, the two Majorana points have the same topological charge.

Let us illustrate the difference in the summation law for charges N_3 and $N_{3\mathrm{M}}$ by introducing the perturbation $M\sigma_1\tau_2$ to the Hamiltonian (15):

$$H = c\tau_1 p_x + c\tau_2\sigma_2 p_y + c\tau_3 p_z + M\sigma_1\tau_2 \ . \tag{16}$$

Due to this perturbation the spectrum of fermions is fully gapped: $E^2(p)=c^2p^2+M^2$. In the description in terms of the chiral fermions, the perturbation mixes left and right fermions. This leads to formation of the Dirac mass M. The annihilation of Fermi points with opposite charges illustrates the summation law $1-1=0$ for the topological charge N_3.

Let us now consider the same process using the description in terms of real fermions. The added term $M\sigma_1\tau_2$ is imaginary. It mixes the real and imaginary components of the complex fermions, and thus it mixes two Majorana fermions. Since the two Majorana fermions have the same topological charge, $N_{3\mathrm{M}} = 1$, the formation of the gap means that the like charges of the Majorana points annihilate each other. This illustrates the summation law $1 + 1 = 0$ for the Majorana fermions.

In both descriptions of the Hamiltonian (15), the total topological charge of the Fermi or Majorana point at $\mathbf{p} = 0$ is zero. We call such topologically trivial point the marginal Fermi point. The topology does not protect the marginal Fermi point, and the small perturbation can lead to formation of the fully gapped vacuum, unless there is a symmetry which prohibits this.

3.2 Quantum Phase Transition in BCS–BEC Crossover Region

Splitting of Marginal Fermi Point

Let us consider some examples of quantum phase transition goverened by the momentum-space topology of gap nodes, between a fully-gapped vacuum state and a vacuum state with topologically-protected point nodes. In the context of condensed-matter physics, such a quantum phase transition may occur in a system of ultracold fermionic atoms in the region of the BEC–BCS crossover, provided Cooper pairing occurs in the non-s-wave channel. For elementary particle physics, such transitions are related to CPT violation, neutrino oscillations, and other phenomena [18].

Let us start with the topological quantum phase transition involving topologically stable Fermi points [16, 17]. Let us consider what happens with the Fermi points in Eq. (12), when one varies the chemical potential μ. For $\mu > 0$, there are two Fermi points, and the density of fermionic states in the vicinity of Fermi points is $\nu(\omega) \propto \omega^2$. For $\mu < 0$, Fermi points are absent and the spectrum is fully-gapped [Fig. 6]. In this topologically-stable fully-gapped vacuum, the density of states is drastically different from that in the topologically-stable gapless regime: $\nu(\omega) = 0$ for $\omega < |\mu|$. This demonstrates that the quantum phase transition considered is of purely topological origin. The transition occurs at $\mu = 0$, when two Fermi points with $N_3 = +1$ and $N_3 = -1$ merge and form one topologically-trivial Fermi point with $N_3 = 0$, which disappears at $\mu < 0$.

The intermediate state at $\mu = 0$ is marginal: the momentum-space topology is trivial ($N_3 = 0$) and cannot protect the vacuum against decay into one of the two topologically-stable vacua unless there is a special symmetry which stabilizes the marginal node. As we shall see in the Sect. 3.3, the latter takes place in the Standard Model with marginal Fermi point.

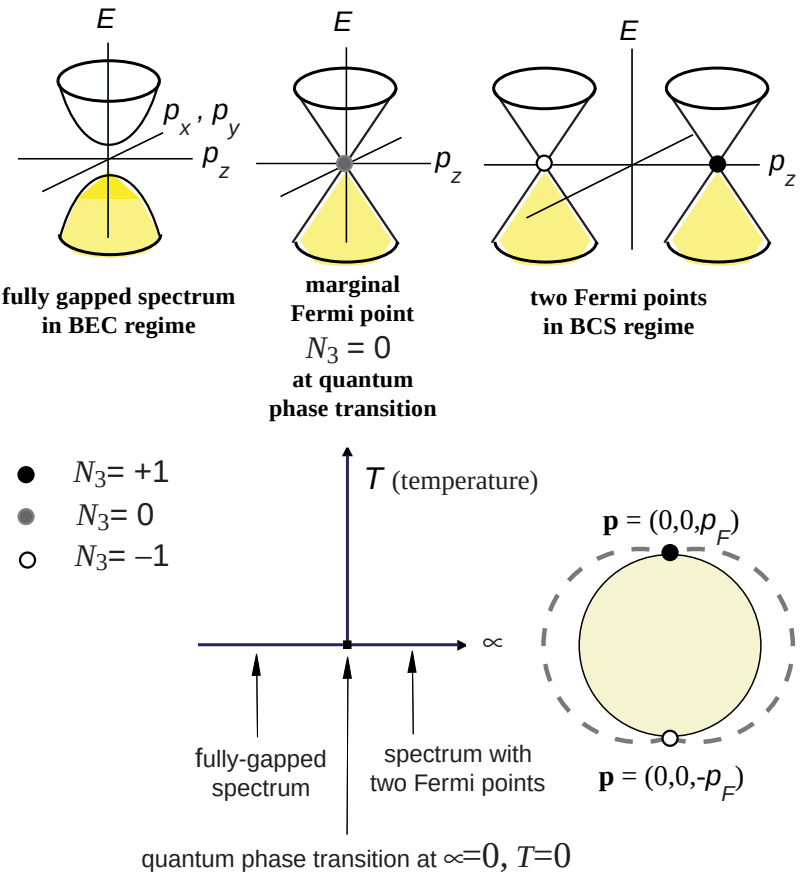

Fig. 6. Quantum phase transition between two p-wave vacua with the same symmetry but of different universality class. In the bottom right corner you find the gap (*dashed line*) in the p-wave state. It depends on the direction in momentum space and becomes zero when $\mathbf{p}$ is along the $\hat{\mathbf{l}}$-vector (along z-axis). At $\mu > 0$, two gap nodes give rise to two zeroes in the spectrum – Fermi points: one with winding number $N_3 = +1$ (*filled circle*) and another with winding number $N_3 = -1$ (*open circle*). The transition occurs when the chemical potential μ in Eq. (11) crosses zero value. The Fermi points merge at $\mu = 0$ forming the marginal (topologically trivial) gap node with $N_3 = 0$ (*grey circle*) and annihilate each other. At $\mu < 0$ the Green's function has no singularities and the quantum vacuum is fully gapped

Transition Involving Multiple Nodes

The Standard Model contains 16 chiral fermions in each generation. The multiple Fermi point may occur in condensed matter too. For systems of cold atoms, an example is provided by another spin-triplet p-wave state, the so-called α-phase. The Bogoliubov–Nambu Hamiltonian which qualitatively describes fermionic quasiparticles in the α–state is given by [3, 4]:

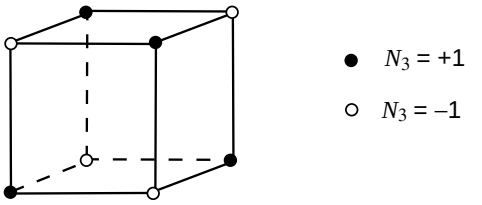

Fig. 7. Fermi points in the α-phase of triplet superfluid/superconductor in the BCS regime

$$H = \begin{pmatrix} p^2/2m - \mu & (\mathbf{\Sigma} \cdot \mathbf{p})\ c_\perp/\sqrt{3} \\ (\mathbf{\Sigma} \cdot \mathbf{p})^\dagger\ c_\perp/\sqrt{3} & -p^2/2m + \mu \end{pmatrix}, \tag{17}$$

with $\mathbf{\Sigma} \cdot \mathbf{p} \equiv \sigma_x p_x + \exp(2\pi i/3)\, \sigma_y p_y + \exp(-2\pi i/3)\, \sigma_z p_z$.

On the BEC side ($\mu < 0$), fermions are again fully-gapped, while on the BCS side ($\mu > 0$), there are 8 topologically protected Fermi points with charges $N_3 = \pm 1$, situated at the vertices of a cube in momentum space [3] [Fig. 7]. The fermionic excitations in the vicinity of these points are left- and right-handed Weyl fermions. At the transition point at $\mu = 0$ these 8 Fermi points merge forming the marginal Fermi point at $\mathbf{p} = 0$.

3.3 Quantum Phase Transitions in Standard Model

Marginal Fermi Point in Standard Model

It is assumed that the Standard Model above the electroweak transition contains 16 chiral fermions in each generation: 8 right-handed fermions with $N_3 = +1$ each and 8 left-handed fermions with $N_3 = -1$ each. If so, then the vacuum of the Standard Model above the electroweak transition is marginal: there is a multiply degenerate Fermi point at $\mathbf{p} = 0$ with the total topological charge $N_3 = +8 - 8 = 0$ [Fig. 8(**a**)]. This vacuum is therefore the intermediate state between two topologically-stable vacua: the fully-gapped vacuum in Fig. 8(**b**); and the vacuum with topologically-nontrivial Fermi points in Fig. 8(**c**).

The absence of the topological stability means that even the small mixing between the fermions leads to annihilation of the Fermi point. In the Standard Model, the proper mixing which leads to the fully gapped vacuum is prohibited by symmetries, namely the continuous electroweak $U(1) \times SU(2)$ symmetry (or the discrete symmetry discussed in Sect. 12.3.2 of [6]) and the CPT symmetry. (Marginal gapless fermions emerging in spin systems were discussed in [36]. These massless Dirac fermions protected by symmetry differ from the chiral fermions of the Standard Model. The latter cannot be represented in terms of massless Dirac fermions, since there is no symmetry between left and right fermions in Standard Model.)

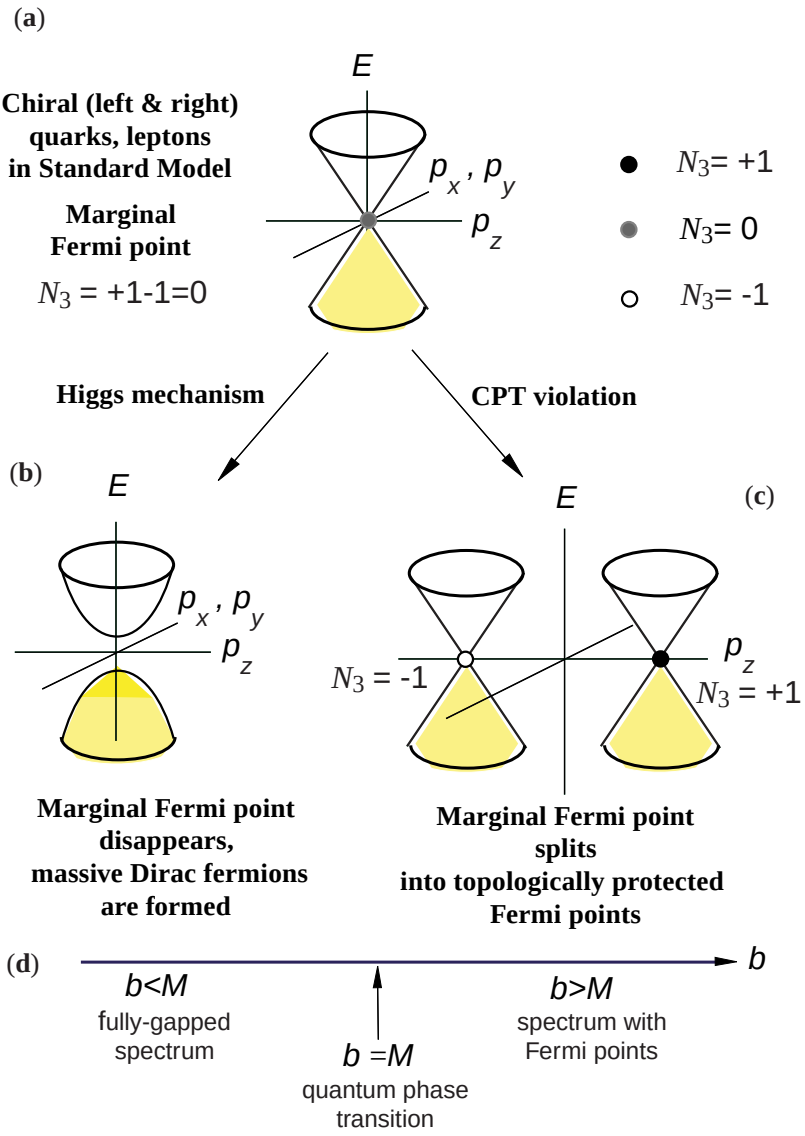

Fig. 8. Two scenarios of annihilation of marginal Fermi point (**a**) in Standard Model of strong and electroweak interactions. Higgs mechanism leads to Dirac mass and thus to the fully gapped vacuum (**b**), while CPT violation leads to splitting of Fermi points (**c**). In the bottom edge you find the quantum phase transition in the model in Eq. (20) when the CPT violating parameter $b \equiv |\mathbf{b}|$ crosses the Dirac mass M

Explicit violation or spontaneous breaking of electroweak or CPT symmetry transforms the marginal vacuum of the Standard Model into one of the two topologically-stable vacua. If, for example, the electroweak symmetry is broken, the marginal Fermi point disappears and the fermions become massive [Fig. 8(**b**)]. This is assumed to happen below the symmetry breaking electroweak transition caused by Higgs mechanism where quarks and charged leptons acquire the Dirac masses. If, on the other hand, the CPT symmetry is violated, the marginal Fermi point splits into topologically-stable Fermi points which protect chiral fermions [Fig. 8(**c**)]. One can speculate that in the Standard Model the latter happens with the electrically neutral leptons, the neutrinos [18, 37].

Quantum Phase Transition with Splitting of Fermi Points

Let us consider this scenario on a simple example of a marginal Fermi point describing a *single* pair of relativistic chiral fermions, that is, one right-handed fermion and one left-handed fermion. These are Weyl fermions with Hamiltonians $H_{\rm right} = \boldsymbol{\sigma} \cdot \mathbf{p}$ and $H_{\rm left} = -\boldsymbol{\sigma} \cdot \mathbf{p}$, where $\boldsymbol{\sigma}$ denotes the triplet of spin Pauli matrices. Each of these Hamiltonians has a topologically-stable Fermi point at $\mathbf{p} = 0$. The corresponding inverse Green's functions are given by

$$G^{-1}_{\rm right}(i\omega, \mathbf{p}) = i\omega - \boldsymbol{\sigma} \cdot \mathbf{p} \,,$$

$$G^{-1}_{\rm left}(i\omega, \mathbf{p}) = i\omega + \boldsymbol{\sigma} \cdot \mathbf{p} \,. \tag{18}$$

The positions of the Fermi points coincide, $\mathbf{p}_1 = \mathbf{p}_2 = 0$, but their topological charges (13) are different. For this simple case, the topological charge equals the chirality of the fermions, $N_3 = C_a$ (i.e., $N_3 = +1$ for the right-handed fermion and $N_3 = -1$ for the left-handed one). The total topological charge of the Fermi point at $\mathbf{p} = 0$ is therefore zero.

The splitting of this marginal Fermi point can be described by the Hamiltonians $H_{\rm right} = \boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{p}_1)$ and $H_{\rm left} = -\boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{p}_2)$, with $\mathbf{p}_1 = -\mathbf{p}_2 \equiv \mathbf{b}$ from momentum conservation. The real vector $\mathbf{b}$ is assumed to be odd under CPT, which introduces CPT violation into the physics. The 4×4 matrix of the combined Green's function has the form

$$G^{-1}(i\omega, \mathbf{p}) = \begin{pmatrix} i\omega - \boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{b}) & 0 \\ 0 & i\omega + \boldsymbol{\sigma} \cdot (\mathbf{p} + \mathbf{b}) \end{pmatrix} . \tag{19}$$

Equation (13) shows that $\mathbf{p}_1 = \mathbf{b}$ is the Fermi point with topological charge $N_3 = +1$ and $\mathbf{p}_2 = -\mathbf{b}$ the Fermi point with topological charge $N_3 = -1$.

Let us now consider the more general situation with both the electroweak and CPT symmetries broken. Due to breaking of the electroweak symmetry the Hamiltonian acquires the off-diagonal term (mass term) which mixes left and right fermions

$$H = \begin{pmatrix} \boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{b}) & M \\ M & -\boldsymbol{\sigma} \cdot (\mathbf{p} + \mathbf{b}) \end{pmatrix} . \tag{20}$$

The energy spectrum of Hamiltonian (20) is

$$E_{\pm}^2(\mathbf{p}) = M^2 + |\mathbf{p}|^2 + b^2 \pm 2\, b \sqrt{M^2 + \left(\mathbf{p} \cdot \hat{\mathbf{b}}\right)^2} , \tag{21}$$

with $\hat{\mathbf{b}} \equiv \mathbf{b}/|\mathbf{b}|$ and $b \equiv |\mathbf{b}|$.

Allowing for a variable parameter b, one finds a quantum phase transition at $b = M$ between the fully-gapped vacuum for $b < M$ and the vacuum with two isolated Fermi points for $b > M$ [Fig. 8(**d**)]. These Fermi points are situated at

$$\mathbf{p}_1 = +\hat{\mathbf{b}} \sqrt{b^2 - M^2} ,$$

$$\mathbf{p}_2 = -\hat{\mathbf{b}} \sqrt{b^2 - M^2} . \tag{22}$$

Equation (13), now with a trace over the indices of the 4×4 Dirac matrices, shows that the Fermi point at $\mathbf{p}_1$ has topological charge $N_3 = +1$ and thus the right-handed chiral fermions live in the vicinity of this point. Near the Fermi point at $\mathbf{p}_2$ with the charge $N_3 = -1$, the left-handed fermions live. The magnitude of the splitting of the two Fermi points is given by $2\sqrt{b^2 - M^2}$. At the quantum phase transition $b = M$, the Fermi points with opposite charge annihilate each other and form a marginal Fermi point at $\mathbf{p} = 0$. The momentum-space topology of this marginal Fermi point is trivial (the topological invariant $N_3 = +1 - 1 = 0$).

Fermi Surface with Global Charge N_3 and Quantum Phase Transition with Transfer of N_3

Extension of the model (20) by introducing the time like parameter b_0

$$H = \begin{pmatrix} \boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{b}) - b_0 & M \\ M & -\boldsymbol{\sigma} \cdot (\mathbf{p} + \mathbf{b}) + b_0 \end{pmatrix} , \tag{23}$$

demonstrates another type of quantum phase transitions [18] shown in Fig. 9.

At $b_0 \neq 0$, Fermi points which exist at $b_0 = 0$, $b > M$ transform to the closed Fermi surfaces. These Fermi surfaces in addition to the local charge N_1 have the global topological invariant N_3 inherited from the original Fermi points. The global charge N_3 is defined by the same Eq. (13), but with a three-dimensional surface Σ_a around the whole Fermi surface. On the line of the quantum phase transition, $b^2 - b_0^2 = M^2$ (*dashed line*), two Fermi surfaces contact each other at the point $\mathbf{p} = 0$. At that moment, the topological charge N_3 is transferred between the Fermi surfaces through the point of the contact. Above the transition line, the global charges of Fermi surfaces are zero. At the quantum phase transition at $b = M$ (*thick vertical line*) these Fermi surfaces shrink to the points; and since the N_3 topology of these points is trivial they disappear at $b < M$ where the state is fully gapped.

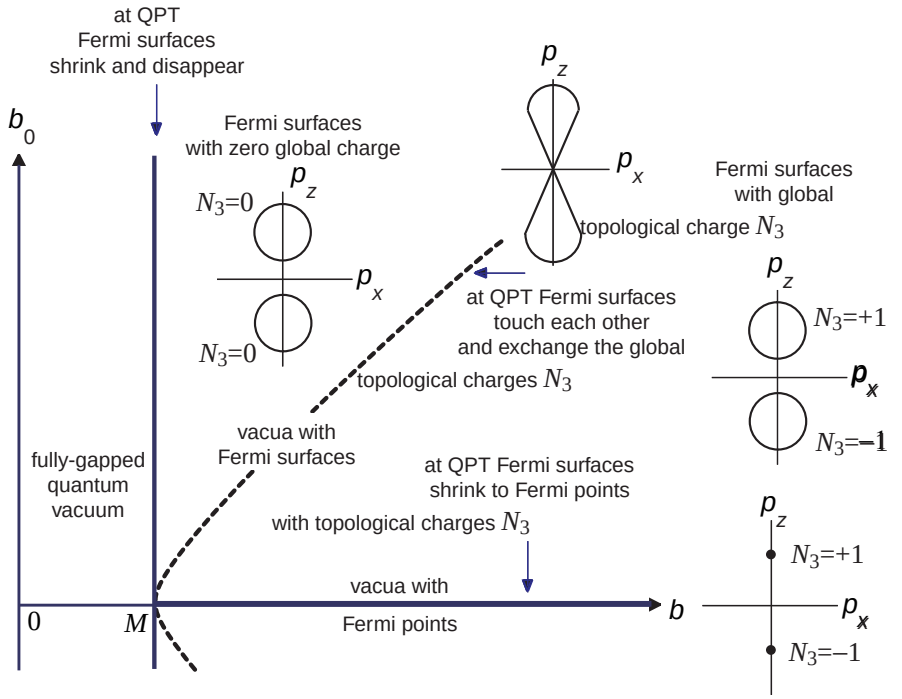

Fig. 9. Topological quantum phase transitions in the model (23). The vacua at $b_0 \neq 0$ and $b > M$ have Fermi surfaces. At $b^2 > b_0^2 + M^2$, these Fermi surfaces have nonzero global topological charges $N_3 = +1$ and $N_3 = -1$. At the quantum phase transition occurring on the line $b_0 = 0$, $b > M$ (*thick horizontal line*) the Fermi surfaces shrink to the Fermi points with nonzero N_3. At $M^2 < b^2 < b_0^2 + M^2$ the global topology of the Fermi surfaces is trivial, $N_3 = 0$. At the quantum phase transition occurring on the line $b = M$ (*thick vertical line*), the Fermi surfaces shrink to the points; and since their global topology is trivial the zeroes disappear at $b < M$ where the vacuum is fully gapped. The quantum phase transition between the Fermi surfaces with and without topological charge N_3 occurs at $b^2 = b_0^2 + M^2$ (*dashed line*). At this transition, the Fermi surfaces touch each other, and their topological charges annihilate each other

Standard Model with Chiral Fermi Point

In the above consideration we assumed that the Fermi point in the Standard Model above the electroweak energy scale is marginal, i.e. its total topological charge is $N_3 = 0$. Since the topology does not protect such a point, everything depends on symmetry, which is a more subtle issue. In principle, one may expect that the vacuum is always fully gapped. This is supported by the Monte-Carlo simulations which suggest that in the Standard Model there is no second-order phase transition at finite temperature, instead one has either the first-order electroweak transition or crossover depending on the ratio of

masses of the Higgs and gauge bosons [38]. This would actually mean that the fermions are always massive.

Such scenario does not contradict to the momentum-space topology, only if the total topological charge N_3 is zero. However, from the point of view of the momentum-space topology there is another scheme of the description of the Standard Model. Let us assume that the Standard Model follows from the GUT with $SO(10)$ group. In this scheme, the 16 Standard Model fermions form at high energy the 16-plet of the $SO(10)$ group. All the particles of this multiplet are left-handed fermions. These are: four left-handed $SU(2)$ doublets (neutrino-electron and 3 doublets of quarks) + eight left $SU(2)$ singlets of anti-particles (antineutrino, positron and 6 anti-quarks). The total topological charge of the Fermi point at $\mathbf{p} = 0$ is $N_3 = -16$, and thus such a vacuum is topologically stable and is protected against the mass of fermions. This topological protection works even if the $SU(2) \times U(1)$ symmetry is violated perturbatively, say, due to the mixing of different species of the 16-plet. Mixing of left leptonic doublet with left singlets (antineutrino and positron) violates $SU(2) \times U(1)$ symmetry, but this does not lead to annihilation of Fermi points and mass formation since the topological charge N_3 is conserved.

We discussed the similar situation in the Sect. 2.3 for the case of the Fermi surface, and found that if the total topological charge of the Fermi surfaces is non-zero, the gap cannot appear perturbatively. It can only arise due to the crucial reconstruction of the fermionic spectrum with effective doubling of fermions. In the same manner, in the $SO(10)$ GUT model the mass generation can only occur non-perturbatively. The mixing of the left and right fermions requires the introduction of the right fermions, and thus the effective doubling of the number of fermions. The corresponding Gor'kov's Green's function in this case will be the $(16 \times 2) \times (16 \times 2)$ matrix. The nullification of the topological charge $N_3 = -16$ occurs exactly in the same manner, as in superconductors. In the extended (Gor'kov) Green's function formalism appropriate below the transition, the topological charge of the original Fermi point is annihilated by the opposite charge $N_3 = +16$ of the Fermi point of "holes"(right-handed particles).

This demonstrates that the mechanism of generation of mass of fermions essentially depends on the momentum space topology. If the Standard Model originates from the $SO(10)$ group, the vacuum belongs to the universality class with the topologically non-trivial chiral Fermi point (i.e. with $N_3 \neq 0$), and the smooth crossover to the fully-gapped vacuum is impossible. On the other hand, if the Standard Model originates from the left-right symmetric Pati–Salam group such as $SU(2)_L \times SU(2)_R \times SU(4)$, and its vacuum has the topologically trivial (marginal) Fermi point with $N_3 = 0$, the smooth crossover to the fully-gapped vacuum is possible.

Chiral Anomaly

Since chiral Fermi points in condensed matter and in Standard Model are described by the same momentum-space topology, one may expect common properties. An example of such a common property would be the axial or chiral anomaly. For quantum anomalies in (3+1)–dimensional systems with Fermi points and their dimensional reduction to (2+1)–dimensional systems, see, e.g., Ref. [6] and references therein. In superconducting and superfluid fermionic systems the chiral anomaly is instrumental for the dynamics of vortices. In particular, one of the forces acting on continuous vortex-skyrmions in superfluid ^{3}He-A is the result the anomalous production of the fermionic charge from the vacuum decsribed by the Adler-Bell-Jackiw equation [39].

4 Fermi Lines

In general the zeroes of co-dimension 2 (nodal lines in 3D momentum space or point nodes in 2D momentum space) do not have the topological stability. However, if the Hamiltonian is restricted by some symmetry, the topological stability of these nodes is possible. The nodal lines do not appear in spin-triplet superconductors, but they may exist in spin-singlet superconductors [3, 40]. The analysis of topological stability of nodal lines in systems with real fermions was done by Horava [7].

4.1 Nodes in High-T_c Superconductors

An example of point nodes in 2D momentum space is provided by the layered quasi-2D high-T_c superconductor. In the simplest form the 2D Bogoliubov–Nambu Hamiltonian is

$$H = \tau_3 \left(\frac{p_x^2 + p_y^2}{2m} - \mu \right) + a\tau_1 (p_x^2 - \lambda p_y^2) \ . \tag{24}$$

In case of tetragonal crystal symmetry one has either the pure s-wave state with $\lambda = -1$ $(p_x^2 + p_y^2)$ or the pure d-wave state with $\lambda = +1$ $(p_x^2 - p_y^2)$. But in case of orthorhombic crystal these two states are not distinguishible by symmetry and thus the general order parameter is represented by the $s + d$ combination, i.e. in the orthorhombic crystal one always has $|\lambda| \neq 1$. For example, experiments in high-T_c cuprate $YBa_2Cu_3O_7$ suggest that $\lambda \sim 0.7$ in this compound [41].

At $\mu > 0$ and $\lambda > 0$, the energy spectrum contains 4 point nodes in 2D momentum space (or four Fermi-lines in the 3D momentum space):

$$p_x^a = \pm p_F \sqrt{\frac{\lambda}{1+\lambda}} \ , \quad p_y^a = \pm p_F \sqrt{\frac{1}{1+\lambda}} \ , \quad p_F^2 = 2\mu m \ . \tag{25}$$

The problem is whether these nodes survive or not if we extend Eq. (24) to the more general Hamiltonian obeying the same symmetry. The important property of this Hamiltonian is that, as distinct from the Hamiltonian (11), it obeys the time reversal symmetry which prohibits the imaginary τ_2-term. In the spin singlet states the Hamiltonian obeying the time reversal symmetry must satisfy the equation $H^*(-\mathbf{p}) = H(\mathbf{p})$. The general form of the 2×2 Bogoliubov-Nambu spin-singlet Hamiltonian satisfying this equation can be expressed in terms of the 2D vector $\mathbf{m}(\mathbf{p}) = (m_x(\mathbf{p}), m_y(\mathbf{p}))$:

$$H = \tau_3 m_x(\mathbf{p}) + \tau_1 m_y(\mathbf{p}) \ . \tag{26}$$

Using this vector one can construct the integer valued topological invariant – the contour integral around the point node in 2D momentum space or around the nodal line in 3D momentum space:

$$N_2 = \frac{1}{2\pi} \oint dl \ \hat{\mathbf{z}} \cdot \left(\hat{\mathbf{m}} \times \frac{d\hat{\mathbf{m}}}{dl} \right) \ , \tag{27}$$

where $\hat{\mathbf{m}} \equiv \mathbf{m}/|\mathbf{m}|$. This is the winding number of the plane vector $\mathbf{m}(\mathbf{p})$ around a vortex line in 3D momentum space or around a point vortex in 2D momentum space. The winding number is robust to any change of the Hamiltonian respecting the time reversal symmetry, and this is the reason why the node is stable.

All four nodes in the above example of Eq. (24) are topologically stable, since nodes with equal signs ($++$ and $--$) have winding number $N_2 = +1$, while the other two nodes have winding number $N_2 = -1$ [Fig. 10].

4.2 Z_2-Lines

Now let us consider the stability of these nodes using the general topological analysis (the so-called K-theory, see [7]). For the general $n \times n$ real matrices the classification of the topologically stable nodal lines in 3D momentum space (zeroes of co-dimension 2) is given by the homotopy group $\pi_1(GL(n, \mathbf{R}))$ [7]. It determines classes of mapping of a contour S^1 around the nodal line (or around a point in the 2D momentum space) to the space of non-degenerate real matrices. The topology of nodes depends on n. If $n = 2$, the homotopy group for lines of nodes is $\pi_1(GL(2, \mathbf{R})) = Z$, it is the group of integers in Eq. (27) obeying the conventional summation $1 + 1 = 2$. However, for larger $n \geq 3$ the homotopy group for lines of nodes is $\pi_1(GL(n, \mathbf{R})) = Z_2$, which means that the summation law for the nodal lines is now $1 + 1 = 0$, i.e. two nodes with like topological charges annihilate each other. These nodes of co-dimension 2 are similar to the points of degeneracy of the energy spectrum of the real-valued Hamiltonian which depends on the external parameters (see Ref. [29, 31, 32]).

The equation (24) is the 2×2 Hamiltonian for the complex fermionic field. But each complex field consists of two real fermionic fields. In terms of the real

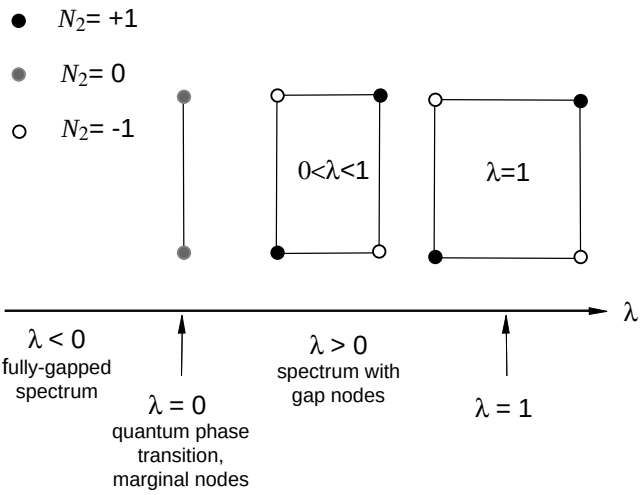

Fig. 10. Quantum phase transition by change of anisotropy parameter λ in Eq. (24) for superconductors in the $d+s$ state. At $\lambda > 0$ the 2D spectrum has 4 nodes: two with topological charge $N_2 = +1$ (*filled circles*) and two with topological charge $N_2 = -1$ (*open circles*). At $\lambda = 0$, points with opposite charges merge forming two marginal nodes with $N_2 = 0$ (*grey circles*). The marginal (topologically trivial) nodes disappear at $\lambda < 0$ leaving the fully gapped vacuum

fermions, this Hamiltonian is the 4×4 matrix and thus all the nodes must be topologically unstable. What keeps them alive is the time reversal symmetry, which does not allow to mix real and imaginary components of the complex field. As a result, the two components are independent; they are described by the same 2×2 Hamiltonian (24); they have zeroes at the same points; and these zeroes are described by the same topological invariants.

If we allow mixing between real and imaginary components of the spinor by introducing the imaginary perturbation to the Hamiltonian, such as $M\tau_2$, the summation law $1 + 1 = 0$ leads to immediate annihilation of the zeroes situated at the same points. As a result the spectrum becomes fully gapped:

$$E^2(\mathbf{p}) = \left(\frac{p_x^2 + p_y^2}{2m} - \mu \right)^2 + a^2(p_x^2 - \lambda p_y^2)^2 + M^2 \ . \tag{28}$$

Thus to destroy the nodes of co-dimension 2 occurring for 2×2 real-valued Hamiltonian (24) describing complex fermions it is enough to violate the time reversal symmetry.

How to destroy the nodes if the time reversal symmetry is obeyed which prohibits mixing? One possibility is to deform the order parameter in such a way that the nodes with opposite N_2 merge and then annihilate each other forming the fully gapped state. In this case, at the border between the state with nodes and the fully gapped state the quantum phase transition occurs

(see Sect. 4.5). This type of quantum phase transition which involves zeroes of co-dimension 2 was also discussed in Ref. [24].

Another possibility is to increase the dimension of the matrix from 2×2 to 4×4. Let us consider this case.

4.3 Gap Induced by Interaction Between Layers

High-T_c superconductors typically have several superconducting cuprate layers per period of the lattice, that is why the consideration of two layers which are described by 4×4 real Hamiltonians is well justified. Let us start again with 2×2 real matrix H, and choose for simplicity the easiest form for the vector $\mathbf{m}(\mathbf{p})$. For $\mathbf{m}(\mathbf{p}) = \mathbf{p} = (p_x, p_y)$ the Hamiltonian is

$$H = \tau_3 p_x + \tau_1 p_y \ . \tag{29}$$

The node which we are interested in is at $p_x = p_y = 0$ and has the topological charge (winding number) $N_2 = 1$ in Eq. (27). The Dirac-type Hamiltonian (29) and the corresponding nodes of co-dimension 2 are relevant for electrons leaving in the 2D carbon sheet known as graphene [42–45].

Let us now introduce two bands or layers whose Hamiltonians have opposite signs:

$$H_{11} = \tau_3 p_x + \tau_1 p_y \ , \quad H_{22} = -\tau_3 p_x - \tau_1 p_y \ , \tag{30}$$

Each Hamiltonian has a node at $p_x = p_y = 0$. In spite of the different signs of the Hamiltonian, the nodes have same winding number $N_2 = 1$: in the second band one has $\mathbf{m}_2(\mathbf{p}) = -\mathbf{m}_1(\mathbf{p})$, but $N_2(\mathbf{m}) = N_2(-\mathbf{m})$ according to Eq. (27).

The Hamiltonians (29) and (30) can be now combined in the 4×4 real Hamiltonian:

$$H = \sigma_3(\tau_3 p_x + \tau_1 p_y) \ , \tag{31}$$

where $\boldsymbol{\sigma}$ matrices operate in the 2-band space. The Hamiltonian (31) has two nodes: one is for projection $\sigma_3 = 1$ and another one – for the projection $\sigma_3 = -1$. Their positions in momentum space and their topological charges coincide. Let us now add the term with σ_1, which mixes the two bands without violation of the time reversal symmetry:

$$H = \sigma_3(\tau_3 p_x + \tau_1 p_y) + \sigma_1 m \ . \tag{32}$$

The spectrum becomes fully gapped, $E^2 = p^2 + m^2$, i.e. the two nodes annililate each other. Since the nodes have the same winding number N_2, this means that the summation law for these nodes is $1+1 = 0$. Thus the zeroes of co-dimension 2 (nodal points in 2D systems or the nodal lines in the 3D systems) which appear in the 4×4 (and higher) real Hamiltonians are described by the Z_2-group. The discussion of the Z_2 nodes in high-T_c materials, polar state in p-wave pairing and mixed singlet-triplet superconducting states can be found in Ref. [46].

The above example demonstrated how in the two band systems (or in the double layer systems) the interaction between the bands (layers) induces the annihilation of likewise nodes and formation of the fully gapped state. Experiments on the graphite film with two graphene layers demonstrate that the spectrum of quasiparticles is essentially different from that in a single carbon sheet [44]. From the detailed calculations [45] it follows that the gap in the spectrum emerges in the graphite bilayer at the neutrality point, illustrating the rule $1 + 1 = 0$ for the Z_2 nodes of co-dimension 2.

Applying this to the high-T_c materials with 2, 3 or 4 cuprate layers per period, one concludes that the interaction between the layers can in principle induce a small gap even in a pure d-wave state. However, this does not mean that such destruction of the Fermi lines necessarily occurs. The interaction between the bands (layers) can lead to splitting of nodes, which then will occupy different positions in momentum space and thus cannot annihilate (this splitting of nodes has been observed in the bilayer cuprate $Bi_2Sr_2CaCu_2O_{8+\delta}$ [47]). Which of the two scenarios occurs – gap formation or splitting of nodes – depends on the parameters of the system. Changing these parameters one can produce the topological quantum phase transition from the fully gapped vacuum state to the vacuum state with pairs of nodes, as we discussed for the case of nodes with co-dimension 3 in Sect. 3.

4.4 Reentrant Violation of "Special Relativity" in Bilayer Graphene

There still can be some discrete symmetry which forbids the annihilation of nodes of co-dimension 2, say, the symmetry between the two layers which forbids the rule $1+1 = 0$. For example, if the Hamiltonian still anti-commutes with some matrix, say, with τ_2-matrix, there is a generalization of the integer valued invariant in Eq. (27) to the $2n \times 2n$ real Hamiltonian (see also [24]):

$$N_2 = -\frac{1}{4\pi i} \text{ tr } \oint dl \ \tau_2 H^{-1} \nabla_l H \ . \tag{33}$$

Since the summation law for this N_2 charge is $1 + 1 = 2$, the nodes with $N_2 = 1$ present at each layer do not annihilate each other if the interaction term preserves the symmetry. In this case the spectrum of the bilayer system remains gapless.

Let us now consider the gapless spectrum in such bilayer material. We start again with the Hamiltonian in Eq. (31), which describes gapless Dirac quasiparticles living in two independent layers, and add the interaction between them which does not violate the τ_2-symmetry:

$$H = \sigma_3(\tau_3 p_x + \tau_1 p_y) + m(\tau_1\sigma_1 - \tau_3\sigma_2) \ . \tag{34}$$

The energy spectrum becomes

$$E_+ = \pm\left(\sqrt{m^2+p^2}+m\right)\ , \quad E_- = \pm\left(\sqrt{m^2+p^2}-m\right)\ . \tag{35}$$

Without interaction, i.e. at $m = 0$, the quasiparticles represent two Dirac fermions with the topological charges $N_2 = 1$ each. Since the Hamiltonian (34) anti-commutes with the τ_2-matrix, the total topological charge N_2 must be conserved even at $m \neq 0$. Thus the total charge for quasiparticles must be $N_2 = 2$. However it is now distributed between the branches of quasiparticle spectrum in the following manner. For $m > 0$, the quasiparticles with energy E_+ acquire the trivial topological charge $N_2 = 0$, that is why their spectrum becomes fully gapped: $E_+(p \ll m) \approx \pm 2m$. The quasiparticles with energy E_- have the rest nonzero topological charge $N_2 = 2$, and thus they must be gapless. The energy spectrum of these gapless fermions with $N_2 = 2$ is exotic: at $p \ll m$ the spectrum becomes that of classical particles with positive and negative masses, $E_- \approx \pm p^2/2m$; in the region $p \gg m$ it is relativistic $E \approx \pm p$; and finally the relativistic invariance is violated again at high p of order of inverse inter-atomic distance. When the parameter m crosses zero, the quantum phase transition occurs.

It is important that the exotic branch with $N_2 = 2$ contains only single fermionic species, i.e. it cannot split into two fermions with $N_2 = 1$ each. That is why the quadratic law for the spectrum of exotic fermions is generic, provided that the proper symmetry of the Hamiltonian is obeyed. The same spectrum (35) takes place for quasiparticles in the carbon film consisting of two graphene sheets: it occurs in some range of parameters of the system where terms in the Hamiltonian, which violate the τ_2-symmetry and induce the gap in the spectrum, are small and can be neglected [45]. Exotic fermions with parabolic spectrum lead to the unconventional quantum Hall effect [45], which has been observed in the bilayer graphene [44].

All this shows that the stability of and the summation law for the nodal lines depend on the type of discrete symmetry which protects the topological stability. The integer valued topological invariants protected by discrete or continuous symmetry were discussed in Chap. 12 of the book [6].

If the symmetry is obeyed we have the following situation. Fermions with the elementary topological charge, $N_2 = \pm 1$, are necessarily relativistic in the low-energy corner, according to the Atiyah-Bott-Shapiro construction. However, even a very small interaction between two species with $N_2 = +1$ each may produce the exotic fermions, which are classical. In this scenario the Lorentz invariance is violated both at very high and at very low energies, therefore the term "reentrant violation of special relativity'.

Similar reentrant violation of Lorentz invariance in the 3D vacua may occur for the Fermi points of co-dimension 3 described by the topological charge N_3 [6,48]. Let us suppose that the Standard Model is an effective theory, and that the right-handed neutrinos are absent in this theory. The left-handed neutrino, which has $N_3 = -1$, is necessarily massless. Its spectrum is necessarily relativistic in the low-energy corner, and thus the Lorentz invariance emerges at low energies. Now let us consider two flavors of the left-handed neutrino

– electron and muon neutrinos with $N_3 = -1$ each. Since the theory is effective, mixing between the flavors is not prohibited, though it is very small. The mixing may lead to the formation of the exotic non-relativistic neutrino with $N_3 = -2$ and the massive neutrino with $N_3 = 0$. In the particular model discussed in Refs. [6, 48], the corresponding spectrum of two neutrino flavors is

$$E_{\pm}^2 = p_z^2 + \left(\sqrt{m^2 + p_x^2 + p_y^2} \pm m\right)^2 . \tag{36}$$

At $p_z = 0$, this 3D spectrum transforms to the 2D spectrum in Eq. (35). The magnitude m of the splitting of the neutrino spectrum has been discussed in Ref. [49].

4.5 Quantum Phase Transition in High-T_c Superconductor

Let us return to the 2×2 real Hamiltonian (24) and consider what happens with gap nodes when one changes the asymmetry parameter λ. When λ crosses zero there is a quantum phase transition at which nodes in the spectrum annihilate each other and then the fully gapped spectrum develops [Fig. 10]. Note that there is no symmetry change across the phase transtion.

The similar quantum phase transition from gapless to gapped state without change of symmetry also occurs when μ crosses zero. This scenario can be realized in the BEC–BCS crossover region, see [21–23].

The presence of the gap nodes in high-T_c superconductors is indicated by the measurement of the field dependence of electronic specific heat C at low temperatures. If the superconducting state is fully gapped, then $C \propto H$; while if there are point nodes in 2D momentum space then the heat capacity is nonlnear, $C \propto \sqrt{H}$ [50]. An unusual behavior of C in high-T_c cuprate $Pr_{2-x}Ce_xCuO_{4-\delta}$ has been reported in Ref. [51]. It was found that the field dependence of electronic specific heat is linear at $T = 2\,\mathrm{K}$, and non-linear at $T \geq 3\mathrm{K}$. If so, this behavior could be identified with the quantum phase transition from gapped to gapless state, which is smeared due to finite temperature. However, the more accurate measurements have not confirmed the change of the regime: the nonlinear behavior $C \propto \sqrt{H}$ continues below $T = 2\,\mathrm{K}$ [52].

5 Topological Transitions in Fully Gapped Systems

5.1 Skyrmion in 2-Dimensional Momentum Space

The fully gapped ground states (vacua) in 2D systems or in quasi-2D thin films, though they do not have zeroes in the energy spectrum, can also be topologically non-trivial. They are characterized by the invariant which is the dimensional reduction of the topological invariant for the Fermi point in Eq. (13) [53, 54]:

$$\tilde{N}_3 = \frac{1}{24\pi^2} e_{\mu\nu\lambda}\ \mathbf{tr} \int d^2p d\omega\ G\partial_{p_\mu} G^{-1} G\partial_{p_\nu} G^{-1} G\partial_{p_\lambda} G^{-1}\ . \qquad (37)$$

For the fully gapped vacuum, there is no singularity in the Green's function, and thus the integral over the entire 3-momentum space $p_\mu = (\omega, p_x, p_y)$ is well determined. If a crystalline system is considered the integration over (p_x, p_y) is bounded by the Brillouin zone.

An example is provided by the 2D version of the Hamiltonian (11) with $\hat{\mathbf{l}} = \hat{\mathbf{z}}$, $\hat{\mathbf{e}}_1 = \hat{\mathbf{x}}$, $\hat{\mathbf{e}}_2 = \hat{\mathbf{y}}$. Since for 2D case one has $p^2 = p_x^2 + p_y^2$, the quasiparticle energy (12)

$$E^2(\mathbf{p}) = \left(\frac{p_x^2 + p_y^2}{2m} - \mu \right)^2 + c^2(p_x^2 + p_y^2) \qquad (38)$$

is nowhere zero except for $\mu = 0$. The Hamiltonian (11) can be written in terms of the three-dimensional vector $\mathbf{g}(p_x, p_y)$:

$$\mathcal{H} = \tau_i g_i(\mathbf{p})\ , \quad g_3 = \frac{p_x^2 + p_y^2}{2m} - \mu\ ,\ g_1 = cp_x\ ,\ g_2 = -cp_y\ . \qquad (39)$$

For $\mu > 0$ the distribution of the unit vector $\hat{\mathbf{g}}(p_x, p_y) = \mathbf{g}/|\mathbf{g}|$ in the momentum space has the same structure as the skyrmion in real space (see Fig. 11). The topological invariant for this momentum-space skyrmion is given by Eq. (37) which can be rewritten in terms of the unit vector $\hat{\mathbf{g}}(p_x, p_y)$:

$$\tilde{N}_3 = \frac{1}{4\pi} \int dp_x dp_y\ \hat{\mathbf{g}} \cdot \left(\frac{\partial \hat{\mathbf{g}}}{\partial p_x} \times \frac{\partial \hat{\mathbf{g}}}{\partial p_y} \right)\ . \qquad (40)$$

Since at infinity the unit vector field $\hat{\mathbf{g}}$ has the same value, $\hat{\mathbf{g}}_{p\to\infty} \to (0, 0, 1)$, the 2-momentum space (p_x, p_y) becomes isomoprhic to the compact S^2 sphere. The function $\hat{\mathbf{g}}(\mathbf{p})$ realizes the mapping of this S^2 sphere to the S^2 sphere of the unit vector $\hat{\mathbf{g}}$ with winding number $\tilde{N}_3$. For $\mu > 0$ one has $\tilde{N}_3 = -1$ and for $\mu < 0$ one has $\tilde{N}_3 = 0$.

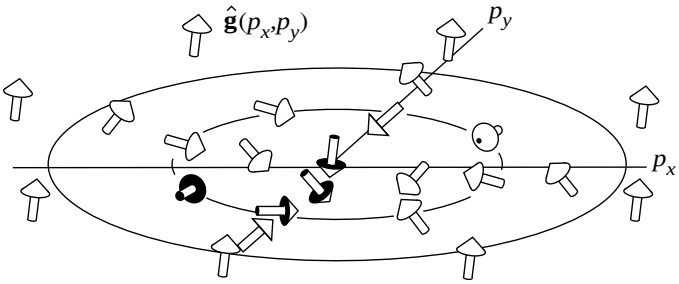

Fig. 11. Skyrmion in **p**-space with momentum space topological charge $\tilde{N}_3 = -1$. It describes topologically non-trivial vacua in 2+1 systems with a fully non-singular Green's function

5.2 Quantization of Physical Parameters

The topological charge $\tilde{N}_3$ and other similar topological charges in $2+1$ systems give rise to quantization parameters. In particular, they are responsible for quantization of Hall and spin-Hall conductivities, which occurs without applied magnetic field (the so-called intrinsic or anomalous quantum Hall and spin quantum Hall effects). There are actually 4 responses of currents to transverse forces which are quantized under appropriate conditions. These are: (i) quantized response of the mass current (or electric current in electrically charged systems) to transverse gradient of chemical potential $\nabla\mu$ (transverse electric field $\mathbf{E}$); (ii) quantized response of the mass current (electric current) to transverse gradient of magnetic field interacting with Pauli spins; (iii) quantized response of the spin current to transverse gradient of magnetic field; and (iv) quantized response of the spin current to transverse gradient of chemical potential (transverse electric field) [55].

Chern–Simons Term and p-Space Topology

All these responses can be described using the generalized Chern–Simons term which mixes different gauge fields (see Eq. (21.20) in [6]):

$$F_{\rm CS}\{\mathbf{A}\} = \frac{1}{16\pi} N_{IJ} e_{\mu\nu\lambda} \int d^2x dt A^I_\mu F^J_{\nu\lambda} \,. \tag{41}$$

Here A^I_μ is the set of the real or auxiliary (fictituous) gauge fields. In electrically neutral systems, instead of the gauge field A_μ one introduces the auxiliary $U(1)$ field, so that the current is given by variation of the action with respect to A_μ: $\delta S/\delta A_\mu = J^\mu$. The auxiliary $SU(2)$ gauge field A^i_μ is convenient for the description of the spin-Hall effect, since the variation of the action with respect to A^a_μ gives the spin current: $\delta S/\delta A^i_\mu = J^\mu_i$. Some components of the field $A_{\mu a}$ are physical, being represented by the real physical quantities which couple to the fermionic charges. Example is provided by the external magnetic field in neutral system, which play the role of A^i_0 (see Sect. 21.2 in Ref. [6]). After the current is calculated the values of the auxiliary fields are fixed. The latest discussion of the mixed Chern–Simons term can be found in Ref. [56]. For the related phenomenon of axial anomaly, the mixed action in terms of different (real and fictituous) gauge fields has been introduced in Ref. [57].

The important fact is that the matrix N_{IJ} of the prefactors in the Chern–Simons action is expressed in terms of the momentum-space topological invariants:

$$N_{IJ} = \frac{1}{24\pi^2} e_{\mu\nu\lambda}\, \mathbf{tr}\, Q_I Q_J \int d^2p d\omega\; G\partial_{p_\mu} G^{-1} G\partial_{p_\nu} G^{-1} G\partial_{p_\lambda} G^{-1} \,, \tag{42}$$

where Q_I is the fermionic charge interacting with the gauge field A^I_μ (in case of several fermionic species, Q_I is a matrix in the space of species).

Intrinsic Spin Quantum Hall Effect

To obtain, for example, the response of the spin current J_z^i to the electric field E_i, one must consider two fermionic charges: the electric charge $Q_1 = e$ interacting with $U(1)$ gauge field, and the spin along z as another charge, $Q_2 = s_z = \hbar\sigma_z/2$, which interacts with the fictituous $SU(2)$ field A_μ^z. This gives the quantized spin current response to the electric field $J_z^i = e^{ij}\sigma_{\rm spin-Hall}E_j$, where $\sigma_{\rm spin-Hall} = (e\hbar/8\pi)N$ and N is integer:

$$N = \frac{1}{24\pi^2} e_{\mu\nu\lambda}\ \mathbf{tr}\ \sigma_z \int d^2p d\omega\ G\partial_{p_\mu}G^{-1}G\partial_{p_\nu}G^{-1}G\partial_{p_\lambda}G^{-1}\ . \qquad (43)$$

Quantization of the spin-Hall conductivity in the commensurate lattice of vortices can be found in Ref. [58].

The above consideration is applicable, when the momentum (or quasi-momentum in solids) is the well defined quantity, otherwise (for example, in the presence of impurities) one cannot construct the invariant in terms of the Green's function $G(\mathbf{p}, \omega)$. However, it is not excluded that in some cases the perturbative introduction of impurities does not change the prefactor N_{IJ} in the Chern–Simons term (41) and thus does not influence the quantization: this occurs if there is no spectral flow under the adiabatic introduction of impurities. In this case the quantization is determined by the reference system – the fully gapped system from which the considered system can be obtained by the continuous deformation without the spectral flow (analogueous phenomenon for the angular momentum paradox in ^{3}He-A was discussed in [59]). The most recent review paper on the spin current can be found in [60].

Momentum Space Topology and Hall Effect in 3D Systems

The momentum space topology is also important for the Hall effect in some 3+1 systems. The contribution of Fermi points to the intrinsic Hall effect is discussed in the Appendix of Ref. [18]. For metals with Fermi surfaces having the global topological charge N_3 (see Sect. 3.3) the anomalous Hall effect is caused by the Berry curvature on the Fermi surface [61]. The magnitude of the Hall conductivity is related to the volume of the Fermi surface in a similar way as the number of particles and the volume of the Fermi surface are connected by the Luttinger theorem [61]. Another "partner" of the Luttinger theorem emerges for the Hall effect in superconductors, where topology enters via the spectral flow of fermion zero modes in the cores of topological defects – Abrikosov vortices [62].

5.3 Quantum Phase Transitions

Plateau Transitions

The integer topological invariant $\tilde{N}_3$ of the ground state cannot follow the continuous parameters of the system. That is why when one changes such a

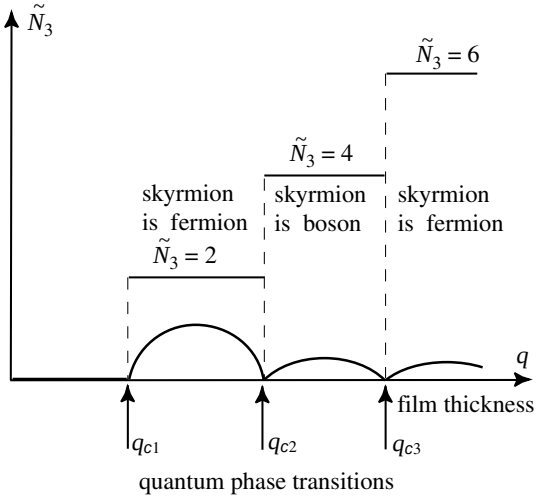

Fig. 12. Quantum phase transitions occurring when one increases the thickness q of the ^{3}He-A film. The transitions at $q = q_{c2}$ and $q = q_{c3}$ are plateau-plateau transitions between vacua with different values of integer topological invariant $\tilde{N}_3$ in Eq. (37). At these transitions the quantum statistics of real-space skyrmions living in thin films changes. The change in the quasiparticle spectrum across the transitions may be seen from the minimum value of the quasiparticle energy, $\min_{\mathbf{p}} E(\mathbf{p})$, at given q (*thick lines*). The transitions at $q = q_{c2}$ and $q = q_{c3}$ between the fully gapped states occur through the gapless states. At $q = q_{c1}$ the transition is between gapless and fully gapped states

parameter, for example, the chemical potential in the model (39), one obtains the quantum phase transition at $\mu = 0$ at which $\tilde{N}_3$ jumps from 0 to -1. The film thickness is another relevant parameter. In the film with finite thickness the matrix of Green's function acquires indices of the levels of transverse quantization. If one increases the thickness of the film, one finds a set of quantum phase transitions between vacua with different integer values of the invariant [Fig. 12], and thus between the plateaus in Hall or spin-Hall conductivity.

The abrupt change of the topological charge cannot occur adiabatically, that is why at the points of quantum transitions fermionic quasiparticles become gapless.

Topological Zero Modes and Edge States

If two vacua with different $\tilde{N}_3$ coexist in space [Fig. 13(**a**)], the phase boundary between them must also contain gapless fermions. This is an example of the so-called fermion zero modes living on different topological objects such as 3D monopole, 2D soliton wall, and 1D vortex/string (see Ref. [63] and references therein). The number of the gapless fermion zero modes obeys the index theorem: in our case the number of the $1 + 1$ fermions living at the

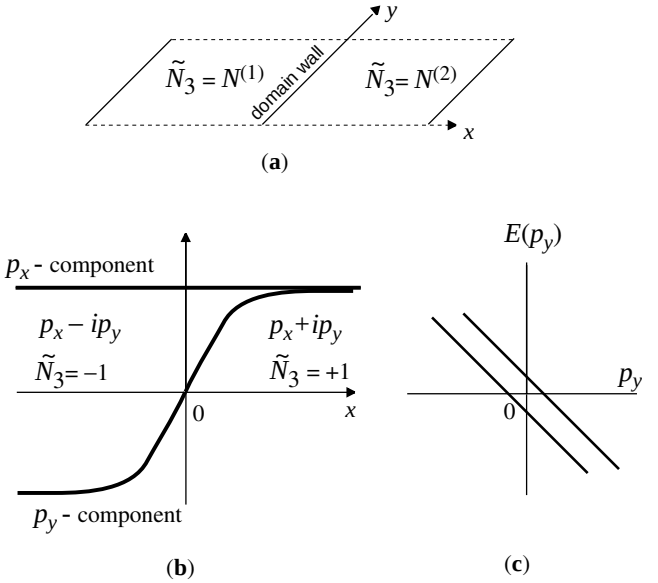

Fig. 13. (**a**) Domain wall between two 2+1 vacua with different topological charges $\tilde{N}_3$. (**b**) Structure of the phase boundary between vacua with charges $\tilde{N}_3 = \pm 1$ in Eq. (44). The prefactor in front of p_y changes sign at $x = 0$, which leads to the change of sign of the topological charge in Eq. (37). (**c**) Fermion zero modes – anomalous branches of fermions living at the interface whose spectrum crosses zero energy level. The number of anomalous branches is determined by the difference of the topological charges $\tilde{N}_3$ across the wall

phase boundary is determined by the difference of the topological charges of the two vacua, $\tilde{N}_3^{(1)} - \tilde{N}_3^{(2)}$ (see Chap. 22 in Ref. [6]).

The boundary of the condensed matter system can be considered as the phase boundary between the state with nonzero $\tilde{N}_3$ and the state with $\tilde{N}_3 = 0$. The corresponding fermion zero modes are the edge states well known in physics of the QHE.

Example of the phase boundary between two vacua with $\tilde{N}_3 = \pm 1$ is shown in Fig. 13(**b**) for the $p_x + ip_y$ superfluids and superconductors. Here the p_y component of the order parameter changes sign across the wall. The simplest structure of such boundary is given by Hamiltonian

$$H = \begin{pmatrix} \frac{p^2}{2m} - \mu & c\left(p_x + ip_y \tanh \frac{x}{\xi}\right) \\ c\left(p_x - ip_y \tanh \frac{x}{\xi}\right) & -\frac{p^2}{2m} + \mu \end{pmatrix} . \tag{44}$$

Let us first consider fermions in semiclassical approach, when the coordinates x and p_x are independent. At $x = 0$ the time reversal symmetry is restored, and the spectrum becomes gapless. At $x = 0$ there are two zeroes of co-dimension 2 at points $p_x = 0$ and $p_y = \pm p_F$. They are similar to zeroes

discussed in Sect. 4.2. These zeroes are marginal, and disappear at $x \neq 0$ where the time reversal symmetry is violated. The topological charge is well defined only at $x \neq 0$. When x crosses zero, the topological charge in Eq. (37) changes sign.

In the quantum mechanical description, x and p_x do not commute. The quantum-mechanical spectrum $E(p_y)$ contains fermion zero modes – branches of spectrum which cross zero. According to the index theorem there are two anomalous branches in Fig. 13(**c**).

The index theorem together with the connection between the topological charge and quantization of physical parameters discussed in Sect. 5.2 implies that the quantization of Hall and/or spin-Hall conductance is determined by the number of edge states in accordance with Refs. [64]. The detailed discussion of the edge modes in $p_x + ip_y$ superfluids and superconductors and their contribution to the effective action can be found in Ref. [65]. These edge modes are Majorana fermions.

"Higgs" Transition in p-Space

Note that the energy spectrum in Eq. (38) experiences an analogue of the Higgs phase transition at $\mu = mc^2$ [Fig. 14]: if $\mu < mc^2$ the quasiparticle energy has a single minimum at $p = 0$, while at $\mu > mc^2$ the minimum is at the circumference with radius $p_0 = \sqrt{2m(\mu - mc^2)}$. There is no symmetry breaking at this transition, since the vacuum state has the same rotational

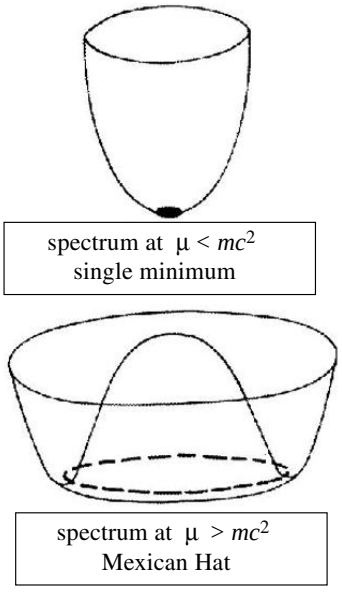

Fig. 14. "Higgs" transition in momentum space.

symmetry above and below the transition, while the asymptotic behavior of the thermodynamic quantities ($\propto T^n \exp\left(-E_{\min}/T\right)$) experiences discontinuity across the transition: the power n changes. That is why the point $\mu = mc^2$ marks the quantum phase transition, at which the topology of the minima of the energy spectrum changes.

However, this transition does not belong to the class of transitions which we discuss in the present review, since the topological invariant of the ground state $\tilde{N}_3$ does not change across this transition and thus at the transition point $\mu = mc^2$ the spectrum remains fully gapped. Moreover, such a transition does not depend on dimension of space-time and occurs in 3+1 systems as well. Example is provided by the s-wave superconductor or s-wave Fermi superfluid, whose spectrum in Eq. (5) experiences the same Higgs-like transition at $\mu = 0$, i.e. in the BSC–BEC crossover region.

5.4 Quantum Phase Transition in 1D Quantum Ising Model

The momentum-space topology is applicable not only to fermionic systems, but to any system which can be expressed in terms of auxiliary fermions.

Fermionization and Topological Invariant

Example is provided by the 1-dimensional quantum Ising model where the topological quantum phase transition between the fully gapped vacua can be described in terms of the invariants for the fermionic Green's function. The original Hamiltonian of this 1D chain of spins is:

$$H = -J \sum_{n=1}^{N} \left(h\sigma_n^x + \sigma_n^z \sigma_{n+1}^z \right) \ , \tag{45}$$

where σ^x and σ^z are Pauli matrices, and h is the parameter describing the external magnetic field. After the standard Jordan-Wigner transformation this system can be represented in terms of the non-interacting fermions with the following Hamiltonian in the continuous $N \to \infty$ limit (see Ref. [66] and references therein):

$$H = 2J \left(h - \cos(pa)\right) \tau_3 + 2J \sin(pa) \tau_1 \ , \quad -\frac{\pi}{a} < p < \frac{\pi}{a} \ . \tag{46}$$

It is periodic in the one-dimensional momentum space p with period $2\pi/a$ where a is the lattice spacing. The integer valued topological invariant here is the same as in Eq. (33) but now the integration is along the closed path in p-space, i.e. from 0 to $2\pi/a$:

$$\tilde{N}_2 = -\frac{1}{4\pi i} \ \mathrm{tr} \oint dp \ \tau_2 H^{-1} \nabla_p H \ . \tag{47}$$

This invariant can be represented in terms of the Green's function

$$G^{-1} = ig_z - g_x\tau_3 + g_y\tau_1 \ , \tag{48}$$

where for the particular case of the model (46), the components of the 3D vector $\mathbf{g}(p,\omega)$ are:

$$g_x(p,\omega) = 2J\,(h - \cos(pa)) \ , \quad g_y(p,\omega) = 2J\sin(pa) \ , \quad g_z(p,\omega) = \omega \ . \tag{49}$$

Then the invariant (47) becomes:

$$\tilde{N}_2 = \frac{1}{4\pi}\int_{-\pi/a}^{\pi/a} dp \int_{-\infty}^{\infty} d\omega \ \hat{\mathbf{g}} \cdot \left(\frac{\partial \hat{\mathbf{g}}}{\partial p} \times \frac{\partial \hat{\mathbf{g}}}{\partial \omega}\right) \ . \tag{50}$$

The invariant is well defined for the fully gapped states, when $\mathbf{g} \neq 0$ and thus the unit vector $\hat{\mathbf{g}} = \mathbf{g}/|\mathbf{g}|$ has no singularity. In the model under discussion, one has for $h \neq 1$:

$$\tilde{N}_2(h<1) = 1 \ , \quad \tilde{N}_2(h>1) = 0 \ . \tag{51}$$

Instanton in (p,ω)-Space

The state with $\tilde{N}_2 = 1$ is the "instanton" in the (ω, p)-space, which is similar to the skyrmion in (p_x, p_y)-space in Fig. 11. The real space-time counterpart of such instanton can be found in Refs. [67]. It describes the periodic phase slip process occurring in superfluid ^{3}He-A [68]. In the model, the topological structure of the instanton at $h < 1$ can be easily revealed for $h = 0$. Introducing "space-time" coordinates $t = p$ and $z = \omega/2J$ one obtains that the unit vector $\hat{\mathbf{g}}$ precesses sweeping the whole unit sphere during one period $\Delta t = 2\pi/a$ [Fig. 15]:

$$\hat{\mathbf{g}}(z,t) = \hat{\mathbf{z}}\cos\theta(z) + \sin\theta(z)\,(\hat{\mathbf{x}}\cos(at) + \hat{\mathbf{y}}\sin(at)) \ , \quad \cot\theta(z) = z \ . \tag{52}$$

This state can be referred to as "ferromagnetic", since in terms of spins it is the quantum superposition of two ferromagnetic states with opposite magnetization.

At $h > 1$, i.e. in the "paramagnetic"phase, the momentum-space topology is trivial, $\tilde{N}_2(h>1) = 0$. The transition at $h = 1$ at which the topological charge $\tilde{N}_2$ of the ground state changes is the quantum phase transition, it only occurs at $T = 0$.

Phase Diagram for Anisotropic XY-Chain

The phase diagram for the extension of the Ising model to the case of the anisotropic XY spin chain in a magnetic field with Hamiltonian (see e.g. [69])

$$H = -J\sum_{n=1}^{N}\left(h\sigma_n^x + \frac{1+\gamma}{2}\sigma_n^z\sigma_{n+1}^z + \frac{1-\gamma}{2}\sigma_n^y\sigma_{n+1}^y\right) \ , \tag{53}$$

is shown in Fig. 16 in terms of the topological charge $\tilde{N}_2$. The lines $h = 1$, $h = -1$ and ($\gamma = 0$, $-1 < h < 1$), which separate regions with different $\tilde{N}_2$, are lines of quantum phase transitions.

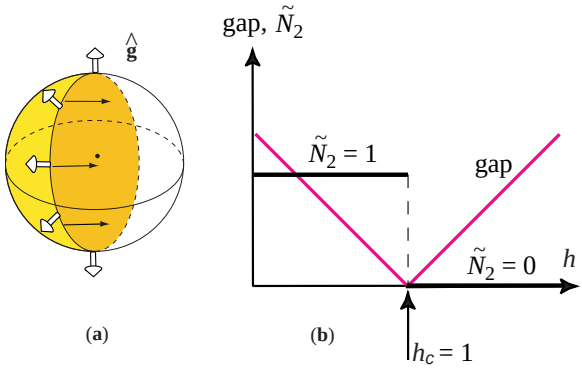

Fig. 15. (**a**) Illustration of the topological invariant $\tilde{N}_2 = 1$ for "instanton" in momentum space for $h = 0$. According to Eq. (52) one has the domain wall in $z = \omega/2J$ space across which the direction of the vector $\mathbf{g}$ changes from $\hat{\mathbf{z}}$ at $z = \infty$ to $-\hat{\mathbf{z}}$ at at $z = -\infty$. The structure is periodic in p and thus is precessing in "time" $t = p$ (*black arrows*). During one period of precession $\Delta t = 2\pi/a$ the unit vector $\hat{\mathbf{g}}(t, z)$ sweeps the whole unit sphere giving $\tilde{N}_2 = 1$ in Eq. (50). (**b**) At the transition point $h_c = 1$ the gap in the energy spectrum of fermions vanishes, because the transition between two vacuua with different topological charge cannot occur adiabatically

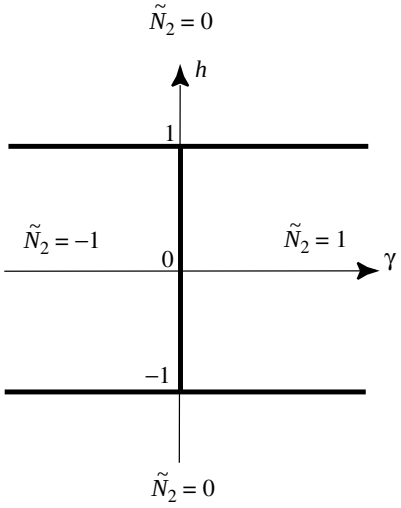

Fig. 16. Phase diagram for anisotropic XY-chain in Eq. (53) in the plane (γ, h). The regions with different topological charge $\tilde{N}_2$ are separated by the lines of topological quantum phase transitions (*thick lines*)

Nullification of Gap at Quantum Transition

Because of the jump in $\tilde{N}_2$ [Fig. 15(**b**)], the transition cannot occur adiabatically. That is why the energy gap must tend to zero at the transition, in the same way as it occurs at the plateau-plateau transition in Fig. 12. In the Ising model, the energy spectrum $E^2(p) = g_x^2(p) + g_y^2(p) = 4J^2\left((h - \cos(pa))^2 + \sin^2(pa)\right)$ has a gap $E(0) = 2J|h-1|$ which tends to zero at $h \to 1$ [Fig. 15 (**b**)]. However, the nullification of the gap at the topological transition between the fully gapped states with different topological charges is the general property, which does not depend on the details of the underlying spin system and is robust to interaction between the auxiliary fermions.

The special case, when the gap does not vanish at the transition because the momentum space is not compact, is discussed in Sect. 11.4 of [6].

Dynamics of Quantum Phase Transition and Superposition of Macroscopic States

In the quantum Ising model of Eq. (45) the ground state at $h < 1$ represents the quantum superposition of two ferromagnetic states with opposite magnetization. However, in the limit of infinite number of spins $N \to \infty$ this becomes the Schrödinger's Cat – the superposition of two macroscopically different states. According to Ref. [70] such superposition cannot be resolved by any measurements, because in the limit $N \to \infty$ no observable has matrix elements between the two ferromagnetic states, which are therefore disjoint. In general, the disjoint states form the equivalence classes emerging in the limit of infinite volume or infinite number of elements.

Another property of the disjoint macroscopic states is that their superposition, even if it is the ground state of the Hamiltonian, can never be achieved. For example, let us try to obtain the superposition of the two ferromagnetic states at $h < 1$ starting from the paramagnetic ground state at $h > 1$ and slowly crossing the critical point $h = 1$ of the quantum phase transition. The dynamics of the time-dependent quantum phase transition in this model has been discussed in Refs. [66, 71]. It is characterized by the transition time τ_Q which shows how fast the transition point is crossed: $1/\tau_Q = \dot{h}|_{h=1}$. One may expect that if the transition occurs adiabatically, i.e. in the limit $\tau_Q \to \infty$, the ground state at $h > 1$ transforms to the ground state at $h < 1$. However, in the limit $N \to \infty$ the adiabatic condition cannot be satisfied. If $\tau_Q \to \infty$ but $\tau_Q \ll N^2/J$, the transition becomes non-adiabatic and the level crossing occurs with probability 1. Instead of the ground state at $h < 1$ one obtains the excited state, which represents two (or several) ferromagnetic domains separated by the domain wall(s). Thus in the $N = \infty$ system instead of the quantum superposition of the two ferromagnetic states the classical coexistence of the two ferromagnetic states is realized.

In the obtained excited state the translational and time reversal symmetries are broken. This example of spontaneous symmetry breaking occurring at $T = 0$ demonstrates the general phenomenon that in the limit of the infinite system one can never reach the superposition of macroscopically different states. On the connection between the process of spontaneous symmetry breaking and the measurement process in quantum mechanics see Ref. [72] and references therein. Both processes are emergent phenomena occurring in the limit of infinite volume V of the whole system. In finite systems the quantum mechanics is reversible. For general discussion of the symmetry breaking phase transition in terms of the disjoint limit Gibbs distributions emerging at $V \to \infty$ see the book by Sinai [73].

6 Conclusion

Here we discussed the quantum phase transitions which occur between the vacuum states with the same symmetry above and below the transition. Such a transition is essentially different from conventional phase transition which is accompanied by the symmetry breaking. The discussed zero temperature phase transition is not the termination point of the line of the conventional 2-nd order phase transition: it is either an isolated point $(q_c, 0)$ in the (q, T) plane, or the termination line of the 1-st order transition. This transition is purely topological – it is accompanied by the change of the topology of fermionic Green's function in **p**-space without change in the vacuum symmetry. The **p**-space topology, in turn, depends on the symmetry of the system. The interplay between symmetry and topology leads to variety of vacuum states and thus to variety of emergent physical laws at low energy, and to variety of possible quantum phase transitions. The more interesting situations are expected for spatially inhomogeneous systems, say for systems with topological defects in **r**-space, where the **p**-space topology, the **r**-space topology, and symmetry are combined (see Refs. [7, 74] and Chap. 23 in [6]).

I thank Frans Klinkhamer for collaboration and Petr Horava for discussions. This work is supported in part by the Russian Foundation for Basic Research #06-02-16002, by the Russian Ministry of Education and Science, through the Leading Scientific School grant #2338.2003.2, and by the European Science Foundation COSLAB Program.

References

1. H. Georgi, S.L. Glashow: Phys. Rev. Lett. **32**, 438 (1974)
2. H. Georgi, H.R. Quinn, S. Weinberg: Phys. Rev. Lett. **33**, 451 (1974)
3. G.E. Volovik, L.P. Gorkov: Sov. Phys. JETP **61**, 843 (1985)
4. D. Vollhardt, P. Wölfle: *The Superfluid Phases of Helium 3* (Taylor and Francis, London, 1990)

5. N.D. Mermin: Rev. Mod. Phys. **51**, 591 (1979)
6. G.E. Volovik: *The Universe in a Helium Droplet* (Clarendon Press, Oxford, 2003)
7. P. Horava: Phys. Rev. Lett. **95**, 016405 (2005)
8. V.A. Khodel, V.R. Shaginyan: JETP Lett. **51**, 553 (1990)
9. G.E. Volovik: JETP Lett. **53**, 222 (1991)
10. V.R. Shaginyan, A.Z. Msezane, M.Ya. Amusia: Phys. Lett. A **338**, 393 (2005)
11. V.A. Khodel, J.W. Clark, M.V. Zverev: 'Thermodynamic properties of Fermi systems with flat single-particle spectra' (cond-mat/0502292)
12. V.A. Khodel, M.V. Zverev, V.M. Yakovenko: Phys. Rev. Lett. **95**, 236402 (2005)
13. G.E. Volovik: JETP Lett. **59**, 830 (1994)
14. S. Sachdev: *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 2003)
15. I.M. Lifshitz: Sov. Phys. JETP **11**, 1130 (1960); I.M. Lifshitz, M.Y. Azbel, M.I. Kaganov: *Electron Theory of Metals* (Consultant Press, New York, 1972)
16. G.E. Volovik: *Exotic Properties of Superfluid* ^{3}*He* (World Scientific, Singapore, 1992)
17. F.R. Klinkhamer, G.E. Volovik: JETP Lett. **80**, 343 (2004)
18. F.R. Klinkhamer, G.E. Volovik: Int. J. Mod. Phys. A **20**, 2795 (2005)
19. V. Gurarie, L. Radzihovsky, A. V. Andreev: Phys. Rev. Lett. **94**, 230403 (2005)
20. S.S. Botelho, C.A.R. Sa de Melo: J. Low Temp. Phys. **140**, 409 (2005)
21. S.S. Botelho, C.A.R. Sa de Melo: Phys. Rev. B **71**, 134507 (2005)
22. L.S. Borkowski, C.A.R. Sa de Melo: "From BCS to BEC superconductivity: Spectroscopic consequences" (cond-mat/9810370)
23. R.D. Duncan, C.A.R. Sa de Melo: Phys. Rev. B **62**, 9675 (2000)
24. X.G. Wen, A. Zee: Phys. Rev. B **66**, 235110 (2002)
25. E. Gubankova: "Conditions for existence of neutral strange quark matter" (hep-ph/0507291) E. Gubankova, E. Mishchenko, F. Wilczek: Phys. Rev. Lett. **94**, 110402 (2005) K. Rajagopal, A. Schmitt: Phys. Rev. D**73**, 045003 (2006) R. Casalbuoni: "Color Superconductivity in High Density QCD' (hep-ph/0512198)
26. N. Bergeal, J. Lesueur, M. Aprili, G. Faini, J. P. Contour, B. Leridon: "Direct test of pairing fluctuations in the pseudogap phase of underdoped cuprates" (cond-mat/0601265)
27. M.V. Sadovskii: "Models of the pseudogap state in high temperature superconductors" (cond-mat/0408489)
28. E.Z. Kuchinskii, M.V. Sadovskii: JETP **103**, 415 (2006)
29. J. von Neumann, E. Wigner: Phys. Zeit. **30**, 467 (1929)
30. A.J. Stone: Proc. R. Soc. London A **351**, 141 (1976)'
31. V.I. Arnold: *Mathematical Method in Classical Mechanics* (Nauka, Moscow, 1979, in Russian) (Springer-Verlag, 1989)
32. M.V. Berry: Proc. R. Soc. London A **392**, 45 (1984)
33. S.P. Novikov: Sov. Phys. Math. Dokl. **23**, 298 (1981)
34. G.E. Volovik: JETP Lett. **46**, 98 (1987)
35. H.B. Nielsen, M. Ninomiya: Nucl. Phys. B **185**, 20 (1981); Nucl. Phys. B **193**, 173 (1981)
36. X.G. Wen: Phys. Rev. Lett. **88**, 011602 (2002)
37. F.R. Klinkhamer: Int. J. Mod. Phys. A**21**, 161 (2006); F.R. Klinkhamer: Nucl. Phys. B (Proc. Suppl.) **149**, 209 (2005)

38. K. Kajantie, M. Laine, K. Rummukainen, M. Shaposhnikov: Phys. Rev. Lett. **77**, 2887 (1996)
39. S.L. Adler: "Anomalies to all orders". In: *Fifty Years of Yang-Mills Theory*, ed. by G. 't Hooft (World Scientific, 2006)
40. E.I. Blount: Phys. Rev. B **32**, 2935 (1985)
41. H.J.H. Smilde, A.A. Golubov, Ariando, G. Rijnders, J.M. Dekkers, S. Harkema, D.H.A. Blank, H. Rogalla, H. Hilgenkamp, Phys. Rev. Lett. **95**, 257001 (2005)
42. K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, M.I. Katsnelson, I.V. Grigorieva, S.V. Dubonos, A.A. Firsov: Nature **438**, 197 (2005)
43. S.G. Sharapov, V.P. Gusynin, H. Beck: Phys. Rev. B **69**, 075104 (2004)
44. K.S. Novoselov, E. McCann, S.V. Morozov, V.I. Falko, M.I. Katsnelson, U. Zeitler, D. Jiang, F. Schedin, A.K. Geim: Nature Physics **2**, 177 (2006)
45. E. McCann, V.I. Fal'ko: Phys. Rev. Lett. **96**, 086805 (2006)
46. M. Sato: Phys. Rev. B**73**, 214502 (2006)
47. A.A. Kordyuk, S.V. Borisenko, A.N. Yaresko, S.-L. Drechsler, H. Rosner, T.K. Kim, A. Koitzsch, K.A. Nenkov, M. Knupfer, J. Fink, R. Follath, H. Berger, B. Keimer, S. Ono, Yoichi Ando: Phys. Rev. B **70**, 214525 (2004); S.V. Borisenko, A.A. Kordyuk, V. Zabolotnyy, J. Geck, D. Inosov, A. Koitzsch, J. Fink, M. Knupfer, B. Buechner, V. Hinkov, C.T. Lin, B. Keimer, T. Wolf, S.G. Chiuzbaian, L. Patthey, R. Follath: Phys. Rev. Lett. **96**, 117004 (2006); T. Yamasaki, K. Yamazaki, A. Ino, M. Arita, H. Namatame, M. Taniguchi, A. Fujimori, Z.-X. Shen, M. Ishikado, S. Uchida: "Unmasking the nodal quasiparticle dynamics in cuprate superconductors using low-energy photoemission," (cond-mat/0603006)
48. G.E. Volovik: JETP Lett. **73**, 162 (2001)
49. F.R. Klinkhamer: Phys. Rev. D **73**, 057301 (2006)
50. G.E. Volovik: JETP Lett. **58**, 469 (1993)
51. H. Balci, R.L. Greene: Phys. Rev. Lett. **93**, 067001 (2004)
52. W. Yu, B. Liang, R.L. Greene: Phys. Rev. B **72**, 212512 (2005)
53. K. Ishikawa, T. Matsuyama: Z. Phys. C **33**, 41 (1986) K. Ishikawa, T. Matsuyama: Nucl. Phys. B **280**, 523 (1987)
54. G.E. Volovik, V.M. Yakovenko: J. Phys.: Condens. Matter **1**, 5263 (1989)
55. G.E. Volovik: "Fractional statistics and analogues of quantum Hall effect in superfluid ^{3}He films". In: *Quantum Fluids and Solids – 1989* ed. by G.G. Ihas, Y. Takano (AIP Conference Proceedings, 1989) **194**, pp. 136–146
56. Su-Peng Kou, Xiao-Liang Qi, Zheng-Yu Weng: Phys. Rev. B **72**, 165114 (2005)
57. D.T. Son, A.R. Zhitnitsky: Phys. Rev. D **70**, 074018 (2004)
58. O. Vafek and A. Melikyan: Phys. Rev. Lett. **96**, 167005 (2006)
59. G.E. Volovik: JETP Lett. **61**, 958 (1995)
60. E.I. Rashba: "Spin-orbit coupling and spin transport" (cond-mat/0507007)
61. S.D.M. Haldane: Phys. Rev. Lett. **93**, 206602 (2004)
62. G.E. Volovik: JETP Lett. **64**, 845 (1996)
63. R. Jackiw, P. Rossi: Nucl. Phys. B **190**, 681, (1981)
64. M. Stone: Ann. Phys. **207**, 38 (1991) X.G. Wen: Phys. Rev. B **43**, 11025 (1991)
65. M. Stone, R. Roy: Phys. Rev. B **69**, 184511 (2004)
66. J. Dziarmaga: Phys. Rev. Lett. **95** 245701 (2005)
67. J.R. Hook, H.E. Hall: J. Phys. C **12**, 783 (1979) G.E. Volovik: JETP Lett. **27**, 573 (1978)
68. D.N. Paulson, M. Krusius, J.C. Wheatley: Phys. Rev. Lett. **36**, 1322 (1976)
69. A.G. Abanov, F. Franchini: Phys. Lett. A **316**, 342 (2003)

70. M. Fioroni, G. Immirzi: "How and why the wave function collapses after a measurement" (gr-qc/9411044)
71. W.H. Zurek, U. Dorner, P. Zoller: Phys. Rev. Lett. **95**, 105701 (2005)
72. M. Grady: "Spontaneous symmetry breaking as the mechanism of quantum measurement" (hep-th/9409049)
73. Ya.G. Sinai: *Theory of Phase Transitions, International series in natural philosophy* (Pergamon Press, 1983)
74. P.G. Grinevich, G.E. Volovik: J. Low Temp. Phys. **72**, 371 (1988); M.M. Salomaa, G.E. Volovik: Phys. Rev. **37**, 9298 (1988); M.M. Salomaa, G.E. Volovik: J. Low Temp. Phys. **74**, 319 (1989)

Superfluid ^{3}He as a Model System for Cosmology – Experimental Point of View

P. Skyba

Centre of Low Temperature Physics, Institute of Experimental Physics, Slovak Academy of Sciences, Watsonova 47, 04001 Košice, Slovakia
skyba@saske.sk

1 Introduction

In spite of the progress in the science during the last hundred years, mankind, in principle, does not know the exact answers to the fundamental questions related to the beginning and the evolution of the Universe (or properties of the quantum vacuum - ether) and mass. Theoretical models of the Universe's evolution from the Big Bang up to present day are supported e.g. by the discovery of cosmic microwave background and by the latest astronomical and astrophysical observations. However, due to the impossibility to realize a controlled cosmological experiment, mankind can only act as a passive observer of the stage of a theater called the Universe. This impossibility leads theoretical and experimental physicists to look for the analogues between the physical processes studied in cosmology with similar processes of condensed matter physics which can be studied experimentally. It is obvious that there are not one to one analogues, however, the possibility of the study of these processes experimentally under various controlled conditions has obvious advantages. Moreover, almost all properties of condensed matter are known from first principles and nearly all of them were verified experimentally.

During its evolution, the Universe has passed through several phase transitions at which the symmetry of its states had been spontaneously broken as individual interactions were separated or formed. Forgetting about, the so called Planck scale at energies of 10^{19} GeV, at very high energy, the Grand Unification Theory (GUT) suggests an initial group of symmetry SO(10) which unifies the strong, weak and hypercharge interactions. The phase transitions which followed at various energies reduced a degree of this symmetry. At energies about 10^{15} GeV, the symmetry group SO(10) was spontaneously broken into subgroups of the Standard model: SU(3) × SU(2) × U(1) corresponding to the separate symmetry for each of three interactions: electro-weak, strong and electromagnetic. At an energy of 200 GeV the electro-weak interaction was formed and the symmetry U(1)× SU(2) was broken. What has been left - we are today - electromagnetic and strong interactions U(1)× SU(3). In the

P. Skyba: *Superfluid ^{3}He as a Model System for Cosmology – Experimental Point of View*, Lect. Notes Phys. **718**, 75–92 (2007)
DOI 10.1007/3-540-70859-6_4

framework of the GUT, the Universe starting from the state with symmetry SU(10) passed through several phase transitions: SU(10)→ SU(3) × SU(2) × U(1) → SU(3)×U(1), simultaneously the energy and symmetries were reduced as well [1].

The idea to use condensed matter to study analogues of the physical processes studied in cosmology is very simple in principle: a symmetry of low energy state is a relic of the symmetry of a higher energy state, both separated from each other by a phase transition. In general, at higher temperatures condensed matter occupies an energy state having a higher symmetry. Then, when condensed matter is cooled down, it undergoes a phase transition which reduces its symmetry. A question which we may ask here is whether we can find condensed matter system which can pass through phase transitions and violate as many symmetries as possible to be an analogue of the predictions mentioned above? It seems that there are few good candidates of condensed matter systems. One of the best candidates for such analogues is the superfluid phases of ^{3}He (for comprehensive overview see [1, 2]).

In fact, about ten years ago experiments at Lancaster University showed that using neutron irradiation one can overheat small regions of superfluid ^{3}He-B above the temperature of superfluid transition [3]. Consecutive experiments performed by the Grenoble and the Helsinki groups demonstrated that the subsequent rapid cooling of such localized heated regions back into the superfluid state are accompanied by the creation of vortices - topological defects [4, 5]. These defects, however, can be regarded as analogues of the cosmological defects formed during the early evolution of the Universe as predicted by a Kibble-Zurek model [6, 7]. Thus, these experiments gave the first quantitative tests of the cosmological Kibble-Zurek model.

Therefore, in the context of the above discussion, the superfluid helium-3 research provides a wider means for offering the possibility of experimental investigations of quantum field/cosmological theories via analogues with the superfluid phases of ^{3}He. To show why the superfluid ^{3}He phases are a good model system to study the cosmological analogues, first it is necessary to describe its basic properties. Basic theory of the superfluid ^{3}He had been developed by A. Leggett [8] and for a comprehensive theoretical and experimental ^{3}He overview see [9] and [10], respectively.

2 Basic Properties of the Superfluid ^{3}He

At higher temperatures ^{3}He is a gas, while below temperature of 3K - due to van der Walls forces – ^{3}He is a normal liquid with all symmetries which a condensed matter system can have: translation, gauge symmetry U(1) and two SO(3) symmetries for the spin ($SO^S(3)$) and orbital ($SO^L(3)$) rotations. At temperatures below 100 mK, ^{3}He behaves as a strongly interacting Fermi liquid. Its physical properties are well described by Landau's theory. Quasiparticles of the ^{3}He (i.e. ^{3}He atoms "dressed" into mutual interactions) have

spin equal to 1/2 and similar to the electrons, they can create Cooper pairs as well. However, different from electrons in a metal, ^{3}He is a liquid without a lattice and the electron-phonon interaction, responsible for superconductivity, can not be applied here. As the ^{3}He quasiparticles have spin, the magnetic interaction between spins rises up when the temperature falls down until, at a certain temperature, Cooper pairs are created - the coupled pairs of ^{3}He quasiparticles - and the normal ^{3}He liquid becomes a superfluid. The Cooper pairs produce a superfluid component and the rest, unpaired ^{3}He quasiparticles, generate a normal component (N -phase).

A physical picture of the superfluid ^{3}He is more complicated than for superconducting electrons. First, the ^{3}He quasiparticles are bare atoms and creating the Cooper pair they have to rotate around its common center of mass, generating an orbital angular momentum of the pair (L=1). Secondly, the spin of the Cooper pair is equal to one (S=1), thus superfluid ^{3}He has magnetic properties (see Fig. 1). Thirdly, the orbital and spin angular momenta of the pair are coupled via a dipole-dipole interaction.

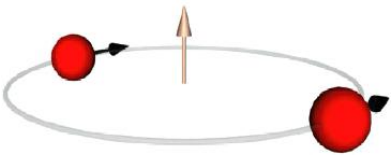

Fig. 1. Schematic picture of the Cooper pair formed from two ^{3}He quasiparticles. Quasiparticles orbit around a common center of mass creating the orbital angular momentum. The spins of individual quasiparticles are aligned perpendicularly to the axis of rotation

It is evident that the phase transition of ^{3}He into the superfluid state is accompanied by spontaneously broken symmetry: orbital, spin and gauge: $\mathrm{SO}^L(3)\times \mathrm{SO}^S(3) \times \mathrm{U}(1)$, except the translational symmetry, as the superfluid ^{3}He is still a liquid. Although this phase transition is not the same as discussed above ($\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$), a similarity is evident. Finally, an energy gap Δ appears in the energy spectrum separating the Cooper pairs (ground state) from unpaired quasiparticles - Fermi excitations.

In superfluid ^{3}He the density of Fermi excitations decreases upon further cooling. For temperatures below around $0.25\mathrm{T}_c$ (where T_c is the superfluid transition temperature), the density of the Fermi excitations is so low that the excitations can be regarded as a non-interacting gas because almost all of them are paired and occupy the ground state. Therefore, at these very low temperatures, the superfluid phases of helium-3 represent well defined models of the quantum vacua (see below) which allows us to study any influences of various external forces on the ground state and excitations from this state as well.

As it was mentioned above, the ground state of superfluid ^{3}He is formed by the Cooper pairs having both spin (S=1) and orbital momentum (L=1). As a consequence of this spin-triplet, orbital p-wave pairing, the order parameter (or wave function) is far more complicated than that of conventional superconductors and superfluid ^{4}He. The order parameter of the superfluid ^{3}He joins two spaces: the orbital (or k space) and spin and can be expressed as:

$$\Psi(\boldsymbol{k}) = \Psi_{\uparrow\uparrow}(\hat{k})|\uparrow\uparrow\rangle + \Psi_{\downarrow\downarrow}(\hat{k})|\downarrow\downarrow\rangle + \sqrt{2}\Psi_{\uparrow\downarrow}(\hat{k})(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \ , \tag{1}$$

where $\hat{k}$ is a unit vector in k space defining a position on the Fermi surface, $\Psi_{\uparrow\uparrow}(\hat{k})$, $\Psi_{\downarrow\downarrow}(\hat{k})$ a $\Psi_{\uparrow\downarrow}(\hat{k})$ are amplitudes of the spin sub-states operators determined by its projection $|\uparrow\uparrow\rangle$, $|\downarrow\downarrow\rangle$ a $(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$ on a quantization axis z.

The order parameter is more often written in a vector representation as a vector $\boldsymbol{d}(\boldsymbol{k})$ in spin space. For any orientation of the $\boldsymbol{k}$ on the Fermi surface, $\boldsymbol{d}(\boldsymbol{k})$ is in the direction for which the Cooper pairs have zero spin projection. Moreover, the amplitude of the superfluid condensate at the same point is defined by $|\boldsymbol{d}(\boldsymbol{k})|^2 = 1/2tr(\Psi\Psi^H)$. The vector form of the order parameter $\boldsymbol{d}(\boldsymbol{k})$ for its components can be written as:

$$d_\nu(\boldsymbol{k}) = \sum_\mu A_{\nu\mu}\boldsymbol{k}_\mu \ , \tag{2}$$

where ν (1,2,3) are orthogonal directions in spin space and μ (x,y,z) are those for orbital space. The matrix components $A_{\nu\mu}$ are complex and theoretically each of them represents possible superfluid phase of ^{3}He. Experimentally, however, only three are stable (see Fig. 2).

Looking at the phase diagram of ^{3}He we can see the presence of two main superfluid phases: A - phase and B - phase. While B - phase consists of all three spin components, the A - phase does not have the component $(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$. There is also a narrow region of the A1 superfluid phase which exists only at higher pressures and temperatures and in nonzero magnetic field. The A1 -

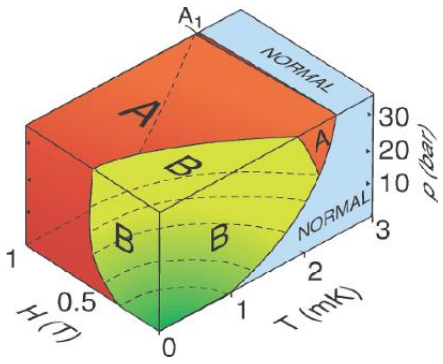

Fig. 2. Phase diagram of the ^{3}He

phase has only one spin component $|\uparrow\uparrow\rangle$. The phase transition from N - phase to the A or B - phase is a second order transition, while the phase transition between the superfluid A and B phases is of first order.

As one can see from Fig. 2, the B - phase occupies a low field region and it is stable down to the lowest temperatures. In zero field, the B - phase is a pure manifestation of p-wave superfluidity suggested by Balian and Werthamer (know as the BW state as well) [11]. Having equal numbers of all possible spin and angular momentum projections, the energy gap separating ground state from excitation is isotropic in k space.

The A - phase is preferable at higher pressures and temperatures in zero field. In limit T$\rightarrow$ 0K, the A - phase can exist at higher magnetic fields (above 340 mT) at zero pressure and this critical field needed for creation of the A - phase rises up as the pressure increases. The A - phase is theoretically described by the ABM state discussed by Anderson, Brinkman and Morel [12]. As it was mentioned above, the absence of $S_z = 0$ component of the Cooper pairs leads to anisotropy in the energy gap of the quasiparticle excitations. In this phase, all Cooper pairs have orbital momenta orientated in a common direction defined by the vector $\hat{l}$, that is the direction in which the energy gap is reduced to zero. It results in a remarkable difference between these superfluid phases, as it can be seen from a spectrum of excitation (see Fig. 3). The B - phase has an isotropic gap, while the A - phase energy spectrum consists of two Fermi points i.e. points with zero energy gap. The difference in the gap structure leads to the different thermodynamic properties of quasiparticle excitations in the limit T$\rightarrow$ 0K. The density of excitation in the B - phase falls down exponentially with temperature as $\exp(-\Delta/k_BT)$, where k_B is the Boltzman constant. At the lowest temperatures their density is so low that the excitations can be regarded as a non-interacting gas with a mean free path of the order of kilometers. On the other hand, in A - phase the Fermi points (or nodes) are far more populated with quasiparticle excitations. The nodes orientation in the $\hat{l}$ direction make the A - phase excitations almost perfectly one-dimensional. The presence of the nodes in the energy spectrum leads to a T^3 temperature dependence of the density of excitations and entropy. As a result, as T$\rightarrow$ 0K, the specific heat of the A - phase is far greater than that of the B - phase. In this limit, the A - phase represents a model system

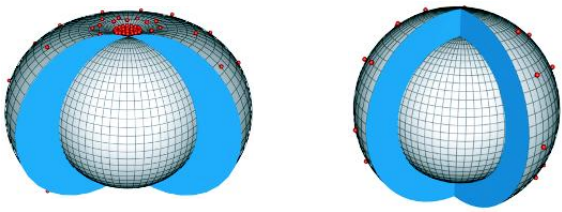

Fig. 3. Energy spectrum of quasiparticles in A - and B - phase. While the B - phase has isotropic energy gap, the A - phase consists of Fermi points

for a vacuum of the Standard model and B - phase is a model system for a Dirac vacuum. Therefore, it is extremely interesting to study the properties of these two different superfluid states in one liquid, as well as, when they are in contact.

As an example of a few experiments which recently were carried out with the aim of studying the properties of the A-B boundary under various conditions, we can mention the Lancaster [13–15] and the Helsinki [16] experiments. The Lancaster's experiments studied the thermodynamics of the A-B phase transition, a structure of A - phase gap nodes and measured the surface tension between these two different states of vacua. The experiment realized by Helsinki group showed that the penetration of the vortices from the A - phase to B - phase starts with the aid of the Kelvin-Helmholtz instability due to counterflow just on the A-B boundary.

In experiments with superfluid ^{3}He phases, application of different external forces can excite the collective modes of the order parameter representing so called Bose excitations, while the Fermi excitations are responsible for the energy dissipation. Coexistence and mutual interactions of these excitations in the limit $T \to 0K$ (in limit of low energies), can be described by quantum field theory, where Bose and Fermi excitations represent Bose and Fermi quantum fields. Thus, ^{3}He research has a much broader impact by offering the possibility of experimentally investigating quantum field/cosmological theories via their analogies with the superfluid phases of ^{3}He.

Now we would like to show and discuss some already experimentally observed phenomena in superfluid ^{3}He phases which could be considered as analogues of cosmological phenomena.

3 States with Coherent Spin Precession in ^{3}He-B and their Cosmological Analogues

Another, but fundamental, signature of any superfluid (or superconducting) system is the presence of supercurrents if the system is not in a coherent state. In general, the supercurrents are created by the phase gradients of the order parameter. The order parameter describing the superfluid ^{3}He phase is complex and consists of an orbital and spin part. Therefore, any inhomogeneity in the spin part of the order parameter leads to the generation of spin supercurrents. This inhomogeneity may simply be created by the application of an inhomogeneous magnetic field on a sample of superfluid ^{3}He and by applying NMR techniques. The amplitudes of the spin currents are proportional to the inhomogeneity tensor in spin space $\Omega_{j\beta}$ [8]:

$$J_{i\alpha} = \rho_{ij\alpha\beta}\Omega_{j\beta} \ , \tag{3}$$

where $\rho_{ij\alpha\beta}$ is the spin density tensor. The spin supercurrents transfer magnetization over the sample much faster than the magnetic relaxation processes can do and in such a way mask them.

However, these spin supercurrents together with the aid of the dipole-dipole interaction which works as a negative feedback, can lead to the spontaneous creation of various types of coherently precessing spin states in ^{3}He-B [17–19]. The crucial property of these dynamic states is that the coherent phase of the spin precession is established spontaneously in the presence of an inhomogeneous external magnetic field. As the spin precession is established spontaneously, the precessing spin system of the superfluid ^{3}He-B is in a dynamic equilibrium state corresponding to the minimum of the total energy. This is considered as an equivalent of the phase-coherent Bose condensate and it is referred to as a magnetic or spin superfluidity. The macroscopic states with coherent spin precession are regarded as its experimental evidence.

The dynamic states with coherent spin precession can also be considered as a model system to study their analogues to cosmology (see below).

The first state with coherent spin precession, a homogeneously precessing domain (HPD), was discovered by the Moscow group in 1984 (for a comprehensive overview see [17]) and theory of HPD was simultaneously developed by I. Fomin [20]. The HPD can be created using both a pulse and continuous NMR technique. To explain how the HPD is generated, let's assume that a sample of superfluid ^{3}He-B occupies a cylindrical experimental cell in which a magnetic field $\boldsymbol{B_0}$ together with a linear field gradient ∇ B are applied. Under conditions of a cw-NMR experiment, the presence of a linear magnetic field gradient ∇B creates inhomogeneity in the initially homogeneous spin part of the order parameter since the spins are precessing with various rates. In an attempt to relax the order parameter gradient, the condensate responds by generating the spin supercurrents which redistribute spins through the sample of ^{3}He-B. The supercurrents have the effect of increasing the tipping angle at the low field side of the sample and reducing it on the opposite side. The spins on the high field side of the sample have their tipping angle reduced, i.e. they are brought closer to the field direction, and once parallel there is no longer any precession, thus creating a stationary domain (SD). Conversely, the spins on the low field side continue tipping until a magic - Leggett angle ($\sim 104^\circ$) is reached, above which a dipole torque increases the rate of spin precession. This generates a spin current in the opposite direction to that due to field inhomogeneity, compensating it, and finally giving rise to a dynamic equilibrium state with coherent spin precession, the homogeneously precessing domain. These two domains are separated by a domain wall of thickness $\lambda_F \sim (1/\nabla B)^{1/3}$ which sits at the position z_0, where the frequency of the spin precession ω_l i.e. the angular frequency of the rf-field, fulfills a Larmor resonance condition $\omega_l = \gamma(B_0 - z_0 \nabla B)$, where γ is the gyromagnetic ratio. When the domain is excited continuously, the exciting rf-field compensates for the energy losses from magnetic relaxation processes in the precessing domain, and the HPD can be maintained indefinitely. Because the resonance frequency ω_l is constant, the position of the domain wall z_0 can easily be controlled by a steady magnetic field B_0.

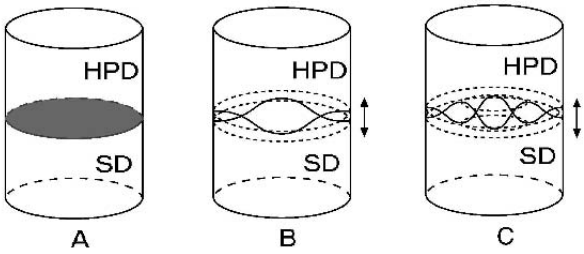

Fig. 4. A schematic visualization of the HPD in the equilibrium state (**A**), first (**B**) and the second (**C**) axial surface oscillation modes of the HPD

As it was mentioned above, the HPD occupies a state of minima of energy. As a result, any disturbances or external forces applied to this system will deviate the HPD from its equilibrium (or ground) state, thus creating the gradients in the phase of the spin precession, i.e. the phase gradient of the spin part of the order parameter. These gradients will again give rise to the spin supercurrents, a powerful mechanism which, in general, will restore the system to a new dynamical equilibrium state. The restoring process may be accompanied by generation of the spin precession waves - oscillations of the spin distribution around the equilibrium state (see Fig. 4). This phenomenon of the spin precession waves together with spin supercurrents can be, in principle, used to create and study a spin wave analogue of a black hole.

3.1 A Spin Wave Analogue of a Black Hole

In their paper [22], R. Schuetzhold and W. Unruh have suggested the realization of an experiment to observe the gravity wave analogues of a black hole on surface waves of a flowing liquid. If one can control the flow velocity of the liquid and simultaneously can excite its surface waves, one can reach a point at which the velocity of liquid flow exceeds that of the surface waves. For observers who get the information by means of the surface waves, at this point an event horizon forms and the wave (or information) never gets to them.

In paper [23] it had been shown that using the HPD generated by cw-NMR technique one can excite the spin precession waves. These waves are analogues to those on the surface of liquid in gravitational field. Following [21], the dispersion relation of the HPD surface oscillation modes can be expressed as:

$$\Omega_{Sm}^2 = Q_{im}\frac{c_1 c_2}{RB}\tanh\left(Q_{im}L\frac{c_2}{R2c_1}\right)\nabla B\,, \tag{4}$$

where R is the radius of the cell, B is the resonance field, Q_{im} are the m-th nonzero roots of equation $J_i'(x) = 0$ (where $J_i(x)$ is the Bessel function for integer i; for example, with $i = 0$ for an axial and $i = 1$ for planar modes), L is the HPD length, $c_1^2 = (5c_\perp^2 - c_\parallel^2)/4$ and $c_2^2 = (5c_\perp^2 + 3c_\parallel^2)/4$, where $c_\perp$ and $c_\parallel$ are the spin wave velocities with respect to the direction of the magnetic field.

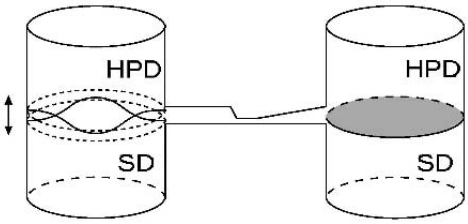

Fig. 5. A schematic visualization of the experimental arrangement for the creation a spin wave analogue of a black hole. Two HPD's are connected through a channel with variable depth. One HPD serves as a source of the spin precession waves which propagate through the channel on a spin flow controlled by the difference in the phase between the rf-excitation field of individual HPD's

This dispersion relation is similar to that of the surface waves on a liquid in the gravitational field as used in [22]. In this case, however, the magnetic field gradient ∇B plays a role of the gravitational field.

A spin wave analogue of the black hole can be, in principle, experimentally observed using two HPD's connected with a channel with variable depth and simultaneously create both: a continuous spin flow and spin precession waves (see Fig. 5). The continuous spin flow between HPD's can simply be adjusted and controlled by setting of the phase difference between the rf excitation fields which are used for the HPD's generation i.e. by setting the phase difference of the spin precession in individual HPD's. Simultaneously, one of the HPD serves as a source of the spin precession wave excited on its resonance frequency. The wave will propagate through the channel on the spin flow, the velocity of which changes due to variable depth of channel. This arrangement - two HPD's connected by the channel with variable depth - is similar to that suggested in [22] and, in principle, could be used for the creation and the study of the spin wave analogue of a black hole.

3.2 A Persistent Precessing Domain as an Analogue of Q-Ball

A Q-ball is a long-lived state with coherent precession. The precession frequency (or rotation frequency) and stability are determined by the conservation of the global charge, e.g. a baryon charge in the case of a cosmological Q-ball, or conservation of the spin projection into the direction of the applied field in the case of a condensed matter Q-ball analogue. Based on this description of a Q-ball, a persistent precessing domain (PPD) excited in superfluid ^{3}He-B seems to be (or can be considered as) a clear representative.

The homogeneously precessing domain and a domain with spin 1/2 [18] are well know, however, the physical origin of a PPD is not very well understood. The PPD was observed by the Lancaster group for the first time in 1992 [19]. It seems that the PPD is a homogeneously spin precessing structure kept insulated from the wall of the cell by the texture of the order parameter

(or a surface energy). Due to a negligibly small density of Fermi excitations, the dissipation is very small and a PPD life time may be as long as several tens of minutes. As a PPD slowly decays, the frequency of the induced signal rises. Moreover, if the PPD is excited using a cw-NMR technique, one has to use frequencies higher than the Larmor frequency (so called off resonant excitation) [26]. However, experiments made during the last five years showed that very reproducible PPD's can be generated in a local magnetic field minimum and to excite them it is enough to apply only a noise signal from a noise generator instead of the rf-pulse [27, 28].

What kind of spin structure is the PPD? This is still an open question. Based on the fact that reproducible PPD's are excited in a field minimum, the Leggett equations of the spin dynamics of ^{3}He-B at local field minima were solved [29]. The solutions were found to resemble some of the properties of the PPD, e.g. when the PPD shrinks its precessing frequency rises.

As a PPD can only be excited in the limit of a vanishingly small quasiparticle density, it seems that the orbital degrees of freedom are defrosted. A motion of the orbital momenta gives rise to a orbital viscosity that leads to energy dissipation. In their paper [30], S. Fisher and N. Suramlishvilli calculated the orbital viscosity of the B - phase in both the hydrodynamic and ballistic regime. They showed that orbital viscosity in the limit $T \rightarrow 0$K becomes vanishingly small, which gives rise to the possibility of coupled spin-orbital dynamics. The orbital angular momentum motion may be induced via the spin-orbit coupling by the spin precession. It looks like this could be the case for a PPD. Taking into account a temperature dependence of the orbital viscosity together with the fact that the mean free path of quasiparticles is limited by the size of an experimental cell D, they found the relaxation time to be:

$$\tau \sim \frac{D}{\pi v_f} \sqrt{\frac{k_B T}{\Delta}} \exp\left(\Delta / k_B T\right), \tag{5}$$

where v_f is the Fermi velocity.

In Fig. 6 the temperature dependence scaled by T_c of the PPD lifetime is shown [31]. The PPD was excited away from the vertical wall of the cell. The line represents a fit of the experimental data using equation (5). Instead of a typical size of the cell D, a longer mean free path was taken into account because quasiparticles can travel a longer distance due to reflections from the wall of the cell. However, when the PPD was pushed closer to the hemispherical end cap of the cell, the PPD lifetime dramatically shortened by a few orders of magnitude [27]. In this region the PPD lifetime was observed to be temperature independent at the lowest temperatures, showing that additional relaxation processes are temperature independent [31]. A possible explanation of this effect can be found if one again considers the orbital degree of freedom in superfluid ^{3}He-B at ultralow temperatures i.e. a coherent precession of both the spin and orbital angular momentum. Such a precession was theoretically

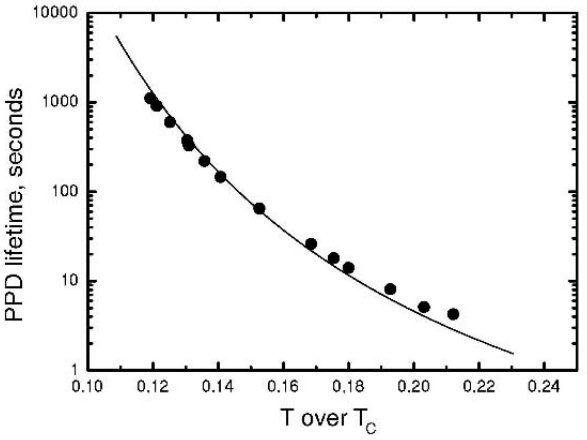

Fig. 6. A temperature dependence scaled by T_c of the PPD lifetime. The line represents a fit of the experimental data using equation (5)

predicted to have a temperature independent relaxation mechanism due to pair-breaking, analogous to Schwinger mechanism [32].

4 Search for an Unruh Effect Analogue in ^{3}He-B

W. Unruh [24] studied the behaviors of a uniformly accelerated detector which is moving in a physical vacuum (vacuum of a quantum field). He showed that an accelerated detector "feels or sees" this vacuum state as a thermal bath and detects it as black body radiation with an effective temperature T_U:

$$T_U = \frac{\hbar a}{2\pi k_B c} , \tag{6}$$

where a is the detector acceleration, $\hbar$ and k_B is the Planck and Boltzman constant, respectively, and c is the velocity of light in vacuum.

The superfluid ^{3}He-B phase can be regarded as the mixture of two components: the superfluid component at temperature T=0K represents the vacuum state and a normal component i.e. the gas of excitations whose total energy represents the real temperature background T of the liquid. The overall density of excitations ρ_n falls down very rapidly with temperature as (see Fig.7):

$$\rho_n = 2N(0) \int_{\Delta}^{\infty} \frac{\epsilon}{\sqrt{\epsilon^2 - \Delta^2(T)}} f(\epsilon, T) d\epsilon , \tag{7}$$

where $f(\epsilon, T)$ is the Fermi-Dirac distribution function, $N(0)$ is the density of states of the normal liquid at the Fermi level per unit volume and per spin and $\Delta(T)$ is the energy gap in the spectrum of excitations which below $0.25T_c$ can be approximated as $\Delta(0) \sim \Delta = 1.76 k_B T_c$.

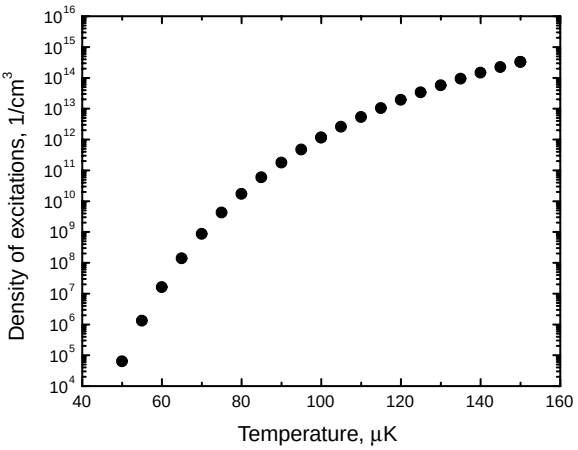

Fig. 7. The temperature dependence of the density of excitations in ^{3}He-B at 0 bar calculated using expression (7)

4.1 Vibrating Wire – An Accelerated Detector in Superfluid ^{3}He-B

The physical properties of the excitations from the vacuum state of superfluid ^{3}He-B can be investigated using a vibrating wire resonator. The vibrating wire is a simple device widely used to study of quantum liquids (^{4}He and ^{3}He). Usually it is a superconducting wire - a very thin fiber - bowed into a semicircle (or other shape e.g. rectangular) and immersed in the liquid (see Fig. 8). With the aid of a static magnetic field **B** orientated along the plane of the wire loop and an AC current I_0 driving the wire, the wire vibrates on resonance as a consequence of an acting Lorentz force. The motion of the wire per unit of mass and wire length is described by following equation:

$$d^2x/dt^2 + \gamma dx/dt + \omega_0^2 x = f\exp(i\omega t)\ , \tag{8}$$

where parameter $f = I_0 B/m$ describes the driving force per mass, $m = \pi r^2 \rho_w$ is the mass per unit length of wire with r and ρ_w being the radius and density of the wire, respectively. The second term in equation (8) characterizes a damping force of the fluid acting against the wire motion and it is assumed that it is a linear function of the wire velocity v. The constant γ ($\gamma = \gamma_2 + i\gamma_1$) is the damping constant, where γ_2 refers to the dissipative component of the damping force. The γ_1 characterizes its reactive component associated with the fluid backflow around the wire and effectively gives the wire a greater mass. The last term in equation (8) is the restoring force of the wire, where ω_0 is the wire resonance frequency in vacuum. The steady state solution of equation (8) is very well known and leads to a Lorentz shape of the absorption and dispersion curves dependent on the frequency. The width of the absorption curve $\gamma_2 = 2\pi\Delta f_2$ is related to the damping force acting on the wire motion:

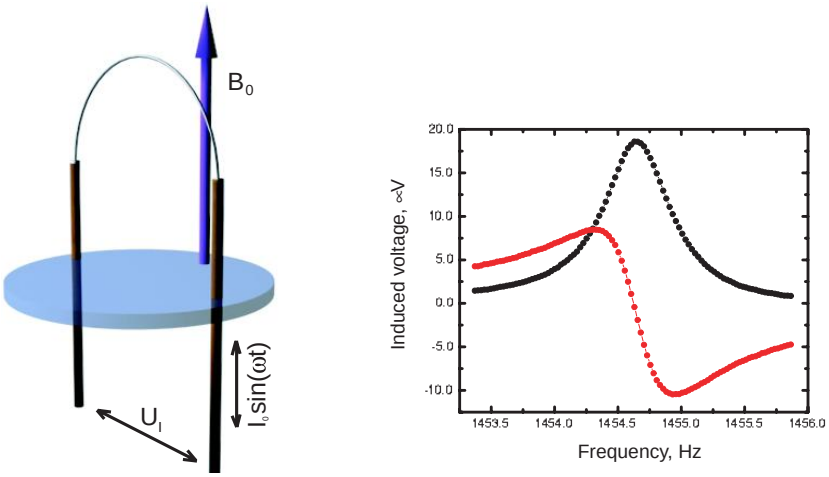

Fig. 8. A typical experimental arrangement of one fiber vibrating wire made from a multifilament superconductor. On the right: the resonance characteristics of the wire

$$F = 2\pi\Delta f_2 v. \tag{9}$$

On the other hand, the frequency shift of the resonance frequency from its vacuum value Δf_1 is associated with γ_1 as $\gamma_1 = 4\pi\Delta f_1$.

The vibrating wire moves in the direction which is almost perpendicular to the applied magnetic field $\mathbf{B}$, so an additional voltage U_i is induced. Its magnitude can be found by applying Faraday's law in form $U_i = kBlv$, where l is the wire length and the constant k characterizes the geometric shape of the wire.

Thus, the vibrating wire may be considered as an accelerated detector which moves with an effective acceleration through the vacuum of excitations and, in principle, the wire could be used for the generation and detection of a condense matter analog of the Unruh effect in superfluid ^{3}He-B.

Assuming that a vibrating wire moves at the resonance frequency Ω with velocity v then the corresponding Unruh temperature T_U can be calculated using expression [2]:

$$T_U(v) = \frac{\hbar\Omega}{4k_B[\ln(v_L/v + \sqrt{v_L^2/v^2 - 1}) - \sqrt{1 - v^2/v_L^2}]}\,. \tag{10}$$

The role of the velocity of light in the superfluid ^{3}He-B is given by the Landau velocity v_L ($v_L = \Delta/p_f$, where p_f is the Fermi momentum). That is the velocity at which a microscopic object moving through the B - phase creates excitations assuming that the object does not: (i) perturb the local gap parameter and (ii) create a backflow. In Fig. 9, the dependence of the Unruh temperature T_U on the vibrating wire velocity for two resonance frequencies

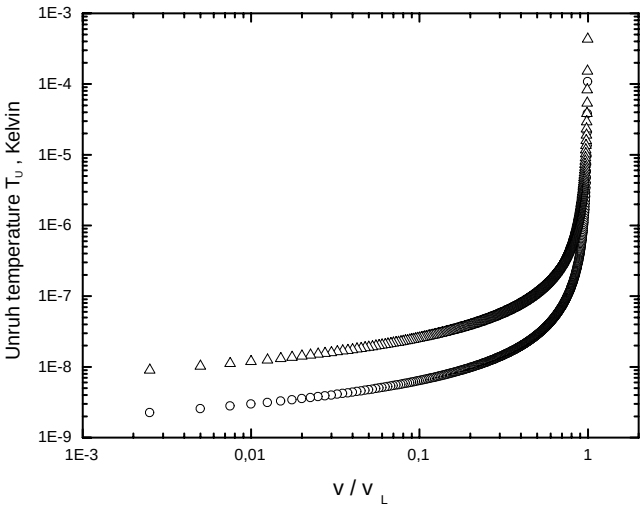

Fig. 9. Calculated Unruh temperature as function of the vibrating wire velocity for two resonance frequencies 1000 Hz (○) and 4000 Hz (△)

Ω is presented. As can be seen from Fig. 9, the Unruh temperature is highly enhanced at wire velocities very close to the Landau velocity. This means that before the Landau velocity is reached, one could observe the Unruh effect. Unfortunately, the motion of the macroscopic object like a vibrating wire in superfluid ^{3}He-B represents a more complex problem. The energy gap is suppressed within a few coherence lengths of the wire surface and the superfluid backflow around the wire must also be taken into account. Indeed, as the velocity of the wire rises, the energy of the superfluid component close to the wire is lifted by $2p_f v$ due to backflow until it reaches a threshold value $\Delta - p_f v$, i.e. the energy value of the excitations approaching the wire from behind (see Fig. 10). Having this energy at a critical velocity $v_c = \Delta/3p_f$, the superfluid component can be scattered on the wire roughnesses into empty energy states which cause a rapid increase of the wire damping [25]. This effect limits the wire velocity, with respect to the Landau velocity, and the ability to observe an additional friction due to the Unruh effect. An ordinary macroscopic vibrating wire should be driven at the velocities lower than $v_L/3$.

The motion of the wire in superfluid ^{3}He-B is restricted due to a mutual interaction between the excitations and the wire. This interaction leads to a damping of the wire motion. A total ordinary damping force acting on the wire, in general, consists of three terms: $F = F_I + F_C + F_T$, where F_I is the intrinsic damping force of the wire, F_C is the damping force due to pair breaking if the wire velocity is above the critical one, and F_T is thermal damping force due to a collision of the wire with existing excitations. Here we shall assume that we have an ideal wire with $F_I = 0$. The F_C term is close to zero because the wire velocity v is always less than the velocity v_c. The last

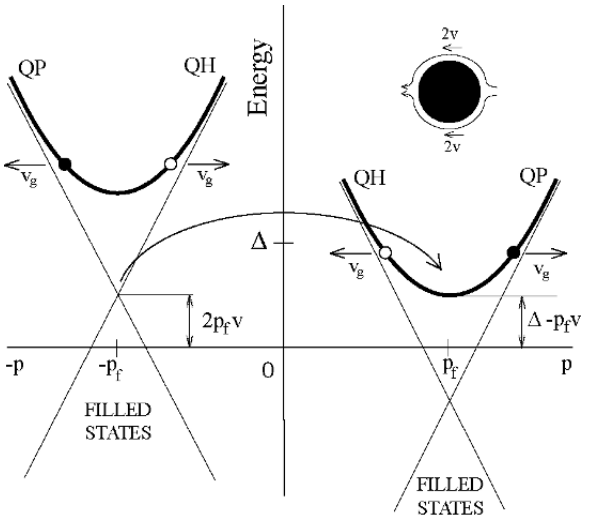

Fig. 10. Quasiparticle energy spectrum as viewed from the reference frame of the wire. An observer connected to the wire sees the energy of the quasiparticles moving in his direction as lower by $p_f v$, whereas those moving opposite to him are lifted by $p_f v$. He also sees that the energy of the superfluid component due to backflow is lifted by $2p_f v$. In the laboratory frame the wire (the black dot) moves from left to right

term F_T per unit area can be expressed as: [25]:

$$F_T = p_f \langle n v_g \rangle \left[1 - \exp\left(\frac{-p_f v}{k_B T} \right) \right], \tag{11}$$

where v_g and n are the group velocity and the number of excitations per unit volume, respectively, v is the wire velocity and

$$\langle n v_g \rangle = n(p_f) k_B T \exp(-\Delta / k_B T) \tag{12}$$

represents the quasiparticle flux with $n(p_f)$ being the density of states in momentum space. The origin of this non-linearity with respect to the wire velocity arises due to an Andreev scattering of the excitations from the flow field around the moving wire. At low velocity, the damping force per unit area of the wire becomes linear in velocity v:

$$F_T = p_f \langle n v_g \rangle \frac{p_f v}{k_B T} = n(p_f) v p_f^2 \exp(-\Delta / k_B T). \tag{13}$$

For very low temperatures the damping force F_T depends only on temperature and is constant at constant velocity and at constant temperature. The total damping force acting on the wire can easily be determined from the width Δf_2 of the vibrating wire resonance curve.

Vibrating Wire and Unruh Effect Analogue

Let an ideal vibrating wire ($F_I = 0$) move through superfluid ^{3}He-B as a consequence of the Lorentz force. The wire motion is damped due to the interaction with excitations representing the thermal background and with the new ("Unruh") excitations generated via tunnelling processes because of the wire acceleration. This additional Unruh damping, per unit area of the wire, can be expressed as:

$$\dot{Q_U} = (\Delta + k_B T_U)\langle n v_g \rangle_U \approx \Delta \langle n v_g \rangle_U, \tag{14}$$

where $< n v_g >_U$ represents the quasiparticle flux of 'Unruh' excitations each having energy Δ because $\Delta >> k_B T_U$. The quasiparticle flux of the 'Unruh' excitations is proportional to $n(p_f) k_B T_U \exp(-\Delta / k_B T_U)$ and contributes an additional damping force, $F_U = n(p_f)\Delta k_B T_U \exp(-\Delta/k_B T_U)/v$, acting on the wire. The total damping force is now $F = F_T + F_U = F_T(1 + F_U/F_T)$, where the second term in brackets represents the contribution to the damping force due to the Unruh effect scaled by thermal damping:

$$\frac{F_U}{F_T} = \frac{k_B T_U}{p_f v} \frac{v_L}{v} \exp\left(-\frac{\Delta}{k_B} \frac{(T - T_U)}{T T_U}\right). \tag{15}$$

Let's assume that we can neglect the backflow and the ideal wire moves with a velocity very close to v_L i.e. $v \simeq v_L$. Then the Unruh temperature T_U can simply be adjusted by the wire velocity (see Fig. 9) and we can express it as $T_U = aT$ (in this case the pre-factor in equation (15) will be $aT/1.76T_c$ within 10% accuracy). Figure 11 shows a dependence of F_U/F_T for various values of the Unruh temperatures, $T_U = aT$, as a function of temperature T.

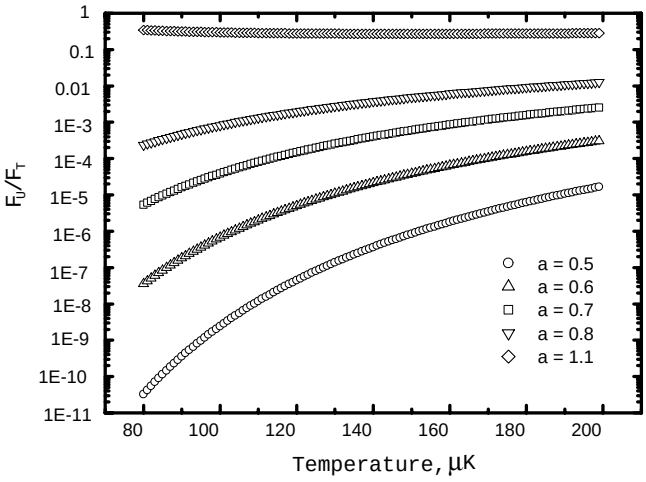

Fig. 11. A dependence of F_U/F_T for various values of the Unruh temperatures, $T_U = aT$, as a function of the temperature. The dependence is calculated for 0 bar

The sensitivity of vibrating wire techniques can be expressed in terms of a relative change of the resonance curve width as $\delta\Delta f_2/\Delta f_2$ and is of order of 10^{-4}. The ratio F_U/F_T is proportional to the relative change $\delta\Delta f_2/\Delta f_2$ and thus the above sensitivity of the measurement gives a limit for detection of the additional damping force caused by "Unruh" excitations.

In a real experiment the vibrating wire oscillates at resonance frequencies up to a few kHz. If the wire velocity is $\sim 0.25 v_L$ with a resonance frequency of 4000 Hz (see Fig. 9), then the corresponding Unruh temperature T_U is slightly above 40 nK. However, nowadays the superfluid ^{3}He-B can be cooled to low temperatures of about 80-100 μK, which is $\sim 10^4$ times higher than the Unruh temperature discussed above. Due to the small values of the wire velocities (less than v_c) and low resonance frequencies Ω as well, the number of "Unruh" excitations (and the Unruh temperature) are very small in comparison with those representing the background temperature T. Therefore, the contribution of the "Unruh" excitations to the total damping force acting on the wire is well below the sensitivity of present day detection techniques.

A question which one could ask is what should be done to observe the condense matter analog of the Unruh effect in superfluid ^{3}He-B by means of a vibrating wire or, perhaps by the use of any other method and in such a way confirm the theoretical prediction? Apart from an 'easy to suggest and hard to make' solution of cooling of the superfluid ^{3}He-B well below 100 μK, in case of the vibrating wire, the main problem seems to be the reduction of the backflow around the wire. The reduction of the backflow would allow one to drive the wire at velocities closer to the Landau velocity and, by approaching this velocity, the wire could detect the presence of the Unruh effect. It is a matter of speculation, however, whether the progress in (nano) technology will allow the use of a nano-wire (or nano-object) of diameter comparable to the coherence length of the order parameter of ^{3}He-B. Another possible way to observe the condense matter analog of the Unruh effect in superfluid ^{3}He-B could be an application of NMR methods operating at much higher frequencies.

5 Discussion

I hope that this is not end of the story and in near future many interesting condensed matter experiments related to cosmology will be performed. Not only with superfluid ^{3}He phases, but also with superconductors, Bose-Einstein condensates and other condensed matter systems. Due to the impossibility to make an active, a controllable cosmological experiment, it seems that this is only the way we can experimentally verify our imagination about fundamental cosmological phenomena, of course with limitations. So, we shall keep watching this analogous space.

I would like to thank to Ian Bradley, Shaun Fisher, Rich Haley, Martin Kupka, George Pickett, Ray Simmonds, Ralf Schuetzhold, William Unruh

and Grisha Volovik for stimulating discussions. This work is supported by ESF Project COSLAB and by grant agencies: APVV under grant 51-016604, VEGA 02/6168/06 and CE I-2/2003.

References

1. G. E. Volovik: Universe in Helium-3 Droplet, Oxford University Press 2003
2. G. E. Volovik: Exotic properties of superfluid ^{3}He, World Scientific 1992
3. D. I. Bradley et al., Phys. Rev. Lett. **75** (10) 1887 (1995)
4. C. Bauerle, et al., Nature **382** 332 (1995)
5. V. M. H. Ruutu, et al., Nature **382** 334 (1995)
6. T. W. B. Kibble, J. Phys. A 9, 1387 (1976)
7. W. H. Zurek, Nature 317, 505 (1985)
8. A. J. Leggett, Rev. Mod. Phys. **47** 331 (1975)
9. D. Vollhardt and P. Wölfle: The Superlfuid Phases of Helium 3, Taylor and Francis (1992)
10. E. R. Dobbs: Helium Three, Oxford Science Publications (2000)
11. R. Balian, N. R. Werthamer, Phys. Rev. **131** 1553 (1963)
12. P. W. Anderson and P. Morel Physica **26** 671 (1960). P. W. Anderson and P. Morel Phys. Rev. **123** 1911 (1961). P. W. Anderson and W. F. Brinkman Phys.Rev.Lett. 30 1108 (1973)
13. M. Bartkowiak et al. Phys. Rev. Lett. 83 3462 (1999)
14. M. Bartkowiak et al. Phys. Rev. Lett. 85 4321 (2000)
15. M. Bartkowiak et al. Phys. Rev. Lett. 93 (4) 045301 (2004)
16. R. Blaauwgeers et al., Phys. Rev. Lett. 89 (15) 155301 (2002)
17. Yu. M. Bunkov, in *Progress Low Temp. Physics* Vol. 14 edited by W. P. Halperin (Elsevier Science Publisher, Amsterdam (1995) p. 69
18. V. V. Dmitriev et al., Phys. Rev. Lett. **78**, 86 (1997)
19. Yu. M. Bunkov et al. Phys. Rev. Lett. **69**, 3092 (1992)
20. I. A. Fomin, in *Helium three* edited by W. P. Halperin and L. P. Pitaevskii, Elsevier Science Publishers, Amsterdam (1990) p. 609
21. I. A. Fomin, JETP Lett. **43**, 171 (1986)
22. R. Schuetzhold, W. Unruh, Phys. Rev. D 66 044019 (2002)
23. E. Gažo et al., Phys. Rev. Lett. **91** (5) 055301 (2003)
24. W. G. Unruh, Phys. Rev. D **14** 870 (1976)
25. S. N. Fisher, et al., Phys. Rev. Lett. **63** 2566 (1989)
26. D. J. Cousins et al., Phys. Rev. Lett. **82**, 4484 (1999)
27. S. N. Fisher et al., Physica B, **329** 80 (2003)
28. D. I. Bradley et al., J. Low Temp. Phys. **134** 351 (2004)
29. M. Kupka and P. Skyba, Physics Letters A **317** 324 (2003)
30. S. N. Fisher and N. Suramlishvilli, J. Low Temp. Phys. **138** Nos.3/4 771 (2005)
31. D.I. Bradley, S. N. Fisher et al., to be published
32. N. Schopohl, G.E. Volovik, Ann. of Phys. **215** 372 (1991)

Dynamical Aspects of Analogue Gravity: The Backreaction of Quantum Fluctuations in Dilute Bose-Einstein Condensates

U. R. Fischer

Eberhard-Karls-Universität Tübingen, Institut für Theoretische Physik,
Auf der Morgenstelle 14, D-72076 Tübingen, Germany

1 Analogue Gravity: An Overview

1.1 The Concept of an Effective Space-time Metric

Curved space-times are familiar from Einstein's theory of gravitation [1], where the metric tensor $\mathsf{g}_{\mu\nu}$, describing distances in a curved space-time with local Lorentz invariance, is determined by the solution of the Einstein equations. A major problem for an experimental investigation of the (kinematical as well as dynamical) properties of curved space-times is that generating a significant curvature, equivalent to a (relatively) small curvature radius, is a close to impossible undertaking in manmade laboratories. For example, the effect of the gravitation of the whole Earth is to cause a deviation from flat space-time on this planet's surface of only the order of 10^{-8} (the ratio of Schwarzschild and Earth radii). The fact that proper gravitational effects are intrinsically small is basically due to the smallness of Newton's gravitational constant $G = 6.67 \times 10^{-11}\,\mathrm{m^3 kg^{-1} sec^{-2}}$. Various fundamental classical and quantum effects in strong gravitational fields are thus inaccessible for Earth-based experiments. The realm of strong gravitational fields (or, equivalently, rapidly accelerating a reference frame to simulate gravity according to the equivalence principle), is therefore difficult to reach. However, Earth-based gravity experiments are desirable, because they have the obvious advantage that they can be prepared and, in particular, repeated under possibly different conditions at will.

A possible way out of this dilemma, at least inasmuch the kinematical properties of curved space-times are concerned, is the realization of *effective* curved space-time geometries to mimic the effects of gravity. Among the most suitable systems are Bose-Einstein condensates, i.e., the dilute matter-wave-coherent gases formed if cooled to ultralow temperatures, where the critical temperatures are of order $T_c \sim 1\,\mathrm{nK} \cdots 1\,\mu\mathrm{K}$; for reviews of the (relatively) recent status of this rapidly developing field see [2–4]. In what follows, it will

U.R. Fischer: *Dynamical Aspects of Analogue Gravity: The Backreaction of Quantum Fluctuations in Dilute Bose–Einstein Condensates*, Lect. Notes Phys. **718**, 93–113 (2007)
DOI 10.1007/3-540-70859-6_5

be of some importance that Bose-Einstein condensates belong to a special class of quantum perfect fluids, so-called superfluids [5].

The curved space-times we have in mind in the following are experienced by sound waves propagating on the background of a spatially and temporally inhomogeneous perfect fluid. Of primary importance is, first of all, to realize that the identification of sound waves propagating on an inhomogeneous background, which is itself determined by a solution of Euler and continuity equations, and photons propagating in a curved space-time, which is determined by a solution of the Einstein equations, is of a *kinematical* nature. That is, the space-time metric is fixed externally by a background field obeying the laws of hydrodynamics (which is prepared by the experimentalist), and not self-consistently by a solution of the Einstein equations.

As a first introductory step to understand the nature of the kinematical identity, consider the wave equation for the velocity potential of the sound field ϕ, which in a homogeneous medium at rest reads

$$\left[\frac{1}{c_s^2}\frac{\partial^2}{\partial t^2} - \Delta\right]\phi = 0\,, \tag{1}$$

where c_s is the sound speed, which is a constant in space and time for such a medium at rest. This equation has Lorentz invariance, that is, if we replace the speed of light by the speed of sound, it retains the form shown above in the new space-time coordinates, obtained after Lorentz-transforming to a frame moving at a constant speed less than the sound speed. Just as the light field *in vacuo* is a proper relativistic field, sound is a "relativistic" field.[1] The Lorentz invariance can be made more manifest by writing equation (1) in the form $\Box\phi \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu\phi = 0$, where $\eta^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$ is the (contravariant) flat space-time metric (we choose throughout the signature of the metric as specified here), determining the fundamental light-cone-like structure of Minkowski space [6]; we employ the summation convention over equal greek indices $\mu, \nu, \cdots$. Assuming, then, the sound speed $c_s = c_s(\boldsymbol{x}, t)$ to be local in space and time, and employing the curved space-time version of the $3+1$D Laplacian $\Box$ [1], one can write down the sound wave equation in a *spatially and temporally inhomogeneous medium* in the generally covariant form [7,8]

$$\frac{1}{\sqrt{-\mathsf{g}}}\partial_\mu(\sqrt{-\mathsf{g}}\mathsf{g}^{\mu\nu}\partial_\nu\phi) = 0\,. \tag{2}$$

Here, $\mathsf{g} = \det[\mathsf{g}_{\mu\nu}]$ is the determinant of the (covariant) metric tensor. It is to be emphasized at this point that, because the space and time derivatives ∂_μ are covariantly transforming objects in (2), the primary object in the condensed-matter identification of space-time metrics via the wave equation (2) is the contravariant metric tensor $\mathsf{g}^{\mu\nu}$ [9]. In the condensed-matter understanding of

[1] More properly, we should term this form of Lorentz invariance *pseudorelativistic* invariance. We will however use for simplicity "relativistic" as a generic term if no confusion can arise therefrom.

analogue gravity, the quantities $\mathsf{g}^{\mu\nu}$ are *material-dependent* coefficients. They occur in a dispersion relation of the form $\mathsf{g}^{\mu\nu}k_\mu k_\nu = 0$, where $k_\mu = (\omega/c_s, \boldsymbol{k})$ is the covariant wave vector, with $\hbar\boldsymbol{k}$ the ordinary spatial momentum (or quasi-momentum in a crystal).

The contravariant tensor components $\mathsf{g}^{\mu\nu}$ for a perfect, irrotational liquid turn out to be [7,8,10]

$$\mathsf{g}^{\mu\nu} = \frac{1}{A_c c_s^2}\begin{pmatrix} 1 & \boldsymbol{v} \\ \boldsymbol{v} & -c_s^2\mathbf{1} + \boldsymbol{v}\otimes\boldsymbol{v}\end{pmatrix}, \tag{3}$$

where $\mathbf{1}$ is the unit matrix and A_c a space and time dependent function, to be determined from the equations of motion for the sound field (see below). Inverting this expression according to $\mathsf{g}^{\beta\nu}\mathsf{g}_{\nu\alpha} = \delta^\beta{}_\alpha$, to obtain the covariant metric $\mathsf{g}_{\mu\nu}$, the fundamental tensor of distance reads

$$\mathsf{g}_{\mu\nu} = A_c\begin{pmatrix} c_s^2 - \boldsymbol{v}^2 & \boldsymbol{v} \\ \boldsymbol{v} & -\mathbf{1}\end{pmatrix}, \tag{4}$$

where the line element is $ds^2 = \mathsf{g}_{\mu\nu}dx^\mu dx^\nu$. This form of the metric has been derived by Unruh for an irrotational perfect fluid described by Euler and continuity equations [7]; its properties were later on explored in more detail in particular by M. Visser [8]. We also mention that an earlier derivation of Unruh's form of the metric exists, from a somewhat different perspective; it was performed by Trautman [10]. The metric belongs to the Painlevé-Gullstrand class of metrics, historically introduced in Refs. [11].

The conformal factor A_c in (4) depends on the spatial dimension of the fluid. It may be unambiguously determined by considering the effective action of the velocity potential fluctuations above an inhomogeneous background, identifying this action with the action of a minimally coupled scalar field in $D+1$-dimensional space-time

$$\begin{aligned}\mathcal{A}_{\text{eff}} &= \int d^{D+1}x \frac{1}{2g}\left[\left(\frac{\partial}{\partial t}\phi - \boldsymbol{v}\cdot\nabla\phi\right)^2 - c_s^2(\nabla\phi)^2\right] \\ &\equiv \frac{1}{2}\int d^{D+1}x\sqrt{-\mathsf{g}}\mathsf{g}^{\mu\nu}\partial_\mu\phi\partial_\nu\phi\,,\end{aligned} \tag{5}$$

where the prefactor $1/g$ in front of the square brackets in the first line is identified with the compressibility of the (barotropic) fluid, $1/g = d(\ln\rho)/dp$, with p the pressure and ρ the mass density of the fluid; we assume here and in what follows that g is a constant independent of space and time so that $c_s^2 = g\rho$, as valid for a dilute Bose gas (see the following subsection). Using the above identification, it may easily be shown that the conformal factor is given by $A_c = (c_s/g)^{2/(D-1)} = (\rho/g)^{1/(D-1)})$, while the square root of the negative determinant is $\sqrt{-\mathsf{g}} = c_s(c_s/g)^{D+1/(D-1)} = (\rho^D/g)^{1/(D-1)}$. The case of one spatial dimension ($D = 1$) is special, in that the conformal invariance in two space-time dimensions implies that the classical equations of motion

are invariant (take the same form) for any space and time dependent choice of the conformal factor A_c, explaining the singular character of the conformal factor at the special value $D = 1$.

The line element $ds^2 = \mathrm{g}_{\mu\nu} dx^\mu dx^\nu$ gives us the distances travelled by the phonons in an effective space-time world in which the scalar field ϕ "lives". In particular, quasiclassical (large momentum) phonons will follow *light*-like, that is, here, *sound*-like geodesics in that space-time, according to $ds^2 = 0$. Noteworthy is the simple fact that the constant time slices obtained by setting $dt = 0$ in the line element are conformally flat, i.e. the quasiparticle world looks on constant time slices like the ordinary (Newtonian) lab space, with a Euclidean metric in the case of Cartesian spatial co-ordinates we display. All the intrinsic curvature of the effective quasiparticle space-time is therefore encoded in the metric tensor elements g_{00} and g_{0i} [12].

1.2 The Metric in Bose-Einstein Condensates

We assumed in Eq. (5) that the compressibility $1/g$ is a constant. This entails that the (barotropic) equation of state reads $p = \frac{1}{2} g \rho^2$. We then have, in the microscopic terms of the interaction between the particles (atoms) constituting the fluid, a contact interaction (pseudo-)potential, $V(\boldsymbol{x} - \boldsymbol{x}') = g\delta(\boldsymbol{x} - \boldsymbol{x}')$. This is indeed the case for the dilute atomic gases forming a Bose-Einstein condensate. Well below the transition temperature, they are described, to lowest order in the gas parameter $(\rho a_s^3)^{1/2} \ll 1$ [where a_s is the s-wave scattering length, assumed positive] by the Gross-Pitaevskiĭ mean-field equation for the order parameter $\Psi \equiv \langle \hat{\Psi} \rangle$, the expectation value[2] of the quantum field operator $\hat{\Psi}$ [3]:

$$i\hbar \frac{\partial}{\partial t} \Psi(\boldsymbol{x}, t) = \left[-\frac{\hbar^2}{2m} \Delta + V_{\mathrm{trap}}(\boldsymbol{x}, t) + g|\Psi(\boldsymbol{x}, t)|^2 \right] \Psi(\boldsymbol{x}, t) \; . \tag{6}$$

Here, V_{trap} denotes the one-particle trapping potential and the coupling constant g is related to the s-wave scattering length a_s via $g = 4\pi\hbar^2 a_s/m$ (in three spatial dimensions). The Madelung transformation decomposing the complex field Ψ into modulus and phase reads $\Psi = \sqrt{\rho} \exp[i\phi]$, where ρ yields the condensate density and ϕ is the velocity potential. It allows for an interpretation of quantum theory in terms of hydrodynamics [13]. Namely, identifying real and imaginary parts on left- and right-hand sides of (6), respectively, gives us the two equations

[2] Observe that $\langle \hat{\Psi} \rangle \neq 0$ breaks particle number conservation (the global U(1) invariance). We will come back to this point in Sect. 2, where we introduce a particle-number-conserving mean-field ansatz for the full quantum field operator, which has the number-conserving property that $\langle \hat{\Psi} \rangle = 0$, and for which the order parameter therefore does not equal $\langle \hat{\Psi} \rangle$.

$$-\hbar\frac{\partial}{\partial t}\phi = \frac{1}{2}m\boldsymbol{v}^2 + V_{\rm trap} + g\rho - \frac{\hbar^2}{2m}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\ , \tag{7}$$

$$\frac{\partial}{\partial t}\rho + \nabla\cdot(\rho\boldsymbol{v}) = 0\ . \tag{8}$$

The first of these equations is the Josephson equation for the superfluid phase. This Josephson equation corresponds to the Bernoulli equation of classical hydrodynamics, where the usual velocity potential of irrotational hydrodynamics equals the superfluid phase ϕ times $\hbar/m$, such that $\boldsymbol{v} = \hbar\nabla\phi/m$. The latter equation implies that the flow is irrotational save on singular lines, around which the *wave function phase* ϕ is defined only modulo 2π. Therefore, circulation is quantized [14], and these singular lines are the center lines of quantized vortices. The usual classical terms in Eq. (7) are augmented by the "quantum pressure" $p_Q \equiv -\frac{\hbar^2}{2m}(\Delta\sqrt{\rho})/\sqrt{\rho}$. The second equation (8) is the continuity equation for conservation of particle number, i.e., atom number in the superfluid gas. The dynamics of the weakly interacting, dilute ensemble of atoms is thus that of a perfect Euler fluid with quantized circulation of singular vortex lines. This is true except for regions in which the density rapidly varies and the quantum pressure term p_Q becomes relevant, which happens on scales of order the coherence length[3] $\xi = \hbar/\sqrt{gm\rho}$ where ρ is a constant (asymptotic) density far away from the density-depleted (or possibly density-enhanced) region. This is the case in the density-depleted cores of quantized vortices, or at the low-density boundaries of the system. The quantum pressure is negligible outside these domains of rapidly varying and/or low density. The whole armoury of space-time metric description of excitations, explained in the last section, and based on the Euler and continuity equations, is then valid for phonon excitations of a Bose-Einstein condensate, with the space-time metric (4), as long as we are outside the core of quantized vortices and far from the boundaries of the system, where both the flow is irrotational and the quantum pressure is negligible.

We mention here in passing that the form (2) of the wave equation is valid in quite general physical contexts. That is, a generally covariant curved space-time wave equation can be formulated not just for the velocity perturbation potential in an irrotational Euler fluid, for which we have introduced the effective metric concept. If the spectrum of excitations (in the local rest frame) is linear, $\omega = c_{\rm prop}k$, where $c_{\rm prop}$ is the propagation speed of *some* collective excitation, the statement that an effective space-time metric exists is true, provided we only consider wave perturbations of a single scalar field Φ constituting a fixed classical background. More precisely, given the generic requirement that the action density $\mathcal{L}$ is a functional of Φ and its space-time derivatives $\partial_\mu\Phi$, i.e. $\mathcal{L} = \mathcal{L}[\Phi, \partial_\mu\Phi]$, the fluctuations $\phi \equiv \delta\Phi$ around some classical background solution Φ_0 of the Euler-Lagrange equations always satisfy a wave equation of the form of Eq. (2), with a possible additional scalar po-

[3] Note that the coherence or "healing" length is also frequently defined in the literature, see e.g. [4], with an additional factor of $1/\sqrt{2}$, i.e., as $\xi_c = \hbar/\sqrt{2gm\rho}$.

tential term [15] comprising, for example, a mass term for the scalar field. As a consequence, the effective metric description also applies, *inter alia*, to the quasiparticle excitations around the gap nodes in the superfluid ^{3}He-A [16], photon propagation in dielectrics [17], and surface waves in shallow water [18].

The analogy between photons propagating on given (curved) space-time backgrounds and phonons in spatially and temporally inhomogeneous superfluids or, more generally, quantized quasiparticles with linear quasiparticle dispersion in some background, allows us to apply many tools and methods developed for quantum fields in curved space-times [19]. We can therefore conclude that the associated phenomena occur (provided the fundamental commutation relations of these quantum fields are fulfilled [20]). Among these phenomena are Hawking radiation [21], the Gibbons-Hawking effect in de Sitter spacetime [22], and cosmological particle production [23,24]. Furthermore, cosmic inflation and the freezing-in of quantum vacuum fluctuations by the horizon crossing of the corresponding modes may be simulated [25]. A comprehensive recent review of the subject of analogue gravity in its broadest sense is given in [26].

1.3 Pseudo-energy-momentum Tensor

An important quantity characterizing the dynamics of the field ϕ is the pseudo-energy-momentum tensor, cf. [27]. Since the equation of motion for the scalar mode, $\nabla_\mu \mathrm{g}^{\mu\nu}\nabla_\nu\phi = 0$, where ∇_μ denotes the space-time-covariant derivative, is equivalent to covariant energy-momentum balance, expressed by $\nabla_\mu T^{\mu\nu} = 0$, the classical pseudo-energy-momentum tensor reads [27]

$$T_{\mu\nu} = (\partial_\mu\phi)(\partial_\nu\phi) - \frac{1}{2}\,\mathrm{g}_{\mu\nu}(\partial_\rho\phi)(\partial_\sigma\phi)\mathrm{g}^{\rho\sigma}\,. \tag{9}$$

We already stressed that the identification of field theoretical effects in curved space-time by analogy (including the *existence* of the pseudo-energy-momentum tensor), is of a kinematical nature. An important question concerns the dynamics, that is, the *backreaction* of the quantum fluctuations of the scalar field onto the classical background. Extending the analogy to curved space-times a bit further, one is tempted to apply the effective-action method (see, e.g., [19] and [28]). In the effective-action method, one integrates out fluctuations of the quantum fields to one-loop order, and then determines the expectation value of the energy-momentum-tensor by the canonical identification $\delta\mathcal{A}_{\rm eff}/\delta\mathrm{g}^{\mu\nu} \equiv \frac{1}{2}\sqrt{-\mathrm{g}}\,\langle\hat{T}_{\mu\nu}\rangle$. Since the dependence of the effective action $\mathcal{A}_{\rm eff}$ on the degrees of freedom of the background η enters via the effective metric $\mathrm{g}^{\mu\nu} = \mathrm{g}^{\mu\nu}(\eta)$, one finds the backreaction contribution to the equations of motion of the η by differentiation of the effective action according to

$$\frac{\delta\mathcal{A}_{\rm eff}}{\delta\eta} = \frac{\delta\mathcal{A}_{\rm eff}}{\delta\mathrm{g}^{\mu\nu}}\,\frac{\delta\mathrm{g}^{\mu\nu}}{\delta\eta} = \frac{1}{2}\,\sqrt{-\mathrm{g}}\,\left\langle\hat{T}_{\mu\nu}\right\rangle\,\frac{\delta\mathrm{g}^{\mu\nu}}{\delta\eta}\,. \tag{10}$$

The precise meaning of the expectation value of the pseudo-energy-momentum tensor, $\langle \hat{T}_{\mu\nu} \rangle$, is difficult to grasp in general, due to the non-uniqueness of the vacuum state in a complicated curved space-time background and the ultraviolet (UV) renormalization procedure. Adopting a covariant renormalization scheme, the results for $\langle \hat{T}_{\mu\nu} \rangle$ can be classified in terms of geometrical quantities (cf. the trace anomaly [19]).

However, in calculating the quantum backreaction using the effective-action method, one is implicitly making two essential assumptions: first, that the leading contributions to the backreaction are completely determined by the effective action in Eq. (5), and, second, that deviations from the low-energy effective action at high energies do not affect the (renormalized) expectation value of the pseudo-energy-momentum tensor, $\langle \hat{T}_{\mu\nu} \rangle$. Since the effective covariance in Eq. (5) is only a low-energy property, the applicability of a covariant renormalization scheme is not obvious in general. In the following, we critically examine the question of whether the two assumptions mentioned above are justified, e.g., whether $\langle \hat{T}_{\mu\nu} \rangle$ completely determines the backreaction of the linearized quantum fluctuations.

The remainder of these lecture notes, which is based on the results of the publication [29], is organized as follows. In Sect. 2, we give a brief introduction to Bose-Einstein condensates, and introduce a particle-number-conserving ansatz for the field operator separated into condensate, single-particle and multi-particle excitation parts. In the subsequent Sect. 3, based on this ansatz, the backreaction onto the motion of the fluid using the full current will be calculated, and it is shown that this yields a different result than that obtained by the effective-action method. Afterwards, the failure of the effective-action technique is discussed in more detail in Sect. 4. The cutoff dependence of the pseudo-energy-momentum tensor (9) is addressed in Sect. 5. As a simple example, we consider the influence of the backreaction contribution on a static quasi-1D condensate in section 6.

2 Excitations in Bose-Einstein Condensates

2.1 Particle-number-conserving Mean-field Expansion

In the s-wave scattering approximation, a dilute many-particle system of interacting bosons is described, on a "microscopic" level corresponding to distance scales much larger than the true range of the interaction potential, by the field operator equation of motion in the Heisenberg picture (we set from now on $\hbar = m = 1$)

$$i\frac{\partial}{\partial t}\hat{\Psi} = \left(-\frac{1}{2}\boldsymbol{\nabla}^2 + V_{\text{trap}} + g\,\hat{\Psi}^\dagger\hat{\Psi}\right)\hat{\Psi}\,. \tag{11}$$

In the limit of many particles $N \gg 1$, in a finite trap at zero temperature with almost complete condensation, the full field operator $\hat{\Psi}$ can be represented in terms of the particle-number-conserving mean-field ansatz [30, 31]

$$\hat{\Psi} = \left(\psi_{\rm c} + \hat{\chi} + \hat{\zeta}\right) \hat{A}\hat{N}^{-1/2} \,. \tag{12}$$

Here, the order parameter $\psi_{\rm c} = \mathcal{O}(\sqrt{N})$ [note that $\psi_c \neq \langle\hat{\Psi}\rangle$, as opposed to the Ψ in Eq. (6)]. The one-particle excitations are denoted $\hat{\chi} = \mathcal{O}(N^0)$, where *one-particle* here means that the Fourier components of $\hat{\chi}$ are linear superpositions of annihilation and creation operators of quasiparticles $\hat{a}_{\boldsymbol{k}}$ and $\hat{a}^\dagger_{\boldsymbol{k}}$, cf. Eq. (38) below. The remaining higher-order, multi-particle corrections are described by $\hat{\zeta} = \mathcal{O}(1/\sqrt{N})$. The above mean-field ansatz can be derived in the dilute-gas limit by formally setting $g = \mathcal{O}(1/N)$ [31–34]; we shall use this formal definition of the dilute-gas limit in what follows. The dilute-gas limit should be compared and contrasted with the usual thermodynamic limit, in which the density and particle interaction remains constant, while the size of the (trapped) system increases with $N \to \infty$, adjusting the harmonic trapping potential $V_{\rm trap}$ correspondingly (in D spatial dimensions, the thermodynamic limit corresponds to keeping $N\omega^D$ constant for $N \to \infty$, where ω is the geometric mean of the trapping frequencies [4]). In the presently used dilute-gas limit, on the other hand, the trapping potential remains constant, but the interaction and the density change. The advantage of the limit gN constant is that in this limit we have a well-defined prescription to implement the mean-field approximation, keeping one power of g for each factor of N, cf. [36].

2.2 Gross-Pitaevskiĭ and Bogoliubov-de Gennes Equations

Insertion of Eq. (12) into (11) yields to $\mathcal{O}(N)$ the Gross-Pitaevskiĭ equation [35] for the order parameter $\psi_{\rm c}$

$$i\frac{\partial}{\partial t}\psi_{\rm c} = \left(-\frac{1}{2}\boldsymbol{\nabla}^2 + V_{\rm trap} + g|\psi_{\rm c}|^2 + 2g\left\langle\hat{\chi}^\dagger\hat{\chi}\right\rangle\right)\psi_{\rm c} + g\left\langle\hat{\chi}^2\right\rangle\psi_{\rm c}^* \,. \tag{13}$$

The Bogoliubov-de Gennes equations [37] for the one-particle fluctuations $\hat{\chi}$ are obtained to $\mathcal{O}(N^0)$

$$i\frac{\partial}{\partial t}\hat{\chi} = \left(-\frac{1}{2}\boldsymbol{\nabla}^2 + V_{\rm trap} + 2g|\psi_{\rm c}|^2\right)\hat{\chi} + g\psi_{\rm c}^2\hat{\chi}^\dagger \,, \tag{14}$$

Finally, the time evolution of the remaining higher-order corrections in the expansion (12), $\hat{\zeta} = \mathcal{O}(1/\sqrt{N})$, neglecting the $\mathcal{O}(1/N)$ terms, is given by:

$$\begin{aligned} i\frac{\partial}{\partial t}\hat{\zeta} \approx{}& \left(-\frac{1}{2}\boldsymbol{\nabla}^2 + V_{\rm trap} + 2g|\psi_{\rm c}|^2\right)\hat{\zeta} + g\psi_{\rm c}^2\,\hat{\zeta}^\dagger \\ &+2g(\hat{\chi}^\dagger\hat{\chi} - \langle\hat{\chi}^\dagger\hat{\chi}\rangle)\psi_{\rm c} + g(\hat{\chi}^2 - \langle\hat{\chi}^2\rangle)\psi_{\rm c}^* \,. \end{aligned} \tag{15}$$

The Gross-Pitaevskiĭ equation in the form (13) ought to be compared with the simple-minded form of (6). The additional terms $2g\langle\hat{\chi}^\dagger\hat{\chi}\rangle$ and $g\langle\hat{\chi}^2\rangle$ in the Gross-Pitaevskiĭ equation in the form of Eq. (13) above ensure that the expectation value of the multi-particle operator, $\hat{\zeta} = \mathcal{O}(1/\sqrt{N})$, vanishes in leading

order, $\langle\hat{\zeta}\rangle = \mathcal{O}(1/N)$. Without these additional terms, the mean-field expansion (12) would still be valid with $\hat{\zeta} = \mathcal{O}(1/\sqrt{N})$, but without $\langle\hat{\zeta}\rangle = \mathcal{O}(1/N)$. The proper incorporation of the so-called "anomalous" fluctuation average $\langle\hat{\chi}^2\rangle$ (the "normal" fluctuation average is $\langle\hat{\chi}^\dagger\hat{\chi}\rangle$) into the description of Bose-Einstein condensates has also been discussed from various points of view in [31, 38, 39].

3 Quantum Backreaction

3.1 Calculation of Backreaction Force From Microscopic Physics

The observation that the Gross-Pitaevskiĭ equation (13) yields an equation correct to leading order $\mathcal{O}(\sqrt{N})$, using either $|\psi_\mathrm{c}|^2$ or $|\psi_\mathrm{c}|^2 + 2\langle\hat{\chi}^\dagger\hat{\chi}\rangle$ in the first line of (13), hints at the fact that quantum backreaction effects correspond to next-to-leading, i.e. quadratic order terms in the fluctuations and cannot be derived *ab initio* in the above manner without additional assumptions. Therefore, we shall employ an alternative method: In terms of the exact density and current given by

$$\varrho = \left\langle\hat{\Psi}^\dagger\hat{\Psi}\right\rangle\,,\ \boldsymbol{j} = \frac{1}{2i}\left\langle\hat{\Psi}^\dagger\boldsymbol{\nabla}\hat{\Psi} - \mathrm{H.c.}\right\rangle\,, \tag{16}$$

the time-evolution is governed by the equation of continuity for ϱ and an Euler type equation for the current $\boldsymbol{j}$. After insertion of Eq. (11), we find that the equation of continuity is not modified by the quantum fluctuations but satisfied exactly (i.e., to all orders in $1/N$ or $\hbar$)

$$\frac{\partial}{\partial t}\varrho + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0\,, \tag{17}$$

in accordance with the $U(1)$ invariance of the Hamiltonian and the Noether theorem, cf. [28]. However, if we insert the mean-field expansion (12) and write the full density as a sum of condensed and non-condensed parts

$$\varrho = \varrho_\mathrm{c} + \langle\hat{\chi}^\dagger\hat{\chi}\rangle + \mathcal{O}(1/\sqrt{N})\,, \tag{18}$$

with $\varrho_\mathrm{c} = |\psi_\mathrm{c}|^2$, we find that neither part is conserved separately in general. Note that this split requires $\langle\hat{\zeta}\rangle = \mathcal{O}(1/N)$, i.e., the modifications to the Gross-Pitaevskiĭ equation (13) discussed above. Similarly, we may split up the full current [with $\varrho_\mathrm{c}\boldsymbol{v}_\mathrm{c} = \Im(\psi_\mathrm{c}^*\boldsymbol{\nabla}\psi_\mathrm{c})$]

$$\boldsymbol{j} = \varrho_\mathrm{c}\boldsymbol{v}_\mathrm{c} + \frac{1}{2i}\left\langle\hat{\chi}^\dagger\boldsymbol{\nabla}\hat{\chi} - \mathrm{H.c.}\right\rangle + \mathcal{O}(1/\sqrt{N})\,, \tag{19}$$

and introduce an average velocity $\boldsymbol{v}$ via $\boldsymbol{j} = \varrho\boldsymbol{v}$. This enables us to *unambiguously* define the quantum backreaction $\boldsymbol{Q}$ as the following additional contribution in an equation of motion for $\boldsymbol{j}$ analogous to the Euler equation:

$$\frac{\partial}{\partial t}\boldsymbol{j} = \boldsymbol{f}_{\rm cl}(\boldsymbol{j},\varrho) + \boldsymbol{Q} + \mathcal{O}(1/\sqrt{N})\,, \tag{20}$$

where the classical force density term

$$\boldsymbol{f}_{\rm cl}(\boldsymbol{j},\varrho) = -\boldsymbol{v}\left[\boldsymbol{\nabla}\cdot(\varrho\boldsymbol{v})\right] - \varrho(\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{v} + \varrho\boldsymbol{\nabla}\left(\frac{1}{2}\frac{\boldsymbol{\nabla}^2\sqrt{\varrho}}{\sqrt{\varrho}} - V_{\rm trap} - g\varrho\right)\,. \tag{21}$$

Here, "classical" means that the force density contains no explicit quantum fluctuation terms (i.e., only those absorbed in the full density and full current), and in addition just the "quantum pressure", which already occurs on the mean-field level. Formulation in terms of the conventional Euler equation, i.e., using a convective derivative of the velocity defined by $\boldsymbol{j} = \rho\boldsymbol{v}$ giving the acceleration, yields

$$\left(\frac{\partial}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\right)\boldsymbol{v} = -\boldsymbol{\nabla}\left(V_{\rm trap} + g\varrho - \frac{1}{2}\frac{\boldsymbol{\nabla}^2\sqrt{\varrho}}{\sqrt{\varrho}}\right) + \frac{\boldsymbol{Q}}{\varrho} + \mathcal{O}(N^{-3/2})\,. \tag{22}$$

The quantum backreaction force density $\boldsymbol{Q}$ can now be calculated by comparing the two equations above and expressing $\partial\boldsymbol{j}/\partial t$ in terms of the field operators via Eqs. (11) and (16)

$$\begin{aligned}\frac{\partial}{\partial t}\boldsymbol{j} = {}& \frac{1}{4}\left\langle\hat\Psi^\dagger\boldsymbol{\nabla}^3\hat\Psi - (\boldsymbol{\nabla}^2\hat\Psi^\dagger)\boldsymbol{\nabla}\hat\Psi + \text{H.c.}\right\rangle \\ & - \left\langle\hat\Psi^\dagger\hat\Psi\right\rangle\boldsymbol{\nabla}V_{\rm trap} - \frac{1}{2g}\boldsymbol{\nabla}\left\langle g^2(\hat\Psi^\dagger)^2\hat\Psi^2\right\rangle\,.\end{aligned} \tag{23}$$

After insertion of the mean-field expansion (12) , we obtain the leading contributions in the Thomas-Fermi limit

$$\begin{aligned}\boldsymbol{Q} = {}& \boldsymbol{\nabla}\cdot\left(\boldsymbol{v}\otimes\boldsymbol{j}_\chi + \boldsymbol{j}_\chi\otimes\boldsymbol{v} - \varrho_\chi\boldsymbol{v}\otimes\boldsymbol{v}\right) \\ & -\frac{1}{2g}\boldsymbol{\nabla}\left(g^2\left\langle 2|\psi_{\rm c}|^2\hat\chi^\dagger\hat\chi + \psi_{\rm c}^2(\hat\chi^\dagger)^2 + (\psi_{\rm c}^*)^2\hat\chi^2\right\rangle\right) \\ & -\frac{1}{2}\boldsymbol{\nabla}\cdot\left\langle(\boldsymbol{\nabla}\hat\chi^\dagger)\otimes\boldsymbol{\nabla}\hat\chi + \text{H.c.}\right\rangle\,,\end{aligned} \tag{24}$$

with $\varrho_\chi = \langle\hat\chi^\dagger\hat\chi\rangle$ and $\boldsymbol{j}_\chi = \Im\langle\hat\chi^\dagger\boldsymbol{\nabla}\hat\chi\rangle$. Under the assumption that the relevant length scales λ for variations of, e.g., ϱ and g, are much larger than the healing length $\xi = (g\varrho)^{-1/2}$, we have neglected terms containing quantum pressure contributions $\boldsymbol{\nabla}^2\varrho$ and $[\boldsymbol{\nabla}\varrho]^2$, which amounts to the Thomas-Fermi or local-density approximation. These contributions would, in particular, spoil the effective (local) geometry in Eq. (5) (the inclusion of the quantum pressure to derive a "nonlocal metric" has been discussed in [40]).

3.2 Comparison with Effective-action Technique

To compare the expression for the backreaction force density derived from the full dynamics of $\hat\Psi$ (24) with the force obtained from Eq. (10), we have

to identify the scalar field ϕ and the contravariant metric $\mathsf{g}^{\mu\nu}$. We already know that phonon modes with wavelength $\lambda \gg \xi$ are described by the action in Eq. (5) in terms of the phase fluctuations ϕ provided that $\mathsf{g}^{\mu\nu}$ is given by (3), where $A_c = c_s/g$ and $\sqrt{-\mathsf{g}} = c_s^3/g^2 = \rho^2/c_s$ in three spatial dimensions. The density fluctuations $\delta\varrho$ are related to the phase fluctuations ϕ via $\delta\varrho = -g^{-1}(\partial/\partial t + \boldsymbol{v}\cdot\boldsymbol{\nabla})\phi$.

The variables $\eta = \{\rho_\mathrm{b}, \phi_\mathrm{b}\}$ or alternatively $\eta = \{\rho_\mathrm{b}, \nabla\phi_\mathrm{b}\} = \{\rho_\mathrm{b}, \boldsymbol{v}_\mathrm{b}\}$ in (10) are then defined by the expectation values of density and phase operators according to

$$\begin{aligned}\hat{\Phi} &= \langle\hat{\Phi}\rangle + \hat{\phi} = \phi_\mathrm{b} + \hat{\phi}\,,\\ \hat{\varrho} &= \langle\hat{\varrho}\rangle + \delta\hat{\varrho} = \varrho_\mathrm{b} + \delta\hat{\varrho}\,.\end{aligned} \tag{25}$$

The phase operator can formally be introduced via the following ansatz for the full field operator

$$\hat{\Psi} = e^{i\hat{\Phi}}\sqrt{\hat{\varrho}}\,. \tag{26}$$

Since $\hat{\Phi}$ and $\hat{\varrho}$ do not commute, other forms such as $\hat{\Psi} = \sqrt{\hat{\varrho}}\,e^{i\hat{\Phi}}$ would not generate a self-adjoint $\hat{\Phi}$ (and simultaneously satisfy $\hat{\Psi}^\dagger\hat{\Psi} = \hat{\varrho}$). Note that, in contrast to the full density which is a well-defined and measurable quantity, the velocity potential, $\hat{\Phi}$, is not [41]. This can be seen as follows. The commutator between density and phase operators

$$[\hat{\varrho}(\boldsymbol{r}), \hat{\Phi}(\boldsymbol{r}')] = i\delta(\boldsymbol{r}-\boldsymbol{r}')\,, \tag{27}$$

yields, if one takes, on both sides, matrix elements in the number basis of the space integral over $\boldsymbol{r}$ for a given volume $\mathcal{V}$ in which $\boldsymbol{r}'$ is situated,

$$(N-N')\langle N|\hat{\Phi}(\boldsymbol{r}')|N'\rangle = i\delta_{NN'}\,, \tag{28}$$

where N and N' are two possible values for the number of particles in $\mathcal{V}$. This is an inconsistent relation for N, N' positive semidefinite and discrete (which the very existence of particles requires), most obviously for $N = N'$. The commutator therefore makes sense only if it is understood to be effectively coarse-grained over a sub-volume $\mathcal{V}$ with large enough number of particles $N \gg 1$, such that the inconsistency inherent in (28) becomes asymptotically irrelevant. It cannot be defined consistently locally, i.e., in arbitrarily small volumina, where there is just one particle or even none, or when the number fluctuations in larger volumina $\mathcal{V}$ are large, such that the probability to have a very small number of particles in $\mathcal{V}$ is not negligible [42].

The action in terms of the total density ϱ and the variable Φ reads (neglecting the quantum pressure term, i.e., in the Thomas-Fermi limit)

$$\mathcal{L} = -\varrho\left(\frac{\partial}{\partial t}\Phi + \frac{1}{2}(\boldsymbol{\nabla}\Phi)^2\right) - \epsilon[\varrho] - V_\mathrm{trap}\varrho\,, \tag{29}$$

with $\epsilon[\varrho]$ denoting the internal energy density. The quantum corrections to the Bernoulli equation up to second order in the fluctuations, using the effective-action method in Eq. (10) are then incorporated by writing (the background density here equals the full density, $\varrho = \varrho_{\rm b}$)

$$\frac{\partial}{\partial t}\phi_{\rm b} + \frac{1}{2}(\nabla\phi_{\rm b})^2 + h[\varrho] - \frac{\delta\mathcal{A}_{\rm eff}}{\delta\varrho} = 0\,, \tag{30}$$

where $h[\varrho] = d\epsilon/d\varrho + V_{\rm trap}$. We obtain, using (5),

$$\frac{\delta\mathcal{A}_{\rm eff}}{\delta\varrho} = \frac{\delta\mathcal{A}_{\rm eff}}{\delta \mathrm{g}^{\mu\nu}}\frac{\delta \mathrm{g}^{\mu\nu}}{\delta\varrho} = \frac{\sqrt{-\mathrm{g}}}{2}\langle\hat{T}_{\mu\nu}\rangle\frac{\delta \mathrm{g}^{\mu\nu}}{\delta\varrho} = -\frac{1}{2}\langle(\boldsymbol{\nabla}\hat{\phi})^2\rangle\,. \tag{31}$$

Clearly, taking the gradient of this result we obtain a backreaction force which markedly differs from the expression (24) derived in the previous subsection. Moreover, it turns out that the backreaction force density in Eq. (24) contains contributions which are not part of the expectation value of the pseudo-energy-momentum tensor, $\langle\hat{T}_{\mu\nu}\rangle$. For example, the phonon density ϱ_χ contains $\langle(\delta\hat{\varrho})^2\rangle_{\rm ren}$ (where $\langle\ldots\rangle_{\rm ren}$ means that the divergent c-number $\hat{\chi}\hat{\chi}^\dagger - \hat{\chi}^\dagger\hat{\chi} = \delta(0)$ has been subtracted already) which is part of $\langle\hat{T}_{\mu\nu}\rangle_{\rm ren}$, but also $\langle\hat{\phi}^2\rangle_{\rm ren}$ which is not. [Note that $\langle\hat{\phi}^2\rangle_{\rm ren}$ cannot be cancelled by the other contributions.] The expression $\langle(\boldsymbol{\nabla}\hat{\chi}^\dagger)\otimes\boldsymbol{\nabla}\hat{\chi} + \text{H.c.}\rangle$ in the last line of Eq. (24) contains $\langle\boldsymbol{\nabla}\hat{\phi}\otimes\boldsymbol{\nabla}\hat{\phi}\rangle_{\rm ren}$ which does occur in $\langle\hat{T}_{\mu\nu}\rangle_{\rm ren}$, but also $\langle\boldsymbol{\nabla}\delta\hat{\varrho}\otimes\boldsymbol{\nabla}\delta\hat{\varrho}\rangle_{\rm ren}$, which does not. One could argue that the latter term ought to be neglected in the Thomas-Fermi or local-density approximation since it is on the same footing as the quantum pressure contributions containing $\boldsymbol{\nabla}^2\varrho$ and $[\boldsymbol{\nabla}\varrho]^2$ [which have been neglected in (24)], but it turns out that this expectation value yields cutoff dependent contributions of the same order of magnitude as the other terms, see Sect. 5 below.

4 Failure of Effective-action Technique

After having demonstrated the failure of the effective-action method for deducing the quantum backreaction, let us study the reasons for this failure in more detail. The full action governing the dynamics of the fundamental field Ψ reads

$$\mathcal{L}^\Psi = i\Psi^*\frac{\partial}{\partial t}\Psi - \Psi^*\left(-\frac{1}{2}\,\boldsymbol{\nabla}^2 + V_{\rm trap} + \frac{g}{2}\,\Psi^*\Psi\right)\Psi\,. \tag{32}$$

Linearization according to $\Psi = \psi_{\rm c} + \chi$ yields the effective second-order action generating the Bogoliubov-de Gennes equations (14)

$$\mathcal{L}^\chi_{\rm eff} = i\chi^*\frac{\partial}{\partial t}\chi - \chi^*\left(-\frac{1}{2}\boldsymbol{\nabla}^2 + V_{\rm trap} + 2g|\psi_{\rm c}^2|\right)\chi - \frac{1}{2}\left[g(\psi_{\rm c}^*)^2\chi^2 + \text{H.c.}\right]\,. \tag{33}$$

If we start with the action (29) in terms of the total density ϱ and the nonfundamental variable Φ, the quantum corrections to the equation of continuity $\delta\mathcal{A}_{\rm eff}/\delta\phi_{\rm b}$ are reproduced correctly but the derived quantum backreaction contribution to the Bernoulli equation, $\delta\mathcal{A}_{\rm eff}/\delta\varrho_{\rm b}$, and therefore the backreaction force, is wrong.

One now is led to the question why the effective-action method works for the fundamental field Ψ and gives the correct expression for the backreaction force, but fails for the non-fundamental variable Φ. The quantized fundamental field $\hat{\Psi}$ satisfies the equation of motion (11) as derived from the above action and possesses a well-defined linearization via the mean-field expansion (12). One of the main assumptions of the effective-action method is a similar procedure for the variable Φ, i.e., the existence of a well-defined and linearizable full quantum operator $\hat{\Phi}$ satisfying the quantum Bernoulli equation (for large length scales), cf. Eq. (30)

$$\frac{\partial}{\partial t}\hat{\Phi} + \frac{1}{2}(\boldsymbol{\nabla}\hat{\Phi})^2 + h[\hat{\varrho}] \stackrel{?}{=} 0\,. \tag{34}$$

The problem is that the commutator of $\hat{\varrho}$ and $\hat{\Phi}$ at the same position diverges, cf. Eq. (27), and hence the quantum Madelung ansatz in Eq. (26) is singular. As a result, the above quantum Bernoulli equation is not well-defined (in contrast to the equation of continuity), i.e., insertion of the quantum Madelung ansatz in Eq. (26) into Eq. (11) generates divergences [41].

In order to study these divergences by means of a simple example, let us consider a generalized Bose-Hubbard Hamiltonian [43], which considers bosons sitting on a lattice with sites i, which can hop between nearest neighbor sites and interact if at the same site:

$$\hat{H} = -\frac{\alpha}{2}\sum_{\langle ij\rangle}(\hat{\Psi}_i^\dagger\hat{\Psi}_j + \text{H.c.}) + \sum_i\left(\beta_i\hat{n}_i + \frac{\gamma}{2}\hat{n}_i^2\right)\,, \tag{35}$$

where $\hat{\Psi}_i$ is the annihilation operator for bosons at a given lattice site i, and $\langle ij\rangle$ denotes summation over nearest neighbors; $\hat{n}_i = \hat{\Psi}_i^\dagger\hat{\Psi}_i$ is the so-called filling factor (operator), equal to the number of bosons at the lattice site i. The quantities β_i multiplying the filling factor depend on the site index. In the continuum limit, the lattice Hamiltonian (35) generates a version of Eq. (11). Setting $a^{D/2}\hat{\Psi}(\boldsymbol{x}_i) = \hat{\Psi}_i$, where the $\hat{\Psi}(\boldsymbol{x}_i)$ are the continuum field operators and a is the lattice spacing taken to zero (we consider for simplicity a simple cubic lattice in D spatial dimensions), the *effective mass* is given by $1/m^* = \alpha a^2$: The bosons moving through the lattice obviously become the heavier the smaller the hopping amplitude α becomes at given a. The coupling constant is determined by $g = \gamma a^D$, and the trap potential is governed by $V_{\rm trap}(\boldsymbol{x}) = g/2 + \beta(\boldsymbol{x}) - \alpha$.

On the other hand, inserting the quantum Madelung ansatz employing a phase operator, Eq. (26) in its lattice version, the problem of operator ordering arises and the (for the Bernoulli equation) relevant kinetic energy term reads

$$\hat{H}_{\Phi} = \frac{1}{4} \sum_i \sqrt{\hat{n}_i(\hat{n}_i+1)}\, (\boldsymbol{\nabla}\hat{\Phi})_i^2 + \text{H.c.}\,, \tag{36}$$

with the replacement $\hat{n}_i + 1$ instead of $\hat{n}_i$ being one effect of the non-commutativity. In the superfluid phase with large filling $n \gg 1$, we therefore obtain the following leading correction to the equation of motion

$$\frac{\partial}{\partial t}\hat{\Phi} + \frac{1}{2}(\boldsymbol{\nabla}\hat{\Phi})^2 + h[\hat{\varrho}] + \frac{1}{\hat{n}}\,\frac{(\boldsymbol{\nabla}\hat{\Phi})^2}{16}\,\frac{1}{\hat{n}} = \mathcal{O}\left(\frac{1}{n^3}\right). \tag{37}$$

The Bernoulli equation in its quantum version thus receives corrections depending on microscopic details like the filling of a particular site and is not of the (conjectured) form (34).

By means of this simple example, we already see that the various limiting procedures such as the quantization and subsequent mean-field expansion, the variable transformation $(\Psi^*, \Psi) \leftrightarrow (\varrho, \Phi)$, and the linearization for small fluctuations, as well as continuum limit do not commute in general – which explains the failure of the effective-action method for deducing the quantum backreaction. The variable transformation $(\Psi^*, \Psi) \leftrightarrow (\varrho, \Phi)$ is applicable to the zeroth-order equations of motion for the classical background as well as to the first-order dynamics of the linearized fluctuations – but the quantum backreaction is a second-order effect, where the aforementioned difficulties, such as the question of the choice of fundamental variables and their operator ordering, arise.

5 Cutoff Dependence of Effective Action

Another critical issue for the applicability of the effective-action method is the UV divergence of $\langle \hat{T}_{\mu\nu} \rangle$. Extrapolating the low-energy effective action in Eq. (5) to large momenta k, the expectation values $\langle \delta\hat{\varrho}^2 \rangle$ and $\langle \hat{\phi}^2 \rangle$ entering ϱ_χ would diverge. For Bose-Einstein condensates, we may infer the deviations from Eq. (5) at large k from the Bogoliubov-de Gennes equations (14). Assuming a static and homogeneous background (which should be a good approximation at large k), a normal-mode expansion yields a Bogoliubov transformation between the bare bosonic operators $\hat{\chi}_{\boldsymbol{k}}$ and the quasiparticle operators $\hat{a}_{\boldsymbol{k}}$, $\hat{a}^\dagger_{\boldsymbol{k}}$:

$$\hat{\chi}_{\boldsymbol{k}} = \sqrt{\frac{\boldsymbol{k}^2}{2\omega_{\boldsymbol{k}}}} \left[\left(\frac{\omega_{\boldsymbol{k}}}{\boldsymbol{k}^2} - \frac{1}{2} \right) \hat{a}^\dagger_{\boldsymbol{k}} + \left(\frac{\omega_{\boldsymbol{k}}}{\boldsymbol{k}^2} + \frac{1}{2} \right) \hat{a}_{\boldsymbol{k}} \right], \tag{38}$$

where the frequency $\omega_{\boldsymbol{k}}$ is determined by the Bogoliubov dispersion relation for the dilute Bose gas, $\omega^2_{\boldsymbol{k}} = g\varrho\,\boldsymbol{k}^2 + \boldsymbol{k}^4/4$. The above form of the Bogoliubov transformation results, after inversion, in the usual phonon quasiparticle operators at low momenta, and gives $\hat{\chi}_{\boldsymbol{k}} = \hat{a}_{\boldsymbol{k}}$ at $\boldsymbol{k} \to \infty$, i.e., the quasiparticles and the bare bosons become, as required, identical at large momenta.

Using a linear dispersion $\omega_{\boldsymbol{k}}^2 \propto \boldsymbol{k}^2$ instead of the full Bogoliubov dispersion, expectation values such as $\langle\hat{\chi}^\dagger\hat{\chi}\rangle$ would be UV divergent, but the correct dispersion relation implies $\hat{\chi}_{\boldsymbol{k}} \sim \hat{a}_{\boldsymbol{k}}^\dagger\, g\varrho/\boldsymbol{k}^2 + \hat{a}_{\boldsymbol{k}}$ for large $\boldsymbol{k}^2$, and hence $\langle\hat{\chi}^\dagger\hat{\chi}\rangle$ is UV finite in three and lower spatial dimensions. Thus the healing length ξ acts as an effective UV cutoff, $k_\xi^{\mathrm{cut}} \equiv 1/\xi$.

Unfortunately, the quadratic decrease for large k in Eq. (38), giving asymptotically $\hat{\chi}_{\boldsymbol{k}} \sim \hat{a}_{\boldsymbol{k}}^\dagger\, g\varrho/\boldsymbol{k}^2 + \hat{a}_{\boldsymbol{k}}$, is not sufficient for rendering the other expectation values (i.e., apart from ϱ_χ and $\boldsymbol{j}_\chi$) in Eq. (24) UV finite in three spatial dimensions. This UV divergence indicates a failure of the s-wave pseudopotential $g\delta^3(\boldsymbol{r}-\boldsymbol{r'})$ in Eq. (11) at large wavenumbers k and can be eliminated by replacing $g\delta^3(\boldsymbol{r}-\boldsymbol{r'})$ by a more appropriate two-particle interaction potential $V_{\mathrm{int}}(\boldsymbol{r}-\boldsymbol{r'})$, see [36]. Introducing another UV cutoff wavenumber k_s^{cut} related to the breakdown of the s-wave pseudo-potential, we obtain $\langle(\boldsymbol{\nabla}\hat{\chi}^\dagger)\otimes\boldsymbol{\nabla}\hat{\chi} + \mathrm{H.c.}\rangle \sim g^2\varrho^2\, k_s^{\mathrm{cut}}$ and $\langle\hat{\chi}^2\rangle \sim g\varrho\, k_s^{\mathrm{cut}}$.

In summary, there are two different cutoff wavenumbers: The first one, k_ξ^{cut}, is associated to the breakdown of the effective Lorentz invariance (change of dispersion relation from linear to quadratic) and renders some – but not all – of the naively divergent expectation values finite. The second wavenumber, k_s^{cut}, describes the cutoff for all (remaining) UV divergences. In dilute Bose-Einstein condensates, these two scales are vastly different by definition. Because the system is dilute, the inverse range of the true potential $k_s^{\mathrm{cut}} \equiv 1/r_0$, must be much larger than the inverse healing length. Thus the following condition of scale separation must hold:

$$k_{\mathrm{UV}}^{\mathrm{cut}} = k_s^{\mathrm{cut}} \gg k_\xi^{\mathrm{cut}} = k_{\mathrm{Lorentz}}^{\mathrm{cut}}\,. \tag{39}$$

In terms of the length scales governing the system, using that $k_\xi^{\mathrm{cut}} = \sqrt{4\pi a_s \rho}$, in the dilute gas we must have the following condition fulfilled, $4\pi a_s r_0^2 \ll d^3$, where $d = \varrho^{-1/3}$ is the interparticle separation. Note that the opposite scale separation relation, $k_{\mathrm{Lorentz}}^{\mathrm{cut}} \gg k_{\mathrm{UV}}^{\mathrm{cut}}$ is very unnatural in the sense that every quantum field theory which has the usual properties such as locality and Lorentz invariance must have UV divergences (e.g., in the two-point function).

The renormalization of the cutoff-dependent terms is different for the two cases: The k_s^{cut}-contributions can be absorbed by a ϱ-independent renormalization of the coupling g [34, 36], whereas the k_ξ^{cut}-contributions depend on the density in a nontrivial way and thus lead to a quantum renormalization of the effective equation of state. We supply an example for this renormalization in the section to follow.

6 Static Example for the Backreaction Force

In order to provide an explicit example for the quantum backreaction term in Eq. (24), without facing the above discussed UV problem, let us consider a quasi-one-dimensional (quasi-1D) condensate [44, 45], where all the involved

quantities are UV finite. In a quasi-1D condensate the perpendicular harmonic trapping $\omega_\perp$ is much larger than the axial trapping ω_z such that the condensate assumes the shape of a strongly elongated cigar with all atomic motion in the perpendicular direction frozen out, $\omega_\perp$ being much larger than the mean energy per particle.

In accordance with general considerations [46], the phonon density ϱ_χ is infrared (IR) divergent in one spatial dimension, therefore inducing finite-size effects, i.e., a dependence of the various quantities of interest on the system size. Nevertheless, in certain situations, we are able to derive a closed local expression for the quantum backreaction term Q: Let us assume a completely static condensate $\boldsymbol{v} = 0$ in effectively one spatial dimension, still allowing for a spatially varying density ϱ and possibly also coupling g. Furthermore, since we require that spatial variations of ϱ and g occur on length scales λ much larger than the healing length (Thomas-Fermi approximation), we keep only the leading terms in $\xi/\lambda \ll 1$, i.e., the variations of ϱ and g will be neglected in the calculation of the expectation values. In this case, the quantum backreaction term Q simplifies considerably and yields (in effectively one spatial dimension, where $g \equiv g_{1\mathrm{D}}$ and $\varrho \equiv \varrho_{1\mathrm{D}}$ now both refer to the 1D quantities)

$$\begin{aligned} Q &= -\nabla \left\langle (\nabla\hat{\chi}^\dagger)\nabla\hat{\chi} \right\rangle - \frac{1}{2g}\nabla \left(g^2 \varrho \left\langle 2\hat{\chi}^\dagger\hat{\chi} + (\hat{\chi}^\dagger)^2 + \hat{\chi}^2 \right\rangle \right) \\ &= -\nabla \left(\frac{1}{3\pi} (g\varrho)^{3/2} \right) + \frac{1}{2\pi g}\nabla \left(g^{5/2}\varrho^{3/2} \right) + \mathcal{O}(\xi^2/\lambda^2) \\ &= \frac{\varrho}{2\pi}\nabla\sqrt{g^3\varrho} + \mathcal{O}(\xi^2/\lambda^2)\,. \end{aligned} \tag{40}$$

It turns out that the IR divergences of $2\langle\hat{\chi}^\dagger\hat{\chi}\rangle$ and $\langle(\hat{\chi}^\dagger)^2 + \hat{\chi}^2\rangle$ cancel each other such that the resulting expression is not only UV but also IR finite. Note that the sign of Q is positive and hence opposite to the contribution of the pure phonon density $\langle\hat{\chi}^\dagger\hat{\chi}\rangle$, which again illustrates the importance of the "anomalous" term $\langle(\hat{\chi}^\dagger)^2 + \hat{\chi}^2\rangle$.

A possible experimental signature of the quantum backreaction term Q calculated above, is the change incurred on the static Thomas-Fermi solution of the Euler equation (22) for the density distribution (cf. [35, 36])

$$\varrho_{1\mathrm{D}} = \frac{\mu - V_{\mathrm{trap}}}{g_{1\mathrm{D}}} + \frac{\sqrt{\mu - V_{\mathrm{trap}}}}{2\pi} + \mathcal{O}(1/\sqrt{N})\,, \tag{41}$$

with μ denoting the (constant) chemical potential. The classical $[\mathcal{O}(N)]$ density profile $\varrho_{\mathrm{cl}} = (\mu - V_{\mathrm{trap}})/g_{1\mathrm{D}}$ acquires nontrivial quantum $[\mathcal{O}(N^0)]$ corrections $\varrho_Q = \sqrt{\mu - V_{\mathrm{trap}}}/2\pi$, where the small parameter is the ratio of the interparticle distance $1/\varrho = \mathcal{O}(1/N)$ over the healing length $\xi = \mathcal{O}(N^0)$. Note that the quantum backreaction term ϱ_Q in the above split $\varrho = \varrho_{\mathrm{cl}} + \varrho_Q$ should neither be confused with the phonon density ϱ_χ in $\varrho = \varrho_{\mathrm{c}} + \varrho_\chi$ (remember that ϱ_χ is IR divergent and hence contains finite-size effects) nor with the quantum pressure contribution $\propto \boldsymbol{\nabla}^2\sqrt{\varrho}$ in the Euler type Eq. (22).

Evaluating the change ΔR of the Thomas-Fermi size (half the full length), where $\mu = V_{\rm trap}$, of a quasi-1D Bose-Einstein condensate induced by backreaction, from Eq. (41) we get $\Delta R = -2^{-5/2}(\omega_\perp/\omega_z)a_s$. Here, the quasi-1D coupling constant $g_{\rm 1D}$ is related to the 3D s-wave scattering length a_s and the perpendicular harmonic trapping $\omega_\perp$ by $g_{\rm 1D} = 2a_s\omega_\perp$ (provided $a_s \ll a_\perp = 1/\sqrt{\omega_\perp}$ [44]). In units of the classical size $R_{\rm cl} = (3a_sN\omega_\perp/\omega_z^2)^{1/3}$, we have

$$\frac{\Delta R}{R_{\rm cl}} = -\frac{1}{4\sqrt{2}}\left(\frac{1}{3N}\right)^{1/3}\left(\frac{\omega_\perp}{\omega_z}\frac{a_s}{a_z}\right)^{2/3}, \tag{42}$$

where $a_z = 1/\sqrt{\omega_z}$ describes the longitudinal harmonic trapping length. In quasi-1D condensates, backreaction thus leads to a *shrinking* of the cloud relative to the classical expectation – whereas in three spatial dimensions we have the opposite effect [4,36]. In one dimension, we thus obtain a *softening* of the quantum renormalized equation of state of the gas; conversely, in three spatial dimensions the effective equation of state becomes *stiffer* due to quantum fluctuations.

For reasonably realistic experimental parameters, the effect of quantum backreaction on the equation of state should be measurable; for $N \simeq 10^3$, $\omega_\perp/\omega_z \simeq 10^3$, and $a_s/a_z \simeq 10^{-3}$, we obtain $|\Delta R/R_{\rm cl}| \simeq 1\%$.

7 Conclusion

By explicit analysis of the analytically tractable case of a dilute Bose gas in the mean-field approximation, we have demonstrated the following. Even given that the explicit form of the quantum backreaction terms depends on the definition of the classical background, the effective-action method does not yield the correct result in the general case (which is, in particular, independent of the choice of variables). The knowledge of the classical (macroscopic) equation of motion – such as the Bernoulli equation – may be sufficient for deriving the first-order dynamics of the linearized quantum fluctuations (phonons), but the quantum backreaction as a second-order effect cannot be obtained without further knowledge of the microscopic structure (which reflects itself, for example, in the operator ordering imposed). It is tempting to compare these findings to gravity, where we also know the classical equations of motion only

$$R_{\mu\nu} - \frac{1}{2}\,\mathsf{g}_{\mu\nu}\,R = \frac{8\pi G}{c^4}\,T_{\mu\nu}\,, \tag{43}$$

which – in analogy to the Bernoulli equation – might yield the correct first-order equations of motion for the linearized gravitons, but perhaps not their (second-order) quantum backreaction. Another potentially interesting point of comparison is the existence of two different high-energy scales – one associated to the breakdown of Lorentz invariance $k_\xi^{\rm cut} = k_{\rm Lorentz}^{\rm cut}$ and the other one, $k_{\rm UV}^{\rm cut} = k_s^{\rm cut}$, to the UV cutoff introduced by the true interaction potential

range. The question then poses itself whether one of the two cutoff scales (or in some sense both of them) correspond to the Planck scale in gravity.

The dominant $\mathcal{O}(\xi/\lambda)$ quantum backreaction contributions like those in Eq. (41) depend on the healing length as the lower UV cutoff and hence cannot be derived from the low-energy effective action in Eq. (5) using a covariant (i.e., cutoff independent) regularization scheme, which does not take into account details of microscopic physics (represented, for example, in the quasiparticle dispersion relation). Note that the leading $\mathcal{O}(\xi/\lambda)$ quantum correction to the pressure could be identified with a cosmological term, $\langle \hat{T}_{\mu\nu} \rangle = \Lambda \, \mathsf{g}_{\mu\nu}$ in Eq. (10), provided that the cosmological "constant" Λ is not constant but depends on g and ϱ. Note that in general relativity, the Einstein equations demand Λ to be constant, due to the equivalence principle and the resulting requirement that the metric be parallel-transported, $\nabla^\mu \mathsf{g}_{\mu\nu} = 0$.

As became evident, the knowledge of the expectation value of the pseudo-energy-momentum tensor $\langle \hat{T}_{\mu\nu} \rangle$ is not sufficient for determining the quantum backreaction effects in general. Even though $\langle \hat{T}_{\mu\nu} \rangle$ is a useful concept for describing the phonon kinematics (at low energies), we have seen that it does not represent the full dynamics of the fluid dynamical variables defined in terms of the fundamental quantum field $\hat{\Psi}$. Related limitations of the *classical* pseudo-energy-momentum tensor, in particular the background choice dependence of the description of the second-order effect of the exchange of energy and momentum between excitations and that background, and the resulting form of the conservation laws, have been discussed in [27].

In general, the quantum backreaction corrections to the Euler equation in Eq. (22) cannot be represented as the gradient of some local potential, cf. Eq. (24). Hence they may effectively generate vorticity and might serve as the seeds for vortex nucleation from the vortex vacuum.

In contrast to the three-dimensional case (see, e.g., [35, 36]), the quantum backreaction corrections given by Eq. (40) *diminish* the pressure in condensates that can be described by Eq. (11) in one spatial dimension (quasi-1D case). This is a direct consequence of the so-called "anomalous" term $\langle (\hat{\chi}^\dagger)^2 + \hat{\chi}^2 \rangle$ in Eq. (40), which – together with the cancellation of the IR divergence – clearly demonstrates that it cannot be neglected in general. We emphasize that even though Eqs. (40)–(42) describe the *static* quantum backreaction corrections to the ground state, which can be calculated by an alternative method [36] as well, the expression in Eq. (24) is valid for more general dynamical situations, such as rapidly expanding condensates. Quantum backreaction can thus generally not be incorporated by rewriting the Euler equation in terms of a renormalized chemical potential. While the static quantum backreaction corrections to the ground state can be absorbed by a redefinition of the chemical potential $\mu(\varrho)$ determining a quantum renormalized (barotropic) equation of state $p(\varrho)$, this is not possible for the other terms in Eq. (24), like the quantum friction-type terms depending on $\boldsymbol{j} \otimes \boldsymbol{v}$.

We have derived, from the microscopic physics of dilute Bose-Einstein condensates, the backreaction of quantum fluctuations onto the motion of the

full fluid and found a quantum backreaction force that is potentially experimentally observable in existing condensates. We observed a failure of the effective-action technique to fully describe the backreaction force in Eq. (24), and a cutoff dependence of backreaction due to the breakdown of covariance at high energies. Whether similar problems, in particular the question of the correct choice of the fundamental variables and the related operator ordering issues, beset the formulation of a theory of "real" (quantum) gravity remains an interesting open question.

References

1. C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, 1973)
2. J. R. Anglin and W. Ketterle, "Bose-Einstein condensation of atomic gases", Nature **416**, 211 (2002)
3. A. J. Leggett, "Bose-Einstein condensation in the alkali gases: Some fundamental concepts", Rev. Mod. Phys. **73**, 307 (2001)
4. F. Dalfovo, S. Giorgini, L. P. Pitaevskiĭ, and S. Stringari, "Theory of Bose-Einstein condensation in trapped gases", Rev. Mod. Phys. **71**, 463 (1999)
5. I. M. Khalatnikov, *An Introduction to the Theory of Superfluidity* (Addison Wesley, Reading, MA, 1965)
6. H. Minkowski, "Raum und Zeit", Physik. Zeitschr. **10**, 104 (1909)
7. W. G. Unruh, "Experimental Black-Hole Evaporation?", Phys. Rev. Lett. **46**, 1351 (1981)
8. M. Visser, "Acoustic black holes: horizons, ergospheres, and Hawking radiation", Class. Quantum Grav. **15**, 1767 (1998)
9. G. E. Volovik, *The Universe in a Helium Droplet* (Oxford University Press, Oxford, 2003)
10. A. Trautman, "Comparison of Newtonian and relativistic theories of spacetime", in: *Perspectives in Geometry and Relativity* (Indiana University Press, Bloomington, 1966)
11. P. Painlevé, "La mécanique classique et la théorie de la relativité", C. R. Hebd. Acad. Sci. (Paris) **173**, 677 (1921); A. Gullstrand, "Allgemeine Lösung des statischen Einkörperproblems in der Einsteinschen Gravitationstheorie", Arkiv. Mat. Astron. Fys. **16**, 1 (1922)
12. U. R. Fischer and M. Visser, "On the space-time curvature experienced by quasiparticle excitations in the Painlevé-Gullstrand effective geometry", Ann. Phys. (N.Y.) **304**, 22 (2003)
13. E. Madelung, "Quantentheorie in hydrodynamischer Form", Z. Phys. **40**, 322 (1927)
14. L. Onsager, "Statistical Hydrodynamics", Nuovo Cimento Suppl. **6**, 279 (1949)
15. C. Barceló, S. Liberati, and M. Visser, "Analogue gravity from field theory normal modes?", Class. Quantum Grav. **18**, 3595 (2001)
16. T. A. Jacobson and G. E. Volovik, "Event horizons and ergoregions in ^{3}He", Phys. Rev. D **58**, 064021 (1998)
17. W. Gordon, "Zur Lichtfortpflanzung nach der Relativitätstheorie", Ann. Phys. (Leipzig) **72**, 421 (1923); U. Leonhardt, "Space-time geometry of quantum dielectrics", Phys. Rev. A **62**, 012111 (2000); R. Schützhold, G. Plunien, and G. Soff, "Dielectric Black Hole Analogs", Phys. Rev. Lett. **88**, 061101 (2002)

18. R. Schützhold and W. G. Unruh, "Gravity wave analogues of black holes", Phys. Rev. D **66**, 044019 (2002)
19. N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, 1984)
20. W. G. Unruh and R. Schützhold, "On slow light as a black hole analogue", Phys. Rev. D **68**, 024008 (2003)
21. M. Visser, "Hawking radiation without black hole entropy", Phys. Rev. Lett. **80**, 3436 (1998); L. J. Garay, J. R. Anglin, J. I. Cirac, and P. Zoller, "Sonic Analog of Gravitational Black Holes in Bose-Einstein Condensates", Phys. Rev. Lett. **85**, 4643 (2000)
22. P. O. Fedichev and U. R. Fischer, "Gibbons-Hawking Effect in the Sonic de Sitter Space-Time of an Expanding Bose-Einstein-Condensed Gas", Phys. Rev. Lett. **91**, 240407 (2003); "Observer dependence for the phonon content of the sound field living on the effective curved space-time background of a Bose-Einstein condensate", Phys. Rev. D **69**, 064021 (2004)
23. P. O. Fedichev and U. R. Fischer, "'Cosmological' quasiparticle production in harmonically trapped superfluid gases", Phys. Rev. A **69**, 033602 (2004)
24. C. Barceló, S. Liberati, and M. Visser, "Probing semiclassical analog gravity in Bose-Einstein condensates with widely tunable interactions", Phys. Rev. A **68**, 053613 (2003)
25. U. R. Fischer and R. Schützhold, "Quantum simulation of cosmic inflation in two-component Bose-Einstein condensates", Phys. Rev. A **70**, 063615 (2004); R. Schützhold, "Dynamical Zero-Temperature Phase Transitions and Cosmic Inflation or Deflation", Phys. Rev. Lett. **95**, 135703 (2005); M. Uhlmann, Y. Xu, and R. Schützhold, "Aspects of Cosmic Inflation in Expanding Bose-Einstein Condensates", New J. Phys. **7**, 248.1–248.17 (2005)
26. C. Barceló, S. Liberati, and M. Visser, "Analogue Gravity", Living Rev. Relativity **8**, 12.1–12.113 (2005); URL: http://www.livingreviews.org/lrr-2005-12
27. M. Stone, "Acoustic energy and momentum in a moving medium", Phys. Rev. E **62**, 1341 (2000); "Phonons and forces: Momentum *versus* pseudomomentum in moving fluids", p. 335 in M. Novello, M. Visser, and G. Volovik (editors), *Artificial Black Holes* (World Scientific, Singapore, 2002)
28. R. Balbinot, S. Fagnocchi, A. Fabbri, and G. P. Procopio, "Backreaction in acoustic black holes", Phys. Rev. Lett. **94**, 161302 (2005); R. Balbinot, S. Fagnocchi, and A. Fabbri, "Quantum effects in acoustic black holes: The backreaction", Phys. Rev. D **71**, 064019 (2005)
29. R. Schützhold, M. Uhlmann, Y. Xu, and U. R. Fischer, "Quantum backreaction in dilute Bose-Einstein condensates", Phys. Rev. D **72**, 105005 (2005)
30. M. Girardeau and R. Arnowitt, "Theory of Many-Boson Systems: Pair Theory", Phys. Rev. **113**, 755 (1959); C. W. Gardiner, "Particle-number-conserving Bogoliubov method which demonstrates the validity of the time-dependent Gross-Pitaevskiĭ equation for a highly condensed Bose gas", Phys. Rev. A **56**, 1414 (1997); M. D. Girardeau, *ibid.* **58**, 775 (1998)
31. Y. Castin and R. Dum, "Low-temperature Bose-Einstein condensates in time-dependent traps: Beyond the U(1) symmetry-breaking approach", Phys. Rev. A **57**, 3008 (1998)
32. E. H. Lieb, R. Seiringer, and J. Yngvason, "Bosons in a trap: A rigorous derivation of the Gross-Pitaevskiĭ energy functional", Phys. Rev. A **61**, 043602 (2000)
33. The total particle number operator $\hat{N} = \hat{A}^\dagger \hat{A}$ and the corresponding creation and annihilation operators satisfy $[\hat{A}, \hat{A}^\dagger] = 1$ and

$[\hat{\chi}, \hat{A}\hat{N}^{-1/2}] = [\hat{\zeta}, \hat{A}\hat{N}^{-1/2}] = 0$, i.e., the excitations $\hat{\chi}$ and $\hat{\zeta}$ are particle-number-conserving, and thus the full Hamiltonian can be written in terms of these operators, cf. [30,31]. The mean-field ansatz in Eq. (12) can be motivated by starting with N free particles, $g = 0$, in the same single-particle state ψ_c with $\hat{\zeta} = 0$ and subsequently switching on the coupling $g > 0$ by following the evolution in Eqs. (13)–(15) such that the corrections $\hat{\zeta} = \mathcal{O}(1/\sqrt{N})$ remain small [34]

34. R. Schützhold, M. Uhlmann, Y. Xu, and U. R. Fischer, "Mean-field expansion in Bose-Einstein condensates with finite-range interactions", Int. J. Mod. Phys. B **20,** 3555 (2006)
35. E. P. Gross, "Structure of a Quantized Vortex in Boson Systems", Nuovo Cimento **20**, 454 (1961); "Hydrodynamics of a superfluid condensate", J. Math. Phys. **4**, 195 (1963); L. P. Pitaevskiĭ, "Vortex lines in an imperfect Bose gas", Sov. Phys. JETP **13**, 451 (1961)
36. T. D. Lee and C. N. Yang, "Many-Body Problem in Quantum Mechanics and Quantum Statistical Mechanics", Phys. Rev. **105**, 1119 (1957); T. D. Lee, K. Huang, and C. N. Yang, "Eigenvalues and Eigenfunctions of a Bose System of Hard Spheres and Its Low-Temperature Properties", *ibid.* **106**, 1135 (1957); E. Timmermans, P. Tommasini, and K. Huang, "Variational Thomas-Fermi theory of a nonuniform Bose condensate at zero temperature", Phys. Rev. A **55**, 3645 (1997)
37. N. N. Bogoliubov, "On the Theory of Superfluidity", J. Phys. (USSR) **11**, 23 (1947); P. G. de Gennes, *Superconductivity of Metals and Alloys* (W. A. Benjamin, New York, 1966)
38. A. Griffin, "Conserving and gapless approximations for an inhomogeneous Bose gas at finite temperatures", Phys. Rev. B **53**, 9341 (1996); E. Zaremba, A. Griffin, and T. Nikuni, "Two-fluid hydrodynamics for a trapped weakly interacting Bose gas", Phys. Rev. A **57**, 4695 (1998)
39. V. I. Yukalov and E. P. Yukalova, "Normal and Anomalous Averages for Systems with Bose-Einstein Condensate", Laser Phys. Lett. **2**, 506 (2005)
40. C. Barceló, S. Liberati, and M. Visser, "Analogue gravity from Bose-Einstein condensates", Class. Quantum Grav. **18**, 1137 (2001)
41. We note that the problem that the "canonical" commutator between density and phase operators leads to fundamental inconsistencies was first pointed out in the context of superfluid hydrodynamics by H. Fröhlich, "A contradiction between quantum hydrodynamics and the existence of particles", Physica **34**, 47 (1967)
42. Y. Castin, "Simple theoretical tools for low dimension Bose gases", J. Phys. IV France **116**, 89 (2004)
43. D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner, and P. Zoller, "Cold Bosonic Atoms in Optical Lattices", Phys. Rev. Lett. **81**, 3108 (1998)
44. M. Olshaniĭ, "Atomic Scattering in the Presence of an External Confinement and a Gas of Impenetrable Bosons", Phys. Rev. Lett. **81**, 938 (1998)
45. A. Görlitz et al., "Realization of Bose-Einstein condensates in lower dimensions", Phys. Rev. Lett. **87**, 130402 (2001)
46. L. Pitaevskiĭ and S. Stringari, "Uncertainty principle and off-diagonal long-range order in the fractional quantum Hall effect", Phys. Rev. B **47**, 10915 (1993)

Analogue Space-time Based on 2-Component Bose–Einstein Condensates

S. Weinfurtner[1], S. Liberati[2], and M. Visser[3]

[1] Victoria University, School of Mathematics, Statistics and Computer Science, PO Box 600, Wellington, New Zealand
silke.weinfurtner@mcs.vuw.ac.nz
[2] International School for Advanced Studies, Via Beirut 2-4, 34014 Trieste, Italy and INFN, Trieste
liberati@sissa.it
[3] Victoria University, School of Mathematics, Statistics and Computer Science, PO Box 600, Wellington, New Zealand
matt.visser@mcs.vuw.ac.nz

Abstract. Analogue space-times are powerful models for probing the fundamental physical aspects of geometry – while one is most typically interested in ultimately reproducing the pseudo–Riemannian geometries of interest in general relativity and cosmology, analogue models can also provide useful physical probes of more general geometries such as pseudo–Finsler space-times. In this chapter we shall see how a 2-component Bose–Einstein condensate can be used to model a specific class of pseudo–Finsler geometries, and after suitable tuning of parameters, both bi-metric pseudo–Riemannian geometries and standard single metric pseudo–Riemannian geometries, while independently allowing the quasi-particle excitations to exhibit a "mass". Furthermore, when extrapolated to extremely high energy the quasi-particles eventually leave the phononic regime and begin to act like free bosons. Thus this analogue space-time exhibits an analogue of the "Lorentz violation" that is now commonly believed to occur at or near the Planck scale defined by the interplay between quantum physics and gravitational physics. In the 2-component Bose–Einstein analogue space-time we will show that the mass generating mechanism for the quasi-particles is related to the size of the Lorentz violations. This relates the "mass hierarchy" to the so-called "naturalness problem". In short the analogue space-time based on 2-component Bose–Einstein condensates exhibits a very rich mathematical and physical structure that can be used to investigate many issues of interest to the high-energy physics, cosmology, and general relativity communities.

1 Introduction and Motivation

Analogue models of curved space-time are interesting for a number of reasons [1]: Sometimes the analogue space-time helps us understand an aspect

S. Weinfurtner et al.: *Analogue Space-Time Based on 2-Component Bose–Einstein Condensates*, Lect. Notes Phys. **718**, 115–163 (2007)
DOI 10.1007/3-540-70859-6_6

of general relativity, sometimes general relativity helps us understand the physics of the analogue space-time, and sometimes we encounter somewhat unusual mathematical structures not normally part of the physics mainstream, with the payoff that one might now develop new opportunities for exploiting the traditional cross-fertilization between theoretical physics and mathematics [2–4].

Specifically, in this chapter we will discuss an analogue space-time based on the propagation of excitations in a 2-component Bose–Einstein condensate (BEC) [5–10]. This analogue space-time has a very rich and complex structure. In certain portions of parameter space the most natural interpretation of the geometry is in terms of a specific class of pseudo–Finsler space-times, and indeed we will see how more generally it is possible to associate a pseudo–Finsler space-time with the leading symbol of a wide class of hyperbolic partial differential equations. In other parts of parameter space, the most natural interpretation of the geometry is in terms of a bi-metric space-time, where one has a manifold that is simultaneously equipped with two distinct pseudo-Riemannian metric tensors. Further specialization in parameter space leads to a region where a single pseudo-Riemannian metric tensor is encountered – this mono-metric regime corresponds to Lorentzian space-times of the type encountered in standard general relativity and cosmology [11–14, 23]. Thus the analogue space-time based on 2-component BECs provides models not just for standard general relativistic space-times, but also for the more general bi-metric, and even more general pseudo–Finsler space-times.

Additionally, the 2-BEC system permits us to provide a mass-generating mechanism for the quasi-particle excitations [5,6]. The specific mass-generating mechanism arising herein is rather different from the Higgs mechanism of the standard model of particle physics, and provides an interesting counterpoint to the more usual ways that mass-generation is achieved. Furthermore, at short distances, where the "quantum pressure" term can no longer be neglected, then even in the mono-metric regime one begins to see deviations from "Lorentz invariance" – and these deviations are qualitatively of the type encountered in "quantum gravity phenomenology", with the interesting property that the Lorentz violating physics is naturally suppressed by powers of the quasi-particle mass divided by the mass of the fundamental bosons that form the condensate [7–10]. So in these analogue systems the mass-generating mechanism is related to the "hierarchy problem" and the suppression of Lorentz-violating physics. The 2-BEC model also allows us to probe the "universality" (or lack thereof) in the Lorentz violating sector [7–10]. More generally, as one moves beyond the hydrodynamic limit in generic pseudo–Finsler parts of parameter space, one can begin to see hints of geometrical structure even more general than the pseudo–Finsler geometries.

While we do not wish to claim that the 2-BEC analogue space-time of this chapter is necessarily a good model for the real physical space-time arising from the putative theory of "quantum gravity" (be it string-model, loop-variable, or lattice based), it is clear that the 2-BEC analogue space-time is an

extraordinarily rich mathematical and physical structure that provides many interesting hints regarding the sort of kinematics and dynamics that one might encounter in a wide class of models for "quantum gravity phenomenology". This is the fundamental reason for our interest in this model, and we hope we can likewise interest the reader in this system and its relatives.

2 Theory of the 2-Component BEC

The basis for our analogue model is an ultra-cold dilute atomic gas of N bosons, which exist in two single-particle states $|\mathrm{A}\rangle$ and $|\mathrm{B}\rangle$. For example, we consider two different hyperfine states, $|F = 1, m_F = -1\rangle$ and $|F = 2, m_F = 1\rangle$ of ^{87}Rb [15,16]. They have different total angular momenta F and therefore slightly different energies. That permits us, from a theoretical point of view, to keep $m_\mathrm{A} \neq m_\mathrm{B}$, even if they are very nearly equal (to about one part in 10^{16}). At the assumed ultra-cold temperatures and low densities the atoms interact only via low-energy collisions, and the 2-body atomic potential can be replaced by a contact potential. That leaves us with with three atom-atom coupling constants, U_AA, U_BB, and U_AB, for the interactions within and between the two hyperfine states. For our purposes it is essential to include an additional laser field, that drives transition between the two single-particle states.

In Fig. 1 the energy levels for different hyperfine states of ^{87}Rb, and possible transitions involving three-level processes, are schematically explained. A more detailed description on how to set up an external field driving the required transitions can be found in [17].

2.1 Gross–Pitaevskii Equation

The rotating-frame Hamiltonian for our closed 2-component system is given by:[1]

$$\hat{H} = \int \mathrm{d}\mathbf{r} \left\{ \sum_{i=\mathrm{A,B}} \left(-\hat{\Psi}_i^\dagger \frac{\hbar^2 \nabla^2}{2m_i} \hat{\Psi}_i + \hat{\Psi}_i^\dagger V_{ext,i}(\mathbf{r}) \hat{\Psi}_i \right) \right.$$
$$\left. + \frac{1}{2} \sum_{i,j=\mathrm{A,B}} \left(U_{ij} \hat{\Psi}_i^\dagger \hat{\Psi}_j^\dagger \hat{\Psi}_i \hat{\Psi}_j + \lambda \hat{\Psi}_i^\dagger (\sigma_x)_{ij} \hat{\Psi}_j \right) \right\}, \tag{1}$$

with the transition rate λ between the two hyperfine states. Here $\hat{\Psi}_i(\mathbf{r})$ and $\hat{\Psi}_i^\dagger(\mathbf{r})$ are the usual boson field annihilation and creation operators for a single-particle state at position $\mathbf{r}$, and σ_x is the usual Pauli matrix. For temperatures

[1] In general, it is possible that the collisions drive coupling to other hyperfine states. Strictly speaking the system is not closed, but it is legitimate to neglect this effect [18].

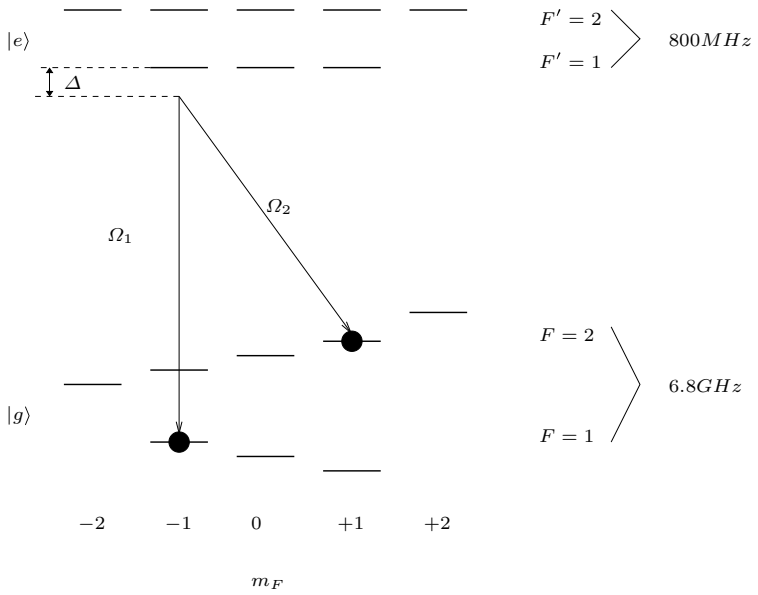

Fig. 1. The *horizontal lines* indicate the hyperfine states of ^{87}Rb. The *arrows* represent two laser fields – with the two frequencies Ω_1 and Ω_2 – necessary to drive transitions between the two trapped states $|F = 1, m_F = -1\rangle$ and $|F = 2, m_F = 1\rangle$, where the frequency difference corresponds to the energy difference of the two hyperfine states. This is realized by a three-level atomic system, because the hyperfine states must be coupled over an intermediate level, that has to lie somewhat below the excited $|e\rangle$ states, as indicated by Δ

at or below the critical BEC temperature, almost all atoms occupy the spatial modes $\Psi_{\mathrm{A}}(\mathbf{r})$ and $\Psi_{\mathrm{B}}(\mathbf{r})$. The mean-field description for these modes,

$$i\,\hbar\,\partial_t \Psi_i = \left[-\frac{\hbar^2}{2\,m_i}\nabla^2 + V_i - \mu_i + U_{ii}\,|\Psi_i|^2 + U_{ij}|\Psi_j|^2 \right] \Psi_i + \lambda\,\Psi_j\,, \qquad (2)$$

are a pair of coupled Gross–Pitaevskii equations (GPE): $(i, j) \to (\mathrm{A}, \mathrm{B})$ or $(i, j) \to (\mathrm{B}, \mathrm{A})$.

2.2 Dynamics

In order to use the above 2-component BEC as an analogue model, we have to investigate small perturbations (sound waves) in the condensate cloud.[2] The excitation spectrum is obtained by linearizing around some background densities ρ_{i0} and phases θ_{i0}, using:

[2] The perturbations amplitude has to be small compared to the overall size of the condensate cloud, so that the system remains in equilibrium.

$$\Psi_i = \sqrt{\rho_{i0} + \varepsilon\,\rho_{i1}}\, e^{i(\theta_{i0}+\varepsilon\,\theta_{i1})} \quad \text{for} \quad i = \text{A, B} \ . \tag{3}$$

To keep the analysis as general as possible, we allow the two initial background phases to be independent from each other, and define

$$\delta_{\text{AB}} \equiv \theta_{\text{A0}} - \theta_{\text{B0}} \ , \tag{4}$$

as their difference.

A tedious calculation [5–7] shows that it is convenient to introduce the following 2×2 matrices: An effective coupling matrix,

$$\hat{\Xi} = \Xi + \hat{X}, \tag{5}$$

where we introduced the energy-independent matrix

$$\Xi \equiv \frac{1}{\hbar} \begin{bmatrix} \tilde{U}_{\text{AA}} & \tilde{U}_{\text{AB}} \\ \tilde{U}_{\text{AB}} & \tilde{U}_{\text{BB}} \end{bmatrix} \ . \tag{6}$$

This matrix contains the quantities

$$\tilde{U}_{\text{AA}} \equiv U_{\text{AA}} - \frac{\lambda \cos\delta_{\text{AB}}\, \sqrt{\rho_{\text{A0}}\rho_{\text{B0}}}}{2} \frac{1}{\rho_{\text{A0}}^2}, \tag{7}$$

$$\tilde{U}_{\text{BB}} \equiv U_{\text{BB}} - \frac{\lambda \cos\delta_{\text{AB}}\, \sqrt{\rho_{\text{A0}}\rho_{\text{B0}}}}{2} \frac{1}{\rho_{\text{B0}}^2}, \tag{8}$$

$$\tilde{U}_{\text{AB}} \equiv U_{\text{AB}} + \frac{\lambda \cos\delta_{\text{AB}}\, \sqrt{\rho_{\text{A0}}\rho_{\text{B0}}}}{2} \frac{1}{\rho_{\text{A0}}\,\rho_{\text{B0}}} \ . \tag{9}$$

A second matrix, denoted $\hat{X}$, contains differential operators $\hat{Q}_{X1}$ – these are the second-order differential operators obtained from linearizing the quantum potential:

$$V_{\text{Q}}(\rho_X) \equiv -\frac{\hbar^2}{2m_X} \left(\frac{\nabla^2\sqrt{\rho_X}}{\sqrt{\rho_X}} \right) = -\frac{\hbar^2}{2m_X} \left(\frac{\nabla^2\sqrt{\rho_{X0} + \varepsilon\rho_{X1}}}{\sqrt{\rho_{X0} + \varepsilon\rho_{X_1}}} \right) \tag{10}$$

$$= -\frac{\hbar^2}{2m_X} \left(\hat{Q}_{X0}(\rho_{X0}) + \varepsilon\, \hat{Q}_{X1}(\rho_{X1}) \right). \tag{11}$$

The quantity $\hat{Q}_{X0}(\rho_{X0})$ corresponds to the background value of the quantum pressure, and contributes only to the background equations of motion – it does not affect the fluctuations. Now in a general background

$$\hat{Q}_{X1}(\rho_{X1}) = \frac{1}{2} \left\{ \frac{(\nabla\rho_{X0})^2 - (\nabla^2\rho_{X0})\rho_{X0}}{\rho_{X0}^3} - \frac{\nabla\rho_{X0}}{\rho_{X0}^2}\nabla + \frac{1}{\rho_{X0}}\nabla^2 \right\} \rho_{X1} \ , \tag{12}$$

and we define the matrix $\hat{X}$ to be

$$\hat{X} \equiv -\frac{\hbar}{2} \begin{bmatrix} \frac{\hat{Q}_{\text{A1}}}{m_{\text{A}}} & 0 \\ 0 & \frac{\hat{Q}_{\text{B1}}}{m_{\text{B}}} \end{bmatrix}. \tag{13}$$

Given the background homogeneity that will be appropriate for later parts of the current discussion, this will ultimately simplify to

$$\hat{Q}_{X1}(\rho_{X1}) = \frac{1}{2\rho_{X0}}\nabla^2\rho_{X1}\ , \tag{14}$$

in which case

$$\hat{X} = -\frac{\hbar}{4}\begin{bmatrix} \frac{1}{m_{\rm A}\,\rho_{\rm A0}} & 0 \\ 0 & \frac{1}{m_{\rm B}\,\rho_{\rm B0}} \end{bmatrix}\nabla^2 = -X\,\nabla^2\,. \tag{15}$$

Without transitions between the two hyperfine states, when $\lambda = 0$, the matrix Ξ only contains the coupling constants $\Xi_{ij} \to U_{ij}/\hbar$. While Ξ is independent of the energy of the perturbations, the importance of $\hat{X}$ increases with the energy of the perturbation. In the so-called hydrodynamic approximation $\hat{X}$ can be neglected, effectively $\hat{X} \to 0$ and $\hat{\Xi} \to \Xi$.

Besides the interaction matrix, we also introduce a transition matrix,

$$\Lambda \equiv -\frac{2\lambda\,\cos\delta_{\rm AB}\,\sqrt{\rho_{i0}\,\rho_{j0}}}{\hbar}\begin{bmatrix} +1 & -1 \\ -1 & +1 \end{bmatrix} \tag{16}$$

and a mass-density matrix,

$$D \equiv \hbar\begin{bmatrix} \frac{\rho_{\rm A0}}{m_{\rm A}} & 0 \\ 0 & \frac{\rho_{\rm B0}}{m_{\rm B}} \end{bmatrix} \equiv \hbar\begin{bmatrix} d_{\rm A} & 0 \\ 0 & d_{\rm B} \end{bmatrix}\ . \tag{17}$$

The final step is to define two column vectors,

$$\bar{\theta} \equiv [\theta_{\rm A1}, \theta_{\rm B1}]^T\ , \tag{18}$$

and

$$\bar{\rho} \equiv [\rho_{\rm A1}, \rho_{\rm B1}]^T\ . \tag{19}$$

We then obtain two compact equations for the perturbation in the phases and densities:

$$\dot{\bar{\theta}} = -\,\hat{\Xi}\,\bar{\rho} - \boldsymbol{V}\cdot\nabla\bar{\theta} + \Theta\,\bar{\theta}, \tag{20}$$

$$\dot{\bar{\rho}} = -\nabla\cdot\left(D\,\nabla\bar{\theta} + \bar{\rho}\,\boldsymbol{V}\right) - \Lambda\,\bar{\theta} - \Theta^T\bar{\rho}\,. \tag{21}$$

Here the background velocity matrix simply contains the two background velocities of each condensate,

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{v}_{\rm A0} & 0 \\ 0 & \boldsymbol{v}_{\rm B0} \end{bmatrix}\ , \tag{22}$$

with two possibly distinct background velocities,

$$\begin{aligned} \boldsymbol{v}_{\rm A0} &= \frac{\hbar}{m_{\rm A}}\nabla\theta_{\rm A0}\ , \\ \boldsymbol{v}_{\rm B0} &= \frac{\hbar}{m_{\rm B}}\nabla\theta_{\rm B0}\ . \end{aligned} \tag{23}$$

Additionally we also introduce the matrix Θ, which depends on the difference of the initial phases and is defined as

$$\Theta \equiv \frac{\lambda \sin \delta_{\rm AB}}{\hbar} \begin{bmatrix} +\sqrt{\frac{\rho_{\rm B0}}{\rho_{\rm A0}}} & -\sqrt{\frac{\rho_{\rm B0}}{\rho_{\rm A0}}} \\ +\sqrt{\frac{\rho_{\rm A0}}{\rho_{\rm B0}}} & -\sqrt{\frac{\rho_{\rm A0}}{\rho_{\rm B0}}} \end{bmatrix} . \tag{24}$$

Now combine these two equations into one:

$$\begin{aligned} \partial_t(\hat{\Xi}^{-1}\,\dot{\bar{\theta}}) = &- \partial_t\left(\hat{\Xi}^{-1}\,\boldsymbol{V}\cdot\nabla\bar{\theta}\right) - \nabla(\boldsymbol{V}\,\hat{\Xi}^{-1}\,\dot{\bar{\theta}}) \\ &+ \nabla\cdot\left[\left(D - \boldsymbol{V}\,\hat{\Xi}^{-1}\,\boldsymbol{V}\right)\nabla\bar{\theta}\right] + \Lambda\,\bar{\theta} \\ &+ \mathrm{K}\,\bar{\theta} + \frac{1}{2}\left\{\Gamma^a\partial_a\bar{\theta} + \partial_a(\Gamma^a\bar{\theta})\right\}, \end{aligned} \tag{25}$$

where the index a runs from 0–3 (that is, over both time and space), and we now define

$$\Gamma^t = \hat{\Xi}^{-1}\Theta - \Theta^T\hat{\Xi}^{-1}, \tag{26}$$

$$\Gamma^i = \boldsymbol{V}\hat{\Xi}^{-1}\Theta - \Theta^T\hat{\Xi}^{-1}\boldsymbol{V}, \tag{27}$$

and

$$\mathrm{K} = \Theta^T\hat{\Xi}^{-1}\Theta + \frac{1}{2}\partial_t(\hat{\Xi}^{-1}\Theta + \Theta^T\hat{\Xi}^{-1}) + \frac{1}{2}\nabla(\boldsymbol{V}\hat{\Xi}^{-1}\Theta + \Theta^T\hat{\Xi}^{-1}\boldsymbol{V}) . \tag{28}$$

Note that the Γ^a matrices are antisymmetric in field-space (A $\leftrightarrow$ B), while the matrix K is symmetric. Also, both $\Gamma^a \to 0$ and $\mathrm{K} \to 0$ as $\delta_{\rm AB} \to 0$.

Our first goal is to show that equation (25), which fundamentally describes quasi-particle excitations interacting with a condensed matter system in the mean-field approximation, can be given a physical and mathematical interpretation in terms of a classical background geometry for massless and massive particles propagating through an analogue space-time [2–4, 19]. This analogy only holds (at least in its cleanest form) in the so-called hydrodynamic limit $\hat{\Xi} \to \Xi$, which limit is directly correlated with the healing length which we shall now introduce.

2.3 Healing Length

The differential operator $\hat{Q}_{X1}$ that underlies the origin of the $\hat{X}$ contribution above is obtained by linearizing the quantum potential

$$V_{\rm Q}(\rho_X) \equiv -\frac{\hbar^2}{2m_X}\left(\frac{\nabla^2\sqrt{\rho_X}}{\sqrt{\rho_X}}\right) \tag{29}$$

which appears in the Hamilton–Jacobi equation of the BEC flow. This quantum potential term is suppressed by the smallness of $\hbar$, the comparative largeness of m_X, and for sufficiently uniform density profiles. But of course in any real system the density of a BEC must go to zero at the boundaries of its electro-magnetic trap (given that $\rho_X = |\psi_X(\boldsymbol{x}, t)|^2$). In a 1-component BEC the healing length characterizes the minimal distance over which the order parameter goes from zero to its bulk value. If the condensate density grows from zero to ρ_0 within a distance ξ the quantum potential term (non local) and the interaction energy (local) are respectively $E_{\text{kinetic}} \sim \hbar^2/(2m\xi^2)$ and $E_{\text{interaction}} \sim 4\pi\hbar^2 a\rho_0/m$. These two terms are comparable when

$$\xi = (8\pi\rho_0 a)^{-1/2}, \tag{30}$$

where a is the s-wave scattering length defined as

$$a = \frac{m\, U_0}{4\pi\hbar^2}\,. \tag{31}$$

Note that what we call U_0 in the above expression is just the coefficient of the non-linear self-coupling term in the Gross–Pitaevskii equation, i.e., just U_{AA} or U_{BB} if we completely decouple the 2 BECs ($U_{\text{AB}} = \lambda = 0$).

Only for excitations with wavelengths much larger than the healing length is the effect of the quantum potential negligible. This is called the hydrodynamic limit because the single–BEC dynamics is then described by the continuity and Hamilton–Jacobi equations of a super-fluid, and its excitations behave like massless phononic modes. In the case of excitations with wavelengths comparable with the healing length this approximation is no longer appropriate and deviations from phononic behaviour will arise.

Such a simple discrimination between different regimes is lost once one considers a system formed by two coupled Bose–Einstein condensates. One is forced to introduce a generalization of the healing ξ length in the form of a "healing matrix". If we apply the same reasoning used above for the definition of the "healing length" to the 2-component BEC system we again find a functional form like that of equation (30) however we now have the crucial difference that both the density and the scattering length are replaced by matrices. In particular, we generalize the scattering length a to the matrix $\mathcal{A}$:

$$\mathcal{A} = \frac{1}{4\pi\hbar^2}\begin{bmatrix}\sqrt{m_{\text{A}}} & 0\\ 0 & \sqrt{m_{\text{B}}}\end{bmatrix}\begin{bmatrix}\tilde{U}_{\text{AA}} & \tilde{U}_{\text{AB}}\\ \tilde{U}_{\text{AB}} & \tilde{U}_{\text{BB}}\end{bmatrix}\begin{bmatrix}\sqrt{m_{\text{A}}} & 0\\ 0 & \sqrt{m_{\text{B}}}\end{bmatrix}. \tag{32}$$

Furthermore, from (30) a healing length matrix Y can be defined by

$$Y^{-2} = \frac{2}{\hbar^2}\begin{bmatrix}\sqrt{\rho_{\text{A}0} m_{\text{A}}} & 0\\ 0 & \sqrt{\rho_{\text{B}0} m_{\text{B}}}\end{bmatrix}\begin{bmatrix}\tilde{U}_{\text{AA}} & \tilde{U}_{\text{AB}}\\ \tilde{U}_{\text{AB}} & \tilde{U}_{\text{BB}}\end{bmatrix}\begin{bmatrix}\sqrt{\rho_{\text{A}0} m_{\text{A}}} & 0\\ 0 & \sqrt{\rho_{\text{B}0} m_{\text{B}}}\end{bmatrix}. \tag{33}$$

That is, in terms of the matrices we have so far defined:

$$Y^{-2} = \frac{1}{2}\, X^{-1/2}\, \Xi\, X^{-1/2}; \qquad Y^2 = 2\, X^{1/2}\, \Xi^{-1}\, X^{1/2}. \tag{34}$$

Define "effective" scattering lengths and healing lengths for the 2-BEC system as

$$a_{\text{eff}} = \frac{1}{2}\ \text{Tr}[\mathcal{A}] = \frac{m_{\text{A}}\tilde{U}_{\text{AA}} + m_{\text{B}}\tilde{U}_{\text{BB}}}{8\pi\hbar^2}\ , \tag{35}$$

and

$$\xi^2_{\text{eff}} = \frac{1}{2}\ \text{Tr}[Y^2] = \text{Tr}[X\Xi^{-1}] = \frac{\hbar^2[\tilde{U}_{\text{BB}}/(m_{\text{A}}\rho_{\text{A0}}) + \tilde{U}_{\text{AA}}/(m_{\text{B}}\rho_{\text{B0}})]}{4(\tilde{U}_{\text{AA}}\tilde{U}_{\text{BB}} - \tilde{U}^2_{\text{AB}})}\ . \tag{36}$$

That is

$$\xi^2_{\text{eff}} = \frac{\hbar^2[m_{\text{A}}\rho_{\text{A0}}\tilde{U}_{\text{AA}} + m_{\text{B}}\rho_{\text{B0}}\tilde{U}_{\text{BB}}]}{4m_{\text{A}}m_{\text{B}}\rho_{\text{A0}}\rho_{\text{B0}}\ (\tilde{U}_{\text{AA}}\tilde{U}_{\text{BB}} - \tilde{U}^2_{\text{AB}})}\ . \tag{37}$$

Note that if the two components are decoupled and tuned to be equivalent to each other, then these effective scattering and healing lengths reduce to the standard one-component results.

3 Emergent Space-time at Low Energies

The basic idea behind analogue models is to re-cast the equation for excitations in a fluid into the equation describing a massless or massive scalar field embedded in a pseudo–Riemannian geometry. Starting from a two component superfluid we are going to show that it is not only possible to obtain a massive scalar field from such an analogue model, in addition we are also able to model much more complex geometries. In Fig. 2 we illustrate how excitations in a 2-component BEC are associated with various types of emergent geometry.

Most generally, we show that excitations in a 2-component BEC (in the hydrodynamic limit) can be viewed as propagating through a specific class of pseudo–Finsler geometry. As additional constraints are placed on the BEC parameter space, the geometry changes from pseudo–Finsler, first to bi-metric, and finally to mono-metric (pseudo–Riemannian, Lorentzian) geometry. This can be accomplished by tuning the various BEC parameters, such as the transition rate λ, the background velocities $\boldsymbol{v}_{\text{A0}}$, $\boldsymbol{v}_{\text{B0}}$, the background densities ρ_{A0}, ρ_{B0}, and the coupling between the atoms U_{AA}, U_{BB} and U_{AB}.

At first, it might seem to be quite an artificial thing to impose such constraints onto the system. But if one considers that the two macroscopic wave functions represent two interacting classical fields, it is more or less obvious that this is the *only* way in which to enforce physical constraints onto the fields themselves, and on the way they communicate with each other.

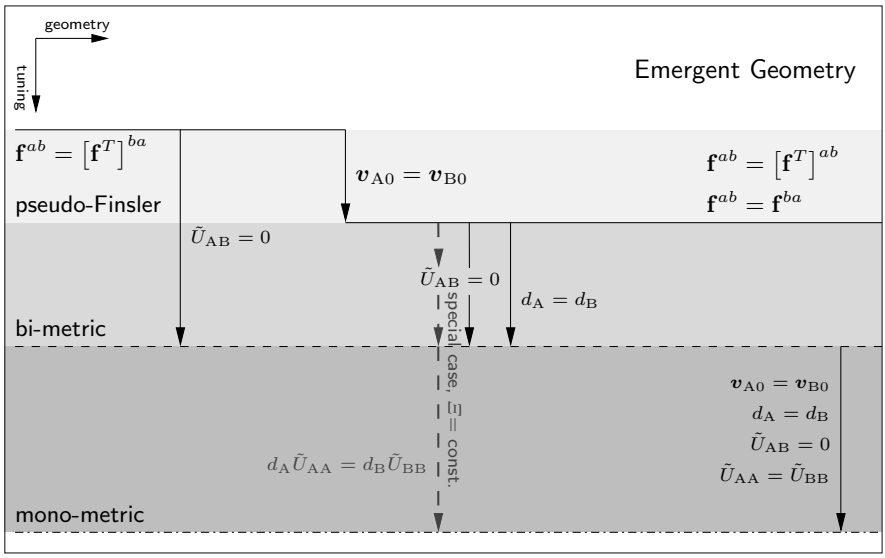

Fig. 2. The dependence of the emergent geometry on the 2-component BEC parameters

3.1 Pseudo-Finsler Geometry

In the hydrodynamic limit ($\hat{\Xi} \to \Xi$), it is possible to simplify equation (25) – without enforcing any constraints on the BEC parameters – if we adopt a (3+1)-dimensional "space-time" notation by writing $x^a = (t, x^i)$, with $i \in \{1, 2, 3\}$ and $a \in \{0, 1, 2, 3\}$. Then equation (25) can be very compactly rewritten as [2, 3]:

$$\partial_a \left(\mathbf{f}^{ab}\, \partial_b \bar{\theta}\right) + (\Lambda + \mathrm{K})\ \bar{\theta} + \frac{1}{2}\left\{\Gamma^a\, \partial_a \bar{\theta} + \partial_a(\Gamma^a \bar{\theta})\right\} = 0\,. \tag{38}$$

The object $\mathbf{f}^{ab}$ is a 4×4 space-time matrix (actually a tensor density), each of whose components is a 2×2 matrix in field-space – equivalently this can be viewed as a 2×2 matrix in field-space each of whose components is a 4×4 space-time tensor density. By inspection this is a self-adjoint second-order linear system of PDEs. The space-time geometry is encoded in the leading-symbol of the PDEs, namely the $\mathbf{f}^{ab}$, without considering the other subdominant terms. That this is a sensible point of view is most easily seem by considering the usual curved-space-time d'Alembertian equation for a charged particle interacting with a scalar potential in a standard pseudo–Riemannian geometry

$$\frac{1}{\sqrt{-g}}[\partial_a - iA_a]\left(\sqrt{-g}g^{ab}[\partial_b - iA_b]\theta\right) + V\theta = 0 \tag{39}$$

from which it is clear that we want to make the analogy

$$\mathbf{f}^{ab} \sim \sqrt{-g}\, g^{ab} \tag{40}$$

as the key quantity specifying the geometry. In addition

$$\Gamma^a \sim iA^a \qquad \text{and} \qquad \Lambda + \mathrm{K} \sim V - g^{ab} A_a A_b \tag{41}$$

so that Γ^a is analogous to a vector potential and Λ (plus corrections) is related to the scalar potential V – in a translation invariant background this will ultimately provide a mass term.

Specifically in the current 2-BEC system we have

$$\mathbf{f}^{ab} = \left(\begin{array}{c|c} -\Xi^{-1} & -(\boldsymbol{V}\Xi^{-1})^T \\ \hline -\boldsymbol{V}\Xi^{-1} & D - \boldsymbol{V}\Xi^{-1}\boldsymbol{V}^T \end{array}\right) , \tag{42}$$

where

$$\boldsymbol{V}^T = \begin{bmatrix} \boldsymbol{v}_{\mathrm{A}0}^T & 0 \\ 0 & \boldsymbol{v}_{\mathrm{B}0}^T \end{bmatrix} \tag{43}$$

is a 2×2 matrix in field space that is also a row vector in physical 3-space. Overall, this does look like a rather complicated object. However, it is possible to re-write the 4×4 geometry containing 2×2 matrices as its elements, in form of a single $(2 \cdot 4) \times (2 \cdot 4)$ matrix.[3] Explicitly we have[4]

$$\mathbf{f}^{ab} = \left[\begin{array}{c|c} \Xi_{11}^{-1}\left(\begin{array}{c|c} -1 & -\boldsymbol{v}_{\mathrm{A}0}^T \\ \hline -\boldsymbol{v}_{\mathrm{A}0} & \frac{d_{\mathrm{A}}}{\Xi_{11}^{-1}}\delta_{ij} - \boldsymbol{v}_{\mathrm{A}0}\boldsymbol{v}_{\mathrm{A}0}^T \end{array}\right) & \Xi_{12}^{-1}\left(\begin{array}{c|c} 1 & \boldsymbol{v}_{\mathrm{B}0}^T \\ \hline \boldsymbol{v}_{\mathrm{A}0} & \boldsymbol{v}_{\mathrm{A}0}\boldsymbol{v}_{\mathrm{B}0}^T \end{array}\right) \\ \hline \Xi_{21}^{-1}\left(\begin{array}{c|c} 1 & \boldsymbol{v}_{\mathrm{A}0}^T \\ \hline \boldsymbol{v}_{\mathrm{B}0} & \boldsymbol{v}_{\mathrm{B}0}\boldsymbol{v}_{\mathrm{A}0}^T \end{array}\right) & \Xi_{22}^{-1}\left(\begin{array}{c|c} -1 & -\boldsymbol{v}_{\mathrm{B}0}^T \\ \hline -\boldsymbol{v}_{\mathrm{B}0} & \frac{d_{\mathrm{B}}}{\Xi_{22}^{-1}}\delta_{ij} - \boldsymbol{v}_{\mathrm{B}0}\boldsymbol{v}_{\mathrm{B}0}^T \end{array}\right) \end{array}\right] , \tag{44}$$

which we can re-write as

$$\mathbf{f} = \left[\begin{array}{c|c} -\Xi_{11}^{-1}\,\mathcal{V}_1\mathcal{V}_1^T + D_{11}h & -\Xi_{12}^{-1}\,\mathcal{V}_1\mathcal{V}_2^T + D_{12}h \\ \hline -\Xi_{21}^{-1}\,\mathcal{V}_2\mathcal{V}_1^T + D_{21}h & -\Xi_{22}^{-1}\,\mathcal{V}_2\mathcal{V}_2^T + D_{22}h \end{array}\right] , \tag{45}$$

where

$$\mathcal{V}_1^a := \left(1, \boldsymbol{v}_{\mathrm{A}0}^i\right) , \tag{46}$$

$$\mathcal{V}_2^a := \left(1, \boldsymbol{v}_{\mathrm{B}0}^i\right) , \tag{47}$$

and

$$h^{ab} := \mathrm{diag}(0, 1, 1, 1) . \tag{48}$$

Even simpler is the form

[3] This result can be generalized for n-component systems. Any 4×4 geometry obtained from a n-component system can be re-written as a single $(n \cdot 4) \times (n \cdot 4)$ matrix.

[4] Note $\Xi_{12} = \Xi_{21}$, so $\Xi_{12}^{-1} = \Xi_{21}^{-1}$.

$$\mathbf{f} = \left[\begin{array}{c|c} -\Xi_{11}^{-1}\,\mathcal{V}_1\mathcal{V}_1^T & -\Xi_{12}^{-1}\,\mathcal{V}_1\mathcal{V}_2^T \\ \hline -\Xi_{21}^{-1}\,\mathcal{V}_2\mathcal{V}_1^T & -\Xi_{22}^{-1}\,\mathcal{V}_2\mathcal{V}_2^T \end{array}\right] + D\otimes h\ . \tag{49}$$

The key point is that this allows us to write

$$\mathbf{f}^{ab} = \begin{bmatrix} f_{11}^{ab} & f_{12}^{ab} \\ f_{21}^{ab} & f_{22}^{ab} \end{bmatrix}\ , \tag{50}$$

where

$$\begin{aligned} f_{11}^{ab} &= -\,\Xi_{11}^{-1}\,\mathcal{V}_1^a\mathcal{V}_1^b + D_{11}h^{ab}\ , \\ f_{12}^{ab} &= -\,\Xi_{12}^{-1}\,\mathcal{V}_1^a\mathcal{V}_2^b\ , \\ f_{21}^{ab} &= -\,\Xi_{12}^{-1}\,\mathcal{V}_2^a\mathcal{V}_1^b\ , \\ f_{22}^{ab} &= -\,\Xi_{22}^{-1}\,\mathcal{V}_2^a\mathcal{V}_2^b + D_{22}h^{ab}. \end{aligned} \tag{51}$$

It is also possible to separate the representation of $\mathbf{f}^{ab}$ into field space and position space as follows

$$\begin{aligned} \mathbf{f}^{ab} = &\begin{bmatrix} \Xi_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}\mathcal{V}_1^a\mathcal{V}_1^b + \begin{bmatrix} 0 & 0 \\ 0 & \Xi_{22}^{-1} \end{bmatrix}\mathcal{V}_2^a\mathcal{V}_2^b \\ &+ \begin{bmatrix} 0 & \Xi_{12}^{-1} \\ 0 & 0 \end{bmatrix}\mathcal{V}_1^a\mathcal{V}_2^b + \begin{bmatrix} 0 & 0 \\ \Xi_{21}^{-1} & 0 \end{bmatrix}\mathcal{V}_2^a\mathcal{V}_1^b + Dh^{ab}. \end{aligned} \tag{52}$$

Why do we assert that the quantity $\mathbf{f}^{ab}$ defines a pseudo–Finsler geometry? (Rather than, say, simply a 2×2 matrix of ordinary Lorentzian geometries?) To see the reason for this claim, recall the standard result [20] that the leading symbol of a system of PDEs determines the "signal speed" (equivalently, the characteristics, or the causal structure) [3]. Indeed if we consider the eikonal approximation (while still remaining in the realm of validity of the hydrodynamic approximation) then the causal structure is completely determined by the leading term in the Fresnel equation

$$\det[\mathbf{f}^{ab}k_ak_b] = 0\ , \tag{53}$$

where the determinant is taken in field space. (The quantity $\mathbf{f}^{ab}k_ak_b$ is exactly what is called the leading symbol of the system of PDEs, and the vanishing of this determinant is the statement that high-frequency modes can propagate with wave vector k_a, thereby determining both characteristics and causal structure.) In the 2-BEC case we can explicitly expand the determinant condition as

$$(f_{11}^{ab}k_ak_b)(f_{22}^{cd}k_ck_d) - (f_{12}^{ab}k_ak_b)(f_{21}^{cd}k_ck_d) = 0\ . \tag{54}$$

Define a completely symmetric rank four tensor

$$Q^{abcd} \equiv f_{11}^{(ab}f_{22}^{cd)} - f_{12}^{(ab}f_{21}^{cd)}\ , \tag{55}$$

then the determinant condition is equivalent to

$$Q^{abcd}k_a k_b k_c k_d = 0 \ , \tag{56}$$

which now defines the characteristics in terms of the vanishing of the pseudo-co-Finsler structure

$$Q(k) = Q^{abcd}k_a k_b k_c k_c \ , \tag{57}$$

defined on the cotangent bundle. As explained in appendix A, this pseudo-co-Finsler structure can be Legendre transformed to provide a pseudo–Finsler structure, a Finslerian notion of distance

$$ds^4 = g_{abcd}\ dx^a dx^b dx^c dx^d \ . \tag{58}$$

Here the completely symmetric rank 4 tensor g_{abcd} determines the "sound cones" through the relation $ds = 0$. It is interesting to note that a distance function of the form

$$ds = \sqrt[4]{g_{abcd}\ dx^a dx^b dx^c dx^d} \tag{59}$$

first made its appearance in Riemann's inaugural lecture of 1854 [21], though he did nothing further with it, leaving it to Finsler to develop the branch of geometry now bearing his name [22]. The present discussion is sufficient to justify the use of the term "pseudo–Finsler" in the generic 2-BEC situation, but we invite the more mathematically inclined reader to see appendix A for a sketch of how much further these ideas can be taken.

The pseudo–Finsler geometry implicit in (50) is rather complicated compared with the pseudo-Riemannian geometry we actually appear to be living in, at least as long as one accepts standard general relativity as a good description of reality. To mimic real gravity, we need to simplify our model. It is now time to use the major advantage of our analogue model, the ability to tune the BEC parameters, and with it the 2-field background configuration. The first order of business is to decouple $\mathbf{f}^{ab}$ in field space.

3.2 Bi-metric Geometry

The reduction of equation (52) to a diagonal representation in field space (via an orthogonal rotation on the fields),

$$\mathbf{f}^{ab} \to \mathrm{diag}\left[f_{11}^{ab}, f_{22}^{ab}\right] = \mathrm{diag}\left[\sqrt{-g_{11}}\, g_{11}^{ab}, \sqrt{-g_{22}}\, g_{22}^{ab}\right] \ , \tag{60}$$

enforces a bi-metric structure onto the condensate. There are two ways to proceed.

Distinct Background Velocities

For

$$\mathcal{V}_1 \neq \mathcal{V}_2 \ , \tag{61}$$

we require all five 2×2 matrices appearing in (52) to commute with each other. This has the unique solution $\Xi_{12}^{-1} = 0$, whence

$$\tilde{U}_{\rm AB} = 0 \,. \tag{62}$$

We then get

$$\mathbf{f}^{ab} = \begin{bmatrix} \Xi_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \mathcal{V}_1^a \mathcal{V}_1^b + \begin{bmatrix} 0 & 0 \\ 0 & \Xi_{22}^{-1} \end{bmatrix} \mathcal{V}_2^a \mathcal{V}_2^b + D h^{ab} \,. \tag{63}$$

Since D is a diagonal matrix this clearly represents a bi-metric geometry. The relevant parameters are summarized in Table 1.

Equal Background Velocities

For

$$\mathcal{V}_1 = \mathcal{V}_2 \equiv \mathcal{V} \,, \tag{64}$$

we are still dealing with a pseudo–Finsler geometry, one which is now independently symmetric in field space ($\mathbf{f}^{ab} = [\mathbf{f}^T]^{ab}$), and position space $\mathbf{f}^{ab} = \mathbf{f}^{ba}$.[5] In terms of the BEC parameters that means we must set equal the two background velocities, $\boldsymbol{v}_{\rm A0} = \boldsymbol{v}_{\rm B0} \equiv \boldsymbol{v}_0$, and equation (52) is simplified to:

$$\mathbf{f}^{ab} = -\Xi^{-1} \mathcal{V}^a \mathcal{V}^b + D h^{ab} \,. \tag{65}$$

From the above, diagonalizability in field space now additionally requires the commutator of the interaction and mass-density matrix to vanish:

$$[\Xi, D] = 0 \quad \Longrightarrow \quad \tilde{U}_{\rm AB}(d_{\rm A} - d_{\rm B}) = 0 \,. \tag{66}$$

Here, we have a choice between two tuning conditions that do the job:

$$\tilde{U}_{\rm AB} = 0 \qquad \text{or} \qquad d_{\rm A} = d_{\rm B} \,. \tag{67}$$

Under the first option, where $\tilde{U}_{\rm AB} = 0$, the two off-diagonal elements in equation (65) are simply zero, and we get the desired bi-metricity in the form[6]

$$\mathbf{f}^{ab} = \begin{bmatrix} \Xi_{11}^{-1} & 0 \\ 0 & \Xi_{22}^{-1} \end{bmatrix} \mathcal{V}^a \mathcal{V}^b + D h^{ab} \,. \tag{68}$$

Under the second option, for $d_{\rm A} = d_{\rm B} \equiv d$, we have $D = d\,\mathbf{I}$. The situation is now a bit trickier, in the sense that one has to diagonalize Ξ^{-1}:

$$\begin{aligned} \tilde{\Xi}^{-1} &= O^T \, \Xi^{-1} O \\ &= \mathrm{diag}\left[\tfrac{\tilde{U}_{\rm AA}+\tilde{U}_{\rm BB}+\sqrt{(\tilde{U}_{\rm AA}-\tilde{U}_{\rm BB})^2+4\tilde{U}_{\rm AB}^2}}{2\,(\tilde{U}_{\rm AA}\tilde{U}_{\rm BB}-\tilde{U}_{\rm AB}^2)}, \tfrac{\tilde{U}_{\rm AA}+\tilde{U}_{\rm BB}-\sqrt{(\tilde{U}_{\rm AA}-\tilde{U}_{\rm BB})^2+4\tilde{U}_{\rm AB}^2}}{2\,(\tilde{U}_{\rm AA}\tilde{U}_{\rm BB}-\tilde{U}_{\rm AB}^2)} \right]. \end{aligned} \tag{69}$$

Once this is done, the way to proceed is to use the elements of $\tilde{\Xi}^{-1}$ instead of Ξ^{-1} in equation (68). The relevant parameters are summarized in Table 1.

[5] The most general pseudo–Finsler geometry is symmetric under simultaneous exchange of field space and position space: $\mathbf{f}^{ab} = [\mathbf{f}^T]^{ba}$.

[6] We would like to stress that this constraint can be easily fulfilled, at least in the special case $\delta_{\rm AB} = 0$, by tuning the transition rate λ, see equation (9).

Table 1. If the pseudo–Finsler geometry decouples into two independent Lorentzian geometries $f^{ab}_{11} = \sqrt{-g_{11}}g_{11}$ and $f^{ab}_{11} = \sqrt{-g_{11}}g_{11}$, with two distinct speed of sounds c_{11} and c_{22}, we are effectively dealing with a bi-metric Lorentzian metric. The table shows the results from three different tuning scenarios, that are sufficient to drive the 2-component BEC from Finsler to bi-Lorentzian space-time. The rightmost column $d_A = d_B$ is addressed in [23] where the authors analyze cosmic inflation in such a bi-metric system

Bi-metric Tuning Scenarios			
	$\boldsymbol{v}_{\mathrm{A}0} \neq \boldsymbol{v}_{\mathrm{B}0}$	$\boldsymbol{v}_{\mathrm{A}0} = \boldsymbol{v}_{\mathrm{B}0}$	
	$\tilde{U}_{\mathrm{AB}} = 0$	$\tilde{U}_{\mathrm{AB}} = 0$	$d_{\mathrm{A}} = d_{\mathrm{B}}$
$f^{ab}_{11} \propto$	$\left(\begin{array}{c\|c} -1 & -\boldsymbol{v}^T_{\mathrm{A}0} \\ \hline -\boldsymbol{v}_{\mathrm{A}0} & \tilde{U}_{\mathrm{A}0}\, d_{\mathrm{A}}\, h^{ij} - \boldsymbol{v}_{\mathrm{A}0}\boldsymbol{v}^T_{\mathrm{A}0} \end{array}\right)$	$\left(\begin{array}{c\|c} -1 & -\boldsymbol{v}^T_{0} \\ \hline -\boldsymbol{v}_{0} & \tilde{U}_{\mathrm{A}0}\, d_{\mathrm{A}}\, h^{ij} - \boldsymbol{v}_{0}\boldsymbol{v}^T_{0} \end{array}\right)$	$\left(\begin{array}{c\|c} -1 & -\boldsymbol{v}^T_{0} \\ \hline -\boldsymbol{v}_{0} & \tilde{\Xi}^{-1}_{11}\, d\, h^{ij} - \boldsymbol{v}_{0}\boldsymbol{v}^T_{0} \end{array}\right)$
$f^{ab}_{22} \propto$	$\left(\begin{array}{c\|c} -1 & -\boldsymbol{v}^T_{\mathrm{B}0} \\ \hline -\boldsymbol{v}_{\mathrm{B}0} & \tilde{U}_{\mathrm{B}0}\, d_{\mathrm{B}}\, h^{ij} - \boldsymbol{v}_{\mathrm{B}0}\boldsymbol{v}^T_{\mathrm{B}0} \end{array}\right)$	$\left(\begin{array}{c\|c} -1 & -\boldsymbol{v}^T_{0} \\ \hline -\boldsymbol{v}_{0} & \tilde{U}_{\mathrm{B}0}\, d_{\mathrm{B}}\, h^{ij} - \boldsymbol{v}_{0}\boldsymbol{v}^T_{0} \end{array}\right)$	$\left(\begin{array}{c\|c} -1 & -\boldsymbol{v}^T_{0} \\ \hline -\boldsymbol{v}_{0} & \tilde{\Xi}^{-1}_{22}\, d\, h^{ij} - \boldsymbol{v}_{0}\boldsymbol{v}^T_{0} \end{array}\right)$
$g_{11\,ab} \propto$	$\left(\begin{array}{c\|c} -(c^2_{11} - v^2_{\mathrm{A}0}) & -\boldsymbol{v}^T_{\mathrm{A}0} \\ \hline -\boldsymbol{v}_{\mathrm{A}0} & h^{ij} \end{array}\right)$	$\left(\begin{array}{c\|c} -(c^2_{11} - v^2_{0}) & -\boldsymbol{v}^T_{0} \\ \hline -\boldsymbol{v}_{0} & h^{ij} \end{array}\right)$	
$g_{22\,ab} \propto$	$\left(\begin{array}{c\|c} -(c^2_{22} - v^2_{\mathrm{B}0}) & -\boldsymbol{v}^T_{\mathrm{B}0} \\ \hline -\boldsymbol{v}_{\mathrm{B}0} & h^{ij} \end{array}\right)$	$\left(\begin{array}{c\|c} -(c^2_{22} - v^2_{0}) & -\boldsymbol{v}^T_{0} \\ \hline -\boldsymbol{v}_{0} & h^{ij} \end{array}\right)$	
$c^2_{11} =$	$\tilde{U}_{\mathrm{AA}} d_{\mathrm{A}} = \frac{U_{\mathrm{AA}}\rho_{\mathrm{A}0} + U_{\mathrm{AB}}\rho_{\mathrm{B}0}}{m_{\mathrm{A}}}$		$\tilde{\Xi}^{-1}_{11}\, d$
$c^2_{22} =$	$\tilde{U}_{\mathrm{BB}} d_{\mathrm{B}} = \frac{U_{\mathrm{BB}}\rho_{\mathrm{B}0} + U_{\mathrm{AB}}\rho_{\mathrm{A}0}}{m_{\mathrm{B}}}$		$\tilde{\Xi}^{-1}_{22}\, d$

There is a subtlety implicit in setting the background velocities equal that should be made explicit. If $\mathcal{V}_1 = \mathcal{V}_2$ so that $\boldsymbol{v}_{\mathrm{A}0} = \boldsymbol{v}_{\mathrm{B}0}$, then since the masses appear in the relationship between phase and velocity we deduce

$$m_{\mathrm{B}}\theta_{\mathrm{A}0}(t, \boldsymbol{x}) - m_{\mathrm{A}}\theta_{\mathrm{B}0}(t, \boldsymbol{x}) = f(t)\ . \tag{70}$$

If $m_{\mathrm{A}} \neq m_{\mathrm{B}}$, and if the background velocity is nonzero, we must deduce that $\delta_{\mathrm{AB}}(t, x)$ will be at the very least be position dependent, and we will be unable to set it to zero. Alternatively, if we demand $\delta_{\mathrm{AB}} = 0$, and have $\nabla\theta_{\mathrm{A}0}(t, \boldsymbol{x}) = \nabla\theta_{\mathrm{B}0}(t, \boldsymbol{x}) \neq 0$, then we cannot set $\boldsymbol{v}_{\mathrm{A}0} = \boldsymbol{v}_{\mathrm{B}0} \neq 0$. Fortunately this will not seriously affect further developments.

Last, but certainly not least, we present the conditions for a mono-metric geometry in a 2-component BEC.

3.3 Mono-metric Geometry

Despite the fact that there are three different routes to bi-metricity, once one demands mono-metricity, where

$$\mathbf{f}^{ab} = \mathrm{diag}\left[f_{11}^{ab}, f_{11}^{ab}\right] = \mathrm{diag}\left[\sqrt{-g_{11}}\, g_{11}^{ab}, \sqrt{-g_{11}}\, g_{11}^{ab}\right] \ , \tag{71}$$

then one ends up with one set unique of constraints to reduce from pseudo–Finsler to a single-metric Lorentzian geometry, namely:

$$\begin{aligned} &\boldsymbol{v}_{\mathrm{A}0} = \boldsymbol{v}_{\mathrm{B}0} = \boldsymbol{v}_0 \ ; \\ &\tilde{U}_{\mathrm{AB}} = 0 \ ; \\ &\tilde{U}_{\mathrm{AA}} = \tilde{U}_{\mathrm{BB}} = \tilde{U} \ ; \\ &d_{\mathrm{A}} = d_{\mathrm{B}} = d \ . \end{aligned} \tag{72}$$

This tuning completely specifies the space-time geometry, in that

$$f_{11}^{ab} = f_{22}^{ab} \propto \left(\begin{array}{c|c} -1 & -\boldsymbol{v}_0^T \\ \hline -\boldsymbol{v}_0 & \tilde{U} d\, h^{ij} - \boldsymbol{v}_0 \boldsymbol{v}_0^T, \end{array}\right) \tag{73}$$

and after a small calculation we get

$${g_{11}}^{ab} = {g_{22}}^{ab} \propto \left(\begin{array}{c|c} -(c^2 - v_0^2) & -\boldsymbol{v}_0^T \\ \hline -\boldsymbol{v}_0 & h^{ij}, \end{array}\right) \tag{74}$$

where we have defined

$$c^2 = \tilde{U}\, d \ , \tag{75}$$

as the speed of sound.[7]

Throughout the preceding few pages we have analyzed in detail the first term in equation (38), and identified different condensate parameters with different emergent geometries. Since there is more than one term in the wave equation describing excitations in a two-component system, this is not the end of the story. The remaining terms in equation (38), which we might generically view as "mass" and "vector potential" terms, do not directly affect the space-time geometry as such. But when an excitation propagates through a specific analogue space-time geometry, these terms will contribute to the kinematics. It then becomes useful to consider the "mass eigenmodes" in field-space.

3.4 Merging Space-time Geometry with Mass Eigenmodes

The eigenmodes we are interested in are eigenmodes of the field-space matrices occurring in the sub-dominant terms of the wave equation. These eigenmodes (when they exist) do not notice the presence of multiple fields – in our specific case a 2-field system – and therefore propagate nicely through the effective curved space-time. As promised in the abstract and the motivation, we are striving for an analogue model representing a massive scalar field in a mono-metric Lorentzian structure. By using the results from Sect. 3.3 we are able to decouple the first term of equation (38).

[7] The speed of sound for quasi-particle excitations is of course our analogue for the speed of light in real gravity.

In the following we are focusing on two issues: First, we decouple the remaining terms in Eq. (38), and subsequently we check that these eigenmodes do not recouple the geometric term. There is however one more (technical) problem, and that is the fact that the terms we want to associate with the effective mass of the scalar field still contain partial derivatives in time and space, which ultimately implies a dependence on the energy of the propagating modes.[8] Luckily, this problem can be easily circumvented, for equal background phases,[9]

$$\theta_{\mathrm{A}0} = \theta_{\mathrm{B}0} \,, \tag{76}$$

in which case

$$K = \Gamma^t = \Gamma^i = 0 \,. \tag{77}$$

This has the effect of retaining only the matrix Λ among the sub-dominant terms, so that the wave equation becomes

$$\partial_a(\mathbf{f}^{ab}\partial_b\bar{\theta}) + \Lambda\,\bar{\theta} = 0 \,. \tag{78}$$

Due to the fact that the structure of the coupling matrix Λ cannot be changed, its eigenmodes determine the eigenmodes of the overall wave equation. The eigenvectors of Λ are given by

$$\begin{aligned} \mathrm{EV1} &:= [+1, +1] \\ \mathrm{EV2} &:= [-1, +1] \end{aligned} \tag{79}$$

The final step is to make sure that our space-time geometry commutes with the eigenvectors of Λ, that is

$$\left[\mathbf{f}^{ab}, \Lambda\right] = 0 \,. \tag{80}$$

This constraint is only fulfilled in the mono-metric case, where we are dealing with two identical classical fields, that effectively do not communicate with each other.[10] That is, all field matrices are proportional to the identity matrix.

3.5 Special Case: $\Xi = constant$

There is one specific class of geometries we are particularly interested it, and that is when Ξ is a position independent and time independent constant. In the next section we will focus exclusively on this case, and apply it to quantum gravity phenomenology. This case is however, also of interest as an example of an alternate interplay between fine tuning and emergent geometry. Under the

[8] This can be easily be seen by going to the eikonal approximation where $\nabla - i\boldsymbol{k}$ and $\partial_t \to i\omega$.

[9] Note that $\delta_{\mathrm{AB}} = 0$ plus mono-metricity implies either $m_{\mathrm{A}} = m_{\mathrm{B}}$ with arbitrary $\boldsymbol{v}_0 \neq 0$, or $m_{\mathrm{A}} \neq m_{\mathrm{B}}$ with zero $\boldsymbol{v}_0 = 0$. These are exactly the two situations we shall consider below.

[10] While $\tilde{U}_{\mathrm{AB}} = 0$, $U_{\mathrm{AB}} \neq 0$.

assumption that Ξ is position and time independent, we are able to directly manipulate the overall wave equation for the excitations and as a consequence obtain slightly milder tuning conditions for mono-metricity.

Let us define

$$\tilde{\theta} = \Xi^{-1/2}\,\bar{\theta}\,, \tag{81}$$

and multiply the whole wave equation (38) with $\Xi^{1/2}$ from the left. What we are doing is a transformation in field space onto a new basis $\tilde{\theta}$, and in the new basis the wave equation is given by,

$$\partial_a\left(\tilde{\mathbf{f}}^{ab}\,\partial_b\tilde{\theta}\right) + \left(\tilde{\Lambda} + \tilde{\mathrm{K}}\right)\,\tilde{\theta} + \frac{1}{2}\left\{\tilde{\Gamma}^a\,\partial_a\tilde{\theta} + \partial_a(\tilde{\Gamma}^a\tilde{\theta})\right\} = 0\,, \tag{82}$$

where the matrices in field space transform as: $\tilde{\Lambda} = \Xi^{1/2}\Lambda\Xi^{1/2}$, $\tilde{K} = \Xi^{1/2}K\Xi^{1/2}$, $\tilde{\Gamma}^a = \Xi^{1/2}\Gamma^a\Xi^{1/2}$, and the tensor-density as

$$\tilde{\mathbf{f}}^{ab} = \Xi^{1/2}\,\mathbf{f}^{ab}\,\Xi^{1/2}\,. \tag{83}$$

In general, the transformation matrix $\Xi^{1/2}$ is non-diagonal, though always symmetric:[11]

$$\Xi^{1/2} = \frac{\Xi + \sqrt{\det\Xi}\;\mathbf{I}}{\sqrt{\mathrm{Tr}[\Xi] + 2\sqrt{\det\Xi}}}\,. \tag{84}$$

A close look at equation (52), now using the tensor-density $\tilde{\mathbf{f}}^{ab}$, makes it obvious that for

$$\tilde{U}_{\mathrm{AB}} = 0\,, \tag{85}$$

the geometry reduces from pseudo-Finsler to bi-metric. For the sake of keeping the discussion short and easy to follow, we set the background velocities equal, and now get

$$\tilde{\mathbf{f}}^{ab} = \mathcal{V}^a\mathcal{V}^b + \tilde{D}h^{ab}\,. \tag{86}$$

In view of the tuning, $\tilde{U}_{\mathrm{AB}} = 0$, we see

$$\tilde{D} = \mathrm{diag}(\tilde{U}_{\mathrm{AA}}\,d_{\mathrm{A}}, \tilde{U}_{\mathrm{BB}}\,d_{\mathrm{B}})\,. \tag{87}$$

The new mass-density matrix, and therefore the overall geometry, is diagonal in field space; hence we are now dealing with the required bi-metric structure.

So far we are in complete agreement with what we have obtained in our previous analysis, see Fig. 2. However, if we now ask for mono-metricity, we obtain a slightly milder constraint:

$$\tilde{U}_{\mathrm{AA}}\,d_{\mathrm{A}} = \tilde{U}_{\mathrm{BB}}\,d_{\mathrm{B}}\,. \tag{88}$$

Last but not least, we show in detail the results we obtain for this tuning scenario when including the Λ term (the mass term). To avoid confusion, we re-define a few matrices,

[11] See appendix B.

$$C_0^2 = \Xi^{1/2}\, D\, \Xi^{1/2}\,; \qquad \text{and} \qquad \Omega^2 = \Xi^{1/2}\, \Lambda\, \Xi^{1/2}\,. \tag{89}$$

Both C_0^2 and Ω^2 are symmetric matrices. If $[C_0^2,\ \Omega^2] = 0$, which is equivalent to the matrix equation $D\ \Xi\ \Lambda = \Lambda\ \Xi\ D$, and is certainly satisfied in view of the above constraint, then they have common eigenvectors. Decomposition onto the eigenstates of the system results in a pair of independent Klein-Gordon equations

$$\frac{1}{\sqrt{-g_{\mathrm{I/II}}}}\,\partial_a\left\{\sqrt{-g_{\mathrm{I/II}}}\,(g_{\mathrm{I/II}})^{ab}\,\partial_b\tilde{\theta}_{\mathrm{I/II}}\right\} + \omega^2_{\mathrm{I/II}}\,\tilde{\theta}_{\mathrm{I/II}} = 0\ , \tag{90}$$

where the "acoustic metrics" are given by

$$(g_{\mathrm{I/II}})_{ab} \propto \left[\begin{array}{c|c} -\left(c^2 - v_0^2\right) & -\boldsymbol{v_0}^T \\ \hline -\boldsymbol{v_0} & \mathbf{I}_{d\times d} \end{array}\right]\ . \tag{91}$$

The metric components depend only on the background velocity $\boldsymbol{v}_0$ and the common speed of sound c. It is also possible to calculate the eigenfrequencies of the two phonon modes,

$$\omega_I^2 = 0; \qquad \omega_{II}^2 = \mathrm{Tr}[\Omega^2]\ . \tag{92}$$

A zero/ non-zero eigenfrequency corresponds to a zero/ non-zero mass for the phonon mode.

In the eikonal limit we see that the in-phase perturbation will propagate with the speed of sound,

$$\boldsymbol{v}_s = \boldsymbol{v}_0 + \hat{k}\ c\ , \tag{93}$$

while the anti-phase perturbations propagates with a lower group velocity given by:

$$\boldsymbol{v}_g = \frac{\partial\omega}{\partial\boldsymbol{k}} = \boldsymbol{v}_0 + \hat{k}\ \frac{c^2}{\sqrt{\omega_{II}^2\ + c^2\ k^2}}\ . \tag{94}$$

Here k is the usual wave number. The dispersion relation we obtain for the mono-metric structure is Lorentz invariant.

The fact that we have an analogue model representing both massive and massless particles is promising for quantum gravity phenomenology if we now extend the analysis to high-energy phonon modes where the quantum pressure term is significant, and where we consequently expect a breakdown of Lorentz invariance. For the following, we concentrate on the generalization of flat Minkowski space-time, which implies a constant Ξ and zero background velocities, $\boldsymbol{v}_0$. In the language of condensed matter physics, we are thinking of a uniform condensate at rest.

4 Application to Quantum Gravity Phenomenology

In using this 2-BEC model to probe issues of interest to the "quantum gravity phenomenology" community it behooves us to simplify as much as possible

the parts of the model not of direct interest for current considerations. Specifically, we wish to use the "quantum pressure" term as a model for the type of Lorentz violating physics that might occur in the physical universe at or near the Planck scale [24]. Since we are then interested in high energies, and consequently short distances, one might expect the average space-time curvature to be negligible – that is, we will be interested in looking for "quantum pressure" induced deviations from special relativity, and can dispense with the notion of curved space-times for now. ("Flat" pseudo–Finsler spaces are already sufficiently complicated to lead to interesting physics.) In terms of the BEC condensates this means that in this section of the chapter we will concentrate on a spatially-homogeneous time-independent background, so that in particular all the matrices $\mathbf{f}^{ab}$ will be taken to be position-independent. (And similarly, Ξ, Λ, D, etc. are taken to be position independent and we set $\boldsymbol{v}_0 = 0$, so the background is at rest.) We now consider Lorentz invariance breakdown in a flat mono-metric spacetime geometry as indicated in Fig. 3. This greatly simplifies the calculations (though they are still relatively messy), but without sacrificing the essential pieces of the physics we are now interested in.

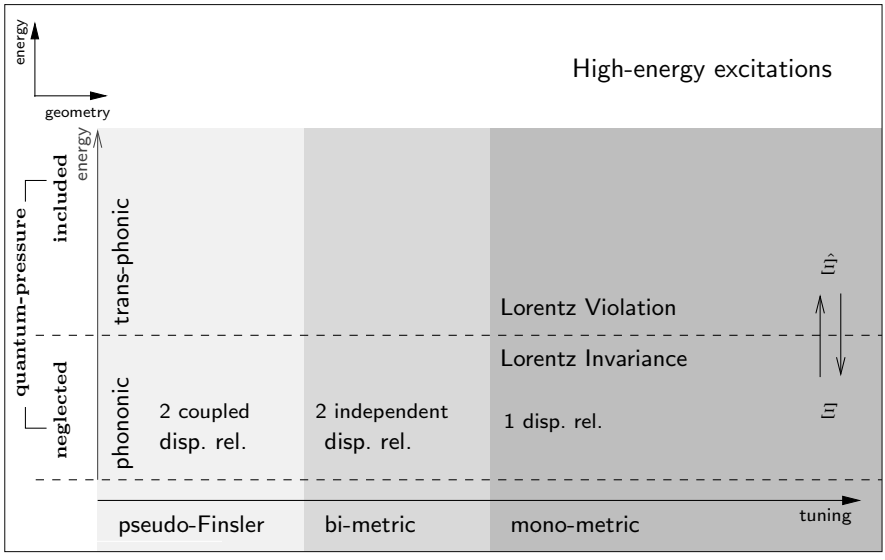

Fig. 3. How to tune the system to exhibit breakdown of Lorentz symmetry

Now the purpose of quantum gravity phenomenology is to analyze the physical consequences arising from various models of quantum gravity. One hope for obtaining an experimental grasp on quantum gravity is the generic prediction arising in many (but not all) quantum gravity models that ultraviolet

physics at or near the Planck scale, $M_{\rm Planck} = 1.2 \times 10^{19}$ GeV/c^2, (or in some models the string scale), typically induces violations of Lorentz invariance at lower scales [25, 26]. Interestingly most investigations, even if they arise from quite different fundamental physics, seem to converge on the prediction that the breakdown of Lorentz invariance can generically become manifest in the form of modified dispersion relations

$$\omega^2 = \omega_0^2 + (1+\eta_2)\ c^2\ k^2 + \eta_4 \left(\frac{\hbar}{M_{\rm Lorentz\ violation}} \right)^2 k^4 + \dots \ , \qquad (95)$$

where the coefficients η_n are dimensionless (and possibly dependent on the particle species considered), and we have restricted our expansion to CPT invariant terms (otherwise one would also get odd powers in k). The particular inertial frame for these dispersion relations is generally specified to be the frame set by cosmological microwave background, and $M_{\rm Lorentz\ violation}$ is the scale of Lorentz symmetry breaking which furthermore is generally assumed to be of the order of $M_{\rm Planck}$.

Although several alternative scenarios have been considered in the literature in order to justify the modified kinematics discussed above, to date the most commonly explored avenue is an effective field theory (EFT) approach. In the present chapter we focus on the class of non-renormalizable EFTs with Lorentz violations associated to dispersion relations like equation (95). Relaxing our CPT invariance condition this class would include the model developed in [27], and subsequently studied by several authors, where an extension of quantum electrodynamics including only mass dimension five Lorentz-violating operators was considered. (That ansatz leads to order k^3 Lorentz and CPT violating terms in the dispersion relation.) Very accurate constraints have been obtained for this model using a combination of experiments and observations (mainly in high energy astrophysics). See e.g. [26, 28–30]. In spite of the remarkable success of this framework as a "test theory", it is interesting to note that there are still significant open issues concerning its theoretical foundations. Perhaps the most pressing one is the so called *naturalness problem* which can be expressed in the following way: Looking back at our ansatz (95) we can see that the lowest-order correction, proportional to η_2, is not explicitly Planck suppressed. This implies that such a term would always be dominant with respect to the higher-order ones and grossly incompatible with observations (given that we have very good constraints on the universality of the speed of light for different elementary particles). Following the observational leads it has been therefore often assumed either that some symmetry (other than Lorentz invariance) enforces the η_2 coefficients to be exactly zero, or that the presence of some other characteristic EFT mass scale $\mu \ll M_{\rm Planck}$ (e.g., some particle physics mass scale) associated with the Lorentz symmetry breaking might enter in the lowest order dimensionless coefficient η_2 – which will be then generically suppressed by appropriate ratios of this characteristic mass to the Planck mass: $\eta_2 \propto (\mu/M_{\rm Planck})^\sigma$ where $\sigma \geq 1$ is some positive power (often taken as one or two). If this is the case then one has two

distinct regimes: For low momenta $p/(M_{\rm Planck}c) \ll (\mu/M_{\rm Planck})^\sigma$ the lower-order (quadratic in the momentum) deviations in (95) will dominate over the higher-order ones, while at high energies $p/(M_{\rm Planck}c) \gg (\mu/M_{\rm Planck})^\sigma$ the higher order terms will be dominant.

The naturalness problem arises because such a scenario is not well justified within an EFT framework; in other words there is no natural suppression of the low-order modifications in these models. In fact we implicitly assumed that there are no extra Planck suppressions hidden in the dimensionless coefficients η_n with $n > 2$. EFT cannot justify why *only* the dimensionless coefficients of the $n \leq 2$ terms should be suppressed by powers of the small ratio $\mu/M_{\rm Planck}$. Even worse, renormalization group arguments seem to imply that a similar mass ratio, $\mu/M_{\rm Planck}$ would implicitly be present also in *all* the dimensionless $n > 2$ coefficients – hence suppressing them even further, to the point of complete undetectability. Furthermore it is easy to show [31] that, without some protecting symmetry, it is generic that radiative corrections due to particle interactions in an EFT with only Lorentz violations of order $n > 2$ in (95) for the free particles, will generate $n = 2$ Lorentz violating terms in the dispersion relation, which will then be dominant. Observational evidence [25] suggests that for a variety of standard model particles $|\eta_2| \lesssim 10^{-21}$. Naturalness in EFT would then imply that the higher order terms are at least as suppressed as this, and hence beyond observational reach.

A second issue is that of "universality", which is not so much a "problem", as an issue of debate as to the best strategy to adopt. In dealing with situations with multiple particles one has to choose between the case of universal (particle-independent) Lorentz violating coefficients η_n, or instead go for a more general ansatz and allow for particle-dependent coefficients; hence allowing different magnitudes of Lorentz symmetry violation for different particles even when considering the same order terms (same n) in the momentum expansion. The two choices are equally represented in the extant literature (see e.g. [32] and [28] for the two alternative ansätze), but it would be interesting to understand how generic this universality might be, and what sort of processes might induce non-universal Lorentz violation for different particles.

4.1 Specializing the Wave Equation

For current purposes, where we wish to probe violations of Lorentz invariance in a flat analogue space-time, we start with our basic wave equation (25) and make the following specializations: $\delta_{\rm AB} \to 0$ (so that $\Gamma^a \to 0$ and $K \to 0$). We also set all background fields to be homogeneous (space and time independent), and use the formal operators $\hat{\Xi}^{1/2}$ and $\hat{\Xi}^{-1/2}$ to define a new set of variables

$$\tilde{\theta} = \hat{\Xi}^{-1/2}\, \bar{\theta}\,, \tag{96}$$

in terms of which the wave equation becomes

$$\partial_t^2 \tilde{\theta} = \left\{ \hat{\Xi}^{1/2}\, [D\nabla^2 - \Lambda]\, \hat{\Xi}^{1/2} \right\} \tilde{\theta}\,, \tag{97}$$

or more explicitly

$$\partial_t^2 \tilde{\theta} = \left\{ [\Xi - X\nabla^2]^{1/2} \, [D\nabla^2 - \Lambda] \, [\Xi - X\nabla^2]^{1/2} \right\} \tilde{\theta} \, . \tag{98}$$

This is now a (relatively) simple PDE to analyze. The objects $\hat{\Xi}^{1/2}$ and $\hat{\Xi}^{-1/2}$ are 2×2 matrices whose elements are pseudo-differential operators, but to simplify things it is computationally efficient to go directly to the eikonal limit where[12]

$$\hat{\Xi} \to \Xi + X \, k^2 \, . \tag{99}$$

This finally leads to a dispersion relation of the form

$$\det \left\{ \omega^2 \, \mathbf{I} - [\Xi + Xk^2]^{1/2} \, [Dk^2 + \Lambda] \, [\Xi + Xk^2]^{1/2} \right\} = 0 \, , \tag{100}$$

and "all" we need to do for the purposes of this chapter, is to understand this quasiparticle excitation spectrum in detail.

4.2 Hydrodynamic Approximation

The hydrodynamic limit consists of formally setting $\hat{X} \to 0$ so that $\hat{\Xi} \to \Xi$. (That is, one is formally setting the healing length matrix to zero: $Y \to 0$. More precisely, all components of the healing length matrix are assumed small compared to other length scales in the problem.) The wave equation (98) now takes the form:

$$\partial_t^2 \tilde{\theta} = \{ \Xi^{1/2} \, [D\nabla^2 - \Lambda] \, \Xi^{1/2} \} \tilde{\theta} \, . \tag{101}$$

Since this is second-order in both space and time derivatives, we now have at least the possibility of obtaining an exact "Lorentz invariance". We can now define the matrices

$$\Omega^2 = \Xi^{1/2} \, \Lambda \, \Xi^{1/2} \, ; \qquad C_0^2 = \Xi^{1/2} \, D \, \Xi^{1/2} \, ; \tag{102}$$

so that after Fourier transformation

$$\omega^2 \tilde{\theta} = \left\{ C_0^2 \, k^2 + \Omega^2 \right\} \, \tilde{\theta} \equiv H(k^2) \, \tilde{\theta} \, , \tag{103}$$

leading to the Fresnel equation

$$\det\{\omega^2 \, \mathbf{I} - H(k^2)\} = 0 \, . \tag{104}$$

That is

$$\omega^4 - \omega^2 \, \mathrm{tr}[H(k^2)] + \det[H(k^2)] = 0 \, , \tag{105}$$

[12] Once we are in the eikonal approximation the pseudo-differential operator $\hat{\Xi}^{1/2} \to \sqrt{\Xi + k^2 X}$ can be given a simple and explicit meaning in terms of the Hamilton–Cayley theorems of appendix B.

whence

$$\omega^2 = \frac{\mathrm{tr}[H(k^2)] \pm \sqrt{\mathrm{tr}[H(k^2)]^2 - 4\ \det[H(k^2)]}}{2} \ . \tag{106}$$

Note that the matrices Ω^2, C_0^2, and $H(k^2)$ have now carefully been arranged to be *symmetric*. This greatly simplifies the subsequent matrix algebra. Also note that the matrix $H(k^2)$ is a function of k^2; this will forbid the appearance of odd powers of k in the dispersion relation – as should be expected due to the parity invariance of the system.

Masses

We read off the "masses" by looking at the special case of space-independent oscillations for which

$$\partial_t^2 \bar{\theta} = -\Omega^2\ \bar{\theta} \ , \tag{107}$$

allowing us to identify the "mass" (more precisely, the natural oscillation frequency) as

$$\text{"masses"} \propto \text{eigenvalues of } (\Xi^{1/2}\ \Lambda\ \Xi^{1/2}) = \text{eigenvalues of } (\Xi\ \Lambda) \ . \tag{108}$$

Since Λ is a singular 2×2 matrix this automatically implies

$$\omega_I^2 = 0; \qquad \omega_{II}^2 = \mathrm{tr}\,(\Xi\ \Lambda) \ . \tag{109}$$

So we see that one mode will be a massless phonon while the other will have a non zero mass. Explicitly, in terms of the elements of the underlying matrices

$$\omega_I^2 = 0; \qquad \omega_{II}^2 = -\frac{2\sqrt{\rho_{\mathrm{A}0}\,\rho_{\mathrm{B}0}}\,\lambda}{\hbar^2}\{\tilde{U}_{\mathrm{AA}} + \tilde{U}_{\mathrm{BB}} - 2\tilde{U}_{\mathrm{AB}}\} \tag{110}$$

so that (before any fine-tuning or decoupling)

$$\omega_{II}^2 = -\frac{2\sqrt{\rho_{\mathrm{A}0}\,\rho_{\mathrm{B}0}}\,\lambda}{\hbar^2} \times \left\{ U_{\mathrm{AA}} + U_{\mathrm{BB}} - 2U_{\mathrm{AB}} - \frac{\lambda}{2\sqrt{\rho_{\mathrm{A}0}\,\rho_{\mathrm{B}0}}} \left[\sqrt{\frac{\rho_{\mathrm{A}0}}{r\rho_{\mathrm{B}0}}} + \sqrt{\frac{\rho_{\mathrm{B}0}}{\rho_{\mathrm{A}0}}} \right]^2 \right\} \ . \tag{111}$$

It is easy to check that this quantity really does have the physical dimensions of a frequency.

Mono-Metricity Conditions

In order for our system to be a perfect analogue of special relativity:

- we want each mode to have a quadratic dispersion relation;
- we want each dispersion relation to have the same asymptotic slope.

Let us start by noticing that the dispersion relation (106) is of the form

$$\omega^2 = [\text{quadratic}_1] \pm \sqrt{[\text{quartic}]}\ . \tag{112}$$

The first condition implies that the quartic must be a perfect square

$$[\text{quartic}] = [\text{quadratic}_2]^2\ , \tag{113}$$

but then the second condition implies that the slope of this quadratic must be zero. That is

$$[\text{quadratic}_2](k^2) = [\text{quadratic}_2](0)\ , \tag{114}$$

and so

$$[\text{quartic}](k^2) = [\text{quartic}](0) \tag{115}$$

must be constant independent of k^2, so that the dispersion relation is of the form

$$\omega^2 = [\text{quadratic}_1](k^2) \pm [\text{quadratic}_2](0)\ . \tag{116}$$

Note that this has the required form (two hyperbolae with the same asymptotes, and possibly different intercepts). Now let us implement this directly in terms of the matrices C_0^2 and M^2.

Step 1: Using the results of the appendix, specifically equation (255):

$$\det[H^2(k)] = \det[\Omega^2 + C_0^2\ k^2] \tag{117}$$

$$= \det[\Omega^2] - \operatorname{Tr}\left\{\Omega^2\ \bar{C}_0^2\right\}\ k^2 + \det[C_0^2]\ (k^2)^2\ . \tag{118}$$

(This holds for any linear combination of 2×2 matrices. Note that we apply trace reversal to the squared matrix C_0^2, we do not trace reverse and then square.) Since in particular $\det[\Omega^2] = 0$, we have:

$$\det[H^2(k)] = -\operatorname{Tr}\left\{\Omega^2\ \bar{C}_0^2\right\}\ k^2 + \det[C_0^2]\ (k^2)^2\ . \tag{119}$$

Step 2: Now consider the discriminant (the quartic)

$$\text{quartic} \equiv \operatorname{tr}[H(k^2)]^2 - 4\ \det[H(k^2)] \tag{120}$$

$$= (\operatorname{Tr}[\Omega^2] + \operatorname{Tr}[C_0^2]\ k^2)^2 - 4\big[- \operatorname{Tr}\left\{\Omega^2\ \bar{C}_0^2\right\}\ k^2 + \det[C_0^2]\ (k^2)^2\big] \tag{121}$$

$$= \operatorname{Tr}[\Omega^2]^2 + \{2\operatorname{Tr}[\Omega^2]\operatorname{Tr}[C_0^2] + 4\operatorname{Tr}\left\{\Omega^2\ \bar{C}_0^2\right\}\}k^2 + \left\{\operatorname{Tr}[C_0^2]^2 - 4\det[C_0^2]\right\}(k^2)^2 \tag{122}$$

$$= \operatorname{Tr}[\Omega^2]^2 + 2\{2\operatorname{Tr}\left\{\Omega^2\ C_0^2\right\} - \operatorname{Tr}[\Omega^2]\operatorname{Tr}[C_0^2]\}k^2 + \left\{\operatorname{Tr}[C_0^2]^2 - 4\det[C_0^2]\right\}(k^2)^2\ . \tag{123}$$

So in the end the two conditions above for mono-metricity take the form

$$\text{mono-metricity} \iff \begin{cases} \operatorname{Tr}[C_0^2]^2 - 4\ \det[C_0^2] = 0; \\ 2\operatorname{Tr}\left\{\Omega^2\ C_0^2\right\} - \operatorname{Tr}[\Omega^2]\ \operatorname{Tr}[C_0^2] = 0. \end{cases} \tag{124}$$

Once these two conditions are satisfied the dispersion relation is

$$\omega^2 = \frac{\mathrm{tr}[H(k^2)] \pm \mathrm{Tr}[\Omega^2]}{2} = \frac{\mathrm{Tr}[\Omega^2] \pm \mathrm{Tr}[\Omega^2] + \mathrm{Tr}[C_0^2]\; k^2}{2} \tag{125}$$

whence

$$\omega_1^2 = \frac{1}{2}\,\mathrm{Tr}[C_0^2]\; k^2 = c_0^2 k^2 \qquad \omega_2^2 = \mathrm{Tr}[\Omega^2] + \frac{1}{2}\,\mathrm{Tr}[C_0^2]\; k^2 = \omega_{II}^2 + c_0^2 k^2\ , \tag{126}$$

as required. One mode is massless, one massive with exactly the "mass" previously deduced. One can now define the quantity

$$m_{II} = \hbar\omega_{II}/c_0^2\ , \tag{127}$$

which really does have the physical dimensions of a mass.

Interpretation of the Mono-Metricity Conditions

But now we have to analyse the two simplification conditions

$$C1: \qquad \mathrm{Tr}[C_0^2]^2 - 4\ \det[C_0^2] = 0\ ; \tag{128}$$

$$C2: \qquad 2\ \mathrm{Tr}\left\{\Omega^2\ C_0^2\right\} - \mathrm{Tr}[\Omega^2]\,\mathrm{Tr}[C_0^2] = 0\ ; \tag{129}$$

to see what they tell us. The first of these conditions is equivalent to the statement that the 2×2 matrix C_0^2 has two identical eigenvalues. But since C_0^2 is symmetric this then implies $C_0^2 = c_0^2\ \mathbf{I}$, in which case the second condition is automatically satisfied. (In contrast, condition $C2$ does not automatically imply condition $C1$.) Indeed if $C_0^2 = c_0^2\ \mathbf{I}$, then it is easy to see that (in order to make C_0^2 diagonal)

$$\tilde{U}_{\mathrm{AB}} = 0\ , \tag{130}$$

(which is sufficient, by itself, to imply bi-metricity) and furthermore that

$$\frac{\tilde{U}_{\mathrm{AA}}\ \rho_{\mathrm{A0}}}{m_{\mathrm{A}}} = c_0^2 = \frac{\tilde{U}_{\mathrm{BB}}\ \rho_{\mathrm{B0}}}{m_{\mathrm{B}}}\ . \tag{131}$$

Note that we can now solve for λ to get

$$\lambda = -2\sqrt{\rho_{\mathrm{A0}}\ \rho_{\mathrm{B0}}}\ U_{\mathrm{AB}}\ , \tag{132}$$

whence

$$c_0^2 = \frac{U_{\mathrm{AA}}\ \rho_{\mathrm{A0}} + U_{\mathrm{AB}}\ \rho_{\mathrm{B0}}}{m_{\mathrm{A}}} = \frac{U_{\mathrm{BB}}\ \rho_{\mathrm{B0}} + U_{\mathrm{AB}}\ \rho_{\mathrm{A0}}}{m_{\mathrm{B}}}\ , \tag{133}$$

and

$$\omega_{II}^2 = \frac{4\rho_{\mathrm{A0}}\rho_{\mathrm{B0}}U_{\mathrm{AB}}}{\hbar^2}\left\{U_{\mathrm{AA}} + U_{\mathrm{BB}} - 2U_{\mathrm{AB}} + U_{\mathrm{AB}}\left[\sqrt{\frac{\rho_{\mathrm{A0}}}{\rho_{\mathrm{B0}}}} + \sqrt{\frac{\rho_{\mathrm{B0}}}{\rho_{\mathrm{A0}}}}\right]^2\right\}\ . \tag{134}$$

Note that (134) is equivalent to (112) with (132) enforced. But this then implies

$$\omega_{II}^2 = \frac{4\rho_{\rm A0}\rho_{\rm B0}U_{\rm AB}}{\hbar^2}\left\{U_{\rm AA}+U_{\rm BB}+U_{\rm AB}\left[\frac{\rho_{\rm A0}}{\rho_{\rm B0}}+\frac{\rho_{\rm B0}}{\rho_{\rm A0}}\right]\right\}. \tag{135}$$

Interpretation: Condition $C2$ forces the two low-momentum "propagation speeds" to be the same, that is, it forces the two $O(k^2)$ coefficients to be equal. Condition $C1$ is the stronger statement that there is no $O(k^4)$ (or higher order) distortion to the relativistic dispersion relation.

4.3 Beyond the Hydrodynamical Approximation

At this point we want to consider the deviations from the previous analogue for special relativity. Our starting point is again equation (98), now retaining the quantum pressure term, which we Fourier transform to get:

$$\omega^2\tilde{\theta} = \left\{\sqrt{\Xi+X\,k^2}\;\;[D\,k^2+\Lambda]\;\;\sqrt{\Xi+X\,k^2}\right\}\;\tilde{\theta} \equiv H(k^2)\;\tilde{\theta}\,. \tag{136}$$

This leads to the Fresnel equation

$$\det\{\omega^2\,\mathbf{I}-H(k^2)\}=0\,. \tag{137}$$

That is

$$\omega^4-\omega^2\;\mathrm{tr}[H(k^2)]+\det[H(k^2)]=0\,, \tag{138}$$

whence

$$\omega^2=\frac{\mathrm{tr}[H(k^2)]\pm\sqrt{\mathrm{tr}[H(k^2)]^2-4\;\det[H(k^2)]}}{2}\,, \tag{139}$$

which is now of the form

$$\omega^2=[\mathrm{quartic}_1]\pm\sqrt{[\mathrm{octic}]}\,. \tag{140}$$

Masses

The "masses", defined as the zero momentum oscillation frequencies, are again easy to identify. Just note that the k-independent term in the Fresnel equation is exactly the same mass matrix $\Omega^2 = \Xi^{1/2}\;\Lambda\;\Xi^{1/2}$ that was present in the hydrodynamical limit. (That is, the quantum potential term X does not influence the masses.)

Dispersion Relations

Differently from the previous case, when the hydrodynamic approximation held, we now have that the discriminant of (139) generically can be an eighth-order polynomial in k. In this case we cannot hope to recover an exact analogue

of special relativity, but instead can at best hope to obtain dispersion relations with vanishing or *suppressed* deviations from special relativity at low k; possibly with large deviations from special relativity at high momenta. From the form of our equation it is clear that the Lorentz violation suppression should be somehow associated with the masses of the atoms $m_{\mathrm{A/B}}$. Indeed we will use the underlying atomic masses to define our "Lorentz breaking scale", which we shall then assume can be identified with the "quantum gravity scale". The exact form and relative strengths of the higher-order terms will be controlled by tuning the 2–BEC system and will eventually decide the manifestation (or not) of the naturalness problem and of the universality issue.

Our approach will again consist of considering derivatives of (139) in growing even powers of k^2 (recall that odd powers of k are excluded by the parity invariance of the system) and then setting $k \to 0$. We shall compute only the coefficients up to order k^4 as by simple dimensional arguments one can expect any higher order term will be further suppressed with respect to the k^4 one.

We can greatly simplify our calculations if before performing our analysis we rearrange our problem in the following way. First of all note that by the cyclic properties of trace

$$\begin{aligned}
\mathrm{Tr}[H(k^2)] &= \mathrm{Tr}[(Dk^2 + \Lambda)\,(\Xi + k^2 X)] && (141)\\
&= \mathrm{Tr}[\Lambda\Xi + k^2(D\Xi + \Lambda X) + (k^2)^2 DX] && (142)\\
&= \mathrm{Tr}[\Xi^{1/2}\Lambda\Xi^{1/2} + k^2(\Xi^{1/2}D\Xi^{1/2} + X^{1/2}\Lambda X^{1/2})\\
&\quad + (k^2)^2 X^{1/2} D X^{1/2}]\,. && (143)
\end{aligned}$$

Putting this all together, we can now define symmetric matrices

$$\Omega^2 = \Xi^{1/2}\Lambda\Xi^{1/2}\,; \tag{144}$$

$$C_0^2 = \Xi^{1/2}D\Xi^{1/2}; \qquad \Delta C^2 = X^{1/2}\Lambda X^{1/2}\,; \tag{145}$$

$$C^2 = C_0^2 + \Delta C^2 = \Xi^{1/2}D\Xi^{1/2} + X^{1/2}\Lambda X^{1/2}\,; \tag{146}$$

$$Z^2 = 2X^{1/2}DX^{1/2} = \frac{\hbar^2}{2}M^{-2}\,. \tag{147}$$

With all these definitions we can then write

$$\mathrm{Tr}[H(k^2)] = \mathrm{Tr}\left[\Omega^2 + k^2(C_0^2 + \Delta C^2) + \frac{1}{2}(k^2)^2 Z^2\right]\,, \tag{148}$$

where everything has been done inside the trace. If we now define

$$H_s(k^2) = \Omega^2 + k^2(C_0^2 + \Delta C^2) + \frac{1}{2}(k^2)^2 Z^2\,, \tag{149}$$

then $H_s(k^2)$ is by definition both polynomial and symmetric and satisfies

$$\mathrm{Tr}[H(k^2)] = \mathrm{Tr}[H_s(k^2)]\,, \tag{150}$$

while in contrast,

$$\det[H(k^2)] \neq \det[H_s(k^2)] \,. \tag{151}$$

But then

$$\omega^2 = \frac{1}{2}\left[\mathrm{Tr}[H_s(k^2)] \pm \sqrt{\mathrm{Tr}[H_s(k^2)]^2 - 4\det[H(k^2)]}\right] \,. \tag{152}$$

Whence

$$\frac{\mathrm{d}\omega^2}{\mathrm{d}k^2} = \frac{1}{2}\left[\mathrm{Tr}[H_s'(k^2)] \pm \frac{\mathrm{Tr}[H_s(k^2)]\,\mathrm{Tr}[H_s'(k^2)] - 2\det'[H(k^2)]}{\sqrt{\mathrm{Tr}[H_s(k^2)]^2 - 4\det[H(k^2)]}}\right] \,, \tag{153}$$

and at $k = 0$

$$\left.\frac{\mathrm{d}\omega^2}{\mathrm{d}k^2}\right|_{k\to 0} = \frac{1}{2}\left[\mathrm{Tr}[C^2] \pm \frac{\mathrm{Tr}[\Omega^2]\,\mathrm{Tr}[C^2] - 2\det'[H(k^2)]_{k\to 0}}{\mathrm{Tr}[\Omega^2]}\right] \,. \tag{154}$$

But now let us consider

$$\begin{aligned}
\det[H(k^2)] &= \det[(Dk^2 + \Lambda)\,(\Xi + k^2 X)] && (155)\\
&= \det[Dk^2 + \Lambda]\;\det[\Xi + k^2 X] && (156)\\
&= \det[\Xi^{1/2}(Dk^2 + \Lambda)\Xi^{1/2}]\;\det[I + k^2\Xi^{-1/2}X\Xi^{-1/2}] && (157)
\end{aligned}$$

where we have repeatedly used properties of the determinant. Furthermore

$$\begin{aligned}
\det[I + k^2\Xi^{-1/2}X\Xi^{-1/2}] &= \det[I + k^2\Xi^{-1}X] && (158)\\
&= \det[I + k^2 X^{1/2}\Xi X^{1/2}] && (159)\\
&= \det[I + k^2 Y^2/2] \,, && (160)
\end{aligned}$$

so that we have

$$\det[H(k^2)] = \det[\Omega^2 + C_0^2 k^2]\;\det[I + k^2 Y^2/2] \,. \tag{161}$$

Note the the matrix Y^2 is the "healing length matrix" we had previously defined, and that the net result of this analysis is that the full determinant is the product of the determinant previously found in the hydrodynamic limit with a factor that depends on the product of wavenumber and healing length.

But now, given our formula (255) for the determinant, we see

$$\begin{aligned}
\det'[H(k^2)] = {} & (-\,\mathrm{Tr}(\Omega^2\bar{C}_0^2) + 2k^2\det[C_0^2])\;\det[I + k^2 Y^2/2] \\
& +\;\det[\Omega^2 + C_0^2 k^2]\;(-\,\mathrm{Tr}[\bar{Y}^2] + k^2\det[Y^2])/2 \,, \qquad (162)
\end{aligned}$$

whence

$$\det'[H(k^2)]_{k\to 0} = -\,\mathrm{Tr}(\Omega^2\bar{C}_0^2) \,, \tag{163}$$

and so

$$\left.\frac{\mathrm{d}\omega^2}{\mathrm{d}k^2}\right|_{k\to 0} = \frac{1}{2}\left[\mathrm{Tr}[C^2] \pm \frac{\mathrm{Tr}[\Omega^2]\,\mathrm{Tr}[C^2] + 2\,\mathrm{Tr}(\Omega^2\bar{C}_0^2)}{\mathrm{Tr}[\Omega^2]}\right] . \tag{164}$$

That is:

$$\left.\frac{\mathrm{d}\omega^2}{\mathrm{d}k^2}\right|_{k\to 0} = \frac{1}{2}\left[\mathrm{Tr}[C^2] \pm \left\{\mathrm{Tr}[C^2] + 2\frac{\mathrm{Tr}(\Omega^2\bar{C}_0^2)}{\mathrm{Tr}[\Omega^2]}\right\}\right] . \tag{165}$$

Note that all the relevant matrices have been carefully symmetrized. Also note the important distinction between C_0^2 and C^2. Now define

$$c^2 = \frac{1}{2}\,\mathrm{Tr}[C^2] , \tag{166}$$

then

$$\left.\frac{\mathrm{d}\omega^2}{\mathrm{d}k^2}\right|_{k\to 0} = c^2(1 \pm \eta_2) , \tag{167}$$

with

$$\eta_2 = \left\{\frac{\mathrm{Tr}[C^2]\,\mathrm{Tr}[\Omega^2] + 2\,\mathrm{Tr}(\Omega^2\bar{C}_0^2)}{\mathrm{Tr}[\Omega^2]\,\mathrm{Tr}[C^2]}\right\} = \left\{1 + \frac{\mathrm{Tr}(\Omega^2\bar{C}_0^2)}{\omega_{II}^2\, c^2}\right\} . \tag{168}$$

Similarly, consider the second derivative:

$$\begin{aligned}\frac{\mathrm{d}^2\omega^2}{\mathrm{d}(k^2)^2} = \frac{1}{2}\Bigg[&\mathrm{Tr}[H_s''(k^2)] \\ &\pm\frac{\mathrm{Tr}[H_s(k^2)]\,\mathrm{Tr}[H_s''(k^2)] + \mathrm{Tr}[H_s'(k^2)]\,\mathrm{Tr}[H_s'(k^2)] - 2\det{}''[H(k^2)]}{\sqrt{\mathrm{Tr}[H_s(k^2)]^2 - 4\det[H(k^2)]}} \\ &\mp\frac{(\mathrm{Tr}[H_s(k^2)]\,\mathrm{Tr}[H_s'(k^2)] - 2\det{}'[H(k^2)])^2}{(\mathrm{Tr}[H_s(k^2)]^2 - 4\det[H(k^2)])^{3/2}}\Bigg] ,\end{aligned} \tag{169}$$

whence

$$\begin{aligned}\left.\frac{\mathrm{d}^2\omega^2}{\mathrm{d}(k^2)^2}\right|_{k\to 0} = \frac{1}{2}\Bigg[&\mathrm{Tr}[Z^2] \pm \frac{\mathrm{Tr}[\Omega^2]\,\mathrm{Tr}[Z^2] + \mathrm{Tr}[C^2]^2 - 2\det{}''[H(k^2)]_{k\to 0}}{\mathrm{Tr}[\Omega^2]} \\ &\mp\frac{(\mathrm{Tr}[\Omega^2]\,\mathrm{Tr}[C^2] - 2\det{}'[H(k^2)]_{k\to 0})^2}{\mathrm{Tr}[\Omega^2]^3}\Bigg].\end{aligned} \tag{170}$$

The last term above can be related to $\mathrm{d}\omega^2/\mathrm{d}k^2$, while the determinant piece is evaluated using

$$\begin{aligned}\det{}''[H(k^2)] = {}&(2\det[C_0^2])\ \det[I + k^2Y^2/2] \\ &+(-\,\mathrm{Tr}(\Omega^2\bar{C}_0^2) + 2k^2\det[C_0^2])\ (-\,\mathrm{Tr}[\bar{Y}^2] + k^2\det[Y^2])/2 \\ &+\det[\Omega^2 + C_0^2k^2]\ (\det[Y^2]/2) \\ &+(-\,\mathrm{Tr}(\Omega^2\bar{C}_0^2) + 2k^2\det[C_0^2])\ (-\,\mathrm{Tr}[\bar{Y}^2] + k^2\det[Y^2])/2 .\end{aligned} \tag{171}$$

Therefore

$$\begin{aligned}\det{}''[H(k^2)]_{k\to 0} = {} & (2\det[C_0^2]) \\ & +(-\operatorname{Tr}(\Omega^2\bar{C}_0^2))\,(-\operatorname{Tr}[\bar{Y}^2])/2 + \det[\Omega^2]\,(\det[Y^2])/2 \\ & +(-\operatorname{Tr}(\Omega^2\bar{C}_0^2))\,(-\operatorname{Tr}[\bar{Y}^2])/2\,. \end{aligned} \tag{172}$$

That is, (recalling $\operatorname{Tr}[\bar{A}] = -\operatorname{Tr}[A]$),

$$\det{}''[H(k^2)]_{k\to 0} = (2\det[C_0^2]) - (\operatorname{Tr}(\Omega^2\bar{C}_0^2))\,(\operatorname{Tr}[Y^2])\,, \tag{173}$$

or

$$\det{}''[H(k^2)]_{k\to 0} = -\operatorname{Tr}[C_0^2\bar{C}_0^2] - \operatorname{Tr}[\Omega^2\bar{C}_0^2]\;\operatorname{Tr}[Y^2]\,. \tag{174}$$

Now assembling all the pieces, a little algebra yields

$$\begin{aligned}\left.\frac{\mathrm{d}^2\omega^2}{\mathrm{d}(k^2)^2}\right|_{k\to 0} = \frac{1}{2}\Bigg[& \operatorname{Tr}[Z^2] \pm \operatorname{Tr}[Z^2] \pm 2\frac{\operatorname{Tr}[\Omega^2\bar{C}_0^2]}{\operatorname{Tr}[\Omega^2]}\operatorname{Tr}[Y^2] \\ & \pm \frac{\operatorname{Tr}[C^2]^2 - 4\det[C_0^2]}{\operatorname{Tr}[\Omega^2]} \mp \frac{\operatorname{Tr}[C^2]^2}{\operatorname{Tr}[\Omega^2]}\eta_2^2\Bigg]. \end{aligned} \tag{175}$$

With the above formula we have completed our derivation of the lowest-order terms of the generic dispersion relation of a coupled 2-BEC system – including the terms introduced by the quantum potential at high wavenumber – up to terms of order k^4. From the above formula it is clear that we do not generically have Lorentz invariance in this system: Lorentz violations arise both due to mode-mixing interactions (an effect which can persist in the hydrodynamic limit where $Z \to 0$ and $Y \to 0$) and to the presence of the quantum potential (signaled by $Z \neq 0$ and $Y \neq 0$). While the mode-mixing effects are relevant at all energies the latter effect characterizes the discrete structure of the effective space-time at high energies. It is in this sense that the quantum potential determines the analogue of quantum gravity effects in our 2-BEC system.

4.4 The Relevance for Quantum Gravity Phenomenology

Following this physical insight we can now easily identify a regime that is potentially relevant for simulating the typical ansätze of quantum gravity phenomenology. We demand that any violation of Lorentz invariance present should be due to the microscopic structure of the effective space-time. This implies that one has to tune the system in order to cancel exactly all those violations of Lorentz invariance which are solely due to mode-mixing interactions in the hydrodynamic limit.

We basically follow the guiding idea that a good analogue of quantum-gravity-induced Lorentz violations should be characterized only by the ultraviolet physics of the effective space-time. In the system at hand the ultraviolet

physics is indeed characterized by the quantum potential, whereas possible violations of the Lorentz invariance in the hydrodynamical limit are low energy effects, even though they have their origin in the microscopic interactions. We therefore start by investigating the scenario in which the system is tuned in such a way that no violations of Lorentz invariance are present in the hydrodynamic limit. This leads us to again enforce the conditions $C1$ and $C2$ which corresponded to "mono-metricity" in the hydrodynamic limit.

In this case (165) and (175) take respectively the form

$$\left.\frac{\mathrm{d}\omega^2}{\mathrm{d}k^2}\right|_{k\to 0} = \frac{1}{2}\left[\mathrm{Tr}[C_0^2] + (1\pm 1)\ \mathrm{Tr}[\Delta C^2]\right] = c_0^2 + \frac{1\pm 1}{2}\,\mathrm{Tr}[\Delta C^2]\,, \quad (176)$$

and

$$\begin{aligned}\left.\frac{\mathrm{d}^2\omega^2}{\mathrm{d}(k^2)^2}\right|_{k\to 0} &= \frac{\mathrm{Tr}[Z^2]\pm\mathrm{Tr}[Z^2]}{2} \mp \mathrm{Tr}[C_0^2]\,\mathrm{Tr}[Y^2] \\ &\quad \pm\frac{1}{2}\frac{\mathrm{Tr}[\Delta C^2]^2 + 2\,\mathrm{Tr}[C_0^2]\,\mathrm{Tr}[\Delta C^2]}{\mathrm{Tr}[\Omega^2]} \mp \frac{1}{2}\frac{\mathrm{Tr}[\Delta C^2]^2}{\mathrm{Tr}[\Omega^2]} \\ &= \frac{\mathrm{Tr}[Z^2]\pm\mathrm{Tr}[Z^2]}{2} \pm \mathrm{Tr}[C_0^2]\left(-\mathrm{Tr}[Y^2] + \frac{\mathrm{Tr}[\Delta C^2]}{\mathrm{Tr}[\Omega^2]}\right)\,. \quad (177)\end{aligned}$$

Recall (see Sect. 4.2) that the first of the physical conditions $C1$ is equivalent to the statement that the 2×2 matrix C_0^2 has two identical eigenvalues. But since C_0^2 is symmetric this then implies $C_0^2 = c_0^2\ \mathbf{I}$, in which case the second condition is automatically satisfied. This also leads to the useful facts

$$\tilde{U}_{\mathrm{AB}} = 0 \quad\Longrightarrow\quad \lambda = -2\sqrt{\rho_{\mathrm{A}0}\,\rho_{\mathrm{B}0}}\,U_{\mathrm{AB}}\,; \quad (178)$$

$$c_0^2 = \frac{\tilde{U}_{\mathrm{AA}}\,\rho_{\mathrm{A}0}}{m_{\mathrm{A}}} = \frac{\tilde{U}_{\mathrm{BB}}\,\rho_{\mathrm{B}0}}{m_{\mathrm{B}}}\,. \quad (179)$$

Now that we have the fine tuning condition for the laser coupling we can compute the magnitude of the effective mass of the massive phonon and determine the values of the Lorentz violation coefficients. In particular we shall start checking that this regime allows for a real positive effective mass as needed for a suitable analogue model of quantum gravity phenomenology.

Effective Mass

Remember that the definition of m_{II} reads

$$m_{II}^2 = \hbar^2\omega_{II}^2/c_0^4\,. \quad (180)$$

Using equation (178) and equation (179) we can rewrite c_0^2 in the following form

$$c_0^2 = [m_{\mathrm{B}}\rho_{\mathrm{A}0}U_{\mathrm{AA}} + m_{\mathrm{A}}\rho_{\mathrm{B}0}U_{\mathrm{BB}} + U_{\mathrm{AB}}(\rho_{\mathrm{A}0}m_{\mathrm{A}} + \rho_{\mathrm{B}0}m_{\mathrm{B}})]/(2m_{\mathrm{A}}m_{\mathrm{B}})\,. \quad (181)$$

Similarly equation (178) and equation (179) when inserted in equation (135) give

$$\omega_{II}^2 = \frac{4U_{\rm AB}(\rho_{\rm A0}m_{\rm B} + \rho_{\rm B0}m_{\rm A})c_0^2}{\hbar^2} \,. \tag{182}$$

We can now estimate m_{II} by simply inserting the above expressions in equation (180) so that

$$m_{II}^2 = \frac{8U_{\rm AB}(\rho_{\rm A0}m_{\rm A} + \rho_{\rm B0}m_{\rm B})m_{\rm A}m_{\rm B}}{[m_{\rm B}\rho_{\rm A0}U_{\rm AA} + m_{\rm A}\rho_{\rm B0}U_{\rm BB} + U_{\rm AB}(\rho_{\rm A0}m_{\rm A} + \rho_{\rm B0}m_{\rm B})]} \,. \tag{183}$$

This formula is still a little clumsy but a great deal can be understood by doing the physically reasonable approximation $m_{\rm A} \approx m_{\rm B} = m$ and $\rho_{\rm A} \approx \rho_{\rm B}$. In fact in this case one obtains

$$m_{II}^2 \approx m^2 \, \frac{8U_{\rm AB}}{[U_{\rm AA} + 2U_{\rm AB} + U_{\rm BB}]} \,. \tag{184}$$

This formula now shows clearly that, as long as the mixing term $U_{\rm AB}$ is small compared to the "direct" scattering $U_{\rm AA}+U_{\rm BB}$, the mass of the heavy phonon will be "small" compared to the mass of the atoms. Though experimental realizability of the system is not the primary focus of the current article, we point out that there is no obstruction in principle to tuning a 2-BEC system into a regime where $|U_{\rm AB}| \ll |U_{\rm AA} + U_{\rm BB}|$. For the purposes of this paper it is sufficient that a small effective phonon mass (small compared to the atomic masses which set the analogue quantum gravity scale) is obtainable for some arrangement of the microscopic parameters. We can now look separately at the coefficients of the quadratic and quartic Lorentz violations and then compare their relative strength in order to see if a situation like that envisaged by discussions of the naturalness problem is actually realized.

Coefficient of the Quadratic Deviation

One can easily see from (176) that the η_2 coefficients for this case take the form

$$\eta_{2,I} = 0 \,; \tag{185}$$

$$\begin{aligned} \eta_{2,II}\, c_0^2 &= {\rm Tr}[\Delta C^2] = {\rm Tr}[X^{1/2}\Lambda X^{1/2}] = {\rm Tr}[X\Lambda] \\ &= -\frac{1}{2}\frac{\lambda}{m_{\rm A}m_{\rm B}}\left(\frac{m_{\rm A}\rho_{\rm A0} + m_{\rm B}\rho_{\rm B0}}{\sqrt{\rho_{\rm A0}\rho_{\rm B0}}}\right). \end{aligned} \tag{186}$$

So if we insert the fine tuning condition for λ, equation (178), we get

$$\eta_{2,II} = \frac{U_{\rm AB}\,(m_{\rm A}\rho_{\rm A0} + m_{\rm B}\rho_{\rm B0})}{m_{\rm A}m_{\rm B}c_0^2} \,. \tag{187}$$

Remarkably we can now cast this coefficient in a much more suggestive form by expressing the coupling $U_{\rm AB}$ in terms of the mass of the massive quasiparticle m_{II}^2. In order to do this we start from equation (182) and note that it enables us to express $U_{\rm AB}$ in (187) in terms of ω_{II}^2, thereby obtaining

$$\eta_{2,II} = \frac{\hbar^2}{4c_0^4}\,\frac{\rho_{\rm A0}m_{\rm A}+\rho_{\rm B0}m_{\rm B}}{\rho_{\rm A0}m_{\rm B}+\rho_{\rm B0}m_{\rm A}}\,\frac{\omega_{II}^2}{m_{\rm A}m_{\rm B}}\,. \tag{188}$$

Now it is easy to see that

$$\frac{\rho_{\rm A0}m_{\rm A}+\rho_{\rm B0}m_{\rm B}}{\rho_{\rm A0}m_{\rm B}+\rho_{\rm B0}m_{\rm A}} \approx \mathcal{O}(1)\,, \tag{189}$$

and that this factor is identically unity if either $m_{\rm A}=m_{\rm B}$ or $\rho_{\rm A0}=\rho_{\rm B0}$. All together we are left with

$$\eta_{2,II} = \bar{\eta}\left(\frac{m_{II}}{\sqrt{m_{\rm A}m_{\rm B}}}\right)^2, \tag{190}$$

where $\bar{\eta}$ is a dimensionless coefficient of order unity.

The product in the denominator of the above expression can be interpreted as the geometric mean of the fundamental bosons masses $m_{\rm A}$ and $m_{\rm B}$. These are mass scales associated with the microphysics of the condensate – in analogy with our experience with a 1-BEC system where the "quantum gravity scale" is set by the mass of the BEC atoms. It is then natural to define an analogue of the scale of the breakdown of Lorentz invariance as $M_{\rm eff}=\sqrt{m_{\rm A}m_{\rm B}}$. (Indeed this "analogue Lorentz breaking scale" will typically do double duty as an "analogue Planck mass".)

Using this physical insight it should be clear that equation (190) effectively says

$$\eta_{2,II} \approx \left(\frac{m_{II}}{M_{\rm eff}}\right)^2, \tag{191}$$

which, given that $m_I=0$, we are naturally lead to generalize to

$$\eta_{2,X} \approx \left(\frac{m_X}{M_{\rm eff}}\right)^2 = \left(\frac{\text{mass scale of quasiparticle}}{\text{effective Planck scale}}\right)^2; \qquad X=I,II\,. \tag{192}$$

The above relation is exactly the sort of dimensionless ratio $(\mu/M)^\sigma$ that has been very often *conjectured* in the literature on quantum gravity phenomenology in order to explain the strong observational constraints on Lorentz violations at the lowest orders. (See earlier discussion.) Does this now imply that this particular regime of our 2-BEC system will also show an analogue version of the naturalness problem? In order to answer this question we need to find the dimensionless coefficient for the quartic deviations, η_4, and check if it will or won't itself be suppressed by some power of the small ratio $m_{II}/M_{\rm eff}$.

Coefficients of the Quartic Deviation

Let us now consider the coefficients of the quartic term presented in equation (177). For the various terms appearing in (177) we get

$$\mathrm{Tr}[Z^2] = 2\,\mathrm{Tr}[DX] = \frac{\hbar^2}{2}\left(\frac{m_{\mathrm{A}}^2 + m_{\mathrm{B}}^2}{m_{\mathrm{A}}^2 m_{\mathrm{B}}^2}\right) ; \tag{193}$$

$$\mathrm{Tr}[\Delta C^2] = \mathrm{Tr}[X\Lambda] = -\frac{\lambda}{2}\frac{m_{\mathrm{A}}\rho_{\mathrm{A}0} + m_{\mathrm{B}}\rho_{\mathrm{B}0}}{m_{\mathrm{A}} m_{\mathrm{B}}\sqrt{\rho_{\mathrm{A}0}\rho_{\mathrm{B}0}}} = U_{\mathrm{AB}}\frac{m_{\mathrm{A}}\rho_{\mathrm{A}0} + m_{\mathrm{B}}\rho_{\mathrm{B}0}}{m_{\mathrm{A}} m_{\mathrm{B}}} ; \tag{194}$$

$$\mathrm{Tr}[Y^2] = 2\,\mathrm{Tr}[X\Xi^{-1}] = \frac{\hbar^2}{2}\frac{\rho_{\mathrm{A}0} m_{\mathrm{A}}\tilde{U}_{\mathrm{AA}} + \rho_{\mathrm{B}0} m_{\mathrm{B}}\tilde{U}_{\mathrm{BB}}}{\rho_{\mathrm{A}0} m_{\mathrm{A}}\rho_{\mathrm{B}0} m_{\mathrm{B}}\tilde{U}_{\mathrm{AA}}\tilde{U}_{\mathrm{BB}}} ; \tag{195}$$

where in the last expression we have used the fact that in the current scenario $\tilde{U}_{\mathrm{AB}} = 0$. Now by definition

$$\eta_4 = \frac{1}{2}(M_{\mathrm{eff}}^2/\hbar^2)\left[\frac{\mathrm{d}^2\omega^2}{(\mathrm{d}k^2)^2}\right]_{k=0} \tag{196}$$

is the dimensionless coefficient in front of the k^4. So

$$\eta_4 = \frac{M_{\mathrm{eff}}^2}{2\hbar^2}\left[\frac{\mathrm{Tr}[Z^2] \pm \mathrm{Tr}[Z^2]}{2} \pm \mathrm{Tr}[C_0^2]\left(-\frac{\mathrm{Tr}[Y^2]}{2} + \frac{\mathrm{Tr}[\Delta C^2]}{\mathrm{Tr}[\Omega^2]}\right)\right] \tag{197}$$

$$= \frac{M_{\mathrm{eff}}^2\, c_0^2}{\hbar^2}\left[\frac{\mathrm{Tr}[Z^2] \pm \mathrm{Tr}[Z^2]}{2\,\mathrm{Tr}[C_0^2]} \pm \left(-\frac{\mathrm{Tr}[Y^2]}{2} + \frac{\mathrm{Tr}[\Delta C^2]}{\mathrm{Tr}[\Omega^2]}\right)\right] . \tag{198}$$

Whence

$$\eta_{4,I} = \frac{M_{\mathrm{eff}}^2\, c_0^2}{\hbar^2}\left[\frac{\mathrm{Tr}[Z^2]}{\mathrm{Tr}[C_0^2]} + \left(-\frac{\mathrm{Tr}[Y^2]}{2} + \frac{\mathrm{Tr}[\Delta C^2]}{\mathrm{Tr}[\Omega^2]}\right)\right] ; \tag{199}$$

$$\eta_{4,II} = \frac{M_{\mathrm{eff}}^2\, c_0^2}{\hbar^2}\left[\left(\frac{\mathrm{Tr}[Y^2]}{2} - \frac{\mathrm{Tr}[\Delta C^2]}{\mathrm{Tr}[\Omega^2]}\right)\right] . \tag{200}$$

Let us compute the two relevant terms separately:

$$\frac{\mathrm{Tr}[Z^2]}{\mathrm{Tr}[C_0^2]} = \frac{\hbar^2}{4c_0^2}\left(\frac{m_{\mathrm{A}}^2 + m_{\mathrm{B}}^2}{m_{\mathrm{A}}^2 m_{\mathrm{B}}^2}\right) = \frac{\hbar^2}{4c_0^2 M_{\mathrm{eff}}^2}\left(\frac{m_{\mathrm{A}}^2 + m_{\mathrm{B}}^2}{m_{\mathrm{A}} m_{\mathrm{B}}}\right) ; \tag{201}$$

$$-\mathrm{Tr}[Y^2]/2 + \frac{\mathrm{Tr}[\Delta C^2]}{\mathrm{Tr}[\Omega^2]} = -\frac{\hbar^2}{4M_{\mathrm{eff}}^2}\left[\frac{\rho_{\mathrm{A}0} m_{\mathrm{A}}\tilde{U}_{\mathrm{AA}}^2 + \rho_{\mathrm{B}0} m_{\mathrm{B}}\tilde{U}_{\mathrm{BB}}^2}{\rho_{\mathrm{A}0}\rho_{\mathrm{B}0}\tilde{U}_{\mathrm{AA}}\tilde{U}_{\mathrm{BB}}\left(\tilde{U}_{\mathrm{AA}} + \tilde{U}_{\mathrm{BB}}\right)}\right]$$

$$= -\frac{\hbar^2}{4M_{\mathrm{eff}}^2\, c_0^2}\left[\frac{m_{\mathrm{A}}^2\tilde{U}_{\mathrm{AA}} + m_{\mathrm{B}}^2\tilde{U}_{\mathrm{BB}}}{m_{\mathrm{A}} m_{\mathrm{B}}\left(\tilde{U}_{\mathrm{AA}} + \tilde{U}_{\mathrm{BB}}\right)}\right] ; \tag{202}$$

where we have used $\rho_{X0}\tilde{U}_{XX} = m_X c_0^2$ for $X = A, B$ as in equation (179). Note that the quantity in square brackets in the last line is dimensionless. So in the end:

$$\eta_{4,I} = \frac{1}{4}\left[\left(\frac{m_{\rm A}^2 + m_{\rm B}^2}{m_{\rm A} m_{\rm B}}\right) - \frac{m_{\rm A}^2\tilde{U}_{\rm AA} + m_{\rm B}^2\tilde{U}_{\rm BB}}{m_{\rm A} m_{\rm B}\left(\tilde{U}_{\rm AA} + \tilde{U}_{\rm BB}\right)}\right] \tag{203}$$

$$= \frac{1}{4}\left[\frac{m_{\rm A}^2\tilde{U}_{\rm BB} + m_{\rm B}^2\tilde{U}_{\rm AA}}{m_{\rm A} m_{\rm B}\left(\tilde{U}_{\rm AA} + \tilde{U}_{\rm BB}\right)}\right]; \tag{204}$$

$$\eta_{4,II} = \frac{1}{4}\left[\frac{m_{\rm A}^2\tilde{U}_{\rm AA} + m_{\rm B}^2\tilde{U}_{\rm BB}}{m_{\rm A} m_{\rm B}\left(\tilde{U}_{\rm AA} + \tilde{U}_{\rm BB}\right)}\right]. \tag{205}$$

Note: In the special case $m_{\rm A} = m_{\rm B}$ we recover identical quartic deviations $\eta_{4,I} = \eta_{4,II} = 1/4$, indicating in this special situation a "universal" deviation from Lorentz invariance. Indeed we also obtain $\eta_{4,I} = \eta_{4,II}$ if we demand $\tilde{U}_{\rm AA} = \tilde{U}_{\rm BB}$, even without fixing $m_{\rm A} = m_{\rm B}$.

Thus in the analogue space-time we have developed the issue of *universality* is fundamentally related to the complexity of the underlying microscopic system. As long as we keep the two atomic masses $m_{\rm A}$ and $m_{\rm B}$ distinct we generically have distinct η_4 coefficients (and the η_2 coefficients are unequal even in the case $m_{\rm A} = m_{\rm B}$). However we can easily recover identical η_4 coefficients, for instance, as soon as we impose identical microphysics for the two BEC systems we couple.

Avoidance of the Naturalness Problem

We can now ask ourselves if there is, or is not, a naturalness problem present in our system. Are the dimensionless coefficients $\eta_{4,I/II}$ suppressed below their naive values by some small ratio involving $M_{\rm eff} = \sqrt{m_{\rm A} m_{\rm B}}$? Or are these ratios unsuppressed? Indeed at first sight it might seem that further suppression is the case, since the square of the "effective Planck scale" seems to appear in the denominator of both the coefficients (204) and (205). However, the squares of the atomic masses also appear in the numerator, rendering both coefficients of order unity.

It is perhaps easier to see this once the dependence of (204) and (205) on the effective coupling $\tilde{U}$ is removed. We again use the substitution $\tilde{U}_{XX} = m_X c_0^2/\rho_{X0}$ for $X = {\rm A}, {\rm B}$, so obtaining:

$$\eta_{4,I} = \frac{1}{4}\left[\frac{m_{\rm A}\rho_{\rm A0} + m_{\rm B}\rho_{\rm B0}}{m_{\rm A}\rho_{\rm B0} + m_{\rm B}\rho_{\rm A0}}\right]; \tag{206}$$

$$\eta_{4,II} = \frac{1}{4}\left[\frac{m_{\rm A}^3\rho_{\rm B0} + m_{\rm B}^3\rho_{\rm A0}}{m_{\rm A} m_{\rm B}\left(m_{\rm A}\rho_{\rm B0} + m_{\rm B}\rho_{\rm A0}\right)}\right]. \tag{207}$$

From these expressions is clear that the $\eta_{4,I/II}$ coefficients are actually of order unity.

That is, if our system is set up so that $m_{II} \ll m_{\mathrm{A/B}}$ – which we have seen in this scenario is equivalent to requiring $U_{\mathrm{AB}} \ll U_{\mathrm{AA/BB}}$ – no naturalness problem arises as for $p > m_{II}\, c_0$ the higher-order, energy-dependent Lorentz-violating terms ($n \geq 4$) will indeed dominate over the quadratic Lorentz-violating term.

It is quite remarkable that the quadratic coefficients (192) are *exactly* of the form postulated in several works on non-renormalizable EFT with Lorentz invariance violations (see e.g. [26]). They are indeed the squared ratio of the particle mass to the scale of Lorentz violation. Moreover we can see from (204) and (205) that there is no further suppression – after having pulled out a factor $(\hbar/M_{\text{Lorentz violation}})^2$ – for the quartic coefficients $\eta_{4,\mathrm{I/|I}}$. These coefficients are of order one and generically non-universal, (though if desired they can be forced to be universal by additional and specific fine tuning).

The suppression of η_2, combined with the *non-suppression* of η_4, is precisely the statement that the "naturalness problem" does not arise in the current model. We stress this is not a "tree level" result as the dispersion relation was computed directly from the fundamental Hamiltonian and was not derived via any EFT reasoning. Moreover avoidance of the naturalness problem is not directly related to the tuning of our system to reproduce special relativity in the hydrodynamic limit. In fact our conditions for recovering special relativity at low energies do not *a priori* fix the the η_2 coefficient, as its strength after the "fine tuning" could still be large (even of order one) if the typical mass scale of the massive phonon is not well below the atomic mass scale. Instead the smallness of η_2 is directly related to the mass-generating mechanism.

The key question is now: Why does our model escape the naive predictions of dominant lowest-dimension Lorentz violations? (In fact in our model for any $p \gg m_{II}$ the k^4 Lorentz violating term dominates over the order k^2 one.) We here propose a nice interpretation in terms of "emergent symmetry": Non-zero λ *simultaneously* produces a non-zero mass for one of the phonons, *and* a corresponding non-zero Lorentz violation at order k^2. (Single BEC systems have only k^4 Lorentz violations as described by the Bogoliubov dispersion relation.) Let us now drive $\lambda \to 0$, but keep the conditions $C1$ and $C2$ valid at each stage. (This also requires $U_{\mathrm{AB}} \to 0$.) One gets an EFT which at low energies describes two non-interacting phonons propagating on a common background. (In fact $\eta_2 \to 0$ and $c_I = c_{II} = c_0$.) This system possesses a $SO(2)$ symmetry. Non-zero laser coupling λ softly breaks this $SO(2)$, the mass degeneracy, and low-energy Lorentz invariance. Such soft Lorentz violation is then characterized (as usual in EFT) by the ratio of the scale of the symmetry breaking m_{II}, and that of the scale originating the Lorentz violation in first place $M_{\text{Lorentz violation}}$. We stress that the $SO(2)$ symmetry is an "emergent symmetry" as it is not preserved beyond the hydrodynamic limit: the η_4 coefficients are in general different if $m_{\mathrm{A}} \neq m_{\mathrm{B}}$, so $SO(2)$ is generi-

cally broken at high energies. Nevertheless this is enough for the protection of the *lowest*-order Lorentz violating operators. The lesson to be drawn is that emergent symmetries are sufficient to minimize the amount of Lorentz violation in the lowest-dimension operators of the EFT. In this regard, it is intriguing to realise that an interpretation of SUSY as an accidental symmetry has indeed been considered in recent times [33], and that this is done at the cost of renouncing attempts to solve the hierarchy problem in the standard way. It might be that in this sense the smallness of the particle physics mass scales with respect to the Planck scale could be directly related to smallness of Lorentz violations in renormalizable operators of the low-energy effective field theory we live in. We hope to further investigate these issues in future work.

5 Outlook, Summary and Discussion

So where can (and should) we go from here? If 2-component BECs provide such a rich mathematical and physical structure, are 3-component BECs, or general multi-component BECs even better? That depends on what you are trying to do:

- If one wishes to actually *build* such an analogue space-time in the laboratory, and perform actual experiments, then iteration through 1-BEC and 2-BEC systems seems the most promising route in terms of our technological capabilities.
- For n-component BECs we sketch the situation in Fig. 4. The key point is that due to overall translation invariance one again expects to find one massless quasi-particle, with now $n-1$ distinct massive modes. Unfortunately the matrix algebra is now considerably messier – not intrinsically difficult (after all we are only dealing with $n \times n$ matrices in field space) – but extremely tedious. Physical insight remains largely intact, but (except

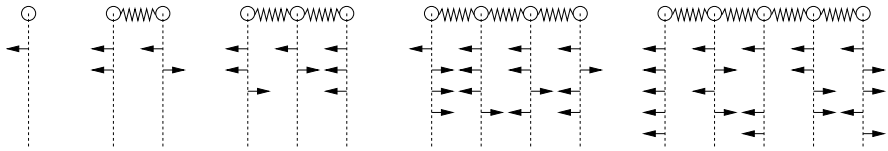

Fig. 4. The figure captures the key features of possible eigenmodes for a small perturbation (*circles*) in a 1 (*left side*), 2, 3, 4, and 5-component (*right side*) BEC. In a 1-component system only one kind of perturbation is allowed, which corresponds to a massless particle propagating through an effective curved space-time, while in the 2-component case two different kinds of mode appear, the one in phase (massless particle) and one in anti-phase (massive particle). For a three-component system we again expect to find one massless particle, when all perturbations are in phase, and now in addition to that two massive particles

in some specific particularly simple cases), computations rapidly become lost in a morass of technical detail.
- However, if one wishes to draw general theoretical lessons from the analogue space-time programme, then multi-component systems are definitely the preferred route – though in this case it is probably better to be even more abstract, and to go beyond the specific details of BEC-based systems to deal with general hyperbolic systems of PDEs.
- In appendix A we have sketched some of the key features of the pseudo–Finsler space-times that naturally emerge from considering the leading symbol of a hyperbolic system of PDEs. While it is clear that much more could be done based on this, and on extending the field theory "normal modes" of [2,3], such an analysis would very much move outside the scope of the COSLAB programme.

In short the 2-BEC system is a good compromise between a system complex enough to exhibit a mass-generating mechanism, and still simple enough to be technologically tractable, with good prospects for laboratory realization of this system in the not too distant future.

The key features we have emphasised in this chapter have been:

- A general analysis of the 2-BEC system to see how perturbations on a 2-BEC background lead to a system of coupled wave equations.
- Extraction of the geometric notion of pseudo–Finsler space-time from this wave equation, coupled with an analysis of how to specialize pseudo–Finsler geometry first to a bi-metric Lorentzian geometry and finally to the usual mono-metric Lorentzian geometry of most direct interest in general relativity and cosmology.
- The mass-generating mechanism we have identified in suitably coupled 2-component BECs is an essential step in making this analogue space-time more realistic; whatever one's views on the ultimate theory of "quantum gravity", any realistic low-energy phenomenology must contain *some* mass-generating mechanism.
- Use of the "quantum pressure" term in the 2-BEC system to mimic the sort of Lorentz violating physics that (based on the relatively young field of "quantum gravity phenomenology") is widely expected to occur at or near the Planck scale.
- Intriguingly, we have seen that in our specific model the mass-generating mechanism interacts with the Lorentz violating mechanism, naturally leading to a situation where the Lorentz violations are suppressed by powers of the quasi-particle mass scale divided by the analogue of the Planck scale.

In summary, while we do not personally believe that the real universe *is* an analogue space-time, we are certainly intrigued by the fact that so much of what is normally viewed as being specific to general relativity and/or particle physics can be placed in this much wider context. We should also be forthright about the key weakness of analogue models as they currently

stand: As we have seen, obtaining an analogue space-time geometry (including space-time curvature) is straightforward – but what is not straightforward is obtaining the Einstein equations. The analogue models are currently analogue models of quantum field theory on curved space-time, but not (yet?) true analogue models of Einstein gravity. Despite this limitation, what can be achieved through the analogue space-time programme is quite impressive, and we expect interest in this field, both theoretical and hopefully experimental, to continue unabated.

Appendix A Finsler and co-Finsler Geometries

Finsler geometries are sufficiently unusual that a brief discussion is in order – especially in view of the fact that the needs of the physics community are often somewhat at odds with what the mathematical community might view as the most important issues. Below are some elementary results, where we emphasise that for the time being we are working with ordinary "Euclidean signature" Finsler geometry. For general references, see [34].

A.1 Basics

Euler theorem: If $H(z)$ is homogeneous of degree n then

$$z^i \, \frac{\partial H(z)}{\partial z^i} = n \, H(z) \, . \tag{208}$$

Finsler function: Defined on the "slit tangent bundle" $T_{\neq 0}(M)$ such that $F : T_{\neq 0}(M) \to [0, +\infty)$ where

$$F(x,t): \quad F(x, \lambda t) = \lambda \, F(x,t) \, , \tag{209}$$

and

$$T_{\neq 0}(M) = \bigcup_{x \in M} \left[T_x - \{\mathbf{0}\} \right] \, . \tag{210}$$

That is, the Finsler function is a defined only for nonzero tangent vectors $t \in [T_x - \{\mathbf{0}\}]$, and for any fixed direction is linear in the size of the vector.
Finsler distance:

$$d_\gamma(x,y) = \int_x^y F(x(\tau), dx/d\tau) \, d\tau \, ; \qquad \tau = \text{ arbitrary parameter} \, . \tag{211}$$

Finsler metric:

$$g_{ij}(x,t) = \frac{1}{2} \frac{\partial^2 [F^2(x,t)]}{\partial t^i \, \partial t^j} \, . \tag{212}$$

The first slightly unusual item is the introduction of co-Finsler structure:
co-Finsler function: Define a co-Finsler structure on the cotangent bundle by Legendre transformation of $F^2(x,t)$. That is:

$$G^2(x,p) = t^j(p)\, p_j - F^2(x,t(p)) \tag{213}$$

where $t(p)$ is defined by the Legendre transformation condition

$$\frac{\partial[F^2]}{\partial t^j}(x,t) = p_j \ . \tag{214}$$

Note

$$\frac{\partial p_j}{\partial t^k} = \frac{\partial[F^2]}{\partial t^j\, \partial t^k} = 2g_{jk}(x,t) \ , \tag{215}$$

which is why we demand the Finsler metric be nonsingular.

Lemma: $G(x,p)$ defined in this way is homogeneous of degree 1.

Proof: Note

$$z^i\, \frac{\partial H(z)}{\partial z^i} = n\, H(z) \tag{216}$$

implies

$$z^i\, \frac{\partial}{\partial z^i} \left[\frac{\partial^m}{(\partial z)^m} H(z) \right] = (n-m)\, \left[\frac{\partial^m}{(\partial z)^m} H(z) \right] \ . \tag{217}$$

In particular:

- F^2 is homogeneous of degree 2.
- g_{ij} is homogeneous of degree 0.
- $\partial[F^2]/\partial t$ is homogeneous of degree 1.
- Therefore $p(t)$ is homogeneous of degree 1
  and $t(p)$ is homogeneous of degree 1.
- Therefore $t(p)p - F^2(t(p))$ is homogeneous of degree 2.
- Therefore $G(p)$ is homogeneous of degree 1.

Thus from a Finsler function $F(x,t)$ we can always construct a co-Finsler function $G(x,p)$ which is homogeneous of degree 1 on the cotangent bundle.

From the way the proof is set up it is clearly reversible – if you are given a co-Finsler function $G(x,p)$ on the cotangent bundle this provides a *natural* way of extracting the corresponding Finsler function:

$$F^2(x,t) = t\, p(t) - G^2(x,p(t)) \ . \tag{218}$$

A.2 Connection with the Quasi-particle PDE Analysis

From the PDE-based analysis we obtain the second-order system of PDEs

$$\partial_a \left(f^{ab}{}_{\rm AB}\, \partial_b \theta^{\rm B} \right) + \text{lower order terms} = 0 \ . \tag{219}$$

We are now generalizing in the obvious manner to any arbitrary number n of interacting BECs, but the analysis is even more general than that – it applies to any field-theory normal-mode analysis that arises from a wide class of Lagrangian based systems [2,3].

Going to the eikonal approximation this becomes

$$f^{ab}_{\rm AB}\, p_a p_b\, \epsilon^{\rm B} + \text{lower-order terms} = 0\ , \tag{220}$$

which leads (neglecting lower order terms for now) to the Fresnel-like equation

$$\det[f^{ab}_{\rm AB}\, p_a p_b] = 0\ . \tag{221}$$

But by expanding the $n \times n$ determinant (n is the number of fields, not the dimension of space-time) we have

$$\det[f^{ab}_{\rm AB}\, p_a p_b] = Q^{abcd\ldots}\, p_a p_b p_c p_d \ldots \tag{222}$$

where if there are n fields there will be $2n$ factors of p.

Now define

$$Q(x,p) = Q^{abcd\ldots}\, p_a p_b p_c p_d \ldots\ , \tag{223}$$

and

$$G(x,p) = \sqrt[2n]{Q(x,p)} = [Q(z,p)]^{1/(2n)}\ , \tag{224}$$

then

- $Q(x,p)$ is homogeneous of degree $2n$.
- $G(x,p)$ is homogeneous of degree 1, and hence is a co-Finsler function.
- We can now Legendre transform $G \to F$, providing a chain

$$Q(x,p) \to G(x,p) \to F(x,t)\ . \tag{225}$$

Can this route be reversed?

Step 1: We can always reverse $F(x,t) \to G(x,p)$ by Legendre transformation.
Step 2: We can always define

$$g^{ab}(x,p) = \frac{1}{2}\frac{\partial}{\partial p_a}\frac{\partial}{\partial p_b}[G(x,p)^2]\ , \tag{226}$$

this is homogeneous of degree 0, but is generically not smooth at $p = 0$.

In fact, *if* $g^{ab}(x,p)$ is smooth at $p = 0$ then there exits a limit

$$g^{ab}(x, p \to 0) = \bar{g}^{ab}(x)\ , \tag{227}$$

but since $g^{ab}(x,p)$ is homogeneous of degree 0 this implies

$$g^{ab}(x,p) = \bar{g}^{ab}(x) \qquad [\forall p]\ , \tag{228}$$

and so the geometry simplifies Finsler $\to$ Riemann.

This observation suggests the following definition.

Definition: A co-Finsler function $G(x,p)$ is $2n$-smooth iff the limit

$$\frac{1}{(2n)!}\lim_{p\to 0}\left(\frac{\partial}{\partial p}\right)^{2n} G(x,p)^{2n} = \bar{Q}^{abcd\ldots} \tag{229}$$

exists independent of the direction p in which you approach zero.

Lemma: If $G(x,p)$ is $2n$-smooth then

$$G(x,p)^{2n} = \bar{Q}^{abcd\dots}\, p_a p_b p_c p_d \dots \,, \tag{230}$$

and indeed

$$G(x,p) = \sqrt[2n]{\bar{Q}^{abcd\dots}\, p_a p_b p_c p_d \dots} \,. \tag{231}$$

Proof: G^{2n} is homogeneous of degree $2n$, so $(\partial/\partial p)^{2n} G^{2n}$ is homogeneous of degree 0. Therefore if the limit

$$\frac{1}{(2n)!} \lim_{p\to 0} \left(\frac{\partial}{\partial p}\right)^{2n} G(x,p)^{2n} = \bar{Q}^{abcd\dots} \tag{232}$$

exists, it follows that

$$\frac{1}{(2n)!} \left(\frac{\partial}{\partial p}\right)^{2n} G(x,p)^{2n} = \bar{Q}^{abcd\dots} \qquad [\forall p] \,, \tag{233}$$

and so the result follows.

Special case $n = 1$*:* If $G(x,p)$ is 2-smooth then

$$\frac{1}{2} \frac{\partial^2}{\partial p^a\, \partial p_b} G(x,p)^2 = \bar{Q}^{ab} = g^{ab}(x,p) \,, \tag{234}$$

and co-Finsler $\to$ Riemann.
These observations have a number of implications:

- For all those co-Finsler functions that are $2n$ smooth we *can* recover the tensor $Q^{abcd\dots}$.
- Not all co-Finsler functions are $2n$ smooth, and for those functions we *cannot* extract $Q^{abcd\dots}$ in any meaningful way.
- But those specific co-Finsler functions that arise from the leading symbol of a 2nd-order system of PDEs are naturally $2n$-smooth, and so for the specific co-Finsler structures we are physically interested in

$$Q(x,p) \leftrightarrow G(x,p) \leftrightarrow F(x,t) \,. \tag{235}$$

- Therefore, in the physically interesting case the Finsler function $F(x,t)$ encodes all the information present in $Q^{abcd\dots}$.

Special case $n = 2$*:* For two fields (appropriate for our 2-BEC system), we can follow the chain

$$\mathbf{f}^{ab} \to Q(x,p) \leftrightarrow G(x,p) \leftrightarrow F(x,t) \tag{236}$$

to formally write

$$ds^4 = g_{abcd}\, dx^a dx^b dx^c dx^d \,, \tag{237}$$

or

$$ds = \sqrt[4]{g_{abcd}\, dx^a dx^b dx^c dx^d} \,. \tag{238}$$

This is one of the "more general" cases Riemann alludes to in his inaugural lecture of 1854 [21].

This discussion makes it clear that the general geometry in our 2-BEC system is a 4-smooth Finsler geometry. It is only for certain special cases that the Finsler geometry specializes first to "multi-metric" and then to "mono-metric" Riemannian geometries.

A.3 Lorentzian Signature Finsler Geometries

The distinction between Finsler and pseudo–Finsler geometries has to do with the distinction between elliptic and hyperbolic PDEs. Elliptic PDEs lead to ordinary Finsler geometries, hyperbolic PDEs lead to *pseudo*–Finsler geometries.

Remember that in special relativity we typically define

$$d_\gamma(x,y) = \int_x^y \sqrt{g_{ab}(dx^a/d\tau)(dx^b/d\tau)}d\tau \,, \tag{239}$$

then

- $d_\gamma(x,y) \in I\!R^+$ for spacelike paths;
- $d_\gamma(x,y) = 0$ for null paths;
- $d_\gamma(x,y) \in I\!I^+$ for timelike paths;

The point is that even in special relativity (and by implication in general relativity) "distances" do not have to be real numbers. This is why physicists deal with *pseudo*–Riemannian [Lorentzian] geometries, not (strictly speaking) Riemannian geometries.

To see how this generalizes in a Finsler situation let us first consider a co-Finsler structure that is multi-metric, that is:

$$Q(x,p) = \Pi_{i=1}^n (g_i^{ab} p_a p_b) \,, \tag{240}$$

where each one of these n factors contains a Lorentzian signature matrix and so can pass through zero. Then

$$G(x,p) = \sqrt[2n]{\Pi_{i=1}^n (g_i^{ab} p_a p_b)} \,, \tag{241}$$

and

$$G(x,p) \in \exp\left(\frac{i\pi\ell}{2n}\right) \; I\!R^+ \,, \tag{242}$$

where

- $\ell = 0 \rightarrow G(x,p) \in I\!R^+ \rightarrow$ outside all n signal cones;
- $\ell = n \rightarrow G(x,p) \in I\!I^+ \rightarrow$ inside all n signal cones.

So we can now define

- Spacelike $\leftrightarrow$ outside all n signal cones $\leftrightarrow$ G real;
- Null $\leftrightarrow$ on any one of the n signal cones $\leftrightarrow$ G zero;
- Timelike $\leftrightarrow$ inside all n signal cones $\leftrightarrow$ G imaginary;
- plus the various "intermediate" cases:

$$\text{"intermediate"} \leftrightarrow \text{inside } \ell \text{ of } n \text{ signal cones} \leftrightarrow G \in i^{\ell/n} \times I\!R^{+} \,. \quad (243)$$

Now this basic idea survives even if we do not have a multi-metric theory. The condition $Q(x,p) = 0$ defines a polynomial of degree $2n$, and so defines n nested sheets (possibly crossing in places). Compare with Courant and Hilbert's discussion of the Monge cone [20].

That is:

$$\begin{aligned} Q(x,p) = 0 &\Leftrightarrow Q(x,(E,\boldsymbol{p})) = 0; \\ &\Leftrightarrow \text{polynomial of degree } 2n \text{ in } E \text{ for any fixed } \boldsymbol{p}; \\ &\Leftrightarrow \text{in each direction } \exists\ 2n \text{ roots in } E; \\ &\Leftrightarrow \text{corresponds to } n \text{ [topological] cones.} \end{aligned}$$

(These are topological cones, not geometrical cones, and the roots might happen to be degenerate.)

Question: Should we be worried by the fact that the co-metric g^{ab} is singular on the signal cone? (In fact on all n of the signal cones.) Not really. We have

$$G(x,p) = \sqrt[2n]{\bar{Q}^{abcd\dots}\, p_a p_b p_c p_d \dots}\,, \quad (244)$$

so

$$g^{ab}(x,p) = \frac{1}{2}\frac{\partial^2}{\partial p^a\, \partial p_b}\left(\sqrt[n]{Q(x,p)}\right) = \frac{1}{2n}\frac{\partial}{\partial p_b}\left\{Q^{\frac{1}{n}-1}\, Q^{abcd\dots}\, p_b p_c p_d \dots\right\}\,, \quad (245)$$

whence

$$\begin{aligned} g^{ab}(x,p) = &\frac{1}{2n} Q^{\frac{1}{n}-1}\, Q^{abcd\dots}\, p_c p_d \dots \quad (246) \\ &+\frac{1}{2n}\left(\frac{1}{n}-1\right) Q^{\frac{1}{n}-2}\left[Q^{acde\dots}\, p_c p_d p_e \dots\right]\, \left[Q^{bfgh\dots}\, p_f p_g p_h \dots\right], \end{aligned}$$

which we can write as

$$\begin{aligned} g^{ab}(x,p) = &\frac{1}{2n} Q^{-(n-1)/n}\, Q^{abcd\dots}\, p_c p_d \dots \quad (247) \\ &-\frac{1}{2n}\frac{n-1}{n} Q^{-(2n-1)/n}\left[Q^{acde\dots}\, p_c p_d p_e \dots\right]\, \left[Q^{bfgh\dots}\, p_f p_g p_h \dots\right]\,. \end{aligned}$$

Yes, this naively looks like it's singular on the signal cone where $Q(x,p) = 0$. But no, this is not a problem: Consider

$$g^{ab}p_a p_b = \frac{1}{2n} Q^{-(2n-1)/n} Q - \frac{1}{2n}\frac{n-1}{n} Q^{-(2n-1)/n} Q^2 \,, \tag{248}$$

then

$$g^{ab}p_a p_b = \frac{1}{2n}\left(1 - \frac{n-1}{n}\right) Q^{1/n} = \frac{1}{2n^2}\, Q^{1/n} = 0 \,, \tag{249}$$

and this quantity is definitely non-singular.

A.4 Summary

In short:

- pseudo–Finsler functions arise naturally from the leading symbol of hyperbolic systems of PDEs;
- pseudo–Finsler geometries provide the natural "geometric" interpretation of a multi-component PDE before fine tuning;
- In particular the natural geometric interpretation of our 2-BEC model (before fine tuning) is as a 4-smooth pseudo–Finsler geometry.

Appendix B Some Matrix Identities

To simplify the flow of argument in the body of the paper, here we collect a few basic results on 2×2 matrices that are used in our analysis.

B.1 Determinants

Theorem: For any two 2×2 matrix A:

$$\det(A) = \frac{1}{2}\left\{\mathrm{Tr}[A]^2 - \mathrm{Tr}[A^2]\right\}. \tag{250}$$

This is best proved by simply noting

$$\det(A) = \lambda_1\lambda_2 = \frac{1}{2}\left[(\lambda_1+\lambda_2)^2 - (\lambda_1^2+\lambda_2^2)\right] = \frac{1}{2}\left\{\mathrm{Tr}[A]^2 - \mathrm{Tr}[A^2]\right\}. \tag{251}$$

If we now define 2×2 "trace reversal" (in a manner reminiscent of standard GR) by

$$\bar{A} = A - \mathrm{Tr}[A]\,\mathbf{I}; \qquad \bar{\bar{A}} = A\,; \tag{252}$$

then this looks even simpler

$$\det(A) = -\frac{1}{2}\,\mathrm{Tr}[A\;\bar{A}] = \det(\bar{A})\,. \tag{253}$$

A simple implication is now:

Theorem: For any two 2×2 matrices A and B:

$$\det(A+\lambda\,B) = \det(A) + \lambda\ \{\mathrm{Tr}[A]\,\mathrm{Tr}[B] - \mathrm{Tr}[A\,B]\} + \lambda^2\ \det(B)\ . \quad (254)$$

which we can also write as

$$\det(A+\lambda\,B) = \det(A) - \lambda\ \mathrm{Tr}[A\,\bar{B}] + \lambda^2\ \det(B)\ . \quad (255)$$

Note that $\mathrm{Tr}[A\,\bar{B}] = \mathrm{Tr}[\bar{A}\,B]$.

B.2 Hamilton–Cayley Theorems

Theorem: For any two 2×2 matrix A:

$$A^{-1} = \frac{\mathrm{Tr}[A]\ \mathbf{I} - A}{\det[A]} = -\frac{\bar{A}}{\det[\bar{A}]}\ . \quad (256)$$

Theorem: For any two 2×2 matrix A:

$$A^{1/2} = \pm\left\{\frac{A \pm \sqrt{\det A}\ \mathbf{I}}{\sqrt{\mathrm{Tr}[A] \pm 2\sqrt{\det A}}}\right\}\ . \quad (257)$$

References

1. C. Barceló, S. Liberati and M. Visser, "Analogue gravity", Living Rev. Rel. **8** 12 (2005), [arXiv:gr-qc/0505065]
2. C. Barceló, S. Liberati and M. Visser, "Analog gravity from field theory normal modes?", Class. Quant. Grav. **18**, 3595 (2001), [arXiv:gr-qc/0104001]
3. C. Barceló, S. Liberati and M. Visser, "Refringence, field theory, and normal modes", Class. Quant. Grav. **19**, 2961 (2002), [arXiv:gr-qc/0111059]
4. M. Visser, C. Barceló and S. Liberati, "Bi-refringence versus bi-metricity", [arXiv:gr-qc/0204017]
5. M. Visser and S. Weinfurtner, "Massive phonon modes from a BEC-based analog model", (2004), [arXiv:cond-mat/0409639]
6. M. Visser and S. Weinfurtner, "Massive Klein-Gordon equation from a BEC-based analogue spacetime", Phys. Rev. D **72** 044020 (2005), [arXiv:gr-qc/0506029]
7. S. Liberati, M. Visser and S. Weinfurtner 2006, "Analogue quantum gravity phenomenology from a two-component Bose-Einstein condensate" Class. Quant. Grav. **23** 3129 (2006), [arXiv:gr-qc/0510125]
8. S. Weinfurtner, S. Liberati and M. Visser, "Analogue model for quantum gravity phenomenology", J. Phys. A **39** 6807 (2006), [arXiv:gr-qc/0511105]
9. S. Weinfurtner, S. Liberati and M. Visser, "Modelling Planck-scale Lorentz violation via analogue models", J. Phys. Conf. Ser. **33** 373 (2006), [arXiv:gr-qc/0512127]
10. S. Liberati, M. Visser and S. Weinfurtner, "Naturalness in emergent spacetime", Phys. Rev. Lett. **96** 151301 (2006), [arXiv:gr-qc/0512139]
11. R. Schutzhold, "Dynamical zero-temperature phase transitions and cosmic inflation / deflation", Phys. Rev. Lett. **95** 135703 (2005), [arXiv:quant-ph/0505196]

12. U. R. Fischer and R. Schutzhold, "Quantum simulation of cosmic inflation in two-component Bose–Einstein" Phys. Rev. A **70** 063615 (2004), [arXiv:cond-mat/0406470]
13. S. Weinfurtner, "Analogue model for an expanding universe", General Relativity and Gravitation **37** 9 1549–1554 (2005), [arXiv:gr-qc/0404063]
14. C. Barceló, S. Liberati and M. Visser, "Analogue models for FRW cosmologies", Int. J. Mod. Phys. D **12** 1641 (2003), [arXiv:gr-qc/0305061]
15. S. D. Jenkins and T. A. B. Kennedy, "Dynamic stability of dressed condensate mixtures", Phys. Rev. A **68**, 053607 (2003)
16. M. Trippenbach, K. Góral, K. Rzażewski, B. Malomed, and Y. B. Band, "Structure of binary Bose–Einstein condensates", J. Phys. B **33** 4017 (2000), [arXiv:cond-mat/0008255]
17. Bloch I 2000, "Atomlaser und Phasenkohärenz atomarer Bose–Einstein–Kondensate", (in German), [http://edoc.ub.uni-muenchen.de/archive/00000208/]
18. Jenkins S D and Kennedy T A B "Spin squeezing in a driven Bose-Einstein condensate", Phys. Rev. A **66** 043621 (2002)
19. C. Barceló, S. Liberati and M. Visser, "Analogue gravity from Bose-Einstein condensates", Class. Quant. Grav. **18** 1137 (2001), [arXiv:gr-qc/0011026]
20. R. Courant and D. Hilbert, "Methods of Mathematical Physics", Vol II, Wiley, John and Sons, (1990)
21. B. Riemann, "Ueber die Hypothesen, welche der Geometrie zu Grunde liegen", 1854.
    "On the Hypotheses which lie at the Bases of Geometry", translated by William Kingdon Clifford, Nature, **8**, pp, 14–17, 36, 37
22. P. Finsler, "Uber Kurven und Flachen in allgemeinen Raumen", [Curves and surfaces in general spaces], PhD thesis (1918)
23. U. R. Fischer and R. Schutzhold, "Quantum simulation of cosmic inflation in two-component Bose-Einstein condensates", Phys. Rev. A **70** (2004) 063615 [arXiv:cond-mat/0406470]
24. M. Visser, C. Barceló and S. Liberati, "Acoustics in Bose–Einstein condensates as an example of broken Lorentz symmetry", [arXiv:hep-th/0109033]
25. D. Mattingly, "Modern tests of Lorentz invariance", Living Rev. Rel. **8** 5 (2005), [arXiv:gr-qc/0502097]
26. T. Jacobson, S. Liberati and D. Mattingly, "Lorentz violation at high energy: Concepts, phenomena and astrophysical constraints", Annals Phys. **321** 150 (2006), [arXiv:astro-ph/0505267]
27. R. C. Myers and M. Pospelov, "Experimental challenges for quantum gravity", Phys. Rev. Lett. **90** 211601 (2003), [arXiv:hep-ph/0301124]
28. T. Jacobson, S. Liberati and D. Mattingly, "Threshold effects and Planck scale Lorentz violation: Combined constraints from high energy astrophysics", Phys. Rev. D **67** 124011 (2003), [arXiv:hep-ph/0209264]
    T. Jacobson, S. Liberati and D. Mattingly, "TeV astrophysics constraints on Planck scale Lorentz violation", Phys. Rev. D **66** (2002) 081302 [arXiv:hep-ph/0112207]
29. T. A. Jacobson, S. Liberati, D. Mattingly and F. W. Stecker, "New limits on Planck scale Lorentz violation in QED", Phys. Rev. Lett. **93** (2004) 021101, [arXiv:astro-ph/0309681]

30. T. Jacobson, S. Liberati and D. Mattingly, "A strong astrophysical constraint on the violation of special relativity by quantum gravity", Nature **424** 1019 (2003), [arXiv:astro-ph/0212190]
31. J. Collins, A. Perez, D. Sudarsky, L. Urrutia and H. Vucetich, "Lorentz invariance: An additional fine-tuning problem", Phys. Rev. Lett. **93** 191301 (2004), [arXiv:gr-qc/0403053]
32. G. Amelino-Camelia and T. Piran, "Planck-scale deformation of Lorentz symmetry as a solution to the UHECR and the TeV-gamma paradoxes", Phys. Rev. D **64** 036005 (2001), [arXiv:astro-ph/0008107]
33. H. S. Goh, M. A. Luty and S. P. Ng, "Supersymmetry without supersymmetry", JHEP **0501** 040 (2005), [arXiv:hep-th/0309103]
34. E. Cartan, "Les Espaces de Finsler", Actualites Scientifiques et Industrielles no. 79, Paris, Hermann (1934)
H. Rund, "The Differential geometry of Finsler spaces", Springer (1959).
D. Bao, S. S. Chern and Z. Shen (eds.), "Finsler geometry", A.M.S. Contemporary Mathematics 196 (1996)
D. Bao, S. S. Chern and Z. Shen, "An Introduction to Riemann-Finsler Geometry", Spring-Verlag (2000)
Z. Shen, "Lectures on Finsler Geometry", World Scientific Publishers (2001)

Links. Relating Different Physical Systems Through the Common QFT Algebraic Structure

G. Vitiello

Dipartimento di Fisica "E.R. Caianiello" and INFN, Universitá di Salerno, 84100 Salerno, Italia
vitiello@sa.infn.it

Abstract. In this report I review some aspects of the algebraic structure of QFT related with the doubling of the degrees of freedom of the system under study. I show how such a doubling is related to the characterizing feature of QFT consisting in the existence of infinitely many unitarily inequivalent representations of the canonical (anti-)commutation relations and how this is described by the q-deformed Hopf algebra. I consider several examples, such as the damped harmonic oscillator, the quantum Brownian motion, thermal field theories, squeezed states, classical-to-quantum relation, and show the analogies, or links, among them arising from the common algebraic structure of the q-deformed Hopf algebra.

1 Introduction

Since several years I am pursuing the study of the vacuum structure in quantum field theory (QFT) through a number of physical problems such as boson condensation and the infrared effects in spontaneously broken symmetry gauge theories, coherent domain formation and defect formation, soliton solutions, particle mixing and oscillation, the canonical formalism for quantum dissipation and unstable states, the quantization in curved background, thermal field theories, quantum-to-classical relationship. In this paper I would like to share with the reader the satisfying feeling of a unified view of several distinct physical phenomena emerging from such a study of the QFT vacuum structure. Besides such a pleasant feeling, there is a concrete interest in pointing out the analogies ("links") among these phenomena, which arises since these links provide a great help not only in the formulation of their mathematical description, but also in the understanding of the physics involved in them. Such a "compared study" also reflects back to a deeper understanding of structural aspects of the same QFT formalism.

G. Vitiello: *Links. Relating Different Physical Systems Through the Common QFT Algebraic Structure*, Lect. Notes Phys. **718**, 165–205 (2007)
DOI 10.1007/3-540-70859-6_7

Quite often QFT is presented as an extension of quantum mechanics (QM) to the relativistic domain. Sometimes it is referred to as "second quantization". Of course, the reasons for that come from the historical developments in the formulation of the quantum theory of elementary particle physics and solid state physics. However, a closer view to the formalism of QFT shows that it is not necessarily related with the relativistic domain and it is not simply a "second" quantization recipe subsequent the quantization procedure in QM. For example, the QFT formalism is widely used, with great success, in condensed matter physics, e.g. in the formulation of superconductivity, of ferromagnetism, etc., where typically one does not refer to the relativistic domain. On the other hand, in dealing with fermion fields one cannot rely on the quantization scheme adopted in QM for boson creation and annihilation operators.

As it will appear in the following, QFT is drastically different from QM. The main reason for this resides in the fact that the well known von Neumann theorem, which characterizes in a crucial way the structure of QM [1,2], does not hold in QFT. In QM the von Neumann theorem states that for systems with a finite number of degrees of freedom all the representations of the canonical commutation relations (ccr) are unitarily equivalent. This means that they are physically equivalent; namely, the representations of the ccr are related by unitary operators and, as well known, physical observables are invariant under the action of unitary operators. Their value is therefore the same independently of the representation one chooses to work in. Such a choice is thus completely arbitrary and does not affect the physics one is going to describe. The situation is quite different in QFT where the von Neumann theorem does not hold. Indeed, the hypothesis of finite number of degrees of freedom on which the theorem rests is not satisfied since fields involve by definition infinitely many degrees of freedom. As a consequence, infinitely many unitarily inequivalent (ui) representations of the ccr are allowed to exist [3–5]. The existence of ui representations is thus a characterizing feature of QFT and a full series of physically relevant consequences follows.

One of the aspects I will discuss below is related with the algebraic structure of QFT. I will show that the relevant algebra underlying the QFT formalism is the Hopf algebra, and this underlies the existence of the ui representations. It manifests in the doubling of the system degrees of freedom and its q-deformation bears deep physical meaning. In the first part of the paper, I will start by considering some aspects of the two-slit experiment. This is a typical subject in QM where quantum features fully show up. The discussion turns out to be useful for the subsequent discussion of the q-deformed Hopf algebra structure of QFT [6,7] and it also provides a good example where the quantum-to-classical relation manifests itself.

The q-deformation of the Hopf algebra will be shown to be also related with quantum dissipation and with thermal field theory, where the description of statistical thermal averages of observables in operatorial terms is made possible by exploiting the existence of infinitely many ui representations [5,8–10].

Recognizing that a symplectic structure with classical dynamics is embedded in the space of the ui representations of ccr in QFT [10] leads to show that trajectories (i.e. a sequence of phase transitions) in such a space may satisfy, under convenient conditions, the criteria for chaoticity prescribed by nonlinear classical dynamics. In a figurate way one could say that a *classical blanket* covers the space of the QFT ui representations. Moving on such a blanket describes (phase) transitions among the representations.

The problem of the interplay between "classical and quantum" is indeed another topic on which I will comment on in this paper and I will show that it is intrinsic to the mathematical structure of QFT [10,11]. The phenomenon of decoherence in QM and the related emergence of classicality from the quantum realm is analyzed in detail in the literature [12]. Similarly, although based on different formal and conceptual frame, the emergence of macroscopic ordered patterns and classically behaving structures out of a QFT (not QM!) dynamics via the spontaneous breakdown of symmetry is since long well known [5,13]. Examples of such classically behaving *macroscopic quantum systems* are crystals, ferromagnets, superconductors, superfluids. These are quantum systems not in the trivial sense that they, as all other systems, are made of quantum components, but in the sense that their macroscopic behavior, characterized by the classical (c-number) observable called order parameter, cannot be explained without recourse to the underlying quantum field dynamics.

On the other hand, in recent years the problem of quantization of a classical theory has attracted much attention in gravitation theories and in nonhamiltonian dissipative system theories, where a novel perspective has been proposed [14] according to which the "emergence" of the quantum-like behavior from a classical frame may occur. I will comment in particular on classical deterministic systems with dissipation (information loss) which are found to exhibit quantum behavior under convenient conditions [14–16]. The paper is organized as follows: the doubling the degrees of freedom is discussed in Sect. 2, the two-slit experiment in Subsect. 2.1, unitarily inequivalent representations in QFT in Sect. 3, quantum dissipation in Subsect. 3.1, the thermal connection and the arrow of time in Subsect. 3.2, two-mode squeezed coherent states in Sect. 4, quantum Brownian motion in Sect. 5, the dissipative noncommutative plane in Sect. 6. Thermal field theory in the operatorial formalism (TFD) is presented in Sect. 7. In Sect. 8 the q-deformed Hopf algebra is shown to be a basic feature of QFT. Entropy as a measure of entanglement and the trajectories in the space of the ui representations are discussed in Sects. 9 and 10 respectively. Deterministic dissipative systems are considered in Sect. 11 with respect to the quantization problem. Section 12 is devoted to conclusions. In this paper I have not considered the doubling of the degrees of freedom in inflationary models and in the problem of the quantization of the matter field in a curved background. The interested reader is referred to the papers [17–19].

2 Doubling the Degrees of Freedom

One of the main features underlying the QFT formalism is the doubling of the degrees of freedom of the system under study. Such a doubling is not simply a mathematical tool useful to describe our system. On the contrary, it bears a physical meaning. It also appears to be an essential feature of QM, as I will show in the examples I am going to discuss in this paper.

The standard formalism of the density matrix [20,21] and of the associated Wigner function [22] suggests that one may describe a quantum particle by splitting the single coordinate $x(t)$ into two coordinates $x_+(t)$ (going forward in time) and $x_-(t)$ (going backward in time). Indeed, the standard expression for the Wigner function is [22],

$$W(p,x,t) = \frac{1}{2\pi\hbar}\int \psi^*\left(x-\frac{1}{2}y,t\right)\psi\left(x+\frac{1}{2}y,t\right)e^{-i\frac{py}{\hbar}}dy \ , \tag{1}$$

where

$$x_\pm = x \pm \frac{1}{2}y \ . \tag{2}$$

By employing the Schwinger quantum operator action principle, or recalling the mean value of a quantum operator

$$\begin{aligned}\bar{A}(t) &= (\psi(t)|A|\psi(t)) \\ &= \int\int \psi^*(x_-,t)\,(x_-|A|x_+)\,\psi(x_+,t)dx_+dx_- \\ &= \int\int (x_+|\rho(t)|x_-)(x_-|A|x_+)dx_+dx_- \ ,\end{aligned} \tag{3}$$

one requires the density matrix

$$W(x,y,t) = (x_+|\rho(t)|x_-) = \psi^*(x_-,t)\psi(x_+,t) \tag{4}$$

to follow two copies of the Schrödinger equation: the forward in time motion and the backward in time motion, respectively. These motions are controlled by the two Hamiltonian operators $H_\pm$:

$$i\hbar\frac{\partial\psi(x_\pm,t)}{\partial t} = H_\pm\psi(x_\pm,t) \ , \tag{5}$$

which gives

$$i\hbar\frac{\partial(x_+|\rho(t)|x_-)}{\partial t} = \mathcal{H}\ (x_+|\rho(t)|x_-) \ , \tag{6}$$

where

$$\mathcal{H} = H_+ - H_- \ . \tag{7}$$

Using two copies of the Hamiltonian (i.e. $H_\pm$) operating on the outer product of two Hilbert spaces $\mathcal{F}_+ \otimes \mathcal{F}_-$ has been implicitly required in QM since the

very beginning of the theory. For example, from Eqs. (6), (7) one finds immediately that the eigenvalues of $\mathcal{H}$ are directly the Bohr transition frequencies $\hbar\Omega_{nm} = E_n - E_m$ which was the first clue to the explanation of spectroscopic structure.

The notion that a quantum particle has two coordinates $x_\pm(t)$ moving at the same time is therefore central [23].

In conclusion, the density matrix and the Wigner function *require* the introduction of a "doubled" set of coordinates, $(x_\pm, p_\pm)$ (or (x, p_x) and (y, p_y)).

Let me show how the doubling of the coordinates works in the remarkable example of the two-slit diffraction experiment. Here I will shortly summarize the discussion reported in [24].

2.1 The Two-slit Experiment

In order to derive the diffraction pattern it is required to know the wave function $\psi_0(x)$ of the particle when it "passes through the slits" at time zero. In other words, one searches for the density matrix

$$(x_+|\rho_0|x_-) = \psi_0^*(x_-)\psi_0(x_+)\ . \tag{8}$$

The probability density for the electron to be found at position x at the detector screen at a later time t is written as

$$P(x,t) = (x|\rho(t)|x) = \psi^*(x,t)\psi(x,t) \tag{9}$$

in terms of the solution $\psi(x,t)$ to the free particle Schrödinger equation

$$\psi(x,t) = \Big(\frac{M}{2\pi\hbar i t}\Big)^{1/2} \int_{-\infty}^{\infty} e^{[\frac{i}{\hbar}A(x-x',t)]}\psi_0(x')dx', \tag{10}$$

where

$$A(x-x',t) = \frac{M(x-x')^2}{2t} \tag{11}$$

is the Hamilton-Jacobi action for a classical free particle to move from x' to x in a time t. Eqs. (8)–(11) then imply that

$$P(x,t) = \frac{M}{2\pi\hbar t}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\Big[iM\frac{(x-x_+)^2-(x-x_-)^2}{2\hbar t}\Big]}(x_+|\rho_0|x_-)dx_+dx_-\ . \tag{12}$$

Equation (12) shows that $P(x,t)$ would not oscillate in x, i.e. there would not be the usual quantum diffraction, if $x_+ = x_-$. In Eq. (12), in order to have quantum interference the forward in time action $A(x-x_+,t)$ must be different from the backward in time action $A(x-x_-,t)$: the non-trivial dependence of the density matrix $(x_+|\rho_0|x_-)$ when the electron "passes through the slits" on the difference $(x_+ - x_-)$ crucially determines the quantum nature of the phenomenon.

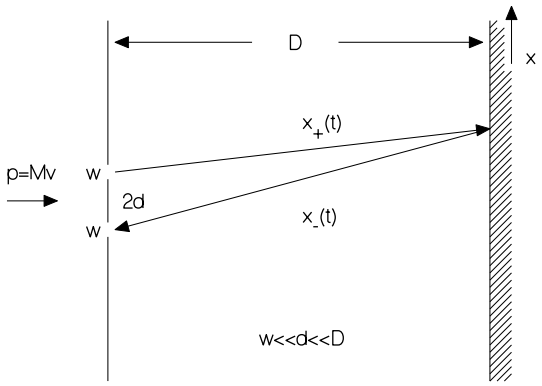

Fig. 1. Two slit experiment

In the quantum diffraction experiment the experimental apparatus is prepared so that $w \ll d \ll D$, with w the opening of the slits which are separated by a distance $2d$. D is the distance between the slits and the screen (Fig. 1). The diffraction pattern is described by $|x| \gg |x_\pm|$. By defining $K = \frac{Mvd}{\hbar D}$, $\beta = \frac{w}{d}$, with $v = D/t$ the velocity of the incident electron, Eq. (12) leads [24] to the usual result

$$P(x, D) \approx \frac{4}{\pi\beta K x^2} \cos^2(Kx) \sin^2(\beta K x) \,, \tag{13}$$

where the initial wave function

$$\psi_0(x) = \frac{1}{\sqrt{2}}\Big[\phi(x-d) + \phi(x+d)\Big], \tag{14}$$

with $\phi(x) = \frac{1}{\sqrt{w}}$ if $|x| \leq \frac{w}{2}$ and zero otherwise, has been used. From Eqs. (8) and (14) we have

$$\begin{aligned}(x_+|\rho_0|x_-) = \; & \frac{1}{2}\{\phi(x_+ - d)\phi(x_- - d) + \phi(x_+ + d)\phi(x_- + d) \\ & + \phi(x_+ - d)\phi(x_- + d) + \phi(x_+ + d)\phi(x_- - d)\} \,. \end{aligned} \tag{15}$$

In the rhs of Eq. (15) the first and the second terms describe the classical processes of the particle going forward and backward in time through slit 1 and going forward and backward in time through slit 2, respectively. In these processes it is $x_+(t) = x_-(t)$ and in such cases no diffraction is observed on the screen. The third term and the fourth term describe the particle going forward in time through slit 1 and backward in time through slit 2, or forward in time through slit 2 and backward in time through slit 1, respectively. These are the terms generating quantum interference since $|x_+(t) - x_-(t)| > 0$.

In conclusion, the doubling the system coordinates, $x(t) \to (x_+(t), x_-(t))$ plays a crucial rôle in the description of the quantum system. If $x(t) \equiv x_+(t) \equiv$

$x_-(t)$, then the system behavior appears to be a classical one. When forward in time and backward in time motions are (at the same time) unequal $x_+(t) \neq x_-(t)$, then the system is behaving in a quantum mechanical fashion and interference patterns appear in measured position probability densities.

I will not comment further on the two-slit experiment. In the following section I go back to the general discussion of the doubling of the degrees of freedom and of its meaning in QFT.

3 Unitarily Inequivalent Representations in QFT

The mathematical rôle and the physical meaning of the doubling of the degrees of freedom fully appears in dealing with phase transitions, with equilibrium and non-equilibrium thermal field theories and with dissipative, open systems. In these cases the doubling of the degrees of freedom appears to be a structural feature of QFT since it strictly relates with the existence of the unitarily inequivalent representations of the ccr in QFT.

Let me consider the case of dissipation [25–27]. I will discuss the canonical quantization of the damped (simple) harmonic oscillator (dho), which is a simple prototype of dissipative system.

3.1 Quantum Dissipation

Dissipation enters into our considerations if there is a coupling to a thermal reservoir yielding a mechanical resistance R. According to the discussion in Sect. 2, the equation of motion for the density matrix is given by Eq. (6), where now the Hamiltonian $\mathcal{H}$ for motion in the (x_+, x_-) plane is [23, 24]

$$\mathcal{H} = \frac{1}{2M}\left(p_+ - \frac{R}{2}\,x_-\right)^2 - \frac{1}{2M}\left(p_- + \frac{R}{2}\,x_+\right)^2 + U(x_+) - U(x_-)\;, \quad (16)$$

where $p_\pm = -i\hbar\frac{\partial}{\partial x_\pm}$. In order to simplify the discussion, it is convenient, without loss of generality, to make an explicit (simple) choice for $U(x_\pm)$, say $U(x_\pm) = \frac{1}{2}\kappa x_\pm^2$. By choosing as doubled coordinates the pair (x, y) with

$$y = x_+ - x_-\;, \quad (17)$$

the Hamiltonian (16) can be derived from the Lagrangian (see [25]– [29])

$$L = M\dot{x}\dot{y} + \frac{1}{2}R(x\dot{y} - \dot{x}y) - \kappa xy\;. \quad (18)$$

The system described by (18) is sometimes called Bateman's dual system [29]. I observe that the doubling imposed by the density matrix and the Wigner function formalism, as seen in Sect. 2, here finds its physical justification in the fact that the canonical quantization scheme can only deal with an isolated

system. In the present case our system has been assumed to be coupled with a thermal reservoir and it is then necessary to *close* the system by including the reservoir. This is achieved by doubling the phase-space dimensions [25, 26]. Equation (18) is indeed the closed system Lagrangian.

By varying Eq. (18) with respect to y gives

$$M\ddot{x} + R\dot{x} + \kappa x = 0 \,, \tag{19}$$

whereas variation with respect to x gives

$$M\ddot{y} - R\dot{y} + \kappa y = 0 \,, \tag{20}$$

which is the *time reversed* ($R \to -R$) of Eq. (19). The physical meaning of the doubled degree of freedom y is now manifest: y may be thought of as describing an effective degree of freedom for the reservoir to which the system (19) is coupled. The canonical momenta are given by $p_x \equiv \frac{\partial L}{\partial \dot{x}} = M\dot{y} - \frac{1}{2}Ry$; $p_y \equiv \frac{\partial L}{\partial \dot{y}} = M\dot{x} + \frac{1}{2}Rx$. For a discussion of Hamiltonian systems of this kind see also [30,31]. Canonical quantization is performed by introducing the commutators

$$[x, p_x] = i\,\hbar = [y, p_y], \quad [x, y] = 0 = [p_x, p_y] \,, \tag{21}$$

and the corresponding sets of annihilation and creation operators

$$\alpha \equiv \left(\frac{1}{2\hbar\Omega}\right)^{\frac{1}{2}} \left(\frac{p_x}{\sqrt{M}} - i\sqrt{M}\Omega x\right), \; \alpha^\dagger \equiv \left(\frac{1}{2\hbar\Omega}\right)^{\frac{1}{2}} \left(\frac{p_x}{\sqrt{M}} + i\sqrt{M}\Omega x\right), \tag{22}$$

$$\beta \equiv \left(\frac{1}{2\hbar\Omega}\right)^{\frac{1}{2}} \left(\frac{p_y}{\sqrt{M}} - i\sqrt{M}\Omega y\right), \; \beta^\dagger \equiv \left(\frac{1}{2\hbar\Omega}\right)^{\frac{1}{2}} \left(\frac{p_y}{\sqrt{M}} + i\sqrt{M}\Omega y\right), \tag{23}$$

$$[\,\alpha, \alpha^\dagger\,] = 1 = [\,\beta, \beta^\dagger\,] \quad , \quad [\,\alpha, \beta\,] = 0 = [\,\alpha, \beta^\dagger\,] \,. \tag{24}$$

I have introduced $\Omega \equiv [\frac{1}{M}(\kappa - \frac{R^2}{4M})]^{\frac{1}{2}}$, the common frequency of the two oscillators Eq. (19) and Eq. (20), assuming Ω real, hence $\kappa > \frac{R^2}{4M}$ (case of no overdamping).

In Sect. 5 I show that, at quantum level, the β modes allow quantum noise effects arising from the imaginary part of the action [23]. Moreover, in Sect. 8 the modes α and β will be shown to be the modes involved in the coproduct operator of the underlying q-deformed Hopf algebra structure. The q-deformation parameter turns out to be a function of R, M and t.

By using the canonical linear transformations $A \equiv \frac{1}{\sqrt{2}}(\alpha + \beta)$, $B \equiv \frac{1}{\sqrt{2}}(\alpha - \beta)$, the quantum Hamiltonian H is then obtained [25, 26] as

$$H = H_0 + H_I \quad , \tag{25}$$

$$H_0 = \hbar\Omega(A^\dagger A - B^\dagger B) \,, \quad H_I = i\hbar\Gamma(A^\dagger B^\dagger - AB) \,, \tag{26}$$

where the decay constant for the classical variable $x(t)$ is denoted by $\Gamma \equiv \frac{R}{2M}$.

In conclusion, *the states generated by $B^\dagger$ represent the sink where the energy dissipated by the quantum damped oscillator flows: the B-oscillator represents the reservoir or heat bath coupled to the A-oscillator.*

The dynamical group structure associated with the system of coupled quantum oscillators is that of $SU(1,1)$. The two mode realization of the algebra $su(1,1)$ is indeed generated by $J_+ = A^\dagger B^\dagger, \quad J_- = J_+^\dagger = AB, \quad J_3 = \frac{1}{2}(A^\dagger A + B^\dagger B + 1), \ [\, J_+, J_- \,] = -2J_3, \quad [\, J_3, J_\pm \,] = \pm J_\pm$. The Casimir operator $\mathcal{C}$ is $\mathcal{C}^2 \equiv \frac{1}{4} + J_3^2 - \frac{1}{2}(J_+ J_- + J_- J_+) = \frac{1}{4}(A^\dagger A - B^\dagger B)^2$.

I also observe that $[\, H_0, H_I \,] = 0$. The time evolution of the vacuum $|0\rangle \equiv |n_A = 0, n_B = 0\rangle = |0\rangle \otimes |0\rangle$, $(A \otimes 1)|0\rangle \otimes |0\rangle \equiv A|0\rangle = 0$; $(1 \otimes B)|0\rangle \otimes |0\rangle \equiv B|0\rangle = 0$, is controlled by H_I

$$
\begin{aligned}
|0(t)\rangle &= \exp\left(-it\frac{H}{\hbar}\right)|0\rangle = \exp\left(-it\frac{H_I}{\hbar}\right)|0\rangle \\
&= \frac{1}{\cosh(\Gamma t)} \exp\left(\tanh(\Gamma t) A^\dagger B^\dagger\right)|0\rangle \ ,
\end{aligned} \tag{27}
$$

$$
\langle 0(t)|0(t)\rangle = 1 \quad \forall t \ , \tag{28}
$$

$$
\lim_{t\to\infty} \langle 0(t)|0\rangle \ \propto \lim_{t\to\infty} \exp(-t\Gamma) = 0 \ . \tag{29}
$$

Notice that once one sets the initial condition of positiveness for the eigenvalues of H_0, such a condition is preserved by the time evolution since H_0 is the Casimir operator (it commutes with H_I). In other words, there is no danger of dealing with energy spectrum unbounded from below. Time evolution for creation and annihilation operators is given by

$$
A \mapsto A(t) = \mathrm{e}^{-i\frac{t}{\hbar}H_I} A \ \mathrm{e}^{i\frac{t}{\hbar}H_I} = A\cosh(\Gamma t) - B^\dagger \sinh(\Gamma t) \ , \tag{30}
$$

$$
B \mapsto B(t) = \mathrm{e}^{-i\frac{t}{\hbar}H_I} B \ \mathrm{e}^{i\frac{t}{\hbar}H_I} = B\cosh(\Gamma t) - A^\dagger \sinh(\Gamma t) \tag{31}
$$

and h.c.. I note that Eqs. (30) and (31) are Bogoliubov transformations: they are canonical transformations preserving the ccr. Equation (29) expresses the instability (decay) of the vacuum under the evolution operator $\exp\left(-it\frac{H_I}{\hbar}\right)$. In other words, time evolution leads out of the Hilbert space of the states. *This means that the QM framework is not suitable for the canonical quantization of the damped harmonic oscillator.* A way out from such a difficulty is provided by QFT [25]: the proper way to perform the canonical quantization of the dho turns out to be working in the framework of QFT. In fact, for many degrees of freedom the time evolution operator $\mathcal{U}(t)$ and the vacuum are formally (at finite volume) given by

$$
\mathcal{U}(t) = \prod_\kappa \exp\left(\Gamma_\kappa t \left(A_\kappa^\dagger B_\kappa^\dagger - A_\kappa B_\kappa\right)\right) \ , \tag{32}
$$

$$
|0(t)\rangle = \prod_\kappa \frac{1}{\cosh(\Gamma_\kappa t)} \exp\left(\tanh(\Gamma_\kappa t) A_\kappa^\dagger B_\kappa^\dagger\right)|0\rangle \ , \tag{33}
$$

with $\langle 0(t)|0(t)\rangle = 1, \forall t$. Using the continuous limit relation $\sum_\kappa \mapsto \frac{V}{(2\pi)^3}\int d^3\kappa$, in the infinite-volume limit we have (for $\int d^3\kappa\ \Gamma_\kappa$ finite and positive)

$$\langle 0(t)|0\rangle \to 0 \ \text{ as } \ V \to \infty \quad \forall\ t \ , \tag{34}$$

and in general, $\langle 0(t)|0(t')\rangle \to 0$ as $V \to \infty \ \ \forall\ t$ and $t',\ t' \neq t$. At each time t a representation $\{|0(t)\rangle\}$ of the ccr is defined and turns out to be ui to any other representation $\{|0(t')\rangle,\ \forall t' \neq t\}$ in the infinite volume limit. In such a way the quantum dho evolves in time through ui representations of ccr (*tunneling*). I remark that $|0(t)\rangle$ is a two-mode time dependent generalized coherent state [32,33]. Also note that

$$\mathcal{N}_{A_\kappa}(t) = \langle 0(t)|A_\kappa^\dagger A_\kappa|0(t)\rangle = \sinh^2\Gamma t \ . \tag{35}$$

The Bogoliubov transformations, Eqs. (30) and (31) can be implemented for every κ as inner automorphism for the algebra $su(1,1)_\kappa$. At each time t one has a copy $\{A_\kappa(t), A_\kappa^\dagger(t), B_\kappa(t), B_\kappa^\dagger(t)\,;\,|0(t)\rangle\,|\,\forall\kappa\}$ of the original algebra induced by the time evolution operator which can thus be thought of as a generator of the group of automorphisms of $\bigoplus_\kappa su(1,1)_\kappa$ parameterized by time t (we have a realization of the operator algebra at each time t, which can be implemented by Gel'fand-Naimark-Segal construction in the C*-algebra formalism [3,34]). Notice that the various copies become unitarily inequivalent in the infinite-volume limit, as shown by Eq. (34): the space of the states splits into ui representations of the ccr each one labeled by time parameter t. As usual, one works at finite volume and only at the end of the computations the limit $V \to \infty$ is performed.

Finally, I note that the "negative" kinematic term in the Hamiltonian (26) (or (16)) also appears in two-dimensional gravity models where, in general, two different strategies are adopted in the quantization procedure [35]: the Schrödinger representation approach, where no negative norm appears, and the string/conformal field theory approach where negative norm states arise as in Gupta-Bleuler electrodynamics.

3.2 The Thermal Connection and the Arrow of Time

It is useful [25] to introduce the functional $\mathcal{F}_A$ for the A-modes

$$\mathcal{F}_A \equiv \langle 0(t)|\left(H_A - \frac{1}{\beta}S_A\right)|0(t)\rangle \ , \tag{36}$$

where β is a non-zero c-number, H_A is the part of H_0 relative to A- modes only, namely $H_A = \sum_\kappa \hbar\Omega_\kappa A_\kappa^\dagger A_\kappa$, and the S_A is given by

$$S_A \equiv -\sum_\kappa \Big\{A_\kappa^\dagger A_\kappa \ln\sinh^2\big(\Gamma_\kappa t\big) - A_\kappa A_\kappa^\dagger \ln\cosh^2\big(\Gamma_\kappa t\big)\Big\} \ . \tag{37}$$

One then considers the extremal condition $\frac{\partial \mathcal{F}_{\rm A}}{\partial \vartheta_\kappa} = 0 \quad \forall \kappa \ , \ \vartheta_\kappa \equiv \Gamma_\kappa t$ to be satisfied in each representation, and using the definition $E_\kappa \equiv \hbar \Omega_\kappa$, one finds

$$\mathcal{N}_{{\rm A}_\kappa}(t) = \sinh^2\left(\Gamma_\kappa t\right) = \frac{1}{{\rm e}^{\beta(t) E_\kappa} - 1} \ , \tag{38}$$

which is the Bose distribution for A_κ at time t, *provided* $\beta(t)$ is the (time-dependent) inverse temperature. Inspection of Eqs. (36) and (37) then suggests that $\mathcal{F}_{\rm A}$ and $S_{\rm A}$ can be interpreted as the *free energy* and the *entropy*, respectively. I will comment more about this in Sects. 7 and 9.

$\{|0(t)\rangle\}$ is thus recognized to be a representation of the ccr at finite temperature (it turns out to be equivalent to the thermo field dynamics (TFD) representation $\{|0(\beta)\rangle\}$ [5,8], see Sect. 7). Use of Eq. (37) shows that

$$\frac{\partial}{\partial t}|0(t)\rangle = -\left(\frac{1}{2}\frac{\partial \mathcal{S}_{\rm A}}{\partial t}\right)|0(t)\rangle \ . \tag{39}$$

One thus see that $i\left(\frac{1}{2}\hbar\frac{\partial \mathcal{S}_{\rm A}}{\partial t}\right)$ is the generator of time translations, namely time evolution is controlled by the entropy variations [36]. It is remarkable that the same dynamical variable $\mathcal{S}_{\rm A}$ whose expectation value is formally the entropy also controls time evolution: damping (or, more generally, dissipation) implies indeed the choice of a privileged direction in time evolution (*arrow of time*) with a consequent breaking of time-reversal invariance.

One may also show that $d\mathcal{F}_{\rm A} = dE_{\rm A} - \frac{1}{\beta} d\mathcal{S}_{\rm A} = 0$, which expresses the first principle of thermodynamics for a system coupled with environment at constant temperature and in absence of mechanical work. As usual, one may define heat as $dQ = \frac{1}{\beta} dS_{\rm A}$ and see that the change in time $d\mathcal{N}_{\rm A}$ of particles condensed in the vacuum turns out into heat dissipation dQ:

$$dE_{\rm A} = \sum_\kappa \hbar \Omega_\kappa \dot{\mathcal{N}}_{{\rm A}_\kappa}(t) dt = \frac{1}{\beta} dS_{\rm A} = dQ \ . \tag{40}$$

Here $\dot{\mathcal{N}}_{{\rm A}_\kappa}$ denotes the time derivative of $\mathcal{N}_{{\rm A}_\kappa}$.

It is interesting to observe that the thermodynamic arrow of time, whose direction is defined by the increasing entropy direction, points in the same direction of the cosmological arrow of time, namely the inflating time direction for the expanding Universe. This can be shown by considering indeed the quantization of inflationary models [17] (see also [18]). The concordance between the two arrows of time (and also with the psychological arrow of time, see Refs. [37]) is not at all granted and is a subject of an ongoing debate (see, e.g., [38]).

In Sect. 6 I will show that quantum dissipation induces a dissipative phase interference [24], analogous to the Aharonov-Bohm phase [39], and a noncommutative geometry in the plane (x_+, x_-) [40].

The quantum dissipation Lagrangian model discussed above is strictly related with the squeezed coherent states in quantum optics and with the quantum Brownian motion. I will briefly discuss these two topics in follolwing sections.

4 Two-mode Squeezed Coherent States

Here I will only mention that in the quantum damped oscillator treatment presented above the time evolution operator $\mathcal{U}(t)$ written in terms of the α and β modes (Eqs. (22) and (23)) is given by

$$\mathcal{U}(t) \equiv \exp\left(-it\frac{H_I}{\hbar}\right) = \prod_\kappa \exp\left(-\frac{\theta_\kappa}{2}\left(\alpha_\kappa^2 - \alpha_\kappa^{\dagger 2}\right)\right)\exp\left(\frac{\theta_\kappa}{2}\left(\beta_\kappa^2 - \beta_\kappa^{\dagger 2}\right)\right)$$
$$\equiv \prod\nolimits_\kappa \hat{S}_\alpha(\theta_\kappa)\hat{S}_\beta(-\theta_\kappa)\ , \tag{41}$$

with $\hat{S}_\alpha(\theta_\kappa) \equiv \exp\left(-\frac{\theta_\kappa}{2}\left({\alpha_\kappa}^2 - {\alpha_\kappa}^{\dagger 2}\right)\right)$ and similar expression for $\hat{S}_\beta(-\theta_\kappa)$ with β and $\beta^\dagger$ replacing α and $\alpha^\dagger$, respectively. The operators $\hat{S}_\alpha(\theta_\kappa)$ and $\hat{S}_\beta(-\theta_\kappa)$ are the squeezing operators for the α_κ and the β_κ modes, respectively, as well known in quantum optics [41]. The set $\theta \equiv \{\theta_\kappa \equiv \Gamma_\kappa t\}$ as well as each θ_κ for all κ is called the squeezing parameter. The state $|0(t)\rangle$ is thus a squeezed coherent states at each time t.

To illustrate the effect of the squeezing, let me focus the attention only on the α_κ modes for sake of definiteness. For the β modes one can proceed in a similar way. As usual, for given κ I express the α mode in terms of conjugate variables of the corresponding oscillator. By using dimensionless quantities I thus write $\alpha = X + iY$, with $[X, Y] = \frac{i}{2}$. The uncertainty relation is $\Delta X \Delta Y = \frac{1}{4}$, with $\Delta X^2 = \Delta Y^2 = \frac{1}{4}$ for (minimum uncertainty) coherent states. The squeezing occurs when $\Delta X^2 < \frac{1}{4}$ and $\Delta Y^2 > \frac{1}{4}$ (or $\Delta X^2 > \frac{1}{4}$ and $\Delta Y^2 < \frac{1}{4}$) in such a way that the uncertainty relation remains unchanged. Under the action of $\mathcal{U}(t)$ the variances ΔX and ΔY are indeed squeezed as

$$\Delta X^2(\theta) = \Delta X^2 \exp(2\theta)\ , \qquad \Delta Y^2(\theta) = \Delta Y^2 \exp(-2\theta)\ . \tag{42}$$

For the tilde-mode similar relations are obtained for the corresponding variances, say $\tilde{X}$ and $\tilde{Y}$:

$$\Delta\tilde{X}^2(\theta) = \Delta\tilde{X}^2 \exp(-2\theta)\ , \qquad \Delta\tilde{Y}^2(\theta) = \Delta\tilde{Y}^2 \exp(2\theta)\ . \tag{43}$$

For positive θ, squeezing then reduces the variances of the Y and $\tilde{X}$ variables, while the variances of the X and $\tilde{Y}$ variables grow by the same amount so to keep the uncertainty relations unchanged. This reflects, in terms of the A and B modes, the constancy of the difference $\mathcal{N}_{\mathrm{A}_\kappa} - \mathcal{N}_{\mathrm{B}_\kappa}$ against separate, but equal, changes of $\mathcal{N}_{\mathrm{A}_\kappa}$ and $\mathcal{N}_{\mathrm{B}_\kappa}$ (degeneracy of the states $|0(t)\rangle$ *labelled* by different $\mathcal{N}_{\mathrm{A}_\kappa}$, or different $\mathcal{N}_{\mathrm{B}_\kappa}$, cf. Eq. (35)).

In conclusion, the θ-set $\{\theta_\kappa(\mathcal{N}_\kappa)\}$, is nothing but the squeezing parameter classifying the squeezed coherent states in the hyperplane $(X, \tilde{X}; Y, \tilde{Y})$. Note that to different squeezed states (different θ-sets) are associated unitarily inequivalent representations of the ccr's in the infinite volume limit. Also note that in the limit $t \to \infty$ the variances of the variables Y and $\tilde{X}$ become infinity making them completely spread out.

Further details on the squeezing states and their relation with deformed algebraic structures in QFT can be found in Refs. [28, 42, 43].

5 Quantum Brownian Motion

By following Schwinger [20], the description of a Brownian particle of mass M moving in a potential U(x) with a damping resistance R, interacting with a thermal bath at temperature T is provided by [23,24]

$$\mathcal{H}_{\rm Brownian} = \mathcal{H} - \frac{ik_BTR}{\hbar}(x_+ - x_-)^2 \ . \tag{44}$$

Here $\mathcal{H}$ is given by Eq. (16) and the evolution equation for the density matrix is

$$i\hbar\frac{\partial(x_+|\rho(t)|x_-)}{\partial t} = \mathcal{H}(x_+\,|\rho(t)|\,x_-) - (x_+\,|N[\rho]|\,x_-) \ , \tag{45}$$

where $N[\rho] \approx (ik_BTR/\hbar)[x,[x,\rho]]$ describes the effects of the reservoir random thermal noise [23,24]. In general the density operator in the above expression describes a mixed statistical state. The thermal bath contribution to the right hand side of Eq. (44), proportional to fluid temperature T, can be shown [24] to be equivalent to a white noise fluctuation source coupling the forward and backward motions according to

$$\langle y(t)y(t')\rangle_{\rm noise} = \frac{\hbar^2}{2Rk_BT}\delta(t-t') \ , \tag{46}$$

so that thermal fluctuations are always occurring in the difference $y = x_+ - x_-$ between forward in time and backward in time coordinates.

The correlation function for the random force f on the particle due to the bath is given by $G(t-s) = (i/\hbar)\langle f(t)f(s)\rangle$. The retarded and advanced Greens functions are studied in Ref. [23] and for brevity I omit here their discussion. The mechanical resistance is defined by $R = lim_{\Omega\to 0}\mathcal{R}eZ(\Omega + i0^+)$, with the mechanical impedance $Z(\zeta)$ (analytic in the upper half complex frequency plane $\mathcal{I}m\ \zeta > 0$) determined by the retarded Greens function $-i\zeta Z(\zeta) = \int_0^\infty dtG_{ret}(t)e^{i\zeta t}$. The time domain quantum noise in the fluctuating random force is $N(t-s) = (1/2)\langle f(t)f(s) + f(s)f(t)\rangle$.

The interaction between the bath and the particle is evaluated by following Feynman and Vernon and one finds [23] for the real and the imaginary part of the action

$$\mathcal{R}e\mathcal{A}[x,y] = \int_{t_i}^{t_f} dtL \ , \tag{47}$$

$$\mathcal{I}m\mathcal{A}[x,y] = (1/2\hbar)\int_{t_i}^{t_f}\int_{t_i}^{t_f} dtdsN(t-s)y(t)y(s) \ , \tag{48}$$

respectively, where L is defined in Eq. (18) for the given choice of $U(x_\pm)$ there adopted (without loss of generality).

I observe that at the classical level the "extra" coordinate y, is usually constrained to vanish. Note that $y(t) = 0$ is a true solution to Eq. (20) so that the constraint is *not* in violation of the equations of motion. From Eqs. (47)

and (48) one sees that *at quantum level nonzero y allows quantum noise effects arising from the imaginary part of the action.* On the contrary, in the classical "$\hbar \to 0$" limit nonzero y yields an "unlikely process" in view of the large imaginary part of the action implicit in Eq. (48). Thus, the meaning of the constraint $y = 0$ at the classical level is the one of avoiding such "unlikely process".

The rôle of the doubled y coordinate (the quantum β, or B mode in the discussion of the previous section) is thus shown again to be absolutely crucial in the quantum regime. There it accounts for the quantum noise in the fluctuating random force in the system-environment coupling [23]: in the limit of $y \to 0$ (i.e. for $x_+ = x_-$) quantum effects are lost and the classical limit is obtained.

It is interesting to remark that the forward and backward in time velocity components $v_\pm = \dot{x}_\pm$ in the (x_+, x_-) plane

$$v_\pm = \frac{\partial \mathcal{H}}{\partial p_\pm} = \pm \frac{1}{M} \left(p_\pm \mp \frac{R}{2} x_\mp \right) \tag{49}$$

do not commute

$$[v_+, v_-] = i\hbar \frac{R}{M^2}, \tag{50}$$

and it is thus impossible to fix these velocities v_+ and v_- as being identical. Equation (50) is similar to the usual commutation relations for the quantum velocities $\mathbf{v} = (\mathbf{p} - (e\mathbf{A}/c))/M$ of a charged particle moving in a magnetic field $\mathbf{B}$; i.e. $[v_1, v_2] = (i\hbar e B_3/M^2 c)$. Just as the magnetic field $\mathbf{B}$ induces an Aharonov-Bohm phase interference for the charged particle, the Brownian motion friction coefficient R induces an analogous phase interference between forward and backward motion which expresses itself as mechanical damping. Equation (50) will be also discussed in connection with noncommutative geometry induced by quantum dissipation [40]. I will comment more on this in the next section.

In the discussion above I have considered the low temperature limit: $T \ll T_\gamma$ where $k_B T_\gamma = \hbar\gamma = \frac{\hbar R}{2M}$. At high temperature, $T \gg T_\gamma$, the thermal bath motion suppresses the probability for $x_+ \neq x_-$ due to the thermal term $(k_B T R/\hbar)(x_+ - x_-)^2$ in Eq. (44) (cf. also Eq. (46)). By writing the diffusion coefficient $D = \frac{k_B T}{R}$ as

$$D = \frac{T}{T_\gamma} \Big(\frac{\hbar}{2M} \Big) , \tag{51}$$

the condition for classical Brownian motion for high mass particles is that $D \gg (\hbar/2M)$, and the condition for quantum interference with low mass particles is that $D \ll (\hbar/2M)$. In colloidal systems, for example, classical Brownian motion for large particles would appear to dominate the motion. In a fluid at room temperature it is typically $D \sim (\hbar/2M)$ for a single atom, or, equivalently, $T \sim T_\gamma$, so that the rôle played by quantum mechanics, although perhaps not dominant, may be an important one in the Brownian motion.

6 Dissipative Noncommutative Plane

The harmonic oscillator on the noncommutative plane, the motion of a particle in an external magnetic field and the Landau problem on the noncommutative sphere are only few examples of systems whose noncommutative geometry has been studied in detail. Noncommutative geometries are also of interest in Chern–Simons gauge theories, the usual gauge theories and string theories, in gravity theory [44, 45]. Here I show that quantum dissipation induces noncommutative geometry in the (x_+, x_-) plane [40].

The velocity components $v_\pm = \dot{x}_\pm$ in the (x_+, x_-) plane are given Eq. (49). Similarly,

$$\dot{p}_\pm = -\frac{\partial \mathcal{H}}{\partial x_\pm} = \mp U'(x_\pm) \mp \frac{Rv_\mp}{2} \ . \tag{52}$$

From Eqs. (49) and (52) it follows that

$$M\dot{v}_\pm + Rv_\mp + U'(x_\pm) = 0 \ . \tag{53}$$

When the choice $U(x_\pm) = \frac{1}{2}\kappa x_\pm^2$ is made, these are equivalent to the equations Eqs. (19) and (20). The classical equation of motion including dissipation thereby holds true if $x_+(t) \approx x_-(t) \approx x(t)$:

$$M\dot{v} + Rv + U'(x) = 0 \ . \tag{54}$$

If one defines

$$Mv_\pm = \hbar K_\pm \ , \tag{55}$$

then Eq. (50) gives

$$[K_+, K_-] = \frac{iR}{\hbar} \equiv \frac{i}{L^2} \ , \tag{56}$$

and a canonical set of conjugate position coordinates (ξ_+, ξ_-) may be defined by

$$\begin{aligned} \xi_\pm &= \mp L^2 K_\mp \\ [\xi_+, \xi_-] &= iL^2 \ . \end{aligned} \tag{57}$$

Another independent canonical set of conjugate position coordinates (X_+, X_-) is defined by (Fig. 2)

$$\begin{aligned} x_+ = X_+ + \xi_+ \ , \ x_- &= X_- + \xi_- \\ [X_+, X_-] &= -iL^2 \ . \end{aligned} \tag{58}$$

Note that $[X_a, \xi_b] = 0$, where $a = \pm$ and $b = \pm$.

The commutation relations Eqs. (57) and (58) characterize the noncommutative geometry in the plane (x_+, x_-). It is interesting to consider the case of pure friction in which the potential $U = 0$. Equations (16), (55) and (57) then imply

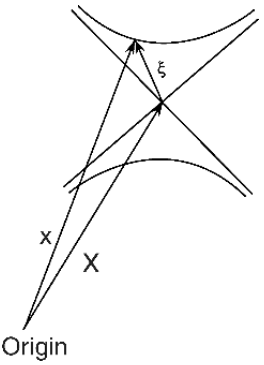

Fig. 2. The hyperbolic path of a particle moving in the $x = (x_+, x_-)$ plane. The noncommuting coordinate pairs $X = (X_+, X_-)$, which points from the origin to hyperbolic center, and $\xi = (\xi_+, \xi_-)$, which points from the center of the orbit to the position on the hyperbola, are shown. $x = X + \xi$

$$\mathcal{H}_{\rm friction} = \frac{\hbar^2}{2M}(K_+^2 - K_-^2) = -\frac{\hbar^2}{2ML^4}(\xi_+^2 - \xi_-^2)\ . \tag{59}$$

The equations of motion are

$$\dot{\xi}_\pm = \frac{i}{\hbar}\,[\mathcal{H}_{\rm friction}, \xi_\pm] = -\frac{\hbar}{ML^2}\xi_\mp = -\frac{R}{M}\xi_\mp = -\Gamma\xi_\mp\ , \tag{60}$$

with the solution

$$\begin{pmatrix} \xi_+(t) \\ \xi_-(t) \end{pmatrix} = \begin{pmatrix} \cosh(\Gamma t) & -\sinh(\Gamma t) \\ -\sinh(\Gamma t) & \cosh(\Gamma t) \end{pmatrix} \begin{pmatrix} \xi_+ \\ \xi_- \end{pmatrix}\ . \tag{61}$$

Equation (61) describes the hyperbolic orbit

$$\xi_-(t)^2 - \xi_+(t)^2 = \frac{2L^2}{\hbar\Gamma}\mathcal{H}_{\rm friction}\ . \tag{62}$$

The hyperbolae are defined by $(x - X)^2 - c^2(t - T)^2 = \Lambda^2$, where $\Lambda^2 = (\frac{mc}{\hbar}L^2)^2$, the hyperbolic center is at (X, cT) and one branch of the hyperbolae is a particle moving forward in time while the other branch is the same particle moving backward in time as an anti-particle.

Now I observe that a quantum phase interference of the Aharonov-Bohm type can always be associated with the noncommutative plane where

$$[X, Y] = iL^2\ , \tag{63}$$

with L denoting the geometric length scale in the plane. Suppose that a particle can move from an initial point in the plane to a final point in the plane via one of two paths, say $\mathcal{P}_1$ or $\mathcal{P}_2$. Since the paths start and finish at the same point, if one transverses the first path in a forward direction and the second

path in a backward direction, then the resulting closed path encloses an area $\mathcal{A}$. The phase interference ϑ between these two points is determined by the difference between the actions for these two paths $\hbar\vartheta = \mathcal{S}(\mathcal{P}_1) - \mathcal{S}(\mathcal{P}_2)$, and I show below it may be written as

$$\vartheta = \frac{\mathcal{A}}{L^2} \, . \tag{64}$$

A physical realization of the mathematical noncommutative plane is present in every laboratory wherein a charged particle moves in a plane with a normal uniform magnetic field $\mathbf{B}$. For this case, there are two canonical pairs of position coordinates which do not commute: (i) the position $\mathbf{R}$ of the center of the cyclotron circular orbit and (ii) the radius vector ρ from the center of the circle to the charged particle position $\mathbf{r} = \mathbf{R} + \rho$ (Fig. 3). The magnetic length scale of the noncommuting geometric coordinates is due to Landau [46],

$$L^2 = \frac{\hbar c}{eB} = \frac{\phi_0}{2\pi B} \quad \text{(magnetic)} \, . \tag{65}$$

Here ϕ_0 is the magnitude of the magnetic flux quantum associated with a charge e.

For motion at fixed energy one may (in classical mechanics) associate with each path $\mathcal{P}$ (in phase space) a phase space action integral

$$\mathcal{S}(\mathcal{P}) = \int_{\mathcal{P}} p_i dq^i \, . \tag{66}$$

As said, the phase interference ϑ between the two paths $\mathcal{P}_1$ and $\mathcal{P}_2$ is determined by the action difference

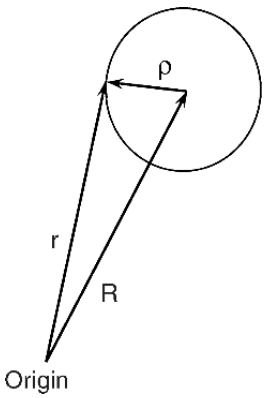

Fig. 3. A charge e moving in a circular cyclotron orbit. Noncommuting coordinate pairs are $\mathbf{R} = (X, Y)$, which points from the origin to the orbit center, and $\rho = (\rho_x, \rho_y)$, which points from the center of the orbit to the charge position $\mathbf{r} = \mathbf{R} + \rho$

$$\hbar\vartheta = \int_{\mathcal{P}_1} p_i dq^i - \int_{\mathcal{P}_2} p_i dq^i = \oint_{\mathcal{P}=\partial\Omega} p_i dq^i \tag{67}$$

wherein $\mathcal{P}$ is the closed path which goes from the initial point to the final point via path $\mathcal{P}_1$ and returns back to the initial point via $\mathcal{P}_2$. The closed $\mathcal{P}$ path may be regarded as the boundary of a two-dimensional surface Ω; i.e. $\mathcal{P} = \partial\Omega$. Stokes theorem yields

$$\vartheta = \frac{1}{\hbar} \oint_{\mathcal{P}=\partial\Omega} p_i dq^i = \frac{1}{\hbar} \int_{\Omega} (dp_i \wedge dq^i) \; . \tag{68}$$

The quantum phase interference ϑ between two alternative paths is thereby proportional to an "area" $\mathcal{A}$ of a surface Ω in phase space $(p_1, \ldots, p_f; q^1, \ldots, q^f)$ as described by the right hand side of Eq. (68).

If one reverts to the operator formalism and writes the commutation Eq. (63) in the noncommutative plane as

$$[X, P_X] = i\hbar \quad \text{where} \quad P_X = \left(\frac{\hbar Y}{L^2}\right) \; , \tag{69}$$

then back in the path integral formalism Eq. (68) reads

$$\vartheta = \frac{1}{\hbar} \int_{\Omega} (dP_X \wedge dX) = \frac{1}{L^2} \int_{\Omega} (dY \wedge dX) \tag{70}$$

and Eq. (64) is proved, i.e. the quantum phase interference between two alternative paths in the plane is determined by the noncommutative length scale L and the enclosed area $\mathcal{A}$.

I also remark that the existence of a phase interference is intimately connected to the zero point fluctuations in the coordinates; e.g. Eq. (63) implies a zero point uncertainty relation $\Delta X \Delta Y \geq (L^2/2)$.

Resorting back to Eq. (56) for the quantum dissipative case, i.e.

$$L^2 = \frac{\hbar}{R} \quad \text{(dissipative)} \; , \tag{71}$$

one then concludes that, provided $x_+ \neq x_-$, the quantum dissipative phase interference $\vartheta = \frac{\mathcal{A}}{L^2} = \frac{\mathcal{A}R}{\hbar}$ is associated with two paths in the noncommutative plane, starting at the same point $\mathcal{P}_1$ and ending to the same point $\mathcal{P}_2$ so to enclose the surface of area $\mathcal{A}$.

A comparison can be made between the noncommutative dissipative plane and the noncommutative Landau magnetic plane as shown in Fig. 3. The circular orbit in Fig. 3 for the magnetic problem is replaced by the hyperbolic orbit and it may be shown that the magnetic field is replaced by the electric field. The hyperbolic orbit in Fig. 2 is reflected in the classical orbit for a charged particle moving along the x-axis in a uniform electric field. For more details on this comparison see [40].

Finally, I recall that the Lagrangian for the system of Eq. (53) has been found [27] to be the same as the Lagrangian for three-dimensional topological massive Chern–Simons gauge theory in the infrared limit. It is also the same as for a Bloch electron in a solid which propagates along a lattice plane with a hyperbolic energy surface [27]. In the Chern–Simons case one has $\theta_{CS} = R/M = (\hbar/ML^2)$, with θ_{CS} the "topological mass parameter". In the Bloch electron case, $(eB/\hbar c) = (1/L^2)$, with B denoting the z-component of the applied external magnetic field. In Ref. [27] it has been considered the symplectic structure for the system of Eq. (53) in the case of strong damping $R \gg M$ (the so-called reduced case) in the Dirac constraint formalism as well as in the Faddeev and Jackiw formalism [47] and in both formalisms a non-zero Poisson bracket for the $(x_+,\ x_-)$ coordinates has been found.

Below I will consider the algebraic structure of the space of the physical states emergent from the doubling of the degrees of freedom discussed in the present and in the previous section. Before that I will discuss thermal field theory in the following section.

7 Thermal Field Theory

In this section I discuss the doubling of the degrees of freedom in connection with thermal field theory. Specifically, I will comment on the formalism of thermo field dynamics (TFD) [5, 8, 48]. In Sect. 8 it will be shown that the algebraic structure on which the TFD formalism is based is naturally provided by the q-deformed Hopf algebras for bosons and for fermions (usually called $h_q(1)$ and $h_q(1|1)$, respectively).

The central point in the TFD formalism is the possibility to express the statistical average $\langle \mathcal{A} \rangle$ of an observable $\mathcal{A}$ as the expectation value in the temperature dependent vacuum $|0(\beta)\rangle$:

$$\langle \mathcal{A} \rangle \equiv \frac{Tr[\mathcal{A}\ e^{-\beta\mathcal{H}}]}{Tr[e^{-\beta\mathcal{H}}]} = \langle 0(\beta)|\mathcal{A}|0(\beta)\rangle\ , \tag{72}$$

where $\mathcal{H} = H - \mu N$, with μ the chemical potential.

The first problem is therefore to construct a suitable temperature dependent state $|0(\beta)\rangle$ which satisfies Eq. (72), namely

$$\langle 0(\beta)|\mathcal{A}|0(\beta)\rangle = \frac{1}{Tr[e^{-\beta\mathcal{H}}]} \sum_n \langle n|\mathcal{A}|n\rangle e^{-\beta E_n}\ , \tag{73}$$

for an arbitrary variable $\mathcal{A}$, with

$$\mathcal{H}|n\rangle = E_n|n\rangle\ , \quad \langle n|m\rangle = \delta_{nm}\ . \tag{74}$$

Such a state cannot be constructed as long as one remains in the original Fock space $\{|n\rangle\}$. To see this, let me closely follow [8]. One can expand $|0(\beta)\rangle$ in terms of $|n\rangle$ as

$$|0(\beta)\rangle = \sum_n f_n(\beta)|n\rangle \ . \tag{75}$$

Then, use of this equation into (73) gives

$$f_n^*(\beta) f_m(\beta) = \frac{1}{Tr[e^{-\beta\mathcal{H}}]} e^{-\beta E_n}\delta_{nm} \ , \tag{76}$$

which is impossible to be satisfied by c-number functions $f_n(\beta)$. However, Eq. (76) can be regarded as the orthogonality condition in a Hilbert space in which the expansion coefficient $f_n(\beta)$ is a vector. In order to realize such a representation it is convenient to introduce a dynamical system identical to the one under study, namely to double the given system. The quantities associated with the doubled system are denoted by the tilde in the usual notation of TFD [8]. Thus the tilde-system is characterized by the Hamiltonian $\tilde{H}$ and the states are denoted by $|\tilde{n}\rangle$, with

$$\tilde{\mathcal{H}}|\tilde{n}\rangle = E_n|\tilde{n}\rangle \ , \quad \langle\tilde{n}|\tilde{m}\rangle = \delta_{nm} \ , \tag{77}$$

where E_n is the same as the one appearing in Eq. (74) by definition. It is also assumed that non-tilde and tilde operators are commuting (anti-commuting) boson (fermion) operators. One then considers the space spanned by the direct product $|n\rangle \otimes \ |\tilde{m}\rangle \equiv |n, \tilde{m}\rangle$. The matrix element of a bose-like operator $\mathcal{A}$ is then

$$\langle\tilde{m}, n|\mathcal{A}|n', \tilde{m}'\rangle = \langle n|\mathcal{A}|n'\rangle\delta_{mm'} \ , \tag{78}$$

and the one of the corresponding $\tilde{\mathcal{A}}$ is

$$\langle\tilde{m}, n|\tilde{\mathcal{A}}|n', \tilde{m}'\rangle = \langle\tilde{m}|\tilde{\mathcal{A}}|\tilde{m}'\rangle\delta_{nn'} \ . \tag{79}$$

In TFD it turns out to be convenient to identify

$$\langle m|\mathcal{A}|n\rangle = \langle\tilde{n}|\tilde{\mathcal{A}}^\dagger|\tilde{m}\rangle \ . \tag{80}$$

Eq. (76) is satisfied if one defines

$$f_n(\beta) = \frac{1}{\sqrt{Tr[e^{-\beta\mathcal{H}}]}} e^{\frac{-\beta E_n}{2}}|\tilde{n}\rangle \ , \tag{81}$$

and Eq. (73) is obtained by using the definition (81) in $|0(\beta)\rangle$ given by (75):

$$|0(\beta)\rangle = \frac{1}{\sqrt{Tr[e^{-\beta\mathcal{H}}]}} \sum_n e^{\frac{-\beta E_n}{2}}|n, \tilde{n}\rangle \ . \tag{82}$$

The vectors $|n\rangle$ and $|\tilde{n}\rangle$ thus appear as a pair in $|0(\beta)\rangle$. I remark that the formal rôle of the "doubled" states $|\tilde{n}\rangle$ is merely to pick up the diagonal matrix elements of $\mathcal{A}$. In this connection, thinking of the rôle of the environment, which is able to reduce the system density matrix to its diagonal form in the QM decoherence processes [12], it is remarkable that the doubled degrees of

freedom in TFD are indeed susceptible of being interpreted as the environment degrees of freedom, as better specified in the following.

It is useful to consider, as an example, the case of the number operator. Let $\mathcal{A} \equiv N = a^\dagger a$. For definiteness I consider the boson case. Then the statistical average of N is the Bose-Einstein distribution $f_B(\Omega)$, where Ω denotes the mode energy, $H = \Omega a^\dagger a$,

$$\langle N \rangle \equiv \frac{Tr[N\ e^{-\beta H}]}{Tr[e^{-\beta H}]} = \langle 0(\beta)|N|0(\beta)\rangle = \frac{1}{e^{\beta\Omega} - 1} = f_B(\Omega) \ . \tag{83}$$

One then can show [8] that, by setting

$$u(\beta) \equiv \sqrt{1 + f_B(\Omega)}, \quad v(\beta) \equiv \sqrt{f_B(\Omega)}\ , \tag{84}$$

$$u^2(\beta) - v^2(\beta) = 1\ , \tag{85}$$

so that

$$u(\beta) = \cosh\theta(\beta)\ , \quad v(\beta) = \sinh\theta(\beta)\ , \tag{86}$$

and defining

$$\mathcal{G} \equiv -i(a^\dagger \tilde{a}^\dagger - a\tilde{a})\ , \tag{87}$$

the state $|0(\beta)\rangle$ is formally given (at finite volume) by

$$|0(\beta)\rangle = e^{i\theta(\beta)\mathcal{G}}|0\rangle = \frac{1}{u(\beta)} \exp\Big(\frac{v(\beta)}{u(\beta)}\Big) a^\dagger \tilde{a}^\dagger |0\rangle\ . \tag{88}$$

It is clear that the state $|0(\beta)\rangle$ is not annihilated by a and $\tilde{a}$. However, it is annihilated by the "new" set of operators $a(\theta)$ and $\tilde{a}(\theta)$,

$$a(\theta)|0(\beta)\rangle = 0 = \tilde{a}(\theta)|0(\beta)\rangle\ , \tag{89}$$

with

$$\begin{aligned} a(\theta) &= \exp(i\theta\mathcal{G})\ a\ \exp(-i\theta\mathcal{G}) = a\cosh\theta - \tilde{a}^\dagger \sinh\theta\ , \\ \tilde{a}(\theta) &= \exp(i\theta\mathcal{G})\ \tilde{a}\ \exp(-i\theta\mathcal{G})\ = \tilde{a}\cosh\theta - a^\dagger \sinh\theta\ , \end{aligned} \tag{90}$$

$$[a(\theta), a^\dagger(\theta)] = 1\ , \quad [\tilde{a}(\theta), \tilde{a}^\dagger(\theta)] = 1\ . \tag{91}$$

All other commutators are equal to zero and $a(\theta)$ and $\tilde{a}(\theta)$ commute among themselves. Equation (90) are nothing but the Bogoliubov transformations of the $(a, \tilde{a})$ pair into a new set of creation, annihilation operators. I will show in Sect. 8 that the Bogoliubov-transformed operators $a(\theta)$ and $\tilde{a}(\theta)$ are linear combinations of the deformed coproduct operators.

The state $|0(\beta)\rangle$ is not the vacuum (zero energy eigenstate) of H and of $\tilde{H}$. It is, however, the zero energy eigenstate for the "Hamiltonian" $\hat{H}$, $\hat{H}|0(\beta)\rangle = 0$, with

$$\hat{H} \equiv H - \tilde{H} = \Omega(a^\dagger a - \tilde{a}^\dagger \tilde{a})\ . \tag{92}$$

The state $|0(\beta)\rangle$ is called the thermal vacuum.

I note that in the boson case $J_1 \equiv \frac{1}{2}(a^\dagger \tilde{a}^\dagger + a\tilde{a})$ together with $J_2 \equiv \frac{1}{2}\mathcal{G}$ and $J_3 \equiv \frac{1}{2}(N + \tilde{N} + 1)$ close the algebra $su(1,1)$. Moreover, $\frac{\delta}{\delta\theta}(N(\theta) - \tilde{N}(\theta)) = 0$, with $(N(\theta) - \tilde{N}(\theta)) \equiv (a^\dagger(\theta)a(\theta) - \tilde{a}^\dagger(\theta)\tilde{a}(\theta))$, consistently with the fact that $\frac{1}{4}(N - \tilde{N})^2$ is the $su(1,1)$ Casimir operator.

In the fermion case $J_1 \equiv \frac{1}{2}\mathcal{G}$, $J_2 \equiv \frac{1}{2}(a^\dagger \tilde{a}^\dagger + a\tilde{a})$ and $J_3 \equiv \frac{1}{2}(N + \tilde{N} - 1)$ close the algebra $su(2)$. Also in this case $\frac{\delta}{\delta\theta}(N(\theta) - \tilde{N}(\theta)) = 0$, with $(N(\theta) - \tilde{N}(\theta)) \equiv (a^\dagger(\theta)a(\theta) - \tilde{a}^\dagger(\theta)\tilde{a}(\theta))$, again consistently with the fact that $\frac{1}{4}(N - \tilde{N})^2$ is related to the $su(2)$ Casimir operator.

Summarizing, the vacuum state for $a(\theta)$ and $\tilde{a}(\theta)$ is formally given (at finite volume) by

$$|0(\theta)\rangle = \exp\ (i\theta\mathcal{G})\ \ |0,0\rangle = \sum_n c_n(\theta)\,|n,\tilde{n}\rangle\ , \tag{93}$$

with $n, \tilde{n} = 0, \ldots, \infty$ for bosons and $n, \tilde{n} = 0, 1$ for fermions, and it appears therefore to be an $SU(1,1)$ or $SU(2)$ generalized coherent state [33], respectively for bosons or for fermions.

In the infinite volume limit $|0(\theta)\rangle$ becomes orthogonal to $|0,0\rangle$ and we have that the whole Hilbert space $\{|0(\theta)\rangle\}$, constructed by operating on $|0(\theta)\rangle$ with $a^\dagger(\theta)$ and $\tilde{a}^\dagger(\theta)$, is asymptotically (i.e. in the infinite volume limit) orthogonal to the space generated over $\{|0,0\rangle\}$. In general, for each value of $\theta(\beta)$, i.e. for each value of the temperature $T = \frac{1}{k_B\beta}$, one obtains in the infinite volume limit a representation of the canonical commutation relations unitarily inequivalent to the others, associated with different values of T. In other words, the parameter $\theta(\beta)$ (or the temperature T) acts as a label for the inequivalent representations [25].

The TFD formalism is a fully developed QFT formalism [5,8,48] and it has been applied to a rich set of problems of physical interest, in condensed matter physics, high energy physics, quantum optics, etc. (see [5,8,17,18,25,28,37,48] and references therein quoted). I will show in Sect. 8 that the doubling of the degrees of freedom on which the TFD formalism is based finds its natural realization in the coproduct map.

Let me recall the so-called "tilde-conjugation rules" which are defined in TFD. For any two bosonic (respectively, fermionic) operators $\mathcal{O}$ and $\mathcal{O}'$ and any two c-numbers α and β the tilde-conjugation rules of TFD are postulated to be the following [8]:

$$(\mathcal{O}\mathcal{O}')\tilde{} = \tilde{\mathcal{O}}\tilde{\mathcal{O}}'\ , \tag{94}$$

$$(\alpha\mathcal{O} + \beta\mathcal{O}')\tilde{} = \alpha^*\tilde{\mathcal{O}} + \beta^*\tilde{\mathcal{O}}'\ , \tag{95}$$

$$(\mathcal{O}^\dagger)\tilde{} = \tilde{\mathcal{O}}^\dagger\ , \tag{96}$$

$$(\tilde{\mathcal{O}})\tilde{} = \mathcal{O}\ . \tag{97}$$

According to (94) the tilde-conjugation does not change the order among operators. Furthermore, it is required that tilde and non-tilde operators are

mutually commuting (or anti-commuting) operators and that the thermal vacuum $|0(\beta)\rangle$ is invariant under tilde-conjugation:

$$[\mathcal{O}, \tilde{\mathcal{O}'}]_{\mp} = 0 = [\mathcal{O}, \tilde{\mathcal{O}'}^{\dagger}]_{\mp} \ , \tag{98}$$

$$\tilde{|0(\beta)\rangle} = |0(\beta)\rangle \ . \tag{99}$$

In order to use a compact notation it is useful to introduce the label σ defined by $\sqrt{\sigma} \equiv +1$ for bosons and $\sqrt{\sigma} \equiv +i$ for fermions. I shall therefore simply write commutators as $[\mathcal{O}, \mathcal{O}']_{-\sigma} \doteq \mathcal{O}\mathcal{O}' - \sigma\mathcal{O}'\mathcal{O}$, and $(1 \otimes \mathcal{O})(\mathcal{O}' \otimes 1) \equiv \sigma(\mathcal{O}' \otimes 1)(1 \otimes \mathcal{O})$, without further specification of whether $\mathcal{O}$ and $\mathcal{O}'$ (which are equal to a, $a^{\dagger}$ in all possible ways) are fermions or bosons.

Upon identifying from now on $a_1 \equiv a$, $a_1^{\dagger} \equiv a^{\dagger}$, one easily checks that the TFD tilde-operators (consistent with (94)–(99)) are straightforwardly recovered by setting $a_2 \equiv \tilde{a}$, $a_2^{\dagger} \equiv \tilde{a}^{\dagger}$. In other words, according to such identification, it is the action of the $1 \leftrightarrow 2$ permutation π: $\pi a_i = a_j, i \neq j, i, j = 1, 2$, that realizes the operation of "tilde-conjugation" defined in (94–97):

$$\pi a_1 = \pi(a \otimes \mathbf{1}) = \mathbf{1} \otimes a = a_2 \equiv \tilde{a} \equiv \tilde{(a)} \tag{100}$$

$$\pi a_2 = \pi(\mathbf{1} \otimes a) = a \otimes \mathbf{1} = a_1 \equiv a \equiv \tilde{(\tilde{a})} \ . \tag{101}$$

In particular, since the permutation π is involutive, also tilde-conjugation turns out to be involutive, as in fact required by the rule (97). Notice that, as $(\pi a_i)^{\dagger} = \pi(a_i{}^{\dagger})$, it is also $(\tilde{(a_i)}\,)^{\dagger} = \tilde{((a_i)^{\dagger})}$, i.e. tilde-conjugation commutes with hermitian conjugation. Furthermore, from (100)–(101), one has

$$\tilde{(ab)} = \tilde{[(a \otimes \mathbf{1})(b \otimes \mathbf{1})]} = \tilde{(ab \otimes \mathbf{1})} = \mathbf{1} \otimes ab = (\mathbf{1} \otimes a)(\mathbf{1} \otimes b) = \tilde{a}\tilde{b} \ . \tag{102}$$

Rules (96) and (94) are thus obtained. (98) is insured by the σ-commutativity of a_1 and a_2. The vacuum of TFD, $|0(\beta)\rangle$, is a condensed state of equal number of tilde and non-tilde particles [8], thus (99) requires no further conditions: Eqs. (100)–(101) are sufficient to show that the rule (99) is satisfied.

TFD appears equipped with a set of canonically conjugate "thermal" variables: θ and $p_{\theta} \equiv -i\dfrac{\delta}{\delta\theta}$. p_{θ} can be regarded as the momentum operator "conjugate" to the "thermal degree of freedom" θ. The notion of thermal degree of freedom [48] thus acquires formal definiteness in the sense of the canonical formalism. It is remarkable that the "conjugate thermal momentum" p_{θ} generates transitions among inequivalent (in the infinite volume limit) representations: $\exp(i\bar{\theta}p_{\theta}) \ \ |0(\theta)\rangle = |0(\theta + \bar{\theta})\rangle$. Notice that derivative with respect to the θ parameter is actually a derivative with respect to the system temperature T. This sheds some light on the rôle of θ in thermal field theories for non-equilibrium systems and phase transitions. I shall comment more on this point in the following section.

Finally, when the proper field description is taken into account, a and $\tilde{a}$ carry dependence on the momentum $\mathbf{k}$. The Bogoliubov transformation analogously, should be thought of as inner automorphism of the algebra $su(1,1)_{\mathbf{k}}$

(or $su(2)_{\mathbf{k}}$). This shows that one is globally dealing with $\oplus_{\mathbf{k}} su(1,1)_{\mathbf{k}}$ (or $\oplus_{\mathbf{k}} su(2)_{\mathbf{k}}$). Therefore one is lead to consider $\mathbf{k}$-dependence also for the θ parameter.

As a final comment, I observe that the "analogies" with the formalism of quantum dissipation presented in Sect. 3.1 are evident.

8 The *q*-deformed Hopf Algebra and QFT

In this section I want to point out that the doubling of the degrees of freedom is intimately related to the structure of the space of the states in QFT [9]. This brings us to consider the q-deformed Hopf algebra [6,7].

One key ingredient of Hopf algebra [7] is the coproduct operation, i.e. the operator doubling implied by the coalgebra. The coproduct operation is indeed a map $\Delta : \mathcal{A} \to \mathcal{A} \otimes \mathcal{A}$ which duplicates the algebra $\mathcal{A}$. Coproducts are commonly used in the familiar addition of energy, momentum, angular momentum and of other so-called primitive operators. The coproduct of a generic operator $\mathcal{O}$ is a homomorphism defined as $\Delta\mathcal{O} = \mathcal{O} \otimes \mathbf{1} + \mathbf{1} \otimes \mathcal{O} \equiv \mathcal{O}_1 + \mathcal{O}_2$, with $\mathcal{O} \in \mathcal{A}$. Since additivity of observables such as energy, momentum, angular momentum, etc. is an essential requirement, the coproduct, and therefore the Lie-Hopf algebra structure, appears to provide an essential algebraic tool in QM and in QFT.

The systems discussed in the sections above, where the duplication of the degrees of freedom has revealed to be central, are thus natural candidates to be described by the Lie-Hopf algebra. The remarkable result holds [9] according to which the infinitely many ui representations of the ccr, whose existence characterizes QFT, are classified by use of the *q-deformed* Hopf algebra. Quantum deformations of Hopf algebra have thus a deeply non-trivial physical meaning in QFT.

In the following I consider boson operators. The discussion and the conclusions can be easily extended to the case of fermion operators [9]. For notational simplicity I will omit the momentum suffix κ.

The bosonic algebra $h(1)$ is generated by the set of operators $\{a, a^\dagger, H, N\}$ with commutation relations:

$$[\, a\,,\, a^\dagger\,] = \; 2H\,, \quad [\, N\,,\, a\,] = -a\,, \quad [\, N\,,\, a^\dagger\,] = a^\dagger\,, \quad [\, H\,,\, \bullet\,] = 0\,. \tag{103}$$

H is a central operator, constant in each representation. The Casimir operator is given by $\mathcal{C} = 2NH - a^\dagger a$. $h(1)$ is an Hopf algebra and is therefore equipped with the coproduct operation, defined by

$$\Delta a = a \otimes \mathbf{1} + \mathbf{1} \otimes a \equiv a_1 + a_2\,, \quad \Delta a^\dagger = a^\dagger \otimes \mathbf{1} + \mathbf{1} \otimes a^\dagger \equiv a_1^\dagger + a_2^\dagger\,, \tag{104}$$

$$\Delta H = H \otimes \mathbf{1} + \mathbf{1} \otimes H \equiv H_1 + H_2\,, \quad \Delta N = N \otimes \mathbf{1} + \mathbf{1} \otimes N \equiv N_1 + N_2\,. \tag{105}$$

Note that $[a_i, a_j] = [a_i, a_j^\dagger] = 0,\ i,j = 1,2,\ i \neq j$. The coproduct provides the prescription for operating on two modes. As mentioned, one familiar example

of coproduct is the addition of the angular momentum J^α, $\alpha = 1, 2, 3$, of two particles: $\Delta J^\alpha = J^\alpha \otimes \mathbf{1} + \mathbf{1} \otimes J^\alpha \equiv J_1^\alpha + J_2^\alpha,\ J^\alpha \in su(2)$.

The q-deformation of $h(1)$ is the Hopf algebra $h_q(1)$:

$$[\, a_q\, ,\, a_q^\dagger\,] = \ [2H]_q\,, \quad [\, N\ ,\, a_q\,] = -a_q\,, \quad [\, N\,,\, a_q^\dagger\,] = a_q^\dagger, \quad [\, H\ ,\, \bullet\,] = 0\,, \tag{106}$$

where $N_q \equiv N$ and $H_q \equiv H$. The Casimir operator $\mathcal{C}_q$ is given by $\mathcal{C}_q = N[2H]_q - a_q^\dagger a_q$, where $[x]_q = \frac{q^x - q^{-x}}{q - q^{-1}}$. The deformed coproduct is defined by

$$\Delta a_q = a_q \otimes q^H + q^{-H} \otimes a_q\,, \qquad \Delta a_q^\dagger = a_q^\dagger \otimes q^H + q^{-H} \otimes a_q^\dagger\,, \tag{107}$$

$$\Delta H = H \otimes \mathbf{1} + \mathbf{1} \otimes H\,, \quad \Delta N = N \otimes \mathbf{1} + \mathbf{1} \otimes N\,, \tag{108}$$

whose algebra is isomorphic with (106): $[\Delta a_q, \Delta a_q^\dagger] = [2\Delta H]_q$, etc. . Note that $h_q(1)$ is a structure different from the commonly considered q-deformation of the harmonic oscillator [49] that does not have a coproduct and thus cannot allow for the duplication of the state space.

I denote by $\mathcal{F}_1$ the single mode Fock space, i.e. the fundamental representation $H = 1/2$, $\mathcal{C} = 0$. In such a representation $h(1)$ and $h_q(1)$ coincide as it happens for $su(2)$ and $su_q(2)$ for the spin-$\frac{1}{2}$ representation. The differences appear in the coproduct and in the higher spin representations.

As customary, I require that a and $a^\dagger$, and a_q and ${a_q}^\dagger$, are adjoint operators. This implies that q can only be real (or of modulus one in the fermionic case. In the two mode Fock space $\mathcal{F}_2 = \mathcal{F}_1 \otimes \mathcal{F}_1$, for $|q| = 1$, the hermitian conjugation of the coproduct must be supplemented by the inversion of the two spaces for consistency with the coproduct isomorphism).

Summarizing, one can write for both bosons (and fermions) on $\mathcal{F}_2 = \mathcal{F}_1 \otimes \mathcal{F}_1$:

$$\Delta a = a_1 + a_2\,, \quad \Delta a^\dagger = a_1^\dagger + a_2^\dagger\,, \tag{109}$$

$$\Delta a_q = a_1 q^{1/2} + q^{-1/2} a_2\,, \quad \Delta a_q^\dagger = a_1^\dagger q^{1/2} + q^{-1/2} a_2^\dagger\,, \tag{110}$$

$$\Delta H = 1, \quad \Delta N = N_1 + N_2\,. \tag{111}$$

Now, the key point is [9] that the full set of infinitely many unitarily inequivalent representations of the ccr in QFT are classified by use of the q-deformed Hopf algebra. Since, as well known, the Bogoliubov transformations relate different (i.e. unitary inequivalent) representations, it is sufficient to show that the Bogoliubov transformations are directly obtained by use of the deformed copodruct operation. I consider therefore the following operators (cf. (107) with $q(\theta) \equiv e^{2\theta}$ and $H = 1/2$):

$$\alpha_{q(\theta)} \equiv \frac{\Delta a_q}{\sqrt{[2]_q}} = \frac{1}{\sqrt{[2]_q}}(e^\theta a_1 + e^{-\theta} a_2)\,, \tag{112}$$

$$\beta_{q(\theta)} \equiv \frac{1}{\sqrt{[2]_q}} \frac{\delta}{\delta\theta} \Delta a_q = \frac{2q}{\sqrt{[2]_q}} \frac{\delta}{\delta q} \Delta a_q = \frac{1}{\sqrt{[2]_q}}(e^\theta a_1 - e^{-\theta} a_2)\,, \tag{113}$$

and h.c.. A set of commuting operators with canonical commutation relations is given by

$$\alpha(\theta) \equiv \frac{\sqrt{[2]_q}}{2\sqrt{2}}[\alpha_{q(\theta)} + \alpha_{q(-\theta)} - \beta^\dagger_{q(\theta)} + \beta^\dagger_{q(-\theta)}] \,, \tag{114}$$

$$\beta(\theta) \equiv \frac{\sqrt{[2]_q}}{2\sqrt{2}}[\beta_{q(\theta)} + \beta_{q(-\theta)} - \alpha^\dagger_{q(\theta)} + \alpha^\dagger_{q(-\theta)}] \,, \tag{115}$$

and h.c. One then introduces [9]

$$A(\theta) \equiv \frac{1}{\sqrt{2}}\,(\alpha(\theta) + \beta(\theta)) = A\,\cosh\,\theta - B^\dagger\,\sinh\,\theta \,, \tag{116}$$

$$B(\theta) \equiv \frac{1}{\sqrt{2}}\,(\alpha(\theta) - \beta(\theta)) = B\,\cosh\,\theta - A^\dagger\,\sinh\,\theta \,, \tag{117}$$

with

$$[A(\theta), A^\dagger(\theta)] = 1 \,, \quad [B(\theta), B^\dagger(\theta)] = 1 \,. \tag{118}$$

All other commutators are equal to zero and $A(\theta)$ and $B(\theta)$ commute among themselves. Equations (116) and (117) are nothing but the Bogoliubov transformations for the (A, B) pair (see the corresponding transformations, e.g. in the case of the dho, Eqs. (30) and (31)). In other words, Eqs. (116), (117) show that the Bogoliubov-transformed operators $A(\theta)$ and $B(\theta)$ are linear combinations of the coproduct operators defined in terms of the deformation parameter $q(\theta)$ and of their θ-derivatives.

From this point on one can re-obtain the results discussed in the previous sections, for example for the dho provided one sets $\theta \equiv \Gamma t$.

The generator of (116) and (117) is $\mathcal{G} \equiv -i(A^\dagger B^\dagger - AB)$:

$$-\,i\frac{\delta}{\delta\theta}A(\theta) = [\mathcal{G}, A(\theta)] \,, \quad -\,i\frac{\delta}{\delta\theta}B(\theta) = [\mathcal{G}, B(\theta)] \,, \tag{119}$$

and h.c.. Compare this generator with H_I in Eq. (26).

Let $|0\rangle \equiv |0\rangle \otimes |0\rangle$ denote the vacuum annihilated by A and B, $A|0\rangle = 0 = B|0\rangle$. By introducing the suffix κ (till now omitted for simplicity), at finite volume V one obtains

$$|0(\theta)\rangle \,= e^{i\sum_\kappa \theta_\kappa \mathcal{G}_\kappa}|0\rangle \,= \prod_k \frac{1}{\cosh\theta_k} e^{\tanh\theta_k A^\dagger_k B^\dagger_k}|0\rangle \,, \tag{120}$$

to be compared with Eq. (33). θ denotes the set $\{\theta_\kappa = \frac{1}{2}\ln q_\kappa, \forall\kappa\}$ and $\langle 0(\theta)|0(\theta)\rangle = 1$. The underlying group structure is $\bigotimes_\kappa SU(1,1)_\kappa$ and the vacuum $|0(\theta)\rangle$ is an $SU(1,1)$ generalized coherent state [33]. The q-deformed Hopf algebra is thus intrinsically related to coherence and to the vacuum structure in QFT.

In the infinite volume limit, the number of degrees of freedom becomes uncountable infinite, and thus one obtains [5, 8, 25] $\langle 0(\theta)|0(\theta')\rangle \to 0$ as

$V \to \infty, \quad \forall\, \theta, \theta',\ \theta \neq \theta'$. By denoting with $\mathcal{H}_\theta$ the Hilbert space with vacuum $|0(\theta)\rangle$, $\mathcal{H}_\theta \equiv \{|0(\theta)\rangle\}$, this means that $\mathcal{H}_\theta$ and $\mathcal{H}_{\theta'}$ become unitarily inequivalent. In this limit, the "points" of the space $\mathcal{H} \equiv \{\mathcal{H}_\theta,\ \forall\, \theta\}$ of the infinitely many ui representations of the ccr are labelled by the deformation parameter θ [9,25]. The space $\mathcal{H} \equiv \{\mathcal{H}_\theta,\ \forall\, \theta\}$ is called the space of the representations.

I note that $p_\theta \equiv -i\dfrac{\delta}{\delta\theta}$ can be regarded [9] as the momentum operator "conjugate" to the "degree of freedom" θ. For an assigned fixed value $\bar\theta$, it is

$$e^{i\bar\theta p_\theta} A(\theta) = e^{i\bar\theta \mathcal{G}} A(\theta) e^{-i\bar\theta \mathcal{G}} = A(\theta + \bar\theta)\ , \tag{121}$$

and similarly for $B(\theta)$.

It is interesting to consider the case of time–dependent deformation parameter. This immediately relates to the dissipative systems considered in the previous sections. The Heisenberg equation for $A(t, \theta(t))$ is

$$\begin{aligned} -i\dot A(t,\theta(t)) &= -i\frac{\delta}{\delta t}A(t,\theta(t)) - i\frac{\delta\theta}{\delta t}\frac{\delta}{\delta\theta}A(t,\theta(t)) \\ &= [H, A(t,\theta(t))] + \frac{\delta\theta}{\delta t}[\mathcal{G}, A(t,\theta(t))] = [H+Q, A(t,\theta(t))]\ , \end{aligned} \tag{122}$$

and $Q \equiv \frac{\delta\theta}{\delta t}\mathcal{G}$ plays the role of the heat–term in dissipative systems. H is the Hamiltonian responsible for the time variation in the explicit time dependence of $A(t, \theta(t))$. $H + Q$ can be therefore identified with the free energy [25]: variations in time of the deformation parameter involve dissipation. In thermal theories and in dissipative systems the doubled modes B play the role of the thermal bath or environment.

Summarizing, QFT is characterized by the existence of ui representations of the ccr [3] which are related among themselves by the Bogoliubov transformations. These, as seen above, are obtained as linear combinations of the deformed coproduct maps which express the doubling of the degrees of freedom. Therefore one may conclude that the intrinsic algebraic structure of QFT (independent of the specificity of the system under study) is the one of the q-deformed Hopf algebra. The ui representations existing in QFT are related and labelled by means of such an algebraic structure.

It should be stressed that the coproduct map is also essential in QM in order to deal with a many mode system (typically, with identical particles). However, in QM all the representations of the ccr are unitarily equivalent and therefore the Bogoliubov transformations induce unitary transformations among the representations, thus preserving their physical content. The q-deformed Hopf algebra therefore does not have that physical relevance in QM, which it has, on the contrary, in QFT. Here, the representations of the ccr, related through Bogoliubov representations, are unitarily *inequivalent* and therefore physically inequivalent: they represent different physical phases of the system corresponding to different boundary conditions, such as, for example, the system temperature. Typical examples are the superconducting and

the normal phase, the ferromagnetic and the non-magnetic (i.e. zero magnetization) phase, the crystal and the gaseous phase, etc. The physical meaning of the deformation parameter q in terms of which ui representations are labelled is thus recognized.

When the above discussion is applied to non-equilibrium (e.g. thermal and/or dissipative) field theories it appears that the couple of conjugate variables θ and $p_\theta \equiv -i\frac{\partial}{\partial\theta}$, with $\theta = \theta(\beta(t))$ $(\beta(t) = \frac{1}{k_B T(t)})$, related to the q-deformation parameter, describe trajectories in the space $\mathcal{H}$ of the representations. In [10] it has been shown that there is a symplectic structure associated to the "degrees of freedom" θ and that the trajectories in the $\mathcal{H}$ space may exhibit properties typical of chaotic trajectories in classical nonlinear dynamics. I will discuss this in the following. In the next section I present further characterizations of the vacuum structure of the ui representations in QFT.

9 Entropy as a Measure of the Entanglement

In Sect. 3 I have shown that the time evolution of the state $|0(t)\rangle$ is actually controlled by the entropy variations (cf. Eq. (39)). I will shortly comment on the entropy in this section from a more general point of view, also in connection with entanglement of the $A-B$ modes, since it appears as a structural aspect of QFT related with the existence of the ui representations of the ccr.

The state $|0(\theta)\rangle$ may be written as:

$$|0(\theta)\rangle = \exp\left(-\frac{1}{2}S_A\right)|\,\mathcal{I}\rangle = \exp\left(-\frac{1}{2}S_B\right)|\,\mathcal{I}\rangle\ , \tag{123}$$

$$S_A \equiv -\sum_\kappa \Big\{A^\dagger_\kappa A_\kappa \ln\sinh^2\theta_\kappa - A_\kappa A^\dagger_\kappa \ln\cosh^2\theta_\kappa\Big\}\ . \tag{124}$$

Here $|\,\mathcal{I}\rangle \equiv \exp\left(\sum_\kappa A^\dagger_\kappa B^\dagger_\kappa\right)|0\rangle$ and S_B is given by an expression similar to S_A, with B_κ and $B^\dagger_\kappa$ replacing A_κ and $A^\dagger_\kappa$, respectively. I simply write S for either S_A or S_B. I can also write [5, 8, 25]:

$$|0(\theta)\rangle = \sum_{n=0}^{+\infty}\sqrt{W_n}\,(|n\rangle\otimes|n\rangle)\ , \tag{125}$$

$$W_n = \prod_k \frac{\sinh^{2n_k}\theta_k}{\cosh^{2(n_k+1)}\theta_k}\ , \tag{126}$$

with n denoting the set $\{n_\kappa\}$ and with $0 < W_n < 1$ and $\sum_{n=0}^{+\infty} W_n = 1$. Then

$$\langle 0(\theta)|S_A|0(\theta)\rangle = \sum_{n=0}^{+\infty} W_n lnW_n\ , \tag{127}$$

which confirms that S can be interpreted as the entropy operator [5, 8, 25].

The state $|0(\theta)\rangle$ in Eq. (120) can be also written as

$$|0(\theta)\rangle = \left(\prod_k \frac{1}{\cosh\theta_k}\right) \times \left(|0\rangle\otimes|0\rangle + \sum_k \tanh\theta_k \left(|A_k\rangle\otimes|B_k\rangle\right) + \dots\right), \qquad (128)$$

which clearly cannot be factorized into the product of two single-mode states. There is thus entanglement between the modes A and B: $|0(\theta)\rangle$ is an entangled state. Equations (125) and (127) then show that S provides a measure of the degree of entanglement.

I remark that the entanglement is truly realized in the infinite volume limit where

$$\langle 0(\theta)|0\rangle = e^{-\frac{V}{(2\pi)^3}\int d^3\kappa \ln\cosh\theta_\kappa} \underset{V\to\infty}{\longrightarrow} 0\,, \qquad (129)$$

provided $\int d^3\kappa \ln\cosh\theta_\kappa$ is not identically zero. The probability of having the component state $|n\rangle\otimes|n\rangle$ in the state $|0(\theta)\rangle$ is W_n. Since W_n is a decreasing monotonic function of n, the contribution of the states $|n\rangle\otimes|n\rangle$ would be suppressed for large n at finite volume. In that case, the transformation induced by the unitary operator $G^{-1}(\theta) \equiv \exp(-i\sum_\kappa \theta_\kappa \mathcal{G}_\kappa)$ could disentangle the A and B sectors. However, this is not the case in the infinite volume limit, where the summation extends to an infinite number of components and Eq. (129) holds (in such a limit Eq. (120) is only a formal relation since $G^{-1}(\theta)$ does not exist as a unitary operator) [19].

It is interesting to note that, although the mode B is related with quantum noise effects (cf. the discussion in Sect. 5), nevertheless the $A-B$ entanglement is not affected by such noise effects. The robustness of the entanglement is rooted in the fact that, once the infinite volume limit is reached, there is no unitary generator able to disentangle the $A-B$ coupling.

10 Trajectories in the $\mathcal{H}$ Space

In this section I want to discuss the chaotic behavior, under certain conditions, of the trajectories in the $\mathcal{H}$ space. Let me start by recalling some of the features of the $SU(1,1)$ group structure (see, e.g., [33]).

$SU(1,1)$ realized on $C\times C$ consists of all unimodular 2×2 matrices leaving invariant the Hermitian form $|z_1|^2 - |z_2|^2$, $z_i \in C, i = 1,2$. The complex z plane is foliated under the group action into three orbits: $X_+ = \{z : |z| < 1\}$, $X_- = \{z : |z|\rangle 1\}$ and $X_0 = \{z : |z| = 1\}$.

The unit circle $X_+ = \{\zeta : |\zeta| < 1\}$, $\zeta \equiv e^{i\phi}\tanh\theta$, is isomorphic to the upper sheet of the hyperboloid which is the set $\mathbf{H}$ of pseudo-Euclidean bounded (unit norm) vectors $\mathbf{n} : \mathbf{n}\cdot\mathbf{n} = 1$. $\mathbf{H}$ is a Kählerian manifold with metrics

$$ds^2 = 4\frac{\partial^2 F}{\partial\zeta\partial\bar{\zeta}} d\zeta \cdot d\bar{\zeta} \,, \tag{130}$$

and

$$F \equiv -\ln(1 - |\zeta|^2) \tag{131}$$

is the Kählerian potential. The metrics is invariant under the group action [33].

The Kählerian manifold **H** is known to have a symplectic structure. It may be thus considered as the phase space for the classical dynamics generated by the group action [33]. The $SU(1,1)$ generalized coherent states are recognized to be "points" in **H** and transitions among these points induced by the group action are therefore classical trajectories [33] in **H** (a similar situation occurs [33] in the $SU(2)$ (fermion) case).

Summarizing, the space of the unitarily inequivalent representations of the ccr, which is the space of the $SU(1,1)$ generalized coherent states, is a Kählerian manifold , $\mathcal{H} \equiv \{\mathcal{H}_\theta,\ \forall\theta\} \approx \mathbf{H}$; it has a symplectic structure and a classical dynamics is established on it by the $SU(1,1)$ action (generated by $\mathcal{G}$ or, equivalently, by p_θ: $\mathcal{H}_\theta \to \mathcal{H}_{\theta'}$). Variations of the θ–parameter induce transitions through the representations $\mathcal{H}_\theta = \{|0(\theta)\rangle\}$, i.e. through the physical phases of the system, the system order parameter being dependent on θ. These transitions are described as trajectories through the "points" in $\mathcal{H}$. One may then assume time-dependent θ: $\theta = \theta(t)$. For example, this is the case of dissipative systems and of non-equilibrium thermal field theories where $\theta_\kappa = \theta_\kappa(\beta(t))$, with $\beta(t) = \frac{1}{k_B T(t)}$.

It is interesting to observe that, considering the transitions $\mathcal{H}_\theta \to \mathcal{H}_{\theta'}$, i.e. $|0(\theta)\rangle \to |0(\theta')\rangle$, we have

$$\langle 0(\theta)|0(\theta')\rangle = e^{-\frac{V}{2(2\pi)^3}\int d^3\kappa F_\kappa(\theta,\theta')} \tag{132}$$

where $F_\kappa(\theta,\theta')$ is given by Eq. (131) with $|\zeta_\kappa|^2 = \tanh^2(\theta_\kappa - \theta'_\kappa)$, which shows the role played by the Kählerian potential in the motion over $\mathcal{H}$.

The result that the group action induces classical trajectories in $\mathcal{H}$ has been also obtained elsewhere [50,51] on the ground of more phenomenological considerations.

With reference to the discussion presented in Sects. 3–5, we may say that on the (classical) trajectories in $\mathcal{H}$ it is $x_+ = x_- = x_{classical}$, i.e. on these trajectories the quantum noise accounted for by y is fully shielded by the thermal bath (cf. Eq. (46)). In Sect. 5 (see [23]) it has been indeed observed that the y freedom contributes to the imaginary part of the action which becomes negligible in the classical regime, but is relevant for the quantum dynamics, namely in each of the "points" in $\mathcal{H}$ (i.e. in each of the spaces $\mathcal{H}_\theta$, for each θ) through which the trajectory goes as θ changes. Upon "freezing" the action of $G(\theta)$ (i.e. upon "freezing" the "motion" through the ui representations) the quantum features of $\mathcal{H}_\theta$, at given θ, become manifest. This relates to the 't Hooft picture [14] and to the results of Refs. [15,16] where dissipation loss in deterministic systems may manifest itself as quantum behavior (see Sect. 11).

Let me use the notation $|0(t)\rangle_\theta \equiv |0(\theta(t))\rangle$. For any $\theta(t) = \{\theta_\kappa(t), \forall\kappa\}$ it is

$$ {}_\theta\langle 0(t)|0(t)\rangle_\theta = 1 \ , \quad \forall t \ . \tag{133}$$

I will now restrict the discussion to the case in which, for any κ, $\theta_\kappa(t)$ is a growing function of time and

$$ \theta(t) \neq \theta(t') \ , \quad \forall t \neq t', \quad \text{and} \quad \theta(t) \neq \theta'(t') \ , \quad \forall t, t' \ . \tag{134}$$

Under such conditions the trajectories in $\mathcal{H}$ satisfy the requirements for chaotic behavior in classical nonlinear dynamics. These requirements are the following [52]:

i) the trajectories are bounded and each trajectory does not intersect itself.
ii) trajectories specified by different initial conditions do not intersect.
iii) trajectories of different initial conditions are diverging trajectories.

Let $t_0 = 0$ be the initial time. The "initial condition" of the trajectory is then specified by the $\theta(0)$-set, $\theta(0) = \{\theta_\kappa(0), \forall\kappa\}$. One obtains

$$ {}_\theta\langle 0(t)|0(t')\rangle_\theta \underset{V\to\infty}{\longrightarrow} 0 \ , \quad \forall\, t\,, t' \ , \quad \text{with} \quad t \neq t' \ , \tag{135}$$

provided $\int d^3\kappa \, \ln\cosh(\theta_\kappa(t) - \theta_\kappa(t'))$ is finite and positive for any $t \neq t'$.

Equation (135) expresses the unitary inequivalence of the states $|0(t)\rangle_\theta$ (and of the associated Hilbert spaces $\{|0(t)\rangle_\theta\}$) at different time values $t \neq t'$ in the infinite volume limit. The non-unitarity of time evolution, implied for example by the damping, is consistently recovered in the unitary inequivalence among representations $\{|0(t)\rangle_\theta\}$'s at different t's in the infinite volume limit.

The trajectories are bounded in the sense of Eq. (133), which shows that the "length" (the norm) of the "position vectors" (the state vectors at time t) in $\mathcal{H}$ is finite (and equal to one) for each t. Equation (133) rests on the invariance of the Hermitian form $|z_1|^2 - |z_2|^2$, $z_i \in C, i = 1, 2$ and I also recall that the manifold of points representing the coherent states $|0(t)\rangle_\theta$ for any t is isomorphic to the product of circles of radius ${r_\kappa}^2 = \tanh^2(\theta_\kappa(t))$ for any κ.

Equation (135) expresses the fact that the trajectory does not crosses itself as time evolves (it is not a periodic trajectory): the "points" $\{|0(t)\rangle_\theta\}$ and $\{|0(t')\rangle_\theta\}$ through which the trajectory goes, for any t and t', with $t \neq t'$, after the initial time $t_0 = 0$, never coincide. The requirement (i) is thus satisfied.

In the infinite volume limit, we also have

$$ {}_\theta\langle 0(t)|0(t')\rangle_{\theta'} \underset{V\to\infty}{\longrightarrow} 0 \qquad \forall\, t\,, t' \quad , \quad \forall\, \theta \neq \theta' \ . \tag{136}$$

Under the assumption (134), Eq. (136) is true also for $t = t'$. The meaning of Eq. (136) is that trajectories specified by different initial conditions $\theta(0) \neq \theta'(0)$ never cross each other. The requirement (ii) is thus satisfied.

In order to study how the "distance" between trajectories in the space $\mathcal{H}$ behaves as time evolves, consider two trajectories of slightly different initial

conditions, say $\theta'(0) = \theta(0) + \delta\theta$, with small $\delta\theta$. A difference between the states $|0(t)\rangle_\theta$ and $|0(t)\rangle_{\theta'}$ is the one between the respective expectation values of the number operator $A^\dagger_\kappa A_\kappa$. For any κ at any given t, it is

$$\begin{aligned} \Delta\mathcal{N}_{A_\kappa}(t) &\equiv \mathcal{N'}_{A_\kappa}\big(\theta'(t)\big) - \mathcal{N}_{A_\kappa}\big(\theta(t)\big) \\ &= {}_{\theta'}\langle 0(t)|A^\dagger_\kappa A_\kappa|0(t)\rangle_{\theta'} - {}_{\theta}\langle 0(t)|A^\dagger_\kappa A_\kappa|0(t)\rangle_{\theta} \\ &= \sinh^2\big(\theta'_\kappa(t)\big) - \sinh^2\big(\theta_\kappa(t)\big) = \sinh\big(2\theta_\kappa(t)\big)\delta\theta_\kappa(t)\ , \end{aligned} \tag{137}$$

where $\delta\theta_\kappa(t) \equiv \theta'_\kappa(t) - \theta_\kappa(t)$ is assumed to be greater than zero, and the last equality holds for "small" $\delta\theta_\kappa(t)$ for any κ at any given t. By assuming that $\frac{\partial\delta\theta_\kappa}{\partial t}$ has negligible variations in time, the time-derivative gives

$$\frac{\partial}{\partial t}\Delta\mathcal{N}_{A_\kappa}(t) = 2\frac{\partial\theta_\kappa(t)}{\partial t}\cosh\big(2\theta_\kappa(t)\big)\delta\theta_\kappa\ . \tag{139}$$

This shows that, provided $\theta_\kappa(t)$ is a growing function of t, small variations in the initial conditions lead to growing in time $\Delta\mathcal{N}_{A_\kappa}(t)$, namely to diverging trajectories as time evolves.

In the assumed hypothesis, at enough large t the divergence is dominated by $\exp\big(2\theta_\kappa(t)\big)$. For each κ, the quantity $2\theta_\kappa(t)$ could be thus thought to play the rôle similar to the one of the Lyapunov exponent.

Since $\sum_\kappa E_\kappa \dot{\mathcal{N}}_{A_\kappa} dt = \frac{1}{\beta} dS_A$, where E_κ is the energy of the mode A_κ and dS_A is the entropy variation associated to the modes A (cf. Eq. (40)) [25], the divergence of trajectories of different initial conditions may be expressed in terms of differences in the variations of the entropy (cf. Eqs. (137) and (139)):

$$\Delta\sum_\kappa E_\kappa \dot{\mathcal{N}}_{A_\kappa}(t)dt = \frac{1}{\beta}\big(dS'_A - dS_A\big)\ . \tag{140}$$

The discussion above thus shows that also the requirement (iii) is satisfied. The conclusion is that trajectories in the $\mathcal{H}$ space exhibit, under the condition (134) and with $\theta(t)$ a growing function of time, properties typical of the chaotic behavior in classical nonlinear dynamics.

11 Deterministic Dissipative Systems and Quantization

In Sect. 3 we have seen that the canonical quantization for the damped oscillator is obtained at the expense of introducing an "extra" coordinate y. The role of the "doubled" y coordinate is absolutely crucial in the quantum regime where it accounts for the quantum noise. When the classical solution $y = 0$ is adopted, the x system appears to be "incomplete"; the loss of information due to dissipation amounts to neglecting the bath and to the ignorance of the bath-system interaction, i.e. the ignorance of "where" and "how" energy flows out of the system. One can thus conclude that the loss of information

occurring at the classical level due to dissipation manifests itself in terms of "quantum" noise effects arising from the imaginary part of the action, to which the y contribution is crucial. This result suggests to consider the approach to dissipation presented above in connection with the proposal put forward by 't Hooft in a series of papers [14]. He proposes that Quantum Mechanics may indeed result from a more fundamental deterministic theory as an effect of a process of information loss. He considers a class of deterministic Hamiltonian systems described by means of Hilbert space techniques. The quantum systems are obtained when constraints implementing the information loss are imposed on the original Hilbert space. The Hamiltonian for such systems is of the form

$$H = \sum_i p_i \, f_i(q) \; , \tag{141}$$

where $f_i(q)$ are non-singular functions of the canonical coordinates q_i. The equations for the q's (i.e. $\dot{q}_i = \{q_i, H\} = f_i(q)$) are decoupled from the conjugate momenta p_i and this implies [14] that the system can be described deterministically even when expressed in terms of operators acting on the Hilbert space. The condition for the deterministic description is the existence of a complete set of observables commuting at all times, called *beables* [53]. For the systems of Eq. (141), such a set is given by the $q_i(t)$ [14].

In order to cure the fact that the Hamiltonians of the type (141) are not bounded from below, one might split H in Eq. (141) as [14]:

$$H = H_I - H_{II} \; , \quad H_I = \frac{1}{4\rho} \left(\rho + H\right)^2 \; , \quad H_{II} = \frac{1}{4\rho} \left(\rho - H\right)^2 \; , \tag{142}$$

where ρ is a time-independent, positive function of q_i. H_I and H_{II} are then positively (semi)definite and $\{H_I, H_{II}\} = \{\rho, H\} = 0$. Then the constraint condition is imposed onto the Hilbert space:

$$H_{II}|\psi\rangle = 0 \; , \tag{143}$$

which ensures that the Hamiltonian is bounded from below. This condition, indeed, projects out the states responsible for the negative part of the spectrum. In other words, one gets rid of the unstable trajectories [14]. In Refs. [15] and [16] it has been shown that the system of damped-antidamped oscillators discussed in Sect. 3 does provide an explicit realization of 't Hooft mechanism. In addition, it has been also shown that there is a connection between the zero point energy of the quantum harmonic oscillator and the geometric phase of the (deterministic) system of damped/antidamped oscillators. This can be seen by noticing that the Hamiltonian Eq. (25) is of the type (141) with $i = 1, 2$ and with $f_1(q) = 2\Omega$, $f_2(q) = -2\Gamma$, provided one uses a set of canonical transformations which for brevity I do not report here (see [15]). By using $J_2 = -\frac{i}{2}(J_+ - J_-)$ and $\mathcal{C} = \frac{1}{2}(A^\dagger A - B^\dagger B)$ one may write Eq. (26) as

$$H = H_I - H_{II} \quad , \quad H_I = \frac{1}{2\Omega\mathcal{C}} (2\Omega\mathcal{C} - \Gamma J_2)^2 \quad , \quad H_{II} = \frac{\Gamma^2}{2\Omega\mathcal{C}} J_2^2 \, . \tag{144}$$

Note that $\mathcal{C}$, being the Casimir operator, is a constant of motion, which ensures that once it has been chosen to be positive it will remain such at all times. The constraint (143) is now imposed by putting

$$J_2|\psi\rangle = 0\,, \tag{145}$$

and the physical states $|\psi\rangle$ are by this defined. It is now convenient to introduce

$$x_1 = \frac{x+y}{\sqrt{2}}\,, \quad x_2 = \frac{x-y}{\sqrt{2}}\,,$$

and

$$x_1 \;=\; r\cosh u\,, \quad x_2 \;=\; r\sinh u\,, \tag{146}$$

in terms of which [27]

$$\mathcal{C} \;=\; \frac{1}{4\Omega m}\left[p_r^2 - \frac{1}{r^2}p_u^2 + m^2\Omega^2 r^2\right]\,, \quad J_2 \;=\; \frac{1}{2}\,p_u\,. \tag{147}$$

Of course, only nonzero r^2 should be taken into account in order for $\mathcal{C}$ to be invertible. If one does not use the operatorial formalism, then the constraint $p_u = 0$ implies $u = -\frac{\gamma}{2m}t$. Equation (145) implies

$$H|\psi\rangle = H_I|\psi\rangle = 2\Omega\mathcal{C}|\psi\rangle = \left(\frac{1}{2m}p_r^2 + \frac{K}{2}r^2\right)|\psi\rangle\,, \tag{148}$$

where $K \equiv m\Omega^2$. H_I thus reduces to the Hamiltonian for the linear harmonic oscillator $\ddot{r} + \Omega^2 r = 0$. The physical states are even with respect to time-reversal ($|\psi(t)\rangle = |\psi(-t)\rangle$) and periodical with period $\tau = \frac{2\pi}{\Omega}$.

I will now introduce the states $|\psi(t)\rangle_H$ and $|\psi(t)\rangle_{H_I}$ satisfying the equations:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle_H = H\,|\psi(t)\rangle_H\,, \tag{149}$$

$$i\hbar\frac{d}{dt}|\psi(t)\rangle_{H_I} = 2\Omega\mathcal{C}|\psi(t)\rangle_{H_I}\,. \tag{150}$$

Equation (150) describes the two-dimensional "isotropic" (or "radial") harmonic oscillator. $H_I = 2\Omega\mathcal{C}$ has the spectrum $\mathcal{H}_I^n = \hbar\Omega n$, $n = 0, \pm1, \pm2, ...$ According to the choice for $\mathcal{C}$ to be positive, only positive values of n will be considered. The generic state $|\psi(t)\rangle_H$ can be written as

$$|\psi(t)\rangle_H = \hat{T}\left[\exp\left(\frac{i}{\hbar}\int_{t_0}^{t} 2\Gamma J_2 dt'\right)\right]|\psi(t)\rangle_{H_I}\,, \tag{151}$$

where $\hat{T}$ denotes time-ordering. Of course, here $\hbar$ is introduced on purely dimensional grounds and its actual value cannot be fixed by the present analysis.

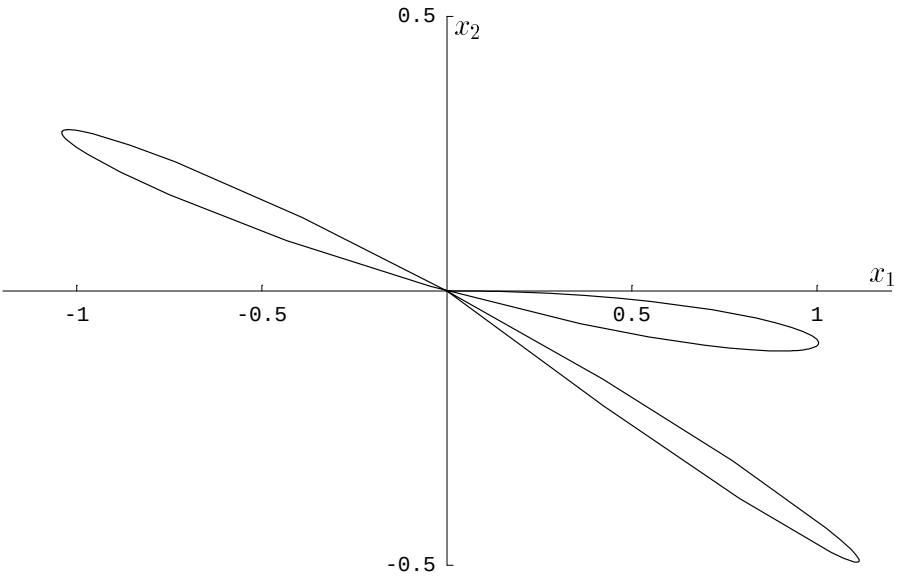

Fig. 4. Trajectories for $r_0 = 0$ and $v_0 = \Omega$, after three half-periods for $\kappa = 20$, $\gamma = 1.2$ and $m = 5$. The ratio $\int_0^{\tau/2}(\dot{x}_1 x_2 - \dot{x}_2 x_1)dt/\mathcal{E} = \pi\frac{\gamma}{m\Omega^3}$ is preserved. $\mathcal{E}$ is the initial energy: $\mathcal{E} = \frac{1}{2}mv_0^2 + \frac{1}{2}m\Omega^2 r_0^2$

One obtains [15]:

$$ {}_H\langle\psi(\tau)|\psi(0)\rangle_H = {}_{H_I}\langle\psi(0)|\exp\left(i\int_{C_{0\tau}} A(t')dt'\right)|\psi(0)\rangle_{H_I} \equiv e^{i\phi}\ , \quad (152) $$

where the contour $C_{0\tau}$ is the one going from $t' = 0$ to $t' = \tau$ and back and $A(t) \equiv \frac{\Gamma m}{\hbar}(\dot{x}_1 x_2 - \dot{x}_2 x_1)$. Note that $(\dot{x}_1 x_2 - \dot{x}_2 x_1)dt$ is the area element in the (x_1, x_2) plane enclosed by the trajectories (see Fig. 4) (cf. Sect. 6). Notice also that the evolution (or dynamical) part of the phase does not enter in ϕ, as the integral in Eq. (152) picks up a purely geometric contribution [39].

Let me consider the periodic physical states $|\psi\rangle$. Following [39], one writes

$$ |\psi(\tau)\rangle = e^{i\phi - \frac{i}{\hbar}\int_0^\tau \langle\psi(t)|H|\psi(t)\rangle dt}|\psi(0)\rangle = e^{-i2\pi n}|\psi(0)\rangle\,, \quad (153) $$

i.e. $\frac{\langle\psi(\tau)|H|\psi(\tau)\rangle}{\hbar}\tau - \phi = 2\pi n$, $n = 0, 1, 2, \ldots$, which by using $\tau = \frac{2\pi}{\Omega}$ and $\phi = \alpha\pi$, gives

$$ \mathcal{H}^n_{I,eff} \equiv \langle\psi_n(\tau)|H|\psi_n(\tau)\rangle = \hbar\Omega\left(n + \frac{\alpha}{2}\right)\ . \quad (154) $$

The index n has been introduced to exhibit the n dependence of the state and the corresponding energy. $\mathcal{H}^n_{\mathrm{i},eff}$ gives the effective nth energy level of the physical system, i.e. the energy given by $\mathcal{H}^n_{\mathrm{i}}$ corrected by its interaction with

the environment. One thus see that the dissipation term J_2 of the Hamiltonian is actually responsible for the "zero point energy" ($n = 0$): $E_0 = \frac{\hbar}{2}\Omega\alpha$.

I recall that the zero point energy is the "signature" of quantization since in Quantum Mechanics it is formally due to the non-zero commutator of the canonically conjugate q and p operators. Thus dissipation manifests itself as "quantization". In other words, E_0, which appears as the "quantum contribution" to the spectrum, signals the underlying dissipative dynamics. If one wants to match the Quantum Mechanics zero point energy, has to fix $\alpha = 1$, which gives [15] $\Omega = \frac{\gamma}{m}$.

In connection with the discussion presented in Sect. 3.1, the thermodynamical features of the dynamical rôle of J_2 can be revealed by rewriting Eq. (151) as

$$|\psi(t)\rangle_H = \hat{T}\left[\exp\left(i\frac{1}{\hbar}\int_{u(t_0)}^{u(t)} 2J_2 du'\right)\right]|\psi(t)\rangle_{H_I}\,, \tag{155}$$

where $u(t) = -\Gamma t$ has been used. Thus,

$$-i\hbar\frac{\partial}{\partial u}|\psi(t)\rangle_H = 2J_2|\psi(t)\rangle_H\,. \tag{156}$$

$2J_2$ appears then to be responsible for shifts (translations) in the u variable, as it has to be expected since $2J_2 = p_u$ (cf. Eq. (147)). One can write indeed: $p_u = -i\hbar\frac{\partial}{\partial u}$. Then, in full generality, Eq. (145) defines families of physical states, representing stable, periodic trajectories (cf. Eq. (148)). $2J_2$ implements transition from family to family, according to Eq. (156)Equation (149) can be then rewritten as

$$i\hbar\frac{d}{dt}|\psi(t)\rangle_H = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle_H + i\hbar\frac{du}{dt}\frac{\partial}{\partial u}|\psi(t)\rangle_H\,. \tag{157}$$

The first term on the r.h.s. denotes of course derivative with respect to the explicit time dependence of the state. The dissipation contribution to the energy is thus described by the "translations" in the u variable. Now I consider the derivative

$$\frac{\partial S}{\partial U} = \frac{1}{T}\,. \tag{158}$$

From Eq. (144), by using $S \equiv \frac{2J_2}{\hbar}$ and $U \equiv 2\Omega\mathcal{C}$, one obtains $T = \hbar\Gamma$. Equation (158) is the defining relation for temperature in thermodynamics (with $k_B = 1$) so that one could formally regard $\hbar\Gamma$ (which dimensionally is an energy) as the temperature, provided the dimensionless quantity S is identified with the entropy. In such a case, the "full Hamiltonian" Eq. (144) plays the role of the free energy $\mathcal{F}$: $H = 2\Omega\mathcal{C} - (\hbar\Gamma)\frac{2J_2}{\hbar} = U - TS = \mathcal{F}$. Thus $2\Gamma J_2$ represents the heat contribution in H (or $\mathcal{F}$). Of course, consistently, $\frac{\partial\mathcal{F}}{\partial T}\big|_{\Omega} = -\frac{2J_2}{\hbar}$. In conclusion $\frac{2J_2}{\hbar}$ behaves as the entropy, which is not surprising since it

controls the dissipative (thus irreversible) part of the dynamics. In this way the conclusions of Sect. 3 are reobtained. It is also suggestive that the temperature $\hbar\Gamma$ is actually given by the background zero point energy: $\hbar\Gamma = \frac{\hbar\Omega}{2}$.

Finally, I observe that

$$\left.\frac{\partial\mathcal{F}}{\partial\Omega}\right|_T = \left.\frac{\partial U}{\partial\Omega}\right|_T = mr^2\Omega\,, \tag{159}$$

which is the angular momentum, as expected since it is the conjugate variable of the angular velocity Ω.

The above results may suggest that the condition (145) can be then interpreted as a condition for an adiabatic physical system. $\frac{2J_2}{\hbar}$ might be viewed as an analogue of the Kolmogorov–Sinai entropy for chaotic dynamical systems.

Finally, I note that a reparametrization-invariant time technique in a specific model [54] also may lead to a quantum dynamics emerging from a deterministic classical evolution.

12 Conclusions

In this report I have reviewed some aspects of the algebraic structure of QFT related with the doubling of the degrees of freedom of the system under study. I have shown how such a doubling is related to the characterizing feature of QFT consisting in the existence of infinitely many unitarily inequivalent representations of the canonical (anti-)commutation relations and how this is described by the q-deformed Hopf algebra. I have considered several examples of systems and shown the analogies, or links, among them arising from the common algebraic structure of the q-deformed Hopf algebra.

I have considered the Wigner function and the density matrix formalism and shown that it requires the doubling of the degrees of freedom, which thus appears to be a basic formal feature also in Quantum Mechanics. In this connection I have considered the two-slit experiment and shown that quantum interference effects disappear in the limit of coincidence of the doubled variable $x_\pm$. Then I have shown how in QFT it is the q-deformed coproduct which is relevant and how Bogoliubov transformations are constructed in terms of it. I have considered quantum dissipation by studying the damped harmonic oscillator and the quantum Brownian motion and commented on how the arrow of time emerges from the intrinsic thermodynamic nature of dissipation. The vacuum structure is the one of the generalized coherent states. The connection (links) with the two-mode squeezed states and the noncommutative geometry in the plane emerges in a natural way in the discussion of these systems. In view of the similarity of some features of the coherent states with those of the fractals, it is an interesting question to ask whether fractal properties enter the QFT structure. A study on this point is in progress.

The relation with thermal field theory, in the thermo field dynamics formalism, reveals one further formal analogy with the systems mentioned above.

In such a contest entropy appears to be a measure of the degree of entanglement between the system and the thermal bath in which it is embedded. This also relates with the connection between the doubled variables and quantum noise effects.

For brevity, here I have not considered the doubling of the degrees of freedom in expanding geometry problems (inflationary models) and in the quantization of the matter field in a curved background. For this I refer to the papers [17–19].

Finally, I have discussed how 't Hooft proposal, according to which the loss of information due to dissipation in a classical deterministic system manifests itself in the quantum features of the system, finds a possible description in the formal frame common to the systems mentioned above. In particular, I have shown that the quantum spectrum of the harmonic oscillator can be obtained from the dissipative character of the underlying deterministic system. In recent years, the problem of quantization of a classical theory has attracted much attention in gravitation theories and in non-hamiltonian dissipative system theories, also in relation with noncommutative space-time structures involving deformation theory (see for example [45]). By taking advantage of the fact that the manifold of the QFT unitarily inequivalent representations is a Kählerian manifold, I have shown that classical trajectories in such a manifold, which may exhibit chaotic behavior under some conditions, describe (phase) transitions among the inequivalent representations. The space of the QFT representations appears thus covered by a *classical blanket*.

Acknowledgements

I thank the MIUR, INFN and the ESF Program COSLAB for partial financial support. I am grateful to the organizers of the COSLAB Dresden Workshop in July 2005 and in particular to Ralf Schuetzhold for giving to me the opportunity to report in that occasion on the matter here presented.

References

1. J. von Neumann: *Mathematical foundations of Quantum Mechanics*, (Princeton University Press, Princeton 1955)
2. H. Umezawa and G. Vitiello: *Quantum Mechanics*, (Bibliopolis, Napoli 1985)
3. O. Bratteli and D. W. Robinson: *Operator Algebras and Quantum Statistical Mechanics*, (Springer, Berlin 1979)
4. G.L. Sewell: *Quantum theory of collective phenomena*, (Clarendon Press, Oxford 1986)
5. H. Umezawa: *Advanced field theory: micro, macro and thermal concepts*, (AIP, N.Y. 1993); U. Umezawa, M. Matsumoto and M. Tachiki: *Thermo Field Dynamics and Condensed States*, (North-Holland, Amsterdam 1982)

6. V.G. Drinfeld: In *Proc. ICM Berkeley, CA*, ed. by A.M. Gleason, (AMS, Providence, R.I. 1986) p. 798; M. Jimbo, Int. J. Mod. Phys., A **4**, 3759 (1989); Yu.I. Manin: *Quantum Groups and Non-Commutative Geometry*, (CRM, Montreal 1988)
7. E. Celeghini, T.D. Palev and M. Tarlini: Mod. Phys. Lett. B **5**, 187 (1991); P.P. Kulish and N.Y. Reshetikhin: Lett. Math. Phys. **18**, 143 (1989); V. Chari and A. Pressley: *A Guide to Quantum Groups*, (Cambridge University Press, Cambridge 1994).
8. Y. Takahashi and H. Umezawa: Collective Phenomena **2** 55 (1975); reprint in Int. J. Mod. Phys. **B10**, 1599 (1996)
9. E. Celeghini, S. De Martino, S. De Siena, A. Iorio, M. Rasetti and G. Vitiello: Phys. Lett. A **244**, 455 (1998); S. De Martino, S. De Siena and G. Vitiello: Int. J. Mod. Phys. B **10**, 1615 (1996)
10. G. Vitiello: Int. J. Mod. Phys. B **18**, 785 (2004)
11. G. Vitiello: Braz. J. Phys. **35**, 351 (2005)
12. W.H. Zurek: Phys. Rev. D **24** 1516 (1981); W.H. Zurek: Phys. Rev. D **26** 1862 (1982); E. Joos and H.D. Zeh: Z. Phys. B **59** 223 (1985); D. Giulini, E. Joss, C. Kiefer, J. Kupsch, I.O. Stamatescu and H.D. Zeh: *Decoherence and the appearance of a classical world in quantum theory* (Springer, Berlin, 1996)
13. C. Itzykson and J.B. Zuber: *Quantum Field Theory*, (McGraw-Hill Inc., 1980); P.W. Anderson: *Basic Notions of Condensed Matter Physics*, (Menlo Park, Benjamin, 1984)
14. G. 't Hooft: In: *Basics and Highlights of Fundamental Physics*, (Erice 1999); G. 't Hooft: Class. Quant. Grav. **16**, 3263 (1999); G. 't Hooft: arXiv:hep-th/0105105 G. 't Hooft: Int. J. Theor. Phys., **42**, 355 (2003)
15. M. Blasone, P. Jizba and G. Vitiello: Phys. Lett., A **287**, 205 (2001)
16. M. Blasone, E. Celeghini, P. Jizba and G. Vitiello: Phys. Lett., A **310**, 393 (2003)
17. E. Alfinito, R. Manka and G. Vitiello: Class. Quant. Grav. **17**, 93 (2000); E. Alfinito and G. Vitiello: Phys. Lett. A **252**, 5 (1999)
18. M. Martellini, P. Sodano and G. Vitiello: Nuovo Cim. A **48**, 341 (1978); A. Iorio, G. Lambiase and G. Vitiello: Annals of Phys. **294**, 234 (2001)
19. A. Iorio, G. Lambiase and G. Vitiello: Annals Phys. **309**, 151 (2004)
20. J. Schwinger: J. Math. Phys. **2**, 407 (1961)
21. R.P. Feynman and F.L. Vernon: Annals of Phys., **24**, 118 (1963)
22. R.P. Feynman: *Statistical Mechanics*, (W.A. Benjamin Publ. Co., Reading, Ma. 1972)
23. Y.N. Srivastava, G. Vitiello and A. Widom: Annals Phys. **238**, 200 (1995)
24. M. Blasone, Y.N. Srivastava, G. Vitiello and A. Widom: Annals Phys. **267**, 61 (1998)
25. E. Celeghini, M. Rasetti and G. Vitiello: Annals Phys. **215**, 156 (1992)
26. H. Feshbach and Y. Tikochinsky: Trans. New York Acad. Sci., (Ser.II) **38**, 44 (1977)
27. M. Blasone, E. Graziano, O.K. Pashaev and G. Vitiello: Annals Phys., **252**, 115 (1996)
28. E. Celeghini, M. Rasetti, M. Tarlini and G. Vitiello: Mod. Phys. Lett. B **3**, 1213 (1989)
29. H. Bateman: Phys. Rev. **38**, 815 (1931); P.M. Morse and H. Feshbach: *Methods of Theoretical Physics*, Vol. I, (McGraw–Hill, New York 1953) pag. 298;

H. Dekker: Phys. Rept. **80**, 1 (1981); M. Blasone and P. Jizba, Annals Phys. **312**, 354 (2004)
30. Y. Tsue, A. Kuriyama and M. Yamamura: Prog. Theor. Phys. **91**, 469 (1994); S. Nishiyama, C. Providencia, J. da Providencia, Y. Tsue and M. Yamamura, Prog. Theor. Phys. **113**, 555 (2005)
31. V.E. Tarasov: Phys. Lett. A **288**, 173 (2001); R. Banerjee and P. Mukherjee, J. Phys. A **35** 5591 (2002); R. Banerjee, Mod. Phys. Lett. A **17** 631 (2002); A. Mostafazadeh: arXiv:math-ph/0107001 R.S. Kaushal and H.J. Korsch: Phys. Lett. A **276**, 47 (2000)
32. J.R. Klauder and E.C. Sudarshan: *Fundamentals of Quantum Optics*, (Benjamin, New York 1968)
33. A. Perelomov: *Generalized Coherent States and Their Applications*, (Springer, Berlin 1986)
34. I. Ojima, Annals Phys. (N.Y.) **137**, 1 (1981)
35. D. Cangemi, R. Jackiw and B. Zwiebach: Annals of Phys., **45**, 408 (1996)
36. S. De Filippo and G. Vitiello: Lett. Nuovo Cim. **19**, 92 (1977)
37. G. Vitiello: Int. J. Mod. Phys. B **9**, 973 (1995); E. Alfinito and G. Vitiello: Int. J. Mod. Phys. B **14**, 853 (2000). [Erratum-ibid. B **14**, 1613 (2000)]
G. Vitiello: *My Double Unveiled – The dissipative quantum model of brain*, (John Benjamins Publ. Co., Philadelphia, Amsterdam 2001).
38. S.W. Hawking and R. Penrose: Sci. Am. **275**, 44 (1996)
39. J. Anandan, Y. Aharonov: Phys. Rev. Lett. **65**, 1697 (1990)
40. S. Sivasubramanian, Y.N. Srivastava, G. Vitiello, A. Widom: Phys. Lett., A **311**, 97 (2003)
41. H.P. Yuen: Phys. Rev. **A 13**, 2226 (1976)
42. A. Iorio and G. Vitiello: Mod. Phys. Lett. B **8**, 269 (1994); A. Iorio and G. Vitiello: Annals Phys. **241**, 496 (1995)
43. E. Celeghini, M. Rasetti and G. Vitiello: Phys. Rev. Lett. **66**, 2056 (1991); E. Celeghini, S. De Martino, S. De Siena, M. Rasetti and G. Vitiello: Mod. Phys. Lett. B **7**, 1321 (1993); E. Celeghini, S. De Martino, S. De Siena, M. Rasetti and G. Vitiello: Annals Phys. **241**, 50 (1995)
44. G.V. Dunne, R. Jackiw and C.A. Trugenberger: Phys. Rev. D **41**, 661 (1990); A. Iorio and T. Sykora: Int. J. Mod. Phys. A **17**, 2369 (2002); A. Connes, M.R. Douglas and A. Schwarz: JHEP **9802**, 003 (1998); N. Seiberg and E. Witten: JHEP **9909**, 032 (1999); R. Banerjee and S. Ghosh: Phys. Lett. B **533**, 162 (2002)
45. P. Aschieri, C. Blohmann, M. Dimitrijevic, F. Meyer, P. Schupp and J. Wess: Class. Quant. Grav. **22**, 3511 (2005)
46. L.D. Landau, E.M. Lifshitz: *Quantum Mechanics*, (Pergamon Press, Oxford 1977) pp. 458 and 184
47. L.D. Faddeev, R. Jackiw: Phys. Rev. Lett. **60**, 1692 (1988)
48. H. Umezawa: In: *Banff/CAP Workshop on Thermal Field Theory*, eds. F.C.Khanna et al., (World Sci., Singapore 1994), p. 109
49. L.C. Biedenharn: J. Phys. **A 22**, L873 (1989); A.J. MacFarlane: J. Phys. A **22**, 4581 (1989)
50. R. Manka, J. Kuczynski and G. Vitiello: Nucl. Phys., B **276**, 533 (1986)
51. E. Del Giudice, R. Manka, M. Milani and G. Vitiello: Phys. Lett., A **206**, 661 (1988)

52. R. Hilborn, *Chaos and Nonlinear Dynamics*, (Oxford University Press, Oxford 1994)
53. J.S. Bell: *Speakable and Unspeakable in Quantum Mechanics*, (Cambridge University Press 1987)
54. H.T. Elze and O. Schipper: Phys. Rev. D **66**, 044020 (2002)

The Classical and Quantum Roots of Pauli's Spin-statistics Relation

B. Kuckert

II. Institut für Theoretische Physik, Luruper Chaussee 149, 22761 Hamburg, Germany

1 Introduction

The world would be a different place if spin-one-half particles were not subject to Pauli's exclusion principle. In all fundamental branches of modern (natural) science, the connection between particle spins and multiparticle behavior plays a crucial role, and to date, no physical system violating it has ever been observed.

On the other hand, the framework of quantum physics – be it relativistic or not – does admit particles with a spin-statistics relation opposite to the familiar one. In quantum mechanics, this is evident, since, e.g., scalar multiparticle wave functions that are antisymmetric under particle exchange can be defined without any problems, i.e., scalar fermions fit into the framework of quantum mechanics.

In quantum field theory, the spin-statistics connection has been established by Burgoyne for fields with Lorentz (or more precisely: $SL(2,\mathbb{C})$-) symmetry and a finite number of components (see Ref. [27] and references given there). Fields with infinitely many components, however, may violate Pauli's relation [26], and the confinement to finite-component fields has no proper physical justification.

A next step was taken by Doplicher, Haag, and Roberts [9] and generalized in Ref. [6]. They proved a spin-statistics theorem in the algebraic approach to quantum field theory. The interesting aspect was that they did this in an analysis that evaluated which particle statistics are possible at all. Particle statistics is an output, not an input. They consider single-particle states with sharp, at most finitely degenerate masses $m > 0$. Massless-particle states are excluded.

In this article we review a novel approach to the spin-statistics connection, which dates back to Refs. [16] and [11], but which makes weaker assumptions. Our analysis is based on surprisingly close ties between the classical symmetries of a quantum field and the intrinsic structure of its algebras of field operators. These ties have been discovered by Bisognano and Wichmann [1,2].

B. Kuckert: *The Classical and Quantum Roots of Pauli's Spin-Statistics Relation*, Lect. Notes Phys. **718**, 207–228 (2007)
DOI 10.1007/3-540-70859-6_8

One of them is that a uniformly accelerated observer perceives the field's vacuum state as a thermal state whose temperature is linearly proportional to the observer's acceleration; this has been discovered by Unruh [29] independently and is called the *Unruh effect*. Another one is the fact that a family of antiunitary conjugations generate reflections by spacelike planes combined with a charge conjugation. We call this *modular P_1CT-symmetry*.

This symmetry will be the main input to our analysis. Apart from leading to some most general spin-statistics theorems, this approach exhibits the beauty and impact of the links between classical and quantum physics.

Surprisingly one does not need to assume covariance with respect to a representation of $SL(2,\mathbb{C})$ from the outset. It suffices to assume modular P_1CT-symmetry in order to *construct* a representation of $SL(2,\mathbb{C})$ with respect to which the field is covariant. This representation can easily be seen to satisfy Pauli's spin-statistics relation. If P_1CT symmetry is only known to hold for some given timelike vector, one can still construct a representation of $SU(2)$, which satisfies the spin-statistics relation as well.

Several assumptions that have been needed for the older spin-statistics theorems can be dropped; these include the restriction to finite-component fields, but also the positivity of the energy ("spectrum condition"), covariance (which is proved, not assumed), and a couple of technicalities.

This way, the input to the analysis is reduced to a minimum of physically well motivated and standard assumptions, and at the same time, the output is much more general than the older results.

In Sect. 2, the properties of quantum fields that are relevant for our analysis will be specified. As a concrete (though finite-component) example, the hermitian scalar field in $1+1$ spacetime dimensions is constructed in order to illustrate the assumptions we make. Then the general standing assumptions are formulated; first for finite-component fields, then for general fields.

In Sect. 3, the Unruh effect and the Bisognano-Wichmann theorem are discussed. The general notion of state and KMS-state are introduced, and basic features of the mathematical theory of Tomita and Takesaki are worked out.

In Sect. 4, it is shown how a covariant and "well behaved" representation of $SU(2)$ (or $SL(2,\mathbb{C})$) can be constructed from the modular P_1CT-conjugations. To this end, simply connected covering groups $\mathbf{G}_\mathrm{R}$ and $\mathbf{G}_\mathrm{L}$ of $SO(3)$ and L_1, respectively, are constructed (in a classical setting). The section is concluded with a remark on the (straightforward) construction of a full PCT-operator from modular P_1CT-conjugations.

The chapter ends with a conclusion summarizing the results, commenting on related earlier results in the same spirit, and finishing with some remarks on remaining open problems.

2 Setting

The analysis to follow is a general and abstract one, and this requires the specification of a corresponding setting by means of basic assumptions. These will be formulated in this section. Before doing so, we discuss an elementary example of a quantum field: the hermitian scalar field in $1+1$ spacetime dimensions, which is a solution of the Klein-Gordon equation. The construction of higher dimensional quantum fields works the same way.

2.1 The Klein-Gordon Equation

If $e_0 \in \mathbb{R}^{1+1}$ is a future-directed timelike unit vector, and if e_1 is the unique spacelike unit vector with $e_0 e_1 = 0$ that "points to the right," then coordinates x_0 and x_1 on $\mathbb{R}^{1+1}$ are defined by $x_0(q) := qe_0$ and $x_1(q) := qe_1$. The partial differential operator

$$\Box_x := \partial_{x_0}^2 - \partial_{x_1}^2 \ .$$

does not depend on the choice of e_0.

The Fourier transform of the Klein-Gordon equation

$$(\Box + m^2)u = 0 \ , \tag{1}$$

where $m > 0$ is a given mass, is[1]

$$(-p^2 + m^2)\hat{u}(p) = 0 \ . \tag{2}$$

As a consequence, the support of $\hat{u}$ has to be a subset of the hyperbola $H_m \subset \mathbb{R}^{1+1}$ specified by the condition $p^2 = m^2$. One connected component of H_m consists of positive-energy vectors only; it is called the *upper mass shell* H_m^+. The elements of H_m^+ are the 4-momenta of classical relativistic point particles.

Denote by L_1 the *restricted Lorentz group*, i.e., the connected component of the Lorentz group containing its unit element.[2] In $1+1$ dimensions, L_1 coincides with the one-parameter Abelian group $B(\chi)$, $\chi \in \mathbb{R}$, of boosts. H_m^+ is an orbit of L_1 without fixed points. So if one chooses any point $p' \in H_m^+$, then there is, for each $p \in H_m^+$, a unique $\chi(p) \in \mathbb{R}$ with $p = B(\chi(p))p'$. By construction, $\chi(B(\xi)p) = \chi(p) + \xi$, so the measure $d\chi$ on H_m^+ is invariant under boosts and does note depend on the choice of p'.

For each $p \in H_m^+$, the plane wave $q \mapsto e^{\pm ipq}$ on $\mathbb{R}^{1+1}$ is a classical solution of the Klein-Gordon equation. The Klein-Gordon equation is linear, so if a_+ and a_- are, say, integrable functions on H_m^+, then

$$F(q) := \int_{H_m^+} \left(a_+(p)e^{-ipq} + a_-(p)e^{ipq}\right) \, d\chi(p) \tag{3}$$

[1] $\hat{u}(p) := (2\pi)^{-1} \int_{\mathbb{R}^{1+1}} u(q)e^{-ipq} \, d^2x(q)$, where pq is the Lorentzian inner product of the points $q, p \in \mathbb{R}^{1+1}$.

[2] This definition can be transferred to $1+3$ dimensions.

is a solution of the Klein-Gordon equation as well. If the functions $a_\pm$ are not integrable, the field F may still be well defined as a distribution. As an example, put $a_\pm \equiv (2\pi)^{-1}$, then

$$F(q) = (2\pi)^{-1} \int_{H_m^+} (e^{-ipq} + e^{ipq})\, d\chi(p) = \pi^{-1} \int_{H_m^+} \cos(pq)\, d\chi(p) =: \Phi(q)\ , \tag{4}$$

and for $a_\pm \equiv \pm(2\pi i)^{-1}$, F equals

$$F(q) = (2\pi i)^{-1} \int_{H_m^+} (e^{-ipq} - e^{ipq})\, d\chi(p) = \pi^{-1} \int_{H_m^+} \sin(pq)\, d\chi(p) =: \Delta(q)\ , \tag{5}$$

where, as usual, the integrations are to be read accordingly.[3]

2.2 The Hermitian Scalar Field

Quantum fields are obtained by "plugging" classical field equations and their solutions into the well-known second quantization procedure. This procedure replaces the complex (or, more generally speaking, finite-dimensional vector) field values by linear operators in an infinite-dimensional Hilbert space, namely, a Fock space. We briefly figure this procedure out here in order to illustrate what are the structures we generalize. The Hilbert space of the hermitian scalar field is constructed from wave functions that are considered as the wave functions of one or several particles of mass m. The *single-particle wave functions* are the elements of the Hilbert space $\mathcal{H}_1 := L^2(H_m^+, d\chi)$. Put the vacuum (zero-particle) space $\mathcal{H}_0$ equal to $\mathbb{C}$, define the vacuum vector $\Omega := 1 \in \mathcal{H}_0$, and define the N-particle space $\mathcal{H}_N$ as the Hilbert space of *symmetric* wave functions in $L^2((H_m^+)^N, d^N\chi)$, i.e., all wave functions ψ with

$$\psi(p_{\pi(1)} \cdots p_{\pi(N)}) = \psi(p_1 \cdots p_N)$$

for all permutations $\pi \in S_N$. The bosonic Fock space $\mathcal{H}$ is defined by $\mathcal{H} := \bigoplus_{N\in\mathbb{N}} \mathcal{H}_N$. The subspace $\mathcal{D} := \bigcup_{M\in\mathbb{N}} \bigoplus_{0\le N\le M} \mathcal{H}_N$ is called the *finite-particle space*. The definition of the N-particle wave functions as symmetric functions endows the field with a Bose–Einstein statistics.

To each wave function $\phi \in \mathcal{H}_1$, assign a *creation operator* $a^+(\phi)$ by

$$a^+(\phi)\psi := C_N\, \phi \otimes_s \psi, \qquad \psi \in \mathcal{D}\ ,$$

[3] In precise terms: Φ assigns to each test function φ on $\mathbb{R}^{1+1}$ the value

$$\Phi(\varphi) := \int_{H_m^+} (\hat\varphi(p) + \check\varphi(p)\,)\ d\chi(p) = \pi^{-1} \int_{H_m^+} \int_{\mathbb{R}^{1+1}} \cos(pq)\varphi(q)\, d^2x(q)\, d\chi(p)$$

where $\check\varphi$ is the inverse Fourier transform; both $\hat\varphi$ and $\check\varphi$ are smooth and decrease rapidly at infinity, so their restrictions to H_m^+ are well defined and integrable. Δ is defined the same way.

where $\otimes_s$ denotes the symmetrized tensor product and where C_N is a constant we specify later on. Spelled out,

$$(a^+(\phi)\psi)(p_1 \cdots p_N) = \frac{C_N}{N} \sum_\nu \phi(p_\nu)\psi(p_{\pi(1)} \cdots \hat{p}_\nu \cdots p_{\pi(N)}) \,, \tag{6}$$

where the hat symbol indicates omission of the argument. This defines $a^+(\phi)$ as a linear operator on the finite-particle space $\mathcal{D}$.

The adjoint operator $a(\phi) := a^+(\phi)^*$ is called an *annihilation operator*; it assigns to each $\psi \in \mathcal{H}_N$, $N \geq 1$, the wave function $a(\phi)\psi \in H_{N-1}$ defined by

$$(a(\phi)\psi)(p_1, \ldots, p_{N-1}) := C_N \int_{H_m^+} \overline{\phi(p)}\, \psi(p_1 \cdots p_{N-1}, p)\, d\chi(p) \,;$$

together with $a(\phi)\Omega := 0$, this suffices to specify $a(\phi)$ on $\mathcal{D}$.

Annihilation operators can also be defined for sharp momenta. Namely, one can define to each $p \in H_m^+$ the annihilation operator $a(p)$ assigning to each $\psi \in \mathcal{H}_N$, $N \geq 1$, the wave function $a(p)\psi \in \mathcal{H}_{N-1}$ given by

$$(a(p)\psi)(p_1 \cdots p_{N-1}) := C_N \psi(p, p_1 \cdots p_{N-1}), \qquad \psi \in \mathcal{H}_N \,,$$

and assigning $0 \in \mathcal{H}$ to Ω. $a(p)$ is, like $a(\phi)$, well defined on the finite-particle space $\mathcal{D}$ as an operator, but its hermitian adjoint is ill-defined as an operator, since the symmetric tensor product of a wave function by a delta function is no wave function.

Given any single-particle wave functions $\psi, \phi \in \mathcal{H}_1$, the commutators $[a(\psi), a(\phi)]$ and $[a^+(\psi), a^+(\phi)]$ vanish by construction. It is customary to choose the constants C_N in such a fashion that creation and annihilation operators exhibit the commutation relation

$$[a(\phi), a^+(\psi)] = \langle \phi, \psi \rangle \,, \tag{7}$$

which requires $C_N = \sqrt{N}$. With this choice, all creation and annihilation operators are unbounded, i.e., they are not continuous.

When defining the hermitian scalar field as an operator valued distribution, it must be taken into account that an annihilation operator $a(\phi)$ depends on its argument ϕ in an *antilinear* fashion. The dependence is, however, $\mathbb{R}$-linear, and one can define the scalar field as a $\mathbb{C}$-linear distribution in two steps.

For each *real*-valued test function φ on $\mathbb{R}^{1+1}$, define

$$\Phi(\varphi) := a(\hat{\varphi}|_{H_m^+}) + a^+(\hat{\varphi}|_{H_m^+}) \,,$$

then one can define for an arbitrary complex-valued φ

$$\Phi(\varphi) := \Phi(\mathrm{Re}(\varphi)) + i\Phi(\mathrm{Im}(\varphi))$$

(cf. Eq. (4)). Φ is called the *hermitian scalar field of mass* m.

One verifies by straightforward computation that

$$[\Phi(q), \Phi(q')] = i\Delta(q - q') \tag{8}$$

(cf. Eq. (5)), which is to be read as an equation of distributions.[4] The distribution Δ vanishes outside the light cone, i.e., $\Delta(q) = 0$ if $q^2 < 0$. Namely, the integrand in Eq. (5) is odd with respect to some $p' \in H_m^+$ if q is spacelike. Note that $pq > 0$ for all $p \in H_m^+$ if $q \in \bar{V}_+$. The consequence of this is called *microcausality*: field operators located in spacelike separated regions commute (for the hermitian scalar field).

2.3 The General Setup

The analysis to follow concerns a very large class of quantum field theories. The framework will be fixed by specifying a couple of standing assumptions. There are several standard properties of quantum fields we do not need to assume here from the outset (covariance, positivity of the energy, etc.). For a discussion of the general setting of quantum field theory, see, e.g., [13, 22, 27]. For the reader's convenience, we first spell out our assumptions for finite-component fields.

(a) **The field.** We consider an **n-component quantum field** F consisting of n distinct linear maps $F_1 \cdots F_n$ that assign to each test function $\varphi \in C_0^\infty(\mathbb{R}^{1+3})$ linear operators $F_1(\varphi) \cdots F_n(\varphi)$ in a Hilbert space $\mathcal{H}$ with domains of definition $D(F_1(\varphi)) \cdots D(F_n(\varphi))$. There exists a dense subspace $\mathcal{D}$ of $\mathcal{H}$ with $\mathcal{D} \subset D(F_\nu(\varphi)) \cap D(F_\nu(\varphi)^*)$ and $F_\nu(\varphi)\mathcal{D} \cup F_\nu(\varphi)^*\mathcal{D} \subset \mathcal{D}$ for all indices ν.[5]

In the above case of the hermitian scalar field (like for other fields in Fock spaces), the space $\mathcal{D}$ is the finite-particle space.

For each open region $\mathcal{O} \subset \mathbb{R}^{1+3}$, the field operators $F_\nu(\varphi)$ with $\operatorname{supp}\varphi \subset \mathcal{O}$, their hermitian adjoints, and the unit operator generate a *field algebra* $\mathcal{F}(\mathcal{O})$ of unbounded operators defined on $\mathcal{D}$. In general, an algebra $\mathfrak{A}$ with an antilinear map $* : \mathfrak{A} \to \mathfrak{A}$ with $(A^*)^* = A$ for all $A \in \mathfrak{A}$ (like the hermitian adjoint in the present case) is called a $*$-*algebra*.[6]

In order to make the problem well posed, we need a notion of statistics and a notion of spin. The statistics of $\mathcal{F}$ is assumed to be abelian Bose–Einstein or Fermi-Dirac statistics in the following way.

[4] spelled-out: $[\Phi(\varphi), \Phi(\varphi')] = i(\Delta * \varphi')(\varphi)$; here $\Delta * \varphi'$ is the convolute of the distribution Δ and the test function φ'; this convolute is a distribution one then applies to the test function φ. We refrain from spelling out these expressions from now on.

[5] It is common (and not costly) to assume continuity and the existence of a Fourier transform in addition, but these properties will not be used in what follows, so we omit them here.

[6] A map like $*$ is called an *antilinear involution*; $*$-algebras are also called *involutive algebras*.

(b) **Microcausality (Bose–Fermi alternative).** If φ and ψ are test functions with spacelike separated supports, then

$$F_\nu(\varphi)F_\mu(\psi)|_{\mathcal{D}} = \pm F_\mu(\psi)F_\nu(\varphi)|_{\mathcal{D}} \ ,$$

where the sign on the right-hand side is positive if F has Bose–Einstein statistics and negative if F has Fermi-Dirac statistics.

Microcausality is closely related to Einstein causality, but not even in the bosonic case do the two notions coincide. Einstein causality requires that any two observables located in spacelike separated regions are commensurable in the sense that their *spectral measures* commute.[7] But fields with Fermi-Dirac statistics are not observables, and not even for bosonic fields does microcausality imply Einstein causality. The sign on the right-hand side does, however, specify the statistics of the field. For the hermitian scalar field, both Einstein causality and microcausality hold.

The assumption of microcausality is a crucial difference with the algebraic approach one finds summarized in [13]. There the input is a system of *observables* satisfying Einstein causality in the proper sense, and a field structure and its possible statistics are *constructed* from these data. It should be remarked that not only Bose and Fermi statistics, but also parastatistics are admitted, which is ruled out here from the outset. The present approach can, however, be transferred to the algebraic one in a straightforward fashion.

A standard property of quantum fields is that they describe excitations of a vacuum state vector Ω. One usually characterizes Ω by its property that it is invariant under a unitary representation of $SL(2,\mathbb{C})$ in $\mathcal{H}$ acting on the field in a covariant fashion. However, we will not assume, but *establish* the existence of a covariant representation of a symmetry group. So the usual characterization of the vacuum vector is not appropriate for our purposes.

But the vacuum vector has another important property. Namely, the subspace $\mathcal{F}(\mathcal{O})\Omega$ of $\mathcal{H}$ is dense for each open region $\mathcal{O} \subset \mathbb{R}^{1+1}$ by a well known theorem due to Reeh and Schlieder. One says that Ω is *cyclic* with respect to each $\mathcal{F}(\mathcal{O})$. If $\mathcal{O}$ is an open cone, the assumptions made in the Reeh-Schlieder theorem can be relaxed considerably. So excitations of the vacuum vector by field operators located in $\mathcal{O}$ are not to be considered as state vectors of a particle located in $\mathcal{O}$, since they are not perpendicular to the excitations by field operators located outside $\mathcal{O}$.

We use this important property to characterize the vacuum state from the outset.

(c) **Vacuum vector.** There exists a vector $\Omega \in \mathcal{D}$ that is cyclic with respect to $\mathcal{F}(\mathcal{O})$ assigned to each open set $\mathcal{O} \subset \mathbb{R}^{1+3}$.

Note that there may be many such vectors.

[7] It does, in general *not* suffice that the observables A and B commute in the sense that $AB = BA$ cf., e.g., Nelson's counterexample discussed in [25].

The notion of vacuum we use here is considerably more general than the usual one for the time being, but assuming modular $\mathrm{P_1CT}$-symmetry below in Sect. 4 will, eventually, imply that Ω has the familiar invariance properties.

As announced earlier, our analysis will also apply for infinite-component fields. For such fields, the preceding assumptions look a bit different. Denote by $\mathfrak{C}$ a linear space of arbitrary dimension (the coordinate space, which is finite dimensional for finite-component fields) and by $\mathfrak{D}$ the space of test functions in $\mathbb{R}^{1+3}$. The above assumptions are generalized as follows.

(A) **The field.** The quantum field F with component space $\mathfrak{C}$ is a linear function that assigns to each $\xi \in \mathfrak{C} \otimes \mathfrak{D}$ a linear operator $F(\xi)$ in a separable Hilbert space $\mathcal{H}$.

(A.1) F is free from redundancies in $\mathfrak{C}$, i.e., if $\mathfrak{c}, \mathfrak{d} \in \mathfrak{C}$ and if $F(\mathfrak{c} \otimes \varphi) = F(\mathfrak{d} \otimes \varphi)$ for all test functions $\varphi \in \mathfrak{D}$, then $\mathfrak{c} = \mathfrak{d}$.

(A.2) Each field operator $F(\xi)$ and its adjoint $F(\xi)^*$ are densely defined. There exists a dense subspace $\mathcal{D}$ of $\mathcal{H}$ with $\mathcal{D} \subset D(F(\xi)) \cap D(F(\xi)^*)$ and $F(\xi)\mathcal{D} \cup F(\xi)^*\mathcal{D} \subset \mathcal{D}$ for all $\xi \in \mathfrak{C} \otimes \mathfrak{D}$.

For each open region $\mathcal{O} \subset \mathbb{R}^{1+3}$, denote by $\mathcal{F}(\mathcal{O})$ the algebra generated by all $F(\mathfrak{c} \otimes \varphi)|_{\mathcal{D}}$ and all $F(\mathfrak{c} \otimes \varphi)^*|_{\mathcal{D}}$ with $\mathrm{supp}(\varphi) \subset \mathcal{O}$. Defining an involution $*$ on $\mathcal{F}(\mathcal{O})$ by $A^* := A^*|_{\mathcal{D}}$, each algebra $\mathcal{F}(\mathcal{O})$ is endowed with the structure of a $*$-algebra.

(A.3) $\mathcal{F}(\mathcal{O})$ is nonabelian for each $\mathcal{O}$, and if $\mathcal{O} \neq \mathcal{P}$, then $\mathcal{F}(\mathcal{O}) \neq \mathcal{F}(\mathcal{P})$.

(B) **Microcausality.** There exists a unitary and self-adjoint operator k on $\mathcal{H}$ with $k\Omega = \Omega$ and with $k\mathcal{F}(\mathcal{O})k = \mathcal{F}(\mathcal{O})$ for all $\mathcal{O}$. Define $F_\pm := \frac{1}{2}(F \pm kFk)$. If $\mathfrak{c}$ and $\mathfrak{d}$ are arbitrary elements of $\mathfrak{C}$ and if $\varphi, \psi \in \mathfrak{D}$ have spacelike separated supports, then

$$\begin{aligned} F_+(\mathfrak{c} \otimes \phi)F_+(\mathfrak{d} \otimes \psi) &= F_+(\mathfrak{d} \otimes \psi)F_+(\mathfrak{c} \otimes \phi)\,, \\ F_+(\mathfrak{c} \otimes \phi)F_-(\mathfrak{d} \otimes \psi) &= F_-(\mathfrak{d} \otimes \psi)F_+(\mathfrak{c} \otimes \phi)\,, \quad \text{and} \\ F_-(\mathfrak{c} \otimes \phi)F_-(\mathfrak{d} \otimes \psi) &= -F_-(\mathfrak{d} \otimes \psi)F_-(\mathfrak{c} \otimes \phi) \end{aligned}$$

for all $\mathfrak{c}, \mathfrak{d} \in \mathfrak{C}$. The involution k is the *statistics operator*, and $F_\pm$ are the bosonic and fermionic components of F, respectively. Defining $\kappa := (1 + ik)/(1 + i)$ and $F^t(\mathfrak{d} \otimes \psi) := \kappa F(\mathfrak{d} \otimes \psi)\kappa^*$, the normal commutation relations read

$$\big[F(\mathfrak{c} \otimes \varphi), F^t(\mathfrak{d} \otimes \psi)\big] = 0\,.$$

This property is referred to as *twisted locality*. Denote $\mathcal{F}^t(\mathcal{O}) := \kappa\mathcal{F}(\mathcal{O})\kappa^*$.

(C) **Vacuum vector.** There exists a vector $\Omega \in \mathcal{D}$ that is cyclic with respect to $\mathcal{F}(\mathcal{O})$ for each open set $\mathcal{O} \subset \mathbb{R}^{1+3}$.

3 The Unruh Effect and the Bisognano-Wichmann Theorem

In this section, a property of quantum fields will be discussed that is closely related to the Hawking effect. It has been discovered by Unruh [29] and,

independently, by Bisognano and Wichmann [1,2] that a uniformly accelerated observer in Minkowski space perceives the vacuum state of a quantum field as a thermal state with a temperature that is proportional to his acceleration (Unruh effect). In addition, Bisognano and Wichmann showed that there is a close tie between PCT-symmetry and the intrinsic algebraic structure of the field (modular P_1CT-symmetry). The second phenomenon will play a crucial role in the next section.

3.1 States of Quantum Systems

Consider an arbitrary quantum system described by a $*$-algebra $\mathfrak{A}$ of operators in a Hilbert space $\mathcal{H}$, and suppose that $\mathfrak{A}$ contains the unit operator 1. Typically, the elements of $\mathfrak{A}$ are observables or field operators. A *state* of the system is a linear functional ω on $\mathfrak{A}$ assigning to each $A \in \mathfrak{A}$ an expectation value $\omega(A) \in \mathbb{C}$ in such a fashion that $\omega(A^*A) \geq 0$ and that $\omega(1) = 1$. The states of $\mathfrak{A}$ are a convex set $\mathfrak{S}$; if ω_0 and ω_1 are states and if $0 < t < 1$, then a *mixed state* ω_t can be defined by convex combination:

$$\omega_t(A) := t\omega_0(A) + (1-t)\omega_1(A) \; .$$

States that are not mixtures of two other states are called *pure states* of $\mathfrak{A}$. In elementary quantum mechanics, pure states ω_ψ are assigned to normalized wave functions ψ by

$$\omega_\psi(A) := \langle \psi, A\psi \rangle \; ,$$

and mixed states are assigned to density matrices ρ by

$$\omega_\rho(B) := \frac{\mathrm{tr}(\rho B)}{\mathrm{tr}(\rho)} \; ,$$

where the domain of A should contain ψ and where B should be bounded.

Next endow the quantum system described by $\mathfrak{A}$ with a Hamiltonian H, and denote the time evolution on $\mathfrak{A}$ by $A_t := e^{itH}Ae^{-itH}$. The quantum dynamical system defined this way can have thermal equilibrium states. As an example, consider a finite quantum system with the property that all $A \in \mathfrak{A}$ are bounded. Then the *Gibbs state* ω_β at inverse temperature β assigns to each $A \in \mathfrak{A}$ the expectation value

$$\omega_\beta(A) := \frac{\mathrm{tr}(e^{-\beta H}A)}{\mathrm{tr}(e^{-\beta H})} =: \mathcal{Z}^{-1}\mathrm{tr}(e^{-\beta H}A) \; .$$

This notion, however, is subject to a severe constraint: for *infinitely* extended systems the trace of the operator $e^{-\beta H}$ becomes infinite, so ω_β is ill defined. But there is a characteristic property of Gibbs states that "survives" the transition to infinite systems. Namely,

$$\begin{aligned}\omega_\beta(A_tB) &= \mathcal{Z}^{-1}\mathrm{tr}(e^{-\beta H}\, e^{itH} A e^{-itH}\, B)\\ &= \mathcal{Z}^{-1}\mathrm{tr}(e^{i(t+i\beta)H} A e^{-itH}\, e^{\beta H} e^{-\beta H}\, B)\\ &= \mathcal{Z}^{-1}\mathrm{tr}(e^{-\beta H} B\, e^{i(t+i\beta)H} A e^{-i(t+i\beta)H})\\ &= \omega_\beta(BA_{t+i\beta})\ .\end{aligned}$$

The first and the last term do not depend on how ω_β is defined. In general, a state ω is called a *KMS-state* with inverse temperature β if there is, for any given $A, B \in \mathfrak{A}$, a continuous $\mathbb{C}$-valued function f on the strip $\{-\beta \leq \Im z \leq 0\}$ that is analytic on the interior of the strip and satisfies $f(t) = \omega(A_tB)$ and $f(t - i\beta) = \omega(BA_t)$. This criterion does not refer to the concrete way how ω is defined. Its relevance for physics has been discovered by Kubo, Martin, and Schwinger [15,23] on the one hand and by Haag, Hugenholtz, and Winnink [14] on the other. It has been shown by Pusz and Woronowicz (for C*-algebras) that the KMS-condition describes precisely the thermal equilibrium states in the sense of the fundamental laws of thermodynamics [24]. The zero-temperature states are the ground states of the Hamiltonian.

3.2 N-point Functions

For each $N \in \mathbb{N}$, the state functional induced by the vacuum state vector Ω assigns to each product $\Phi(\varphi_1), \ldots, \Phi(\varphi_N)$ the expectation value

$$w_N(\varphi_1, \ldots, \varphi_N) := \langle \Omega, \Phi(\varphi_1) \cdots \Phi(\varphi_N)\Omega \rangle\ .$$

The distributions w_N defined this way are called the *N-point functions*.

The N-point functions of the hermitian scalar field vanish for all odd N. For $N = 2$, one has

$$w_2(q_1, q_2) = \Delta_+(q_1 - q_2) := \frac{1}{2\pi} \int_{H_m^+} e^{-ip(q_1 - q_2)}\, d\chi(p) \tag{9}$$

and, recursively,

$$w_N(q_1 \cdots q_N) = \sum_{\nu=2}^{N} \Delta_+(q_1 - q_\nu) w_{N-2}(q_2 \cdots \hat{q}_\nu \cdots q_N)\ .$$

These w_N are invariant under boosts.

In general, the N-point functions of quantum fields are the basic input of the Feynman rules and, hence, the data from which scattering cross sections are computed. Conversely, quantum fields are completely fixed by their N-point functions by Wightman's reconstruction theorem (cf., e.g., [22, 27]).

3.3 Rindler Wedges and the Unruh Effect

The Unruh effect is perceived by a uniformly accelerated observer, and the spacetime region W such an observer can interact with – i.e., send a signal

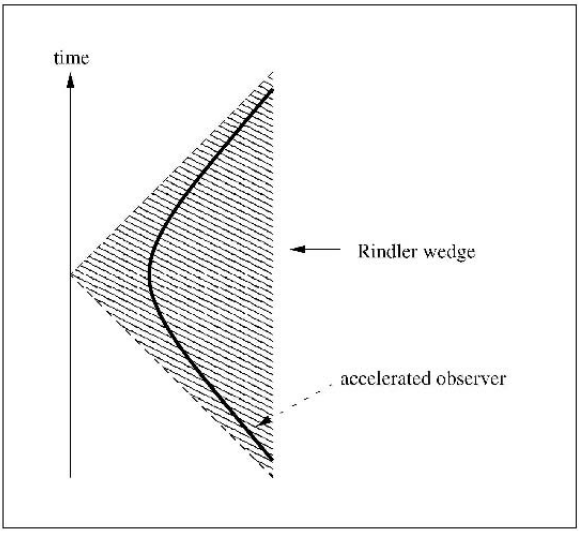

Fig. 1. Space-time diagram of Rindler wedge

and receive a reply – is called a *Rindler wedge* see Fig. 1. The edge of W is a plane S,[8] and the set S' of points that are spacelike separated from S has two connected components, one of which is W. In what follows we consider wedges in the set $\mathcal{W}_0$ of Rindler wedges whose edges contain the origin, and without loss of relevant information, we can and will confine ourselves to the case of $1+1$ dimensions, where $\mathcal{W}_0$ contains precisely two wedges $\pm W$ with the edge $S = \{0\}$. For arbitrary Minkowski coordinates x_0, x_1, the wedge W at the right hand of $\{0\}$ is

$$W := \{q \in \mathbb{R}^{1+1} : x_1(q) > |x_0(q)|\} .$$

The *Rindler coordinates* ρ and σ on W s are obtained from the Cartesian coordinates by $\rho := \sqrt{-x^2}$ and $\sigma := \operatorname{artanh}(x_0/x_1)$, respectively. Conversely, the Cartesian coordinates are obtained from the Rindler coordinates by $x_0 = \rho \sinh \sigma$ and $x_1 = \rho \cosh \sigma$, respectively. Endowing W with the metric $\rho^2 d\sigma^2 - d\rho^2$, one obtains the two-dimensional *Rindler spacetime*. A geodesic in this spacetime has the form $\{q \in W : \rho(q) = \alpha\}$ for some fixed $\alpha > 0$. Embedded into Minkowski space, this is the trajectory of an observer uniformly accelerated by the acceleration $a = \alpha^{-1}$.

Consider two events q and r on this hyperbola with $\sigma(q) = -\sigma(r) =: \sigma$, then $t := 2\sigma/a$ is the time elapsing between these events according to the observer's clock.

Written in Rindler coordinates, the two-point function of the hermitian scalar field reads, by eq. (9),

[8] in other than three spatial dimensions: an affine subspace of codimension 2.

$$w_2(q,r) = \frac{1}{2\pi}\int_{\mathbb{R}} d\chi\, \exp\big(2ima^{-1}\cosh\chi\,\sinh(at/2)\big) =: f(t).$$

f has the property that

$$\begin{aligned} f(t+2\pi i/a) &= \frac{1}{2\pi}\int_{\mathbb{R}} d\chi\, \exp\big(2ima^{-1}\cosh\chi\,\sinh(at/2+i\pi)\big) \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} d\chi\, \exp\big(-2ima^{-1}\cosh\chi\,\sinh(at/2)\big) \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} d\chi\, \exp\big(2ima^{-1}\cosh\chi\,\sinh(-at/2)\big) \\ &= f(-t) = w_2(q_2,q_1). \end{aligned}$$

One can conclude from this that the vacuum state is a KMS-state with inverse temperature $2\pi/a$, which is the inverse Unruh temperature in natural units. In order to make the argument complete and precise, one needs to use the fact that the two-point function is invariant under both spacetime translations and Lorentz boosts. Note that the Unruh temperature does not depend on m. The Unruh effect and the Unruh temperature just obtained are typical features of a large class of quantum fields, as will be discussed now.

3.4 The Bisognano-Wichmann Theorem

The Unruh effect was – independently and at the same time – also discovered by Bisognano and Wichmann, but in a different and more general formulation and with a different scope. Again, we confine ourselves to the hermitian scalar field, but this time, we consider the case of $1+3$ dimensions, since we need it for later use.

As is well known, there is a unitary representation U of the restricted Lorentz group $L_1(\mathbb{R}^{1+3}) \equiv L_+^{\uparrow}(\mathbb{R}^{1+3})$ in $\mathcal{H}$ with the properties

$$U(\lambda)\Phi(\varphi)U(\lambda)^* = \Phi(\lambda\varphi)\ ,$$

where $(\lambda\varphi)(x) = \varphi(\lambda^{-1}x)$, and $U(\lambda)\Omega = \Omega$. It is defined by

$$U(\lambda)\Phi(\varphi)\Omega := \Phi(\lambda(\varphi))\Omega\ ;$$

note that the set of all vectors $\Phi(\varphi)\Omega$, $\varphi \in C_0^\infty(\mathbb{R}^{1+3})$, is dense.

On the other hand, we know from Jost's PCT-theorem that there exists an antilinear operator Θ – called the *PCT-operator* – with the properties

$$\Theta^2 = 1\ , \tag{10}$$

$$\Theta\Omega = \Omega\ , \tag{11}$$

$$\Theta U(\lambda)\Theta = U(\lambda)\ , \qquad \text{and} \tag{12}$$

$$\Theta\Phi(\varphi)\Theta = \Phi(x \mapsto \overline{\varphi(-x)})\ . \tag{13}$$

Consider a Rindler wedge $W \in \mathcal{W}_0$, denote by $B_W(\chi)$, $\chi \in \mathbb{R}$, the unique one-parameter group of boosts with $B_W(\chi)W = W$ and by $R_W(\alpha)$, $\alpha \in S^1$, the Abelian group of rotations with $R_W(\alpha)W = W$. Define $\theta_W := \Theta R_W(\pi)$.

Bisognano and Wichmann discovered that B_W has an analytic continuation to a representation of the complex Lorentz group and that

$$\theta_W B_W(\pi i)\Phi(\varphi)\Omega = \Phi(\varphi)^*\Omega \tag{14}$$

for all test functions φ with support in W.

The bridge of this observation to physics is built by an important branch of the theory of operator algebras, which has been developed by Tomita and Takesaki [3, 28].

Ω is cyclic with respect to $\mathcal{F}(-W)$ by Assumption (C). This has the consequence that no nonzero field operator in the field algebra $\mathcal{F}(W)$ can annihilate Ω, i.e., Ω is *separating* with respect to each $\mathcal{F}(W)$. Namely, if $A \in \mathcal{F}(W)$ satisfies $A\Omega = 0$ and $B \in \mathcal{F}(-W)$, then $AB\Omega = BA\Omega = 0$, whence $A = 0$ follows by Assumption (C).

The fact that Ω is both cyclic and separating with respect to $\mathcal{F}$(W) implies that an antilinear operator $\underline{S}_W$ can be defined on the domain $\mathcal{F}(W)\Omega$ by

$$\underline{S}_W A\Omega := A^*\Omega \ .$$

It is, though nontrivial, not difficult to show that $\underline{S}_W$ can be extended to an operator $S_W := \underline{S}_W^{**}$, which is called the *Tomita operator* of Ω and $\mathcal{F}(W)$. The operator $\Delta_W := S_W^* S_W$ is a selfadjoint operator with positive spectrum; it is called the *modular operator* of Ω and $\mathcal{F}(W)$. Furthermore, the operator $J_W := S_W \Delta_W^{-1/2}$ is antiunitary, and $J_W^2 = 1$; it is called the *modular conjugation* of Ω and $\mathcal{F}(W)$. The modular operator and conjugation yield the *polar decomposition* $S_W = J_W \Delta_W^{1/2}$ of the Tomita operators, i.e., the decompositions into the "length" $\Delta_W^{1/2}$ and the "phase" J_W.

A glance at Eq. (14) shows that $\Delta_W^{1/2} = B_W(\pi i)$ and $\theta_W = J_W$. As a consequence, $B_W(\chi) = \Delta^{-it}$ for all $t \in \mathbb{R}$.

One can now use the properties of J_W and Δ_W in order to show that the one-parameter group $B(\chi)$, $t \in \mathbb{R}$, implements a time evolution with respect to which the vacuum state is a KMS-state. We first prove a simple lemma.

Lemma 1. *Let K be a unitary or antiunitary operator in $\mathcal{H}$ such that $K\mathcal{D} = \mathcal{D}$ and $K\Omega = \Omega$, and suppose there is a Rindler wedge X such that $K\mathcal{F}(W)K^* = \mathcal{F}(X)$. Then $KJ_WK^* = J_X$, and $K\Delta_W K^* = \Delta_X$.*

Proof. If $A \in \mathcal{F}(X)$, then

$$KS_W K^* A\Omega = KS_W \underbrace{K^* A K}_{\in \mathcal{F}(W)} \Omega = A^*\Omega = S_X A\Omega \ .$$

So S_W has the polar decomposition $KS_WK^* = KJK^* \cdot K\Delta^{1/2}K^*$. But the polar decomposition of an operator is unique. □

It follows from this lemma that J_W and Δ_W^{it} commute for all $t \in \mathbb{R}$, and since $J_W\Omega = \Omega = \Delta_W^{it}\Omega$ by construction, one computes

$$\begin{aligned}
\left\langle \Omega, \left(\Delta_W^{-it} A \Delta_W^{it}\right) B\Omega \right\rangle &= \left\langle \Omega, A\Delta_W^{it} B\Omega \right\rangle = \left\langle A^*\Omega, \Delta_W^{it}(B^*)^*\Omega \right\rangle \\
&= \left\langle \left(J_W \Delta_W^{1/2} A\right)\Omega, \Delta_W^{-it}\left(J_W \Delta_W^{1/2} B^*\right)\Omega \right\rangle \\
&= \left\langle J_W\left(\Delta_W^{1/2} A\Omega\right), J_W\left(\Delta_W^{it}\Delta_W^{1/2} B^*\Omega\right) \right\rangle \\
&= \left\langle \Delta_W^{it}\Delta_W^{1/2} B^*\Omega, \Delta_W^{1/2} A\Omega \right\rangle \\
&= \left\langle \Omega, B\left(\Delta_W^{1/2}\Delta_W^{-it}\Delta_W^{1/2}\right) A\Omega \right\rangle \\
&= \left\langle \Omega, B\Delta_W^{-i(t+i)} A\Omega \right\rangle \\
&= \left\langle \Omega, B\left(\Delta_W^{-i(t+i)} A \Delta_W^{i(t+i)}\right)\Omega \right\rangle \ .
\end{aligned}$$

So with respect to the dynamics induced by Δ_W^{-it}, $t \in \mathbb{R}$, the vacuum state is a KMS-state with inverse temperature 1. But by the Bisognano-Wichmann theorem, $\Delta_W^{1/2} = B_W(\pi i)$, so $\Delta_W = B_W(-2\pi i)$, and $\Delta_W^{-it} = B_W(2\pi t)$. As above, this yields the Unruh effect.

In addition, Bisognano and Wichmann have shown that $J_W = \theta_W$, i.e., the operator θ_W, which we call a P_1CT-*operator*, is a modular conjugation. We call this phenomenon *modular* P_1CT-*symmetry*. In a more general setting, where the Bisognano-Wichmann theorem does, as it stands, not hold in general, Guido and Longo have shown that if one *assumes* the Unruh effect, one can *derive* modular P_1CT-symmetry [11]. Accordingly, modular P_1CT-symmetry is the more general principle. As it stands, it makes sense even if the modular group does not implement Lorentz boosts or if the theory is not even Lorentz covariant.

The proofs of Bisognano and Wichmann apply to arbitrary finite-component quantum fields satisfying standard conditions. To be precise: they apply to precisely those Wightman fields for which Burgoyne has proved his spin-statistics theorem.

But like the spin-statistics issue, the Bisognano-Wichmann symmetries and the Unruh effect are meaningful notions in more general settings. The symmetry that will be exploited in what follows is modular P_1CT-symmetry.

4 Modular P_1CT-symmetry, the Spin-statistics Connection, and Modular PCT-symmetry

Assuming modular P_1CT-symmetry in addition to the assumptions (A)–(C) specified in Sect. 2, one can *construct* representations of $SL(2,\mathbb{C})$ in $\mathfrak{C}$ and $\mathfrak{D}$ with respect to which the field is covariant, and these representations are easily seen to satisfy Pauli's spin-statistics relation.

Modular P_1CT-symmetry neither entails the Unruh effect nor covariance with respect to $SL(2,\mathbb{C})$, and the notion of spin is well defined once a representation of $SU(2)$ has been specified. If one assumes P_1CT-symmetry only for wedges whose edge is perpendicular to a distinguished timelike vector e_0, then already this input suffices to construct representations of $SU(2)$ in $\mathcal{H}$ and $\mathfrak{C}$ with respect to which the field is covariant, and like in the above case, these representations can very easily be verified to satisfy Pauli's spin-statistics connection. So in contradistinction to the old spin-statistics theorems, the argument to be outlined now does *not* require full Lorentz symmetry! This is possible because our result is merely an *existence* result. It does not forbid pathological representations in appropriate settings, and there is nothing wrong about this – as mentioned above: there are according examples, even with full Loventz symmetry.

The groups $SL(2,\mathbb{C})$ and $SU(2)$ are relevant because they are simply connected covering groups of the spacetime symmetry groups L_1 and $SO(3)$, respectively. Since all simply connected covering groups are isomorphic, any other simply connected covering groups can replace $SL(2,\mathbb{C})$ and $SU(2)$, and this will be customary in the subsequent analysis.

The notion of covering group and the construction of our alternative covering groups will be the topic of the next subsection. Then a subsection with sketches of the proofs of our two spin-statistics theorems follows, and a subsection with a brief construction of a full PCT-operator from modular P_1CT-conjugations will conclude the section.

4.1 Covering Groups of $SO(3)$ and L_1

Consider arbitrary topological spaces X and Y. A continuous map f from X onto Y is called an *N-sheeted covering map*, $N \in \mathbb{N} \cup \infty$ if for each $y \in Y$, there are precisely N elements x of X with $f(x) = y$ and if for each $x \in X$ there exists a neighborhood U of x with the property that the restriction $f|_U$ has a continuous inverse from $f(U)$ onto U. Endowed with the map f, the space X is called a *covering space* of Y.

Recall that a topological space is called *pathwise connected* if any two of its elements can be connected by a curve and that it is called *simply connected* if in addition, each curve can be continuously contracted to any of its points. The unit circle S^1 is pathwise connected, but not simply connected, whereas the unit sphere S^2 is simply connected. It is also well known (though less evident) that $SO(3)$ is pathwise, but not simply connected, whereas $SU(2)$ is simply connected. $SU(2)$ is a two-sheeted simply connected covering space of $SO(3)$; the covering map f is well known from textbooks. Since $f(AB) = f(A)f(B)$ and $f(1) = 1$, the group $SU(2)$ is called a *covering group* of $SO(3)$.

Each pathwise connected Lie group G has a simply connected covering group $\widetilde{G}$. In order to see this, denote by $\underline{G}$ the class of all curves $c : [0,1] \to G$ starting at $c(0) = 1$ and ending at an arbitrary $c(1) \in G$. Two curves $c, d \in \underline{G}$ are called *homotopic* if $c(1) = d(1)$ and if c can be deformed into d continuously

and keeping the endpoint $c(1)$ fixed.[9] Homotopy is an equivalence relation, so let $\widetilde{G}$ be the quotient space of $\underline{G}$ by this relation, i.e., the set of equivalence classes. Evidently, $\widetilde{G}$ is a group when endowed with the group product $[c][d] = [cd]$, where $[c]$ denotes the equivalence class of $[c]$. It is straightforward to verify that $\widetilde{G}$ is simply connected and that a covering map f can be defined by $f([c]) := c(1)$. If H is another simply connected covering group of G with covering map g, then there is a bijection $\gamma : H \to G$ with $\gamma(AB) = \gamma(A)\gamma(B)$, $\gamma(1) = 1$, and $g(A) = f(\gamma(A))$, i.e., the structures of H and G as covering groups are identical. Given a concrete group H, there typically are many choices for γ, none of which needs to be particularly "natural." Given G, there are, in general, many simply connected covering groups H. Since $\widetilde{G}$ has been constructed in an intrinsic fashion – which also applies for a large class of pathwise connected topological spaces without manifold structure – it is called the *universal covering group* of G.

For the example of $SO(3)$, the group $SU(2)$ is a simply connected covering group, but it is not the universal covering group in this topologists' terminology. Another simply connected covering group is constructed as follows.

For each $a \in S^2 \subset \mathbb{R}^3$, denote by j_a the orthogonal reflection by the two-dimensional subspace $a^\perp$. For each pair $(a, b) \in S^2 \times S^2$, the linear transformation $j_a j_b$ is a rotation. But there are uncountably many other pairs (c, d) with $j_a j_b = j_c j_d$, so neither rotations nor elements of a covering group can be represented by pairs in $S^2 \times S^2$.

There is, however, an equivalence relation that yields a quotient space doing the job. Namely, call elements (a, b) and (c, d) of $S^2 \times S^2$ equivalent if $j_a j_b = j_c j_d$ and if there exists a rotation σ having the same axis as $j_a j_b$ and satisfying $(\sigma a, \sigma b) = (c, d)$. In this case, (a, b) is *not* equivalent to $(c, -d)$ Fig. 2. Denote the equivalence class of (a, b) by $\pi(a, b)$. Denote by $\mathbf{G}_R$ the quotient space $S^2 \times S^2 / \sim$.

Define a map $\tilde{\rho} : \mathbf{G}_R \to SO(3)$ by $\tilde{\rho}(\pi(a, b)) := j_a j_b$.[10]

Theorem 1. *$\tilde{\rho}$ is a covering map, and $\mathbf{G}_R$ is simply connected.*

We refer to Ref. [20] for the proof. The theorem implies that there exists a unique group product on $\mathbf{G}_R$ with $\tilde{\rho}(gh) = \tilde{\rho}(g)\tilde{\rho}(h)$ and $\rho(1) = 1$, so $\mathbf{G}_R$ is a covering group.

The same construction yields a simply connected covering of L_1 as well. First note that the above equivalence relation constituting $\mathbf{G}_R$ can be reformulated as follows: (a, b) and (c, d) are equivalent if $j_a j_b = j_c j_d$ and if there exists a rotation τ commuting with $j_a j_b$ and satisfying $(c, d) = (\tau^2 a, \tau^2 b)$. In this form, it can be transferred to the construction of $\mathbf{G}_L$.

[9] i.e., there exists a continuous map $h : [0, 1] \times [0, 1] \to G$ with $h(0, s) = 1$, $h(1, s) = c(1)$, $h(t, 0) = c(t)$, and $h(t, 1) = d(t)$.

[10] Here and below there occur tilde symbols that appear redundant. This is the notation of the original paper [21], where they are not redundant. The symbol ρ, e.g., is used for the notation $\rho(a, b) := j_a j_b$.

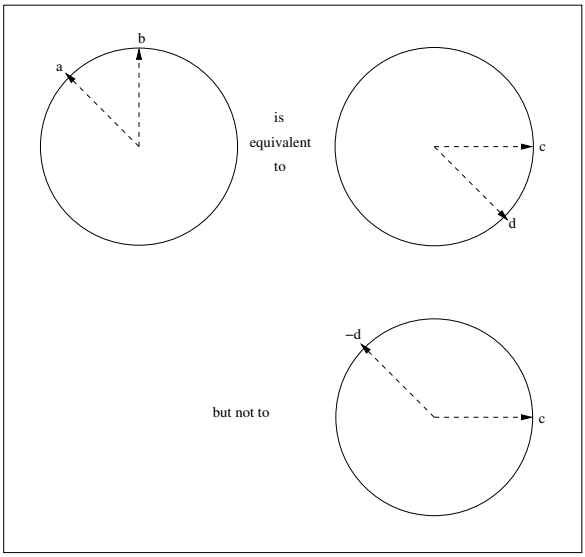

Fig. 2. The equivalence relation $\sim$

Assign to each Rindler wedge $W \in \mathcal{W}_0$ the orthogonal reflection j_W by the edge of W. For each pair $(W, X) \in \mathcal{W}_0 \times \mathcal{W}_0$, the transformation $j_W j_X$ is a restricted Lorentz transformation , and it can be shown that each restricted Lorentz transformation can be written as such a pair (this is less evident than for $SO(3)$). Let pairs (W, X) and (Y, Z) satisfy the relation $(W, X) \sim (Y, Z)$ if $j_W j_X = j_Y j_Z$ and if there exists a Lorentz transformation $\nu \in L_1$ commuting with $j_W j_Y$ and satisfying $(\nu^2 W, \nu^2 X) = (Y, Z)$. This relation can be shown to be an equivalence relation. Let $\mathbf{G}_L$ be the quotient space $\mathbf{G}_L := \mathcal{W}_0 \times \mathcal{W}_0 / \sim$.

A map $\tilde{\lambda} : \mathbf{G}_L \to L_1$ is defined by $\tilde{\lambda}(\pi(W, X)) := j_W j_X$. Like $\tilde{\rho}$, this map can be shown to be a covering map, but the proof is considerably much more involved. $\mathbf{G}_L$ is simply connected and can, like $\mathbf{G}_R$, be endowed with the structure of a simply connected covering group.

4.2 The Spin-statistics Connection

We now first construct representations of $\mathbf{G}_R$ in $\mathfrak{C}$ and $\mathcal{H}$ that can be shown to satisfy the spin-statistics connection by an elementary argument. Some remarks on the corresponding analysis for $\mathbf{G}_L$ will follow at the end of this subsection.

If e_0 is a future-directed timelike vector, then the set of spacelike unit vectors perpendicular to e_0 is a 2-sphere $S^2_{e_0} =: S^2$, and one can assign to each $a \in S^2$ the Rindler wedge

$$W_a := \{q \in \mathbb{R}^{1+3} : qa \geq |qe_0|\} .$$

For each $a \in S^2$, denote by j_a be the orthogonal reflection by the plane $a^\perp \cap e_0^\perp$,[11] and for each test function $\varphi \in \mathfrak{D}$, define $j_a\varphi \in \mathfrak{D}$ by $j_a\varphi(x) := \varphi(j_a x)$.

In addition to the above assumptions, we assume that the modular conjugations J_a belonging to the vacuum vector Ω and the field algebras $\mathcal{F}(W_a)$, $a \in S^2$, act as $\mathrm{P}_1\mathrm{CT}$-operators. In the general formulation, this reads:

(D) **Partial modular $\mathrm{P}_1\mathrm{CT}$-symmetry.** For each $a \in S^2$, there exists an antilinear involution C_a in $\mathcal{C}$ such that for all $\mathfrak{c} \in \mathcal{C}$ and $\varphi \in \mathfrak{D}$, one has

$$J_a F(\mathfrak{c} \otimes \varphi) J_a = F^t(C_a\mathfrak{c} \otimes \overline{j_a\varphi}) \,.$$

The map $a \mapsto J_a$ is continuous, i.e., if $a_\nu \to_{\nu\to\infty} a$ in S^2, then $J_{a_\nu}\Psi \to J_a\Psi$ for each $\Psi \in \mathcal{H}$.

The next task is to establish covariance of the field with respect to $\mathbf{G}_R$ by constructing appropriate representations.

Theorem 2. *(i) If $(a,b) \sim (c,d)$, then $J_aJ_b = J_cJ_d =: \tilde{W}(\pi(a,b))$.*

(ii) The map $\tilde{W}$ from $\mathbf{G}_R$ into the group of unitary operators in $\mathcal{H}$ is a group representation, i.e., $\tilde{W}(g)\tilde{W}(h) = \tilde{W}(gh)$.

(iii) There exists a representation $\tilde{D}$ of $\mathbf{G}_R$ in $\mathfrak{C}$ with

$$\tilde{W}(g)F(\mathfrak{c} \otimes \phi)\tilde{W}(g)^* = F(\tilde{D}(g)\mathfrak{c} \otimes \tilde{\rho}(g)\phi) \quad \forall g, \mathfrak{c}, \phi \,, \tag{15}$$

where $(\tilde{\rho}(g)\phi)(x) := \phi(\tilde{\rho}(g)^{-1}x)$.

Proof. The first two statements are elementary, cf. [20].

It remains to construct $\tilde{D}$. If $(a,b) \sim (c,d)$, then modular $\mathrm{P}_1\mathrm{CT}$-symmetry implies

$$\begin{aligned} F(C_aC_b\mathfrak{c} \otimes j_aj_b\phi) &= J_aJ_bF(\mathfrak{c} \otimes \phi)J_bJ_a \\ &= J_cJ_dF(\mathfrak{c} \otimes \phi)J_dJ_c \\ &= F(C_cC_d\mathfrak{c} \otimes j_cj_d\phi) \\ &= F(C_cC_d\mathfrak{c} \otimes j_aj_b\phi) \end{aligned}$$

for all $\mathfrak{c}$ and all ϕ. Using Assumption (A.1), one obtains $C_aC_b\mathfrak{c} = C_cC_d\mathfrak{c}$ for all $\mathfrak{c}$, so a map $\tilde{D} : \mathbf{G}_R \to GL(\mathcal{C})$ is defined by $\tilde{D}(\pi(a,b)) := C_aC_b$. This map $\tilde{D}$ now inherits the representation property from $\tilde{W}$.

Now the proof of our spin-statistics theorem is elementary algebra.

Theorem 3 (Spin-statistics connection).

$$F_\pm(\mathfrak{c} \otimes \varphi) = \frac{1}{2}(1 \pm F(\tilde{D}(-1)\mathfrak{c} \otimes \varphi))$$

for all $\mathfrak{c}$ and all φ.

[11] I.e., the linear reflection with $j_a a = -a$, $j_a e_0 = -e_0$, and $j_a x = x$ for all $x \in a^\perp \cap e_0^\perp$.

Proof. Lemma 1 implies, by definition of k and κ above, the relations

$$kJ_ak = J_a, \quad \text{whence} \quad J_a\kappa = \kappa^* J_a \ .$$

Using twisted locality, one also obtains

$$\kappa J_a \kappa^* = \kappa^* J_a \kappa = J_{-a} \ .$$

For each $a \in S^2$ one has

$$\tilde{W}(-1) = J_a J_{-a} = J_a \kappa J_a \kappa^* = J_a^2 (\kappa^*)^2 = k \ ,$$

so

$$\begin{aligned} F(\mathfrak{c} \otimes \varphi) &= kF(\mathfrak{c} \otimes \varphi)k = \tilde{W}(-1)F(\mathfrak{c} \otimes \varphi)\tilde{W}(-1) \\ &= \tilde{W}(-1)F(\mathfrak{c} \otimes \varphi)\tilde{W}(-1)^* = F(\tilde{D}(-1)\mathfrak{c} \otimes \varphi) \ . \end{aligned}$$

□

If, in particular, $\tilde{D}$ is irreducible with spin s, then $\tilde{D}(-1) = e^{2\pi i s}$, so $F_- = 0$ for integer s and $F_+ = 0$ for half-integer s.

In order to obtain a symmetry with respect to $\mathbf{G}_L$, one needs a stronger P_1CT-symmetry, which is straightforward to write down:

(D') **Full modular $\mathbf{P}_1$CT-symmetry.** For each $W \in \mathcal{W}_0$, there exists an antilinear involution C_W in $\mathcal{C}$ such that for all $\mathfrak{c} \in \mathcal{C}$ and $\varphi \in \mathfrak{D}$, one has

$$J_W F(\mathfrak{c} \otimes \varphi) J_W = F^t(C_W \mathfrak{c} \otimes \overline{j_W \varphi}) \ .$$

The map $W \mapsto J_W$ is continuous, i.e., if $W_\nu \to_{\nu\to\infty} W$ in $\mathcal{W}_0$, then $J_{W_\nu}\Psi \to J_W\Psi$.

Like above, one can construct representations $\tilde{W}$ and $\tilde{D}$ with the desired properties. For bosonic fields, the unitary representation $\tilde{W}$ in $\mathcal{H}$, which, actually, is a representation of L_1, has been constructed earlier by Buchholz, Dreyer, Florig, and Summers [5, 7, 10].

The proof of the spin-statistics relation is the same.

4.3 PCT-symmetry

In order to justify the term "modular P_1CT-symmetry", one should show that this condition yields, at least in $1+3$ dimensions, a full PCT-operator in a base-independent fashion.

Theorem 4 (PCT-symmetry). *There exists an antiunitary involution Θ with the properties*

(i) $J_a J_b J_c = \Theta$ for each right-handed orthogonal basis (a, b, c) of $e_0^\perp$.

(ii) There exists an antilinear involution C such that

$$\Theta F(\mathfrak{c} \otimes \varphi)\Theta = F^t(C\mathfrak{c} \otimes PT\overline{\varphi}) \ ,$$

where $(PT\overline{\varphi})(x) := \overline{\varphi(-x)}$.

Proof. Let (a', b', c') be a second right-handed orthonormal base, and define $\Theta' := J_{a'} J_{b'} J_{c'}$. Then it follows from modular symmetry that

$$\begin{aligned} \Theta'\Theta F(\mathfrak{c} \otimes \varphi)\Omega &= \Theta'\Theta F(\mathfrak{c} \otimes \varphi)\Theta\Theta'\Omega \\ &= F(C_{a'} C_{b'} C_{c'} C_a C_b C_c \mathfrak{c} \otimes \varphi)\Omega \\ &= F(\tilde{D}(1)\mathfrak{c} \otimes \varphi)\Omega \\ &= F(\mathfrak{c} \otimes \varphi)\Omega \,. \end{aligned}$$

Since Ω is cyclic, this implies the statement. □

If (a, b, c) is right-handed and (a', b', c') is left-handed, then $\tilde{D}(1)$ has to be replaced by $\tilde{D}(-1)$ in the above computation. Since $J_{-a} J_{-b} J_{-c} = \kappa J_a J_b J_c \kappa^*$, this is no surprise.

5 Conclusion

Pauli's spin-statistics relation links the spin of a quantum field, which is a classical quantity, to its statistics, which exhibits quantum theoretical peculiarities.

Each rotation can be obtained by combining two reflections by two-dimensional subspaces, and each restricted Lorentz transformation can be obtained by combining two reflections by two-dimensional *spacelike* subspaces. These facts can be used in order to construct simply connected covering groups $\mathbf{G}_R \sim SU(2)$ and $\mathbf{G}_L \sim SL(2, \mathbb{C})$ of the rotation and the restricted Lorentz group, respectively, in a purely classical fashion.

Considering fields with modular $\mathrm{P}_1\mathrm{CT}$-symmetry (which is a standard symmetry), the constructions of $\mathbf{G}_R$ and $\mathbf{G}_L$ can be "quantized" in order to obtain unitary representations of these groups in the fields' Hilbert spaces by multiplying pairs of $\mathrm{P}_1\mathrm{CT}$-operators. The representations obtained this way can easily be seen to exhibit Pauli's spin-statistics relation. It is the close link of the $\mathrm{P}_1\mathrm{CT}$-operators to the field's algebraic structure that links the classical objects $\mathbf{G}_R$ and $\mathbf{G}_L$ to their covariant representations in the quantum theoretical setting.

A couple of remarks concerning the closely related literature are in place now. The first proof of the spin-statistics connection using modular $\mathrm{P}_1\mathrm{CT}$-symmetry was given by the author in [16]. It uses the Doplicher-Haag-Roberts framework, whose input is a system of observables and whose output is a system of field algebras similar to the above ones and a first-kind gauge symmetry group [8]. Modular $\mathrm{P}_1\mathrm{CT}$-symmetry needs to be assumed only for the observables; then it can be *derived* for the field. On the other hand, the argument only applies to a restricted class of particle states, as already mentioned in the introduction.

In this aspect, the result of Guido and Longo [12] is superior. They just start with a system of field algebras (like we have done above), not caring on how they have been constructed and which masses are incorporated in the theory. They assume the Unruh effect directly for this algebra, and using an earlier result together with Brunetti [4], they prove the existence of a covariant representation with the correct spin-statistics relation. The authors also have succeeded in generalizing this result to conformal field theory [12]. As a remarkable spin-off, they derive modular P_1CT-symmetry from the Unruh effect.

On the other hand, Guido and Longo make the stronger symmetry assumption (as their derivation of modular P_1CT-symmetry from the Unruh effect shows), and they do so for the full field and not only for the observable input. The Unruh effect itself already entails full Lorentz symmetry, whereas partial modular P_1CT-symmetry merely entails rotational symmetry. The result discussed above generalizes the results in Refs. [16] and [11].

The next challenge would be the construction of a simply connected covering $\mathbf{G}_P$ of the restricted *Poincaré* group P_1 not from pairs, but from quadruples of reflections by arbitrary spacelike planes. Another situation where more than two P_1CT-conjugations are needed is met in higher-dimensional spacetimes in order to generalize the above strategy. These problems are under present investigation.

References

1. Bisognano, J.J., Wichmann, E.H.: On the Duality Condition for a Hermitean Scalar Field, *J. Math. Phys.* **16**, 985–1007 (1975)/
2. Bisognano, J.J., Wichmann, E.H.: On the Duality Condition for Quantum Fields, *J. Math. Phys.* **17**, 303 (1976)
3. Bratteli, O. Robinson, D.W.: *Operator Algebras and Quantum Statistical Mechanics*, I, Springer 1987
4. Brunetti, R., Guido, D., Longo, R.: Modular Structure and Duality in Conformal Quantum Field Theory, *Commun. Math. Phys.* **156**, 201–219 (1993)
5. Buchholz, D., Dreyer, O., Florig, M., Summers, S.J.: Geometric Modular Action and spacetime Symmetry Groups, *Rev. Math. Phys.* **12** 475–560 (2000)
6. Buchholz, D., Epstein, H.: Spin and Statistics for Quantum Topological Charges, *Fizika* **17**, 329–343 (1985)
7. Buchholz, D., Summers, S.J.: An algebraic characterization of vacuum states in Minkowski space. III. Reflection maps. Comm. Math. Phys. **246**, 625–641 (2004)
8. Doplicher, S., Roberts, J.E.: Why there is a field algebra with a compact gauge group describing the superselection structure in particle physics, *Commun. Math. Phys.* **131**, 51–107 (1990)
9. Doplicher, S., Haag, R., Roberts, J.E.: Local Observables and Particle Statistics II, *Commun. Math. Phys.* **35**, 49–85 (1974)
10. Florig, M.: Geometric Modular Action, PhD-thesis, University of Florida, Gainesville, 1999

11. Guido, D., Longo, R.: An Algebraic Spin and Statistics Theorem, *Commun. Math. Phys.* **172**, 517–534 (1995)
12. Guido, D., Longo, R.: The Conformal Spin and Statistics Theorem, *Commun. Math. Phys.* **181**, 11–36 (1996)
13. Haag, R.: *Local Quantum Physics*, Berlin, Heidelberg, New York: Springer 1992
14. Haag, R.: Hugenholtz, N.M. and Winnink, M.: On the equilibrium states in quantum statistical mechanics, Commun. Math. Phys. **5**, 215–236 (1967).
15. Kubo, R.: Statistical mechanical theory of irreversible processes I, *J. Phys. Soc. Japan* **12**, 570–586 (1957).
16. Kuckert, B.: A New Approach to Spin & Statistics, *Lett. Math. Phys.* **35**, 319–335 (1995)
17. Kuckert, B.: Covariant Thermodynamics of Quantum Systems: Passivity, Semipassivity, and the Unruh Effect. *Ann. Phys. (N. Y.)* **295**, 216–229 (2002)
18. Kuckert, B.: Spin & Statistics in Nonrelativistic Quantum Mechanics, I, *Phys. Lett. A* **332**, 47–53 (2004)
19. Kuckert, B., Mund, J.: Spin & Statistics in Nonrelativistic Quantum Mechanics, II, *Ann. Phys. (Leipzig)* **14**, 309–311 (2005)
20. Kuckert, B.: Spin, Statistics, and Reflections, I, *Ann. H. Poincaré* **6**, 849–862 (2006)
21. Kuckert, B., Lorenzen, R.: Spin, Statistics, and Reflections, II, *Commun. Math. Phys.* **269**, 809–831
22. Kuckert, B.: Axiomatic Quantum Field Theory, Francoise, J.-P., Naber, G.L. Tsun, T.S. (eds.) *Encyclopedia of Mathematical Physics*, Amsterdam: Elsevier / Academic Press, 2006
23. Martin, P.C., Schwinger, J.: Theory of many-particle systems I, Phys. Rev. **115**, 1342–1373 (1959).
24. Pusz, W., Woronowicz, S.L.: Passive states and KMS-states for general quantum systems, Commun. Math. Phys. **58**, 273–290 (1978).
25. Reed, M., Simon, B.: Methods of Modern Mathematical Physics, I., Functional Analysis, Academic Press 1974
26. Streater, R.F.: Local Field with the Wrong Connection Between Spin and Statistics, *Commun. Math. Phys.* **5**, 88–98 (1967)
27. Streater, R.F., Wightman, A.S.: PCT, Spin & Statistics, and All That, Benjamin, 1964
28. Takesaki, M.: *Tomita's Theory of Modular Hilbert Algebras and Its Applications*, Lecture Notes in Mathematics **128**, Springer 1970 (New York)
29. Unruh, W.G.: Notes on black-hole evaporation. *Phys. Rev.* **D14**, 870–892 (1976)

Black Hole Lasers Revisited

U. Leonhardt and T. G. Philbin

School of Physics and Astronomy, University of St. Andrews, North Haugh, St. Andrews, Fife KY16 9SS, Scotland
ulf@st-andrews.ac.uk, tgp3@st-andrews.ac.uk

Abstract. The production of Hawking radiation by a single horizon is not dependent on the high-frequency dispersion relation of the radiated field. When there are two horizons, however, Corley and Jacobson have shown that superluminal dispersion leads to an amplification of the particle production in the case of bosons. The analytic theory of this "black hole laser" process is quite complicated, so we provide some numerical results in the hope of aiding understanding of this interesting phenomenon. Specifically, we consider sonic horizons in a moving fluid. The theory of elementary excitations in a Bose-Einstein condensate provides an example of "superluminal" (Bogoliubov) dispersion, so we add Bogoliubov dispersion to Unruh's equation for sound in the fluid. A white-hole/black-hole horizon pair will then display black hole lasing. Numerical analysis of the evolution of a wave packet gives a clear picture of the amplification process. By utilizing the similarity of a radiating horizon to a parametric amplifier in quantum optics we also analyze the black hole laser as a quantum-optical network.

1 Introduction

One important contribution of the theoretical study of black hole analogues has been to help clarify the derivation of the Hawking effect [1–3]. This in turn led to a study [4] of Hawking radiation in a more general context, one that involves, among other features, two horizons. The results of [4] are of particular interest because they offer a scenario, perhaps realizable in a black hole analogue, in which the Hawking radiation is amplified. Let us begin by describing the background to these ideas.

There is an apparent contradiction in Hawking's semiclassical derivation of black hole evaporation [5], in that the radiated fields undergo arbitrarily large blue-shifting in the calculation, thus acquiring arbitrarily large masses, which contravenes the underlying assumption that the gravitational effects of the quantum fields may be ignored. This is known as the trans-Plankian problem [1–3,6,7]. A similar issue arises in condensed matter analogues such as

U. Leonhardt and T.G. Philbin: *Black Hole Lasers Revisited*, Lect. Notes Phys. **718**, 229–245 (2007)
DOI 10.1007/3-540-70859-6_9

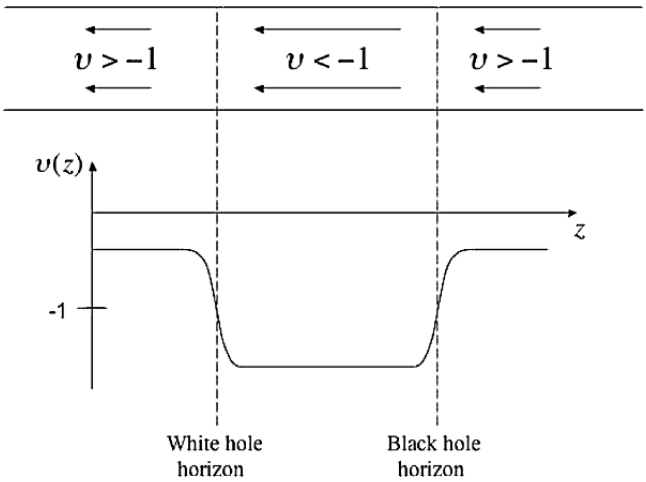

Fig. 1. Sonic horizons in a moving fluid, in which the speed of sound is 1. The velocity profile of the fluid, $v(z)$, attains the value -1 at two values of z; these are horizons for sound waves that are right-moving with respect to the fluid. At the right-hand horizon right-moving waves are trapped, with waves just to the left of the horizon being swept into the supersonic flow region $v < -1$; no sound can emerge from this region through the horizon, so it is reminiscent of a black hole. At the left-hand horizon right-moving waves become frozen and cannot enter the supersonic flow region; this is reminiscent of a white hole

the sonic black hole [1,8], and to make our discussion concrete from the outset we shall consider sonic horizons in one-dimensional fluid flow (see Fig. 1).

When the velocity profile of the fluid is as shown in Fig. 1 two horizons are formed for sound waves that propagate to the right with respect to the fluid. The horizon on the right of the supersonic flow region $v < -1$ behaves like a black hole horizon for right-moving waves, while the horizon on the left of the supersonic flow region behaves like a white hole horizon for these waves. In [8] Unruh showed that in such a system, with some reasonable simplifying assumptions, the equation for a small perturbation ϕ of the velocity potential is

$$(\partial_t + \partial_z v)(\partial_t + v\partial_z)\phi - \partial_z^2 \phi = 0 \,. \tag{1}$$

In terms of a new coordinate τ defined by

$$d\tau := dt + \frac{v}{1-v^2} dz$$

(1) is the equation $\phi_{;\mu}{}^{;\mu} = 0$ of a scalar field in the black-hole-type metric

$$ds^2 = (1-v^2)d\tau^2 - \frac{dz^2}{1-v^2} \,.$$

The results for quantum black holes [5, 9, 10] apply equally well here, so it follows that each horizon will produce a thermal spectrum of phonons with

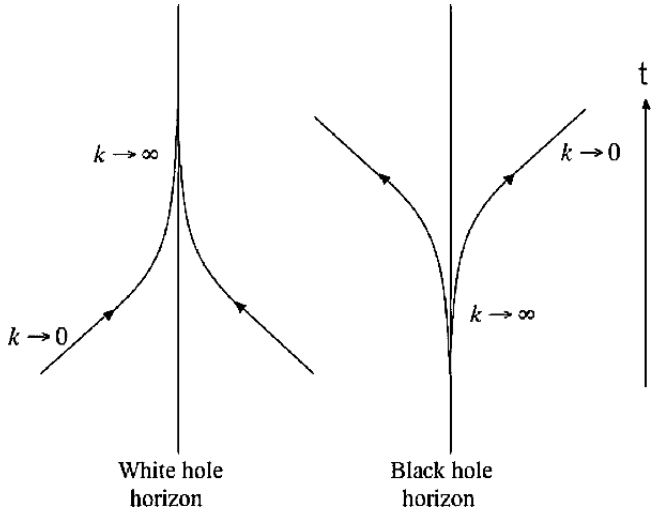

Fig. 2. Hawking phonons in the fluid flow of Fig. 1. Real phonons have positive frequency in the fluid-element frame and (3) shows that for right-moving phonons this frequency $(\omega - vk)$ is $\frac{\omega}{1+v} = k$. Thus in the subsonic-flow regions ω (conserved for each ray) is positive, whereas in the supersonic-flow region it is negative; k is positive for all real phonons. The frequency in the fluid-element frame diverges at the horizons– the trans-Plankian problem

a temperature determined by the quantity that corresponds to the surface gravity at the horizon, namely the absolute value of the slope of the velocity profile:

$$k_B T = \frac{\hbar\alpha}{2\pi} , \qquad \alpha := \left|\frac{dv}{dz}\right|_{v=-1} . \tag{2}$$

The trajectories of the created phonons are readily drawn (see Fig. 2), but they are formally deduced from the dispersion relation of the sound equation (1). Geometrical acoustics applied to (1) gives the dispersion relation

$$\omega - vk = \pm k \tag{3}$$

and the Hamilton equations

$$\frac{dz}{dt} = \frac{\partial\omega}{\partial k} = v \pm 1 , \tag{4}$$

$$\frac{dk}{dt} = -\frac{\partial\omega}{\partial z} = -v'k . \tag{5}$$

The left-hand side of (3) is the frequency in the frame co-moving with a fluid element, whereas ω is the frequency in the laboratory frame; the latter is constant for a time-independent fluid flow ("time-independent Hamiltonian" $d\omega/dt = \partial\omega/\partial t = 0$). Since the Hawking radiation is right-moving with respect

to the fluid, we clearly must choose the positive sign in (3) and hence in (4) also. By approximating $v(z)$ as a linear function near the horizons we obtain from (4) and (5) the ray trajectories of Fig. 2. The disturbing feature of the rays is the behavior of the wave vector k: at the horizons the radiation is exponentially blue-shifted[1], leading to a diverging frequency in the fluid-element frame. These runaway frequencies are unphysical since (1) asserts that sound in a fluid element obeys the ordinary wave equation at all wavelengths, in contradiction with the atomic nature of fluids (see Sect. 2). Moreover the conclusion that this Hawking radiation is actually present in the fluid also assumes that (1) holds at all wavelengths, as exponential blue-shifting of wave packets at the horizon is a feature of the derivation [5, 9, 10]. Similarly, in the black-hole case the equation used in the calculation [5, 9, 10] does not hold at arbitrarily high frequencies because it ignores the gravity of the fields. For the black hole, a complete resolution of this difficulty will require knowledge of the gravitational physics of quantum fields, i.e. quantum gravity, but for the dumb hole the physics is available for a more realistic treatment. The issue to be addressed is the dispersion relation for sound at high frequencies [1], and a consideration of this will lead us to the black hole laser.

2 Dispersion

In reality one would expect the dispersion relation for sound to differ from the linear formula (3) when the wavelength is of the order of the distance between the fluid atoms. What kind of dispersion relation should we expect at these wavelengths? Let us naively picture the fluid atoms as occupying well-defined positions with a separation distance a (see Fig. 3). The equation for small

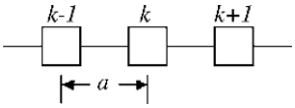

Fig. 3. A "classical" picture of fluid atoms. Each atom is labeled and a is the equilibrium separation

oscillations in this chain of atoms is the wave equation with a discretized second-order spatial derivative:

$$\partial_t^2 \phi_k - \frac{1}{a^2} \left(\phi_{k+1} - 2\phi_k + \phi_{k-1} \right) = 0 \, . \tag{6}$$

A partial differential equation that serves as an approximation to (6) is obtained by substituting the following relation, familiar from numerical analysis:

[1] We trust the reader will not be too distracted by our mixing of acoustic and optical terminology.

$$\frac{1}{a^2}\left(\phi_{k+1} - 2\phi_k + \phi_{k-1}\right) = \partial_z^2\phi + \frac{a^2}{12}\partial_z^4\phi + O(a^4) \; .$$

For small a we thus obtain the approximate dispersion relation

$$\omega^2 \approx k^2 - \frac{a^2}{12}k^4 \; . \tag{7}$$

As the last term in (7) is negative we find that the group velocity $\frac{d\omega}{dk}$ of waves decreases with decreasing wavelength, i.e. we have *subluminal* or *normal* dispersion.

A nonlinear dispersion relation $\omega = \omega(k)$ such as (7) limits the amount of blue-shifting that occurs at a horizon, so the embarrassment of infinite frequencies disappears; there then remains the question of whether one can still derive the Hawking effect. Unruh [1] modified the sound equation (1) to incorporate subluminal dispersion and showed numerically that the prediction of thermal radiation at temperature (2) still holds good, regardless of the details of the high-frequency behavior. Indeed, analysis of generalized models shows that the precise form of the dispersion relation at high frequencies has no effect on the particle production by a single horizon, at least to lowest order [2, 3]. Unsurprisingly, such a general conclusion involves some assumptions about the short-distance physics at the horizon, and since Planck-scale physics is unknown it remains possible that conditions at this scale conspire to invalidate the evaporation result for real black holes [3].

We now turn to the main purpose of this section, namely the introduction of a dispersion relation that, together with the horizon arrangement of Fig. 1, leads to black hole lasing.

2.1 Bogoliubov Dispersion

The dispersion relation used by Corley and Jacobson in the original treatment of the black hole laser [4] is well known from the theory of elementary excitations in a Bose–Einstein condensate [11, 12]. In this formalism the bosonic atoms are described by an a field $\hat{\psi}$ but most of the atoms are assumed to form a condensate with macroscopic wave function

$$\psi_0 = \sqrt{\rho_0}e^{iS_0} \; . \tag{8}$$

The deviations of $\hat{\psi}$ from the mean field ψ_0 are given by a quantum field $\hat{\phi}$ of small fluctuations:

$$\hat{\psi} = \psi_0 + e^{iS_0}\hat{\phi} \; .$$

It is then found that $\hat{\phi}$ satisfies the Bogoliubov–de Gennes equation

$$i\hbar\partial_t\hat{\phi} = \left(-\frac{\hbar^2}{2m}\nabla^2 - \mu + 2mc^2\right)\hat{\phi} + mc^2\hat{\phi}^\dagger \; , \tag{9}$$

where μ is the chemical potential and c is the speed of sound. The field $\hat{\phi}$ is given by the usual mode expansion

$$\hat{\phi} = \sum_{\nu} \left(u_\nu \hat{a}_\nu + v_\nu^* \hat{a}_\nu^\dagger \right) \ . \tag{10}$$

For plane-wave modes

$$u_\nu = u e^{i\mathbf{k}\cdot\mathbf{r} - i\omega t}, \qquad v_\nu = v e^{i\mathbf{k}\cdot\mathbf{r} - i\omega t} \ ,$$

(9) and its complex conjugate give

$$\hbar\omega \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \frac{\hbar^2 k^2}{2m} - \mu + 2mc^2 & mc^2 \\ -mc^2 & -\frac{\hbar^2 k^2}{2m} + \mu - 2mc^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \ .$$

The condition for this matrix equation to have a solution is

$$\left(\frac{\hbar^2 k^2}{2m} - \mu + 2mc^2 - \hbar\omega \right) \left(-\frac{\hbar^2 k^2}{2m} + \mu - 2mc^2 - \hbar\omega \right) + m^2 c^4 = 0 \ ,$$

and assuming the chemical potential μ is equal to the local energy of the condensate, mc^2, this is the Bogoliubov dispersion relation

$$\omega^2 = c^2 k^2 + \frac{\hbar^2}{4m^2} k^4 \ . \tag{11}$$

From (11) we see that the group velocity increases with decreasing wavelength, so it is an example of *superluminal* or *anomalous* dispersion. Comparison with (7) shows that the purely quantum term in (11) corresponds to an imaginary separation distance between the atoms, a reflection of the inapplicability of the classical picture of Fig. 3 to a condensate.

Let us now add Bogoliubov dispersion to our model system of Fig. 1. The appropriate generalization of (1) is

$$(\partial_t + \partial_z v)(\partial_t + v\partial_z)\phi - \partial_z^2 \phi + \frac{1}{k_c^2} \partial_z^4 \phi = 0 \ , \tag{12}$$

where k_c is a constant; this give the dispersion relation

$$\omega - vk = \pm k \sqrt{1 + \frac{k^2}{k_c^2}} \ , \tag{13}$$

so that we have Bogoliubov dispersion in a moving fluid. As with (3) we need only consider the positive sign in (13) since our interest is in phonons that propagate to the right with respect to the fluid.

How does the presence of the k_c-dependent term in (13) alter the ray trajectories of Fig. 2? We numerically solve the Hamilton equations resulting from (13) (with the positive sign) and obtain the trajectories shown in Fig. 4.

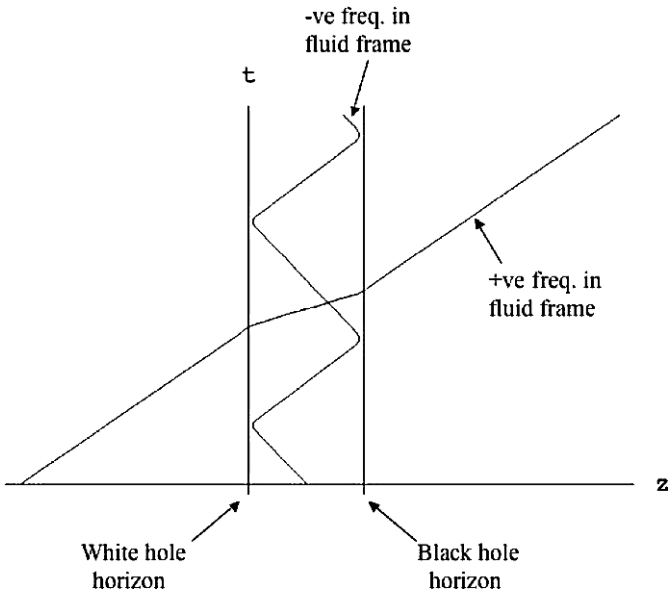

Fig. 4. Numerical solution of ray trajectories for the dispersion relation (13). Both rays are right-moving with respect to the fluid and have the same positive value of ω. The ray that is confined to the supersonic-flow region has negative frequency in a frame co-moving with the fluid, whereas the ray that passes through the system from left to right has positive frequency in the fluid frame

The qualitative behavior of the rays in Fig. 4 is easily understood from the superluminal character of the dispersion relation (13). Consider the ray approaching the white hole horizon from the left: the ray moves at the group velocity and is blue-shifted as it nears the horizon; this increases the group velocity until the ray traverses the horizon into the supersonic-flow region. Similarly, the ray being swept towards the white hole horizon from the right is blue-shifted until the group velocity exceeds the speed of the flow, allowing the ray to move away from the horizon. The behavior at the black hole horizon is simply the time reverse of that at the white hole horizon. Thus, the ray in the supersonic-flow region bounces back and forth between the horizons while the ray incident from the left passes through the system.

In Fig. 4 we have actually solved for the two rays using the same positive value of ω, the frequency in the laboratory frame, which is constant throughout the motion. It is then the case that the ray propagating through the system has positive frequency $\omega - vk$ in a frame co-moving with the fluid, whereas the bouncing ray has negative frequency in the fluid frame and does not therefore represent a real phonon. (A real phonon in the supersonic-flow region that is swept towards the white hole horizon has negative ω and exhibits the same bouncing behavior.) We have chosen the same value of ω for both rays in order to discuss particle creation. As the ray from the left reaches the horizon

geometrical acoustics breaks down, allowing a wave packet centered on this ray to undergo mode conversion wherein it develops a component corresponding to another branch of the dispersion relation having the same conserved value of ω. Thus a wave packet centered on the incoming ray in Fig. 4 will develop a component between the horizons that is centered on the bouncing ray in Fig. 4, which has negative frequency in the fluid frame. In the quantized theory it is a positive frequency in the fluid frame that is associated with an annihilation operator through a mode expansion such as (10). That fact that a positive-frequency wave packet develops a negative-frequency component implies a mixing of creation and annihilation operators in the quantum theory and this means there is particle creation by the system [1, 5, 9, 10].

The significance of the frequency in the fluid frame is apparent from the inner product

$$(\phi_1, \phi_2) = i \int [\phi_1^*(\partial_t + v\partial_z)\phi_2 - \phi_2(\partial_t + v\partial_z)\phi_1^*]\, dz \tag{14}$$

which is conserved for solutions of (12), i.e. $\partial_t(\phi_1, \phi_2) = 0$. The operator $\partial_t + v\partial_z$ is the time derivative in the fluid frame and so is associated with the frequency in this frame. It is easily shown a wave packet ϕ made up of plane-wave solutions with positive (negative) frequency in the fluid frame has positive (negative) conserved norm (ϕ, ϕ). Since it is the frequency in the fluid frame that is important for particle production and for the inner product (14), the terms positive/negative frequency will hereafter refer to the fluid frame.

The above considerations lead to the schematic picture of Fig. 5 for the evolution of an incident positive-frequency (and positive-norm) wave packet ϕ_{in}. At the white hole horizon the packet ϕ_{in} undergoes mode conversion, producing a negative-frequency component $\phi_{-R}^{(1)}$ that propagates to the right together with the positive-frequency part $\phi_{+}^{(1)}$. When these packets reach the black hole horizon there is further mode conversion and a positive-frequency packet $\phi_{\text{out}}^{(1)}$ exits the system leaving behind a red-shifted negative-frequency packet $\phi_{-L}^{(1)}$ that propagates to the left towards the white hole horizon. On reaching the white hole horizon $\phi_{-L}^{(1)}$ generates through mode conversion a positive-frequency packet $\phi_{+}^{(2)}$, which joins the blue-shifted negative-frequency packet $\phi_{-R}^{(2)}$ in propagating towards the black hole horizon. When $\phi_{+}^{(2)}$ and $\phi_{-R}^{(2)}$ reach the black hole horizon mode conversion occurs again and there is another output of a positive-frequency packet $\phi_{\text{out}}^{(2)}$, leaving a negative-frequency packet $\phi_{-L}^{(2)}$ between the horizons. The process continues as outlined. Now the total norm of this evolving solution is conserved, so in each instance of mode conversion amounts of positive- and negative-frequency wave packet are created that have equal and opposite norm. Thus with each bounce the norm of the negative-frequency packet between the horizons increases in magnitude; indeed, it increases exponentially since the norms of the wave packets created by mode conversion are in proportion to the norm of the incident packet. The

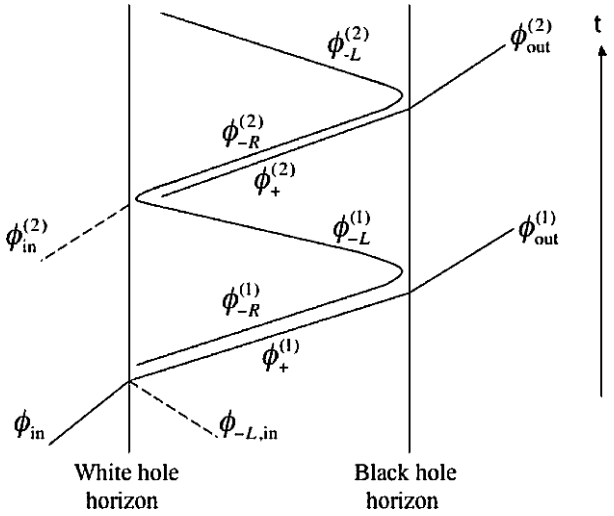

Fig. 5. Schematic picture of the evolution of a positive-frequency wave packet $\phi_{\rm in}$. Lines that begin at the horizons denote wave packets created by mode conversion. All wave packets follow one of the ray trajectories of Fig. 4. The *dotted lines* indicate modes that are empty in this evolution but that feature in the connection formulae for modes at the horizons

size of the negative frequency packet generated by $\phi_{\rm in}$ determines the particle production so the number of particles produced increases exponentially with time – this is the black hole laser [4].

To quantify the above description one can utilize the connection formulae obtained by Corley and Jacobson for modes at the horizons [4]. Let us assume that the slope of the velocity profile is $-\alpha$ at the white hole horizon and γ at the black hole horizon. Then the relation between modes with the same conserved value of ω at the white hole horizon is (see Fig. 5):

$$\phi_{\omega{\rm in}} = [2\sinh(\pi\omega/\alpha)]^{-1/2}\left(e^{\pi\omega/2\alpha}\phi^{(1)}_{\omega+} + e^{-\pi\omega/2\alpha}\phi^{(1)}_{\omega-R}\right), \qquad (15)$$

$$\phi_{\omega-L{\rm in}} = [2\sinh(\pi\omega/\alpha)]^{-1/2}\left(e^{-\pi\omega/2\alpha}\phi^{(1)}_{\omega+} + e^{\pi\omega/2\alpha}\phi^{(1)}_{\omega-R}\right), \qquad (16)$$

with similar formulae for the other modes that meet at this horizon in Fig. 5. At the black hole horizon the connection formulae for modes is (see Fig. 5):

$$\phi^{(1)}_{\omega{\rm out}} = [2\sinh(\pi\omega/\gamma)]^{-1/2}\left(e^{\pi\omega/2\gamma}\phi^{(1)}_{\omega+} + e^{-\pi\omega/2\gamma}\phi^{(1)}_{\omega-L}\right), \qquad (17)$$

$$\phi^{(1)}_{\omega-L} = [2\sinh(\pi\omega/\gamma)]^{-1/2}\left(e^{-\pi\omega/2\gamma}\phi^{(1)}_{\omega+} + e^{\pi\omega/2\gamma}\phi^{(1)}_{\omega-L}\right), \qquad (18)$$

with similar formulae applying for the other modes that interact at this horizon in Fig. 5. For modes that travel between the horizons ($\phi^{(n)}_{\omega+}$, $\phi^{(n)}_{\omega-R}$ and

$\phi^{(n)}_{\omega-L}$) we must in addition specify how their values at the two horizons are related to each other. This can be done using a WKB approximation to propagate the modes from one horizon to the other; the result will be a simple phase relationship between the values at each horizon:

$$\phi^{(n)}_{\omega+}\Big|_{\mathrm{BH}} = e^{i\theta_+}\ \phi^{(n)}_{\omega+}\Big|_{\mathrm{WH}}\,, \tag{19}$$

$$\phi^{(n)}_{\omega-R}\Big|_{\mathrm{BH}} = e^{-i\theta_-}\ \phi^{(n)}_{\omega-R}\Big|_{\mathrm{WH}}\,, \tag{20}$$

$$\phi^{(n)}_{\omega-L}\Big|_{\mathrm{WH}} = e^{-i\theta_0}\ \phi^{(n)}_{\omega-L}\Big|_{\mathrm{BH}}\,, \tag{21}$$

where BH (WH) refers to the black hole (white hole) horizon. Using the above formulae as building blocks one can calculate the evolution of a wave packet and hence compute the number of created particles [4]; in Sect. 4 we shall address this task in a different guise. First, a numerical solution for an evolving wave packet will provide an enlightening picture of the process outlined in Fig. 5.

3 Numerical Results

Let us summarize what we expect to see if we propagate a positive-frequency wave packet towards the white hole horizon. According to Fig. 5 this positive-norm wave packet will plough through the system and emerge out of the black hole horizon ($\phi^{(1)}_{\mathrm{out}}$) somewhat amplified, leaving behind a small negative-norm packet ($\phi^{(1)}_{-L}$) such that the total norm is conserved. After a period of time equal to that required by the negative-frequency ray in Fig. 4 to propagate from the black hole horizon to the white hole horizon and back again we expect another, smaller, positive-norm packet ($\phi^{(2)}_{\mathrm{out}}$) to emerge from the system. At regular intervals there should be further outputs of positive-norm packets ($\phi^{(3)}_{\mathrm{out}}$, etc.) and a build-up of the negative-norm packet between the horizons. The increasing amplitude of the negative-norm packet will cause an increase in the size of the positive-norm outputs produced by mode conversion, and this increase should in fact be exponential.

In Fig. 6 we show the result of a numerical evolution of a wave packet centered on the positive-frequency ray in Fig. 4. Periodic boundary conditions are in force in Fig. 6 and the incident packet has ploughed through the system ($\phi^{(1)}_{\mathrm{out}}$) and is circling round on the left-hand side where it is about to hit the white hole horizon a second time. The successive outputs of positive-norm packets ($\phi^{(2)}_{\mathrm{out}}$, etc.) are clearly visible; five separate emissions are distinguishable, corresponding to the negative-norm packet bouncing off the black hole horizon five times. As the total amount of positive-norm packet increases with each output it is balanced by a growing amount of negative-norm packet between the horizons, leaving the total norm unchanged.

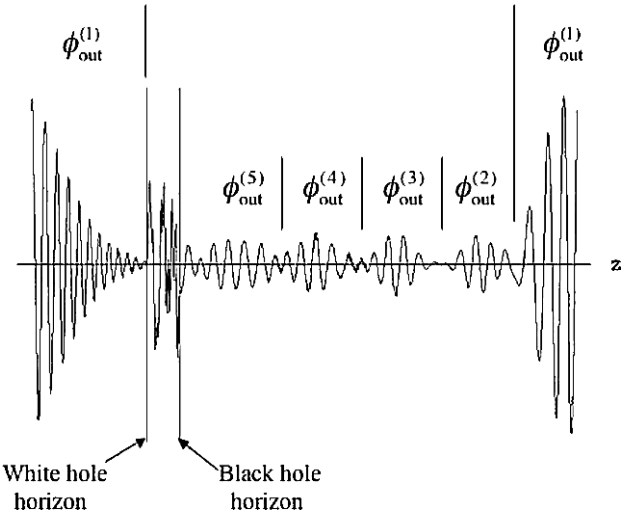

Fig. 6. Numerical solution for a positive-frequency wave packet that has entered the horizon system from the left. The packet is centered on the positive-frequency ray in Fig. 4 so it has propagated through the system ($\phi^{(1)}_{\text{out}}$) and begun to emerge on the left-hand side due to periodic boundary conditions. A sequence of positive-frequency wave packets are seen to follow $\phi^{(1)}_{\text{out}}$, in line with expectations based on Fig. 5

The distance between the horizons has been set quite small relative to the length of the incident wave packet in order to increase the number of bounces that occur before $\phi^{(1)}_{\text{out}}$ collides with the white hole horizon. This makes it difficult to see clearly the motion of the negative-frequency wave packet that is confined to the supersonic-flow region. Nevertheless alternate red- and blue-shifting is observed between the horizons as the solution evolves and the time between successive positive-norm outputs is equal to the time between visits to the black hole horizon of the ray on which the negative-frequency packet is centered (see Fig. 4).

Although the last output $\phi^{(5)}_{\text{out}}$ in Fig. 6 seems to be somewhat larger than $\phi^{(2)}_{\text{out}}$ there is no clear sign of an exponential increase in the size of the output packets after this number of emissions. The analytic formulae for the evolution show that about ten outputs would be required in this simulation before there is a significant increase in their size, as we shall see in the next section.

Before we turn to an analytic treatment of the black hole laser and the calculation of the particle production, we briefly mention the equally remarkable results when the field ϕ is fermionic rather than bosonic [4]. The conserved norm for a fermionic field is positive definite, so both positive- and negative-frequency packets have positive norm. From this fact alone follows the evolution of a positive-frequency fermionic wave packet. When the initial wave packet ploughs through the system as in Fig. 6 it is diminished in

amplitude in order to conserve the total norm, since the negative-frequency packet left behind also has positive norm. As subsequent positive-frequency packets are emitted conservation of norm requires the amount of negative frequency packet between the horizons to decrease rather than increase as in the bosonic case. In the limit of infinite time the amount of negative frequency packet goes to zero and therefore so does the number of particles produced – there is complete suppression of the Hawking effect. This can be understood as a filling up of all the allowed fermionic states between the horizons by the Hawking particles created in this region [4].

4 Black Hole Amplifier

The Bogoliubov transformation that underlies particle creation by a black hole is similar to that which describes the action of a parametric amplifier or a phase-conjugating mirror in quantum optics [13, 14]. An optical device will transform a set of incident quantum light modes, described by annihilation operators $\{\hat{a}_1, \ldots, \hat{a}_n\}$ and their Hermitian conjugates $\{\hat{a}_1^\dagger, \ldots, \hat{a}_n^\dagger\}$, into outgoing modes, described by operators $\{\hat{a}'_1, \ldots, \hat{a}'_n\}$ and $\{\hat{a}_1'^\dagger, \ldots, \hat{a}_n'^\dagger\}$. If the device is linear this transformation is given by the scattering matrix S:

$$\begin{pmatrix} \hat{a}'_1 \\ \vdots \\ \hat{a}'_n \\ \hat{a}_1'^\dagger \\ \vdots \\ \hat{a}_n'^\dagger \end{pmatrix} = S \begin{pmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_n \\ \hat{a}_1^\dagger \\ \vdots \\ \hat{a}_n^\dagger \end{pmatrix} . \tag{22}$$

The S-matrix is constrained by the requirement that the creation and annihilation operators satisfy the usual commutation relations; this imposes the quasi-unitarity condition

$$SGS^\dagger = G, \qquad G = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} . \tag{23}$$

Matrices S satisfying (23) form the group $U(n, n)$ [15]. An example of such a matrix is

$$S = \begin{pmatrix} \cosh\xi & 0 & 0 & \sinh\xi \\ 0 & \cosh\xi & \sinh\xi & 0 \\ 0 & \sinh\xi & \cosh\xi & 0 \\ \sinh\xi & 0 & 0 & \cosh\xi \end{pmatrix} , \tag{24}$$

which is the scattering matrix for a parametric amplifier or a phase-conjugating mirror [13,14]. In this case there are two input and output modes and the device creates or annihilates pairs of photons. The energy for the pair production

or the reservoir for pair annihilation is provided by the pump process of the amplifier, which determines the parameter ξ. For a black hole we may define ξ by

$$\tanh \xi = e^{-\pi\omega/\gamma} \ , \tag{25}$$

where γ is the surface gravity at the horizon and ω is the frequency of the quantum modes under consideration. Then (24) gives the Bogoliubov transformation for the Hawking effect [9,10], which we write concisely as

$$\begin{pmatrix} \hat{a}'_1 \\ \hat{a}'^{\dagger}_2 \end{pmatrix} = \begin{pmatrix} \cosh\xi & \sinh\xi \\ \sinh\xi & \cosh\xi \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}^{\dagger}_2 \end{pmatrix} \ . \tag{26}$$

For a moving fluid, (25) and (26) describe phonon creation by a black hole horizon, where γ denotes the slope of the velocity profile $v(z)$ at the horizon. A white hole horizon in the fluid, where the slope of $v(z)$ is $-\alpha$, is the time reverse of a black hole horizon. The "surface gravity" is α and the roles of input and output modes are reversed. Thus, defining a parameter ζ by

$$\tanh \zeta = e^{-\pi\omega/\alpha} \tag{27}$$

the white hole horizon generates the Bogoliubov transformation

$$\begin{pmatrix} \hat{a}'_1 \\ \hat{a}'^{\dagger}_2 \end{pmatrix} = \begin{pmatrix} \cosh\zeta & -\sinh\zeta \\ -\sinh\zeta & \cosh\zeta \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}^{\dagger}_2 \end{pmatrix} \ . \tag{28}$$

In the black hole laser the particle production by the pair of horizons is not given by (26) and (28) because the horizons interact due to the superluminal dispersion. We can represent the black hole laser as a quantum-optical device by expressing the amplification process of Fig. 5 in terms of the annihilation operators of the modes, and using a more convenient notation (see Fig. 7). In the nth amplification step we have a pair of input modes $\hat{a}_{+n}$ and $\hat{a}_{-n}$ at the white hole horizon that generates an output mode $\hat{a}'_{+n}$ on the right-hand side of the black hole laser as well as a mode $\hat{a}'_{-n}$ that is trapped between the horizons where it serves as a further input to the device, i.e.

$$\hat{a}'_{-n} = \hat{a}_{-(n+1)}. \tag{29}$$

The picture is that of a network of amplifiers, as outlined in Fig. 7.

The output of the black hole laser is completely specified by the matrix B that effects the Bogoliubov transformation for each step in Fig. 7:

$$\begin{pmatrix} \hat{a}'_{-n} \\ \hat{a}'^{\dagger}_{+n} \end{pmatrix} = B \begin{pmatrix} \hat{a}_{-n} \\ \hat{a}^{\dagger}_{+n} \end{pmatrix} \ . \tag{30}$$

There are four distinct processes incorporated in the matrix B, allowing it to be expressed as

$$B = B_4 B_3 B_2 B_1 \ . \tag{31}$$

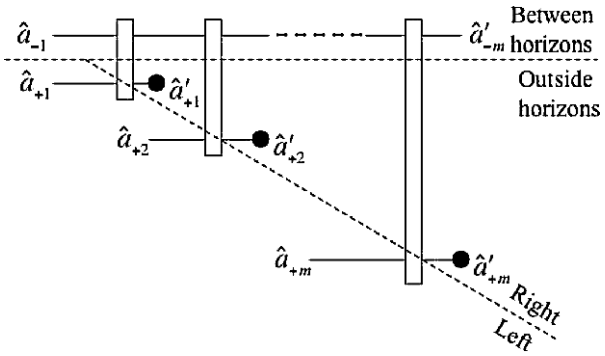

Fig. 7. The black hole laser as a network of amplifiers. The mode operators with negative subscripts correspond to the Hawking radiation trapped between the horizons, while those with positive subscripts describe the Hawking radiation present outside the horizons. An input consisting of a mode on the left of the black hole laser and a mode between the horizons produces an output mode on the right of the black hole laser and another output mode between the horizons, as in Fig. 5. The process repeats, with the output mode between the horizons acting as an input mode for the next cycle, i.e. $\hat{a}'_{-n} = \hat{a}_{-(n+1)}$

First there is the pair production by the white hole horizon, described by (28); we therefore have

$$B_1 = \begin{pmatrix} \cosh\zeta & -\sinh\zeta \\ -\sinh\zeta & \cosh\zeta \end{pmatrix} . \tag{32}$$

Next, the two modes produced by the white hole horizon propagate to the black hole horizon, as in Fig. 5; this imparts separate phases θ_- and θ_+ to each mode operator, so we write

$$B_2 = \begin{pmatrix} e^{-i\theta_-} & 0 \\ 0 & e^{i\theta_+} \end{pmatrix} . \tag{33}$$

At the black hole horizon the modes are transformed according to (26), giving

$$B_3 = \begin{pmatrix} \cosh\xi & \sinh\xi \\ \sinh\xi & \cosh\xi \end{pmatrix} . \tag{34}$$

Finally the mode trapped between the the horizons propagates to the white hole horizon, acquiring a phase θ_0:

$$B_4 = \begin{pmatrix} e^{-i\theta_0} & 0 \\ 0 & 1 \end{pmatrix} . \tag{35}$$

The particle content after m amplification steps is now easily calculated from (30) and (29), using the value (31)–(35) for B. It is worthwhile to consider the restrictions on the 2×2 matrix B resulting from the fact that it determines

a $U(2,2)$ scattering matrix S (see (30) and (22)). Constructing S from B and imposing (23) one finds that the components of B must satisfy

$$|B_{11}|^2-|B_{12}|^2=1\ ,\qquad |B_{21}|^2-|B_{22}|^2=1\ ,\qquad B_{11}^*B_{21}-B_{12}^*B_{22}=0\ , \tag{36}$$

which is the statement that B is a $U(1,1)$ matrix. The constraints (36) impose the following structure on B:

$$B=e^{i\psi}\begin{pmatrix}\mu & \nu^*\\ \nu & \mu^*\end{pmatrix},\quad \psi \text{ real}\ , \tag{37}$$

$$|\mu|^2-|\nu|^2=1\ . \tag{38}$$

Using (30), (29) and (37) we find after m cycles that $\hat{a}'_{-m}$ for the mode trapped between the horizons and $\hat{a}'^{\dagger}_{+m}$ for the mode outside the horizons are given by

$$\hat{a}'_{-m}=e^{im\psi}\mu^m\hat{a}_{-1}+\nu^*\sum_{l=1}^{m}e^{i(m-l+1)\psi}\mu^{m-l}\hat{a}^{\dagger}_{+l}\ , \tag{39}$$

$$\hat{a}'^{\dagger}_{+m}=e^{im\psi}\nu\mu^{m-1}\hat{a}_{-1}+|\nu|^2\sum_{l=1}^{m-1}e^{i(m-l+1)\psi}\mu^{m-l-1}\hat{a}^{\dagger}_{+l}+e^{i\psi}\mu^*\hat{a}^{\dagger}_{+m}\ . \tag{40}$$

The condition that we have vacuum before the formation of the black hole laser gives

$$\hat{a}_{-1}|0\rangle=0\ ,\qquad \hat{a}_{+l}|0\rangle=0\ ,\quad l=1,\dots m\ .$$

We may write the particle content for the two output modes in terms of $|\mu|^2$ because of (38). The particle number for the trapped mode after m cycles is

$$\begin{aligned}\langle\hat{N}_{-m}\rangle&=\langle\hat{a}'^{\dagger}_{-m}\hat{a}'_{-m}\rangle\\ &=|\nu|^2\sum_{l=1}^{m}|\mu|^{2(m-l)}=|\nu|^2\frac{|\mu|^{2m}-1}{|\mu|^2-1}\\ &=|\mu|^{2m}-1\ ,\end{aligned} \tag{41}$$

whereas the number of particles outside the horizons is

$$\begin{aligned}\langle\hat{N}_{+m}\rangle&=\langle\hat{a}'^{\dagger}_{+m}\hat{a}'_{+m}\rangle\\ &=|\nu|^2|\mu|^{2(m-1)}\\ &=|\mu|^{2m}\left(1-|\mu|^{-2}\right).\end{aligned} \tag{42}$$

The explicit form of $|\mu|^2$ is found from (31)–(35) and (37) to be

$$|\mu|^2=\frac{1}{2}\big[1+\cosh(2\xi)\cosh(2\zeta)-\cos(\theta_++\theta_-)\sinh(2\xi)\sinh(2\zeta)\big]\ . \tag{43}$$

Note that the phase θ_0 does not appear in (43) and thus has no effect on the particle content; this is because θ_0 is acquired by a mode traveling to the white hole horizon where it interacts with a mode from outside the device, and the latter has no phase sensitivity. The combination $\theta_+ + \theta_-$, however, is the phase difference between the two modes that propagate across the device from the white hole horizon and this affects how these modes interact at the black hole horizon. We see from (41)–(43) that for frequencies such that $\cos(\theta_+ + \theta_-) = 1$ the particle production is maximally suppressed; indeed if the horizons are symmetric ($\xi = \zeta$) there is no production at this frequency. On the other hand maximal amplification of the Hawking radiation occurs at frequencies for which $\cos(\theta_+ + \theta_-) = 0$, giving a WKB-type condition

$$\theta_+ + \theta_- = 2\pi \left(n + \frac{1}{2} \right), \qquad n \text{ an integer.}$$

In Fig. 8 we use formula (42) to plot the particle number outside the horizons as a function of the number of amplification cycles. The slopes of $v(z)$ at the horizons are those used for the numerical example in Sect. 3 and the frequency chosen is that on which the incident pulse in Fig. 6 is centered. A WKB approximation is used to estimate the phases θ_+ and θ_-. Figure 8 shows the consequence, in terms of Hawking radiation, of the wave-packet evolution in Fig. 6. As discussed in Sect. 3, the stream of positive-frequency pulses emerging from the black hole horizon in Fig. 6 implies a growth in particle production, as in Fig. 8. The solution domain in Fig. 6 is only large enough to contain the result of five amplification cycles. In this time the positive-frequency pulses trailing $\phi^{(1)}_{\text{out}}$ are of approximately the same size; this implies an approximately linear increase in the particle number, as is the case in Fig. 8 with $n \leq 5$. In order for the exponential nature of the increase in particle number to be clearly seen in a wave-packet evolution, one would require at least ten positive-frequency outputs in the evolution, as one sees from Fig. 8.

In summary, the black hole laser is a dramatic quantum-vacuum effect and provides an interesting example of the interplay of Hawking radiation

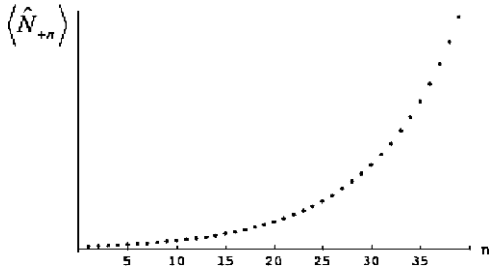

Fig. 8. The number of Hawking particles outside the horizons versus the number of amplification cycles, as given by formula (42)

and nonlinear dispersion. Any cosmological significance of this process remains highly speculative [4], but in contrast a laboratory example of black hole lasing is not nearly so far-fetched. Bose–Einstein condensates have the properties required for the construction of a sonic black hole laser, however formidable the practical difficulties may be, and this is not the only current possibility [16]. As the study of black hole analogues continues the prospect of exponentially amplified Hawking radiation is tantalizing for both theorists and experimentalists.

References

1. Unruh W.G. (1995) Phys Rev D51:2827
2. Brout R., Masser S., Parentani R., Spindel Ph. (1995) Phys Rev D52:4559
3. Unruh W.G., Schützhold R. (2005) Phys Rev D71:024028
4. Corley S., Jacobson T. (1999) Phys Rev D59:124011
5. Hawking S.W. (1975) Commun Math Phys 43:199
6. 't Hooft G. (1995) Nucl Phys B256:727
7. Jacobson T. (1991) Phys Rev D44:1731
8. Unruh W.G. (1981) Phys Rev Lett 46:1351
9. Birrell N.D., Davies P.C.W. (1982) Quantum fields in curved space. Cambridge Univeristy Press, Cambridge
10. Brout R., Masser S., Parentani R., Spindel Ph. (1995) Phys Rep 260:329
11. Dalfovo F., Giorgini S., Pitaevskii L.P., Stringari S. (1999) Rev Mod Phys 71:463
12. Pitaevskii L.P., Stringari S. (2003) Bose–Einstein condensation. Clarendon Press, Oxford
13. Shen Y.R. (1984) The principles of nonlinear optics. Wiley, New York
14. Leonhardt U. (2003) Rep Prog Phys 66:1207
15. Cornwell J.F. (1984) Group theory in Physics. Volume 2. Academic Press, London
16. Volovik G.E. (2005) JETP Lett 82:624

Cosmic Strings

M. Sakellariadou

Department of Physics, King's College London, University of London, Strand WC2R 2LS, United Kingdom
Mairi.Sakellariadou@kcl.ac.uk

Cosmic strings, a hot subject in the 1980's and early 1990's, lost its appeal when it was found that it leads to inconsistencies in the power spectrum of the measured cosmic microwave background temperature anisotropies. However, topological defects in general, and cosmic strings in particular, are deeply rooted in the framework of grand unified theories. Indeed, it was shown that cosmic strings are expected to be generically formed within supersymmetric grand unified theories. This theoretical support gave a new boost to the field of cosmic strings, a boost which has been recently enhanced when it was shown that cosmic superstrings (fundamental or one-dimensional Dirichlet branes) can play the rôle of cosmic strings, in the framework of braneworld cosmologies.

To build a cosmological scenario we employ high energy physics; inflation and cosmic strings then naturally appear. Confronting the predictions of the cosmological scenario against current astrophysical/cosmological data we impose constraints on its free parameters, obtaining information about the high energy physics we employed.

This is a beautiful example of the rich and fruitful interplay between cosmology and high energy physics.

1 Introduction

The basic ingredient in cosmology is general relativity and the choice of a metric. The Friedmann-Lemaître-Robertson-Walker (FLRW) model, known as the *hot big bang* model is a homogeneous and isotropic solution of Einstein's equations; the hyper-surfaces of constant time are homogeneous and isotropic, i.e., spaces of constant curvature. The hot big bang model is based on the FLRW metric

$$ds^2 = dt^2 - a^2(t)\gamma_{ij}dx^i dx^j = a^2(\tau)[d\tau^2 - \gamma_{ij}dx^i dx^j] \; ; \qquad (1)$$

M. Sakellariadou: *Cosmic Strings*, Lect. Notes Phys. **718**, 247–288 (2007)
DOI 10.1007/3-540-70859-6_10

$a(t)$ or $a(\tau)$ is the cosmic scale factor in terms of the cosmological time t or the conformal time τ (with $dt = ad\tau$) respectively, and γ_{ij} is the metric of a space with constant curvature κ. The metric γ_{ij} can be expressed as

$$\gamma_{ij}dx^i dx^j = dr^2 + \chi^2(r)\left(d\vartheta^2 + sin^2\vartheta d\varphi^2\right) \ , \tag{2}$$

$$\text{with} \ \ \chi(r) = \begin{cases} r & \text{for} \quad \kappa = 0 \\ \sin r & \text{for} \quad \kappa = 1 \\ \sinh r & \text{for} \quad \kappa = -1 \ ; \end{cases} \tag{3}$$

the scale factor $a(\tau)$ has been rescaled so that the curvature is $\kappa = \pm 1$ or 0.

The cornerstone of the hot big bang model is the high degree of symmetry of the FLRW metric: there is only one dynamical variable, the cosmic scale factor. The FLRW model is so successful that became the standard cosmological model. The high degree of symmetry of the metric, originally a theorist's simplification, is now an evidence thanks to the remarkable uniformity of temperature of the Cosmic Microwave Background (CMB) measured first by the COBE-DMR satellite [1].

The four pillars upon which the success of standard hot big bang model lie are: (i) the expansion of the Universe, (ii) the origin of the cosmic background radiation, (iii) the synthesis of light elements, and (iv) the formation of large-scale structures. However, there are questions, which mainly concern the initial conditions, to which the hot big bang model is unable to provide an answer. These shortcomings of the FLRW model are: (i) the horizon problem, (ii) the flatness problem, (iii) the exotic relics, (iv) the origin of density fluctuations, (v) the cosmological constant, and (vi) the singularity problem. To address these issues, inflation was proposed [2, 3]. Inflation essentially consists of a phase of accelerated expansion, corresponding to repulsive gravity and an equation of state $3p < -\rho$, which took place at a very high energy scale. Even though inflation is at present the most appealing scenario to describe the early stages of the Universe, the issue of how generic is the onset of inflation is still under discussion [4–6], at least within a large class of inflationary potentials, in the context of classical general relativity and loop quantum cosmology.

From the observational point of view, the remarkable uniformity of the CMB indicates that at the epoch of last scattering, approximately 2×10^5 yr after the big bang, when the Universe was at a temperature of approximately $0.26\,\text{eV} \simeq 3 \times 10^3\,\text{K}$, the Universe was to a high degree or precision (10^{-5}) isotropic and homogeneous. At very large scales, much bigger than 110 Mpc $\approx 10^{21}$ km, the Universe is smooth, while at small scales the Universe is very lumpy. The fractional overdensity at the time of decoupling between baryons and photons was

$$\left(\frac{\delta\rho}{\rho}\right)_{\text{dec}} = \mathcal{C} \times \left(\frac{\delta T}{T}\right) \leq \mathcal{O}(10^{-2} - 10^{-3}) \ ; \tag{4}$$

the constant $\mathcal{C}$ depends on the nature of density perturbations and it is $\mathcal{C} = \mathcal{O}(10 - 100)$. Then one asks the following question: *how does a very smooth Universe at the time of decoupling became very lumpy today?*

In the 1980's and 1990's, cosmologists had the following picture in mind: small, primeval density inhomogeneities grew via gravitational instability into the large inhomogeneities we observe today. To address the question of the origin of the initial density inhomogeneities, one needs to add more ingredients (namely scalar fields) to the cosmological model. This is where high energy physics enter the picture. Clearly, to build a detailed scenario of structure formation one should know the initial conditions, i.e., the total amount of nonrelativistic matter, the composition of the Universe, the spectrum and type of primeval density perturbations.

For almost two decades, two families of models have been considered challengers for describing, within the framework of gravitational instability, the formation of large-scale structure in the Universe. Initial density perturbations can either be due to *freezing in* of quantum fluctuations of a scalar field during an inflationary period, or they may be seeded by a class of topological defects [7], which could have formed naturally during a symmetry breaking phase transition in the early Universe. On the one hand, quantum fluctuations amplified during inflation produce *adiabatic*, or *curvature* fluctuations with a scale-invariant spectrum. It means that there are fluctuations in the local value of the spatial curvature, and that the fractional overdensity in Fourier space behaves as $|\delta_k|^2 \propto k^{-3}$. If the quantum fluctuations of the inflaton field are in the vacuum state, then the statistics of the CMB is Gaussian [8,9]. On the other hand, topological defects trigger *isocurvature*, or *isothermal* fluctuations, meaning that there are fluctuations in the form of the local equation of state, with nongaussian statistics and a scale-invariant spectrum. The CMB anisotropies provide a link between theoretical predictions and observational data, which may allow us to distinguish between inflationary models and topological defects scenarios, by purely linear analysis. The characteristics of the CMB anisotropy multipole moments (position, amplitude of acoustic peaks), and the statistical properties of the CMB are used to discriminate among models, and to constrain the parameters space.

Many particle physics models of matter admit solutions which correspond to a class of topological defects, that are either stable or long-lived. Provided our understanding about unification of forces and the big bang cosmology are correct, it is natural to expect that such topological defects could have formed naturally during phase transitions followed by spontaneously broken symmetries, in the early stages of the evolution of the Universe. Certain types of topological defects (local monopoles and local domain walls) lead to disastrous consequences for cosmology and thus, they are undesired, while others may play a useful rôle. We consider gauge theories, thus we are only interested in cosmic strings, since on the one hand strings are not cosmologically dangerous (monopoles and domain walls are), and on the other hand they can be useful in cosmology (textures decay too fast).

Cosmic strings are linear topological defects, analogous to flux tubes in type-II superconductors, or to vortex filaments in superfluid helium. In the framework of Grand Unified Theories (GUTs), cosmic strings might have been

formed at a grand unification transition, or much later, at the electroweak transition, or even at an intermediate one. These objects carry a lot of energy and they could play a rôle in cosmology and/or astrophysics. In the simplest case, the linear mass density of a cosmic string, denoted by μ, is equal to the string tension. Thus, the characteristic speed of waves on the string is the speed of light. For strings produced at a phase transition characterised by temperature $T_{\rm c}$, one expects roughly $\mu \sim T_{\rm c}^2$. The strength of the gravitational interaction of cosmic strings is given in terms of the dimensionless quantity $G\mu \sim (T_{\rm c}/M_{\rm Pl})^2$, where G and $M_{\rm Pl}$ denote the Newton's constant and the Planck mass, respectively. For grand unification strings, the energy per unit length is $\mu \sim 10^{22}\,{\rm kg/m}$, or equivalently, $G\mu \sim \mathcal{O}(10^{-6} - 10^{-7})$.

Topological defects (global or local) in general, and cosmic strings in particular, are ruled out as the unique source of the measured CMB temperature anisotropies. Clearly, one should then address the following question: *which are the implications for the high energy physics models upon which the cosmological scenario is based?* This leads to the following list of questions: (i) how generic is cosmic string formation? (ii) which is the rôle of cosmic strings, if any? and (iii) which is a natural inflationary scenario (inflation is still a paradigm in search of a model)? These questions will be addressed in what follows. We will see that cosmic strings are generically formed at the end of an inflationary era, within the framework of Supersymmetric Grand Unified Theories (SUSY GUTs). This implies that cosmic strings have to be included as a subdominant partner of inflation. We will thus consider mixed models, where both the inflaton field and cosmic strings contribute to the measured CMB temperature anisotropies. Comparing theoretical predictions against CMB data we will find the maximum allowed contribution of cosmic strings to the CMB measurements. We will then ask whether the free parameters of supersymmetric inflationary models can be adjusted so that the contribution of strings to the CMB is within the allowed window.

Finally, the recent proposal that cosmic superstrings can be considered as cosmic string candidates opens new perspectives on the theoretical point of view. More precisely, in the framework of large extra dimensions, long superstrings may be stable and appear at the same energy scale as GUT scale cosmic strings.

In what follows we first discuss, in Sect. 2, topological defects in GUTs. We classify topological defects and we give the criterion for their formation. We then briefly discuss two simple models leading to the formation of global strings (vortices) and local (gauge) strings, namely the Goldstone and the Abelian-Higgs model, respectively. Next, we present the Kibble and Zurek mechanisms of topological defect formation. We concentrate on local gauge strings (cosmic strings) and give the equations of motion for strings in the limit of zero thickness, moving in a curved spacetime. We subsequently discuss the evolution of a cosmic string network; the results are based on heavy numerical simulations. We then briefly present string statistical mechanics and the Hagedorn phase transition. We end this Section by addressing the

question of whether cosmic strings are expected to be generically formed after an inflationary era, in the context of supersymmetric GUTs. In Sect. 3 we discuss the most powerful tool to test cosmological predictions of theoretical models, namely the spectrum of CMB temperature anisotropies. We then analyse the predictions of models where the initial fluctuations leading to structure formation and the induced CMB anisotropies were triggered by topological defects. In Sect. 4, we study inflationary models in the framework of supersymmetry and supergravity. In Sect. 5, we address the issue of cosmic superstrings as cosmic strings candidates, in the context of braneworld cosmologies. We round up with our conclusions in Sect. 6.

2 Topological Defects

2.1 Topological Defects in GUTs

The Universe has steadily cooled down since the Planck time, leading to a series of Spontaneous Symmetry Breaking (SSB), which may lead to the creation of topological defects [7], false vacuum remnants, such as domain walls, cosmic strings, monopoles, or textures, via the Kibble mechanism [10].

The formation or not of topological defects during phase transitions, followed by SSB, and the determination of the type of the defects, depend on the topology of the vacuum manifold $\mathcal{M}_n$. The properties of $\mathcal{M}_n$ are usually described by the kth homotopy group $\pi_k(\mathcal{M}_n)$, which classifies distinct mappings from the k-dimensional sphere S^k into the manifold $\mathcal{M}_n$. To illustrate that, let us consider the symmetry breaking of a group G down to a subgroup H of G. If $\mathcal{M}_n = \mathrm{G/H}$ has disconnected components, or equivalently if the order k of the nontrivial homotopy group is $k = 0$, then two-dimensional defects, called *domain walls*, form. The spacetime dimension d of the defects is given in terms of the order of the nontrivial homotopy group by $d = 4-1-k$. If $\mathcal{M}_n$ is not simply connected, in other words if $\mathcal{M}_n$ contains loops which cannot be continuously shrunk into a point, then *cosmic strings* form. A necessary, but not sufficient, condition for the existence of stable strings is that the first homotopy group (the fundamental group) $\pi_1(\mathcal{M}_n)$ of $\mathcal{M}_n$, is nontrivial, or multiply connected. Cosmic strings are line-like defects, $d = 2$. If $\mathcal{M}_n$ contains unshrinkable surfaces, then *monopoles* form, for which $k = 1,\ d = 1$. If $\mathcal{M}_n$ contains noncontractible three-spheres, then event-like defects, *textures*, form for which $k = 3,\ d = 0$.

Depending on whether the symmetry is local (gauged) or global (rigid), topological defects are called *local* or *global*. The energy of local defects is strongly confined, while the gradient energy of global defects is spread out over the causal horizon at defect formation. Patterns of symmetry breaking which lead to the formation of local monopoles or local domain walls are ruled out, since they should soon dominate the energy density of the Universe and close it, unless an inflationary era took place after their formation. Local

textures are insignificant in cosmology since their relative contribution to the energy density of the Universe decreases rapidly with time [11].

Even if the nontrivial topology required for the existence of a defect is absent in a field theory, it may still be possible to have defect-like solutions. Defects may be *embedded* in such topologically trivial field theories [12]. While stability of topological defects is guaranteed by topology, embedded defects are in general unstable under small perturbations.

2.2 Spontaneous Symmetry Breaking

The concept of Spontaneous Symmetry Breaking has its origin in condensed matter physics. In field theory, the rôle of the order parameter is played by scalar fields, the Higgs fields. The symmetry is said to be spontaneously broken if the ground state is characterised by a nonzero expectation value of the Higgs field and does not exhibit the full symmetry of the Hamiltonian.

The Goldstone Model

To illustrate the idea of SSB we consider the simple Goldstone model. Let ϕ be a complex scalar field with classical Lagrangian density

$$\mathcal{L} = (\partial_\mu \bar{\phi})(\partial^\mu \phi) - V(\phi) \ , \tag{5}$$

and potential $V(\phi)$:

$$V(\phi) = \frac{1}{4}\lambda[\bar{\phi}\phi - \eta^2]^2 \ , \tag{6}$$

with positive constants λ, η. This potential, Eq. (6), has the symmetry breaking *Mexican hat* shape. The Goldstone model is invariant under the U(1) group of global phase transformations,

$$\phi(x) \rightarrow e^{i\alpha}\phi(x) \ , \tag{7}$$

where α is a constant, i.e., independent of spacetime. The minima of the potential, Eq. (6), lie on a circle with fixed radius $|\phi| = \eta$; the ground state of the theory is characterised by a nonzero expectation value, given by

$$\langle 0|\phi|0\rangle = \eta e^{i\theta} \ , \tag{8}$$

where θ is an arbitrary phase. The phase transformation, Eq. (7), leads to the change $\theta \rightarrow \theta + \alpha$, which implies that the vacuum state $|0\rangle$ is not invariant under the phase transformation, Eq. (7); the symmetry is spontaneously broken. The state of unbroken symmetry with $\langle 0|\phi|0\rangle = 0$ is a local maximum of the Mexican hat potential, Eq. (6). All broken symmetry vacua, each with a different value of the phase θ are equivalent. Therefore, if we select the vacuum with $\theta = 0$, the complex scalar field ϕ can be written in terms of two real scalar fields, ϕ_1, ϕ_2, with zero vacuum expectation values, as

$$\phi = \eta + \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \ . \tag{9}$$

As a consequence, the Lagrangian density, Eq. (5), can be written as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi_1)^2 + \frac{1}{2}(\partial_\mu \phi_2)^2 - \frac{1}{2}\lambda\eta^2\phi_1^2 + \mathcal{L}_{\rm int} \ . \tag{10}$$

The last term, $\mathcal{L}_{\rm int}$, is an interaction term which includes cubic and higher-order terms in the real scalar fields ϕ_1, ϕ_2. Clearly, ϕ_1 corresponds to a massive particle, with mass $\sqrt{\lambda}\eta > 0$, while ϕ_2 corresponds to a massless scalar particle, the Goldstone boson. The appearance of Goldstone bosons is a generic feature of models with spontaneously broken global symmetries.

Going around a closed path L in physical space, the phase θ of the Higgs field ϕ develops a nontrivial winding, i.e., $\Delta\theta = 2\pi$. This closed path can be shrunk continuously to a point, only if the field ϕ is lifted to the top of its potential where it takes the value $\phi = 0$. Within a closed path for which the total change of the phase of the Higgs field ϕ is 2π, a string is trapped. A string must be either a closed loop or an infinitely long (no ends) string, since otherwise one could deform the closed path L and avoid to cross a string.

We should note that we considered above a purely classical potential, Eq. (6), to determine the expectation value of the Higgs field ϕ. In a more realistic case however, the Higgs field ϕ is a quantum field which interacts with itself, as well as with other quantum fields. As a result the classical potential $V(\phi)$ should be modified by radiative corrections, leading to an effective potential $V_{\rm eff}(\phi)$. There are models for which the radiative corrections can be neglected, while there are others for which they play an important rôle.

The Goldstone model is an example of a second-order phase transition leading to the formation of global strings, *vortices*.

The Abelian-Higgs Model

We are interested in local (gauge) strings (cosmic strings), so let us consider the simplest gauge theory with a spontaneously broken symmetry. This is the Abelian-Higgs model with Lagrangian density

$$\mathcal{L} = \bar{D}_\mu \phi \mathcal{D}^\mu \phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - V(\phi) \ , \tag{11}$$

where ϕ is a complex scalar field with potential $V(\phi)$, given by Eq. (6), and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor. The covariant derivative $\mathcal{D}_\mu$ is defined by $\mathcal{D}_\mu = \partial_\mu - ieA_\mu$, with e the gauge coupling constant and A_μ the gauge field. The Abelian-Higgs model is invariant under the group U(1) of local gauge transformations

$$\phi(x) \to e^{i\alpha(x)}\phi(x) \quad ; \quad A_\mu(x) \to A_\mu(x) + \frac{1}{e}\partial_\mu \alpha(x) \ , \tag{12}$$

where $\alpha(x)$ is a real single-valued function.

The minima of the Mexican hat potential, Eq. (6), lie on a circle of fixed radius $|\phi| = \eta$, implying that the symmetry is spontaneously broken and the complex scalar field ϕ acquires a nonzero vacuum expectation value. Following the same approach as in the Goldstone model, we chose to represent ϕ as

$$\phi = \eta + \frac{\phi_1}{\sqrt{2}} \, , \tag{13}$$

leading to the Lagrangian density

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi_1)^2 - \frac{1}{2}\mu^2 \phi_1^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} M^2 A_\mu A^\mu + \mathcal{L}_{\rm int} \, , \tag{14}$$

where the particle spectrum contains a scalar particle (Higgs boson) with mass $m_{\rm s} = \sqrt{\lambda}\eta$ and a vector field (gauge boson) with mass $m_{\rm v} = \sqrt{2}e\eta$. The breaking of a gauge symmetry does not imply a massless Goldstone boson. The Abelian-Higgs model is the simplest model which admits string solutions, the Nielsen-Olesen vortex lines. The width of the string is determined by the Compton wavelength of the Higgs and gauge bosons, which is $\sim m_{\rm s}^{-1}$ and $\sim m_{\rm v}^{-1}$, respectively.

In the Lorentz gauge, $\partial_\mu A^\mu = 0$, the Higgs field ϕ has the same form as in the case of a global string at large distances from the string core, i.e.,

$$\phi \approx \eta e^{in\theta} \, , \tag{15}$$

where the integer n denotes the string winding number. The gauge field asymptotically approaches

$$A_\mu \approx \frac{1}{ie} \partial_\mu \ln \phi \, . \tag{16}$$

The asymptotic forms for the Higgs and gauge fields, Eqs. (15) and (16) respectively, imply that far from the string core, we have

$$\mathcal{D}_\mu \phi \approx 0 \quad , \quad F_{\mu\nu} \approx 0 \, . \tag{17}$$

As a consequence, far from the string core, the energy density vanishes exponentially, while the total energy per unit length is finite. The string linear mass density μ is

$$\mu \sim \eta^2 \, . \tag{18}$$

In the case of a global U(1) string there is no gauge field to compensate the variation of the phase at large distances from the string core, resulting to a linear mass density which diverges at long distances from the string. For a global U(1) string with winding number $n = 1$ one obtains

$$\mu \sim \eta^2 + \int_\delta^R \left[\frac{1}{r} \frac{\partial \phi}{\partial \theta} \right]^2 2\pi r dr \approx 2\pi \eta^2 \ln \left(\frac{R}{\delta} \right) \, , \tag{19}$$

where δ stands for the width of the string core and R is a cut-off radius at some large distance from the string, e.g. the curvature radius of the string, or the distance to the nearest string segment in the case of a string network. The logarithmic term in the expression for the string energy mass density per unit length leads to long-range interactions between global U(1) string segments, with a force $\sim\eta^2/R$.

The field equations arising from the Lagrangian density, Eq. (11), read

$$\begin{aligned}(\partial_\mu - ieA_\mu)(\partial^\mu - ieA^\mu)\phi + \frac{\lambda}{2}\phi(\phi\bar{\phi} - \eta^2) = 0 \\ \partial_\nu F^{\mu\nu} - 2e\mathrm{Im}\left[\bar{\phi}(\partial^\nu - ieA^\nu)\phi\right] = 0\ .\end{aligned} \tag{20}$$

The equations of motion can be easily solved for the case of straight, static strings.

The internal structure of the string is meaningless when we deal with scales much larger than the string width. Thus, for a straight string lying along the z-axis, the effective energy-momentum tensor is

$$\tilde{T}^\mu_{\ \nu} = \mu\delta(x)\delta(y)\mathrm{diag}(1,0,0,1)\ . \tag{21}$$

2.3 Thermal Phase Transitions and Defect Formation

In analogy to condensed matter systems, a symmetry which is spontaneously broken at low temperatures can be restored at higher temperatures. In field theories, the expectation value of the Higgs field ϕ can be considered as a Bose condensate of Higgs particles. If the temperature T is nonzero, one should consider a thermal distribution of particles/antiparticles, in addition to the condensate. The equilibrium value of the Higgs field ϕ is obtained by minimising the free energy $F = E - TS$. Only at high enough temperatures the free energy is effectively temperature-dependent, while at low temperatures the free energy is minimised by the ordered state of the minimum energy.

Let us consider for example the Goldstone model, for which the high-temperature effective potential is

$$V_{\mathrm{eff}}(\phi, T) = m^2(T)|\phi|^2 + \frac{\lambda}{4}|\phi|^4 \quad \text{where} \quad m^2(T) = \frac{\lambda}{12}(T^2 - 6\eta^2)\ . \tag{22}$$

The effective mass-squared term $m^2(T)$ for the Higgs field ϕ in the symmetric state $\langle\phi\rangle = 0$, vanishes at the critical temperature $T_{\mathrm{c}} = \sqrt{6}\eta$. The effective potential is calculated using perturbation theory and the leading contribution comes from one-loop Feynman diagrams. For a scalar theory, the main effect is a temperature-dependent quadratic contribution to the potential. Above the critical temperature, $m^2(T)$ is positive, implying that the effective potential gets minimised at $\phi = 0$, resulting to a symmetry restoration. Below the critical temperature, $m^2(T)$ is negative, implying that the Higgs field has a nonvanishing expectation value.

Even if there is symmetry restoration and the mean value $\langle\phi\rangle$ of the Higgs field vanishes, the actual value of the field ϕ fluctuates around the mean value, meaning that ϕ at any given point is nonzero. The thermal fluctuations have, to a leading approximation, a Gaussian distribution, thus they can be characterised by a two-point correlation function, which typically decays exponentially, with a decay rate characterised by the *correlation length* ξ. The consequence of this is that fluctuations at two points separated by a distance greater than the correlation length ξ are independent.

Kibble [10] was first to estimate the initial density of topological defects formed after a phase transition followed by SSB in the context of cosmology. His criterion was based on the causality argument and the Ginzburg temperature, $T_{\rm G}$, defined as the temperature below which thermal fluctuations do not contain enough energy for regions of the field on the scale of the correlation length to overcome the potential energy barrier and restore the symmetry,

$$\xi^3(T_{\rm G})\Delta F(T_{\rm G}) \sim T_{\rm G} \ ; \tag{23}$$

ΔF is the difference in free energy density between the false and true vacua.

According to the Kibble mechanism, the initial defect network is obtained by the equilibrium correlation length of the Higgs field at the Ginzburg temperature. Consequently, laboratory tests confirmed defect formation at the end of a symmetry breaking phase transition, but they disagree with defect density estimated by Kibble. More precisely, Zurek [13,14] argued that the the relaxation time $\bar{\tau}(T)$, which is the time it takes correlations to establish on the length scale $\xi(T)$, has an important rôle in determining the initial defect density.

Let us describe the *freeze-out* proposal suggested by Zurek to estimate the initial defect density. Above the critical temperature $T_{\rm c}$, the field starts off in thermal equilibrium with a heat bath. Near the phase transition, the equilibrium correlation length diverges

$$\xi(T) = \xi_0 \left(\frac{T - T_{\rm c}}{T_{\rm c}} \right)^{-\nu} , \tag{24}$$

where ν denotes the critical component. At the same time, the dynamics of the system becomes slower, and this can be expressed in terms of the equilibrium relaxation timescale of the field, which also diverges, but with a different exponent μ:

$$\bar{\tau}(T) = \bar{\tau}_0 \left(\frac{T - T_{\rm c}}{T_{\rm c}} \right)^{-\mu} . \tag{25}$$

The values of the critical components μ, ν depend on the theory under consideration. Assuming, for simplicity, that the temperature is decreasing linearly,

$$T(t) = \left(1 - \frac{t}{\bar{\tau}_{\rm Q}} \right) T_{\rm c} \ , \tag{26}$$

where the *quench timescale* $\bar{\tau}_{\rm Q}$ characterises the cooling rate.

As the temperature decreases towards its critical temperature $T_{\rm c}$, the correlation length ξ grows as

$$\xi(t) \sim \left(\frac{|t|}{\bar{\tau}_{\rm Q}} \right)^{-\nu} , \tag{27}$$

but at the same time the dynamics of the system becomes slower,

$$\bar{\tau}(t) \sim \left(\frac{|t|}{\bar{\tau}_{\rm Q}} \right)^{-\mu} . \tag{28}$$

As the system approaches from above the critical temperature, there comes a time $|\hat{t}|$ during the quench when the equilibrium relaxation timescale equals the time that is left before the transition at the critical temperature, namely

$$\bar{\tau}(\hat{t}) = |\hat{t}| . \tag{29}$$

After this time, the system can no longer adjust fast enough to the change of the temperature of the thermal bath, and falls out of equilibrium. At time $\hat{t}$, the dynamics of the correlation length freezes. The correlation length cannot grow significantly after this time, and one can safely state that it freezes to its value at time $\hat{t}$. Thus, according to Zurek's proposal the initial defect density is determined by the freeze-out scale [13, 14]

$$\hat{\xi} \equiv \xi(\hat{t}) \sim \bar{\tau}_{\rm Q}^{\nu/(1+\mu)} . \tag{30}$$

We note that the above discussion is in the framework of second-order phase transitions.

The above prediction, Eq. (30), has been tested experimentally in a variety of systems, as for example, in superfluid ^{4}He [15, 16] and ^{3}He [17, 18], and in liquid crystals [19–21]. Apart the experimental support, the Kibble-Zurek picture is supported by numerical simulations [22–24] and calculations using the methods of nonequilibrium quantum field theory [25].

2.4 Cosmic String Dynamics

The world history of a string can be expressed by a two-dimensional surface in the four-dimensional spacetime, which is called the string worldsheet:

$$x^\mu = x^\mu(\zeta^a) , \quad a = 0, 1 ; \tag{31}$$

the worldsheet coordinates ζ^0, ζ^1 are arbitrary parameters chosen so that ζ^0 is timelike and ζ^1 spacelike ($\equiv \sigma$).

The string equations of motion, in the limit of a zero thickness string, are derived from the Goto-Nambu effective action which, up to an overall factor, corresponds to the surface area swept out by the string in spacetime:

$$S_0[x^\mu] = -\mu \int \sqrt{-\gamma} d^2\zeta \ , \tag{32}$$

where γ is the determinant of the two-dimensional worldsheet metric γ_{ab},

$$\gamma = \det(\gamma_{ab}) = \frac{1}{2}\epsilon^{ac}\epsilon^{bd}\gamma_{ab}\gamma_{cd} \quad , \quad \gamma_{ab} = g_{\mu\nu}x^\mu_{,a}x^\nu_{,b} \ . \tag{33}$$

If the string curvature is small but not negligible, one may consider an expansion in powers of the curvature, leading to the following form for the string action up to second order

$$S = -\int d^2\zeta\sqrt{-\gamma}(\mu - \beta_1 K^A K^A + \beta_2 R) \ , \tag{34}$$

where β_1, β_2 are dimensionless numbers. The Ricci curvature scalar R is a function of the extrinsic curvature tensor K^A_{ab} (with $A = 1, 2$),

$$R = K^{abA}K^A_{ab} - K^A K^A \ ; \tag{35}$$

$K^A = \gamma^{ab}K^A_{ab}$. Finite corrections and their effects to the effective action have been studied by a number of authors [26–29].

By varying the action, Eq. (32), with respect to $x^\mu(\zeta^a)$, and using the relation $d\gamma = \gamma\gamma^{ab}d\gamma_{ab}$, where γ_{ab} is given by Eq. (33b), one gets the string equations of motion:

$$x^\mu_{,a}{}^{;a} + \Gamma^\mu_{\nu\sigma}\gamma^{ab}x^\nu_{,a}x^\sigma_{,b} = 0 \ , \tag{36}$$

where $\Gamma^\mu_{\nu\sigma}$ is the four-dimensional Christoffel symbol,

$$\Gamma^\mu_{\nu\sigma} = \frac{1}{2}g^{\mu\tau}(g_{\tau\nu,\sigma} + g_{\tau\sigma,\nu} - g_{\nu\sigma,\tau}) \ , \tag{37}$$

and the covariant Laplacian is

$$x^\mu_{,a}{}^{;a} = \frac{1}{\sqrt{-\gamma}}\partial_a(\sqrt{-\gamma}\gamma^{ab}x^\mu_{,b}) \ . \tag{38}$$

One can derive the same string equations of motion by using Polyakov's form of the action [30]

$$S[x^\mu, h_{ab}] = -\frac{\mu}{2}\int \sqrt{-h}h^{ab}\gamma_{ab}d^2\zeta \ , \tag{39}$$

where h_{ab} is the internal metric with determinant h.

Including a force of friction $F^{\mu\nu}$ due to the scattering of thermal particles off the string, the equation of motion reads [31]

$$\mu\big[x^\mu_{,a}{}^{;a} + \Gamma^\mu_{\nu\sigma}\gamma^{ab}x^\nu_{,a}x^\sigma_{,b}\big] = F^\mu(u^\lambda_\perp, T, \sigma) \ . \tag{40}$$

The force of friction depends on the temperature of the surrounding matter T, the velocity of the fluid transverse to the world sheet $u^\nu_\perp \equiv u^\nu - x^\nu_{,a}x^{\sigma,\alpha}u_\sigma$,

and the type of interaction between the particles and the string, which we represent by σ. Cosmic strings of mass per unit length μ would have formed at cosmological time

$$t_0 \sim (G\mu)^{-1} t_{\rm Pl} \,, \tag{41}$$

where $t_{\rm Pl}$ is the Planck time. Immediately after the phase transition the string dynamics would be dominated by friction [31], until a time of order

$$t_\star \sim (G\mu)^{-2} t_{\rm Pl} \,. \tag{42}$$

For cosmic strings formed at the grand unification scale, their mass per unit length is of order $G\mu \sim 10^{-6}$ and friction is important only for a very short period of time. However, if strings have formed at a later phase transition, for example closer to the electroweak scale, their dynamics would be dominated by friction through most of the thermal history of the Universe. The evolution of cosmic strings taking into account the frictional force due to the surrounding radiation has been studied in Ref. [32].

The string energy-momentum tensor can be obtained by varying the action, Eq. (32), with respect to the metric $g_{\mu\nu}$,

$$T^{\mu\nu}\sqrt{-g} = -2\frac{\delta S}{\delta g_{\mu\nu}} = \mu \int d^2\zeta \sqrt{-\gamma}\gamma^{ab} x^\mu_{,a} x^\nu_{,b} \delta^{(4)}(x^\sigma - x^\sigma(\zeta^a)) \,. \tag{43}$$

For a straight cosmic string in a flat spacetime lying along the z-axis and choosing $\zeta^0 = t$, $\zeta^1 = z$, the above expression reduces to the one for the effective energy-momentum tensor, Eq. (21).

Cosmic Strings in Curved Spacetime

The equations of motion for strings are most conveniently written in comoving coordinates, where the FLRW metric takes the form

$$ds^2 = a^2(\tau)[d\tau^2 - d\mathbf{r}^2] \,. \tag{44}$$

The comoving spatial coordinates of the string, $\mathbf{x}(\tau, \sigma)$, are written as a function of conformal time τ and the length parameter σ. We have thus chosen the gauge condition $\zeta^0 = \tau$. For a cosmic string moving in a FLRW Universe, the equations of motion, Eq. (36), can be simplified by also choosing the gauge in which the unphysical parallel components of the velocity vanish,

$$\dot{\mathbf{x}} \cdot \mathbf{x}' = 0 \,, \tag{45}$$

where overdots denote derivatives with respect to conformal time τ and primes denote spatial derivatives with respect to σ.

In these coordinates, the Goto-Nambu action yields the following equations of motion for a string moving in a FLRW metric:

$$\ddot{\mathbf{x}} + 2\left(\frac{\dot{a}}{a}\right)\dot{\mathbf{x}}(1-\dot{\mathbf{x}}^2) = \left(\frac{1}{\epsilon}\right)\left(\frac{\mathbf{x}'}{\epsilon}\right)' . \tag{46}$$

The string energy per unit σ, in comoving units, is $\epsilon \equiv \sqrt{\mathbf{x}'^2/(1-\dot{\mathbf{x}}^2)}$, implying that the string energy is $\mu a \int \epsilon d\sigma$. Equation (46) leads to

$$\frac{\dot{\epsilon}}{\epsilon} = -2\frac{\dot{a}}{a}\dot{\mathbf{x}}^2 . \tag{47}$$

One usually fixes entirely the gauge by choosing σ so that $\epsilon = 1$ initially.

Cosmic Strings in Flat Spacetime

In flat spacetime spacetime, the string equations of motion take the form

$$\partial_a(\sqrt{-\gamma}\gamma^{ab}x^{\mu}_{,b}) = 0 . \tag{48}$$

We impose the conformal gauge

$$\dot{x}\cdot x' = 0 \quad , \quad \dot{x}^2 + x'^2 = 0 , \tag{49}$$

where overdots denote derivatives with respect to ζ^0 and primes denote derivatives with respect to ζ^1. In this gauge the string equations of motion is just a two-dimensional wave equation,

$$\ddot{\mathbf{x}} - \mathbf{x}'' = 0 . \tag{50}$$

To fix entirely the gauge, we also impose

$$t \equiv x^0 = \zeta^0 , \tag{51}$$

which allows us to write the string trajectory as the three dimensional vector $\mathbf{x}(\sigma, t)$, where $\zeta^1 \equiv \sigma$, the spacelike parameter along the string. This implies that the constraint equations, Eq. (49), and the string equations of motion, Eq. (50), become

$$\begin{aligned} \dot{\mathbf{x}}\cdot\mathbf{x}' &= 0 \\ \dot{\mathbf{x}}^2 + \mathbf{x}'^2 &= 1 \\ \ddot{\mathbf{x}} - \mathbf{x}'' &= 0 . \end{aligned} \tag{52}$$

The above equations imply that the string moves perpendicularly to itself with velocity $\dot{\mathbf{x}}$, that σ is proportional to the string energy, and that the string acceleration in the string rest frame is inversely proportional to the local string curvature radius. A curved string segment tends to straighten itself, resulting to string oscillations.

The general solution to the string equation of motion in flat spacetime, Eq. (52c), is

$$\mathbf{x} = \frac{1}{2}[\mathbf{a}(\sigma - t) + \mathbf{b}(\sigma + t)] \,, \tag{53}$$

where $\mathbf{a}(\sigma - t)$ and $\mathbf{b}(\sigma + t)$ are two continuous arbitrary functions which satisfy

$$\mathbf{a}'^2 = \mathbf{b}'^2 = 1 \,. \tag{54}$$

Thus, σ is the length parameter along the three-dimensional curves $\mathbf{a}(\sigma), \mathbf{b}(\sigma)$.

Cosmic String Intercommutations

The Goto-Nambu action describes to a good approximation cosmic string segments which are separated. However, it leaves unanswered the issue of what happens when strings cross. Numerical simulations have shown that the ends of strings exchange partners, *intercommute*, with probability equal to 1. These results have been confirmed for global [33], local [34], and superconducting [35] strings.

String-string and self-string intersections leading to the formation of new long strings and loops are drawn in Fig. 1. Clearly string intercommutations produce discontinuities in $\dot{\mathbf{x}}$ and $\mathbf{x}'$ on the new string segments at the intersection point. These discontinuities, *kinks*, are composed of right- and left-moving pieces travelling along the string at the speed of light.

2.5 Cosmic String Evolution

Early analytic work [37] identified the key property of *scaling*, where at least the basic properties of the string network can be characterised by a single length scale, roughly the persistence length or the interstring distance ξ, which grows with the horizon. This result was supported by subsequent numerical work [38]. However, further investigation revealed dynamical processes, including loop production, at scales much smaller than ξ [39, 40].

The cosmic string network can be divided into long (infinite) strings and small loops. The energy density of long strings in the scaling regime is given by (in the radiation era)

$$\rho_{\mathrm{L}} = \tilde{\kappa}\mu t^{-2} \,, \tag{55}$$

where $\tilde{\kappa}$ is a numerical coefficient ($\tilde{\kappa} = 20 \pm 10$). The small loops, their size distribution, and the mechanism of their formation remained for years the least understood parts of the string evolution.

Assuming that the long strings are characterised by a single length scale $\xi(t)$, one gets

$$\xi(t) = \left(\frac{\rho_{\mathrm{L}}}{\mu}\right)^{-1/2} = \tilde{\kappa}^{-1/2} t \,. \tag{56}$$

Thus, the typical distance between the nearest string segments and the typical curvature radius of the strings are both of the order of ξ. Early numerical simulations have shown that indeed the typical curvature radius of long strings

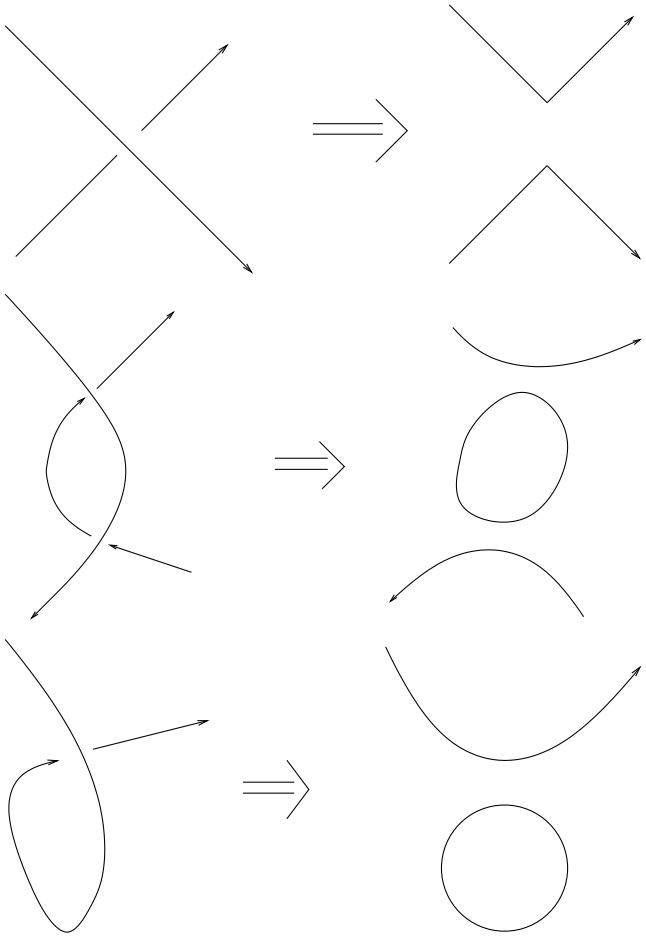

Fig. 1. Illustration of string intersections: (**a**) string-string intersection in one point, leading to the formation of two new long strings via exchange of partners; (**b**) string-string intersections in two points, leading to the formation of two new long strings via exchange of partners, and one closed loop; and (**c**) self-string intersections leading to the formation of one long string and a closed string loop [36]

and the characteristic distance between the strings are both comparable to the evolution time t. Clearly, these results agree with the picture of the scale-invariant evolution of the string network and with the one-scale hypothesis.

However, the numerical simulations have also shown [40, 41] that small-scale processes (such as the production of small closed loops) play an essential rôle in the energy balance of long strings. The existence of an important small scale in the problem was also indicated [40] by an analysis of the string shapes. In response to these findings, a three-scale model was developed [42], which describes the network in terms of three scales, namely the usual energy

density scale ξ, a correlation length $\bar{\xi}$ along the string, and a scale ζ relating to local structure on the string. The small-scale structure (wiggliness), which offers an explanation for the formation of the small sized loops, is basically developed through intersections of long string segments. It seemed likely from the three-scale model that ξ and $\bar{\xi}$ would scale, with ζ growing slowly, if at all, until gravitational radiation effects became important when $\zeta/\xi \approx 10^{-4}$ [43, 44]. Thus, according to the three-scale model, the small length scale may reach scaling only if one considers the gravitational back reaction effect. Aspects of the three-scale model have been checked [45] evolving a cosmic string network in Minkowski spacetime. However, it was found that loops are produced with tiny sizes, which led the authors to suggest [45] that the dominant mode of energy loss of a cosmic string network is particle production and not gravitational radiation as the loops collapse almost immediately. One can find in the literature studies which support [46] this finding, and others which they do not [47, 48].

Very recently, numerical simulations of cosmic string evolution in a FLRW Universe (*see*, Fig. 2), found evidence [49] of a scaling regime for the cosmic string loops in the radiation and matter dominated eras down to the hundredth of the horizon time. It is important to note that the scaling was found without considering any gravitational back reaction effect; it was just the result of string intercommuting mechanism. As it was reported in Ref. [49], the scaling regime of string loops appears after a transient relaxation era,

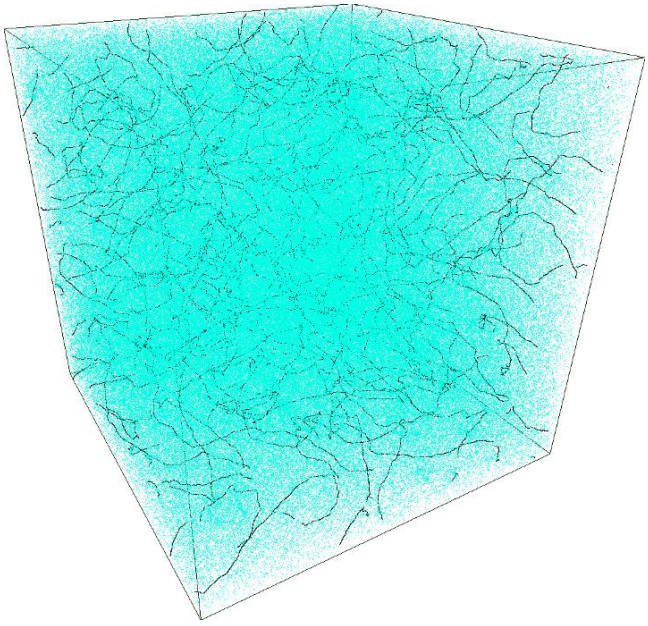

Fig. 2. Snapshot of a network of long strings and closed loops in the matter era [49]

driven by a transient overproduction of string loops with length close to the initial correlation length of the string network. Calculating the amount of energy momentum tensor lost from the string network, it was found [49] that a few percents of the total string energy density disappear in the very brief process of formation of numerically unresolved string loops during the very first timesteps of the string evolution. Subsequently, two other studies support these findings [50], [51].

2.6 String Thermodynamics

It is one of the basic facts about string theory that the degeneracy of string states increases exponentially with energy,

$$d(E) \sim e^{\beta_{\rm H} E} \ . \tag{57}$$

A consequence of this is that there is a maximum temperature $T_{\rm max} = 1/\beta_{\rm H}$, the Hagedorn temperature [52–54]. In the microcanonical ensemble the description of this situation is as follows: Consider a system of closed string loops in a three-dimensional box. Intersecting strings intercommute, but otherwise they do not interact and are described by the Goto-Nambu equations of motion. The statistical properties of a system of strings in equilibrium are characterised by only one parameter, the energy density of strings, ρ,

$$\rho = \frac{E}{L^3} \ , \tag{58}$$

where L denotes the size of the cubical box. The behaviour of the system depends on whether it is at low or high energy densities, and it undergoes a phase transition at a critical energy density, the Hagedorn energy density $\rho_{\rm H}$. Quantisation implies a lower cutoff for the size of the string loops, determined by the string tension μ. The lower cutoff on the loop size is roughly $\mu^{-1/2}$, implying that the mass of the smallest string loops is $m_0 \sim \mu^{1/2}$.

For a system of strings at the low energy density regime ($\rho \ll \rho_{\rm H}$), all strings are chopped down to the loops of the smallest size, while larger loops are exponentially suppressed. Thus, for small enough energy densities, the string equilibrium configuration is dominated by the massless modes in the quantum description. The energy distribution of loops, given by the number dn of loops with energies between E and $E + dE$ per unit volume, is [53–55]

$$dn \propto e^{-\alpha E} E^{-5/2} dE \ \ (\rho \ll \rho_{\rm H}) \ , \tag{59}$$

where $\alpha = (5/2m_0) \ln(\rho_{\rm H}/\rho)$.

However, as we increase the energy density, more and more oscillatory modes of strings get excited. In particular, if we reach a critical energy density, $\rho_{\rm H}$, then long oscillatory string states begin to appear in the equilibrium state. The density at which this happens corresponds to the Hagedorn temperature.

The Hagedorn energy density $\rho_{\rm H}$, achieved when the separation between the smallest string loops is of the order of their sizes, is $\rho_{\rm c} \sim m_0^4$. At the Hagedorn energy density the system undergoes a phase transition characterised by the appearance of infinitely long strings.

At the high energy density regime ($\rho \gg \rho_{\rm c}$), the energy distribution of string loops is [53–55]

$$dn = Am_0^{9/2} E^{-5/2} dE \quad (\rho \gg \rho_{\rm H}) \,, \tag{60}$$

where A is a numerical coefficient independent of m_0 and of ρ. Equation (60) implies that the mean-square radius R of the closed string loops is

$$R \sim m_0^{-3/2} E^{1/2} \,, \tag{61}$$

meaning that large string loops are random walks of step $\sim m_0^{-1}$. Equations (60) and (61) imply

$$dn = A' R^{-4} dR \quad (\rho \gg \rho_{\rm H}) \,, \tag{62}$$

where A' is a numerical constant. From Eq. (62) one concludes that the distribution of closed string loops is scale invariant, since it does not depend on the cutoff parameter m_0.

The total energy density in finite string loops is independent of ρ. Increasing the energy density ρ of the system of strings, the extra energy $E - E_{\rm H}$, where $E_{\rm H} = \rho_{\rm H} L^3$, goes into the formation of infinitely long strings, implying

$$\rho - \rho_{\rm inf} = \text{const} \quad (\rho \gg \rho_{\rm H}) \,, \tag{63}$$

where $\rho_{\rm inf}$ denotes the energy density in infinitely long strings.

Clearly, the above analysis describes the behaviour of a system of strings of low or high energy densities, while there is no analytic description of the phase transition and of the intermediate densities around the critical one, $\rho \sim \rho_{\rm H}$. An experimental approach to the problem has been proposed in Ref. [56] and later extended in Ref. [57].

The equilibrium properties of a system of cosmic strings have been studied numerically in Ref. [56]. The strings are moving in a three-dimensional flat space and the initial string states are chosen to be a *loop gas* consisting of the smallest two-point loops with randomly assigned positions and velocities. This choice is made just because it offers an easily adjustable string energy density. Clearly, the equilibrium state is independent of the initial state. The simulations revealed a distinct change of behaviour at a critical energy density $\rho_{\rm H} = 0.0172 \pm 0.002$. For $\rho < \rho_{\rm H}$, there are no infinitely long strings, thus their energy density, $\rho_{\rm inf}$, is just zero. For $\rho > \rho_{\rm H}$, the energy density in finite strings is constant, equal to $\rho_{\rm H}$, while the extra energy goes to the infinitely long strings with energy density $\rho_{\rm inf} = \rho - \rho_{\rm H}$. Thus, Eqs. (60) and (63) are valid for all $\rho > \rho_{\rm H}$, although they were derived only in the limit $\rho \gg \rho_{\rm H}$. At the critical energy density, $\rho = \rho_{\rm H}$, the system of strings is scale-invariant. At bigger energy densities, $\rho > \rho_{\rm H}$, the energy distribution of closed string loops

at different values of ρ were found [56] to be identical within statistical errors, and well defined by a line $dn/dE \propto E^{-5/2}$. Thus, for $\rho > \rho_{\rm H}$, the distribution of finite strings is still scale-invariant, but in addition the system includes the infinitely long strings, which do not exhibit a scale-invariant distribution. The number distribution for infinitely long strings goes as $dn/dE \propto 1/E$, which means that the total number of infinitely long strings is roughly $\log(E - E_{\rm H})$. So, typically the number of long strings grows very slowly with energy; for $\rho > \rho_{\rm H}$ there are just a few infinitely long strings, which take up most of the energy of the system.

The above numerical experiment has been extended [57] for strings moving in a higher dimensional box. The Hagedorn energy density was found for strings moving in boxes of dimensionality $d_{\rm B} = 3, 4, 5$ [57]:

$$\rho_{\rm H} = \begin{cases} 0.172 \pm 0.002 & \text{for} \quad d_{\rm B} = 3 \\ 0.062 \pm 0.001 & \text{for} \quad d_{\rm B} = 4 \\ 0.031 \pm 0.001 & \text{for} \quad d_{\rm B} = 5 \end{cases} \tag{64}$$

Moreover, the size distribution of closed finite string loops at the high energy density regime was found to be independent of the particular value of ρ for a given dimensionality of the box $d_{\rm B}$. The size distribution of finite closed string loops was found [57] to be well defined by a line

$$\frac{dn}{dE} \sim E^{-(1+d_{\rm B}/2)} \ , \tag{65}$$

where the space dimensionality $d_{\rm B}$ was taken equal to 3, 4, or 5 . The statistical errors indicated a slope equal to $-(1+d_{\rm B}/2)\pm 0.2$. Above the Hagedorn energy density the system is again characterised by a scale-invariant distribution of finite closed string loops and a number of infinitely long strings with a distribution which is not scale invariant.

2.7 Genericity of Cosmic Strings Formation within SUSY GUTs

The Standard Model (SM), even though it has been tested to a very high precision, is incapable of explaining neutrino masses [58–60]. An extension of the SM gauge group can be realised within Supersymmetry (SUSY). SUSY offers a solution to the gauge hierarchy problem, while in the supersymmetric extension of the standard model the gauge coupling constants of the strong, weak and electromagnetic interactions meet at a single point, $M_{\rm GUT} \simeq (2-3) \times 10^{16}$ GeV. In addition, SUSY GUTs provide the scalar field which could drive inflation, explain the matter-antimatter asymmetry of the Universe, and propose a candidate, the lightest superparticle, for cold dark matter. We will address the question of whether cosmic string formation is generic, in the context of SUSY GUTs. Within SUSY GUTs there is a large number of SSB patterns leading from a large gauge group G to the SM gauge group $G_{\rm SM} \equiv SU(3)_{\rm C} \times SU(2)_{\rm L} \times U(1)_{\rm Y}$. The study of the homotopy group of the

false vacuum for each SSB scheme determines whether there is defect formation and identifies the type of the formed defect. Clearly, if there is formation of domain walls or monopoles, one will have to place an era of supersymmetric hybrid inflation to dilute them. To consider a SSB scheme as a successful one, it should be able to explain the matter/anti-matter asymmetry of the Universe and to account for the proton lifetime measurements [58].

In what follows, we consider a mechanism of baryogenesis via leptogenesis, which can be thermal or nonthermal one. In the case of nonthermal leptogenesis, $U(1)_{B-L}$ (B and L, are the baryon and lepton numbers, respectively) is a sub-group of the GUT gauge group, G_{GUT}, and B-L is broken at the end or after inflation. If one considers a mechanism of thermal leptogenesis, B-L is broken independently of inflation. If leptogenesis is thermal and B-L is broken before the inflationary era, then one should check whether the temperature at which B-L is broken – this temperature defines the mass of the right-handed neutrinos – is smaller than the reheating temperature. To have a successful inflationary cosmology, the reheating temperature should be lower than the limit imposed by the gravitino. To ensure the stability of proton, the discrete symmetry Z_2, which is contained in $U(1)_{B-L}$, must be kept unbroken down to low energies. Thus, the successful SSB schemes should end at $G_{SM} \times Z_2$. Taking all these considerations into account we will examine within all acceptable SSB patterns, how often cosmic strings form at the end of the inflationary era.

To proceed, one has to first choose the large gauge group G_{GUT}. In Ref. [61] this study has been done in detail for a large number of simple Lie groups. Considering GUTs based on simple gauge groups, the type of supersymmetric hybrid inflation will be of the F-type. The minimum rank of G_{GUT} has to be at least equal to 4, to contain the G_{SM} as a subgroup. Then one has to study the possible embeddings of G_{SM} in G_{GUT} so that there is an agreement with the SM phenomenology and especially with the hypercharges of the known particles. Moreover, the large gauge group G_{GUT} must include a complex representation, needed to describe the SM fermions, and it must be anomaly free. In principle, $SU(n)$ may not be anomaly free. We thus assume that all $SU(n)$ groups we consider have indeed a fermionic representation which certifies that the model is anomaly free. We set as the upper bound on the rank r of the group, $r \leq 8$. Clearly, the choice of the maximum rank is in principle arbitrary. This choice could, in a sense, be motivated by the Horava-Witten [62] model, based on $E_8 \times E_8$. Concluding, the large gauge group G_{GUT} could be one of the following: SO(10), E_6, SO(14), SU(8), SU(9); flipped SU(5) and $[SU(3)]^3$ are included within this list as subgroups of SO(10) and E_6, respectively.

A detailed study of all SSB schemes which bring us from G_{GUT} down to the SM gauge group G_{SM}, by one or more intermediate steps, shows that cosmic strings are generically formed at the end of hybrid inflation. If the large gauge group G_{GUT} is SO(10) then cosmic strings formation is unavoidable [61].

The genericity of cosmic string formation for E_6 depends whether one considers thermal or nonthermal leptogenesis. More precisely, under the assumption of nonthermal leptogenesis, cosmic string formation is unavoidable. Considering thermal leptogenesis, cosmic strings formation at the end of hybrid inflation arises in 98% of the acceptable SSB schemes [63]. Finally, if the requirement of having Z_2 unbroken down to low energies is relaxed and thermal leptogenesis is considered as the mechanism for baryogenesis, cosmic string formation accompanies hybrid inflation in 80% of the SSB schemes.

The SSB schemes of SU(6) and SU(7) down to the G_{SM} which could accommodate an inflationary era with no defect (of any kind) at later times are inconsistent with proton lifetime measurements. Minimal SU(6) and SU(7) do not predict neutrino masses [61], implying that these models are incompatible with high energy physics phenomenology.

Higher rank groups, namely SO(14), SU(8) and SU(9), should in general lead to cosmic string formation at the end of hybrid inflation. In all these schemes, cosmic string formation is sometimes accompanied by the formation of embedded strings. The strings which form at the end of hybrid inflation have a mass which is proportional to the inflationary scale.

3 Cosmic Microwave Background Temperature Anisotropies

The CMB temperature anisotropies offer a powerful test for theoretical models aiming at describing the early Universe. The characteristics of the CMB multipole moments can be used to discriminate among theoretical models and to constrain the parameters space.

The spherical harmonic expansion of the CMB temperature anisotropies, as a function of angular position, is given by

$$\frac{\delta T}{T}(\mathbf{n}) = \sum_{\ell m} a_{\ell m} \mathcal{W}_\ell Y_{\ell m}(\mathbf{n}) \quad \text{with} \quad a_{\ell m} = \int \mathrm{d}\Omega_{\mathbf{n}} \frac{\delta T}{T}(\mathbf{n}) Y^*_{\ell m}(\mathbf{n}) \; ; \tag{66}$$

$\mathcal{W}_\ell$ stands for the ℓ-dependent window function of the particular experiment. The angular power spectrum of CMB temperature anisotropies is expressed in terms of the dimensionless coefficients C_ℓ, which appear in the expansion of the angular correlation function in terms of the Legendre polynomials P_ℓ:

$$\left\langle 0 \left| \frac{\delta T}{T}(\mathbf{n}) \frac{\delta T}{T}(\mathbf{n}') \right| 0 \right\rangle \Big|_{(\mathbf{n}\cdot\mathbf{n}'=\cos\vartheta)} = \frac{1}{4\pi} \sum_\ell (2\ell+1) C_\ell P_\ell(\cos\vartheta) \mathcal{W}_\ell^2 \; , \tag{67}$$

where we have used the addition theorem of spherical harmonics, i.e.,

$$\sum_{m=-\ell}^{\ell} Y_{\ell m}(\mathbf{n}) Y^\star_{\ell m}(\mathbf{n}') = \frac{2\ell+1}{4\pi} P_\ell(\mathbf{n}\cdot\mathbf{n}') \; . \tag{68}$$

It compares points in the sky separated by an angle ϑ. Here, the brackets denote spatial average, or expectation values if perturbations are quantised. Equation (67) holds only if the initial state for cosmological perturbations of quantum-mechanical origin is the vacuum [8,9]. The value of C_ℓ is determined by fluctuations on angular scales of the order of π/ℓ. The angular power spectrum of anisotropies observed today is usually given by the power per logarithmic interval in ℓ, plotting $\ell(\ell+1)C_\ell$ versus ℓ.

To find the power spectrum induced by topological defects, one has to solve, in Fourier space for each given wave vector $\mathbf{k}$ a system of linear perturbation equations with random sources:

$$\mathcal{D}X = \mathcal{S}\,, \tag{69}$$

where $\mathcal{D}$ denotes a time dependent linear differential operator, X is a vector which contains the various matter perturbation variables, and $\mathcal{S}$ is the random source term, consisting of linear combinations of the energy momentum tensor of the defect. For given initial conditions, Eq. (69) can be solved by means of a Green's function, $\mathcal{G}(\tau,\tau')$, in the form

$$X_j(\tau_0,\mathbf{k}) = \int_{\tau_{in}}^{\tau_0} d\tau \mathcal{G}_{jm}(\tau_0,\tau,\mathbf{k})\mathcal{S}_m(\tau,\mathbf{k})\,. \tag{70}$$

To compute power spectra or, more generally, quadratic expectation values of the form $\langle X_j(\tau_0,\mathbf{k})X_m^*(\tau_0,\mathbf{k}')\rangle$, one has to calculate

$$\begin{aligned}&\langle X_j(\tau_0,\mathbf{k})X_l^\star(\tau_0,\mathbf{k}')\rangle \\ &= \int_{\tau_{in}}^{\tau_0} d\tau \mathcal{G}_{jm}(\tau,\mathbf{k}) \int_{\tau_{in}}^{\tau_0} d\tau' \mathcal{G}_{ln}^\star(\tau',\mathbf{k}') \times \langle \mathcal{S}_m(\tau,\mathbf{k})\mathcal{S}_n^\star(\tau',\mathbf{k}')\rangle\,. \end{aligned} \tag{71}$$

Thus, to compute power spectra, one should know the unequal time two-point correlators $\langle \mathcal{S}_m(\tau,\mathbf{k})\mathcal{S}_n^\star(\tau',\mathbf{k}')\rangle$ in Fourier space [64]. This object is calculated by means of heavy numerical simulations.

The CMB temperature anisotropies provide a powerful tool to discriminate among inflation and topological defects. On large angular scales ($\ell \lesssim 50$), both families of models lead to approximately scale-invariant spectra, with however a different prediction regarding the statistics of the induced perturbations. Provided the quantum fields are initially placed in the vacuum, inflation predicts generically Gaussian fluctuations, whereas in the case of topological defect models, the induced perturbations are clearly nongaussian, at least at sufficiently high angular resolution. This is an interesting fingerprint, even though difficult to test through the data. In the context of inflation, nongaussianity can however also be present, as for example in the case of stochastic inflation [65], or in a class of inflationary models involving two scalar fields leading to nongaussian isothermal fluctuations with a blue spectrum [66]. In addition, allowing nonvacuum initial states for the cosmological perturbations of quantum-mechanical origin, one generically obtains a non-Gaussian spectrum [8,9], in the context of single-field inflation.

On intermediate and small angular scales however, the predictions of inflation are quite different than those of topological defect models, due to the different nature of the induced perturbations. On the one hand, the inflationary fluctuations are coherent, in the sense that the perturbations are initially at the same phase and subsequently evolve linearly and independently of each other. The subsequent progressive phase shift between different modes produces the *acoustic peak* structure. On the other hand, in topological defect models, fluctuations are constantly induced by the sources (defects). Since topological defects evolve in a nonlinear manner, and since the random initial conditions of the source term in the perturbation equations of a given scale leaks into other scales, perfect coherence is destroyed. The predictions of the defects models regarding the characteristics of the CMB spectrum are:

- Global $\mathcal{O}(4)$ textures lead to a position of the first acoustic peak at $\ell \simeq 350$ with an amplitude $\sim$1.5 times higher than the Sachs-Wolfe plateau [67].
- Global $\mathcal{O}(N)$ textures in the large N limit lead to a quite flat spectrum, with a slow decay after $\ell \sim 100$ [68]. Similar are the predictions of other global $\mathcal{O}(N)$ defects [69, 70].
- Local cosmic strings simulations [71] found a broad peak at $\ell \approx 150 - 400$, being produced from both vector and scalar modes, which peaks at $\ell \approx 180$ and $\ell \approx 400$ respcetively.

The position and amplitude of the acoustic peaks, as found by the CMB measurements [72–75], are in disagreement with the predictions of topological defect models. As a consequence, CMB measurements rule out pure topological defect models as the origin of initial density perturbations leading to the observed structure formation.

3.1 Mixed Models

Since cosmic strings are expected to be generically formed in the context of SUSY GUTs, one should consider *mixed perturbation models* where the dominant rôle is played by the inflaton field but cosmic strings have also a contribution, small but not negligible. Restricting ourselves to the angular power spectrum, we can remain in the linear regime. In this case,

$$C_\ell = \alpha C_\ell^{\rm I} + (1-\alpha) C_\ell^{\rm S} \ , \tag{72}$$

where $C_\ell^{\rm I}$ and $C_\ell^{\rm S}$ denote the (COBE normalized) Legendre coefficients due to adiabatic inflaton fluctuations and those stemming from the cosmic string network, respectively. The coefficient α in Eq. (72) is a free parameter giving the relative amplitude for the two contributions. Comparing the C_ℓ, calculated using Eq. (72) – where $C_\ell^{\rm I}$ is taken from a generic inflationary model and $C_\ell^{\rm S}$ from numerical simulations of cosmic string networks – with data obtained from the most recent CMB measurements, one gets that a cosmic string contribution to the primordial fluctuations higher than 14% is excluded up to 95% confidence level [76–78] (*see*, Fig. 3).

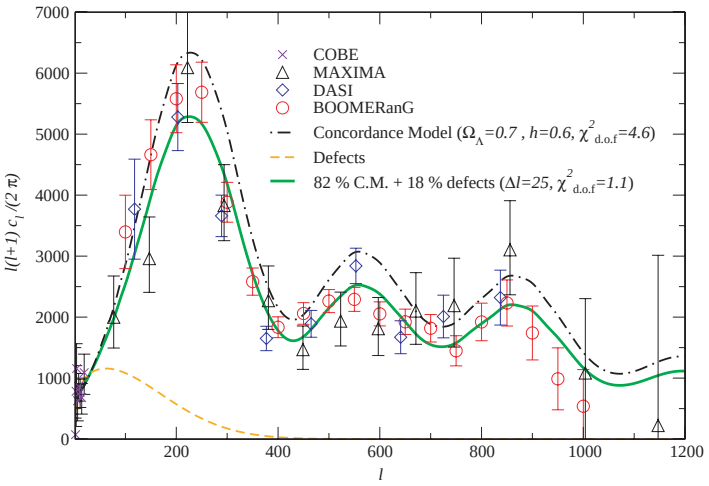

Fig. 3. $\ell(\ell+1)C_\ell$ versus ℓ for three different models. The string contribution turns out to be $\sim$18% of the total [76]

In what follows, we follow a conservative approach and do not allow cosmic strings to contribute more than 10% to the CMB temperature anisotropies.

3.2 Supersymmetric Hybrid Inflation

Inflation offers simple answers to the shortcomings of the standard hot big bang model. In addition, simple inflationary models offer successful candidates for the initial density fluctuations leading to the observed structure formation.

One crucial question though is to answer how generic is the onset of inflation [4–6] and to find consistent and natural models of inflation from the point of view of particle physics. Even though one can argue that the initial conditions which favor inflationary models are the likely outcome of the quantum era before inflation [4], one should then show that inflation will last long enough to solve the shortcomings of the standard hot big bang model [5, 6]. In addition, to find natural ways to guarantee the flatness of the inflaton potential remains a difficult task. Inflation is, unfortunately, still a paradigm in search of a model. It is thus crucial to identify successful but natural inflationary models motivated from high energy physics.

In what follows we discuss two well-studied inflationary models in the framework of supersymmetry, namely F/D-term inflation. Our aim is to check the compatibility of these models – here cosmic string inflation is generic – with the CMB and gravitino constraints.

F-term Inflation

F-term inflation can be naturally accommodated in the framework of GUTs, when a GUT gauge group G_{GUT} is broken down to the G_{SM} at an energy M_{GUT}, according to the scheme

$$G_{GUT} \xrightarrow{M_{GUT}} H_1 \xrightarrow{9} M_{infl} 1 \Phi_+ \Phi_- H_2 \longrightarrow G_{SM} \ ; \tag{73}$$

Φ_+, Φ_- is a pair of GUT Higgs superfields in nontrivial complex conjugate representations, which lower the rank of the group by one unit when acquiring nonzero vacuum expectation value. The inflationary phase takes place at the beginning of the symmetry breaking $H_1 \xrightarrow{M_{infl}} H_2$.

F-term inflation is based on the globally supersymmetric renormalisable superpotential

$$W^F_{infl} = \kappa S(\Phi_+ \Phi_- - M^2) \ , \tag{74}$$

where S is a GUT gauge singlet left handed superfield, Φ_+ and Φ_- are defined above; κ and M are two constants (M has dimensions of mass) which can be taken positive with field redefinition. The chiral superfields S, Φ_+, Φ_- are taken to have canonical kinetic terms. This superpotential is the most general one consistent with an R-symmetry under which $W \to e^{i\beta} W$, $\Phi_- \to e^{-i\beta} \Phi_-$, $\Phi_+ \to e^{i\beta} \Phi_+$, and $S \to e^{i\beta} S$. An R-symmetry can ensure that the rest of the renormalisable terms are either absent or irrelevant.

The scalar potential reads

$$V(\phi_+, \phi_-, S) = |F_{\Phi_+}|^2 + |F_{\Phi_-}|^2 + |F_S|^2 + \frac{1}{2} \sum_a g_a^2 D_a^2 \ . \tag{75}$$

The F-term is such that $F_{\Phi_i} \equiv |\partial W / \partial \Phi_i|_{\theta=0}$, where we take the scalar component of the superfields once we differentiate with respect to $\Phi_i = \Phi_+, \Phi_-, S$. The D-terms are

$$D_a = \bar{\phi}_i \, (T_a)^i{}_j \, \phi^j + \xi_a \ , \tag{76}$$

with a the label of the gauge group generators T_a, g_a the gauge coupling, and ξ_a the Fayet-Iliopoulos term. By definition, in the F-term inflation the real constant ξ_a is zero; it can only be nonzero if T_a generates an extra U(1) group. In the context of F-term hybrid inflation, the F-terms give rise to the inflationary potential energy density, while the D-terms are flat along the inflationary trajectory, thus one may neglect them during inflation.

The potential has one valley of local minima, $V = \kappa^2 M^4$, for $S > M$ with $\phi_+ = \phi_- = 0$, and one global supersymmetric minimum, $V = 0$, at $S = 0$ and $\phi_+ = \phi_- = M$. Imposing initially $S \gg M$, the fields quickly settle down the valley of local minima. Since in the slow roll inflationary valley, the ground state of the scalar potential is nonzero, SUSY is broken. In the tree level, along the inflationary valley the potential is constant, therefore perfectly flat. A slope along the potential can be generated by including the one-loop radiative corrections. Thus, the scalar potential gets a little tilt which helps

the inflaton field S to slowly roll down the valley of minima. The one-loop radiative corrections to the scalar potential along the inflationary valley, lead to an effective potential [79–82]

$$V_{\rm eff}^{\rm F}(|S|) = \kappa^2 M^4 \Bigg\{ 1 + \frac{\kappa^2 \mathcal{N}}{32\pi^2} \Bigg[2 \ln \frac{|S|^2 \kappa^2}{\Lambda^2} + \left(\frac{|S|^2}{M^2} + 1 \right)^2 \ln \left(1 + \frac{M^2}{|S|^2} \right) + \left(\frac{|S|^2}{M^2} - 1 \right)^2 \ln \left(1 - \frac{M^2}{|S|^2} \right) \Bigg] \Bigg\} \quad (77)$$

where Λ is a renormalisation scale and $\mathcal{N}$ stands for the dimensionality of the representation to which the complex scalar components ϕ_+, ϕ_- of the chiral superfields Φ_+, Φ_- belong. For example, $\mathcal{N} = \mathbf{27}, \mathbf{126}, \mathbf{351}$, correspond to realistic SSB schemes in SO(10), or E_6 models.

Considering only large angular scales, i.e., taking only the Sachs-Wolfe contribution, one can get the contributions to the CMB temperature anisotropies analytically. The quadrupole anisotropy has one contribution coming from the inflaton field, splitted into scalar and tensor modes, and one contribution coming from the cosmic string network, given by numerical simulations [83]. The inflaton field contribution is

$$\left(\frac{\delta T}{T} \right)_{\rm Q-infl} = \left[\left(\frac{\delta T}{T} \right)^2_{\rm Q-scal} + \left(\frac{\delta T}{T} \right)^2_{\rm Q-tens} \right]^{1/2} , \quad (78)$$

where the quadrupole anisotropy due to the scalar and tensor Sachs-Wolfe effect is

$$\left(\frac{\delta T}{T} \right)_{\rm Q-scal} = \frac{1}{4\sqrt{45}\pi} \frac{V^{3/2}(\varphi_{\rm Q})}{M_{\rm Pl}^3 V'(\varphi_{\rm Q})}$$
$$\left(\frac{\delta T}{T} \right)_{\rm Q-tens} \sim \frac{0.77}{8\pi} \frac{V^{1/2}(\varphi_{\rm Q})}{M_{\rm Pl}^2} , \quad (79)$$

respectively, with $V' \equiv dV(\varphi)/d\varphi$, $M_{\rm Pl}$ the reduced Planck mass, $M_{\rm Pl} = (8\pi G)^{-1/2} \simeq 2.43 \times 10^{18}$GeV, and $\varphi_{\rm Q}$ the value of the inflaton field when the comoving scale corresponds to the quadrupole anisotropy became bigger than the Hubble radius. It can be calculated using Eqs. (77)–(79).

Fixing the number of e-foldings to 60, the inflaton and cosmic string contribution to the CMB, for a given gauge group $G_{\rm GUT}$, depend on the superpotential coupling κ, or equivalently on the symmetry breaking scale M associated with the inflaton mass scale, which coincides with the string mass scale. The relation between κ and M is

$$\frac{M}{M_{\rm Pl}} = \frac{\sqrt{N_{\rm Q} \mathcal{N}}\, \kappa}{2\, \pi\, y_{\rm Q}} , \quad (80)$$

where

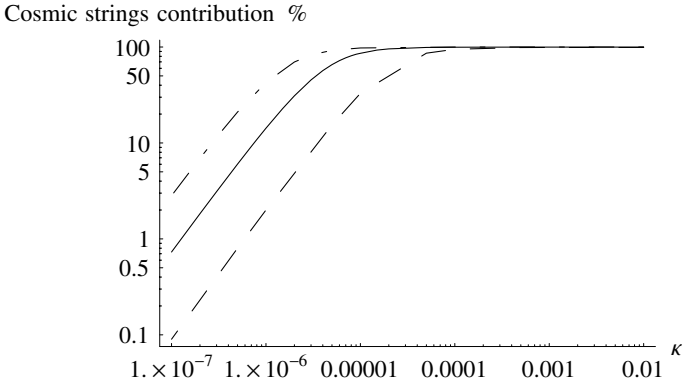

Fig. 4. Contribution of cosmic strings to the quadrupole anisotropy as a function of the superpotential coupling κ. The three curves correspond to $\mathcal{N} = \mathbf{27}$ (*curve with broken line*), $\mathcal{N} = \mathbf{126}$ (*full line*) and $\mathcal{N} = \mathbf{351}$ (*curve with lines and dots*) [82]

$$y_{\rm Q}^2 = \int_1^{\frac{|S_{\rm Q}|^2}{M^2}} dz \big[z \big\{ (z+1)\ln(1+z^{-1}) + (z-1)\ln(1-z^{-1}) \big\} \big] \ . \tag{81}$$

The total quadrupole anisotropy has to be normalised to the COBE data.

A detailed study has been performed in Ref. [82]. It was shown that the cosmic string contribution is consistent with the CMB measurements, provided [82]

$$M \lesssim 2 \times 10^{15} {\rm GeV} \quad \Leftrightarrow \quad \kappa \lesssim 7 \times 10^{-7} \ . \tag{82}$$

In Fig. 4, one can see the contribution of cosmic strings to the quadrupole anisotropy as a function of the superpotential coupling κ [82]. The three curves correspond to $\mathcal{N} = \mathbf{27}$ (curve with broken line), $\mathcal{N} = \mathbf{126}$ (full line) and $\mathcal{N} = \mathbf{351}$ (curve with lines and dots).

The constraint on κ given in Eq. (82) is in agreement with the one found in Ref. [84]. Strictly speaking the above condition was found in the context of SO(10) gauge group, but the conditions imposed in the context of other gauge groups are of the same order of magnitude since M is a slowly varying function of the dimensionality $\mathcal{N}$ of the representations to which the scalar components of the chiral Higgs superfields belong.

The superpotential coupling κ is also subject to the gravitino constraint which imposes an upper limit to the reheating temperature, to avoid gravitino overproduction. The reheating temperature $T_{\rm RH}$ characterises the reheating process via which the Universe enters the high entropy radiation dominated phase at the end of the inflationary era. Within the minimal supersymmetric standard model and assuming a see-saw mechanism to give rise to massive neutrinos, the reheating temperature is [82]

$$T_{\rm RH} \approx \frac{(8\pi)^{1/4}}{7} (\Gamma M_{\rm Pl})^{1/2} \ , \tag{83}$$

where Γ is the decay width of the oscillating inflaton and Higgs fields into right-handed neutrinos

$$\Gamma = \frac{1}{8\pi}\left(\frac{M_i}{M}\right)^2 m_{\rm infl} \ ; \tag{84}$$

with $m_{\rm infl} = \sqrt{2}\kappa M$ the inflaton mass and M_i the right-handed neutrino mass eigenvalue with $M_i < m_{\rm infl}/2$. Equations (80), (83), (84) lead to

$$T_{\rm RH} \sim \frac{1}{12}\left(\frac{60}{N_{\rm Q}}\right)^{1/4}\left(\frac{1}{\mathcal{N}}\right)^{1/4} y_{\rm Q}^{1/2} M_i \ . \tag{85}$$

In order to have successful reheating, it is important not to create too many gravitinos, which imply the following constraint on the reheating temperature [85] $T_{\rm RH} \leq 10^9\,{\rm GeV}$. Since the two heaviest neutrinos are expected to have masses of the order of $M_3 \simeq 10^{15}$ GeV and $M_2 \simeq 2.5 \times 10^{12}$ GeV respectively [86], M_i is identified with $M_1 \sim 6 \times 10^9$ GeV [86]. The gravitino constraint on κ reads [82] $\kappa \lesssim 8 \times 10^{-3}$, which is clearly a weaker constraint than the one imposed from the CMB data.

Concluding, F-term inflation leads generically to cosmic string formation at the end of the inflationary era. The cosmic strings formed are of the GUT scale. This class of models can be compatible with CMB measurements, provided the superpotential coupling is smaller[1] than 10^{-6}. This tuning of the free parameter κ can be softened if one allows for the curvaton mechanism.

According to the curvaton mechanism [88, 89], another scalar field, called the curvaton, could generate the initial density perturbations whereas the inflaton field is only responsible for the dynamics of the Universe. The curvaton is a scalar field, that is sub-dominant during the inflationary era as well as at the beginning of the radiation dominated era which follows the inflationary phase. There is no correlation between the primordial fluctuations of the inflaton and curvaton fields. Clearly, within supersymmetric theories such scalar fields are expected to exist. In addition, embedded strings, if they accompany the formation of cosmic strings, they may offer a natural curvaton candidate, provided the decay product of embedded strings gives rise to a scalar field before the onset of inflation. Considering the curvaton scenario, the coupling κ is only constrained by the gravitino limit. More precisely, assuming the existence of a curvaton field, there is an additional contribution to the temperature anisotropies. The WMAP CMB measurements impose [82] the following limit on the initial value of the curvaton field

$$\psi_{\rm init} \lesssim 5 \times 10^{13}\left(\frac{\kappa}{10^{-2}}\right){\rm GeV} \ , \tag{86}$$

provided the parameter κ is in the range $[10^{-6},\ 1]$ (*see*, Fig. 5).

[1] The linear mass density μ gets a correction due to deviations from the Bogomol'nyi limit, which may enlare [87] the parameter space for F-term inflation. Note that this does not hold for D-term inflation, since then strings are BPS (Bogomol'nyi-Prasad-Sommerfield) states.

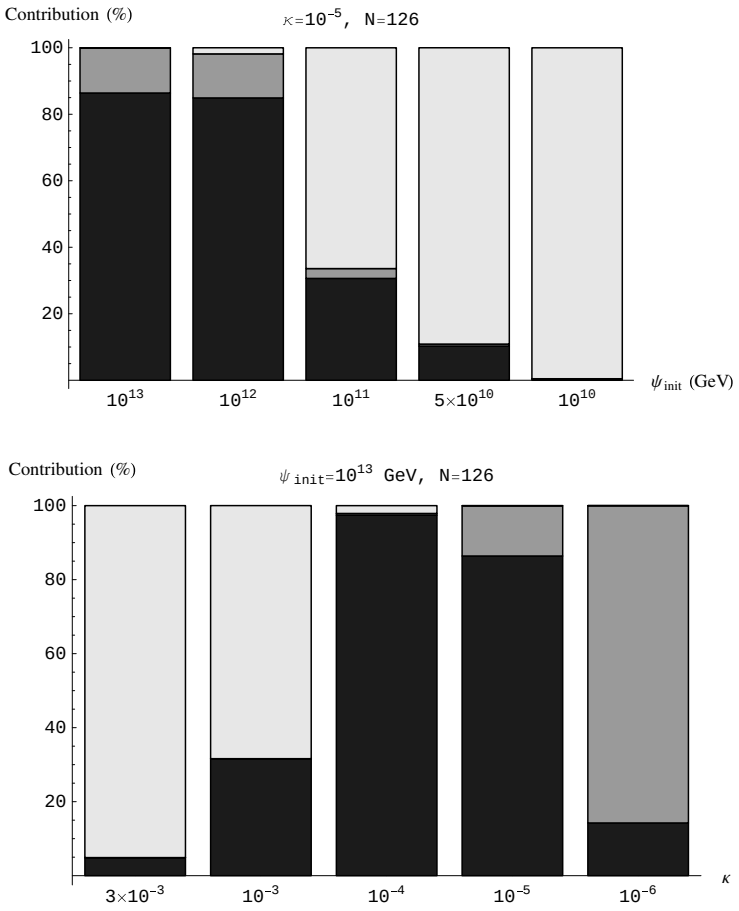

Fig. 5. The cosmic strings (*dark gray*), curvaton (*light gray*) and inflaton (*gray*) contributions to the CMB temperature anisotropies as a function of the the initial value of the curvaton field $\psi_{\rm init}$, and the superpotential coupling κ, for $\mathcal{N} = \mathbf{126}$ [82]

The above results hold also if one includes supergravity corrections. This is expected since the value of the inflaton field is several orders of magnitude below the Planck scale.

D-term Inflation

The early history of the Universe at energies below the Planck scale is described by an effective N = 1 supergravity (SUGRA) theory. Inflation should have taken place at an energy scale $V^{1/4} \lesssim 4 \times 10^{16}$ GeV, implying that inflationary models should be constructed in the framework of SUGRA.

However, it is difficult to implement slow-roll inflation within SUGRA. The positive false vacuum of the inflaton field breaks spontaneously global

supersymmetry, which gets restored after the end of inflation. In supergravity theories, the supersymmetry breaking is transmitted to all fields by gravity, and thus any scalar field, including the inflaton, gets an effective mass of the order of the the expansion rate H during inflation.

This problem, known as the problem of *Hubble-induced mass*, originates from F-term interactions – note that it is absent in the model we have described in the previous subsection – and thus it is resolved if one considers the vacuum energy as being dominated by non-zero D-terms of some superfields [90, 91]. This result led to a dramatic interest in D-term inflation, since in addition, it can be easily implemented within string theory.

D-term inflation is derived from the superpotential

$$W_{\rm infl}^{\rm D} = \lambda S \Phi_+ \Phi_- \ ; \tag{87}$$

S, Φ_-, Φ_+ are three chiral superfields and λ is the superpotential coupling. D-term inflation requires the existence of a nonzero Fayet-Iliopoulos term ξ, which can be added to the Lagrangian only in the presence of an extra U(1) gauge symmetry, under which, the three chiral superfields have charges $Q_S = 0$, $Q_{\Phi_+} = +1$, and $Q_{\Phi_-} = -1$. This extra U(1) gauge symmetry can be of a different origin; hereafter we consider a nonanomalous U(1) gauge symmetry. Thus, D-term inflation requires a scheme, like

$$\mathrm{G_{GUT}} \times \mathrm{U}(1) \xrightarrow{M_{\rm GUT}} \mathrm{H} \times \mathrm{U}(1) \xrightarrow{9} M_{\rm nfl} 1 \Phi_+ \Phi_- \mathrm{H} \rightarrow \mathrm{G_{SM}} \ . \tag{88}$$

The symmetry breaking at the end of the inflationary phase implies that cosmic strings are always formed at the end of D-term hybrid inflation. To avoid cosmic strings, several mechanisms have been proposed which either consider more complicated models or require additional ingredients. For example, one can add a nonrenormalisable term in the potential [92], or add an additional discrete symmetry [93], or consider GUT models based on non-simple groups [94], or introduce a new pair of charged superfields [95] so that cosmic string formation is avoided at the end of D-term inflation. In what follows, we show that standard D-term inflation followed unavoidably by cosmic string production is compatible with CMB data, because the cosmic string contribution to the CMB data is not constant nor dominant. Thus, one does not have to invoke some new physics.

In the global supersymmetric limit, Eqs. (75), (87) lead to the following expression for the scalar potential

$$V^{\rm D}(\phi_+, \phi_-, S) = \lambda^2 \left[|S|^2 (|\phi_+|^2 + |\phi_-|^2) + |\phi_+ \phi_-|^2 \right] + \frac{g^2}{2} (|\phi_+|^2 - |\phi_-|^2 + \xi)^2 \ , \tag{89}$$

where g is the gauge coupling of the U(1) symmetry and ξ is a Fayet-Iliopoulos term, chosen to be positive.

In D-term inflation, as opposed to F-term inflation, the inflaton mass acquires values of the order of Planck mass, and therefore, the correct analysis

must be done in the framework of SUGRA. The SSB of SUSY in the inflationary valley introduces a splitting in the masses of the components of the chiral superfields $\Phi_{\pm}$. As a result, we obtain [96] two scalars with squared masses $m_{\pm}^2 = \lambda^2 |S|^2 \exp\left(|S|^2/M_{\rm Pl}^2\right) \pm g^2 \xi$ and a Dirac fermion with squared mass $m_{\rm f}^2 = \lambda^2 |S|^2 \exp\left(|S|^2/M_{\rm Pl}^2\right)$. Calculating the radiative corrections, the effective scalar potential for minimal supergravity reads [82, 96]

$$V_{\rm eff} = \frac{g^2 \xi^2}{2} \left\{ 1 + \frac{g^2}{16\pi^2} \times \left[2 \ln \frac{|S|^2 \lambda^2}{\Lambda^2} e^{\frac{|S|^2}{M_{\rm Pl}^2}} + \left(\frac{\lambda^2 |S|^2}{g^2 \xi} e^{\frac{|S|^2}{M_{\rm Pl}^2}} + 1 \right)^2 \ln \left(1 + \frac{g^2 \xi}{\lambda^2 |S|^2} e^{-\frac{|S|^2}{M_{\rm Pl}^2}} \right) + \left(\frac{\lambda^2 |S|^2}{g^2 \xi} e^{\frac{|S|^2}{M_{\rm Pl}^2}} - 1 \right)^2 \ln \left(1 - \frac{g^2 \xi}{\lambda^2 |S|^2} e^{-\frac{|S|^2}{M_{\rm Pl}^2}} \right) \right] \right\} \qquad (90)$$

As it was explicitely shown in Refs. [82, 96], D-term inflation can be compatible with current CMB measurements; the cosmic strings contribution to the CMB is model-dependent. The results obtained in Refs. [82, 96] can be summarised as follows: (i) $g \gtrsim 2 \times 10^{-2}$ is incompatible with the allowed cosmic string contribution to the WMAP measurements; (ii) for $g \lesssim 2 \times 10^{-2}$ the constraint on the superpotential coupling λ reads $\lambda \lesssim 3 \times 10^{-5}$; (iii) SUGRA corrections impose in addition a lower limit to λ; (iv) the constraints induced on the couplings by the CMB measurements can be expressed as a single constraint on the Fayet-Iliopoulos term ξ, namely $\sqrt{\xi} \lesssim 2 \times 10^{15}$ GeV. They are shown in Fig. 6.

Assuming the existence of a curvaton field, the fine tuning on the couplings can be avoided provided [82, 96]

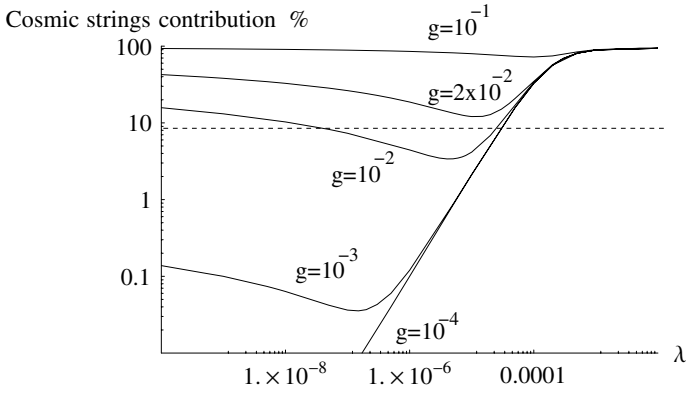

Fig. 6. Cosmic string contribution to the CMB temperature anisotropies as a function of the superpotential coupling λ for different values of the gauge coupling g. The maximal contribution allowed by WMAP is represented by a *dotted line* [82, 96]

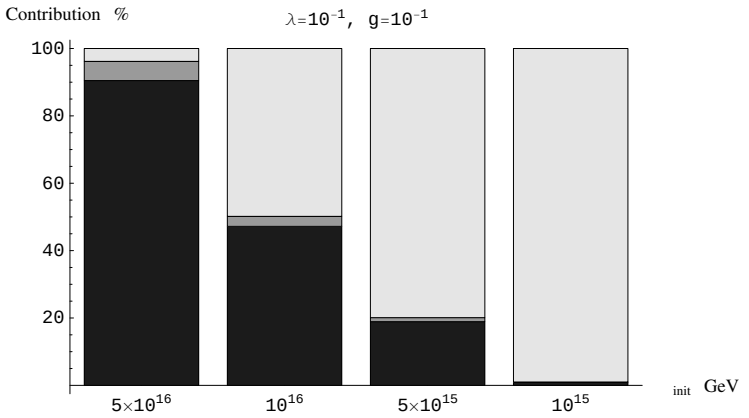

Fig. 7. The cosmic strings (*dark gray*), curvaton (*light gray*) and inflaton (*gray*) contributions to the CMB temperature anisotropies as a function of the initial value of the curvaton field $\psi_{\rm init}$, for $\lambda = 10^{-1}$ and $g = 10^{-1}$ [96]

$$\psi_{\rm init} \lesssim 3 \times 10^{14} \left(\frac{g}{10^{-2}}\right) \text{ GeV} \quad \text{for} \quad \lambda \in [10^{-1}, 10^{-4}] \,. \tag{91}$$

Clearly, for smaller values of λ, the curvaton mechanism is not necessary. We show in Fig. 7 the three contributions as a function of $\psi_{\rm init}$, for $\lambda = 10^{-1}$ and $g = 10^{-1}$. There are values of $\psi_{\rm init}$ which allow bigger values of the superpotential coupling λ and of the gauge coupling g, than the upper bounds obtained in the absence of a curvaton field.

Concluding, standard D-term inflation always leads to cosmic string formation at the end of the inflationary era; these cosmic strings are of the grand unification scale. This class of models is still compatible with CMB measurements, provided the couplings are small enough. As in the case of F-term inflation the fine tuning of the couplings can be softened provided one considers the curvaton mechanism.

The above conclusions are still valid in the revised version of D-term inflation, in the framework of SUGRA with constant Fayet-Iliopoulos terms. In the context of N = 1, 3 + 1 SUGRA, the presence of constant Fayet-Iliopoulos terms shows up in covariant derivatives of all fermions. In addition, since the relevant local U(1) symmetry is a gauged R-symmetry [97], the constant Fayet-Iliopoulos terms also show up in the supersymmetry transformation laws. In Ref. [98] there were presented all corrections of order $g\xi/M_{\rm Pl}^2$ to the classical SUGRA action required by local supersymmetry. Under U(1) gauge transformations in the directions in which there are constant Fayet-Iliopoulos terms ξ, the superpotential must transform as [97]

$$\delta W = -i \frac{g\xi}{M_{\rm Pl}^2} W \,, \tag{92}$$

otherwise the constant Fayet-Iliopoulos term ξ vanishes. This requirement is consistent with the fact that in the gauge theory at $M_{\rm Pl} \to \infty$ the potential is U(1) invariant. To promote the simple SUSY D-term inflation model, Eq. (87), to SUGRA with constant Fayet-Iliopoulos terms, one has to change the charge assignments for the chiral superfields, so that the superpotential transforms under local R-symmetry [98]. In SUSY, the D-term potential is neutral under U(1) symmetry, while in SUGRA the total charge of $\Phi_\pm$ fields does not vanish but is equal to $-\xi/M_{\rm Pl}^2$. More precisely, the D-term contribution to the scalar potential V [see Eq. (89)], should be replaced by $(g^2/2)(q_+|\phi_+|^2+q_-|\phi_-|^2+\xi)^2$ where

$$q_\pm = \pm 1 - \rho_\pm \frac{\xi}{M_{\rm Pl}^2} \quad \text{with} \quad \rho_+ + \rho_- = 1 \; . \tag{93}$$

In addition, the squared masses of the scalar components $\phi_\pm$ become

$$m_\pm^2 = \lambda^2 |S|^2 \exp\left(|S|^2/M_{\rm Pl}^2\right) \pm g^2 \xi q_\pm \; ; \tag{94}$$

the Dirac fermion mass remains unchanged.

For the limits we imposed on the Fayet-Iliopoulos term ξ, the correction $\xi/M_{\rm Pl}^2$ is $\sim 10^{-6}$, implying that the constraints we obtained on g and λ, or equivalently on $\sqrt{\xi}$, as well as the constraint on $\psi_{\rm init}$ still hold in the revised version of D-term inflation within SUGRA [63].

It is important to generalise the above study in the case of nonminimal SUGRA [99], in order to know whether qualitatively the above picture remains valid. A recent study [99] has shown that non-minimal Kähler potential do not avoid the fine tuning, since the cosmic string contribution remains dominant unless the couplings and mass scales are small. For example, taking into account higher order corrections in the Kähler potential, or considering supergravity with shift symmetry, we have obtained [99] that the 9% constraint in the allowed contribution of cosmic strings in the CMB spectrum implies

$$\sqrt{\xi} \leq 2 \times 10^{15} {\rm GeV} \; \Leftrightarrow \; G\mu \leq 8 \times 10^{-7} \; . \tag{95}$$

The *cosmic string problem* can be definitely cured if one considers more complicated models, for example where strings become topologically unstable, namely semi-local strings.

4 Cosmic Superstrings

At first, and for many years, cosmic strings and superstrings were considered as two well separated issues. The main reason for this clear distinction may be considered the Planckian tension of superstrings. If the string mass scale is of the order of the Planck mass, then the four-dimensional F- and D-string self gravity is $G_4\mu_{\rm F} \sim \mathcal{O}(g_{\rm s}^2)$ and $G_4\mu_{\rm D1} \sim \mathcal{O}(g_{\rm s})$ ($g_{\rm s}$ stands for the string coupling), respectively, while current CMB measurements impose an upper

limit on the self gravity of strings of $G\mu < 10^{-6}$. Moreover, heavy superstrings could have only been produced before inflation, and therefore diluted. In addition, Witten showed [100] that, in the context of the heterotic theory, long fundamental BPS strings are unstable, thus they would not survive on cosmic time scales; non-BPS strings were also believed to be unstable.

At present, the picture has been dramatically changed (for a review, see e.g., Ref. [101]). In the framework of braneworld cosmology, our Universe represents a three-dimensional Dirichlet brane (D3-brane) on which open fundamental strings (F-strings) end [102]. Such a D3-brane is embedded in a higher dimensional space, the bulk. Brane interactions can unwind and evaporate higher dimensional branes, leaving behind D3-branes embedded in a higher dimensional bulk; one of these D3-branes could play the rôle of our Universe [103]. Large extra dimensions can be employed to address the hierarchy problem [104], a result which lead to an increasing interest in braneworld scenarios. As it has been argued [105, 106] D-brane-antibrane inflation leads to the production of lower-dimensional D-branes, that are one-dimensional (D-strings) in the noncompact directions. The production of zero- and two-dimensional defects (monopoles and domain walls, respectively) is suppressed. The large compact dimensions and the large warp factors can allow for superstrings of much lower tensions, in the range between $10^{-11} < G\mu < 10^{-6}$. Depending on the model of string theory inflation, one can identify [107] D-strings, F-strings, bound states of p fundamental strings and q D-strings for relatively prime (p, q), or no strings at all.

The probability that two colliding superstrings reconnect can be much less than one. Thus, a reconnection probability $\mathcal{P} < 1$ is one of the distinguishing features of superstrings. D-strings can miss each other in the compact dimension, leading to a smaller $\mathcal{P}$, while for F-strings the scattering has to be calculated quantum mechanically, since these are quantum mechanical objects.

The collisions between all possible pairs of superstrings have been studied in string perturbation theory [108]. For F-strings, the reconnection probability is of the order of $g_{\rm s}^2$. For F-F string collisions, it was found [108] that the reconnection probability $\mathcal{P}$ is $10^{-3} \lesssim \mathcal{P} \lesssim 1$. For D-D string collisions, $10^{-1} \lesssim \mathcal{P} \lesssim 1$. Finally, for F-D string collisions, the reconnection probability can take any value between 0 and 1. These results have been confirmed [109] by a quantum calculation of the reconnection probability for colliding D-strings. Similarly, the string self-intersection probability is reduced. When D- and F-strings meet they can form a three-string junction, with a composite DF-string. In IIB string theory, they may be found bound (p, q) states of p F-strings and q D-strings, where p and q are coprime. This leads to the question of whether there are frozen networks dominating the matter content of the Universe, or whether scaling solutions can be achieved.

The evolution of cosmic superstring networks has been addressed numerically [110–115] and analytically [116].

The first numerical approach [110], studies independent stochastic networks of D- and F-strings, evolving in a flat spacetime. One can either evolve strings in a higher dimensional space keeping the reconnection probability equal to 1, or evolve them in a three-dimensional space with $\mathcal{P} \ll 1$. These two approaches lead to results which are equivalent qualitatively, as it has been shown in Ref. [110]. These numerical simulations have shown that the characteristic length scale ξ, giving the typical distance between the nearest string segments and the typical curvature of strings, grows linearly with time

$$\xi(t) \propto \zeta t \ ; \tag{96}$$

the slope ζ depends on the reconnection probability $\mathcal{P}$, and on the energy of the smallest allowed loops (i.e., the energy cutoff). For reconnection (or intercommuting) probability in the range $10^{-3} \lesssim \mathcal{P} \lesssim 0.3$, it was found [110]

$$\zeta \propto \sqrt{\mathcal{P}} \Rightarrow \xi(t) \propto \sqrt{\mathcal{P}} t \ , \tag{97}$$

in agreement with older results [40].

One can find in the literature statements claiming that $\xi(t)$ should be proportional to $\mathcal{P}t$ instead. If this were correct, then the energy density of cosmic superstrings of a given tension could be considerably higher than that of their field theory analogues (cosmic strings). In Ref. [117] it is claimed that the energy density of longs strings $\rho_{\rm l}$ evolves as $\dot{\rho}_{\rm l} = 2(\dot{a}/a)\rho_{\rm l} - \mathcal{P}(\rho_{\rm l}/\xi)$, where $H = \dot{a}/a$ is the Hubble constant. Substituting the ansatz $\xi(t) = \gamma(t)t$, the authors of Ref. [117] obtain $\dot{\gamma} = -[1/(2t)](\gamma - \mathcal{P})$, during the radiation-dominated era. This equation has a stable fixed point at $\gamma(t) = \mathcal{P}$, implying that [117] $\xi \simeq \mathcal{P}t$. However, Ref. [117] misses out the fact that intersections between two long strings is not the most efficient mechanism for energy loss of the string network. The possible string intersections can be divided into three possible cases (see, Fig. 1): (i) two long strings collide in one point and exchange partners with intercommuting probability $\mathcal{P}_1$; (ii) two strings collide in two points and exchange partners chopping off a small loop with intercommuting probability $\mathcal{P}_1^2$; and (iii) one long string self-intersects in one point and chops off a loop with intercommuting probability $\mathcal{P}_2$, which in general is different than $\mathcal{P}_1$. Only cases (ii) and (iii) lead to a closed loop formation and therefore remove energy from the long string network. Between cases (ii) and (iii), only case (iii) is an efficient way of forming loops and therefore dissipating energy: case (iii) is more frequent than case (ii), and case (ii) has in general a smaller probability, since $\mathcal{P}_1 \sim \mathcal{P}_2$ [110]. However, the heuristic argument employed in Ref. [117] does not refer to self-string intersections (i.e, case (iii)); it only applies to intersections between two long strings, which depend on the string velocity. However self-string intersections should not depend on how fast the string moves, a string can intersect itself even if it does not move but it just oscillates locally.

The findings of Ref. [110] cleared the misconception about the behaviour of the scale ξ, and shown that the cosmic superstring energy density may be higher than the field theory case, but at most only by one order of magnitude.

An important question to be addressed is whether cosmic superstrings can survive for a long time and eventually dominate the energy density of the Universe. This could lead to an overdense Universe with catastrophic cosmological consequences. If the reconnection probability is too low, or equivalently, the strings move in a higher dimensional space and therefore miss each other even if $\mathcal{P}$ is high, then one may fear that the string network does not reach a scaling regime. The string energy density redshifts as $1/a^2$, where a stands for the scale factor. Since for a string network with correlation length ξ, there is about 1 long string per horizon volume, the string energy density is $\sim \mu/a^2$. String interactions leading to loop formation guarantee a scaling regime, in the sense that strings remain a constant fraction of the energy density of the Universe. Loops do not feel the expansion of the Universe, so they are not conformally stretched and they redshift as $1/a^3$. As the loops oscillate, they lose their energy and they eventually collapse. Clearly, scaling is not a trivial issue for cosmic superstrings.

In the first numerical approach [110], where they have been only considered independent stochastic networks of either F- or D-strings, it was shown that each such network reaches a scaling regime. This has been shown by either evolving strings in a higher dimensional space with intercommuting probability equal to 1, or evolving strings in a three-dimensional space with intercommuting probability much smaller than 1.

In a realistic case however, (p, q) strings come in very large number of different types, while a (p, q) string can decay to a loop only if it self-intersects of collide with another (p, q) or $(-p, -q)$ string. A collision between (p, q) and (p', q') strings will lead to a new string $(p \pm p', q \pm q')$, provided the end points of the initial two strings are not attached to other three-string vertices, thus they are not a part of a web. If the collision between two strings can lead to the formation of one new string, on a timescale much shorter than the typical collision timescale, then the creation of a web may be avoided, and the resulting network is composed by strings which are on the average nonintersecting. Then one can imagine the following configuration: A string network, composed by different types of (p, q) strings undergoes collisions and self-intersections. Energy considerations imply the production of lighter daughter strings, leading eventually to one of the following strings: $(\pm 1, 0), (0, \pm 1), \pm(1, 1), \pm(1, -1)$. These ones may then self-intersect, form loops and scale individually. Provided the relative contribution of each of these strings to the energy density of the Universe is small enough, the Universe will not be overclosed.

This result has been confirmed by studying numerically the behavior of a network of interacting Dirichlet-fundamental strings (p, q) in Ref. [112]. To model (p, q) strings arising from compactifications of type IIB string theory, the authors studied [112] the evolution of nonabelian string networks. The positive element of such nonabelian networks is that they contain multiple vertices where many different types of string join together. Such networks have the potential of leading to a string dominated Universe due to tangled networks of interacting (p, q) strings that freeze. It was shown [112] that

such freezing does not take place and the network reaches a scaling limit. In this field theory approach however strings are not allowed to have different tensions, which is a characteristic property of cosmic superstrings. This issue has been addressed later in the context of modelling (p, q) cosmic superstrings [113]. It was found that such networks rapidly approach a stable scaling solution, where once scaling is reached, only a small number of the lowest tension states is populated substantially. An interesting question is to find out whether the field theory approach of Ref. [112] mimics the results of the modelling approach of Ref. [113]. Finally, performing full classical field theory simulations for a model of a string network with junctions, where the junctions can be thought of as global monopoles connected by global strings, it was shown [115] that the evolution is consistent with a late-time scaling regime. Thus, the presence of junctions is not itself inconsistent with scaling.

The cosmic superstring network is characterised [110] by two components: there are a few long strings with a scale-invariant evolution; the characteristic curvature radius of long strings, as well as the typical separation between two long strings are both comparable to the horizon size, $\xi(t) \simeq \sqrt{\mathcal{P}}t$, and there is a large number of small closed loops having sizes $\ll t$. Assuming there are string interactions, the network of long strings will reach an asymptotic energy density, where the energy density in long strings is

$$\rho_{\rm l} = \frac{\mu}{\mathcal{P}t^2} \ . \tag{98}$$

Thus, the fraction of the total density in the form of strings in the radiation-dominated era reads

$$\frac{\rho_{\rm str}}{\rho_{\rm total}} = \frac{32\pi}{3}\frac{G\mu}{\mathcal{P}} \ . \tag{99}$$

Recent numerical investigations [114] of strings evolving in a matter- or radiation-dominated FLRW background claim a weaker power law for the dependence of the scaling string energy density. More precisely, in Ref. [114] it was found that for $\mathcal{P} \gtrsim 0.1$, the function $\rho(1/\mathcal{P})$ is approximately flat, while for $\mathcal{P} \lesssim 0.1$, the function $\rho(1/\mathcal{P})$ is well-fitted by a power-law with exponent $0.6^{+0.15}_{-0.12}$. The behaviour of the string energy density as a function of $\mathcal{P}$ has an important impact for the observational consequences of cosmic superstring networks.

Oscillating string loops loose energy by emitting graviton, dilaton and Ramond-Ramond (RR) fields. Accelerated cosmic strings are sources of gravitational radiation, in particular from the vicinity of the cusps where the string velocity approaches the speed of light. Similarly, cosmic superstrings emit gravity waves but since the intercommutation probability is less than unity, their network is denser with more cusps, resulting in an enhancement of the emitted gravitational radiation. As it was pointed out [118], the gravitational wave bursts emitted from cusps of oscillating string or superstring loops could be detectable with the gravitational-wave interferometers LIGO/VIRGO and LISA.

One can place constraints on the energy scale of cosmic strings from the observational bounds on dilaton decays [119]. Considering that the dilaton lifetime is in the range $10^7\text{s} \lesssim \tau \lesssim t_{\rm dec}$, one can obtain an upper bound $\eta \lesssim \mathcal{P}^{-1/3} \lesssim 10^{11}$ GeV [110] for the energy scale of cosmic superstrings, which determines the critical temperature for the transition leading to string formation. A lower reconnection probability allows a higher energy scale of strings, at most by one order of magnitude.

5 Conclusions

A realistic cosmological scenario necessitates the input of high energy physics, implying that models describing the early stages of the evolution of the Universe have their foundations in both general relativity as well as high energy physics. Comparing the predictions of such models against current astrophysical and cosmological data one concludes to either their acceptance or their rejection, while in the first case one can also fix the free parameters of the models. One of the most beautiful examples in this interplay between cosmology and high energy physics is the case of cosmic strings.

Cosmic strings are expected to be generically formed during the evolution of the Universe, provided the general theoretical picture we have in mind is correct. However, many independent studies concluded in the robust statement that cosmic strings have a limited rôle in the measured CMB temperature anisotropies. Knowing the upper bounds on the contribution of strings to the CMB, one has to examine whether the theoretical models can be adjusted so that there is an agreement between predictions and data. This issue has been addressed in length here. In this respect, cosmology uses high energy physics to build a natural and successful cosmological model, while it offers back some means for testing high energy physics itself.

Cosmic strings are a robust prediction of GUTs, or even M-theory. Even though their rôle in explaining the origin of the observed large-scale structure is sub-dominant, their astrophysical and cosmological implications remain important. Cosmic strings are a small but by no means negligible contribution to any successful cosmological model.

References

1. G. F. Smoot, et al. Astrophys. J. **396**, L1 (1992)
2. A. H. Guth, Phys. Rev. D **23**, 347 (1981)
3. A. D. Linde, Phys. lett. B **108**, 389 (1982)
4. E. Calzetta and M. Sakellariadou, Phys. Rev. D **45**, 2802 (1992); E. Calzetta and M. Sakellariadou, Phys. Rev. D **47**, 3184 (1993)
5. G. W. Gibbons and N. Turok, *The Measure Problem in Cosmology*, [arXiv:hep-th/06090]

6. C. Germani, W. Nelson and M. Sakellariadou, *On the Onset of Inflation in Loop Quantum Cosmology*, [gr-qc/0701172]
7. A. Vilenkin and E. P. S. Shellard, *Cosmic Strings and Other Topological Defects* (Cambridge University Press, Cambridge, England, 2000)
8. J. Martin, A. Riazuelo and M. Sakellariadou, Phys. Rev. D **61**, 083518 (2000)
9. A. Gangui, J. Martin and M. Sakellariadou, Phys. Rev. D **66**, 083502 (2002)
10. T. W. B. Kibble, J. Phys. A **9**, 387 (1976)
11. N. Turok, Phys. Rev. Lett. **63**, 2625 (1989)
12. T. Vachaspati and M. Barriola, Phys. Rev. Lett. **69**, 1867 (1992)
13. W. H. Zurek, Nature **317**, 505 (1985)
14. W. H. Zurek, Phys. Rep. **276**, 178 (1996)
15. P. C.Hendry, Nature **368**, 315 (1994)
16. M. E. Dodd et al., Phys. Rev. Lett. **81**, 3703 (1998)
17. V. M. Ruutu, et al., Nature **382**, 334 (1996)
18. C. Bäuerle, et al., Nature **382**, 332 (1996)
19. I. Chuang, R. Durrer, N. Turok and B. Yurke, Science **251**, 1336 (1991)
20. M. J. Bowick, L. Chandar, E. A. Schiff and A. M. Srivastava, Science **263**, 943 (1994)
21. S. Digal, R. Ray and A. M. Srivastava, Phys. Rev. Lett. **83**, 5030 (1999)
22. P. Laguna and W. H. Zurek, Phys. Rev. Lett. **78**, 2519 (1997)
23. A. Yates and W. H. Zurek, Phys. Rev. Lett. **80**, 5477 (1998)
24. N. D. Antunes, L. M. A. Bettencourt and W. H. Zurek, Phys. Rev. Lett. **82**, 282 (1999)
25. G. J. Stephens, E. A. Calzetta, B. L. Hu and S. A. Ramsey, Phys. Rev. D **59**, 045009 (1999)
26. S. K. Blau, E. I. Guendelman and A. H. Guth, Phys. Rev. D **35**, 1747 (1987)
27. R. Gregory, Phys. Rev. D **43**, 520 (1991); R. Gregory, D. Haws and D. Garfinkle, *ibid.* **42**, 343 (1990)
28. P. S. Letelier, Phys. Rev. D **41**, 1333 (1990)
29. B. Barrabès, B. Boisseau and M. Sakellariadou, Phys. Rev. D **49**, 2734 (1994)
30. A. M. Polyakov, Phys. Lett. B **103** (1981)207
31. A. Vilenkin, Phys. Rev. D **43** (1991) 1061
32. J. Garriga and M. Sakellariadou, Phys. Rev. D **48** (1993) 2502.
33. E. P. S. Shellard, Nucl. Phys. B **283**, 264 (1988)
34. K. Moriarty, E. Myers and C. Rebbi, Phys. Lett. B **207**, 411 (1988)
35. P. Laguna and R. Matzner, Phys. Rev. D **41**, 1751 (1990)
36. M. Sakellariadou, JCAP **0504**, 003 (2005)
37. T. W. B. Kibble, Nucl. Phys. B **252**, 277 (1985)
38. A. Albrecht and N. Turok, Phys. Rev. Lett. **54**, 1868 (1985); A. Albrecht and N. Turok, Phys. Rev. D **40**, 973 (989)
39. D. P. Bennett, in *Formation and Evolution of Cosmic Strings*, edited by G. Gibbons, S. Hawking and T. Vachaspati (Cambridge University Press, Cambridge, England, 1990); F. R. Bouchet, *ibid*; E. P. S. Shellard and B. Allen, *ibid*
40. M. Sakellariadou and A. Vilenkin, Phys. Rev. D **42**, 349 (1990)
41. D. P. Bennett and F. R. Bouchet, Phys. Rev. Lett. **60**, 257 (1988)
42. D. Austin, E. J. Copeland and T. W. B. Kiblle, Phys. Rev. D **48**, 5594 (1993)
43. M. Sakellariadou, Phys. Rev. D **42**, 354 (1990)
44. M. Hindmarsh, Phys. Lett. B **251**, 28 (1990)
45. G. R. Vincent, M. Hindmarsh and M. Sakellariadou, Phys. Rev. D **56**, 637 (1997)

46. G. Vincent, N. D. Antunes and M. Hindmarsh, Phys. Rev. Lett. **80**, 2277 91998)
47. P. P. Avelino, E. P. S. Shellard, J. H. P. Wu and B. Allen, Phys. Rev. D **60**, 023511 (1999)
48. J. N. Moore, E. P. S. Shellard and C. J. A. P. Martins, Phys. Rev. D **65**, 023503 (2002)
49. C. Ringeval, M. Sakellariadou and F. R. Bouchet, *Cosmological evolution of cosmic string loops*, [arXiv:astro-ph/0511646]
50. V. Vanchurin, K. D. Olum and A. Vilenkin, Phys. Rev. D **74**, 063527 (2006)
51. C. J. A. P. Martins and E. P. S. Shellard, Phys. Rev. D **73**, 043515 (2006)
52. R. Hagedorn, Nuovo Cimento Suppl. **3**, 147 (1965)
53. S. Frautschi, Phys. Rev. D **3**, 2821 (1971)
54. R. D. Carlitz, Phys. Rev. D **5**, 3231 (1972)
55. D. Mitchell and N. Turok, Phys. Rev. Lett. **58**, 1577 (1987)
56. M. Sakellariadou and A. Vilenkin, Phys. Rev. D **37**, 885 (1988)
57. M. Sakellariadou, Nucl. Phys. B **468**, 319 (1996)
58. Y. Fukuda, et. al., [Super-Kamiokande Collaboration], Phys. Rev. Lett. **81**, 1562 (1998)
59. Q. R. Ahmad, *et al.* [SNO Collaboration], Phys. Rev. Lett. **87**, 071301 (2001)
60. K. Eguchi, *et al.* [KamLAND Collaboration], Phys. Rev. Lett. **90**, 021802 (2003)
61. R. Jeannerot, J. Rocher and M. Sakellariadou, Phys. Rev. D **68**, 103514 (2003)
62. P. Horava and E. Witten, Nucl. Phys. B **460** (1996) 506
63. M. Sakellariadou, Annalen Phys. **15**, 264 (2006)
64. M. Hindmarsh, Nucl. Phys. Proc. Suppl. **43**, 50 (1995)
65. A. Gangui, Phys. Rev D **50**, 3684 (1994); A. Gangui, F. Lucchin, S. Matarrese and S. Mollerach, Astrophys. J. **430**, 447 (1994); A. Matacz, Phys. Rev. D **55**, 1860 (1997)
66. A. Linde and V. Mukhanov, Phys. Rev. D **56**, 535 (1997)
67. R. Durrer, A. Gangui and M. Sakellariadou, Phys. Rev. Lett. **76**, 579 (1996)
68. R. Durrer, M. Kunz and A. Melchiorri, Phys. Rev. D **59**, 123005 (1999)
69. N. Turok, U.-L. Pen and U. Seljak, Phys. Rev. D **58**, 023506 (1998)
70. U.-L. Pen, U. Seljak and N. Turok, Phys. Rev. Lett. **79**, 1611 (1997)
71. N. Bevis, M. Hindmarsh, M. Kunz and J Urrestilla, *CMB power spectrum contribution from cosmic strings using field-evolution simulations of the Abelian Higgs model*, [arXiv: astro-ph/0605018]
72. A. T. Lee, Astrophys. J. **561**, L1 (2001); R. Stompor, Astrophys. J. **561**, L7 (2001)
73. C. B. Netterfield, et. al., Astrophys. J. **571**, 604 (2002); P. Be Bernardis, et. al., Astrophys. J. **564**, 559 (2002)
74. N. W. Halverson, et. al., Astrophys. J. **568**, 38 (2002); C. Pryke, et. al. , Astrophys. J. **568**, 46 (2002)
75. C. L. Bennett, et al., Astroph. J. Suppl. **148**, 1 (2003)
76. F. R. Bouchet, P. Peter, A. Riazuelo and M. Sakellariadou, Phys. Rev. D **65**, 021301 (2002)
77. L. Pogosian, M. Wyman and I. Wasserman, J. of Cosm. and Astrop. Phys. **09**, 008 (2004)
78. M. Wyman, L. Pogosian and I. Wasserman, Phys. Rev. D **72** (2005) 023513
79. G. Dvali, Q. Shafi and R. Schaefer, Phys. Rev. Lett. **73**, 1886 (1994)

80. G. Lazarides, *Inflationary cosmology*, [arXiv:hep-ph/0111328]
81. V. N. Senoguz and Q. Shafi, Phys. Lett. B **567**, 79 (2003)
82. J. Rocher and M. Sakellariadou, JCAP **0503**, 004 (2005)
83. M. Landriau and E. P. S. Shellard, Phys. Rev. D **69**, 23003 (2004)
84. R. Kallosh and A. Linde, JCAP **0310**, 008 (2003)
85. M. Yu Khlopov, A. Linde, Phys. Lett. B **138**, 265 (1984)
86. J. C. Pati, Int. J. Mod. Phys. A **18**, 4135 (2003)
87. R. Jeannerot and M. Postma, JCAP **0607**, 012 (2006)
88. D. H. Lyth and D. Wands, Phys. Lett. B **524**, 5 (2002)
89. T. Moroi and T. Takahashi, Phys. Lett. B **522**, 215 (2001), Erratum-ibid. B **539**, 303 (2002)
90. E. Halyo, Phys. Lett. B **387**, 43 (1996)
91. P. Binetruy and D. Dvali, Phys. Lett. B **388**, 241 (1996)
92. R. Jeannerot, S. Khalil, G. Lazarides and Q. Shafi, JHEP **0010**,012 (2000)
93. G. Lazarides and C. Panagiotakopoulos, Phys. Rev. D **52**, 559 (1995)
94. T. Watari and T. Yanagida, Phys. Lett. B **589**, 71 (2004)
95. J. Urrestilla, A. Achúcarro and A. C. Davis, Phys. Rev. Lett. **92**, 251302 (2004)
96. J. Rocher and M. Sakellariadou, Phys. Rev. Lett. **94**, 011303 (2005)
97. A. Van Proeyen, Fortsch. Phys. **53**, 997 (2005)
98. P. Binetruy, G. Dvali, R. Kallosh and A. Van Proeyen, Class. Quant. Grav. **21**, 3137 (2004)
99. J. Rocher and M. Sakellariadou, JCAP **0611**, 001 (2006)
100. E. Witten, Phys. Lett. **B153**, 243 (1985)
101. M. Majumdar, *A tutorial on liknks between cosmic string tehory and superstring theory* [arXiv: hep-th/0512062]
102. J. Polchinski, *String theory. Vol. II: Superstring theory and beyond*, Cambridge University Press (1998)
103. R. Durrer, M. Kunz and M. Sakellariadou , Phys. Lett. B **614**, 125 (2005)
104. N. Arkani-Hamed, S. Dimopoulos and G. Dvali, Phys. Lett. B **429**, 263 (1998)
105. N. T. Jones, H. Stoica and S.-H. H. Tye, JHEP **0207**, 051 (2002)
106. H. Stoica and S.-H. H. Tye, Phys. Lett. B **536**, 185 (2003)
107. E. J. Copeland, R. C. Myers and J. Polchinski, JHEP **0406**, 013 (2004)
108. M. G. Jackson, N. T. Jones and J. Polchinski, JHEP **0510**, 013 (2005).
109. A. Hanany and K. Hashimoto K, JHEP **0506**, 021 (2005)
110. M. Sakellariadou, JCAP **0504**, 003 (2005)
111. A. Avgoustidis and E. P. S. Shellard, Phys. Rev. D **71**, 123513 (2005)
112. E. Copeland and P. Shaffin, JHEP **0511**, 023 (2005)
113. S.-H. H. Tye, I. Wasserman and M. Wyman, Phys. Rev. D **71**, 103508 (2005); Erratum-ibid. D **71**, 129906 (2005)
114. A. Avgoustidis and E. P. S. Shellard, Phys. Rev. D **73**, 041301 (2006)
115. M. Hindmarsh and P. M. Saffin, JHEP **0608**, 066 (2006)
116. E. J. Copeland, T. W. B. Kibble and D. A. Steer, Phys. Rev. Lett. **97**, 021602 (2006)
117. N. T. Jones, H. Stoica and S.-H. H. Tye, Phys. Lett. B **563**, 6 (2003)
118. T. Damour and A. Vilenkin, Phys. Rev. D **71**, 063510 (2005)
119. T. Damour and A. Vilenkin, Phys. Rev. Lett. **78**, 2288 (1997)

Index

Lecture Notes in Physics

For information about earlier volumes
please contact your bookseller or Springer
LNP Online archive: springerlink.com

Vol.672: R. Kh. Zeytounian, Topics in Hyposonic Flow Theory

Vol.673: C. Bona, C. Palenzula-Luque, Elements of Numerical Relativity

Vol.674: A. G. Hunt, Percolation Theory for Flow in Porous Media

Vol.675: M. Kröger, Models for Polymeric and Anisotropic Liquids

Vol.676: I. Galanakis, P. H. Dederichs (Eds.), Half-metallic Alloys

Vol.677: A. Loiseau, P. Launois, P. Petit, S. Roche, J.-P. Salvetat (Eds.), Understanding Carbon Nanotubes

Vol.678: M. Donath, W. Nolting (Eds.), Local-Moment Ferromagnets

Vol.679: A. Das, B. K. Chakrabarti (Eds.), Quantum Annealing and Related Optimization Methods

Vol.680: G. Cuniberti, G. Fagas, K. Richter (Eds.), Introducing Molecular Electronics

Vol.681: A. Llor, Statistical Hydrodynamic Models for Developed Mixing Instability Flows

Vol.682: J. Souchay (Ed.), Dynamics of Extended Celestial Bodies and Rings

Vol.683: R. Dvorak, F. Freistetter, J. Kurths (Eds.), Chaos and Stability in Planetary Systems

Vol.684: J. Dolinšek, M. Vilfan, S. Žumer (Eds.), Novel NMR and EPR Techniques

Vol.685: C. Klein, O. Richter, Ernst Equation and Riemann Surfaces

Vol.686: A. D. Yaghjian, Relativistic Dynamics of a Charged Sphere

Vol.687: J. W. LaBelle, R. A. Treumann (Eds.), Geospace Electromagnetic Waves and Radiation

Vol.688: M. C. Miguel, J. M. Rubi (Eds.), Jamming, Yielding, and Irreversible Deformation in Condensed Matter

Vol.689: W. Pötz, J. Fabian, U. Hohenester (Eds.), Quantum Coherence

Vol.690: J. Asch, A. Joye (Eds.), Mathematical Physics of Quantum Mechanics

Vol.691: S. S. Abdullaev, Construction of Mappings for Hamiltonian Systems and Their Applications

Vol.692: J. Frauendiener, D. J. W. Giulini, V. Perlick (Eds.), Analytical and Numerical Approaches to Mathematical Relativity

Vol.693: D. Alloin, R. Johnson, P. Lira (Eds.), Physics of Active Galactic Nuclei at all Scales

Vol.694: H. Schwoerer, J. Magill, B. Beleites (Eds.), Lasers and Nuclei

Vol.695: J. Dereziński, H. Siedentop (Eds.), Large Coulomb Systems

Vol.696: K.-S. Choi, J. E. Kim, Quarks and Leptons From Orbifolded Superstring

Vol.697: E. Beaurepaire, H. Bulou, F. Scheurer, J.-P. Kappler (Eds.), Magnetism: A Synchrotron Radiation Approach

Vol.698: S. Bellucci (Ed.), Supersymmetric Mechanics – Vol. 1

Vol.699: J.-P. Rozelot (Ed.), Solar and Heliospheric Origins of Space Weather Phenomena

Vol.700: J. Al-Khalili, E. Roeckl (Eds.), The Euroschool Lectures on Physics with Exotic Beams, Vol. II

Vol.701: S. Bellucci, S. Ferrara, A. Marrani, Supersymmetric Mechanics – Vol. 2

Vol.702: J. Ehlers, C. Lämmerzahl, Special Relativity

Vol.703: M. Ferrario, G. Ciccotti, K. Binder (Eds.), Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1

Vol.704: M. Ferrario, G. Ciccotti, K. Binder (Eds.), Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 2

Vol.705: P. Bhattacharyya, B.K. Chakrabarti (Eds.), Modelling Critical and Catastrophic Phenomena in Geoscience

Vol.706: M.A.L. Marques, C.A. Ullrich, F. Nogueira, A. Rubio, K. Burke, E.K.U. Gross (Eds.), Time-Dependent Density Functional Theory

Vol.707: A.V. Shchepetilov, Calculus and Mechanics on Two-Point Homogenous Riemannian Spaces

Vol.708: F. Iachello, Lie Algebras and Applications

Vol.709: H.-J. Borchers and R.N. Sen, Mathematical Implications of Einstein-Weyl Causality

Vol.710: K. Hutter, A.A.F. van de Ven, A. Ursescu, Electromagnetic Field Matter Interactions in Thermoelastic Solids and Viscous Fluids

Vol.711: H. Linke, A. Månsson (Eds.), Controlled Nanoscale Motion

Vol.712: W. Pötz, J. Fabian, U. Hohenester (Eds.), Modern Aspects of Spin Physics

Vol.713: L. Diósi, A Short Course in Quantum Information Theory

Vol.714: Günter Reiter and Gert R. Strobl (Eds.), Progress in Understanding of Polymer Crystallization

Vol.715: Stefan Hüfner (Ed.), Very High Resolution Photoelectron Spectroscopy

Vol.716: M. Henkel, R. Sanctuary, M. Pleimling (Eds.), Ageing and the Glass Transition

Vol.717: R. Alicki, K. Lendi, Quantum Dynamical Semigroups and Applications

Vol.718: W.G. Unruh, R. Schützhold (Eds.), Quantum Analogues: From Phase Transitions to Black Holes and Cosmology