Gravity as an emergent phenomenon: fundamentals and applications

Raúl Carballo Rubio

Universidad de Granada
Programa de Doctorado en Física y Matemáticas
Escuela de Doctorado de Ciencias, Tecnologías e Ingenierías
Granada, 2016

Editor: Universidad de Granada. Tesis Doctorales
Autor: Raúl Carballo Rubio
ISBN: 978-84-9125-879-7
URI: http://hdl.handle.net/10481/43711

Gravity as an emergent phenomenon: fundamentals and applications

Raúl Carballo Rubio[1]

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Under the supervision of:
Carlos Barceló Serón[1]
Luis Javier Garay Elizondo[2,3]

[1] Instituto de Astrofísica de Andalucía (IAA-CSIC),
Glorieta de la Astronomía, 18008 Granada, Spain
[2] Departamento de Física Teórica II,
Universidad Complutense de Madrid, 28040 Madrid, Spain
[3] Instituto de Estructura de la Materia (IEM-CSIC),
Serrano 121, 28006 Madrid, Spain

Consejo Superior de Investigaciones Científicas

Compromiso de respeto de derechos de autor

El doctorando Raúl Carballo Rubio y los directores de la tesis Carlos Barceló Serón y Luis Javier Garay Elizondo garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

En Granada, a 22 de diciembre de 2015,

Directores de la tesis

Fdo: Carlos Barceló Serón

Fdo: Luis Javier Garay Elizondo

Doctorando

Fdo: Raúl Carballo Rubio

Estructuración de contenidos en la memoria

Los contenidos exigidos en una tesis doctoral en la Universidad de Granada para los programas de doctorado regulados por el RD99/2011 se encuentran en la presente memoria estructurados en las siguientes secciones:

Título
Portada

Compromiso de respeto de derechos de autor
Compromiso de respeto de derechos de autor

Resumen
Resumen

Introducción
Introduction

Objetivos
Introduction, Secs. 1.1, 2.1, 3.1 and 4.1

Metodología
Secs. 1.2, 1.3, 2.2, 4.2, 4.3 and 4.4

Resultados
Secs. 1.4, 1.5, 1.6, 2.3, 2.4, 2.5, 3.2, 3.3, 4.5 and 4.6

Conclusiones
Secs. 1.7, 2.6, 3.4 and 4.7; Main conclusions and future directions

Bibliografía
Bibliography

Resumen

En esta tesis se ha realizado un estudio de distintos aspectos de la aproximación a la construcción de una teoría de gravedad cuántica conocida como gravedad emergente, con el objetivo de analizar preguntas fundamentales en el marco de este programa de investigación, así como posibles aplicaciones a problemas actuales de la física teórica gravitacional. El objetivo principal de este programa de investigación es la identificación de mecanismos que den lugar de manera robusta a las propiedades observadas en la física de bajas energías conocida, descrita mediante la relatividad general y el modelo estándar de física de partículas. Las técnicas utilizadas comprenden desde física de la materia condensada y sistemas de muchas partículas no relativistas, hasta teoría clásica y cuántica de campos (en espaciotiempos planos y curvos), teorías de campos efectivas, geometría espaciotemporal y relatividad general.

En la primera parte de la tesis, dividida en dos capítulos, se analiza la emergencia de las propiedades emergentes más relevantes para la física de bajas energías: la simetría de Lorentz y simetrías internas. En el primer capítulo se expone en detalle la construcción de un modelo en el que se analiza la emergencia de la electrodinámica cuántica en una clase de universalidad de sistemas de muchas partículas (fermiónicas de espín 1/2) no relativistas. La razón de considerar la interacción electromagnética en lugar de la gravitatoria se debe a la mayor simplicidad de la primera, lo cual hace natural su estudio como paso previo a la comprensión de situaciones más complejas. La elección de modelo no relativista inicial se basa en la aparición de puntos de Fermi en su espectro al producirse una transición a una fase superfluida. Se hace un uso extenso de técnicas de aproximación que han sido desarrolladas en paralelo al estudio experimental de las fases superfluidas del helio-3, en particular aquellas que presentan puntos de Fermi. En el segundo capítulo se estudia la propiedad más importante que diferencia la interacción gravitatoria respecto de la electromagnética: su carácter no lineal. Se revisa la determinación de los vértices de interacción característicos de la relatividad general mediante el denominado problema de autointeracción de gravitones. Por una parte, el interés de reexaminar este conocido problema aparece debido a recientes publicaciones que cuestionan la solución estándar al mismo. Por otra parte, su comprensión en detalle es de fundamental importancia para el programa de gravedad emergente. Las conclusiones obtenidas mediante un desarrollo minucioso del problema son comparadas con la extensa literatura al respecto.

La segunda parte de la tesis versa sobre dos aplicaciones de resultados específicos de la primera parte, a cada una de las cuales se le dedica un capítulo completo. La primera de las aplicaciones concierne al conocido problema de la constance cosmológica. En el capítulo correspondiente se explica cómo la emergencia de estructuras relativistas y de la propia interacción gravitatoria permitiría evadir el problema. Se demuestra explícitamente la existencia de una teoría de gravedad en la que la constante cosmológica puede entenderse como cualquier otra constante fundamental en física, sin incurrir en una tensión entre los principios de la relatividad general y las teorías de campos efectivas. Por su parte, la segunda aplicación atañe a la física de los agujeros negros. En el marco de gravedad emergente, se estudia la formulación geométrica de una nueva propuesta que podría evitar los problemas de pérdida de información y de formación de singularidades en agujeros negros. En términos cualitativos, las geometrías construidas describen la transición entre una geometría de agujero negro y una geometría de agujero blanco en escalas de tiempo cortas, respecto a la escala de tiempo típica de un agujero negro en evaporación. La discusión detallada de la motivación y las propiedades geométricas de esta propuesta es seguida de una exploración de sus posibles implicaciones experimentales.

List of publications

The research activities reported in this thesis have resulted in the following articles and conference proceedings:

Regular articles

12. Carlos Barceló, Raúl Carballo-Rubio and Luis J. Garay
Black holes turn white fast, otherwise stay black: no half measures
arXiv: [gr-qc/1511.00633]

11. Carlos Barceló, Raúl Carballo-Rubio and Luis J. Garay
Where does the physics of extreme gravitational collapse reside?
Invited submission to the special issue "Open Questions in Black Hole Physics" of Universe. [gr-qc/1510.04957]

10. Carlos Barceló, Raúl Carballo-Rubio and Luis J. Garay
Uncovering the effective spacetime (Lessons from the effective field theory rationale)
Honorable Mention in the 2015 GRF Awards for Essays on Gravitation.
Int. J. Mod. Phys. **D24** (2015) 1544019 [hep-th/1505.05315]

9. Raúl Carballo-Rubio
Longitudinal diffeomorphisms obstruct the protection of vacuum energy
Phys. Rev. **D91** (2015) 124071 [gr-qc/1502.05278]

8. Carlos Barceló, Raúl Carballo-Rubio, Luis J. Garay and Gil Jannes
The lifetime problem of evaporating black holes: mutiny or resignation
Class. Quant. Grav. **32** (2015) 035012 [gr-qc/1409.1501]

7. Daniel N. Blaschke, Raúl Carballo-Rubio and Emil Mottola
Fermion pairing and the scalar boson of the 2D conformal anomaly
JHEP **1412** (2014) 153 [hep-th/1407.8523]

6. Carlos Barceló, Raúl Carballo-Rubio, Luis J. Garay and Gil Jannes
Electromagnetism as an emergent phenomenon: a step by step guide
New J. Phys. **16** (2014) 123028 [gr-qc/1407.6532]

5. Carlos Barceló, Raúl Carballo-Rubio and Luis J. Garay
Muting at the white-hole district
Honorable Mention in the 2014 GRF Awards for Essays on Gravitation.
Int. J. Mod. Phys. **D23** (2014) 1442022 [gr-qc/1407.1391]

4. Carlos Barceló, Raúl Carballo-Rubio and Luis J. Garay
Absence of cosmological constant problem in special relativistic field theory of gravity
arXiv: [gr-qc/1401.2941]

3. Carlos Barceló, Raúl Carballo-Rubio and Luis J. Garay
Unimodular gravity and general relativity from graviton self-interactions
Phys. Rev. **D89** (2014) 124019 [gr-qc/1401.2941]

2. Carlos Barceló, Raúl Carballo-Rubio, Luis J. Garay and Ricardo Gómez-Escalante
Hybrid classical-quantum dynamics asks for hybrid notions
Phys. Rev. **A86** (2012) 042120 [quant-ph/1206.7036]

1. Carlos Barceló, Raúl Carballo-Rubio and Luis J. Garay
Two formalisms, one renormalized stress-energy tensor
Phys. Rev. **D85** (2012) 084001 [gr-qc/1112.0489]

Contributions to conferences with proceedings (speaker first)

4. Luis J. Garay, Carlos Barceló, Raúl Carballo-Rubio and Gil Jannes
Do stars die too long?
To appear in the MG14 proceedings volumes (published by World Scientific).

3. Raúl Carballo-Rubio, Carlos Barceló and Luis J. Garay
Some not-so-common ideas about gravity
J. Phys.: Conf. Ser. **626** (2015) 012010

2. Carlos Barceló, Raúl Carballo-Rubio, Luis J. Garay and Gil Jannes
Do transient white holes have a place in nature?
J. Phys.: Conf. Ser. **600** (2015) 012033

1. Raúl Carballo-Rubio, Carlos Barceló and Luis J. Garay
Absence of cosmological constant problem in special relativistic field theory of gravity: one-loop renormalization group
J. Phys.: Conf. Ser. **600** (2015) 012032

Contents

It is my opinion that there is no worthwhile physics
(indeed perhaps no physics at all) in this paper.

Such a counterintuitive way of constructing a theory
can only be called schizophrenic.

This is an enormously long, and on the whole,
unnecessary embroidery on a fully settled subject.

Anonymous referees

Introduction

'The last thing one settles in writing a book,'
Pascal observes, 'is what one should put in first.'
So, having written, collected and arranged these strange tales,
having selected a title and two epigraphs,
I must now examine what I have done - and why.

Oliver Sacks

The evolution of personal knowledge does not always proceed in a strictly linear fashion. Sometimes, reaching a satisfactory understanding requires to reevaluate again and again a given problem, frequently changing your mind about its relevance for the overall picture or even the most important conclusions that should be pulled out from it. This is certainly my experience during the last years. Therefore, writing down a doctoral dissertation is a challenging opportunity to make an overall balance and evaluate the net advance to be reported; it is a time to perform an introspective analysis of the personal view on the subject which has been the field of study during all these years.

The area of knowledge in which this thesis is framed is that of quantum gravity. Under this name, a large number of research programs with different motivations, techniques, and goals are clustered together. As with many other denominations, these words do not even make justice to some of the approaches that find shelter under this linguistic umbrella. In this text, we shall understand that any research program that intends to go further than the theory of general relativity falls within this category, irrespectively of its particular nature.

That it is necessary to find a coherent description of nature that contains, but also expands the regime of applicability of the theory of general relativity, has been acknowledged for decades. It is worth remarking that this is not a practical, or empirical requirement, but rather it arises from formal or conceptual considerations. There is no known experiment, the explanation of which requires a theory of quantum gravity. But the theoretical skeleton of general relativity does not succeed in hiding its imperfections, thus motivating an aesthetic search for a more complete and satisfactory framework. Of course, the ultimate desire is that the construction of a theory of quantum gravity would uncover genuine phenomenological implications, that could eventually be corroborated in future experiments.

The imperfections of general relativity are known as singularities, and are frequently depicted in a qualitative way as regions of spacetime in which the known laws of physics break down. It was Albert Einstein who recognized that the dynamical evolution of the gravitational interaction should be described in terms of the evolution of spacetime. The remarkable fact is that, in the framework of general relativity, initial conditions that are reasonable from a physical perspective could lead generically to the occurrence of singularities. This is for instance the case of the gravitational collapse of massive stars to black holes, a phenomenon that is believed to be frequent in our universe.

The missing piece that causes the incompleteness of general relativity is broadly (but perhaps prematurely [1]) identified with quantum mechanics; an identification that explains the naming of this research area. The general belief is that a successful marriage between the principles of these theories would be enough to overcome, or regularize, the imperfections of general relativity. Nevertheless, merging together quantum mechanics and general relativity encounters a number of difficulties, most of them arguably emanating from their different empathies towards the presence of background structures. Whereas quantum mechanics is easily implementable when a background structure exists, general relativity demands the absence of a background structure that is fixed a priory. Given this situation, one can opt for trying to adapt quantum mechanics so as to elevate background independence, or in other words a geometrical viewpoint, to a fundamental principle. One of the most prominent examples of this approach is the loop quantum gravity program [2, 3, 4].

However, the situation has also led some researchers to ask themselves whether Einstein's theory could be just an emergent theory [5, 6, 7]. From this perspective one does not have to strictly quantize general relativity, but to search for an underlying structure, containing in principle no geometric notions whatsoever, such that general relativity can emerge at a coarse-grained level. In this work we will use the word emergent in this sense: we shall consider the string theory approach to quantum gravity as emergent, but approaches such as causal dynamical triangulations, causal sets or loop quantum gravity, that retain to a greater or lesser extend geometric notions, as non-emergent. Leaving aside the widely developed string theory approach, there exist some much less explored emergent-gravity approaches based on condensed-matter-like systems [8, 9]. Contrarily to string theory [10, 11], these latter approaches keep no relativistic trace at a deeper level, as even special relativity is emergent. Most importantly, the attitude with respect to fundamentality in these different approaches is completely different: while string theory prioritizes a very special fundamental structure of nature, the research based in condensed-matter-like models focuses on the study of universality classes that share the same low-energy physics.

The possibility that familiar non-relativistic phenomena, such as the propagation of acoustic disturbances in inviscid and barotropic fluids with irrotational flow [12, 13], could serve as mimickers of relativistic wave equations in curved spacetimes, has sparked a great deal of interest [14]. Indeed, in the last 10 years it has become clear that the appearance of metric structures controlling the propagation of effective fields within condensed matter systems, mostly in low-energy regimes, is quite simple and ubiquitous [15]. On the other hand, totally independent arguments shaped with the help of the formalism of quantum field theory suggest that Lorentz invariance should be highly fragile when interactions between these effective fields are included, so that even tiny violations at high energies lead to unacceptably large effects at low energies [16, 17]. These disparate results complicate the distillation of a coherent picture.

When one leaves the purely kinematic stage and asks about the effective dynamics of the would-be gravitational degrees of freedom, the situation is even more obscure. In general, the effective metric structures do not follow Einstein's equations, and it turns out to be very difficult to force them to, even at a theoretical level [15, 18]. It is not even clear what the fundamental origin of this difficulty is. In the context of an emergent dynamics *à la* Sakharov, this difficulty has been argued to be a manifestation of the ubiquitous non-relativistic behavior of the effective fields at the scale that plays the role of the Planck scale in these systems [8, 19]. Again, there exists an independent result that adds confusion to the mixture: it is widely quoted that any consistent theory that describes the nonlinear behavior of excitations with the properties of gravitons (or gravitational waves) must be given by general relativity [20, 21]. This seems to indicate a simple path to obtain general relativity in an emergent framework [22]. Delineating the reasons why retrieving general relativity from condensed-matter-like systems remains elusive would probably lead to new insights on the idiosyncrasies of the gravitational interaction.

This is the ambience that sets the motivation for the questions that are addressed in this dissertation. During the last years, we have tried to make sense of our intuitions, making an effort to reconcile our findings with seemingly contradictory results that frequently come from very different research areas, and are therefore expressed in different languages (that could be nevertheless deceptively similar sometimes). The tools that have been used are necessarily scattered between different fields: from condensed matter physics and non-relativistic many-body systems, to relativistic classical and quantum field theory (in general spacetimes), effective field theory, spacetime geometry and general relativity.

The contents of this thesis are presented in two separate parts, containing two chapters each, and differentiated by the nature of the topics that are discussed in them. The first part of the thesis deals with fundamental questions in the emergent gravity program, for instance: Is it possible for Lorentz invariance to be realized as an emergent symmetry? Which kind of mechanism could ensure the emergence and stability of Lorentz invariance? Is it possible to make explicit the connection between the properties of relativistic wave equations and fields and the basic properties of the underlying condensed-matter-like systems? What is the role of gauge invariance in the overall picture? Why obtaining general relativity in these systems as an effective theory remains elusive?

Given the difficulties in constructing an emergent theory of gravity within this setting, we decided to explore in detail all the steps involved in the construction of the much simpler case of emergent electrodynamics. A deep knowledge of this simpler problem would probably result useful as a preparation to the gravitational problem. Moreover, to our knowledge, there does not exist a work of reference in which this construction in condensed-matter-like systems is performed in a step-by-step fashion, making transparent all the hypotheses and approximations involved. It is our intention that this work may serve as well as a study guide for specialists in other approaches to quantum gravity, which may be more accustomed to the language and intuitions that are natural to their respective areas. The step-by-step construction of an emergent theory of (massless) electrodynamics makes up Chap. 1.

Although our current experimental knowledge of quantum electrodynamics has not asked for a revision, it is still interesting to analyze the structure of a possible deeper layer underneath electrodynamics. From an exercising perspective, as we have said it is always helpful to understand simpler systems before embarking in more complicated endeavors. From a more physical perspective, there are partial emergent models that suggest that gravity and electromagnetism might emerge in a unified manner from a single underlying system [8], which is for instance the situation in string theory also. On the other hand, if the very arena in which physics takes place (spacetime) has a discrete underlying structure, it is sensible to think that electrodynamics would share this structure. This observation can be certainly extended to the complete structure of the standard model of particle physics.

We shall present two models of emergent electrodynamics, with the aim of highlighting their coincident features. One of the models is originally due to Maxwell himself [23, 24]. We revise and slightly update Maxwell's hydrodynamical model in the light of the physics we know today. The other model, which constitutes the bulk of the chapter, is more sophisticated and is based on ideas coming from what we know about the superfluid phases of Helium–3. This construction follows the lead of the works of Volovik (see [18, 25, 8, 26] and many other references therein), and among other things intends to make his ideas more accessible to non-specialists in condensed matter. Many steps in the construction have our own perspective though, so that any misjudgment or error can only be blamed on us.

The contents of Chap. 2 deal with the particular properties of the gravitational interaction that mark the difference with respect to electrodynamics, specially its nonlinear (non-abelian) nature. While electrodynamics is an abelian theory in which the force carriers (photons) are not charged, gravitons carry energy, and therefore a nontrivial gravitational field [27]. When gravity is described by means of classical relativistic field theory, this nonlinear character translates into the fact that the classical action is not quadratic in the gravitational field. Indeed, when expanding the Einstein-Hilbert action around Minkowski spacetime one obtains an infinite chain of polynomial terms, or interaction vertices. To obtain this infinite set of polynomial terms without invoking the very fundamental principles of general relativity, which is the goal of the emergent gravity program, seems to be a daunting task.

However, there exists in the literature a well-known route to determine the right set of polynomial terms in the gravitational action, formulated by means of the so-called graviton self-interaction problem. In this picture one starts with the linear representation of gravitons as free particles propagating in Minkowski spacetime, and considers the possible coupling of these particles to other particles describing matter fields or even other known interactions [28]. These simple ingredients lead to the conclusion that gravitons should be described by a non-abelian theory. In principle, it is possible following this path to determine the right set of interaction vertices and reach in this way general relativity or, in other words, to derive diffeomorphism invariance unambiguously [20, 21]. Nevertheless, there have recently appeared a set of works that raise doubts about the uniqueness and the overall legitimacy of this construction [29, 30].

Our in-depth analysis of this topic aims to settle this controversy. Understanding to what extent there exists an alternative, non-geometric derivation of the geometric structure of general relativity is of clear importance for the emergent gravity program. It could also offer suggestions about the root of the difficulties in obtaining general relativity at low energies in condensed-matter-like models.

The second part of the thesis encloses the application of some of the results or ideas that appear in the preceding part. One of the indications of the degree of development of a given research program is its capacity of suggesting solutions to open problems, or even possible phenomenological consequences. That we use of the word suggestion is not accidental; it is frequent in physics that there is no proof (by a mathematician's standard) that the developments at different levels of the effective description of a system are equivalent. Only a fine knowledge of these different layers could permit to reduce the breach between them.

In this thesis we have considered two renowned problems in the framework of quantum gravity. The first of them is the so-called cosmological constant problem [31, 32, 33, 34]. We devote Chap. 3 to a possible way of alleviating this problem in the framework of emergent gravity. Given the overwhelming number of long treatises about this problem, we have opted for a concise writing style in which the root of the problem is unambiguously highlighted. Then, it is directly shown that a specific nonlinear theory of gravity that is motivated by our previous discussion in Chap. 2 escapes the problem. Also we include a brief discussion about a promising way to fix the value of the cosmological constant, proposed by Grigory Volovik [35, 36], that provides a bridge between the contents of this chapter and that of Chap. 1.

The discussion in Chap. 4 deal with a completely different problem, namely the gravitational collapse of massive stars to black holes. The determination of the properties of black holes in quantum gravity represents by itself an independent area of research. Much has been written since the original works of Stephen Hawking [37, 38] about the fate of singularities and horizons in a suitable ultraviolet completion of general relativity, and the consequences that would follow to derived problems such as the information loss in black holes [39]. However, our discussion in this chapter can be considered orthogonal to previous analyses. The emergent gravity program based in condensed-matter-like models presents a strong suggestion about the behavior of the gravitational interaction at high densities that would inevitably affect the gravitational collapse of massive stars [40]. The subsequent picture is hardly reachable in a framework that considers the geometric structure of general relativity (and black holes) to be fundamental. An extensive analysis of the effective geometries corresponding to this process, describing the transition between a black-hole and a white hole-geometry, is presented, together with an exploration of the possible phenomenological consequences.

The discussions in each of the chapters are intended to be self-contained. Consequently, all of them contain a proper introduction and concluding section. These sections also serve as junctions between the contents of the different chapters. The bigger picture that motivates these developments is stressed in both this introduction, and the last section exposing the main conclusions and a brief description of future directions that could be taken from here.

Notation and conventions

The International System of Units (SI) is used throughout the text. Physical dimensions are denoted by integer powers of capital letters, surrounded by brackets. For instance, $[L]$ would correspond to the physical dimension of length, and $[LT^{-1}]$ denotes the physical dimensions of velocity. The list of capital letters being used is the following:

Length	L
Mass	M
Time	T
Charge	Q

When used to relate two physical magnitudes, the symbol $\sim$ implies that these two quantities are of the same order of magnitude. On the other hand, $\simeq$ indicates an approximate value.

When used as superscripts and subscripts, Latin and Greek lowercase letters in italic fonts are assumed to take the integer values $0, 1, 2, 3$. The only exception are the Latin letters i, j and k, which take the values $1, 2, 3$ due to their association with the axis of the three-dimensional Cartesian coordinate system. When the introduction of superscripts or subscripts that are valued in a different range of integers is needed, capital Latin letters will be used, and the corresponding range will be identified.

Frequently, some definitions in the text will make use of specific superscripts or subscripts, or groups of them, in order to keep track of the corresponding magnitudes. These quantities are not regular superscripts or subscripts but just labels, and therefore do not take any particular values. Potential confusions are avoided by using the standard text font for them. To put an example, in p^i the superscript i takes the values $i = 1, 2, 3$, but p^{i} is just a given quantity (for instance, a specific value of the magnitude p that is singled out among others by some physical reason).

We generally work in four spacetime dimensions, making use of the signature convention $(-1, +1, +1, +1)$. It will be explicitly specified when working in a different dimensionality. Besides from this, we follow the conventions of [41] regarding the description of spacetime geometries.

Part I

Fundamentals

Chapter 1

Lessons from emergent electrodynamics

1.1 Emergence and many-body systems

The objective of this chapter is to introduce an emergent theory of electrodynamics that is largely based on the widely developed physics of superfluid helium. Before entering into specific details of the construction, in this section we shall set up some general definitions and notation, which are intended to be a useful guidance for the reader.

From a practical perspective, the basic assumption behind the emergence program is to assume that all the known fundamental particles, and interactions between them, are nothing but excitations of an underlying medium [42, 43, 44]. If this picture is realized in nature, then all the physical processes we can currently describe by using both the standard model of particle physics and general relativity would only be the top of the iceberg, the superficial manifestations of the properties of a hypothetical medium, the physical characteristics of which are still largely unknown.

Given the lack of a compelling fundamental principle defining the nature and characteristics of the constituents of this hypothetical underlying medium, a possible strategy (and the one we want to pursue here) is to study to which extent the properties we observe at low energies may arise in a robust way, i.e., without depending on the fine details within a large equivalence class of systems. As emphasized in Volovik's work [18], the occurrence of certain topological features such as Fermi points may entail an efficient mechanism that ensure the appearance of these features in a wide class of many-body systems. The description of many-body systems and these topological features find a natural language in the formalism of quantum mechanics.

In a very simplified vision, the basics of the model we are going to use can be understood in terms of a single harmonic oscillator. A quantum harmonic oscillator is described in terms of a set of states and observables taking different values on each state. In particular, each state has a definite energy, and there exists a state with minimum energy, the so-called ground state or vacuum $|0\rangle$.[1] Excitations on top of this state are labeled by an integer $n \in \mathbb{N}$, i.e., $|1\rangle$, $|2\rangle$,... and span the entire space of states. These excitations can be obtained by making use of the so-called creation operators:

$$|n\rangle = (a^\dagger)^n |0\rangle, \tag{1.1}$$

where $(a^\dagger)^n$ is defined by induction as $(a^\dagger)^{n+1} = a^\dagger \times (a^\dagger)^n = (a^\dagger)^n \times a^\dagger$, and $(a^\dagger)^0 = 1$.

When we move on to a many-body theory, some of these notions are still useful. There will be again a ground state which, given the additional complexity of the system, can support nontrivial properties such as topological defects and collective excitations [26]. Excited states are constructed by the action on the ground state of creation operators with different labels, that correspond to the classification of these excitations into families with different properties. The functional dependence of two of the observables of these excited states, the energy on the one hand and the momentum on the other, leads to the notion of dispersion relation. This notion permits to classify the excited states of the system in different branches of states. In general, a model of this sort would contain a gapless branch (meaning that these excitations can be created with arbitrarily small energy) as well as gapped branches (only excitations with energies above a certain scale, or *gap*, can be produced).

This framework has its roots in condensed matter physics, in which it is certainly ubiquitous: for example, phonon excitations are gapless excitations, while the bands in semiconductors correspond to gapped branches. Indeed, our goal here is to discuss a condensed-matter-like system in which the ground state contains Fermi points. In this case, the spectrum of excitations would be schematically the one presented in the left-hand side of Fig. 1.1. The first excited branch, corresponding to the gapless excitations, would present the typical parabolic form expected in Galilean physics but for the marked effects of interactions, which enforce that the corresponding gapless excitations have zero energy in two points in momentum space. These points are called Fermi points. The relevant observation which acts as the trigger of all the following discussion is that the excitations near these points in momentum space approximatively follow a relativistic (linear rather than parabolic) dispersion relation. This linear behavior would only be valid below certain characteristic energy $E_{\rm L}$ whose value would depend on the details of the model. The form of the gapped branch would also be model dependent but, as there is no mechanism similar to the occurrence of Fermi points in the gapless branch, it will have all the features of a Galilean branch. In other words, the Planck-scale degrees of freedom are not relativistic at all, which is not surprising given the nature of the model. However, this should not hinder the relativistic invariance of the low-energy excitations, as it is indeed observed in everyday condensed-matter experiments (e.g., measuring dispersion relations of phonons). One of the goals of this chapter is studying in detail the transition between these two regimes, that is, the emergence of a low-energy relativistic phenomenology and the mechanisms that ensure the decoupling of the low-energy degrees of freedom from the non-relativistic underlying physics.

[1] In the incoming discussion, the first notation is probably the most adequate one, as this state will not necessarily be void in the intuitive sense of the word.

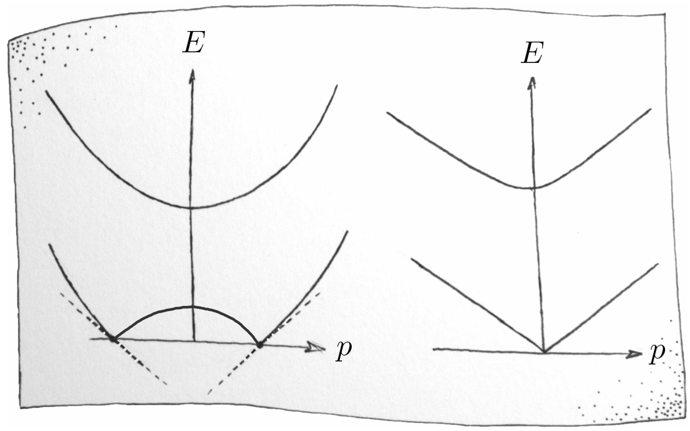

Figure 1.1: Dispersion relations for both gapless and gapped excitations. Vertical axes correspond to the energy E, horizontal axes to the Euclidean norm p of the momentum $\boldsymbol{p}$ in three-dimensional momentum space. On the left we sketch the Galilean (parabolic) dispersion relation with Fermi points, corresponding to the intersection between the lowest branch and the momentum axis. On the right we depict the typical form of the relativistic dispersion relation for both massless (below) and massive excitations (above). While these two figures are globally different, we can always use the linear Taylor approximation close to the Fermi points so that, locally around these points, the figure on the left will be indistinguishable from the lowest branch in the picture on the right.

This discussion may sound familiar to a reader accustomed to string theory. Contrary to our discussion, in string theory the emphasis is put on the guiding principle to construct the fundamental underlying medium [45]. As the reader would probably know, the essential assumption in this approach is that the basic elements of nature are relativistic strings. The ultimate theory constructed by following this principle, which could be depicted in some sense as the theory of a "liquid" of these strings, is not completely understood nowadays. However, it is already known how to evaluate certain properties in this framework, such as the spectrum of excitations of the strings around Minkowski spacetime. As a result, one obtains indeed a branch of gapless (massless) excitations which should correspond to the low-energy fields we have observed so far in nature. Over these states one has a tower of gapped states, with the characteristic string scale controlling the gap value. In the critical spacetime dimension, the dispersion relations of all these branches have the characteristic relativistic form (which is linear for the gapless excitations; see Fig. 1.1).

The goal of this chapter is to develop in detail a model of this sort that, at low energies, leads to electrodynamics. Let us stress that we do not take the Galilean invariance of the model to be constructed or, in other words, its Newtonian character, as a fundamental principle which has to be necessarily realized in nature. For us it is the emergence of Fermi points the essential property to be focused on. This feature, which is robust in the sense that it does not depend heavily on the fine details of the underlying physics [18, 8], effectively erases any trace of the Galilean symmetry at low energies, as well as other features of the underlying system. This is an example of a mechanism that agglutinates a large class of systems (a *universality class*) in what concerns their lowest-energy excitations, and which may be behind the emergence of the properties we observe in our world. One could even imagine the possibility of constructing abstract theories, not based on Newtonian or Galilean notions, but that nevertheless present features that resemble the properties of Fermi points. We do not have knowledge of any such formulation, but it seems to us an interesting direction to consider. It is also important to keep in mind that our ultimate goal is not the construction of a real condensed matter experiment that mimics electrodynamics. The values of the characteristic constants of the model that reproduces electrodynamics in a certain regime may probably imply its unviability as a tabletop experiment.

On the other hand, we are considering electrodynamics as a toy model for the much complex case of gravity. Despite its simplicity, it presents nontrivial features from the perspective of emergence. Understanding these features in this simpler setting could prove very helpful in order to handle more involved scenarios. Remarkably, the history of electrodynamics itself is inextricably tied up to the emergent program: Maxwell arrived to his unification of light and electromagnetism through the development of a mechanical model that could underlie all the electromagnetic phenomena [23, 24]. For both historical and physical reasons we review this model in the following section, before embarking in the superfluid model. In brief, he imagined the electromagnetic aether as consisting of an anisotropic and compressible fluid made of cells, capable of acquiring rotation, separated by a layer of small idle wheels or ball bearings capable of rotating and moving between the cells. The bodies would be immersed in this fluid as an iron ball is immersed in water; they would distort the fluid around them. He did not commit with this specific model as truly representing physical reality, though, but defended it on the grounds of a proof of principle of the possibility of formulating electromagnetism as a mechanical model. Let us stress that our approach to the construction of a specific emergent model is conceptually the same.

It has been argued that the most crucial step made by Maxwell was to abandon his mechanical model and just worry about the properties of the resulting coarse-grained effective field theory [46, 47, 48]. The field-theoretical point of view has since been a central theme in most developments in fundamental physics. Whereas nobody can deny the tremendous power and success of this approach, it assumes many ingredients as a matter of principle, without a deeper explanation. The following set of questions may serve as an example: Why is there a maximum velocity for the propagation of signals? Why is there gauge invariance? Why are elementary particles within a class indistinguishable? Why are there no magnetic monopoles? As we will see, an emergent approach is capable of providing explanations for many of these questions. On the other hand, an emergent perspective puts a stronger accent on the universal characteristics of possible microscopic theories than on the specifics of a particular implementation. We think that the emergent approach complements the field-theoretical approach, together providing a much richer source of understanding.

1.2 An updated Maxwell fluid model

In Maxwell's time people did not have a clear idea of what electric currents really were, not to mention the then unknown atomic structure of matter. Given the present knowledge, we can propose an updated fluid model for electromagnetism following closely Maxwell's proposal [23, 24]. For other modern viewpoints on Maxwell's hydrodynamical model the reader might find interesting, e.g., [49, 50, 51, 52, 53, 54, 55].

Imagine a fluid made of two different elementary constituents: vortical cells and small ball bearings. A vortical cell is made of a topologically spherical and deformable membrane filled with a fluid. The details of this fluid are not very important in what follows so, to simplify matters, let us take it to be incompressible and highly viscous. The membrane provides a fixed constant tension in all its points. It supports tangential as well as normal tensions. In the case in which the membrane rotates around an axis, the filling fluid would rapidly end up rotating with a uniform angular velocity around that axis. The total angular momentum of the vortical cell will be $I\,\Omega$, with I its moment of inertia and Ω its angular velocity, or $Iv_{\rm e}/r_{\rm e}$, with $r_{\rm e}, v_{\rm e}$ its equatorial radius and velocity [56].

The fluid inside the cell has an isotropic hydrostatic pressure. When it is non-rotating, this pressure takes a constant value p_0 throughout the cell. However, rotation provides centrifugal forces that change the pressure pattern: in the equator the pressure will have an excess $p_0 + \frac{1}{2}\rho v_{\rm e}^2$ with respect to the poles. Independently of the precise form of the cell, the cell as a point will exert a pressure excess in the directions orthogonal to the rotation axis, and this pressure excess will be proportional to the rotation velocity squared:

$$p_\parallel = p_0 + C_\parallel \Omega^2, \qquad p_\perp = p_0 + C_\perp \Omega^2, \qquad C_\parallel < C_\perp. \tag{1.2}$$

We can also write this excess as

$$\Delta p = p_\perp - p_\parallel = \mu_{\rm micro}^{-1} B_{\rm micro}^2. \tag{1.3}$$

At this stage the dimensions of $\boldsymbol{B}_{\rm micro}$ and $\mu_{\rm micro}$ are not fixed but only the dimensions of the above product. For later convenience let us choose $\boldsymbol{B}_{\rm micro}$ to denote minus the average density of angular momentum in the vortical cell, multiplied by a typical length scale R in the system,

$$\boldsymbol{B}_{\rm micro} = -\frac{I}{V} R\, \boldsymbol{\Omega}. \tag{1.4}$$

Then, the quantity $\mu_{\rm micro}$ is a constant with units $[ML]$ (mass times length). Although the dimensions have been fixed one can still multiply $\mu_{\rm micro}$ by a dimensionless number N and $\boldsymbol{B}_{\rm micro}$ by $\sqrt{N}$ with no effect, or in other words, one can change the value of the length scale R that defines $\boldsymbol{B}_{\rm micro}$ if one redefines $\mu_{\rm micro}$ accordingly. One could also have defined $\boldsymbol{B}_{\rm micro}$ with a reversed sign with respect to the definition in Eq. (1.4) with no effect; in fact, we have chosen the negative sign for later convenience. A specific definition of $\boldsymbol{B}_{\rm micro}$ will only appear when fixing an operational meaning for it. Let us advance here that when later introducing the unit of charge, it will be natural to define the macroscopic version of $\boldsymbol{B}_{\rm micro}$ with units $[MT^{-1}Q^{-1}]$, and the macroscopic version of $\mu_{\rm micro}$ with units $[ML\,Q^{-2}]$.

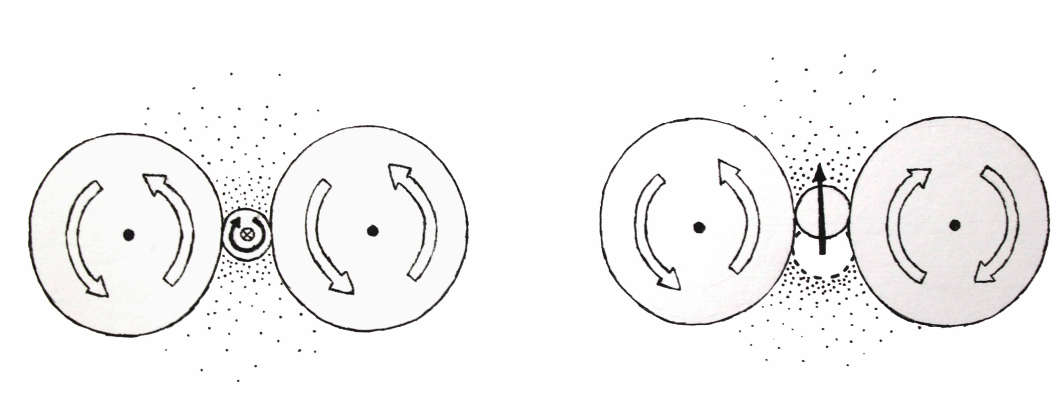

Figure 1.2: Diagram showing the transfer of rotation between the different elements in the fluid (on the left) and tensions due to small displacements of ball bearings (on the right).

On the other hand, a ball bearing is a small spherical ball, much smaller than a vortical cell, that sticks to any membrane in such a way that whereas it can move over it, it cannot slide. That is, any movement has to be accompanied either by rotation of the surrounding vortical cells or by a tangential stretching of the membrane itself (see Fig. 1.2). This fluid of small balls is also endowed with an isotropic hydrostatic pressure. This pressure produces a displacement of the microscopic distribution of ball bearings with respect to the vortical cells that in turn produces microscopic restoration forces. This is due to the fact that most ball bearings will be attached to at least two vortical cells so that the only way to move them is by creating a tangential distortion (and a subsequent tension) on the membranes. Thus, the hydrostatic pressure of the ball bearings combined with their stickiness results in an equilibrium state endowed with tensions, which we will call, in a modern language, ground state.

The complete description of a fluid made of a huge number of vortical cells with an even larger number of ball bearings stuck to their surfaces, all put together in a box, would be tremendously complicated and uncontrollable in practice. However, from a coarse-grained perspective, we could use just a few macroscopic variables to characterize the state of the fluid, as it is done in standard fluid mechanics. Consider one small part of the fluid but still containing a large number of constituents. At any such coarse-grained point the vortical cells will contribute with an overall hydrostatic pressure $p_{\rm H}$ plus some tension acting in a specific direction, the overall rotation axis. This leads to an anisotropic pressure that can be written as

$$p_{ij} = \delta_{ij} p_{\rm H} - \mu_0^{-1} B_i B_j. \tag{1.5}$$

Here, the vector $\boldsymbol{B}$ is the macroscopic version of $\boldsymbol{B}_{\rm micro}$ and is therefore proportional to the angular momentum density (total angular momentum in the coarse-grained point divided by its volume). The quantity μ_0 is a constant with dimensions $[ML]$. The same redefinition ambiguities associated with the microscopic quantities as defined by Eq. (1.3) apply to their macroscopic versions. These $p_{\rm H}$ and $\boldsymbol{B}$ are our first macroscopic variables.

Now, excited states can have tangential displacements of the ball bearings beyond their equilibrium positions, with their associated restoring tensions as illustrated in Fig. 1.2. We can characterize these tensions by a microscopic displacement vector field $\boldsymbol{D}_{\rm micro}$ (displacement of each ball bearing with respect to its position in the ground state). At the coarse-grained level we can construct a displacement-density vector field $\boldsymbol{D}$ and the associated force field $\boldsymbol{E} = \epsilon_0^{-1}\boldsymbol{D}$, with ϵ_0 being for now a free constant with the appropriate dimensions. The real restoration force field is proportional to the displacement and hence to this force field $\boldsymbol{E}$.

To recover classical electrodynamics from a fluid system like the one being described, we still need one more ingredient: something has to play the role of charge. In the ground state, ball bearings are all strongly stuck to vortical cells so that they cannot move from one vortical cell to another. However, out of this ground state, there can be movable ball bearings able to performing macroscopic displacements jumping from cell to cell. To introduce movable ball bearings in the system one could even break some of the strong links of the ball bearings characterizing the ground state. In this way ball bearings can be relocated in space. As we will see, the presence of regions with an overdensity or an underdensity of movable ball bearings with respect to the ground state can be associated with a positive and a negative charge density, respectively. These overdensities and underdensities will in turn be responsible for the redistribution of tensions in excited states that we described before.

Now we are in a position to explain how the entire set of Maxwell's equations arises in this fluid system:

i) Any rotational of the force field $\boldsymbol{E}$ will exert a torque that will increase the angular momentum of the vortical cells, and thus decrease $\boldsymbol{B}$ since we have defined it to be minus an angular momentum density in Eq. (1.4). This proportionality is one of Maxwell's equations. Once a specific meaning for $\boldsymbol{B}$ is given (recall that one can redistribute a constant dimensionless factor N between $\boldsymbol{B}$ and μ, or in other words, one has an initial flexibility in defining the length scale R), one can always find a specific ϵ_0 so as to write
$$\frac{\partial \boldsymbol{B}}{\partial t} = -\boldsymbol{\nabla} \times \boldsymbol{E}. \tag{1.6}$$
In other words, this equation can be interpreted as fixing the value of ϵ_0 with to respect μ_0, and so fixing the relation between the force field $\boldsymbol{E}$ and the displacement field $\boldsymbol{D}$.

ii) The presence of a ball-bearing overdensity or underdensity produces a change in the displacement field. Assuming the displacements to be sufficiently small one can write
$$\boldsymbol{\nabla} \cdot \boldsymbol{D} = \rho_Q, \tag{1.7}$$
or
$$\epsilon_0 \boldsymbol{\nabla} \cdot \boldsymbol{E} = \rho_Q. \tag{1.8}$$

The charge density ρ_Q is tightly related to the density of ball bearings with respect to the ground state. However, there is no need that they perfectly coincide. The only quantity with a macroscopic operational meaning (at least at this linear level) is the charge density. At this stage one could introduce some reference unit of charge, and accordingly change the units of all the quantities by referring them to the effect of this reference charge.

iii) When the ball bearings move they exert torques on the cells. This applies to both ball bearings strongly stuck to the cells (not movable to other cells), that produce a change in the displacement field, and to ball bearings movable between cells (they are associated with charge currents). Reciprocally, when the rotation field of the vortical cells $\boldsymbol{B}$ acquires some rotational, it causes the ball bearings in the region to move within their respective possibilities. This behavior is encoded in an equation of the form

$$\boldsymbol{\nabla} \times (\mu_0^{-1} \boldsymbol{B}) = \boldsymbol{J}_T = \boldsymbol{J}_Q + \frac{\partial \boldsymbol{D}}{\partial t}. \tag{1.9}$$

The first term of the current $\boldsymbol{J}_T$ is due to the movable ball bearings (a proper current of charge) while the second term is due to the displacement of the non-movable cells; hence its name displacement current. This equation fixes the value of μ_0 or, equivalently, the precise definition of $\boldsymbol{B}$ [in other words, the equation determines the value of the length-scale constant R introduced in Eq. (1.4)].

iv) Let us assume that the rotation field is divergenceless in average, although a priori there is no reason why this should be the case:

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0. \tag{1.10}$$

Indeed, the fluid model a priori allows magnetic monopole configurations. However, when looking at the model carefully one realizes that this kind of configurations does not seem to be favored by the system. Microscopically speaking, a magnetic monopole involves rotating cells with their angular momenta distributed radially. Any ball bearing located at the confluence of these cells will produce friction since the cells cause dragging forces incompatible with the no-sliding condition (see Fig. 1.3). It seems reasonable that the system would tend to avoid these configurations; the divergence-free condition (1.10) arises, when applying the coarse-graining procedure, as the result of the low probability of these configurations.

Following these steps one retrieves all of Maxwell's equations from a mechanical fluid system. We would like to make the following observations, that not only are interesting by itself, but some of them will be worth keeping in mind when considering the superfluid model in the next sections:

- As already remarked by Maxwell, the important point here is not the fine details of this specific fluid model but the fact of its very existence. Due to coarse-graining, the equations for the macroscopic fields will not depend strongly on these details. Note

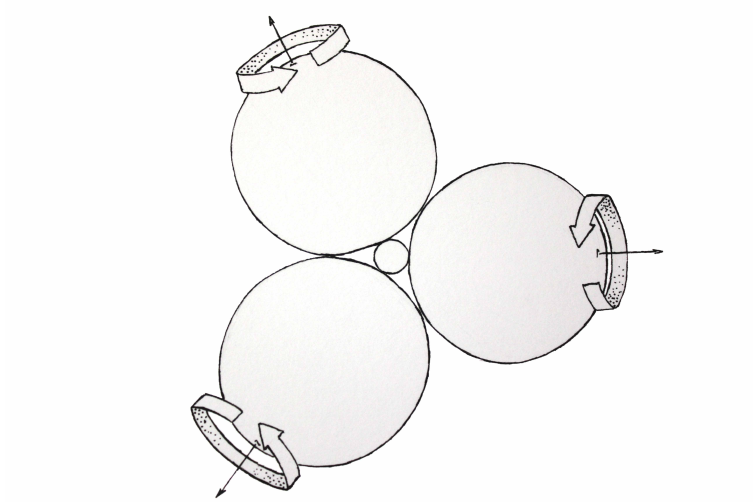

Figure 1.3: Diagram explaining the absence of magnetic monopoles. The system avoids these configuration because they would create friction for the central ball bearing at confluence of the rotating cells.

that in the derivation it has been necessary to make the assumption of smallness of all the perturbations with respect to the ground state. Beyond this regime, one would expect to observe non-linear effects. For instance, one could expect non-linear pressures of the form

$$p_{ij} = \delta_{ij} p_0 - \mu^{-1}(B, E) B_i B_j. \tag{1.11}$$

In the linear limit one can approximate $\mu(B, E)$ by a constant μ_0.

- From Maxwell's equations one immediately deduces that this system admits light-like perturbations. These perturbations propagate with a speed $c = \sqrt{\epsilon_0 \mu_0}$. The fluid system can be perfectly described using Newtonian physics in which there is no limitation to the velocity of the bodies. Nonetheless, the light speed shows up directly from the elastic properties of the body. The crucial ingredient for generalized sound velocities to emerge is that variations in time of local properties depend on the local gradients of the very same properties.

- Given that $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$, one can always write $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$ locally. For instance we could associate $\boldsymbol{A}$ with the macroscopic version of the flow lines of the fluid within the vortical cells. On the other hand, in places in which magnetic fields are stationary, $\boldsymbol{\nabla} \times \boldsymbol{E} = 0$ so that one can write $\boldsymbol{E} = -\boldsymbol{\nabla}\phi$ locally. Then, the field ϕ represents a hydrostatic electric tension. A positively charged body will tend to move to places with smaller electric tension. In more general situations, due to the structure of Maxwell's equations, we can always write $\boldsymbol{E} = -\boldsymbol{\nabla}\phi + \partial_t \boldsymbol{A}$. Knowing the coarse-grained structure of the fluid flow lines and hydrostatic electric tension, one knows ϕ and $\boldsymbol{A}$.

Now, one can realize that regarding the values of $\boldsymbol{E}, \boldsymbol{B}$, the combination $\{\phi, \boldsymbol{A}\}$ and $\{\phi + \partial_t \chi, \boldsymbol{A} + \boldsymbol{\nabla}\chi\}$ are equivalent. From an emergent perspective an appropriate interpretation of this effective gauge invariance is that, although the flow structure of the vortical cells and the electric tension both have a specific reality, different macroscopic configurations related through a gauge transformation are operationally indistinguishable from the effective dynamical theory (see also [53, 54]). Gauge invariance appears because aspects of the system are "invisible" to observers restricted to experience only the low-energy effective fields.

- In Maxwell's version of the fluid model, charges were associated directly with individual ball bearings, and currents with the movement of ball bearings from cell to cell. It seems completely unrealistic to have an electric current of this sort without some resistance or friction. But this resistance was perfectly accommodated in Maxwell's model by considering the currents as existing only within materials (conductors). Maxwell ascribed the ubiquitous resistance in a conducting wire to the collisions of the movable ball bearings when jumping between cells. However, this model will have a hard time to deal with a charged elementary particle in an otherwise empty space. There is no experimental evidence that the vacuum causes friction on a charge moving with uniform velocity. In fact, the presence of an effect of this sort would immediately uncover the existence of a privileged inertial reference frame, against all we know about the relativity principle.

  We wanted our updated fluid framework to encompass also the movement of a free electron in an otherwise empty space. For that we have proposed to associate an electron to an overdense region of ball bearings (this identification is of course not complete as we have not attempted to specify the internal forces responsible for its structural stability). When thinking of an electron as a very localized overdense region, it appears difficult to find it in a uniform-velocity trajectory without dissipation of some sort. The movement of the ball bearings would be very noisy, with multiple collisions involved. However, if one imagines a pure plane-wave distribution of the overdense regions, it appears perfectly plausible for the propagation of the wave to occur without any appreciable friction: the propagation of the wave will not involve the presence of a macroscopic current of ball bearings. It is interesting to point out an analogy between this behavior and that of free quantum particles. A position eigenstate of a wave function can propagate with a certain velocity, but as it travels it diffuses in space. This diffusion might be seen as an analogue in the quantum formulation of the noisy propagation expected in our fluid model. On the other hand, a momentum eigenstate is also an eigenstate of the Hamiltonian and it is not distorted by the propagation.

  It is also interesting to realize that the system allows in principle the creation of pairs of particles from the ground state. If one pulls out ball bearings in one place and moves them to another region, one would have created equivalent overdense and underdense regions. The appearance of these "quantum-like" behaviors in our model might be interpreted as suggesting that quantum mechanism itself could be emergent. Here we just mention this possibility without pursuing it any further.

- It is interesting to estimate how small the constituents of this fluid would have to be in order to pass unnoticed to current experiments. The smallest length scale ever tested is of the order of 10^{-19} m. So, in principle, a fluid structure several orders of magnitude beyond 10^{-19} m would remain undetected. Notice that the Planck length is 10^{-35} m, still 15 orders of magnitude ahead. This would be equivalent to compare the size of a human being with interstellar distances.

- It is also interesting to point out the different nature of light excitations and charged matter, even when considering the latter as wave-like. One can perfectly imagine a charge density with no overall velocity, representing a charged particle or distribution of particles at rest. Light is however similar to a phonon excitation of a lattice, as it always travels with its fixed velocity (of course, at very high energies one would expect to develop some dispersive effects).

1.3 Model based on Helium–3

In the rest of the chapter we shall present a model of emergent electrodynamics based on the well-established theoretical understanding of the physics of the superfluids phases of Helium–3. Our presentation of the model will follow a top-down scheme. Nonetheless, we would like to stress the fact that these theoretical ideas were developed in close feedback with experiments and are proved to a great extent by them. Indeed, it would be far from straightforward to motivate some of these approximations in a purely mathematical framework, a feature that is worth keeping in mind.

Most of the introductory material covered here is well understood nowadays but, as far as we know, it has not been presented in a logical step-by-step order so as to lead to a final picture of emergent electrodynamics. For a general discussion on superfluid He–3, one can draw on the review [57], or the books [58, 59] and references therein. Concerning the low-energy properties of this system and analogies with other branches of physics, including relativistic field theories, the seminal reference is [18]. In the following, references are quoted only in case they are relevant to specific points in the discussion.

1.3.1 Non-interacting fermions: the Fermi surface

The model that we will be eventually discussing is an ensemble of a large number N of interacting, non-relativistic quantum fermions with spin 1/2. In order to introduce the notation we will use, as well as some concepts that will be convenient later, we first introduce briefly the properties of the non-interacting ensemble, the so-called ideal Fermi gas [60, 59, 61].

Let $|0\rangle$ denote the state in which there are no fermions, that is, it describes empty space. Now we can imagine that a single fermion, traveling with constant speed,[2] is placed on this empty space. The creation operator that corresponds to the excited state with a fermion of the family B and momentum $\boldsymbol{p}$ will be denoted as $a^{+}_{\boldsymbol{p}\,B}$.[3] This excited state is then obtained from the vacuum as

$$|\boldsymbol{p}\,B\rangle := a^{\dagger}_{\boldsymbol{p}\,B}|0\rangle. \tag{1.12}$$

This state corresponds to a plane-wave solution of the corresponding Schrödinger equation, with energy

$$E_{\boldsymbol{p}\,B}(\boldsymbol{p}) = \frac{\boldsymbol{p}^2}{2m_B}. \tag{1.13}$$

This is the usual Galilean, or parabolic dispersion relation in which the energy of a particle is proportional to the square of its momentum. The proportionality coefficient depends on the mass m_B of the family to which the fermion belongs. In the following we will consider, without losing any fundamental aspects, that the mass is the same for all the families so that we will drop the B subindex when writing this parameter.

For each state $|\boldsymbol{p}\,B\rangle$ we may define an operator $a_{\boldsymbol{p}\,B}$ so that

$$a_{\boldsymbol{p}\,B}|\boldsymbol{p}\,B\rangle = |0\rangle. \tag{1.14}$$

These operators are dubbed annihilation operators and, by definition, the state $|0\rangle$ is on their kernel. As we are dealing with fermions, the annihilation and creation operators follow the usual anticommutation relations which, in finite volume, take the form [61]:

$$\{a_{\boldsymbol{p}\,B}, a^{\dagger}_{\boldsymbol{p}'C}\} = \delta_{\boldsymbol{p}'\boldsymbol{p}}\delta_{BC}. \tag{1.15}$$

The momentum Kronecker delta will be transformed into a Dirac delta function in the infinite volume limit. The rest of anticommutation relations are $\{a_{\boldsymbol{p}\,B}, a_{\boldsymbol{p}'C}\} = \{a^{\dagger}_{\boldsymbol{p}\,B}, a^{\dagger}_{\boldsymbol{p}'C}\} = 0$. These commutation relations represent the way of encoding the Pauli exclusion principle, so that two fermions cannot share the same quantum numbers.

The exclusion principle has important consequences for the ground state of a collectivity of fermions. While the ground state for an arbitrary number of fermions is $|0\rangle$, representing the absence of fermions, this state would not be attainable if we are interested in the energy regime in which fermions cannot be destroyed (or created). We will restrict our discussion to the set of states with N fermions, and ask the question about the state which minimizes the energy within this family of states, denoted by $|\Phi_0\rangle$. To be more precise, we could indeed permit fluctuations in the number of fermions, as long as these fluctuations are small with respect to N. This situation corresponds, in a statistical mechanics framework,

[2] In the Newtonian framework we are working in, it is assumed the existence of some additional structure that determines the family of inertial reference systems [62, 56].

[3] Following the standard conventions, in the rest of the text we will always write O instead of $\hat{O}$ for any operator, using this notation only for unit vectors.

to the regime in which a thermodynamical description within the grand canonical ensemble makes sense [63].

In the grand canonical ensemble, the chemical potential enters explicitly in the evaluation of the partition function, so that the relevant energy operator is given by

$$H - \mu N := \sum_{\boldsymbol{p}\,B} \left(\frac{\boldsymbol{p}^2}{2m} - \mu \right) a^+_{\boldsymbol{p}\,B} a_{\boldsymbol{p}\,B}. \tag{1.16}$$

In the following we shall refer to this operator as the Hamiltonian or energy operator. While one may intuitively think that excitations with momentum lower than the Fermi momentum,

$$p_{\mathrm{F}} := \sqrt{2m\mu}, \tag{1.17}$$

have negative energy, to correctly interpret this formal statement one has to take into account the properties of the ground state $|\Phi_0\rangle \neq |0\rangle$. This ground state is constructed by progressively placing the N fermions, one by one, in the lowest energy states which are unoccupied. The maximum momentum reached in this way will depend on the number of fermions and corresponds to the Fermi momentum, a quantity that is directly related to the chemical potential by virtue of Eq. (1.17).

The ground state $|\Phi_0\rangle$ behaves differently than the state with no fermions $|0\rangle$. As an example, the ground state is in the kernel of the creation operators with momentum below the Fermi momentum, that is,

$$a^+_{\boldsymbol{p}\,B} |\Phi_0\rangle = 0, \qquad\qquad |\boldsymbol{p}| \leq p_{\mathrm{F}}. \tag{1.18}$$

On the other hand, this state is not annihilated by the corresponding annihilation operators,

$$a_{\boldsymbol{p}\,B} |\Phi_0\rangle \neq 0, \qquad\qquad |\boldsymbol{p}| \leq p_{\mathrm{F}}. \tag{1.19}$$

One may consider the following (canonical) transformation preserving the anticommutation rules [64]:

$$a_{\boldsymbol{p}\,B} = \begin{cases} c_{\boldsymbol{p}\,B}, & |\boldsymbol{p}| > p_{\mathrm{F}} \\ d^+_{-\boldsymbol{p}\,B}, & |\boldsymbol{p}| \leq p_{\mathrm{F}} \end{cases} \tag{1.20}$$

Now the ground state $|\Phi_0\rangle$ is annihilated by all the annihilation operators $c_{\boldsymbol{p}\,B}$ and $d_{\boldsymbol{p}\,B}$, that is,

$$c_{\boldsymbol{p}\,B} |\Phi_0\rangle = d_{\boldsymbol{p}\,B} |\Phi_0\rangle = 0. \tag{1.21}$$

The excitations corresponding to $c^+_{\boldsymbol{p}\,B}$ are called particles, while those corresponding to $d_{\boldsymbol{p}\,B}$ are denominated holes. All these excitations have a positive energy associated: the Hamiltonian operator (1.16) becomes

$$H - \mu N = \sum_{\boldsymbol{p}\,B,\, |\boldsymbol{p}| > p_{\mathrm{F}}} \left(\frac{\boldsymbol{p}^2}{2m} - \mu \right) c^+_{\boldsymbol{p}\,B} c_{\boldsymbol{p}\,B} + \sum_{\boldsymbol{p}\,B,\, |\boldsymbol{p}| \leq p_{\mathrm{F}}} \left(\mu - \frac{\boldsymbol{p}^2}{2m} \right) d^+_{\boldsymbol{p}\,B} d_{\boldsymbol{p}\,B} + \sum_{\boldsymbol{p}\,B,\, |\boldsymbol{p}| \leq p_{\mathrm{F}}} \left(\frac{\boldsymbol{p}^2}{2m} - \mu \right). \tag{1.22}$$

The action of $c^{+}_{\boldsymbol{p}B}$ adds a fermion with momentum greater than the Fermi momentum; the action of $d^{+}_{\boldsymbol{p}B}$ removes a fermion with momentum below the Fermi momentum. If the number of fermions is fixed, particles and holes can only be created in pairs.

The form of the energy operator shows that the excitations with the lowest energies are all associated with the region in momentum space that is close to the so-called Fermi surface, defined as the sphere $|\boldsymbol{p}| = p_{\rm F}$. This observation is in turn very important for the understanding of the properties of the ideal Fermi gas. Most importantly, it can be shown that the Fermi surface is a robust topological feature that survives the introduction of interactions, so that it is preserved in the realm of non-ideal Fermi gases and even Fermi liquids [26]. This is the first example we encounter of the emergence of robust properties for complex systems: the infrared physics is dominated by the existence of the Fermi surface, and thus it is essentially the same for all the systems that fall within this universality class [18, 26].

1.3.2 Microscopic (He–3)-like systems

Let us now consider a quantum liquid composed of a large collectivity of spin-1/2 atoms. From now on we use the word "atom" to stress that these objects need not be elementary objects (they need not be precisely He–3 atoms either); we do not use the word "fermion" that we reserve for the low-energy relativistic fermionic quasiparticles. We require the interactions between these atoms to be short-range but otherwise they can be very complicated, including higher-than-two-body effects. We also require the two-body interactions to be characterized by a potential of Lennard-Jones type (which is rotationally invariant) plus possibly some interaction term involving the spins. Interactions in He–3 display indeed these characteristics.

To solve a system of this sort in full detail is beyond human capacities. We need simpler theories which serve as approximate models of the exact microscopic theory. A first step in this direction is provided by Landau's Fermi-liquid theory. This theory starts from the exact description of the ideal Fermi gas, in which the notions of Fermi surface and particle/hole excitations appear. Landau's hypothesis is that generically, or at least under certain conditions of temperature and pressure, the N-particle ground state and the spectrum of the low-energy excitations (i.e., the spectrum in the surroundings of the Fermi surface) of the above strongly interacting theory is in adiabatic one-to-one correspondence with that of the free theory [59]. Under this hypothesis we can use the same labels for these states. There exists some microscopic justification of Landau's hypothesis [57], although its main backup comes from the experimental study of quantum liquids, such as He–3 in its normal phase or non-superconducting metals at low temperatures [59].

Following this hypothesis, to describe all the features of the physics associated with low energies (again, ground state and excitations close to it), one can then substitute the precise strongly interacting theory by an equivalent weakly interacting theory of quasiatoms. The prefix "quasi" is used to remark that these excitations no longer correspond to the original atoms, but should be understood as collective excitations, also known generically as quasiparticles [57]. For instance, in second quantization language and in a momentum representation, one can write Landau's grand canonical Hamiltonian as [59]

$$H_{\rm L} - \mu N := \sum_{\boldsymbol{p}\,B} \left(\frac{p^2}{2m^*} - \mu\right) a^\dagger_{\boldsymbol{p}\,B} a_{\boldsymbol{p}\,B} + \frac{1}{2} \sum_{\boldsymbol{p}\boldsymbol{p}'\,B\,C} f(\boldsymbol{p}, \boldsymbol{p}', B, C) a^\dagger_{\boldsymbol{p}\,B} a_{\boldsymbol{p}\,B} a^\dagger_{\boldsymbol{p}'C} a_{\boldsymbol{p}'C}. \tag{1.23}$$

Here $a^\dagger_{\boldsymbol{p}\,B}$, $a_{\boldsymbol{p}\,C}$ are respectively creation and annihilation operators of quasiatoms, with $B, C = \uparrow, \downarrow$ representing the spin degree of freedom. The chemical potential is $\mu = p^2_{\rm F}/(2m^*)$ with $p_{\rm F}$ the Fermi momentum and m^* the effective mass of the quasiatoms (this mass does not need to coincide with the mass of the initial atoms; for instance, in He–3 it is a few times smaller). The function $f(\boldsymbol{p}, \boldsymbol{p}', B, C)$ must be symmetric under the exchange $(\boldsymbol{p}, B) \leftrightarrow (\boldsymbol{p}', C)$. Both m^* and f are in principle phenomenological quantities that depend on details of the microscopic interaction. This model Hamiltonian has proved to be very successful, for example for the description of the normal phase of He–3, in the temperature range between 1K and 0.03K.

Again, this hypothesis implies that there exist many systems that are different in the details of their interactions but are, however, indistinguishable from a low-energy point of view. Therefore, when working out a theory of this sort, one is really working with an entire class of theories with the same low-energy behavior. The same operators $a_{\boldsymbol{p}\,B}$ can represent different physical quasiatoms in different strongly-interacting spin-fluid systems. These operators may also represent the proper atoms of a weakly-interacting spin-gas system. In the following, we will analyze the properties of a specific weakly interacting theory, independently of any specific physical realization one could have in mind. Thus, we will speak only of atoms, having always in mind that they could be equivalently quasiatoms.

1.3.3 A weakly interacting gas

Let us henceforth focus on a weakly interacting theory of spin-1/2 atoms. In this framework, one can go one step further than Landau's Fermi-liquid theory and analyze a more general interaction term.

Let us introduce the atom field Ψ, in terms of which the Hamiltonian operator for the system of spin-1/2 atoms with two-body interactions reads

$$\begin{aligned} H_{\rm I} - \mu N := &\sum_{B} \int d^3x\, \Psi^\dagger_B(\boldsymbol{x}) \left(-\frac{\hbar^2}{2m^*}\nabla^2 - \mu\right) \Psi_B(\boldsymbol{x}) \\ &+ \frac{1}{2} \sum_{B\,C} \int d^3x\, d^3x'\, V(\boldsymbol{x} - \boldsymbol{x}') \Psi^\dagger_C(\boldsymbol{x}) \Psi^\dagger_B(\boldsymbol{x}') \Psi_B(\boldsymbol{x}') \Psi_C(\boldsymbol{x}). \end{aligned} \tag{1.24}$$

We have assumed for the time being that the interaction potential does not depend on the spin. We can always come back to the momentum representation, $\Psi_B = \mathscr{V}^{-1/2} \sum_{\boldsymbol{p}} a_{\boldsymbol{p}\,B} e^{i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar}$

with $\mathscr{V}$ the volume of the system, to write the Hamiltonian as

$$H_{\rm I} - \mu N := \sum_{\boldsymbol{p}\,B} \left(\frac{p^2}{2m^*} - \mu \right) a^\dagger_{\boldsymbol{p}\,B} a_{\boldsymbol{p}\,B} + \frac{1}{2} \sum_{\boldsymbol{p}_1\,\boldsymbol{p}_2\,\boldsymbol{p}_3\,\boldsymbol{p}_4\,B\,C} \delta_{\boldsymbol{p}_1+\boldsymbol{p}_2,\boldsymbol{p}_3+\boldsymbol{p}_4} \tilde{V}\left(\frac{\boldsymbol{p}_1 - \boldsymbol{p}_2 + \boldsymbol{p}_3 - \boldsymbol{p}_4}{2} \right) a^\dagger_{\boldsymbol{p}_4\,C}\, a^\dagger_{\boldsymbol{p}_3\,B}\, a_{\boldsymbol{p}_2\,B}\, a_{\boldsymbol{p}_1\,C}, \tag{1.25}$$

with

$$\tilde{V}(\boldsymbol{p}) := \frac{1}{8\mathscr{V}} \int d^3 r e^{i\boldsymbol{p}\cdot\boldsymbol{r}/\hbar} V(\boldsymbol{r}), \qquad V(-\boldsymbol{r}) = V(\boldsymbol{r}). \tag{1.26}$$

Note that $\tilde{V}(\boldsymbol{p})$ has dimensions of energy, and $\boldsymbol{r} := \boldsymbol{x} - \boldsymbol{x}'$. Our notation in what follows assumes a finite box with volume $\mathscr{V}$ and so a discrete sum in momentum space; the infinite-volume limit could be taken if desired. Notice that the potential term in (1.25) is invariant under a Galilean boost transformation of the reference frame. We should keep in mind this property, which can apparently be lost under certain approximations that will be made in the following.

The interaction term contains different interaction channels: the Hartree channel [which contains the previous Landau terms in Eq. (1.23)], the Fock channel, and the pairing channel [59]. Of special relevance in what follows is the pairing channel that appears for interactions satisfying $\boldsymbol{p}_1 = -\boldsymbol{p}_2 =: \boldsymbol{p}$ and $\boldsymbol{p}_3 = -\boldsymbol{p}_4 =: \boldsymbol{p}'$. The pairing terms control the form of the ground state of the theory; see Leggett's discussion in [59], and [65] for additional justification of this feature from the perspective of effective field theory. The pairing Hamiltonian reads

$$H_{\rm P} - \mu N := \sum_{\boldsymbol{p}\,B} \left(\frac{p^2}{2m^*} - \mu \right) a^\dagger_{\boldsymbol{p}\,B} a_{\boldsymbol{p}\,B} + \frac{1}{2} \sum_{\boldsymbol{p}\,\boldsymbol{p}'\,B\,C} \tilde{V}(\boldsymbol{p}' + \boldsymbol{p}) a^\dagger_{-\boldsymbol{p}'\,C}\, a^\dagger_{\boldsymbol{p}'\,B}\, a_{-\boldsymbol{p}\,B}\, a_{\boldsymbol{p}\,C}. \tag{1.27}$$

If the potential does not depend on the orientation, it can only depend on $|\boldsymbol{p}' + \boldsymbol{p}| = p^2 + p'^2 + 2pp'\hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{p}}'$. Then, we can always write it [57] as an expansion of the form

$$\tilde{V}(|\boldsymbol{p} + \boldsymbol{p}'|) = \sum_{l=0}^{\infty} (2l+1) \tilde{V}_l(p, p') P_l(\hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{p}}'), \tag{1.28}$$

where P_l represents Legendre polynomials. The converse assertion is not true: not all expansions can be put in exact correspondence with $V(r)$ potentials. As we are always interested in the surroundings of the Fermi surface, where the low-energy excitations reside, we can take the potential to depend only on the angle $\hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{p}}'$ and not in the norms which will be $p, p' \simeq p_{\rm F}$.

Take now a microscopic interaction such that $|\tilde{V}_1| \gg |\tilde{V}_{l\neq 1}|$. Then $g := -\tilde{V}_1(p_{\rm F}, p'_{\rm F})$ will be a positive constant because of the binding character of the potential. The potential will be then written as

$$\tilde{V} \simeq -g \hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{p}}' \simeq -\frac{g}{p_{\rm F}^2} \boldsymbol{p} \cdot \boldsymbol{p}'. \tag{1.29}$$

The simplest interaction of this kind is the one provided by

$$V(\boldsymbol{r}) = 8g\mathscr{V}\left[\frac{\hbar^2}{2p_{\rm F}^2}\boldsymbol{\nabla}^2\delta(\boldsymbol{r}) + \delta(\boldsymbol{r})\right]. \tag{1.30}$$

This interaction has $\tilde{V}_{l\geq 2} = 0$. Near the Fermi surface, the remaining components verify $|\tilde{V}_1| \gg |\tilde{V}_0|$ so that the potential approximately behaves as Eq. (1.29). Indeed, using Eq. (1.26) one has, close to the Fermi surface,

$$\tilde{V}_0(\boldsymbol{p}+\boldsymbol{p}') = g\left(1 - \frac{p^2 + p'^2}{2p_{\rm F}^2}\right) \simeq 0, \qquad \tilde{V}_1(\boldsymbol{p}+\boldsymbol{p}') = -\frac{g}{p_{\rm F}^2}\boldsymbol{p}\cdot\boldsymbol{p}' \simeq -g\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{p}}'. \tag{1.31}$$

This interaction is the distributional limit of potentials of the form shown in Fig. 1.4. These potentials exhibit a repulsive hard core and an attractive tail (precisely the type of interaction between He–3 atoms). It is not possible to construct a translation-invariant interaction potential with only $\tilde{V}_1 \neq 0$, as it would fail to be invariant under constant shifts in momentum space, so that Eq. (1.30) is indeed the best approximation to an interaction of the form (1.29) one can find.

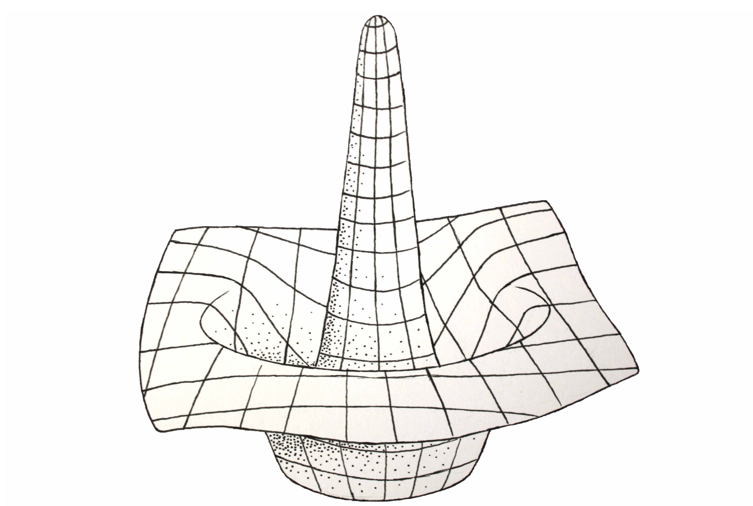

Figure 1.4: Diagram showing the qualitative form of the interaction potential $V(\boldsymbol{r})$.

For properties involving long wavelengths compared with the interparticle distance, the model potential (1.30) will be perfectly appropriate as representative of an entire microscopic class. Taking this potential, the grand canonical pairing Hamiltonian finally reads

$$H_{\rm P} - \mu N = \sum_{\boldsymbol{p}B}\left(\frac{p^2}{2m^*} - \mu\right)a^\dagger_{\boldsymbol{p}B}a_{\boldsymbol{p}B} + \frac{g}{2p_{\rm F}^2}\sum_{\boldsymbol{p}\boldsymbol{p}'BC}(\boldsymbol{p}'\cdot\boldsymbol{p})a^\dagger_{-\boldsymbol{p}'C}a^\dagger_{\boldsymbol{p}'B}a_{\boldsymbol{p}B}a_{-\boldsymbol{p}C}. \tag{1.32}$$

Note that we have reversed the sign of $\boldsymbol{p}$ in the interaction term for later convenience. This is the system we will work with in the next subsections. It is instructive to keep in mind the symmetry group of this pairing Hamiltonian, given by

$$G := \mathrm{SO}(3)_{\mathbf{L}} \times \mathrm{SO}(3)_{\mathbf{S}} \times \mathrm{U}(1)_N, \tag{1.33}$$

where the notation for each symmetry group is the standard one and each subscript denotes the corresponding conserved quantity: the angular momentum $\mathbf{L}$, the spin $\boldsymbol{S}$ and the particle number N. Our previous discussion in this section was intended to motivate the construction of the pairing Hamiltonian (1.32), but the reader could equivalently take the latter as the bona fide starting point that sets the basis for the rest of the chapter. Although we are interested mainly in the formal properties of the system, it is interesting to keep in mind that the very same model provides a satisfactory description of the experimental properties of superfluid He–3.

1.3.4 Condensation and order parameters

The interaction term described in the previous subsection is called a p-wave spin-triplet pairing interaction. Below a critical temperature T_{C} it enforces the formation of anisotropic Cooper pairs, as opposed to the isotropy of the Cooper pairs in standard superconductivity [66, 67, 68]. The spatial anisotropy of these pairs is associated with the fact that they possess angular momentum. Given the antisymmetric structure of the orbital part of the wave function, its spin structure has to be symmetric and thus belongs to the triplet space of the spin product. These pairs condense acquiring a macroscopic occupation. The macroscopic wave function or order parameter associated with the condensed pairs can be described as

$$\boldsymbol{\Psi}_{BC} := \frac{g}{p_{\mathrm{F}}} \left\langle \sum_{\boldsymbol{p}} \boldsymbol{p}\, a_{\boldsymbol{p}\,B} a_{-\boldsymbol{p}\,C} \right\rangle. \tag{1.34}$$

As a consequence of the spin dependence and of the dominance of the anisotropic p-wave interaction, this order parameter is not a scalar, as in the case of classical superconductivity or Bose-Einstein condensation, but a matrix with spin indices $B, C = \uparrow, \downarrow$. There is also an implicit orbital index i because of the $\boldsymbol{p}$-dependence of $\boldsymbol{\Psi}_{BC}$. The normal-liquid phase has as symmetry group (1.33), that is, independent rotations of the coordinate and spin spaces plus a phase-invariance symmetry associated with the conservation of the number of atoms. Pair condensation amounts to the spontaneous (partial) breaking of this symmetry. The order parameters appearing in this p-wave spin-triplet condensation are symmetric in the spin indices and therefore can always be written as:

$$\Psi^i_{BC} = i(\sigma_P \sigma_2)_{BC}\, d^{Pi}, \tag{1.35}$$

where d^{Pi} is in general a complex vector in both spin and position space, $P = 1, 2, 3$ and σ_P are the Pauli matrices:

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.36}$$

The set $\{\sigma_P \sigma_2\}_{P=1,2,3}$ forms a basis of the space of 2×2 symmetric matrices, if allowing for complex coefficients d^{Pi}.

Depending on the details of the interaction, the order parameter can acquire different structures. The precise form of the order parameter is obtained by a minimization principle. In the microscopic theory the quantity to be minimized is the expectation value of the Hamiltonian (1.32) within a suitable family of states. One can alternatively use a minimization within Ginzburg-Landau theory, which is a special case of the phenomenological Landau-Lifshitz theory of second-order phase transitions. This is the way we shall proceed.

1.3.5 Ginzburg-Landau minimization

The validity of Ginzburg-Landau theory [69] is restricted to temperatures near the critical transition temperature $T_{\rm C}$ to the superfluid phase. However, it is much easier to handle the calculations within this restricted setup, and then generalize them to the whole range of temperature by using the microscopic theory. In either case the structure of the order parameter (1.34) is obtained by a minimization principle. In Ginzburg-Landau theory, and for constant temperature T and volume V, the Helmholtz free energy functional of the order parameter, constructed as follows, is the quantity to be minimized.

The main observation is that the order parameter as defined in Eq. (1.34) is zero above a certain critical temperature $T_{\rm C}$ but takes a finite value for $T < T_{\rm C}$. If we suppose that, near $T_{\rm C}$, the free energy is analytic in the order parameter and obeys the symmetries of the microscopic Hamiltonian, then one can write a Taylor expansion near the critical temperature. To have a non-trivial minimization problem of this free energy one just needs to take into account the first two non-zero orders. In our case, as the free energy must be invariant under rotations in both coordinate and spin spaces, these terms will be of second-order and fourth-order. Given the order parameter (1.35), there is one possible second-order term and five fourth-order terms:

$$
\begin{aligned}
I_0 &:= \sum_{Pi} d^{Pi} d^{*Pi}, \\
I_1 &:= \sum_{Pi} \sum_{Qj} d^{Pi} d^{Pi} d^{*Qj} d^{*Qj}, \\
I_2 &:= \sum_{Pi} \sum_{Qj} d^{Pi} d^{Qj} d^{*Pi} d^{*Qj}, \\
I_3 &:= \sum_{Pi} \sum_{Qj} d^{Pi} d^{Pj} d^{*Qi} d^{*Qj}, \\
I_4 &:= \sum_{Pi} \sum_{Qj} d^{Pi} d^{Qj} d^{*Pj} d^{*Qi}, \\
I_5 &:= \sum_{Pi} \sum_{Qj} d^{Pi} d^{Qi} d^{*Pj} d^{*Qj}.
\end{aligned}
\tag{1.37}
$$

In these equations, $Q = 1, 2, 3$ also. The general form of the free energy is then

$$F = F_{\mathrm{n}} + \alpha(T) I_0 + \frac{1}{2}\beta(T)\sum_{s=1}^{5}\beta_s I_s. \tag{1.38}$$

Terms proportional to gradients of the order parameter are neglected because the variations of the order parameter are considered to be smooth enough. These additional terms will be of relevance when discussing inhomogeneities in Sec. 1.5. F_{n} is the free energy of the normal phase, which is independent of the order parameter and so an irrelevant constant for our purposes. The order parameter should be zero above the transition temperature and take a uniform nonzero value for $T < T_{\mathrm{C}}$. This implies that, to the lowest order around the critical temperature, $\alpha(T) \simeq \alpha_0(T - T_{\mathrm{C}})$ and $\beta(T) \simeq \beta(T_{\mathrm{C}}) > 0$ (see Sec. 5.7 in [59] for a detailed discussion). The precise values of the coefficients in Ginzburg-Landau theory depend on the microscopic theory and in principle can be derived from it, for example as in the BCS theory of superconductivity [18, 57, 70].

As it stands, this minimization problem is not analytically solvable. For this reason, the so-called unitarity condition,

$$\sum_{QR} \epsilon_{PQR} d^{*Qi} d^{Rj} = 0, \tag{1.39}$$

is imposed ($P, Q, R = 1, 2, 3$). Although there is no theoretical argument to impose this condition (apart from simplicity of certain expressions), there is some experimental justification since it seems that the states of He–3 that are realized in nature are all unitary in this sense. Consideration of non-unitary states could of course be interesting for other purposes. In our case, although the state we are most interested in has not been observed yet in nature, it is nevertheless unitary.

Under the unitarity condition, four solutions are found to the minimization problem. The so-called BW (Balian-Werthamer) and ABM (Anderson-Brinkman-Morel) states are associated through the confrontation with experiments with the superfluid phases B and A of He–3, respectively. The other two states are the so-called planar and polar states. The planar state and the ABM state are topologically characterized by the presence of Fermi points. The BW state is fully gapped, while the polar state has a Fermi manifold of dimension 1. As we will discuss, Fermi points give rise to relativistic low-energy excitations.

For our purposes, we are specially interested in the planar state; we will justify why is this the case in the next sections. Its order parameter is

$$d^{Pi}_{\mathrm{planar}}(T) := \Delta(T)(\hat{s}^P \hat{m}^i + \hat{s}'^P \hat{n}^i), \tag{1.40}$$

where $\hat{\boldsymbol{m}}$, $\hat{\boldsymbol{n}}$ are unit vectors in position space and $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{s}}'$ unit vectors in spin space subject to the orthogonality conditions $\hat{\boldsymbol{m}} \cdot \hat{\boldsymbol{n}} = 0$, $\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}' = 0$. In this expression, the scalar function $\Delta(T)$ is the gap parameter which contains the dependence of the order parameters on the temperature T and the coupling constant g. At zero temperature its value is approximately $\Delta_0 := \Delta(0) \simeq k_{\mathrm{B}} T_{\mathrm{C}}$, where T_{C} is the critical temperature. In this state the continuous symmetries of the system are reduced from the group G in Eq. (1.33) down to $G_{\rm planar} = {\rm U}(1)_{L_z} \times {\rm U}(1)_{S_z}$.

The planar state has not yet been observed in nature among the superfluid phases of He–3. If one neglects the dipole-dipole interactions this state is never the lowest-energy state of the system. However, when taking into account these interactions, which in He–3 are rather feeble, this state should be the global minimum in a narrow temperature band in phase space (see for example [71]). Here we are not considering real He–3 but a system constructed with atoms adapted to our needs. Thus we will just assume that there exist some additional atom-atom interactions beyond the Lennard-Jones potentials such that they select the planar state as the natural ground state.

It is sometimes instructive to have in mind the other well-known states of this system: the ABM and BW states. Their order parameters are respectively

$$d^{Pi}_{\rm ABM}(T) := \Delta(T)\hat{s}^P(\hat{m}^i + i\hat{n}^i), \tag{1.41}$$

$$d^{Pi}_{\rm BW}(T) := \Delta(T)\delta^{Pi}. \tag{1.42}$$

Here there is also an orthogonality condition $\hat{\boldsymbol{m}} \cdot \hat{\boldsymbol{n}} = 0$.

Before closing this subsection let us comment that, within the interpretation of a strongly interacting system of atoms, the realization of any of these condensed phases takes us beyond the strict limits of applicability of Landau's Fermi-liquid hypothesis. The Fermi surface of the free system has been deformed so strongly that it no longer survives. It has been either completely eliminated (as in the BW state) or reduced to just some points (in the planar and ABM states). However, it is remarkable that a weakly interacting model of quasiatoms is able to describe correctly the condensation and low-energy excitations of these systems. For the interpretation in which one directly starts from a weakly interacting system of atoms, the previous comment is irrelevant: in this case the weakly interacting theory is already the very microscopic theory.

1.3.6 Fermi points

The phenomenon of condensation, being a substantial rearrangement of the ground state of the fermion system, triggers drastic changes in its physical properties, transferring it to a different universality class. A convenient way to notice these differences is studying the behavior of quasiparticle excitations on top of a given order parameter describing the properties of the condensed part of the system. It will be enough for our purposes to consider the so-called Gor'kov factorization [72]: once the system has settled to a condensed state, the pairing interaction can be expanded up to quadratic order in perturbations around the condensed state. The resulting quadratic Hamiltonian reads

$$\begin{aligned} H_{\rm P} - \mu N :=& \sum_{\boldsymbol{p}B} M(\boldsymbol{p}) a^\dagger_{\boldsymbol{p}B} a_{\boldsymbol{p}B} \\ &+ \frac{1}{2p_{\rm F}} \sum_{\boldsymbol{p}BC} \boldsymbol{p}\cdot\boldsymbol{\Psi}_{BC} a^\dagger_{-\boldsymbol{p}C} a^\dagger_{\boldsymbol{p}B} + \frac{1}{2p_{\rm F}} \sum_{\boldsymbol{p}BC} \boldsymbol{p}\cdot\boldsymbol{\Psi}^*_{CB} a_{\boldsymbol{p}B} a_{-\boldsymbol{p}C}, \end{aligned} \tag{1.43}$$

where we have defined $M(\boldsymbol{p}) = p^2/(2m^*) - \mu$. Consider now the order parameter to be a homogeneous planar state characterized by the vectors $\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}', \hat{\boldsymbol{m}}, \hat{\boldsymbol{n}}$. Let us choose a system of coordinates adapted to the pairs-spin-space Cartesian trihedral

$$\hat{\boldsymbol{x}} = \hat{\boldsymbol{s}}, \qquad \hat{\boldsymbol{y}} = \hat{\boldsymbol{s}}', \qquad \hat{\boldsymbol{z}} = \hat{\boldsymbol{s}} \times \hat{\boldsymbol{s}}'. \tag{1.44}$$

The evaluation of the commutator between quasiparticle operators and $H_{\rm P} - \mu N$ shows that the evolution equations of quasiparticle operators particularized to the order parameter of the planar state (1.40) are, using Eqs. (1.35) and (1.43),

$$i\hbar\dot{a}_{\boldsymbol{p}\uparrow} = M(\boldsymbol{p})a_{\boldsymbol{p}\uparrow} - c_\perp \boldsymbol{p}\cdot(\hat{\boldsymbol{m}} - i\hat{\boldsymbol{n}})a^\dagger_{-\boldsymbol{p}\uparrow}, \tag{1.45}$$
$$i\hbar\dot{a}_{\boldsymbol{p}\downarrow} = M(\boldsymbol{p})a_{\boldsymbol{p}\downarrow} + c_\perp \boldsymbol{p}\cdot(\hat{\boldsymbol{m}} + i\hat{\boldsymbol{n}})a^\dagger_{-\boldsymbol{p}\downarrow}. \tag{1.46}$$

Here we have introduced the parameter

$$c_\perp := \Delta_0/p_{\rm F} \tag{1.47}$$

with dimensions of velocity. The two spin populations are decoupled in their evolution. This property permits us to deal with these two populations separately, simplifying the treatment in the following.

These equations are enough to show the replacement, in this particular condensed ground state, of the Fermi surface by a Fermi point structure. Acting with the operator $i\partial_t$ on (1.45), and using Eq. (1.45) again, one finds that the dependence on $a^\dagger_{-\boldsymbol{p}\uparrow}$ vanishes, so that one can write

$$\begin{aligned}(i\hbar\partial_t)^2 a_{\boldsymbol{p}\uparrow} &= iM(\boldsymbol{p})\dot{a}_{\boldsymbol{p}\uparrow} - ic_\perp \boldsymbol{p}\cdot(\hat{\boldsymbol{m}} - i\hat{\boldsymbol{n}})\dot{a}^\dagger_{-\boldsymbol{p}\uparrow} \\ &= \left\{M^2(\boldsymbol{p}) + c_\perp^2[(\boldsymbol{p}\cdot(\hat{\boldsymbol{m}} - i\hat{\boldsymbol{n}})(\boldsymbol{p}\cdot(\hat{\boldsymbol{m}} + i\hat{\boldsymbol{n}})]\right\} a_{\boldsymbol{p}\uparrow} \\ &= \left\{M^2(\boldsymbol{p}) + c_\perp^2(\boldsymbol{p}\times\hat{\boldsymbol{l}})^2\right\} a_{\boldsymbol{p}\uparrow},\end{aligned} \tag{1.48}$$

where $\hat{\boldsymbol{l}} := \hat{\boldsymbol{m}} \times \hat{\boldsymbol{n}}$. Being this action diagonal permits to extract the dispersion relation of quasiparticles directly as

$$E^2(\boldsymbol{p}) = M^2(\boldsymbol{p}) + c_\perp^2(\boldsymbol{p}\times\hat{\boldsymbol{l}})^2. \tag{1.49}$$

The eigenvalues of the evolution operator, given by the two square roots of Eq. (1.49), vanish only in the so-called Fermi points in momentum space,[1]

$$\boldsymbol{p}_{{\rm F},\pm} = \pm p_{\rm F}\hat{\boldsymbol{l}}. \tag{1.50}$$

This is in stark difference with the Fermi surface structure, in which the dispersion relation displays an entire sphere of zeros in momentum space. This implies that the low-energy

[1] Also often called Weyl points. We use the term "Fermi point" in accordance with [18]. "Fermi point" can be understood as the generic term for topological point nodes, which includes Weyl points when the underlying manifold is 3+1 dimensional, and Dirac points for 2+1 dimensions. The A phase of He–3 is then an example of the Weyl category of Fermi points.

excitations will be now concentrated around these two points in momentum space. Let us recall that $p_{\rm F} = \sqrt{2m^*\mu}$.

As a first glimpse at these properties, we can now see that the dispersion relation is relativistic (that is, linear) near these points in momentum space. It is easy to understand why this is the case: near these points, the dispersion relation of quasiparticles is linear to leading order, and is three-dimensional, unlike in the case of Fermi manifolds of higher dimension [73]. Indeed, considering a plane wave with momentum $\boldsymbol{p} = Zp_{\rm F}\hat{\boldsymbol{l}} + \mathfrak{p}$ with $Z = \pm 1$ and $\mathfrak{p}$ a small deviation with respect to the corresponding Fermi point, we obtain the frequency

$$(\hbar\omega)^2 = c_\parallel^2 \mathfrak{p}_{\rm l}^2 + c_\perp^2(\mathfrak{p}_{\rm m}^2 + \mathfrak{p}_{\rm n}^2), \tag{1.51}$$

where

$$c_\parallel = p_{\rm F}/m^*. \tag{1.52}$$

Recall that the set $\{\rm m, n, l\}$ are not regular subscripts (this is the reason for using the standard text font for them), but rather denote the projections on the pairs Cartesian trihedral $\hat{\boldsymbol{m}}, \hat{\boldsymbol{n}}, \hat{\boldsymbol{l}}$; for instance, $\mathfrak{p}_{\rm m} = \mathfrak{p}\cdot\hat{\boldsymbol{m}}$. The parameters $c_\parallel$ and $c_\perp$ correspond to the propagation velocity of low-energy quasiparticles in the directions parallel and perpendicular to the anisotropy axis (or just $\hat{\boldsymbol{l}}$), respectively.

This linearization is only valid for a finite range of momenta around the Fermi points. Given the anisotropic character of the system, it is useful to study independently the excitations that are parallel and orthogonal to the anisotropy axis $\hat{\boldsymbol{l}}$. For parallel quasiparticles, the linearization is valid for momenta much lower than

$$p_\parallel^\star := p_{\rm F}. \tag{1.53}$$

On the other hand, for quasiparticles traveling in the orthogonal plane to $\hat{\boldsymbol{l}}$, the corresponding quantity is given by

$$p_\perp^\star := mc_\perp = p_{\rm F}\frac{c_\perp}{c_\parallel}. \tag{1.54}$$

For general quasiparticles, the momentum linearization scale will be given by a smooth interpolation between these two values depending on the angle θ between the trajectory and the anisotropy axis $\hat{\boldsymbol{l}}$:

$$(p^\star)^2(\theta) := (p_\parallel^\star)^2\cos^2\theta + (p_\perp^\star)^2\sin^2\theta. \tag{1.55}$$

It is the first time we encounter the ratio $c_\perp/c_\parallel$, namely in Eq. (1.54). The value of this dimensionless factor depends on the specific parameters of the model. For superfluid He–3 on its A phase, it takes roughly the value 10^{-5} [18]. For our purposes it will be enough to keep in mind that, for the class of systems we want to consider,

$$\frac{c_\perp}{c_\parallel} \ll 1. \tag{1.56}$$

This quantity is essential in our discussion, in that it sets the hierarchy of energy and momentum scales in the system. For example, we can see that $p_\perp^\star$ in Eq. (1.54) is much lower than $p^\star_\parallel$ in Eq. (1.53), so that $p^\star(\theta)$ in Eq. (1.55) verifies

$$p^\star_\perp \leq p^\star(\theta) \leq p^\star_\parallel. \tag{1.57}$$

The lowest energy scale for which deviations with respect to the linear behavior will appear, corresponding to the energy of orthogonal quasiparticles with momentum (1.54), is then given by

$$E_{\rm L} := m^* c_\perp^2. \tag{1.58}$$

Notice, however, that in general the energy scale of violation of Lorentz invariance is anisotropic as well. In the following section we shall study in detail the emergent properties of the quasiparticles that fall within the relativistic part of the spectrum.

1.4 Fermion fields

We have seen that, when the system settles down on a condensed ground state with the structure of the planar state, it develops a Fermi point structure. The dispersion relation of quasiparticles is then relativistic near these points in momentum space. Given the fermionic nature of the quasiparticles of the system, it is expected that the Dirac equation should emerge as the relevant dynamical equation in this regime. Let us describe in detail this feature and some of its implications.

1.4.1 The emergent Dirac equation

Let us first analyze the dynamical equations that describe the evolution of quasiparticles near the Fermi points on the planar ground state. These are new types of quasiparticles, specific combinations of the atoms of Landau's theory. We will eventually call them effective fermions, or just fermions.

It will be useful to introduce a new set of creation and annihilation operators, with elements $a_{\boldsymbol{p}BZ}$, as follows. As labels for these operators we use the deviation $\mathfrak{p}$ with respect to any of the Fermi points, $\boldsymbol{p} = \pm p_{\rm F}\hat{\boldsymbol{l}} + \mathfrak{p}$, the spin index B, and a subscript $Z = \pm 1$ (or just $\pm$ to simplify) indicating the Fermi point near which it is localized in momentum space: the $+$ Fermi point $(+p_{\rm F}\hat{\boldsymbol{l}})$ or the $-$ Fermi point $(-p_{\rm F}\hat{\boldsymbol{l}})$. The precise definition of these operators is

$$a_{\mathfrak{p}\uparrow +} := a_{p_{\rm F}\hat{\boldsymbol{l}}+\mathfrak{p}\uparrow}, \qquad a_{\mathfrak{p}\uparrow -} := a_{-p_{\rm F}\hat{\boldsymbol{l}}+\mathfrak{p}\uparrow}. \tag{1.59}$$

Focusing first on the $\uparrow$ spin projection, we can write the corresponding equations of motion in terms of the new operators as

$$\begin{aligned} i\hbar\dot{a}_{\mathfrak{p}\uparrow +} &= c_\parallel \mathfrak{p}_{\rm l} a_{\mathfrak{p}\uparrow +} - c_\perp(\mathfrak{p}_{\rm m} - i\mathfrak{p}_{\rm n}) a^\dagger_{-\mathfrak{p}\uparrow -}, \\ i\hbar\dot{a}^\dagger_{-\mathfrak{p}\uparrow -} &= -c_\parallel \mathfrak{p}_{\rm l} a^\dagger_{-\mathfrak{p}\uparrow -} - c_\perp(\mathfrak{p}_{\rm m} + i\mathfrak{p}_{\rm n}) a_{\mathfrak{p}\uparrow +}. \end{aligned} \tag{1.60}$$

In these equations, the equality sign would be strictly speaking an approximately-equal sign, as these relations are not satisfied for momenta too far from the Fermi point. The two previous equations can be written in a compact manner as

$$i\hbar\partial_t\chi_{\mathfrak{p}\uparrow} = \mathscr{H}_{\mathfrak{p}\uparrow}\chi_{\mathfrak{p}\uparrow}, \qquad \chi_{\mathfrak{p}\uparrow} = \begin{pmatrix} a_{\mathfrak{p}\uparrow+} \\ a^\dagger_{-\mathfrak{p}\uparrow-} \end{pmatrix}, \tag{1.61}$$

with

$$\mathscr{H}_{\mathfrak{p}\uparrow} := c_{\|}\mathfrak{p}_{\rm l}\sigma_3 - c_\perp\mathfrak{p}_{\rm m}\sigma_1 - c_\perp\mathfrak{p}_{\rm n}\sigma_2. \tag{1.62}$$

This is a linear spinor equation for a Weyl spinor (usually dubbed Nambu-Gor'kov spinor [18]) with helicity +1. We have defined the helicity as the product of the three factors ±1 that appear in front of the Pauli matrices. Indeed, the two Weyl equations for different helicities can be distinguished by evaluating this quantity [61]. When the full Dirac equation is constructed, this notion will be smoothly matched with the usual definition of helicity.

Before continuing let us notice that the evolution equations for all the $a_{\boldsymbol{p}\uparrow}$ in Eq. (1.45) are not linear in the complex plane due to the presence of complex conjugate terms. However, they have a different quasilinear symmetry. The system is invariant if one multiplies the $a_{\boldsymbol{p}\uparrow}$ with $\boldsymbol{p}$ in the $Z=+1$ hemisphere by a complex constant c and those in the $Z=-1$ hemisphere by its complex conjugate c^*. This symmetry has allowed us to write a linear equation for the previous spinor $\chi_{\mathfrak{p}\uparrow}$. This spinor contains information about both Fermi points.

The same arguments can be applied to the ↓ projection of the real (atomic) spin of Landau's quasiparticles to obtain

$$\chi_{\mathfrak{p}\downarrow} := \begin{pmatrix} a_{\mathfrak{p}\downarrow+} \\ a^\dagger_{-\mathfrak{p}\downarrow-} \end{pmatrix}, \tag{1.63}$$

and Hamiltonian operator

$$\mathscr{H}_{\mathfrak{p}\downarrow} := c_{\|}\mathfrak{p}_{\rm l}\sigma_3 + c_\perp\mathfrak{p}_{\rm m}\sigma_1 - c_\perp\mathfrak{p}_{\rm n}\sigma_2, \tag{1.64}$$

with helicity −1 in this case. Notice that the only difference between the two spin projections is a sign accompanying $\hat{\boldsymbol{m}}$ in the order parameter [see also Eqs. (1.45,1.46)], which translates into a different helicity in the low-energy corner. For this reason the atomic spin projection index can be thought of as a helicity index for the effective fermions. We will explicitly check this later.

As a final step one can arrange the two spinors to form a bispinor that obeys the following evolution equation:

$$i\hbar\partial_t\begin{pmatrix} \chi_{\mathfrak{p}\uparrow} \\ \chi_{\mathfrak{p}\downarrow} \end{pmatrix} = e^i{}_j Y^j\mathfrak{p}_i\begin{pmatrix} \chi_{\mathfrak{p}\uparrow} \\ \chi_{\mathfrak{p}\downarrow} \end{pmatrix}, \tag{1.65}$$

with

$$Y^1 = \begin{pmatrix} -\sigma_1 & 0 \\ 0 & \sigma_1 \end{pmatrix}, \quad Y^2 = \begin{pmatrix} -\sigma_2 & 0 \\ 0 & -\sigma_2 \end{pmatrix}, \quad Y^3 = \begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}. \tag{1.66}$$

The only non-zero components of $e^i{}_j$ are

$$e^1{}_1 := c_\perp, \qquad e^2{}_2 := c_\perp, \qquad e^3{}_3 := c_\parallel. \tag{1.67}$$

Now we shall find a matrix X such that the set $\{X, XY^1, XY^2, XY^3\}$ is a representation of the Dirac matrices. Taking into account that the matrices (1.66) verify the properties

$$(Y^i)^2 = I_4, \qquad \{Y^i, Y^j\} = 2\delta^{ij} I_4, \tag{1.68}$$

(with I_4 the 4×4 identity), which follow directly from the properties of the Pauli matrices, such a matrix X must verify

$$X^2 = I_4, \qquad \{X, Y^i\} = 0. \tag{1.69}$$

One can check that a solution to these equations is given by

$$X = \begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}. \tag{1.70}$$

The corresponding representation of the Dirac matrices is:

$$\gamma^0 = \begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}, \qquad \gamma^1 = \begin{pmatrix} 0 & I_2 \\ -I_2 & 0 \end{pmatrix},$$
$$\gamma^2 = \begin{pmatrix} 0 & -i\sigma_3 \\ -i\sigma_3 & 0 \end{pmatrix}, \qquad \gamma^3 = \begin{pmatrix} 0 & -i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix}. \tag{1.71}$$

This closes our proof that the low-energy excitations on top of the planar phase correspond to massless Dirac spinors in Minkowski spacetime, satisfying the evolution equation

$$e^a{}_b \gamma^b \bar{\mathfrak{p}}_a \psi_{\mathfrak{p}} = 0, \qquad \psi_{\mathfrak{p}} := \begin{pmatrix} \chi_{\mathfrak{p}\uparrow} \\ \chi_{\mathfrak{p}\downarrow} \end{pmatrix}, \tag{1.72}$$

where we have taken the Fourier time transform and defined $\bar{\mathfrak{p}}^a := (\hbar\omega, \mathfrak{p})$. The components of the tetrad $e^a{}_b$ are given by Eq. (1.67), complemented by

$$e^0{}_0 := 1. \tag{1.73}$$

Spacetime is therefore Minkowskian because the effective tetrad field is constant. The constant velocities $c_\parallel$ and $c_\perp$ can be absorbed into a rescaling of the coordinates. This laboratory anisotropy would in any case be unobservable for internal observers [74] (see next section).

The occurrence of four components in the low-energy fermionic object ψ, whose Fourier components are defined in (1.72), is tied up to the existence of two degrees of freedom, one for each Fermi point, for each projection of the spin. The spin projection must be considered as the helicity eigenvalue in the low-energy description: let us evaluate the chirality operator in this representation,

$$\gamma^5 := i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}. \tag{1.74}$$

Its diagonal form implies that $\chi_\uparrow$ and $\chi_\downarrow$ have a well-defined chirality, as

$$\frac{1+\gamma^5}{2}\psi = \psi_\uparrow = \begin{pmatrix} \chi_\uparrow \\ 0 \end{pmatrix}, \qquad \frac{1-\gamma^5}{2}\psi = \psi_\downarrow = \begin{pmatrix} 0 \\ \chi_\downarrow \end{pmatrix}. \tag{1.75}$$

In this thesis we are interested in constructing an effective low-energy world that cannot be distinguished operationally from the world of electrodynamics. To reproduce the behavior of electrons and positrons, one would need to generate some small mass gap for the excitations. This might require a more complicated system, with for example some Yukawa couplings, that is beyond the analysis carried out in this thesis.

1.4.2 Internal observers

What we want to discuss here is the important meaning of internal observer. In relation with the emergence of charge, here we can already discuss the difference between two types of potential internal observers. One can be called an internal Fermi-point observer, the other an internal low-energy observer. These should be contrasted with the external, or laboratory observer which, as its name indicates, is alien to the condensed-matter-like model. While the external observer would be an outsider to our world, the internal observers can only be defined in the framework of the model, as they are made up from the very same substratum that also form particles and fields.

An internal *Fermi-point* observer is an observer living in one specific Fermi point Z, with $Z = \pm 1$. We can associate a momentum $Zp_{\rm F}\hat{\boldsymbol{l}}$ to that observer. He would see the momentum region around him as a low-energy world full of spinor waves (these will not be Weyl spinors but specific superpositions of them). His world would have half the degrees of freedom compared to the Dirac bispinors. In addition, this observer will see quasiparticles coming from the other Fermi point, which will have a tremendous relative momentum essentially given by $2Zp_{\rm F}$, although they will have low energies. To obtain a standard electrodynamic world for these kind of observers this model would lack two ingredients: i) the quasiparticles from different Fermi points should not interact with each other (in the model we are discussing they do); ii) one should duplicate in some way the number of degrees of freedom associated with that Fermi point (maybe through some fragmented condensation [75]).

An internal *low-energy* observer on the other hand is an observer who sees *all* the low-energy excitations. It is reasonable that they will use as a natural momentum label the deviations $\mathfrak{p}$. These momenta can properly describe the scattering events between all the quasiparticles as long as these observers confer an additional property to these quasiparticles, which is conserved in the interaction process. As we will discuss in the following section, this property is charge, even though for the external observers it will be nothing but the difference between the quasiparticle number around both Fermi points.

Let us discuss one final point regarding the nature of the low-energy excitations of the system (the previously described spin waves or Dirac quasiparticles). Consider a spin wave with exactly the Fermi point momentum

$$e^{ip_{\rm F}\hat{\boldsymbol{l}}\cdot\boldsymbol{x}/\hbar}. \tag{1.76}$$

We have seen that this oscillation pattern carries no energy. This oscillation is stationary so it cannot carry momentum either. This oscillation pattern is in reality part of the ground state. If the momentum of these spin waves has a small departure $\mathfrak{p}$ from the Fermi point momentum, then we have seen that they do carry an energy $E \simeq c|\mathfrak{p}|$. As we have explained, the effective spacetime is Minkowskian, so the anisotropic velocities in the laboratory will not have any operational meaning for low-energy internal observers. We will just define a single constant c relating their space and time dimensions. The momentum carried by one such wave is precisely $\mathfrak{p}$. Its direction marks the direction of the propagation of the spin wave. Its modulus can be seen as derived from E/c. Therefore, the momentum $p_{\mathrm{F}}\hat{\boldsymbol{l}} + \mathfrak{p}$ is not the real momentum carried by the spin wave, relevant to experiments measuring impulse transfers within the liquid. The real momentum of the spin wave is $\mathfrak{p}$.

A momentum $|\mathfrak{p}|$ has an associated wavelength $\lambda = 2\pi\hbar/|\mathfrak{p}|$. These wavelengths, which are much larger than the mean interparticle distance $\lambda_{\mathrm{I}} = 2\pi\hbar/p_{\mathrm{F}}$, are the actual "observable" wavelengths of the spin waves in the liquid. In the case of a classical (large amplitude) spin wave this "observability" will match our intuitive sense of observation of a wave.

1.4.3 The emergence of charge

Charge is a fundamental concept in physics. The embedding of this experimental concept in the group-theoretical formalism has been very fruitful, as exemplified by the success of the standard model of particle physics. However, this exploitation of the charge concept does not arguably provide any hint of its origin. In this section we discuss how, in the specific emergent model we are considering, this property arises as the result of the special properties of momentum space.

Let us first discuss scattering processes to illustrate that, for an external observer (or laboratory observer), charge conservation is due to momentum conservation (in three-dimensional momentum space). Imagine the scattering of two quasiparticles from *the same* Fermi point. Momentum conservation only tells us that

$$\boldsymbol{p}_1 + \boldsymbol{p}_2 = 2Zp_{\mathrm{F}}\hat{\boldsymbol{l}} + \mathfrak{p}_1 + \mathfrak{p}_2 = \boldsymbol{p}_3 + \boldsymbol{p}_4. \tag{1.77}$$

As the products of the scattering event must be quasiparticles, the solutions of this equation must be of the form

$$\boldsymbol{p}_3 = Zp_{\mathrm{F}}\hat{\boldsymbol{l}} + \mathfrak{p}_3, \qquad \boldsymbol{p}_4 = Zp_{\mathrm{F}}\hat{\boldsymbol{l}} + \mathfrak{p}_4. \tag{1.78}$$

The momentum conservation condition is thus equivalent to the conservation of the deviations $\mathfrak{p}$:

$$\mathfrak{p}_1 + \mathfrak{p}_2 = \mathfrak{p}_3 + \mathfrak{p}_4. \tag{1.79}$$

If the scattering is instead between quasiparticles from *different* Fermi points, the conservation of momentum reads

$$\boldsymbol{p}_1 + \boldsymbol{p}_2 = \mathfrak{p}_1 + \mathfrak{p}_2 = \boldsymbol{p}_3 + \boldsymbol{p}_4, \tag{1.80}$$

which implies that

$$\boldsymbol{p}_3 = Z p_{\rm F} \hat{\boldsymbol{l}} + \mathfrak{p}_3, \qquad \boldsymbol{p}_4 = -Z p_{\rm F} \hat{\boldsymbol{l}} + \mathfrak{p}_4. \tag{1.81}$$

The resulting quasiparticles have to live in different Fermi points. Again, the momentum conservation condition is then equivalent to the conservation of the deviations $\mathfrak{p}$,

$$\mathfrak{p}_1 + \mathfrak{p}_2 = \mathfrak{p}_3 + \mathfrak{p}_4. \tag{1.82}$$

In both cases, the energy-conservation condition is equivalent at low-energies (i.e., linear on the deviations $\mathfrak{p}$) to the usual massless relativistic condition in terms of $\mathfrak{p}$.

Of course, this argument alone would be valid at linear order for any given theory by specifying a specific value of the momentum $\boldsymbol{p}_0$ and linearizing arbitrarily around $\pm\boldsymbol{p}_0$. The essential point in our construction is that there exists a value of the momentum, the Fermi momentum, which is singled out at low energies. Indeeed, at low energies only excitations around the Fermi points are produced, so that the physics will be described by the deviations $\mathfrak{p}$. These momenta can properly describe the scattering events between all the quasiparticles as long as these observers confer an additional property to these quasiparticles, which is conserved in the interaction process. This property is charge. In other words, if states with momentum $+p_{\rm F}\hat{\boldsymbol{l}} + \mathfrak{p}$ and $-p_{\rm F}\hat{\boldsymbol{l}} + \mathfrak{p}$ are to be identified with respect to the momentum label, an additional label must be attached to these in order to maintain the proper counting of degrees of freedom, as illustrated in Fig. 1.5. The low-energy observer is oblivious to the existence of Fermi points, though he could infer their existence by the low-energy properties he has access to.

From a symmetry perspective, the emergent Dirac equation in (1.94) is invariant under a U(1) transformation of ψ (in fact it is invariant under transformations of $\psi_\uparrow$ and $\psi_\downarrow$ separately; this could lead to an interesting interplay with the axial anomaly; see, e.g., [76]). The corresponding conserved charge is

$$Q := Q_\uparrow + Q_\downarrow = N_+ - N_-, \tag{1.83}$$

where the operators N_+ and N_- represent the number of excitations associated with the positive Fermi point (with positive $Z = 1$ charge) and with the negative Fermi point (with negative $Z = -1$ charge), respectively. A notion of charge has emerged in the low-energy theory owing to the duplicity of Fermi points. In the low-energy description, the conserved charge is associated with an intrinsic property of the Dirac field and its conservation has nothing to do, from this point of view, with momentum conservation. We will see later that, when coupling these Dirac quasiparticles to an effective electromagnetic field, this charge indeed plays the role of an electric charge.

This conceptualization of the fermion charge in terms of topological properties of the momentum space could be of interest for more involved systems such as the standard model of particle physics. However, to be applied to the particle content of the standard model, a more involved Fermi point structure must be considered. This is out of the scope of the present work, but some discussion can be read in [18]. Notice that this interpretation suggests that the emergence of Poincaré invariance is always accompanied

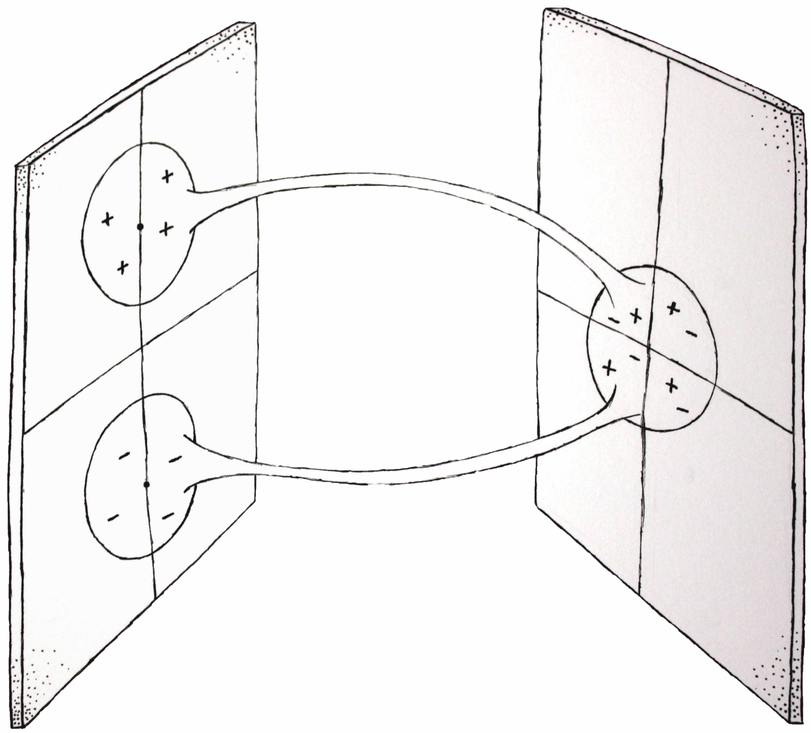

Figure 1.5: Diagram explaining the emergence of charge. The left-hand side shows the quasiparticles living near the two Fermi points. The right-hand side shows the effective low-energy description for an internal observer. An internal observer is insensitive to the origin (upper or lower Fermi point) of the quasiparticles in terms of momentum, but he sees them as having opposite charges.

with the occurrence of charge. Indeed, all the fermions in the standard model are charged with respect to some interaction carried by a spin-1 boson.

1.5 Electromagnetic fields

Electrodynamics is characterized by both bosonic and fermionic degrees of freedom. While we have exhaustively covered the fermion content of the theory, in order to accomplish a full emergent model we should deal with the bosonic excitations. Given that the building blocks of our model are fermions, it is quite natural to obtain fermionic quasiparticles in the low-energy regime. The emergence of bosonic quasiparticles is more involved, but fermion pairing offers a natural way in which these excitations may be produced. Classical electromagnetic fields would then be associated to the condensed phase of Cooper pairs, that is, to the order parameter structure. We will start then our discussion by studying the degrees of freedom associated with local disturbances of the order parameter of the planar state.

1.5.1 Textures

Let us analyze the case in which the order parameter, instead of being completely homogeneous, contains a perturbation in position space with respect to the homogeneous situations previously studied. Such inhomogeneities in this kind of order parameters are typically called textures. These variations develop over a scale which is large compared to the effective size of a Cooper pair, known as *coherence* or *healing* length, whose zero-temperature limit is [57, 59]

$$\xi_0 := \frac{\hbar}{p_{\rm F}}\frac{c_\parallel}{c_\perp} \simeq \frac{\hbar p_{\rm F}}{m^* k_{\rm B} T_{\rm C}}. \tag{1.84}$$

This quantity has a mild dependence on temperature below $T_{\rm C}$, so that its order of magnitude is essentially the one given by (1.84). This implies that the Fourier modes of the variations of the order parameter have wave numbers given by

$$k \ll k_{\rm max} := \frac{2\pi}{\xi_0}. \tag{1.85}$$

Thus all wavelength variations must be much larger than the healing length (1.84), since shorter wavelength variations are not consistent with the very existence of a local order parameter. An equivalent restriction applies to the rapidity of temporal variations of the order parameter. The anisotropy of the system implies that there is no isotropic scale in this case: we may define two natural time scales, $\xi_0/c_\perp$ and $\xi_0/c_\parallel$. The largest of both scales is the first one, so that we may safely define $t_0 := \xi_0/c_\perp$, and demand for consistency that the temporal variations of the order parameter are slower than this time scale.

The textures we will consider are of three kinds. The first one is usually called orbital texture [59] and is given by the bending of the direction $\hat{\boldsymbol{l}}$ of the angular momentum of the pairs. This bending amounts to two degrees of freedom. In addition to this, the planar-phase order parameter has, in general, the possibility of rotating around the angular momentum axis. This leads to one additional degree of freedom. In simple situations, when only this rotation is present, the additional degree of freedom is just a phase from which one can define a superfluid velocity and momentum as

$$\boldsymbol{v}_{\rm S}(\boldsymbol{x}) := \frac{1}{2m^*}\boldsymbol{\nabla}\phi, \qquad \boldsymbol{p}_{\rm S}(\boldsymbol{x}) := m^*\boldsymbol{v}_{\rm S}(\boldsymbol{x}). \tag{1.86}$$

In more complicated situations the superfluid velocity need not be irrotational (see [77]), but the important thing is that, in any case, the superfluid velocity contributes with one single additional degree of freedom to the physics of the system, on top of the two degrees of freedom of $\delta\hat{\boldsymbol{l}}$. In the simplest case in which $\boldsymbol{p}_{\rm S}$ is a constant vector we can again work in momentum space to analyze the form of the low-energy excitations. The last kind of textures we shall consider are density fluctuations, corresponding to local variations of the Fermi momentum $p_{\rm F}$.

The selection of a specific inertial frame in which the Fermi fluid is at rest can be seen as a peculiar example of spontaneously broken symmetry. Two relatively moving states have different energies with respect to a third inertial observer (e.g., the laboratory observer). However, if there were no interaction at all between the fluid and external objects in a particular frame, there would be no physical reason to select one specific uniform fluid velocity rather than another. In practice, the tiny interactions between the fluid and some specific inertial environment (typically the laboratory environment/frame) selects this very frame as the rest frame of the fluid. Then, the condensed ground state incorporates this same frame selection: the pairs are at rest with respect to this specific frame.

In what follows we take the operational view that a specific frame with a constant velocity $\boldsymbol{v}_{\mathrm{S}}$ with respect to the laboratory has been selected, regardless of the origin of this selection. This means that the pairing has occurred between $\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}$ and $-\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}$ atoms. This implies that Eqs. (1.45) and (1.46) should now be written as

$$i\hbar\dot{a}_{\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}\uparrow} = [M(\boldsymbol{p})+\boldsymbol{p}\cdot\boldsymbol{v}_{\mathrm{S}}]a_{\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}\uparrow} - c_{\perp}\boldsymbol{p}\cdot(\hat{\boldsymbol{m}}-i\hat{\boldsymbol{n}})a^{\dagger}_{-\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}\uparrow},$$
$$i\hbar\dot{a}_{\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}\downarrow} = [M(\boldsymbol{p})+\boldsymbol{p}\cdot\boldsymbol{v}_{\mathrm{S}}]a_{\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}\downarrow} + c_{\perp}\boldsymbol{p}\cdot(\hat{\boldsymbol{m}}+i\hat{\boldsymbol{n}})a^{\dagger}_{-\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}\downarrow}. \tag{1.87}$$

To reach these equations one needs to perform an active Galilean transformation under which any momentum label is shifted by $+\boldsymbol{p}_{\mathrm{S}}$, and take into account the transformation laws for the different objects appearing in the evolution operator. This is best understood from the Galilean transformation of the grand canonical Hamiltonian (1.25). Recall that the potential term is invariant under such transformation and that the kinetic term acquires two extra terms: a Doppler contribution $\boldsymbol{p}\cdot\boldsymbol{v}_{\mathrm{S}}$ and a global shift $\boldsymbol{p}_{\mathrm{S}}^2/(2m^*)$, which can be absorbed in the chemical potential for the moving system [78],

$$\bar{\mu} = \mu + \frac{\boldsymbol{p}_{\mathrm{S}}^2}{2m^*}\ ; \tag{1.88}$$

$$M(\boldsymbol{p})+\boldsymbol{p}\cdot\boldsymbol{v}_{\mathrm{S}} = \frac{(\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}})^2}{2m^*} - \bar{\mu}. \tag{1.89}$$

If we take into account that the pairing channel is given by $\boldsymbol{p}_1+\boldsymbol{p}_2 = \boldsymbol{p}_3+\boldsymbol{p}_4 = 0$, we immediately reach the conclusion that the transformed pairing Hamiltonian is given precisely by Eq. (1.32) with a Doppler shift and with all the labels shifted by $\boldsymbol{p}_{\mathrm{S}}$. Finally, the order parameter is just book-keeping the statistics of the pairs and hence depends only on the relative momenta of the members of the pairs and not on their global motion, which means that the order parameter is unchanged by the Galilean transformation.

As an alternative to the treatment in section 1.4.1, we will first combine the excitations in a bispinor and then concentrate on the excitations close to the Fermi points. The results are independent of the order of operations, but in this case it is simpler to proceed this way. Starting with the $\uparrow$ spin projection, the equations of motion are given by:

$$i\hbar\partial_t \begin{pmatrix} a_{\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}\uparrow} \\ a^{\dagger}_{-\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}\uparrow} \end{pmatrix} = H_{\boldsymbol{p}\boldsymbol{p}_{\mathrm{S}}\uparrow} \begin{pmatrix} a_{\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}\uparrow} \\ a^{\dagger}_{-\boldsymbol{p}+\boldsymbol{p}_{\mathrm{S}}\uparrow} \end{pmatrix}, \tag{1.90}$$

where

$$\mathscr{H}_{\boldsymbol{p}\boldsymbol{p}_{\mathrm{S}}\uparrow} := M(\boldsymbol{p})\sigma_3 + \boldsymbol{p}\cdot\boldsymbol{v}_{\mathrm{S}}\,\sigma_0 - c_{\perp}\boldsymbol{p}_{\mathrm{m}}\sigma_1 - c_{\perp}\boldsymbol{p}_{\mathrm{n}}\sigma_2. \tag{1.91}$$

Similar manipulations with the $\downarrow$ spin projection permit us to write the following evolution equation for bispinors:

$$i\hbar\partial_t \begin{pmatrix} a_{\boldsymbol{p}+\boldsymbol{p}_\mathrm{S}\uparrow} \\ a^\dagger_{-\boldsymbol{p}+\boldsymbol{p}_\mathrm{S}\uparrow} \\ a_{\boldsymbol{p}+\boldsymbol{p}_\mathrm{S}\downarrow} \\ a^\dagger_{-\boldsymbol{p}+\boldsymbol{p}_\mathrm{S}\downarrow} \end{pmatrix} = \mathscr{H}_{\boldsymbol{p}\,\boldsymbol{p}_\mathrm{S}} \begin{pmatrix} a_{\boldsymbol{p}+\boldsymbol{p}_\mathrm{S}\uparrow} \\ a^\dagger_{-\boldsymbol{p}+\boldsymbol{p}_\mathrm{S}\uparrow} \\ a_{\boldsymbol{p}+\boldsymbol{p}_\mathrm{S}\downarrow} \\ a^\dagger_{-\boldsymbol{p}+\boldsymbol{p}_\mathrm{S}\downarrow} \end{pmatrix}, \tag{1.92}$$

where the 4×4 evolution operator is

$$\mathscr{H}_{\boldsymbol{p},\boldsymbol{p}_\mathrm{S}} := M(\boldsymbol{p})Y^3 + \boldsymbol{p}\cdot\boldsymbol{v}_\mathrm{S}Y^0 + c_\perp \boldsymbol{p}_\mathrm{m} Y^1 + c_\perp \boldsymbol{p}_\mathrm{n} Y^2. \tag{1.93}$$

The set of matrices $\{Y^1, Y^2, Y^3\}$ were defined in (1.66), and $Y^0 := I_4$. Now if we concentrate on excitations near the Fermi points (i.e., linearize around $\boldsymbol{p} = +p_\mathrm{F}\hat{\boldsymbol{l}}$; this linearization is sufficient as the equation is already encompassing all the degrees of freedom) one obtains a Dirac equation in momentum space:

$$\left(e^a{}_b\gamma^b\bar{\mathfrak{p}}_a + \gamma^0 p_\mathrm{F}\hat{\boldsymbol{l}}\cdot\boldsymbol{v}_\mathrm{S}\right)\psi_\mathfrak{p} = 0, \tag{1.94}$$

where $\psi_\mathfrak{p}$ is the bispinor that appears in Eq. (1.92), $\mathfrak{p} = \boldsymbol{p} - p_\mathrm{F}\hat{\boldsymbol{l}}$, and $\bar{\mathfrak{p}}^\mu = (\omega, \mathfrak{p})$. The non-zero components of the tetrad $e^a{}_b$ are given by

$$e^0{}_0 := 1, \qquad e^1{}_1 := c_\perp, \qquad e^2{}_2 := c_\perp, \qquad e^3{}_3 := c_\parallel, \qquad e^i{}_0 := v^i_\mathrm{S}. \tag{1.95}$$

The corresponding metric components are

$$g^{ab} = \eta^{cd}e^a{}_c e^b{}_d, \quad \longrightarrow \quad g^{ab} = \left(\begin{array}{c|c} -1 & -v^i_\mathrm{S} \\ \hline -v^i_\mathrm{S} & D^{ij} - v^i_\mathrm{S}v^j_\mathrm{S} \end{array}\right), \tag{1.96}$$

with

$$D^{ij} = \begin{pmatrix} c_\perp^2 & 0 & 0 \\ 0 & c_\perp^2 & 0 \\ 0 & 0 & c_\parallel^2 \end{pmatrix}. \tag{1.97}$$

This is an acoustic metric [14] which, given that the we are assuming a uniform background velocity $\boldsymbol{v}_\mathrm{S}$, corresponds to a flat Minkowski spacetime. Eq. (1.94) is completely equivalent to Eq. (1.72) in the homogeneous case. The only difference is a constant shift in the energy of quasiparticles.

In order to discuss inhomogeneities it is better to work in position space. Eq. (1.94) would then be written as:

$$e^a{}_b\gamma^b(i\partial_a - b_a)\psi = 0, \tag{1.98}$$

where $\partial_a := (\partial_t, \boldsymbol{\nabla})$ is the derivative operator including time, and

$$b_a := (p_\mathrm{F}\hat{\boldsymbol{l}}\cdot\boldsymbol{v}_\mathrm{S}, p_\mathrm{F}\hat{\boldsymbol{l}}) \tag{1.99}$$

is a constant background. The content of this equation will be completely equivalent to one with $b_a = 0$, since a constant background value can be absorbed into unobservable offsets of energy and momentum. However, here we want to consider fluctuations of the background $\delta p_{\rm F}$, $\delta\hat{\boldsymbol{l}}$ and $\delta\boldsymbol{v}_{\rm S}$. These fluctuations act as an effective vector field that cannot be absorbed. This implies that, even when we absorb the offset b_μ in the redefinition of energy and momentum, we are left with a modified evolution equation for quasiparticles:

$$e^a{}_b\gamma^b(i\partial_a - \nu A_a)\psi = 0. \tag{1.100}$$

Here ν is a constant which controls the dimensions of the field A_a (recall that in standard electrodynamics the vector potential has dimensions of momentum per unit charge). Let us obtain the form of this effective vector field in terms of inhomogeneities. At the light of (1.99), one may be inclined to write down $\nu A_a = (\delta[p_{\rm F}\hat{\boldsymbol{l}}\cdot\boldsymbol{v}_{\rm S}], \delta[p_{\rm F}\hat{\boldsymbol{l}}])$. This identification is premature, as we shall see now. The kind of coupling of the fermionic quasiparticles to the vector field A^a suggests that we identify it as an effective electromagnetic gauge field, as in other inhomogeneous situations in condensed matter physics (see, e.g., [79]). However, we should put the metric in its standard Minkowskian form before carrying out this identification. To do that, we transform to comoving coordinates so that the $v^i_{\rm S}\partial_i$ term in Dirac's equation (1.100) vanishes. Then the vector field is identified as

$$\boldsymbol{A} := \frac{1}{\nu}\delta(p_{\rm F}\hat{\boldsymbol{l}}), \qquad A_0 = \frac{1}{\nu}p_{\rm F}\hat{\boldsymbol{l}}\cdot\delta\boldsymbol{v}_{\rm S}. \tag{1.101}$$

The object $\boldsymbol{A}$ is a genuine three-dimensional vector, with three degrees of freedom: two originate from the variations $\delta\hat{\boldsymbol{l}}$ of the order parameter, and the other one from density fluctuations $\delta p_{\rm F}$. On the other hand, A_0 contains just one degree of freedom independent of these, which is encoded in the variations of the superfluid velocity $\boldsymbol{v}_{\rm S}$. Let us stress again that not all the variations of the superfluid velocity are permitted, so that there is only one real degree of freedom [recall the discussion below Eq. (1.86)].

To end this subsection, let us also point out that the inhomogeneities in the order parameter make the acoustic metric (1.96) to be non-flat. Thus when considering higher-than-first-order effects, the same degrees of freedom making up this effective electromagnetic potential will be responsible for some additional effects associated with spacetime curvature.

1.5.2 Kinematical gauge symmetry

Our discussion concerning the properties of effective fermions has shown that the low-energy description of the system contains features that are not included in the original theory, such as the notion of electric charge and chirality. In this section we shall deal with another emergent property: gauge symmetry. When gauge fields are emergent entities, the discussion naturally splits in two aspects: on the one hand, the *kinematical* invariance of the theory under gauge transformations and, on the other hand, the *dynamical* preservation of this symmetry [80, 22]. The study of analogue gravity setups, where the relevant gauge group is composed by diffeomorphisms, has shown that condensed-matter analogies usually fail to achieve the second point [14, 81]. This section is devoted to an analysis of these issues in the context of the model developed in this article, where the gauge group is simpler.

By kinematical gauge invariance we refer to a property of the way in which the low-energy quasiparticle excitations, the Nambu-Gor'kov spinors in (1.61) and (1.63), react under the presence of different given fields A_a, independently of their origin. As we have seen, the fields A_a are associated with spatial and temporal variations of the orbital part of the order parameter, which is represented by a trihedral $\{\hat{\boldsymbol{m}}, \hat{\boldsymbol{n}}, \hat{\boldsymbol{l}}\}$. Kinematic gauge invariance occurs when there are equivalent classes of A_a leading to essentially the same effect upon the quasiparticles.

Consider as an example the following static texture:

$$\delta\hat{\boldsymbol{m}} = \delta\hat{\boldsymbol{m}}(\boldsymbol{x}\cdot\hat{\boldsymbol{m}}_0), \qquad \delta\hat{\boldsymbol{n}} = \delta\hat{\boldsymbol{n}}(\boldsymbol{x}\cdot\hat{\boldsymbol{n}}_0), \qquad \delta\boldsymbol{p}_{\rm s} = 0. \tag{1.102}$$

In this case, one can find a local phase transformation of the fermionic fields that transforms the evolution equation (1.100) into a free Dirac equation for the new spinor field. That is, for internal observers, the configuration (1.102) would be equivalent to the absence of textures if they identify the physical objects with equivalence classes defined by these gauge transformations. A spinor field wave packet is not deflected in any way by the previous texture and one could take that as a defining feature of the equivalent class of configurations. As we could have anticipated, two textures differing in the gradient of a scalar, $A'_a - A_a = \partial_a\varphi$, lead to the same type of effects in the spinor field; the function φ can always be absorbed locally into the spinor's local phase:

$$\psi \;\longrightarrow\; \exp[i\varphi(t,\boldsymbol{x})]\psi. \tag{1.103}$$

Recall that, in the same way, in Maxwell's model (section 1.2 and [23, 24]) electromagnetic potentials have also a reality but some of their properties are not relevant at low energies. At this point it is important to remark again that this picture is only partial: the description of the system is simple because we are looking only at low-energy phenomena. In particular, this gauge invariance will be violated at some point when the low-energy description breaks down, for instance, when the effective Lorentz invariance disappears. At some point even the condensation and so the very existence of the field A_a would disappear. On the other hand, we are not considering the excitation of other collective modes (for example, the clapping modes which can be associated with gravitons; see [82] for a general discussion of the different collective modes and their significance, and [83] for the surprising relation between these clapping modes and the effective cosmological constant in the A phase of He–3). The consideration of these modes would lead to additional excitations and therefore to more complicated theories, for instance containing gravity. We have made here the mild assumption that different sectors corresponding to excitations with different properties can be analyzed independently, being eventually combined in a unified description.

Let us now discuss the issue of dynamical gauge invariance. In principle it could be the case that the kinematical gauge invariance was not preserved by the dynamics.

By looking at the interaction of two spinor wave packets, through a mediator field A_a, one could detect differences beyond the introduction of a local phase. This amounts to the possibility of distinguishing between different members of the kinematical equivalence class. The emergence of a dynamical gauge invariance will definitely signal the irrelevance of certain degrees of freedom of A_a in the low-energy effective theory of the system.

The dynamical implementation of gauge invariance has turned out to be an issue much more subtle than expected. Rethinking the problem, we realized that the key may well reside in the very emergence of Lorentz invariance. In the next section we describe the logic of the emergence of dynamical gauge invariance along this line of thought.

1.6 The low-energy effective field theory

We have essentially put on the table all the necessary ingredients to construct the low-energy theory following the effective field theory logic [84, 65]. On the one hand, we have described the relevant degrees of freedom at low energies, and shown that the fermionic part of the spectrum (composed by fermionic quasiparticles) enjoys an emergent Lorentz symmetry. On the other hand, we still need to discuss how this symmetry is materialized in the bosonic sector (corresponding to inhomogeneities of the order parameter), as well as the mechanism of implementation of dynamical gauge invariance. There is an additional issue which is the delimitation of the regime of applicability of the resulting effective field theory. As we argue in the following, these features are intimately related.

1.6.1 Regime of applicability and effective action

A seemingly innocent question, but eminently relevant for the following discussion, is the regime of applicability of the effective field theory we are constructing. Our initial theory was a non-relativistic fermionic theory with four-fermion interaction and Hamiltonian operator (1.32). However, we rapidly resorted to approximations in order to conveniently describe the physical content of the theory. As a result, we have shown that the spectrum of low-energy excitations contains both fermions and bosons. The bosonic degrees of freedom correspond to inhomogeneities of the condensed part of the system. Eventually, we may imagine individual photons as tiny fluctuations of the condensed phase – so tiny that they involve only one (indeed, probably a finite superposition) of these effective bosons composed by a pair of fermions. This picture spotlights the typical size ξ_0 of the Cooper pairs (1.84) as a relevant scale to take into account which, in terms of momentum, corresponds to

$$\Lambda_{\rm C} := \frac{\hbar}{\xi_0} = p_{\rm F} \frac{c_\perp}{c_\parallel}. \tag{1.104}$$

As we argue in the following, this is the momentum scale which divide the degrees of freedom in low- and high-energy modes, meaning that only the modes below this scale should be taken into account in the low-energy effective field theory. This is the reason of the notation $\Lambda_{\rm C}$, typically associated with a cutoff. That this quantity has this interpretation

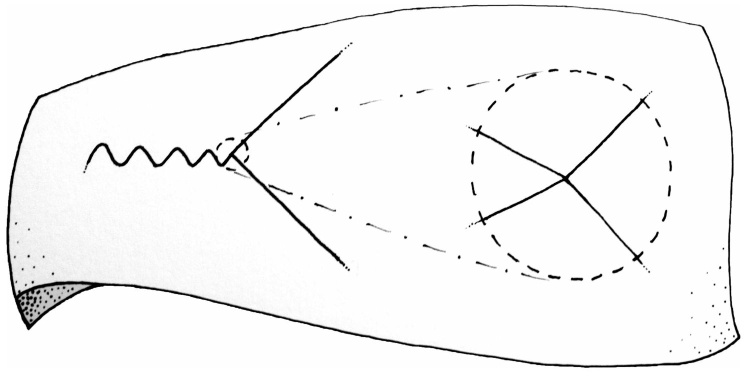

Figure 1.6: Basic scattering events in the low-energy effective theory (on the left) and in the high-energy theory (on the right). If the scattering between an effective boson and a pair of fermions is looked at closely, with spatial resolutions of the order of the Cooper pair size ξ_0, it will be perceived as a scattering event between four fermions. Equivalently, fermions with momenta greater than $\Lambda_{\rm C}$ are able to discern the fermionic content of the effective bosons, signaling the failure of the effective theory.

when considering the bosons is easy to understand: simply there exist no collective bosonic excitations with momenta greater than (1.104). On the other hand, fermionic excitations with momenta greater than (1.104) will be able to discern the fermionic content of the effective boson. In terms of Feynman diagrams, this quantity marks then the separation between processes that can be understood as the low-energy scattering of two fermions with a composite effective boson, or just the interaction of four fermions as described in the framework of the high-energy theory (see Fig. 1.6). In other words, for momenta higher than (1.104), for both on- and off-shell fermions, we should go further than the effective field theory description and consider the four-fermion interaction of the high-energy theory in order to properly describe the scattering events. This clearly marks the breakdown of the low-energy effective field theory description. This discussion is reminiscent to that of chiral effective field theory as a low-energy limit of quantum chromodynamics [85], in which the relevant degrees of freedom are mesons, i.e., effective bosons composed by quarks.

This observation not only fixes the set of low-energy modes, but also the symmetries of the effective theory. Indeed, a quick comparison between the value of $\Lambda_{\rm C}$ in (1.104) and Eqs. (1.54) and (1.57), concerning the maximum (anisotropic) value of the momentum for fermions to display the relativistic dispersion relation, shows that

$$\Lambda_{\rm C} = p^{\star}_{\perp} \leq p^{\star}(\theta). \tag{1.105}$$

This implies that all the low-energy fermionic modes are relativistic. Two effects are combined in order to guarantee this result: (i) the coherence length of the effective bosons acts as the ultraviolet cutoff of the effective theory, and (ii) this effective cutoff falls within the Lorentz-invariant part of the fermionic spectrum. Notice that these two effects are intertwined: both are controlled by the gap parameter, Δ_0, or equivalently $c_{\perp}$. Indeed, both effects have a common physical origin on the occurrence of condensation on the planar ground state. Notice that the scale $\Lambda_{\rm C}$ that marks the breakdown of the low-energy effective theory has to be isotropic. The reason is that disordered Cooper pairs leading to fluctuations of the order parameter will display arbitrary orientations. What about the emergent bosons? Given the nature of the system, it is natural to expect that their dynamical properties are inherited from that of the fermions composing the Cooper pairs. This assumption alone makes us expect a relativistic behavior of these effective bosons. Let us detail the argument in what follows. It will be useful to have in mind the intuitive picture that an effective boson corresponds to the excitation of a pair (or a finite number of pairs; this is not relevant for our discussion) of fermionic quasiparticles. What we need to show is that: (i) the tree-level propagation of the effective bosons is relativistic, and (ii) radiative corrections do not change this behavior.

The tree-level propagator of effective bosons arises as the result of integrating out all the possible perturbative diagrams for momenta greater than $\Lambda_{\rm C}$. Recall that in the ground state, Cooper pairs have zero overall momentum; see, e.g., Eq. (1.34). At the lowest order, we may intuitively think about the effective bosons as the excitations of a ground-state Cooper pair with momentum $(+p_{\rm F}\hat{\boldsymbol{l}}, -p_{\rm F}\hat{\boldsymbol{l}})$ and zero energy into states of nonzero momentum that nevertheless preserve the integrity of the pair, namely $(+p_{\rm F}\hat{\boldsymbol{l}} + \boldsymbol{k}/2, -p_{\rm F}\hat{\boldsymbol{l}} + \boldsymbol{k}/2)$. Any other kind of excitation would not maintain the structure of the pair and will be therefore energetically disfavored. The magnitude of the momentum $\boldsymbol{k}$ has to be smaller than the cutoff (1.104) implying, in particular, that the dispersion relation of the composite state is relativistic and has the same form as for fermionic quasiparticles in Eq. (1.51). This has to be corrected by taking into account diagrams with arbitrary number of fermion loops, with momenta greater than $\Lambda_{\rm C}$, that arise from interactions of the excited pair with fermionic quasiparticles. All these contributions will be materialized in the renormalization of the couplings in the Hamiltonian operator (1.32). In other words, we can integrate out all the modes the momenta of which are greater than $\Lambda_{\rm C}$, and the only difference in all our arguments will be the occurrence of dressed instead of bare constants, e.g., a change in the numeric values of the propagation speeds $c_{\|}$ and $c_{\perp}$. The stability of (1.32) is not a trivial feature, as other interaction channels (recall our discussion in Sec. 1.3.3) could in principle be generated by radiative corrections. The reason why this is not the case is that the pairing channel is the only relevant one in the effective field theory sense [65]. This arguments shows that it is not reasonable to expect deviations from the standard relativistic form in the tree-level propagator of effective bosons when radiative corrections are taken into account.[4]

Now this tree-level propagation in the effective theory may be modified when taking into account low-energy radiative corrections or, in other words, loop corrections for momentum lower than $\Lambda_{\rm C}$. The essential observation is that the value of the cutoff $\Lambda_{\rm C}$ ensures that loop integrations that are performed in order to construct the dressed propagator for

[4] Another possibility is that tree-level terms for the effective bosons do not play an important role so that the logic of induced electrodynamics *à la* Zel'dovich [86] is realized in this system. We discuss this possibility in the appendix.

the effective bosons are restricted to the relativistic region in momentum space. From a technical perspective, we shall use a regularization procedure that preserves Lorentz invariance, such as dimensional or Pauli-Villars regularization [87]. The results of these calculations will be the same as in the standard relativistic field theory, ensuring Lorentz-invariant results. That a criterion that concerns only the properties of the effective bosons, not being directly related to the fermionic excitations of the theory, permits to close up this picture satisfactorily is remarkable. Indeed, we missed this property in the first analysis carried out in reference [88], partly because of the influence of earlier discussions on the problem.

To understand the relevance of this observation let us briefly provide some context to it. It was pointed out in [16, 17] (see also [89, 90, 91] for additional discussions) that the preservation of Lorentz invariance in theories with violations of this symmetry at high energies, such as the one we are considering here, is a subtle issue if we follow the standard recipes of effective field theory. In a few words, the problem comes from the integration to high momenta within loops in Feynman diagrams. Even if the external legs of a certain Feynman diagram are below the Lorenz-violating energy scale $E_{\rm L}$, the loop integration would cover much higher energies, inevitably exploring the region with appreciable violations of the Lorentz symmetry. Using the standard rules of effective field theory, one can show that these effects percolate to low energies, leading to unacceptable (from an experimental perspective) violations of the emergent Lorentz symmetry. This does not imply that any attempt to construct an emergent Lorentz-invariant theory is doomed from the beginning but rather, as the authors of [16, 17] acknowledge, that some nontrivial mechanism should be at work in order to stabilize the emergent symmetry. The discussion on the two paragraphs above is the description of this mechanism in our model.

To be precise, let us take the specific one-loop fermion diagram with external photon legs that, aside from being well-known in electrodynamics, is also the analogous of the example that was considered in [16, 17]. The momentum integration in the loop integral must be extended to the maximum momentum allowed in the effective field theory. If the region of integration in momentum space overlaps with the region in which Lorentz-violating effects are relevant, then one should expect on general grounds modifications of the relativistic behavior of the external photon even at low energies. It is indeed generally assumed that the integration would reach the Lorenz-violating sector of the theory. However, we have shown in this specific model that there exists a natural mechanism that forbids this naive expectation to be realized: the existence of a coherence length, or effective size of the emergent bosonic degrees of freedom. Going further than this scale demands a drastic rearrangement of the degrees of freedom of the effective field theory, instead of a mere change of the dispersion relation at high energies. Thus this coherence length acts as the relevant threshold in both on- and off-shell quantities, cutting off the loop integrals far below the Lorenz-violating scale $E_{\rm L}$ and stabilizing the emergent Lorentz symmetry.[5]

[5] It is interesting to notice that in the standard model of particle physics all the interactions between fermions are mediated by bosons. This implies that this mechanism is a well-posed candidate to apply to all the known interactions in a suitable emergent theory, the low-energy limit of which recovers the standard model.

On the basis of this discussion, in order to construct the low-energy effective Lagrangian density we shall consider the most general form for the low-energy degrees of freedom consistent with the symmetry requirements, namely Lorentz invariance:

$$\mathscr{L}(\psi, A_a) := -\frac{1}{4\mu_0} F^{ab} F_{ab} + \frac{\beta}{2\mu_0} (\partial_a A^a)^2 + \frac{M^2}{2\mu_0} A_a A^a + \mathscr{L}_{\rm D}(\psi, A_a). \tag{1.106}$$

In this expression, $\mathscr{L}_{\rm D}(\psi, A_a)$ is the Dirac Lagrangian density for a fermion coupled to a vector field. The way the different dimensional constants in the problem enter in $\mathscr{L}_{\rm D}(\psi, A_a)$ is not relevant for this section, but it is discussed explicitly in the appendix.

Following the effective field theory logic, the parameters in the Lagrangian density are to be fixed by comparison with experiments, or by matching with the high-energy theory. Indeed, the knowledge of the high-energy theory permits us to examine the possible values of the mass constant in the effective Lagrangian density (1.106). A simple argument shows that M cannot be different from zero. A non-zero value for the mass parameter would mean that to create a texture there would always have to be an energy gap, no matter how smooth the texture may be (i.e., no matter how large the associated wavelengths). But precisely these smooth variations of the order parameter explore the degeneracy manifold of the planar order parameter (1.40). Thus, in the limit of very long wavelengths, the construction of a texture should cost no energy. This is clear when we think of orbital textures, but it might appear less clear for perturbations of the superfluid velocity field. The problem is that from the point of view of the order parameter, a constant velocity already imposes a specific length scale for the variations of the phase. The perturbations of the velocity are encoded in second derivatives of the phase. If perturbations of the velocity field did have a gap, then, among other things, our Lorentz-invariant hypothesis would be broken as the A_a would have an anisotropic mass (indeed, this is what happens in the A phase of He–3 when spin-orbit interactions are taken into account [18]). However, it is well known that assuming a fixed constant velocity background, the extra energy associated to the introduction of acoustic waves is such that its dispersion relation is gapless. So, indeed there is no mass term for A_a in this theory, so that we shall consider $M = 0$ in the following.[6]

As we discuss in the next section, the parameter β has no effects on the dynamics of the system, provided some natural boundary conditions are imposed. Concerning the only remaining meaningful constant in the effective action (1.106), a dimensional analysis shows that the vacuum permeability should be given by the following expression in terms of the properties of the high-energy theory,

$$\mu_0 = \frac{4\pi m^* \hbar}{\nu^2 p_{\rm F}} \alpha, \tag{1.107}$$

where α is a dimensionless constant. The notation is not accidental: it corresponds to the effective fine-structure constant of the theory. Moreover, the only dimensionless quantity

[6] There might be other effects provoking the occurrence of a mass term, but the corresponding contributions are usually polynomial in the temperature T, and thus subdominant for low temperatures [82].

one can construct from the constants in the problem are the quotients $c_\perp/c_\parallel$ and T_C/T. This means that the effective fine-structure constant must be calculable as a function of these quantities, i.e.,

$$\alpha = \alpha\left(\frac{c_\perp}{c_\parallel}, \frac{T_\mathrm{C}}{T}\right). \tag{1.108}$$

This is all we can say with certainty about this quantity. The specific form of the fine-structure constant must be determined by comparing the same process (e.g., scattering of two quasiparticles) as described in both the low-energy relativistic theory and the high-energy theory with Hamiltonian (1.32), in which all the constants are explicit. However, the occurrence of the condensation could hinder this comparison, as it implies a non-trivial resummation of the perturbative contributions at different orders. This is beyond the scope of this thesis, although it is certainly interesting. As we shall discuss in the appendix, if the picture of emergent electrodynamics is realized in this system, this comparison could be much easier to perform.

1.6.2 Gauge symmetry and dynamics

If the situation with Lorentz invariance is clear, we still need to handle the dynamical implementation of gauge invariance. Being this issue inevitably intertwined with the dynamical behavior of the effective vector potential, let us consider its equations of motion,

$$\Box A_a - (1+\beta)\partial_a\partial_b A^b = j_a. \tag{1.109}$$

The source of this equation of motion is the fermionic current

$$j^a := \bar{\psi}\gamma^a\psi, \tag{1.110}$$

which is conserved, $\partial_a j^a = 0$, by virtue of the emergent Dirac equation (1.100). If one takes the divergence in the last equation, and makes use of the conservation of the fermionic current, one finds for $\beta \neq 0$ the following equation ($\beta = 0$ leads to a trivial identity):

$$\Box(\partial^a A_a) = 0. \tag{1.111}$$

In this way, the divergence $\partial_a A^a$ effectively behaves as a free scalar field, not coupled to the rest of fields.[7] The absence of sources in this construction makes it natural to impose a zero value to this field or, in other words, the Lorenz gauge condition

$$\partial^a A_a = 0. \tag{1.112}$$

The remaining degrees of freedom are then invariant under residual gauge transformations,

$$A_a \longrightarrow A_a + \partial_a\chi, \qquad \Box\chi = 0. \tag{1.113}$$

[7] When gravity is included this sentence is no longer strictly true; while the decoupling from fermion fields still holds, the existence of such a scalar degree of freedom could have non-trivial cosmological implications through the interaction with gravity [92].

This implies that we are left with usual electrodynamics in the Lorenz gauge [61, 87]. The lack of gauge invariance of the original Lagrangian density (1.106) is due to the introduction of a spurious degree of freedom that decouples dynamically.

The discussion has been entirely classical. The decoupling seems to rely on conditions which are met only when the equations of motion hold, i.e., on shell. In quantum mechanics some off-shell properties are inevitably explored as a part of the theory. For this reason, one could suspect that this mechanism would breakdown when moving to the quantum theory. Let us detail why this is not the case. Indeed, it is remarkable that the classical description given above is completely parallel to the indefinite, or Gupta-Bleuler quantization of the electromagnetic field (see, e.g., [87] for a textbook discussion). In the framework of fundamental quantum electrodynamics some of the features of this quantization procedure arise somewhat artificially, such as the introduction of a "gauge-breaking" term proportional to $(\partial_a A^a)^2$, and the subsequent imposition of the Lorenz condition on physical states. However, from an emergent perspective they are suggested as the natural way to proceed, given that gauge invariance is not the fundamental input, but just the apparently milder assumptions of Lorentz invariance and the coupling to the fermionic current.

In the indefinite quantization, a complete set of four independent pairs of creation-annihilation operators are introduced for the vector field A^μ. Given that the classical condition $\partial_a A^a = 0$ is a consequence of the classical equations of motion, in the quantum realm it is natural to demand that it holds for transition amplitudes involving on-shell quantum states in the Hilbert space $\mathcal{H}$:

$$\langle\psi|\partial_a A^a|\eta\rangle = 0, \qquad\qquad |\psi\rangle, |\eta\rangle \in \mathcal{H}. \tag{1.114}$$

A state in $\mathcal{H}$ can be decomposed as the tensor product

$$|\psi\rangle = |\psi_F\rangle \otimes |\psi_\gamma\rangle \otimes |\psi_{\gamma'}\rangle, \tag{1.115}$$

where $|\psi_F\rangle$ are fermions, $|\psi_\gamma\rangle$ transverse (physical) photons, and $|\psi_{\gamma'}\rangle$ longitudinal photons. Then we have the following standard results: first, the condition (1.114) enforces the decoupling of longitudinal photon states $|\psi_{\gamma'}\rangle$ from physical processes. Moreover, the entire quantization is independent of the parameter β in Eq. (1.106). As in the classical case, the arbitrariness on the choice of state for the spurious degree of freedom $|\psi_{\gamma'}\rangle$ is directly related to the residual gauge invariance preserving the Lorenz condition [87], i.e.,

$$\langle\psi_{\gamma'}|A_a(x)|\psi_{\gamma'}\rangle = \partial_a\chi(x), \qquad \Box\chi = 0. \tag{1.116}$$

In summary, we have already shown in the previous section that the low-energy effective field theory is Lorentz invariant, a feature that is tied up to the existence of Fermi points and a nonzero coherence length of Cooper pairs. The resulting dynamical theory is effectively gauge invariant, and so indistinguishable from standard electrodynamics, by just taking into account the conservation of the fermion current j^μ. From the perspective of electrodynamics as a fundamental theory, the imposition of Lorentz invariance and the fact that the interaction is mediated by a massless vector field which couples to the conserved fermion current density are necessary and sufficient to obtain a gauge-invariant theory. This extends to the emergent scenario, thus fixing these as the relevant conditions one has to impose in order to reproduce electrodynamics at low energies. In other words: while in the standard view gauge invariance enforces the conservation of the fermion current, what we emphasize here is that, provided natural boundary conditions are imposed, the converse is true. One can wonder whether this specific result is particular to electrodynamics or not. This discussion will be of relevance when me move on to consider the much more complex case of the gravitational interaction, the non-abelian character of which would probably play an important role.

1.7 Conclusions

In this chapter we have made a case for understanding electrodynamics from an emergent point of view. We have also tried to convey the power contained in emergent constructions: very simple elementary components and interactions can lead to an enormously rich phenomenology as described by the effective theory. We do not commit with the specifics of the models presented. Quite on the contrary, we want to turn the emphasis to the generic characteristics of these models, those that would probably be shared by any emergent model of electrodynamics. In order to illustrate this point, let us compare the two models presented here: the Maxwell fluid model and the construction based on the physics of superfluid He–3.

Although the constituents of both models are of different nature, it is not difficult to draw parallels. Maxwell's proposal contains two kinds of elements: vortical cells, whose most salient property is their ability to acquire rotation, and ball bearings, from which one constructs the analogue of charged matter. These two elements are also present in the model inspired by He–3. The role of movable ball bearings is now played by fermionic quasiparticles, low-energy excitations of a fundamental system of fermions subject to particular kinds of interactions. We have shown that these low-energy excitations, or quasiparticles, evolve following Dirac's equation. On the other hand, when the fundamental fermions are paired up and condensed they act as vortical cells, which possess intrinsic rotational properties due to the finite value of angular momentum characteristic of the p-wave condensation. The electromagnetic fields analyzed here arise as the coarse-grained view of these effective bosons, i.e., as perturbations of the condensed portion of the system.

In both models the velocity of light is emergent. Since the underlying theories have been formulated as Galilean theories, there is in principle no obstruction for the elementary component to travel at arbitrarily large velocities. The speed of light appears as a "sound" speed, the velocity of wave-like excitations in the system. In the case of the superfluid model this velocity and its independence of the wavelength is strongly tied to the occurrence of Fermi points where the dispersion relation becomes relativistic. All the physics could be described by a privileged external observer by using Newtonian notions. However, internal low-energy observers would tend to develop ways to understand their low-energy world that do not assume external structures. This epistemological choice is certainly valid, but it seems that the price to pay would be the necessary assumption of some features as fundamental principles, and hence a loss of explanatory power. Another interesting parallelism is that, in both cases, the electromagnetic potential has a physical reality in terms of specific properties of the system under consideration. It is only at low energies that some of these degrees of freedom become effectively invisible and the internal gauge symmetry appears.

Beyond these parallelisms, the superfluid model goes further than Maxwell's model:

- While in Maxwell's construction the two substances making up the system, charged matter and electromagnetic fields, are independently postulated, in the superfluid framework they arise from the same single set of underlying elements.

- The superfluid model can account for the spinorial and quantum-mechanical properties of matter. The notion of charge cannot emerge here from the (quantum mechanical) quasiparticle density that is always non-negative ($\psi^{+}\psi \geq 0$). However, the fact that there exist two signs for the charge is a nice logical consequence of the appearance in pairs of the Fermi points.

- Moreover, it seems possible to include quantum features of the electromagnetic field in the (He–3)-like model. Individual photons would correspond to tiny fluctuations of the condensed phase – so tiny that they involve only one of these effective bosons composed by a pair of fermions. The picture suggested by this model is that photons should not be viewed as fundamental particles, but as composite structures emerging from the same fundamental ingredients as the fermionic quasiparticles (there are other examples in the literature in which photons and electrons arise from the same underlying system, although in those constructions even Fermi-Dirac statistics is emergent [93]).

Emergent views of the kind analyzed in this chapter always imply that the low-energy properties, for instance Lorentz invariance, will eventually break up at some high-energy scale. Thus, it is important to stress here that deviations from Lorentz invariance need not occur at the Planck scale (and indeed Lorentz violations at the Planck scale are almost excluded by experimental observations; see, e.g., [94, 95]). Quite on the contrary, there are strong arguments suggesting that, if general relativity is an emergent theory, then Lorentz symmetry has to be very accurately respected at the Planck scale [25, 19], being the latter defined as the scale that signals the breakdown of the low-energy effective theory. This is indeed what we observe in our construction, in which the concept of Planck scale is identified with the nonzero coherence length of Cooper pairs. This feature stabilizes the emergent Lorentz invariance with respect to radiative corrections, providing interesting insights on the observations that were raised in [16, 17] regarding the percolation of tiny Lorentz violations at high energies to low energies. It is also quite interesting to have a specific model in which a drastic reorganization of the degrees of freedom of the effective theory is required in order to go further than the Planck scale.

Although the models presented here are Newtonian at high energies, we are far from suggesting that high-energy physics should be Newtonian. What these examples show is that high-energy physics will most probably incorporate ingredients rather distinct from those of its low-energy incarnation. The emergent perspective is capable of providing tantalizing explanations of principles of physics without relaying on the specifics of the high-energy theory. We thus think that, in our search of a deeper understanding of nature, an emergent point of view is a useful and probably even necessary complement to an analysis based on fundamental principles.

Most of the points of our discussion would be relevant in order to deal with a model in which the emergence of the gravitational interaction is presented step by step. For example, it should be expected that the degrees of freedom of the gravitational field come from the perturbations of the condensed portion of the superfluid, so that gravitons should correspond to the d-channel pairing of fermions. The gravitational case would of course contain new features that are not covered in this discussion, eminently its nonlinear character. While the detailed study of such a system would be highly interesting, and may be considered if constructed as a proof of principle of the possibility of doing so, we are not going to consider this issue in this thesis due to time constraints. We will, however, use some of the lessons we have learnt in the emergent model of electrodynamics in order to extract information about these genuine properties of the gravitational interaction in a hypothetical emergent framework. In particular, the observation in Secs. 1.5 and 1.6 that gauge invariance is not necessarily fundamental (gauge-dependent quantities may have a reality at energies above the application of the low-energy effective theory) will be of relevance. If this is the case, then one should find some universal mechanism leading to gauge invariance at low energies. The emergent model of electrodynamics suggests that Poincaré invariance, supplemented with the conservation of the quantities that act as sources of the bosonic excitations, should be enough to guarantee that the physical properties of the emergent theory are indistinguishable from the usual gauge-invariant theory.

Indeed, our construction may be considered as a toy-model, being electrodynamics a much simpler setting, for the construction of fundamental theories of the known interactions and gravity such as, e.g., string theory or quantum gravity. One of the conclusions we can draw on from this model is that, in order to achieve the goal of reproducing electrodynamics (or the observed properties of superfluid helium that are described by the same formalism), it is necessary to resort to a series of approximations that may not be completely well founded, or naturally suggested, in a purely mathematical discussion. It has been only due to the constant experimental guidance that the mathematical formulations of these problems were satisfactorily developed, even if the fundamentals of the microscopic theory were already well known. This observation is even more interesting given the lack of success in recovering the known low-energy physics in those fundamental approaches: e.g., recovering the standard model of particles plus gravity in string theory, or the usual continuum picture of spacetime in quantum geometry approaches. Reasoning by analogy with well-known and therefore more mature frameworks, such as the condensed-matter-like model considered here, may prove a powerful approach in order to tackle these involved issues. As a specific example, we have discussed the tension between Lorentz violations at high energies (ubiquitous in both condensed-matter-like models and discrete abstract models) and the preservation of the relativistic symmetry at low energies that arises in the formalism of effective field theory. When this effective framework is embedded in the condensed-matter-like model we have constructed, the tension is naturally alleviated by means of the very mechanism associated with the occurrence of Cooper pairs and thus of superfluidity. These unexpected connections between different branches of physics are insightful, in that they could provide new ways to think about old problems.

Appendix A: Inhomogeneous Ginzburg-Landau free energy and Zel'dovich picture of emergence

In our discussion, we have not followed the usual way in which the issue of the dynamics of the order parameter is approached (see, e.g., [18]). In this appendix, let us shortly describe this approach, as well as its shortcomings that eventually motivated us to look for a consistent alternative. The following discussion may be more demanding in terms of background material, at least in order to understand all the technical points, but it is conceptually simple: it is the comparison of two methods of different nature, in order to check whether their results may be identified as essentially depicting the same reality.

1. From a condensed-matter perspective, it may be natural to expect that the dynamics of the order parameter can be determined by a generalized Ginzburg-Landau approach, analogue to the one sketched in Sec. 1.3.4, but this time retaining terms containing derivatives of the order parameter. In a similar way, the temperature-dependent coefficients accompanying each term of the free action can be calculated from the microscopic theory. The result of this calculation would be an expression quadratic in the derivatives of the textures. If, for the moment, we restrict the discussion to orbital textures of the unit vector, $\delta\hat{\boldsymbol{l}}$, the corresponding inhomogeneous part of the free action was worked out in [96]. At finite temperature T it is given by the following expression:
$$\frac{p_{\rm F} c_{\|}}{12\pi^2\hbar}\left[\ln\left(\frac{\Delta_0}{k_{\rm B}T}\right)[\hat{\boldsymbol{l}}\times(\boldsymbol{\nabla}\times\hat{\boldsymbol{l}})]^2+[\hat{\boldsymbol{l}}\cdot(\boldsymbol{\nabla}\times\hat{\boldsymbol{l}})]^2+(\boldsymbol{\nabla}\cdot\hat{\boldsymbol{l}})^2\right]. \qquad (1.117)$$
In this expression we are keeping the dominant terms in the zero-temperature limit $T\to 0$, as well as the first order in an expansion in the parameter $c_\perp/c_\|$. This expansion is usually carried out in the literature because of the smallness of this parameter in the experimental case of He–3 that we have stressed several times. The reason for the infrared divergence in the first term can be traced back to the existence of Fermi points in the fermionic spectrum. In laboratory realizations the infrared divergence is always regulated by the temperature of the system. However, the other terms have coefficients which are constant in the limit $T\to 0$. Therefore one can always, in principle, lower the temperature sufficiently to make the first term in (1.117) the dominant one. For completeness, let us mention an assumption one

can read in the literature, concerning the existence of the following additional term in this expansion:

$$\frac{p_{\mathrm{F}} c_{\parallel}}{12\pi^2\hbar}\left(\frac{c_{\perp}}{c_{\parallel}}\right)^2 \ln\left(\frac{\Delta_0}{k_{\mathrm{B}}T}\right)[\hat{\boldsymbol{l}}\cdot(\boldsymbol{\nabla}\times\hat{\boldsymbol{l}})]^2. \tag{1.118}$$

It has been claimed [18] that this term has been usually neglected because of being of quadratic order in the parameter $c_{\perp}/c_{\parallel}$ when compared to (1.117), but that it appears in an explicit evaluation that is carried out up to this order [97]. From our point of view, however, there is no conclusive argument in this respect, as the definite relation between the evaluation of the Ginzburg-Landau energy and the technical procedure used in [97] (which indeed seems to be closer to Zel'dovich's approach we discuss in this section) is not clear for us. Notice that the terms presented here, (1.117) plus (1.118), would correspond to the potential energy of a theory with the usual kinetic term for the restricted kind of textures considered, $(\partial_t\delta\hat{\boldsymbol{l}})^2$. To be precise, this would be the result of this procedure:

$$(\partial_t\delta\hat{\boldsymbol{l}})^2+\frac{p_{\mathrm{F}} c_{\parallel}}{12\pi^2\hbar}\left[\ln\left(\frac{\Delta_0}{k_{\mathrm{B}}T}\right)[\hat{\boldsymbol{l}}\times(\boldsymbol{\nabla}\times\hat{\boldsymbol{l}})]^2+[\hat{\boldsymbol{l}}\cdot(\boldsymbol{\nabla}\times\hat{\boldsymbol{l}})]^2+(\boldsymbol{\nabla}\cdot\hat{\boldsymbol{l}})^2\right]$$
$$+\frac{p_{\mathrm{F}} c_{\parallel}}{12\pi^2\hbar}\left(\frac{c_{\perp}}{c_{\parallel}}\right)^2 \ln\left(\frac{\Delta_0}{k_{\mathrm{B}}T}\right)[\hat{\boldsymbol{l}}\cdot(\boldsymbol{\nabla}\times\hat{\boldsymbol{l}})]^2. \tag{1.119}$$

In the following we will check whether or not all these terms can be obtained in a simpler way from the perspective of the low-energy relativistic theory we have been describing in the main text.

2. On the other hand, i.e., within the emergent relativistic field theory framework, a possible way in which an internal observer can determine the dynamics of the gauge fields is by integrating out fluctuations of the relativistic fermionic fields. This is, again, nothing but the suggestion of Sakharov concerning gravity [98], adapted to electrodynamics by Zel'dovich [86]. A recent revision can be read in [99]. The integration over fermionic fluctuations, which technically amounts to an evaluation of a fermionic path integral in the presence of background fields, can be found in the literature carried out in different ways; see for example [100, 97]. A note of caution: these two approaches (Ginzburg-Landau on the one hand and Sakharov-Zel'dovich on the other) are, in principle, very distinct in nature. One is a finite-temperature analysis (implying that we have a thermal distribution of fermionic quasiparticles) while the second is a zero-temperature calculation. We will proceed with the comparison anyway, but keeping this in the back of our minds.

   From a conceptual perspective, the main difference in this approach as compared with that of Sec. 1.6.2 is the assumption that there are no tree-level terms in the effective action describing the propagation of the effective bosons, or at least these terms are negligible (in some sense detailed below). The only terms that are present in the

Lagrangian density describe the propagation of fermionic excitations in presence of "frozen" distributions of bosons. The propagation properties of these bosons in the effective field theory must then arise from radiative corrections that are calculable within the effective field theory, that are assumed to be dominant over any possible tree-level term (*one-loop dominance* [98, 99]). This procedure resonates well with our step-by-step construction.

The one-loop fermion diagram with external photon legs is then understood as the evaluation of the first nonzero contribution to the dressed propagator for a composite boson, understood as a two-fermion object. The only calculation we need is then the evaluation of the one-loop polarization tensor. This is a standard calculation in electrodynamics (see, e.g., [61, 87]) that involves an ultraviolet regularization of the momentum integral with an upper cutoff Λ_+ as well as an infrared regularization by means of a similar quantity Λ_-. At this stage the significance of these quantities is merely formal, although they will gain a clear physical interpretation later. We first need to write down the specific form of the Dirac Lagrangian density $\mathscr{L}_{\rm D}(\psi, A_a)$ in the effective action (1.106). The linearized kinetic term for the fermionic field that leads to the Dirac equation (1.72) is given by

$$\hbar \int dt d^3x\, e^a{}_b \overline{\psi} \gamma^b \partial_a \psi. \tag{1.120}$$

The fermionic field has dimensions $[\psi] = [L^{-3/2}]$ as it might be supposed from the original Hamiltonian operator (1.23). To eliminate the tetrad $e^a{}_b$ one can rescale the spatial coordinates:

$$\hbar c_\parallel c_\perp^2 \int dt d^3x'\, \overline{\psi} \gamma^a \partial'_a \psi; \qquad x'_\perp := \frac{1}{c_\perp} x_\perp, \quad x'_\parallel := \frac{1}{c_\parallel} x_\parallel. \tag{1.121}$$

Now all the integration variables have dimensions of time. One can introduce a constant with dimensions of velocity $[LT^{-1}]$ to convert the units of the integration measure. Physical results should not depend on this quantity and, in fact, when the fermionic fields are suitably rescaled,[8] the kinetic term acquires the usual form

$$\hbar \int d^4x\, \overline{\psi}\, \partial\!\!\!/ \psi. \tag{1.122}$$

The same kind of reasoning permits us to write the interaction term corresponding to the coupling to the effective vector potential in (1.100), so that

$$\int \mathrm{d}^4x\, \mathscr{L}_{\rm D}(\psi, A_a) = \hbar \int d^4x\, \overline{\psi}\, \partial\!\!\!/ \psi + \nu \int d^4x\, \overline{\psi} A\!\!\!/ \psi. \tag{1.123}$$

In the case in which Lorentz and gauge symmetries are preserved by the regularization method (using, e.g., dimensional or Pauli-Villars regularization [87], the divergence

[8] We are interested in the calculation of the path integral of the theory, so this rescaling is innocuous.

is logarithmic in the limit $\Lambda_+/\Lambda_- \to \infty$ and corresponds to a term in the action

$$-\frac{\nu^2}{48\pi^2\hbar} \ln\left(\frac{\Lambda_+}{\Lambda_-}\right)^2 \int d^4x\, F^{ab}F_{ab}. \tag{1.124}$$

In standard quantum field theory, this term would be absorbed by a suitable counterterm before taking the limit $\Lambda_+/\Lambda_- \to \infty$, leading to charge and photon field renormalization [61]. However, in this construction we have no definite tree-level kinetic term for the vector potential. Moreover, the ultraviolet divergence here is clearly an artifact of the extrapolation of the low-energy theory to higher energies. Following Zel'dovich, we can interpret this (finite) term as the leading kinetic term for the electromagnetic field induced at one-loop level:

$$-\frac{1}{4\mu_0} \int d^4x\, F^{ab}F_{ab}. \tag{1.125}$$

In principle, the logarithmic behavior supports this interpretation since it permits this term to be dominant over any unknown, tree-level dynamical terms for the gauge field, at least for certain values of the quotient Λ_+/Λ_-. This is what is usually understood as one-loop dominance [98, 99]. In this regime it is reasonable to think that the dynamics of the order parameter should be well described by (1.125). In the resulting effective theory, as expected, the actual value of ν is not relevant since it can be changed by a redefinition of the vector potential; the only meaningful quantity is the combination

$$\alpha_{\rm ind} := \frac{\nu^2\mu_0}{4\pi\hbar} = \frac{3\pi}{\ln(\Lambda_+/\Lambda_-)}. \tag{1.126}$$

Notice that the value of this dimensionless coupling constant, the effective fine-structure constant, is universal: it only depends on the value of the ultraviolet and infrared cutoffs.

3. The next step would be the comparison of the term (1.124) with the corresponding limit of the Ginzburg-Landau approach, summarized in Eq. (1.119), that contains in principle all the information about the evolution of textures. We are going to do it for restricted textures in which the superfluid velocity plays no role. At low temperatures two terms in (1.119) are the dominant ones. At first order in the texture, for which

$$-\frac{\nu}{p_{\rm F}}\boldsymbol{A} = \delta\hat{\boldsymbol{l}} = \delta\hat{\boldsymbol{m}} \times \hat{\boldsymbol{n}}_0 + \hat{\boldsymbol{m}}_0 \times \delta\hat{\boldsymbol{n}}, \tag{1.127}$$

one can see that the sum of these terms is equivalent to the spatial part of the relativistic term which can be written as:

$$\frac{1}{2}F^{ab}F_{ab} = c_\perp^2(\partial_0 A_1)^2 + c_\perp^2(\partial_0 A_2)^2$$
$$- c_\perp^4(\partial_1 A_2 - \partial_2 A_1)^2 - c_\perp^2 c_\parallel^2(\partial_3 A_1)^2 - c_\perp^2 c_\parallel^2(\partial_3 A_2)^2. \tag{1.128}$$

In this way, one can in principle accept that the picture of induction of dynamics captures the relevant dynamics of the system when the temperature is low enough. As long as Lorentz invariance is kept intact in the present scheme, one can argue in favor of the occurrence of the term (1.118) in the inhomogeneous Ginzburg-Landau free energy (1.119). Then the matching of the low-energy relativistic theory with the Ginzburg-Landau approach tells us the value of the quotient of regulators:

$$\frac{\Lambda_+}{\Lambda_-} = \frac{\Delta_0}{k_{\mathrm{B}} T}. \tag{1.129}$$

This fixes the value of the fine-structure constant (1.126) in terms of parameters of the system and provides an interpretation of the ultraviolet and infrared regulators. They would be given by:

$$\Lambda_+ = \Delta_0, \qquad \Lambda_- = k_{\mathrm{B}} T. \tag{1.130}$$

The value of the infrared regulator simply implies that the energy scale associated with the temperature physically removes the infrared divergence. On the other hand, the value of the ultraviolet regulator is telling us that we are performing the integration over fermions up to energies given by Δ_0. In order to compare this result with our previous discussion in Sec. 1.6.1, we need to say a few words about a property that we have ignored in the latter calculations: the anisotropy of the system. That we did not worry about this issue is not a coincidence. The small quotient $c_\perp/c_\parallel$ makes the system essentially isotropic in what concerns the application of a momentum cutoff. The reason is that the relativistic dispersion relation of quasiparticles (1.51) will be dominated by the longitudinal terms, except essentially in the orthogonal plane $\mathfrak{p}_\parallel = 0$. This implies that, except for small corrections, the energy cutoff that is derived from the application of the momentum cutoff (1.104) is given by

$$c_\parallel \Lambda_{\mathrm{C}} = c_\parallel p_{\mathrm{F}} \frac{c_\perp}{c_\parallel} = \Delta_0 = \Lambda_+. \tag{1.131}$$

Thus both notions coincide! It is remarkable that two different arguments of very different nature lead to the same cutoff for the low-energy theory. Moreover, this implies that the integration in the Zel'dovich picture is performed in the relativistic region of the spectrum, so that it is indeed reasonable that the dynamics is given by the relativistic term (1.125).

Notice that this feature is essential in order for the picture to be consistent: on the contrary, the integration over non-relativistic fermions would make appear terms which are quadratic in the texture but non-relativistic (the relevant symmetry group is ultimately the Galileo group). These terms can be suppressed only by the logarithm in Eq. (1.119) or, equivalently, Eq. (1.124) at low temperatures. The first unsatisfactory feature of this argument is the difficulty of reconciling this logarithmic suppression with the accuracy of known symmetries [101]. But the definite drawback is the following: the value of the fine-structure constant (1.126) shows that the suppression of these terms would be proportional to α. Thus in this situation, the usual perturbative expansion in terms of the fine-structure constant that works perfectly well in electrodynamics would make no sense.

The identification of ultraviolet and infrared cutoffs with Δ_0 and $k_{\rm B}T$, respectively, permits to write down the fine-structure constant (1.126) as a particular case of Eq. (1.108):

$$\alpha_{\rm ind} := \frac{\nu^2\mu_0}{4\pi\hbar} = \frac{3\pi}{\ln(T_{\rm C}/T)}. \tag{1.132}$$

This would imply a nearly zero temperature of the system if we use the experimental value of the fine-structure constant $\alpha \simeq 1/137$:

$$T = T_{\rm C}\exp[-3\pi\alpha_{\rm ind}] \simeq 10^{-434}T_{\rm C}. \tag{1.133}$$

In summary, there are indications that these two different mechanisms are pointing to the same dynamics for the emergent electromagnetic fields. However, there is still a lot of work to do in order to claim a categorical equivalence. First, the term (1.118) has to be shown to appear in the evaluation of the inhomogeneous Ginzburg-Landau energy. On the other hand, Eq. (1.119) does not take into account all kinds of textures that we have discussed; the incorporation of these makes the evaluation of the inhomogeneous Ginzburg-Landau energy much more involved. As we have stressed, the overall consistency arises from the existence of a coherence length that imposes a momentum cutoff (1.104) that falls well below the scale of violation of the emergent Lorentz symmetry. This is one of the main observations we want to make: without this feature it would be difficult to support a comparison between the Ginzburg-Landau free energy and the low-energy action for the emergent gauge fields *à la* Zel'dovich. Note the essential difference between the approach depicted in this appendix and the discussion in Sec. 1.6.2: in the former it is argued that gauge-breaking terms are conveniently suppresed, while in the latter it is shown that these terms are harmless when only conserved currents are allowed as sources of the electromagnetic fields. The discussion above encourages the exploration of these open issues, that can moreover serve to push further the understanding of the present model as a beautiful example of the principles of emergence and effective field theory at work.

Appendix B: Some comments regarding the ABM order parameter

While we have based all our discussion on a specific choice of condensed ground state, the planar state, and the reasons for using it should be clear by now, we would like to illustrate in this appendix what would happen if we would have considered different states (see also the discussion in [18]). Let us carry out a brief analysis of the quasiparticle evolution equations for the ABM state with order parameter (1.41) and basis vectors

$$\hat{\boldsymbol{x}} = \hat{\boldsymbol{m}}, \qquad \hat{\boldsymbol{y}} = \hat{\boldsymbol{n}}, \qquad \hat{\boldsymbol{z}} = \hat{\boldsymbol{s}}, \tag{1.134}$$

along the lines of the analysis performed for the planar state in the main text, to show the differences between both states. The reason for the choice of axes (1.134) is that the alignment of $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{m}} \times \hat{\boldsymbol{n}}$ is favored by the action of nuclear dipole interactions [102, 103], so that we may take them as being mutually orthogonal. The arguments and conclusions in this section do not depend on the choice of basis.

Because of this choice of axes, it is better to write the evolution equations of quasiparticle operators in the spin basis in $\hat{\boldsymbol{x}}$ direction, defined as

$$a_{\boldsymbol{p}\rightarrow} := \frac{a_{\boldsymbol{p}\uparrow} + a_{\boldsymbol{p}\downarrow}}{\sqrt{2}}, \qquad a_{\boldsymbol{p}\leftarrow} := \frac{a_{\boldsymbol{p}\uparrow} - a_{\boldsymbol{p}\downarrow}}{\sqrt{2}}. \tag{1.135}$$

When these equations are linearized around the Fermi point $+p_{\mathrm{F}}\hat{\boldsymbol{l}}$ and represented in position space, one obtains the analogue of Eq. (1.98) but split for both spin projections. One can check that one obtains similar equations directly in the basis $\uparrow, \downarrow$ when $\hat{\boldsymbol{s}} = \hat{\boldsymbol{x}}$. If, instead of that, one takes $\hat{\boldsymbol{s}} = \hat{\boldsymbol{y}}$, the equations are almost the same, with only a change of the sign of the two last terms in the last equation. The evolution operators are given by the following expressions:

$$\begin{aligned} \mathscr{H}_{\rightarrow} &:= c_{\parallel}\hat{\boldsymbol{l}} \cdot (-i\boldsymbol{\nabla} - p_{\mathrm{F}}\hat{\boldsymbol{l}}) + c_{\perp}(\sigma_1\hat{\boldsymbol{m}} - \sigma_2\hat{\boldsymbol{n}}) \cdot (-i\boldsymbol{\nabla}), \\ \mathscr{H}_{\leftarrow} &:= c_{\parallel}\hat{\boldsymbol{l}} \cdot (-i\boldsymbol{\nabla} - p_{\mathrm{F}}\hat{\boldsymbol{l}}) - c_{\perp}(\sigma_1\hat{\boldsymbol{m}} + \sigma_2\hat{\boldsymbol{n}}) \cdot (-i\boldsymbol{\nabla}). \end{aligned} \tag{1.136}$$

One can realize that these equations both have the same chirality, by multiplying the ± 1 factors in front of the Pauli matrices. For this reason it is better to apply a linear transformation to one of the equations, say the second, to change its chirality. Such a transformation is given by

$$\psi_{\leftarrow} \longrightarrow i\sigma_2\psi^*_{\leftarrow}. \tag{1.137}$$

The transformed Hamiltonian is

$$\mathscr{H}'_{\leftarrow} := -\sigma_2\mathscr{H}^*_{\leftarrow}\sigma_2 = -c_{\parallel}\hat{\boldsymbol{l}} \cdot (-i\boldsymbol{\nabla} + p_{\mathrm{F}}\hat{\boldsymbol{l}}) + c_{\perp}(\sigma_1\hat{\boldsymbol{m}} - \sigma_2\hat{\boldsymbol{n}}) \cdot (-i\boldsymbol{\nabla}). \tag{1.138}$$

In the same way as we did with the planar phase, let us define the matrices

$$Z^1 = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_1 \end{pmatrix}, \quad Z^2 = \begin{pmatrix} -\sigma_2 & 0 \\ 0 & -\sigma_2 \end{pmatrix}, \quad Z^3 = \begin{pmatrix} \sigma_3 & 0 \\ 0 & -\sigma_3 \end{pmatrix}. \tag{1.139}$$

The problem of finding a representation of the gamma matrices $\{P, PZ^a\}_{a=1,2,3}$ is similar to the one studied for the case of the planar state. Here, a solution is given by

$$P := \begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix}. \tag{1.140}$$

However, now the two chiralities have a different coupling to the vector potential when the same perturbative analysis of Sec. 1.5.1 is applied to the evolution equations of this section. Such a coupling implies that we cannot write the low-energy evolution equations as a Dirac field coupled to a vector potential. In fact, the coupling is now axial, with a term in the equation of motion proportional to

$$\gamma^5 \gamma^a V_a \psi. \tag{1.141}$$

Here we are using the symbol V_a instead of A_a to denote the inhomogeneities of the orbital part of the order parameter. The reason is that, even if these objects are written in the same way in terms of the inhomogeneities, they should have different transformation properties under the usual symmetry transformations such as parity. This can be understood by looking at the structure of the Cooper pairs in both states, ABM and planar. In the first case the vector $\hat{\boldsymbol{l}}$ shows the direction of the angular momentum of the Cooper pairs with positive as well as negative projection of spin, and so, $\hat{\boldsymbol{l}}$ is an axial vector in this state. On the other hand, in the planar state the two spin populations form Cooper pairs with opposite angular momentum, implying that the planar state is not axial (for a more detailed discussion see Sec. 7.4 of [18]). This picture is consistent with the kind of coupling to a vector field which appears in each state.

This does not contradict the claim that the low-energy quasiparticle excitations of the Fermi liquid are determined by topology in momentum space. Both ABM and planar states are characterized by two Fermi points. In a homogeneous system, the low-energy fermionic excitations are the same in both states, and can be represented by a (free) Dirac field. However, the structure of the order parameter is different in both states, and so is the coupling of the fermionic excitations to the inhomogeneities of this order parameter. In other words, the only difference between these two states is their chirality. This is the reason why the planar state serves better as an analogue of the ground state of electrodynamics.

ABM and planar states are limiting cases of the family of axiplanar states [18]. In these states the two spin populations are decoupled as far as the order parameter is concerned. For these two limiting states, one is considering perturbations with $\delta\hat{\boldsymbol{l}}_\uparrow = \delta\hat{\boldsymbol{l}}_\downarrow$ and $\delta\hat{\boldsymbol{l}}_\uparrow = -\delta\hat{\boldsymbol{l}}_\downarrow$, respectively. General axiplanar states can be analyzed with the same techniques to show that, in general, one has couplings to a polar as well as an axial vector, both constructed by different linear combinations of the independent variations $\delta\hat{\boldsymbol{l}}_\uparrow$ and $\delta\hat{\boldsymbol{l}}_\downarrow$, i.e., $(\delta\hat{\boldsymbol{l}}_\uparrow \pm \delta\hat{\boldsymbol{l}}_\downarrow)/\sqrt{2}$.

Chapter 2

From gravitons to gravity

2.1 Emergent gravitons and gauge invariance

In the first chapter of this thesis we have considered mechanisms that lead generically to emergent Lorentz-invariant theories at low energies. The fact that we have used electrodynamics as the case study does not prevent to extract lessons that could also prove valuable for more complicated scenarios. While a natural step forward in the same direction of the previous discussion would be to construct a specific model in which the gravitational degrees of freedom are also emergent at low energies, this falls outside the scope of this thesis, mainly due to time constraints. Instead, in this chapter we will only look at some aspects of this problem, which are complex and interesting enough to be studied on their own. This section serves as the link between the previous chapter and the contents of the present one.

The essential difference between the electromagnetic and the gravitational cases is the intrinsically nonlinear character of the latter, which has its roots in the fact that all forms of energy gravitate. A priori, it could be possible to construct a more complicated condensed-matter-like model that displays, strictly at the linear level, excitations that contain the degrees of freedom associated with gravitons (in the following, by gravitons we shall understand massless, spin-2 particles).[1] The nontrivial conceptual question that is addressed in this chapter is how the complete nonlinear dynamics of general relativity can arise in this framework or, in other words, which kind of principle can be used to select the Einstein-Hilbert action as describing the nonlinear behavior of the degrees of freedom associated with gravitons. So we will no longer work with a condensed-matter-like model, but we shall make the assumption that there exists a suitable model of this kind that leads to the effective field theory linear description of free gravitons, the mathematical details of which we discuss in the following. While both gravitons and photons are up to date accepted to be massless [105], they have different intrinsic angular momentum. In mathematical terms,

[1] In this regard, the Weinberg-Witten theorem [104] is usually cited as an important obstruction. However, this theorem is not directly applicable when gauge invariance emerges jointly with Lorentz invariance, which is what we expect on the basis of our in-depth discussion of electrodynamics as an emergent theory.

these excitations correspond to different irreducible representations of the Poincaré group [106, 107].

In practical terms, the only difference is that we should use a second rank tensor h^{ab} as the natural description of these excitations, instead of a vector field as we did with photons. As with the construction of the effective action for the vector field A^a, the only condition to be imposed is Lorentz invariance. We do not demand gauge invariance from the beginning, contrary to what is usually done in other contexts. This implies that the construction has a priori more degrees of freedom than those strictly associated with gravitons. But we will argue in this section that the analogue of what we proved in the case of electrodynamics holds: when the coupling of the excitations described by h^{ab} to the rest of fields is carried out through a specific conserved current, some of the degrees of freedom that are formally contained in h^{ab} are effectively frozen at the linear level, leading to the usual linear description of gravitons.

The most general Lagrangian density following these principles, and ignoring for the time being surface terms, turns out to be [108]

$$\mathscr{L}^{(1)} + c_2\mathscr{L}^{(2)} + c_3\mathscr{L}^{(3)} + c_4\mathscr{L}^{(4)}, \tag{2.1}$$

where we have fixed an irrelevant global normalization factor (fixing $c_1 := 1$) and defined the following terms:

$$\mathscr{L}^{(1)} := -\tfrac{1}{4}\partial_a h^{bc}\partial^a h_{bc}, \qquad \mathscr{L}^{(2)} := -\tfrac{1}{2}\partial_b h^{bc}\partial_a h^a{}_c,$$
$$\mathscr{L}^{(3)} := \tfrac{1}{2}\partial^a h\,\partial^b h_{ab}, \qquad \mathscr{L}^{(4)} := -\tfrac{1}{4}\partial_a h\,\partial^a h.$$

This description is not unique, as one can perform the following redefinition of the graviton field (in the following, h will stand for the trace of h^{ab}):

$$h^{ab} \to h^{ab} + \gamma\,\eta^{ab}h. \tag{2.2}$$

Under this transformation, with γ constant, the parameters in the Lagrangian density (2.2) change following certain transformation rules that are not important in our discussion. These transformation rules can be checked in [108]. To this linear description of the effective gravitational degrees of freedom, we add additional matter fields that describe fermionic and bosonic excitations with lower spin. In the following we will collectively call them *matter* fields.

The free equations of motion for the graviton field are given by:

$$\frac{1}{2}\Box h^{ab} + c_2\partial^c\partial^{(a}h^{b)}{}_c - \frac{c_3}{2}\eta^{ab}\partial_c\partial_d h^{cd} - \frac{c_3}{2}\partial^a\partial^b h + \frac{1}{2}c_4\eta^{ab}\Box h = 0. \tag{2.3}$$

Here $\Box := \partial^a\partial_a$ is the d'Alembert operator in flat spacetime. In order to successfully describe the gravitational phenomena, we should couple the graviton field to the matter fields. We need then to select a second-rank tensor field to play the role of the source. In analogy with the case of electrodynamics, we shall take a conserved current: the stress-energy tensor T^{ab}, which is the conserved current associated with invariance under translations in Minkowski spacetime.

So let us consider the inclusion of a term λT_{ab}, where λ stands for the coupling constant with suitable units, to the right-hand side of Eq. (2.3). Here T_{ab} is the stress-energy tensor of the matter fields. The specific functional form of the stress-energy tensor is not relevant for the moment, but just two of its properties. The first one is its conservation

$$\partial_b T^{ab} = 0. \tag{2.4}$$

The second one is less general but necessary for the following arguments to hold: the stress-energy tensor *must be* traceless for the following mechanism to be operative, i.e.,

$$T := \eta_{ab} T^{ab} = 0. \tag{2.5}$$

The traceless condition follows from the massless character of the matter fields or, in other words, from classical scale invariance. The arbitrariness in the addition of identically conserved Belinfante terms to the canonical stress-energy tensor permits to construct the so-called *improved* stress-energy tensor, which is traceless in the presence of scale invariance (see, for instance, the corresponding discussion in [109]). Recall that this is the result that we obtained when discussing the emergence of electrodynamics in the previous chapter; nonzero masses would require a suitable coarse-graining procedure, or radiative symmetry breaking [110], for instance. In other words, we are identifying T^{ab} with the improved stress-energy tensor. With the two Eqs. (2.4) and (2.5) one can show that the degrees of freedom of the field h^{ab} are effectively reduced to those that are usually associated with gravitons.

The demonstration goes as follows. In the general case, namely for arbitrary values of the parameters in the Lagrangian density (2.1), one can show that these five conditions imply the following five conditions on the graviton field:

$$\partial_b h^{ab} = 0, \qquad\qquad h = 0. \tag{2.6}$$

More precisely, the action of the d'Alembert operator on these quantities is zero which, provided natural boundary conditions are chosen, implies Eq. (2.6). This reasoning is completely parallel to the corresponding discussion in electrodynamics.

Indeed, taking the divergence on Eq. (2.3) leads, even in the presence of the source term T_{ab}, to the equations

$$\Box \left(\partial_b h^{ab} + c_2 \partial_b h^{ab} - c_3 \partial^a h + c_4 \partial^a h\right) - (1 + c_3)\, \partial^a \partial_b \partial_c h^{bc} = 0. \tag{2.7}$$

These are four equations that are solved by the four conditions

$$(1 + 2c_2 - c_3)\partial_b h^{ab} + (c_4 - c_3)\partial^a h = 0. \tag{2.8}$$

These are the analogue of [insert the equation in chapter 1]. On the other hand, taking the trace of Eq. (2.3) also with the source term leads to

$$(c_2 - 2c_3)\partial_a \partial_b h^{ab} + \frac{1}{2}\,(1 - c_3 + 4c_4)\,\Box h = 0. \tag{2.9}$$

Given these two equations, (2.8) and (2.9), there are different possible options depending on the specific values of the constants c_2, c_3 and c_4. For instance, the well-known Fierz-Pauli theory corresponds to the left-hand side of Eq. (2.8) being identically zero, with $c_2 = c_3 = c_4 = -1$. Another interesting case, known as Weyl-transverse theory, is given by $c_2 = -1$, $c_3 = -1/2$ and $c_4 = -3/8$, making the left-hand side of Eq. (2.9) identically zero. These two situations have the maximum number of generators of gauge symmetries, and therefore both describe the two degrees of freedom of a graviton [111, 108]. A clear way to see this is that, in these particular situations, it is possible to fix the gauge so as to guarantee that Eqs. (2.6) are verified.

Nonetheless, the decoupling of the unwanted degrees of freedom which is expressed through Eqs. (2.6) holds for any values of these parameters, and whether Eq. (2.9) is independent from Eq. (2.8). In the case it is independent, then the conditions (2.6) follow directly. In the case that Eq. (2.9) is not independent, the reasoning is a little more involved but the conclusion, namely the decoupling of the degree of freedom encoded in h, is kept the same. The reason is that if the trace of the equations of motion is zero as the result of Eq. (2.8), then the action is invariant under the transformations (2.2) but with $\gamma(x)$ a local parameter. Then the trace is a gauge degree of freedom, in the sense that $h = 0$ can be chosen by one of these transformations.

Therefore, one is led generically to the following picture, in the framework of which the conditions (2.6) are known as *transverse* and *traceless* conditions. The basic field variable is a constrained field h^{ab} that satisfies the equations of motion

$$\frac{1}{2}\Box h^{ab} = \lambda T^{ab}. \tag{2.10}$$

The theory is invariant under the gauge symmetry

$$h_{ab} \longrightarrow h'_{ab} = h_{ab} + \partial_a \xi_b + \partial_b \xi_a, \tag{2.11}$$

with the vector field ξ^a satisfying

$$\Box \xi^a, \qquad \partial_a \xi^a = 0. \tag{2.12}$$

It is not difficult to check that the traceless and transverse conditions are indeed preserved by these gauge transformations. These constraints in the definition of the field h^{ab} can be thought as the elimination of the scalar and vector representations of the Poincaré group (see Appendix I in [112]). Thus we recover dynamically the on-shell spin-2 representation of the Poincaré group, coupled to matter. As said above, and we will discuss later in detail, this on-shell representation has two natural extensions in which both the constraints (2.6) on the graviton field h^{ab} and on the generators (2.12) are dropped. We will consider the self-coupling problem for the on-shell picture of gravitons (Sec. 2.3) as well as for these two extensions, known as Fierz-Pauli theory (Sec. 2.4) and Weyl-transverse theory (Sec. 2.5), thus exhausting all the possibilities.

A few words about the corresponding representations of the Poincaré group are in order. The unitary representations of the Poincaré group as first classified by Wigner are determined by the value of the mass m and the eigenvalues of the so-called little group [106, 113, 107]. For a particle with mass $m \neq 0$ the little group is SO(3), so the corresponding label is the angular momentum j and one has $2j + 1$ states in each representation, corresponding to polarizations which range from $\sigma = -j$ to $\sigma = +j$ jumping in units. However, for a massless particle the little group is ISO(2) (the 2-dimensional Euclidean group) and only the states with polarizations $\sigma = \pm j$ are left. This means that massless particles with integer spin carry only two independent degrees of freedom. As linear representation space one would like to construct a tensor-field space using exclusively these degrees of freedom, but it is in this construction where gauge invariance appears inevitably intertwined with Lorentz invariance. A detailed analysis shows that one can always find a vector ξ^a such that the states with helicites $\sigma = 0, \pm 1$ are gauged away or, in other words, the corresponding components h^{00} and h^{0i}, $i = 1, 2, 3$ are set to zero while the remaining components are constrained so that there are two independent degrees of freedom. Another common choice to show this is the light-cone gauge [114].

Let us note that we are treating these degrees of freedom within a classical field theory. In practical terms, this means that our conclusions are expected to be applicable to the long-wavelength limit of theories in which a graviton propagates over Minkowski space in interaction with matter, independently of the ultraviolet completion of the theory (considering always theories with up to second derivatives of the fields; beyond that see [115]). As we have covered in detail in the previous chapter, the very notion of a Minkowski preferred background could be emergent in the sense of being applicable only below some characteristic energy scale, instead of a fundamental structure present in all regimes. If there is a regime in the theory in which gravity functions classically but the matter fields behave quantum-mechanically, so that a semiclassical description is meaningful, it is reasonable to expect that our conclusions would also apply to it as the self-coupling only occurs in the gravitational sector, still described by c-numbers.

2.2 The self-coupling problem

The discussion of the self-coupling problem of gravitons has a long history. The fundamental character of this problem has attracted a large number of physicists. In this section we shall briefly review this history, which we moreover use to place our contributions in context. We shall also derive one of the common (and central) features of different approaches: the iterative equations of the self-coupling problem.

2.2.1 A brief historical account

The geometric nature of general relativity is undoubtedly one of the factors that makes it conceptually beautiful. However, it also makes the theory very different from the formalism developed to describe the other fundamental forces we know about, namely the standard model of particle physics. While the latter is formulated as a quantum field theory in Minkowski spacetime, in general relativity there is no such a notion of preferred, immutable arena in which physics takes place. Instead this environment (spacetime) is also a dynamical object in its own right. This is arguably the root of the conceptual problems concerning the reconciliation between general relativity and quantum mechanics.

Trying to bridge the gap between these two disparate formalisms, Rosen showed that general relativity can be reinterpreted as a nonlinear field theory on Minkowski spacetime [116, 117]. Later Gupta proposed that a consistent theory of self-interacting gravitons should have precisely the structure of general relativity [28]. In brief, Gupta's idea is to start with a free graviton field in Minkowski spacetime, and then make it interact with the rest of fields. General considerations show that the spin of this field implies that this can be done only if the graviton field interacts with itself, making the overall theory nonlinear. Since one can always express general relativity plus matter as a nonlinear theory for the deviations of the metric with respect to some flat reference metric, there is wiggle room to reconcile both visions. Rosen's pioneering idea rapidly fade into oblivion as the geometric vision of gravity gained popularity, though as we will see it is certainly central to the construction to the self-coupling problem.

To make Gupta's program come to fruition, one must be able to determine the nature of the resulting nonlinear theory which arises from the self-coupling of the graviton field. The first subtle point is that the Lagrangian density of such a theory contains, in principle, infinite interaction terms which are obtained consecutively by an iterative process, so one should devise a way to manage them and show that this infinite series converges, at least for specific cases, to the Lagrangian density of general relativity. This question was indirectly addressed in the work of Kraichnan [118] and Feynman [27], but was finally settled by Deser [20]. To do that, he used specific variables which make the series finite, thus avoiding to perform the sum of an infinite series.

The second source of concern is the non-uniqueness of the construction as there are many and, in principle, inequivalent ways to make the graviton field self-interact. This was first raised by Huggins in his 1962 thesis [119]. The central point of his argument is that one needs more information to uniquely fix the stress-energy tensor of the graviton field to which it couples itself. Thus there are potentially many self-interacting theories and, as there is no control of those theories, it is not easy to conclude whether or not they are equivalent to general relativity. Recently, Padmanabhan has raised equivalent arguments [29]. In fact, the work presented in this chapter has been partially motivated by Padmanabhan's paper, a subsequent follow up by Butcher et al. [30], and the reply by Deser [21].

It is our goal in this chapter to consider the self-coupling problem from its very basics. We have found no place in which all the issues concerning this problem have been exposed at this level of detail (though [109] certainly presents a deep discussion on some of the most relevant points); this reason alone is sufficient for us to consider the content of this chapter a valuable contribution to the literature on the subject. Moreover, as we shall see our discussion will serve to settle the dust on some of the issues that have been raised before. On the one hand, our results confirm the concerns of Huggins first [119], and later Padmanabhan [29], in that the self-coupling problem by itself does not uniquely lead to general relativity unless further conditions are imposed along the process. Specifically, one needs to require that the gauge structure of the initial linear theory is preserved, although deformed, in the final outcome. This condition singles out general relativity but in a version that explicitly shows the underlying Minkowski spacetime, in the spirit of Rosen's flat-background bimetric theory; this feature is not optional, but rather a necessity rooted on the very basics of the construction. On the other hand, once the gauge preservation condition is applied, the entire construction can be taken to completion in a natural way using only flat spacetime notions, position which is defended by Deser in [21]. In fact, the presence of the Minkowski background structure permits to clearly separate the gauge transformations from invariance under changes of coordinates. As we discuss in Sec. 2.5, a further source of non-uniqueness comes from the linear representation of gravitons or, in other words, the way the information that is encoded in the corresponding irreducible representation of the Poincaré group is embedded into a tensorial description [111, 120, 108, 121, 122].

2.2.2 The iterative equations of the self-coupling problem

After this brief historical interlude, let us come back to our discussion around Eq. (2.10). We were arguing that, in order to couple the matter fields to the graviton field, we need to define a quantity T_{ab} that is symmetric, traceless and transverse on solutions. Then, we could write:

$$\frac{1}{2}\Box h_{ab} = \lambda T_{ab}. \tag{2.13}$$

While this is consistent at the level of the equations of motion only, a quite natural consistency condition is to impose that this equation can be obtained from an action. In physical terms, this amounts to the imposition of the law of action-reaction: not only matter fields act on the graviton field through Eq. (2.13), but the graviton field reacts on the matter fields. Up to now, the action of the theory consists of two terms,

$$\mathscr{A} := \mathscr{A}_{\mathrm{G},0} + \mathscr{A}_{\mathrm{M},0} = \int \mathrm{d}^4x\, \mathscr{L}_{\mathrm{G},0} + \int \mathrm{d}^4x\, \mathscr{L}_{\mathrm{M},0}, \tag{2.14}$$

where the variation of the first term leads to the left-hand side of Eq. (2.13), while $\mathscr{A}_{\mathrm{M},0}$ describes the matter excitations. It is important to keep in mind that some features of the following discussion will be better understood in terms of the corresponding Lagrangian densities, denoted generically by the calligraphic letter $\mathscr{L}$, instead of the actions themselves.

The right-hand side of Eq. (2.13) can be accommodated in the action by adding a suitable term to Eq. (2.14). Indeed, it appears as the corresponding term in the Euler-Lagrange equation with respect to of h^{ab} of the following piece to be added to the Lagrangian density,

$$\Delta\mathscr{L} := \lambda h^{ab} T_{ab}. \tag{2.15}$$

One immediately realizes that one cannot only use the stress-energy tensor of the free matter theory, that we call T^{M}_{ab} in the following. Indeed, for consistency one must use the total stress-energy tensor of the interacting theory. Had we started tentatively by adding a term $\lambda h^{ab}T^{\rm M}_{ab}$ to the Lagrangian density, this very term would have changed the matter stress-energy tensor, making necessary to add new energetic contributions to the Lagrangian density. This is a general property of the coupling with matter: as the coupling is done through the stress-energy tensor, the transverse condition for the graviton field implies that the total matter traceless stress-energy tensor should be divergenceless.[2] However, this would not be the case when interaction is switched on, as the matter fields no longer behave as an isolated system, being the energy transferred between them and the graviton field through tidal forces. The natural way to remedy this is to realize that the graviton field must also act as a source of itself (the charge/source of the graviton field is the energy and gravitons should possess energy), which leads us to the issue of the graviton self-coupling. Therefore, the iterative procedure has to act also in the spin-2 sector.

An important problem shows up when thinking about the stress-energy tensor of the spin-2 sector: there is no way of constructing a non-trivial conserved stress-energy tensor for the graviton field that is invariant under the gauge transformations (2.11) [123, 124]. By non-trivial we mean that it is not exactly zero for any solution. An indirect way of realizing it could be the Weinberg-Witten theorem [104, 107, 125], which explicitly forbids this possibility. Thus, one cannot associate a local notion of energy with the physical configurations in the free theory.[3] One can live with this fact if the theory is non-interacting, so that there is no operational way to define the energy-momentum of gravitons. Within an interaction scheme this is untenable. We shall see that this feature is preserved in the nonlinear regime, which is reflected in the well-known property of general relativity of not having a well-defined notion of local energy.

So one is led quite naturally to the consideration of the stress-energy tensor of all the fields as the relevant object to which the graviton field couples. While on the one hand this observation is compelling, on the other hand it has several unpleasant features, coming mainly from its non-uniqueness. Let us first consider the specific prescription to obtain the stress-energy tensor. The canonical stress-energy tensor is the conserved current of any free field theory which is Poincaré invariant, associated with invariance under translations:

$$\Theta^a{}_b := \mathscr{L}_0\delta^a_b - \frac{\delta\mathscr{L}_0}{\delta(\nabla_a\psi^\alpha)}\nabla_b\psi^\alpha, \qquad \nabla_a\Theta^a{}_b = 0. \tag{2.16}$$

Here $\mathscr{L}_0$ is the free Lagrangian density of both spin-2 and matter fields, collectively denoted by ψ^α, with α numbering the different matter fields.

However, direct use of this quantity is not possible: in general, the fully covariant or contravariant counterparts of Eq. (2.16) are not symmetric. But we can exploit the ambiguity in the addition of identically conserved tensors, the so-called Belinfante-Rosenfeld terms, to obtain a symmetric stress-energy tensor which leads to the same conserved quantities [109]. This symmetric stress-energy tensor is not unique: one can still add identically

[2] It can be seen that when the transverse condition is relaxed, it is the resulting gauge invariance of the theory the responsible for this feature [109].

[3] Something equivalent happens in non-abelian Yang-Mills theory: one cannot find a Lorentz-covariant conserved current which is also gauge invariant.

conserved tensors keeping the symmetric character. All the manipulations that follow in this chapter could be performed by directly using the symmetrized versions of the canonical stress-energy tensor. Therefore, these manipulations in no way involve any conceptual curved spacetime notion. However, as shown by Belinfante and Rosenfeld [126, 127], these symmetric stress-energy tensors can be equivalently obtained by the simple Hilbert prescription,

$$T_{ab} := -\lim_{\gamma\to\eta} \frac{1}{\sqrt{-\gamma}} \frac{\delta \mathscr{A}_0[\gamma]}{\delta \gamma^{ab}}, \tag{2.17}$$

where the flat metric η_{ab} in $\mathscr{A}_0$ as defined in Eq. (2.14) has been replaced by an auxiliary (generally curved) metric γ_{ab}, being γ^{ab} its inverse.

To apply this procedure one first needs to write down the action (2.14) in curvilinear coordinates in flat space, and then generalize it to curved space. It is in this second step where the ambiguities show up in Hilbert's prescription. In practice, the ambiguities in the stress-energy tensor appear now as the addition of non-minimal couplings of the physical fields to the auxiliary metric γ_{ab}. In fact, as we will see later these non-minimal couplings can be understood as surface terms in the original free action. We will show that these different choices of stress-energy tensor as the source of Eq. (2.13) lead to different solutions to the self-coupling problem. Let us stress again that here we use Hilbert's prescription as a mere calculational device and insist that no conceptual curved spacetime notion is used throughout the calculations. One can obtain the very same result (2.17) by using the canonical prescription in Eq. (2.16) and adding suitable Belinfante terms. Moreover, the very precise form of the source is selected by the iterative equations of the self-coupling problem and the zeroth-order Lagrangian density only, with no additional input. Padmanabhan's objections in this regard [29] are therefore not well founded.

Now that we have discussed the relevant properties of the stress-energy tensor, we shall move to the fundamental point in this section and derive the self-interacting equations of motion. Let us denote the coupling constant by λ. This can be done if we add a term $\lambda\mathscr{A}_1$ of order $\mathscr{O}(\lambda)$ in the action, such that:

$$\frac{\delta \mathscr{A}_1}{\delta h^{ab}} = \lim_{\gamma\to\eta} \frac{\delta \mathscr{A}_0[\gamma]}{\delta \gamma^{ab}}. \tag{2.18}$$

As noticed by Gupta [28], this additional term of order $\mathscr{O}(\lambda)$ in the action would modify the definition of the source by a term of order $\mathscr{O}(\lambda^2)$, which implies that we need to contemplate a term $\lambda^2\mathscr{A}_2$ in the action. This is the iterative procedure we want to solve for. It will generate an action of the form[4]

$$\mathscr{A} := \mathscr{A}_0 + \mathscr{A}_{\rm I} = \sum_{n=0}^{\infty} \lambda^n \mathscr{A}_n, \tag{2.19}$$

[4]Notice that there is no compulsory reason for this series to be infinite. There are two examples in the literature of this kind of series: the first one is the trivial one, in which one solves for the matter part only, therefore with no self-interactions (see next section). This series is infinite. The only example of a self-interacting series is the one constructed by Deser [20] which is finite, with only $\mathscr{A}_1 \neq 0$. Nevertheless, in our discussion we are going to consider always infinite series.

where we have considered the decomposition of the nonlinear part of the action $\mathscr{A}_{\rm I}$ in terms of the set of partial actions $\{\mathscr{A}_n\}_{n=1}^{\infty}$. Given the complete action (2.19), we can obtain its stress-energy tensor by following again the Hilbert prescription. $\mathscr{A}_{\rm I}$ is then fixed by the requirement of leading to this very stress-energy tensor as the source of the equations of motion:

$$\frac{\delta \mathscr{A}_{\rm I}}{\delta h^{ab}} = \lambda \lim_{\gamma\to\eta} \frac{\delta(\mathscr{A}_0 + \mathscr{A}_{\rm I})}{\delta \gamma^{ab}}. \tag{2.20}$$

One just needs to expand $\mathscr{A}_{\rm I} = \sum_{n=1}^{\infty} \lambda^n \mathscr{A}_n$ and compare different orders in the coupling constant λ to obtain the following set of iterative equations:

$$\frac{\delta \mathscr{A}_n}{\delta h^{ab}} = \lim_{\gamma\to\eta} \frac{\delta \mathscr{A}_{n-1}}{\delta \gamma^{ab}}, \qquad n \geq 1. \tag{2.21}$$

These are the fundamental equations of the self-coupling problem. We shall regard these equations, or Eq. (2.20), as a differential equation for the complete action $\mathscr{A}$. This analogy is worth keeping in mind for two reasons. On the one hand, different initial conditions specify different solutions. In this framework, the initial condition is selected by the specific form of the Lagrangian density $\mathscr{L}_0 = \mathscr{L}_{\rm G,0} + \mathscr{L}_{\rm M,0}$. On the other hand, it is much easier to check that a specific function is a solution rather than solving from scratch the set of equations (2.21), a feature that we shall eventually exploit.

2.2.3 Matter and gravitational sectors

We can split the complete action that results from the self-coupling problem as $\mathscr{A} = \mathscr{A}_{\rm G} + \mathscr{A}_{\rm M}$. Moreover, being the iterative equations of the self-coupling problem (2.21) linear, we can solve for these two parts independently. The integration of the iterative equations for the matter part is straightforward, so that the resulting form of the matter sector $\mathscr{A}_{\rm M}$ can be easily written down. By construction, the matter part on the right-hand side of Eq. (2.18) is independent of h^{ab} at the lowest order, linear at first order, and so on, making the integration of this part of the equation trivial. The resulting action is obtained as a Taylor series which can be summed. The formal result of this sum is the free matter action expressed in terms of a curved metric, $\mathscr{A}_{\rm M,0}[g]$ with $g^{ab} := \eta^{ab} + \lambda h^{ab}$ in the case of general relativity [20, 29] (or a suitable combination with fixed determinant in the case of Weyl-transverse gravity). Notice that non-minimal couplings to matter are not ruled out by any consistency condition, so minimal coupling to the physical metric in the resulting matter action does not necessarily hold.

The same recipe cannot be used to deal with the gravitational sector $\mathscr{A}_{\rm G,0}$. Doing so ignores the nonlinear character of the problem, and therefore leads to the considerations of actions that *are not* solutions of the self-coupling problem, such as the ones discussed by Padmanabhan in [29]. Therefore it is the gravitational self-interacting part of the iterative procedure, namely $\mathscr{A}_{\rm G}$, the part that has to be handled carefully. In the following we shall deal only with this sector, so that we will drop the subscript, just writing $\mathscr{A}_{\rm G,0} = \mathscr{A}_0$ and $\mathscr{A}_{\rm G} = \mathscr{A}$. In previous analysis, the infinite set of iterative equations (2.21) for the gravitational part have been indirectly addressed. For example, demanding the preservation of the original gauge invariance in the form of nonlinear diffeomorphism invariance permits to write down unambiguously the Einstein-Hilbert action, that can be then shown to be a solution of these equations [30]. However, here we would like to understand the interplay between the preservation of this symmetry and the iterative self-coupling procedure, instead of taking its existence as an assumption from the beginning. Even in [20], a change of variables is used that render the series finite for a specific solution (corresponding to general relativity); the way in which the author conducted the discussion hides nevertheless the non-uniqueness of the overall construction, as we shall discuss in detail.

2.3 The minimal ("on-shell") construction

Solving the iterative equations from scratch is far from straightforward. Although these equations have been known since at least the 1950s [28], there is no serious attempt in the literature of dealing with them in all their generality. For this reason, we shall first consider a simplified setting in which the simplest description for the graviton field is taken. Even if presenting some unpleasant features due to the imposition of constraints at the linear level, this limited framework will however present all the relevant features of more interesting cases to be discussed later. This discussion will therefore give us invaluable insight into the structure of the space of solutions of these equations even for more complicated scenarios.

2.3.1 Solving the iterative equations

So let us come back to the field-theoretical discussion of gravitons. Working in a covariant fashion with respect to changes of coordinates in Minkowski spacetime will prove useful to distinguish clearly between invariance under changes of coordinates and gauge invariance in the resulting nonlinear theory. Let us therefore summarize the main ingredients of our previous discussion but using this notation. At the linear level, we use tensorial objects h^{ab} such that they are traceless and transverse,

$$\eta_{ab}h^{ab} = 0, \qquad \nabla_b h^{ab} = 0, \tag{2.22}$$

satisfying the equations of motion

$$\Box h^{ab} = 0. \tag{2.23}$$

Moreover, any h^{ab} and h'^{ab} will represent the same physical configuration if they are related by a gauge transformation,

$$h'^{ab} = h^{ab} + \eta^{ac}\nabla_c\xi^b + \eta^{bc}\nabla_c\xi^a, \tag{2.24}$$

with generators verifying the conditions

$$\nabla_a\xi^a = 0, \qquad \Box\xi^a = 0. \tag{2.25}$$

These gauge transformations and the transformations associated with general change of coordinates with generators $\tilde{\xi}^a$,

$$h'^{ab} = h^{ab} + \tilde{\xi}^c \nabla_c h^{ab} + h^{ac} \nabla_c \tilde{\xi}^b + h^{bc} \nabla_c \tilde{\xi}^a, \tag{2.26}$$

are completely different from each other: the space of generators is different and so is their implementation in the symmetry transformation. Moreover, the latter transformation affects the coordinates and the rest of fields. The reader will notice that, through the following calculations, we always keep using the object h^{ab} and never its covariant counterpart $h_{ab} := \eta_{ac}\eta_{bd}h^{cd}$. This simplifies some steps which involve taking variational derivatives with respect to an auxiliary metric γ_{ab} after the replacement $\eta_{ab} \to \gamma_{ab}$ in the action.

Now the particular Lagrangian density coming from the restriction of the Lagrangian density (2.1) to transverse and traceless graviton fields has only two terms. Indeed, there are only two contractions of $\nabla_a h^{bc} \nabla_r h^{st}$ with metric objects (η_{ab}, η^{ab} and δ^a_b) that are not identically zero by virtue of the traceless and transverse conditions (2.22). Hence the corresponding Lagrangian density has the form

$$\mathscr{L}_0 := -\frac{1}{4}\eta^{ar}\eta_{bs}\eta_{ct}\nabla_a h^{bc}\nabla_r h^{st} + c_2 \delta^a_t \delta^r_c \eta_{bs} \nabla_a h^{bc} \nabla_r h^{st}. \tag{2.27}$$

The second term is equivalent to a total divergence (leading to a surface term) because of the transverse condition in Eq. (2.22), so that it does not affect the form of the equations of motion. However, as we shall see its presence can affect the definition of the source of the self-interacting equations, so we will keep it explicitly in the following discussion. Let us introduce a bit of notation to conveniently write the free action as:

$$\mathscr{A}_0[\eta, \theta] := \frac{1}{4} \int \mathrm{d}\mathscr{V}_\eta \, M^{ar}{}_{bcst}(\eta, \theta) \nabla_a h^{bc} \nabla_r h^{st}, \tag{2.28}$$

where $\mathrm{d}\mathscr{V}_\eta := \mathrm{d}^4 x \sqrt{-\eta}$ is the Minkowski volume element and the Lorentz tensor $M^{ar}{}_{bcst}(\eta, \theta)$ is defined as

$$M^{ar}{}_{bcst}(\eta, \theta) := \theta \left[\eta_{b(s}\delta^a_{t)}\delta^r_c + \eta_{c(s}\delta^a_{t)}\delta^r_b \right] - \eta^{ar}\eta_{b(s}\eta_{t)c}. \tag{2.29}$$

The tensorial quantity $M^{ar}{}_{bcst}(\eta, \theta)$ is symmetric under $b \leftrightarrow c$, $s \leftrightarrow t$ and $(a, b, c) \leftrightarrow (r, s, t)$, as one can directly check from its definition. When used in the action we do not need to worry about these symmetries because it is contracted with an object, $\nabla_a h^{bc} \nabla_r h^{st}$, which already has these symmetries. However, to solve the iterative equations of the self-coupling problem it will be necessary to use its symmetric form as it appears in Eq. (2.29). The parameter θ, directly proportional to c_2, controls the surface terms we are considering in the free action. One can easily check that this action is invariant up to a surface term under the gauge transformations (2.24) with generators satisfying the conditions in Eq. (2.25). In fact, the case $\theta = 1$ is special in the sense that one could drop the second condition in Eq. (2.25) and these transformations are still a symmetry. Thus only the first condition in Eq. (2.25) is necessary when considering $\theta = 1$, which leads us to the minimal theory of gravitons considered in [128].

Let us start with the first-order iterative equation (2.18). We shall evaluate the right-hand side of this equation and then integrate the functional form to obtain the corresponding left-hand side. The first step is then to apply Hilbert's prescription to obtain the source of the equations of motion. To do that we have to extend the action (2.28) so that it is evaluated on a general curved metric. It is in this step where the ambiguities in the addition of non-minimal couplings can arise. We will deal with this ambiguity in the following section, thus making here the simplest choice. This corresponds to the minimal coupling prescription:

$$\mathscr{A}_0[\gamma,\theta] = \frac{1}{4}\int \mathrm{d}\mathscr{V}_\gamma\, M^{ar}{}_{bcst}(\gamma,\theta)\nabla'_a h^{bc}\nabla'_r h^{st}. \tag{2.30}$$

Here ∇' is the covariant derivative with respect to the auxiliary metric γ_{ab} and $\mathrm{d}\mathscr{V}_\gamma := \mathrm{d}^4x\sqrt{-\gamma}$ the corresponding volume element. Notice also that we have changed the metric in the argument of the tensor $M^{ar}{}_{bcst}(\eta,\theta)$ defined in Eq. (2.29).

Following Hilbert's prescription, we obtain the stress-energy tensor by performing variations on this metric, and then taking the limit back to flat space. Under such a variation, the action (2.30) changes as:

$$\delta\mathscr{A}_0[\gamma,\theta] = \frac{1}{4}\int \mathrm{d}^4x\,\delta[\sqrt{-\gamma}\,M^{ar}{}_{bcst}(\gamma,\theta)]\nabla'_a h^{bc}\nabla'_r h^{st} + \frac{1}{2}\int \mathrm{d}^4x\sqrt{-\gamma}\,M^{ar}{}_{bcst}(\gamma,\theta)\delta[\nabla'_a h^{bc}]\nabla'_r h^{st}. \tag{2.31}$$

The first term gives two contributions, one coming from the variation of the determinant and the other from the variation of $M^{ar}{}_{bcst}(\gamma,\theta)$. There are two possible ways of dealing with the former. The first one is to notice that the first-order equation must be traceless so the corresponding term is not going to contribute. This observation can be extended to all orders with the following recipe: do not change the measure $\mathrm{d}\mathscr{V}_\eta$ in the partial actions $\mathscr{A}_n$ when writing them in terms of the auxiliary metric γ_{ab}. Although a departure from Hilbert's prescription, this alternative procedure leads to a sensible source to be used in the self-coupling procedure when the constraints on the field h^{ab} are taken into account. We shall follow this approach in this section. A second option is to proceed with no previous knowledge of the restrictions on h^{ab} and integrate the contribution coming from the variation of the determinant. The iterative equations (2.21) are linear, so that we only need to add the corresponding contributions obtained in this way to the result of the calculations of this section. We will study in the next section the result of this procedure, demonstrating that this is just an operational choice that does not affect the physical results at the end of the day, namely when the constraints on the field h^{ab} are considered.

Let us now deal with the second term of Eq. (2.31). There we have the difference of two Levi-civita connections associated with γ_{ab} and $\gamma_{ab}+\delta\gamma_{ab}$, respectively. This difference is characterized (see, e.g., [129]) by the tensor

$$C'^{b}{}_{ad} := \frac{1}{2}(\gamma^{be}+\delta\gamma^{be})\nabla'_\mu(\gamma_{\nu\rho}+\delta\gamma_{\nu\rho})D^{\mu\nu\rho}{}_{aed}, \tag{2.32}$$

where

$$D^{\mu\nu\rho}{}_{aed} := \delta^{\mu}_{a}\delta^{(\nu}_{d}\delta^{\rho)}_{e} + \delta^{\mu}_{d}\delta^{(\nu}_{a}\delta^{\rho)}_{e} - \delta^{\mu}_{e}\delta^{(\rho}_{a}\delta^{\nu)}_{d}. \tag{2.33}$$

Then one can see that the variation $\delta[\nabla'_a h^{bc}]$ is given, at first order, by:

$$\delta C'^{(b}_{ad} h^{c)d} = \gamma^{e(b} h^{c)d} \nabla'_{\mu} \delta\gamma_{\nu\rho} D^{\mu\nu\rho}{}_{aed}. \tag{2.34}$$

The notation $\delta C'^{a}_{bc}$ here means that we only take the terms in Eq. (2.32) that are linear in the variations $\delta\gamma_{ab}$. If we integrate by parts, the contribution of these terms equals to

$$\begin{aligned} &-\frac{1}{2}\int \mathrm{d}^4x\, \delta\gamma_{\nu\rho}\gamma^{e(b} D^{\mu\nu\rho}{}_{aed} M^{ar}{}_{bcst}(\gamma,\theta)\nabla'_{\mu}[h^{c)d}\nabla'_r h^{st}] \\ &= \frac{1}{2}\int \mathrm{d}^4x\, \delta\gamma^{pq}\gamma_{p\nu}\gamma_{q\rho}\gamma^{de} D^{a\nu\rho}{}_{\mu e(b} M^{\mu r}{}_{c)dst}(\gamma,\theta)\nabla'_a(h^{bc}\nabla'_r h^{st}). \end{aligned} \tag{2.35}$$

Therefore taking into account the contributions of the two terms in Eq. (2.31), the corresponding source takes the form:

$$\begin{aligned} T_{pq} :=& -\frac{1}{4}\left.\frac{\delta M^{ar}{}_{bcst}(\gamma,\theta)}{\delta\gamma^{pq}}\right|_{\gamma\to\eta} \nabla_a h^{bc}\nabla_r h^{st} \\ &- \frac{1}{2}\eta_{p\nu}\eta_{q\rho}\eta^{de} D^{a\nu\rho}{}_{\mu e(b} M^{\mu r}{}_{c)dst}(\eta,\theta)\nabla_a(h^{bc}\nabla_r h^{st}). \end{aligned} \tag{2.36}$$

Notice that this expression contains second derivatives of the graviton field. It is important to notice also that it naturally splits into two kinds of terms, proportional to $\nabla_a h^{bc}\nabla_r h^{st}$ and $\nabla_a(h^{bc}\nabla_r h^{st})$, respectively. As it stands, it is symmetric under the exchanges $p \leftrightarrow q$ and $b \leftrightarrow c$.

The objective now is to find a term in the action $\lambda\mathscr{A}_1$ whose variation with respect to h^{ab} gives the desired source term [Eq. (2.36)]. The most general expression that contains no more than two derivatives of the graviton field can be always written as

$$\frac{1}{4}\int \mathrm{d}\mathscr{V}_{\eta}\, N^{ar}{}_{bcstpq}(\eta) h^{pq}\nabla_a h^{bc}\nabla_r h^{st}. \tag{2.37}$$

Then taking the functional derivative with respect to h^{ab} we obtain

$$\frac{1}{4}\int \mathrm{d}\mathscr{V}_{\eta}\left[N^{ar}{}_{bcstpq}(\eta)\nabla_a h^{bc}\nabla_r h^{st} - 2N^{ar}{}_{pqstbc}(\eta)\nabla_a(h^{bc}\nabla_r h^{st})\right]\delta h^{pq}. \tag{2.38}$$

Now we get two equations coming from the comparison of the coefficients accompanying the two independent combinations $\nabla_a h^{bc}\nabla_r h^{st}$ and $\nabla_a(h^{bc}\nabla_r h^{st})$ in both Eqs. (2.36) and (2.38):

$$N^{ar}{}_{bcstpq}(\eta) = \left.\frac{\delta M^{ar}{}_{bcst}(\gamma,\theta)}{\delta\gamma^{pq}}\right|_{\gamma\to\eta}, \tag{2.39}$$

and

$$-N^{ar}{}_{bcstpq}(\eta) = \eta_{p\nu}\eta_{q\rho}\eta^{de} D^{a\nu\rho}{}_{\mu e(b} M^{\mu r}{}_{c)dst}(\eta,\theta). \tag{2.40}$$

The first equation provides the form of the first-order action (2.37). The second equation then becomes a consistency condition that must be satisfied for the whole procedure to be consistent:

$$-\left.\frac{\delta M^{ar}{}_{pqst}(\gamma,\theta)}{\delta\gamma^{bc}}\right|_{\gamma\to\eta} = \eta_{p\nu}\eta_{q\rho}\eta^{de}D^{a\nu\rho}{}_{\mu e(b}M^{\mu r}{}_{c)dst}(\eta,\theta). \tag{2.41}$$

It is this equation which imposes restrictions to the solutions of the iterative equations that, in fact, select $\theta = 1$. To obtain this condition on θ, let us notice that the right-hand side of Eq. (2.41) can be written as

$$\eta_{pb}M^{ar}{}_{qcst}(\eta,\theta) + \delta^a_b\eta_{p\mu}M^{\mu r}{}_{qcst}(\eta,\theta) - \eta_{pb}\eta_{q\mu}\eta^{ad}M^{\mu r}{}_{dcst}(\eta,\theta), \tag{2.42}$$

where we must impose a symmetrization under the exchange of indices $p \leftrightarrow q$ and $b \leftrightarrow c$. It is useful to write this expression explicitly by using Eq. (2.29),

$$\begin{aligned}
&\theta\,\eta_{pb}\left[\eta_{q(s}\delta^a_{t)}\delta^r_c + \eta_{c(s}\delta^a_{t)}\delta^r_q\right] - \eta_{pb}\eta^{ar}\eta_{q(s}\eta_{t)c}\\
&+\theta\,\delta^a_b\left[\eta_{p(s}\eta_{t)q}\delta^r_c + \eta_{p(s}\eta_{t)c}\delta^r_q\right] - \delta^a_b\eta_{q(s}\eta_{t)c}\delta^r_p\\
&-\theta\left[\eta^{ar}\eta_{pb}\eta_{q(s}\eta_{t)c} + \eta_{pb}\eta_{q(s}\delta^a_{t)}\delta^r_c\right] + \delta^r_q\delta^a_{(s}\eta_{pb}\eta_{t)c},
\end{aligned} \tag{2.43}$$

and symmetrize this equation with respect to $p \leftrightarrow q$, so it can be simplified to:

$$\begin{aligned}
&\theta\left[\eta_{pb}\eta_{c(s}\delta^a_{t)}\delta^r_q + \eta_{qb}\eta_{c(s}\delta^a_{t)}\delta^r_p\right] + \delta^a_b\delta^r_c\eta_{p(s}\eta_{t)q}\\
&-\frac{\theta+1}{2}\eta^{ar}\left[\eta_{pb}\eta_{c(s}\eta_{t)q} + \eta_{qb}\eta_{p(s}\eta_{t)c}\right]\\
&+\frac{\theta-1}{2}\delta^a_b\left[\eta_{p(s}\eta_{t)c}\delta^r_q + \eta_{q(s}\eta_{t)c}\delta^r_p\right].
\end{aligned} \tag{2.44}$$

This equation must be compared with the left-hand side of Eq. (2.41), i.e., with

$$\begin{aligned}
&-\left.\frac{\delta M^{ar}{}_{pqst}(\gamma,\theta)}{\delta\gamma^{bc}}\right|_{\gamma\to\eta} = \theta\left[\eta_{pb}\eta_{c(s}\delta^a_{t)}\delta^r_q + \eta_{qb}\eta_{c(s}\delta^a_{t)}\delta^r_p\right]\\
&+\delta^a_b\delta^r_c\eta_{p(s}\eta_{t)q} - \eta^{ar}\left[\eta_{pb}\eta_{c(s}\eta_{t)q} + \eta_{qb}\eta_{p(s}\eta_{t)c}\right],
\end{aligned} \tag{2.45}$$

which must be still symmetrized under the exchange $b \leftrightarrow c$. A direct comparison of these equations tells us that the only solution of Eq. (2.41) is given by $\theta = 1$.

In this way, we have shown how to integrate the first-order iterative equation (2.18). The result is:

$$\begin{aligned}
&\mathscr{A}_0 + \lambda\mathscr{A}_1 + \mathscr{O}(\lambda^2)\\
&= \frac{1}{4}\int \mathrm{d}\mathscr{V}_\eta\, M^{ar}{}_{bcst}(\eta+\lambda h, \theta = 1)\nabla_a h^{bc}\nabla_r h^{st} + \mathscr{O}(\lambda^2).
\end{aligned} \tag{2.46}$$

Now that we have worked out the first order in detail, let us show that the result to the full problem that can be anticipated from this discussion is in fact the correct result:

$$\mathscr{A} = \frac{1}{4}\int \mathrm{d}\mathscr{V}_\eta \, M^{ar}{}_{bcst}(\eta + \lambda h, \theta = 1)\nabla_a h^{bc}\nabla_r h^{st}. \tag{2.47}$$

Decomposing this ansatz in partial actions, $\mathscr{A} = \sum_{n=0}^{\infty}\lambda^n \mathscr{A}_n$, with

$$\mathscr{A}_n = \frac{1}{4n!}\int \mathrm{d}^4x \; \frac{\delta^n M^{ar}{}_{bcst}(\gamma, \theta = 1)}{\delta\gamma^{pq}\delta\gamma^{de}\dots}\bigg|_{\gamma\to\eta} \nabla_a h^{bc}\nabla_r h^{st} h^{pq} h^{de}\dots, \tag{2.48}$$

and applying the iterative equations (2.21) to this sequence, we find the consistency conditions:

$$\begin{aligned} & n\,\eta_{p\nu}\eta_{q\rho}\eta^{fg} D^{a\nu\rho}{}_{\mu g(b} \, \frac{\delta^{n-1} M^{\mu r}{}_{c)fst}(\gamma, \theta = 1)}{\delta\gamma^{de}\dots}\bigg|_{\gamma\to\eta} \\ & = -\frac{\delta^n M^{ar}{}_{pqst}(\gamma, \theta = 1)}{\delta\gamma^{bc}\delta\gamma^{de}\dots}\bigg|_{\gamma\to\eta}. \end{aligned} \tag{2.49}$$

Notice the symmetrization on the pair of indices (b, c). To work better with this expression, we can avoid at first to evaluate it in the limit $\gamma \to \eta$, working thus with the equation

$$n\,\gamma_{p\nu}\gamma_{q\rho}\gamma^{fg} D^{a\nu\rho}{}_{\mu g(b}\frac{\delta^{n-1} M^{\mu r}{}_{c)fst}(\gamma, \theta = 1)}{\delta\gamma^{de}\dots} = -\frac{\delta^n M^{ar}{}_{pqst}(\gamma, \theta = 1)}{\delta\gamma^{bc}\delta\gamma^{de}\dots}, \tag{2.50}$$

which can be viewed as a differential equation with an initial condition imposed in flat space. In fact, if we drop the indices we can write it schematically as

$$n\,\Theta(\gamma)\frac{\partial^{n-1} M(\gamma, \theta = 1)}{\partial\gamma^{n-1}} = -\frac{\partial^n M(\gamma, \theta = 1)}{\partial\gamma^n}, \tag{2.51}$$

with

$$\Theta \sim (\gamma)^{-1}. \tag{2.52}$$

Up to now, we have only shown that Eq. (2.41) is valid, which in this simplified notation becomes

$$\Theta(\gamma) M(\gamma, \theta = 1) = -\frac{\partial M(\gamma, \theta = 1)}{\partial\gamma}. \tag{2.53}$$

Thus to show by induction that Eq. (2.47) represents in fact the solution to the iterative problem we only need, as we have already proved that it holds for $n = 1$, to show that $\Theta(\gamma)$ verifies the differential equation

$$\frac{\partial\Theta(\gamma)}{\partial\gamma} = -\Theta^2(\gamma), \tag{2.54}$$

as it is indeed the case [recall Eq. (2.52)].

Coming back to the full equations, one has:

$$-\frac{\delta^{n+1}M^{ar}{}_{pqst}(\gamma,\theta=1)}{\delta\gamma^{uv}\delta\gamma^{bc}\delta\gamma^{fg}\dots} = n\,\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}D^{a\nu\rho}{}_{\mu e(b}\frac{\delta^{n}M^{\mu r}{}_{c)dst}(\gamma,\theta=1)}{\delta\gamma^{uv}\delta\gamma^{fg}\dots}$$
$$+\,n\,D^{a\nu\rho}{}_{\mu e(b}\frac{\delta^{n-1}M^{\mu r}{}_{c)dst}(\gamma,\theta=1)}{\delta\gamma^{fg}\dots}\frac{\delta}{\delta\gamma^{uv}}\left(\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}\right)$$
$$+\,\{(u,v)\leftrightarrow(b,c)\}. \tag{2.55}$$

Then the consistency condition with the induction can be read as:

$$D^{a\nu\rho}{}_{\mu e(b}\frac{\delta^{n-1}M^{\mu r}{}_{c)dst}(\gamma,\theta=1)}{\delta\gamma^{fg}\dots}\frac{\delta}{\delta\gamma^{uv}}\left(\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}\right)$$
$$+\,D^{a\nu\rho}{}_{\mu e(u}\frac{\delta^{n-1}M^{\mu r}{}_{v)dst}(\gamma,\theta=1)}{\delta\gamma^{fg}\dots}\frac{\delta}{\delta\gamma^{bc}}\left(\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}\right)$$
$$=\frac{1}{n}\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}D^{a\nu\rho}{}_{\mu e(b}\frac{\delta^{n}M^{\mu r}{}_{c)dst}(\gamma,\theta=1)}{\delta\gamma^{uv}\delta\gamma^{fg}\dots}$$
$$+\,\frac{1}{n}\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}D^{a\nu\rho}{}_{\mu e(u}\frac{\delta^{n}M^{\mu r}{}_{v)dst}(\gamma,\theta=1)}{\delta\gamma^{bc}\delta\gamma^{fg}\dots}. \tag{2.56}$$

Because of the symmetrization, we can take only one of the terms in each side of the last equation, thus obtaining the equation:

$$D^{a\nu\rho}{}_{\mu e(b}\frac{\delta^{n-1}M^{\mu r}{}_{c)dst}(\gamma,\theta=1)}{\delta\gamma^{fg}\dots}\frac{\delta}{\delta\gamma^{uv}}\left(\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}\right)$$
$$=\frac{1}{n}\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}D^{a\nu\rho}{}_{\mu e(u}\frac{\delta^{n}M^{\mu r}{}_{v)dst}(\gamma,\theta=1)}{\delta\gamma^{bc}\delta\gamma^{fg}\dots}$$
$$=-\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}D^{a\nu\rho}{}_{\mu e(u}\gamma_{d\alpha}\gamma_{v)\beta}\gamma^{\gamma\delta}D^{\mu\alpha\beta}{}_{\theta\delta(b}\frac{\delta^{n-1}M^{\theta r}{}_{c)\gamma st}(\gamma,\theta=1)}{\delta\gamma^{fg}}. \tag{2.57}$$

In the last line we have used Eq. (2.50). So we arrive at the equation:

$$D^{a\nu\rho}{}_{\mu eb}\frac{\delta}{\delta\gamma^{st}}\left(\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}\right)=-\gamma_{p\nu}\gamma_{q\rho}\gamma^{\gamma e}D^{a\nu\rho}{}_{\theta e(s}\gamma_{\gamma\alpha}\gamma_{t)\beta}\gamma^{d\delta}D^{\theta\alpha\beta}{}_{\mu\delta b}, \tag{2.58}$$

where we have changed the free indices to avoid potential confusions. This is the equation represented schematically by (2.54). The reader can find in Appendix 2.6 the demonstration that this algebraic relation is indeed true and, therefore, the induction proof is finished.

As the construction of the iterative series relies ultimately in the solution of a system that is formally equivalent to a set of ordinary differential equations schematically represented by Eq. (2.53), with an initial condition posed in flat space, the solution is unique.

This is given by

$$\mathscr{A}[g,\theta=1] = \frac{1}{4}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\int \mathrm{d}\mathscr{V}_\eta \left.\frac{\delta^n M^{ai}{}_{bcjk}(\gamma,\theta=1)}{\delta\gamma^{pq}\delta\gamma^{st}\dots}\right|_{\gamma\to\eta}\nabla_a h^{bc}\nabla_i h^{jk}h^{pq}h^{st}\dots$$
$$= \frac{1}{4}\int \mathrm{d}\mathscr{V}_\eta\, M^{ai}{}_{bcjk}(\eta+\lambda h,\theta=1)\nabla_a h^{bc}\nabla_i h^{jk}$$
$$= \frac{1}{4\lambda^2}\int \mathrm{d}\mathscr{V}_\eta\, M^{ai}{}_{bcjk}(g,\theta=1)\nabla_a g^{bc}\nabla_i g^{jk}, \tag{2.59}$$

where we have defined the field

$$g^{ab} := \eta^{ab} + \lambda h^{ab}. \tag{2.60}$$

It is important, in order to avoid confusions, to keep in mind that ∇ is the covariant derivative compatible with η_{ab}.

2.3.2 Non-minimal couplings and surface terms

In this section we shall elucidate the effect of allowing contributions to the stress-energy tensor coming from non-minimal couplings or, what is equivalent in this case, covariant surface terms. To avoid confusions with the surface terms proportional to the parameter θ we shall nevertheless refer to these additional terms unambiguously as non-minimal couplings. These terms fully parametrize the ambiguity inherent to the definition of the source in the equations of motion. They must be considered for the sake of completeness when the free action in flat space is generalized to a general metric space in terms of the auxiliary metric γ_{ab}.

Non-minimal couplings are defined as scalar quantities which can be written in terms of the auxiliary metric γ_{ab} and the graviton field when using Hilbert's prescription to obtain the stress-energy tensor, which vanish in the flat-space limit. The most general form of these terms, as they would be added to Eq. (2.30) for instance, is given by:

$$\int \mathrm{d}\mathscr{V}_\gamma A^r{}_{bcst}(\gamma,\nabla\gamma)h^{bc}\nabla_r h^{st}. \tag{2.61}$$

The function $A^r{}_{bcst}(\gamma,\nabla\gamma)$ must be proportional to $\nabla\gamma$, namely $A^r{}_{bcst} = \nabla_a[B^{ar}{}_{bcst}(\gamma)]$. Using the flat covariant derivative ∇ guarantees that these terms vanish in the flat-space limit. We have also restricted them with the condition of leading to contributions to the stress-energy tensor which are quadratic in the derivatives of the graviton field. These contributions are obtained by varying this expression with respect to γ^{ab} (after integrating by parts) and then taking the flat-space limit.

The reader could find strange the form (2.61) we associate with non-minimal couplings. While the usual representation uses curvature-related tensor quantities, as the Riemann tensor, constructed from specific combinations of the auxiliary metric and its ordinary derivatives $\partial\gamma$, in Eq. (2.61) we are using arbitrary scalar combinations of the metric and its covariant derivatives $\nabla\gamma$. To do that we are exploiting the fact that we have a Minkowski reference metric, which permits us to easily construct scalar quantities that contain the covariant derivatives of the auxiliary metric with respect to the flat reference metric. Let us consider as an example the Riemann tensor: given a generic decomposition of a metric γ_{ab} in the form $\gamma_{ab} = q_{ab}+\epsilon_{ab}$, one can always write its Riemann tensor, ${R^a}_{bcd}(\gamma)$, as

$$ {R^a}_{bcd}(\gamma) = {R^a}_{bcd}(q) + 2\bar{\nabla}_{[c}\bar{C}^a_{d]b} + 2\bar{C}^a_{e[c}\bar{C}^e_{d]b}. \tag{2.62} $$

In this expression, $\bar{C}^a_{bc}$ is the tensor which characterizes the difference between covariant derivatives with respect to the two metrics γ_{ab} and q_{ab}, respectively denoted by ∇' and $\bar{\nabla}$ (see for example [129], Eq. D7 adapted to our sign conventions). Now one can consider the special situation in which $q_{ab} = \eta_{ab}$ to realize that the Riemann tensor of γ_{ab} can be written as a particular case of the integrand in Eq. (2.61).

With this definition of the possible non-minimal couplings, it is not difficult to realize that the same effect can be reproduced by adding a covariant surface term instead. This term would have the following form, after writing the original action in terms of the auxiliary metric:

$$ \int \mathrm{d}\mathscr{V}_\gamma \nabla'_a[{S^{ar}}_{bcst}(\gamma)h^{bc}\nabla'_r h^{st}]. \tag{2.63} $$

As in the case of non-minimal couplings, this is the most general possible expression containing two covariant derivatives of the graviton field. Recall that ∇' is the covariant derivative associated with γ_{ab}.

The question now is whether the results we have obtained in the previous section could change because of the introduction of non-minimal couplings. In other words, we want to know whether there exists a different functional

$$ \mathscr{A}'_1 := \frac{1}{4}\int \mathrm{d}\mathscr{V}_\eta\, {O^{ar}}_{bcstpq}(\eta)h^{pq}\nabla_a h^{bc}\nabla_r h^{st}, \tag{2.64} $$

solution up to order $\mathscr{O}(\lambda)$ of the iterative procedure when certain additional terms in the stress-energy tensor are taken into account.

The effect of the non-minimal couplings would be to add some terms of the form $\nabla_a(h^{bc}\nabla_r h^{st})$ to the stress-energy tensor. Thus, the iterative equations give us two conditions, analogous to Eqs. (2.39) and (2.40): the first one is directly

$$ {O^{ar}}_{bcstpq} = \left.\frac{\delta {M^{ar}}_{bcst}(\gamma,\theta)}{\delta\gamma^{pq}}\right|_{\gamma\to\eta}, \tag{2.65} $$

as in the minimal coupling case, while the second one will notice the effect of non-minimal couplings, being changed to

$$ -{O^{ar}}_{bcstpq} = \eta_{p\nu}\eta_{q\rho}\eta^{de}{D^{a\nu\rho}}_{\mu e(b}{M^{\mu r}}_{c)dst}(\eta,\theta) + {\Delta^{ar}}_{bcstpq}, \tag{2.66} $$

where the term ${\Delta^{ar}}_{bcstpq}$ is the contribution coming from non-minimal couplings, parametrized as depicted in Eq. (2.61), or equivalently Eq. (2.63), in terms of which

$$ {\Delta^{ar}}_{bcstpq} = \frac{2}{\sqrt{\gamma}}\frac{\delta[\sqrt{-\gamma}{S^{ar}}_{bcst}(\gamma)]}{\delta\gamma^{pq}}. \tag{2.67} $$

Eqs. (2.65) and (2.66) must be understood as the conditions which permit the determination of the non-minimal couplings we need in order to make the self-coupling procedure consistent for different values of the parameter θ. Indeed, these equations uniquely determine the tensor $\Delta^{ar}{}_{bcstpq}$ for a given value of θ, which then fixes the non-minimal couplings (or covariant surface terms) through Eq. (2.67). We see that the addition of non-minimal couplings allows in principle to find solutions to the problem for $\theta \neq 1$. It is important that the necessary non-minimal couplings are not put by hand, but are *determined* by using the iterative equations (2.21) and the zeroth-order action (2.28) with θ arbitrary. This observation will be of importance when discussing the naturalness of the construction in Sec. (2.4).

Now Eq. (2.65) implies that the solution, if it exists, will be expressible as the first term of a Taylor expansion in λ of the free action displaced to $\eta^{ab} + \lambda h^{ab}$, for any value of θ. That is,

$$\mathscr{A}_0 + \lambda \mathscr{A}_1 + \mathscr{O}(\lambda^2) = \frac{1}{4} \int \mathrm{d}\mathscr{V}_\eta \, M^{ar}{}_{bcst}(\eta + \lambda h, \theta) \nabla_a h^{bc} \nabla_r h^{st} + \mathscr{O}(\lambda^2). \tag{2.68}$$

The complete iterative procedure would give place to the complete Taylor series in complete analogy with the minimal-coupling case (2.59). The best way to demonstrate that the analogue of Eqs. (2.65), (2.66) and (2.67) for the entire set of iterative equations (2.21) represent well-posed equations for the non-minimal couplings that are necessary for each value of θ is expanding the resulting actions in terms of the coupling constant λ as we sketch below, and explain in more detail in Sec. 2.4.

Let us consider now the issue of the variation of the volume element $\mathrm{d}\mathscr{V}_\gamma$ or, in other words, of the factor $\sqrt{-\gamma}$ in the partial actions $\mathscr{A}_n[\gamma]$. The only difference in the integration of the first-order iterative equation is that the variation of the determinant $\delta\sqrt{-\gamma}$ must be taken into account in Eq. (2.31). This implies that Eq. (2.65) is modified to include this variation, and therefore the necessary non-minimal couplings as determined in conjunction with Eq. (2.66) would be different. If a solution exists, the measure in (2.68) as well as in the final action would be given by $\mathrm{d}\mathscr{V}_g := \mathrm{d}\mathscr{V}_\eta \, \kappa$ instead of $\mathrm{d}\mathscr{V}_\eta$, where

$$\kappa := \sqrt{-g}/\sqrt{-\eta}. \tag{2.69}$$

This motivates the following ansatz for the general solution to the self-coupling problem:

$$\begin{aligned} \mathscr{A} &= \frac{1}{4\lambda^2} \int \mathrm{d}\mathscr{V}_\eta \, \kappa' \, M^{ar}{}_{bcst}(\eta + \lambda h, \theta) \nabla_a h^{bc} \nabla_r h^{st} \\ &= \frac{1}{4\lambda^2} \int \mathrm{d}\mathscr{V}_\eta \, \kappa' \, M^{ar}{}_{bcst}(g, \theta) \nabla_a g^{bc} \nabla_r g^{st} \ . \end{aligned} \tag{2.70}$$

The factor κ' is either $\kappa' = 1$ or $\kappa' = \kappa$ depending on the prescription we follow to obtain the source at different orders.

So far, our arguments in this section have been suggestive, but we have not obtained the form of the non-minimal couplings for the entire set of iterative equations (2.21) for an arbitrary value of θ, or shown that the determination of these is well-posed, independently of whether or not we include the variations of the volume element $\mathrm{d}\mathscr{V}_\gamma$. The main result of these considerations have been writing down the set of actions (2.70). This form of the action (specially the fact that only the combination $g^{ab} = \eta^{ab} + \lambda h^{ab}$ occurs) was not a logical necessity from the beginning, but the analysis shows that this arises as the only possible result. With these actions at hand one can proceed to a constructive proof of all our assertions in this section: by expanding this action with respect to $g^{ab} = \eta^{ab} + \lambda h^{ab}$ in the formalism of [30], one can check that it indeed satisfies the self-coupling problem with the appropriate quadratic (zeroth order) form for each value of the parameter θ, and corresponding non-minimal couplings that can be evaluated explicitly through successive functional differentiation. These therefore correspond to different solutions for different initial conditions: what at the linear level is a surface term proportional to θ, in the final theory is no longer reducible to a surface term, giving place to a complete θ-parameter family of solutions to the problem. Let us recall that our goal in this section was merely to gain an intuition about the structure of the family of solutions to the self-coupling problem when starting from the on-shell picture of gravitons; we shall justify in more depth these manipulations in a most interesting setting in Sec. 2.4.

2.3.3 The nonlinear actions: relation with unimodular gravity

Now that we have written in a compact form the nonlinear theories that arise as solutions to the self-coupling problem, we are in position to investigate some of their physical properties. In particular, we shall consider their gauge symmetries, and argue that this feature can be used to distinguish between them.

It is essential for the following discussion the consideration of the constraints on the field g^{ab}. These corresponds to the finite version of a (possibly) nonlinear equation of the form $f_{ab}\delta g^{ab} = 0$. This constraint guarantees that the resulting theory has the same degrees of freedom as the original linear construction of the graviton field. Being this a scalar constraint, two options arise: $\eta_{ab}h^{ab} = 0$ or $\sqrt{-g} = \sqrt{-\eta}$. The first one is the original constraint imposed at the linear level. However, when considering the self-interacting theory it is natural to expect that a modified nonlinear condition unfolds instead of maintaining the original traceless condition. This is the second case above, which reduces to the former at the lowest nontrivial order in the coupling constant λ. To this freedom one has to add the choice of the parameter θ. These different selections of the parametric and functional freedom lead to different theories with their own peculiarities. Notice that all of these theories are by construction invariant under general changes of coordinates. However, the amount of gauge symmetry that they present can be different.

The case we shall consider most extensively is the deformation $\sqrt{-g} = \sqrt{-\eta}$; some comments about the alternative option are given at the end of the section. Under this condition κ', as defined just below Eq. (2.70), is always $\kappa' = 1$. The first useful thing to do is try to express the action of the theory, Eq. (2.70), in an alternative form. This can be easily done at least for $\theta = 1$. To do that, let us introduce the derivative operator $\tilde{\nabla}$

associated with the field g^{ab} interpreted as a spacetime metric, such that

$$\tilde{\nabla}_a g^{bc} = 0. \tag{2.71}$$

Now we can define a tensor field C^c_{ab} relating the two derivative operators $\tilde{\nabla}$ and ∇.[5] The entire action for $\theta = 1$ can be written in terms of this tensor field. Indeed, expanding the compatibility condition (2.71) one can solve for C^c_{ab} as

$$\begin{aligned} C^c_{ab} &= -\frac{1}{2} g_{al} g_{bm} \left(g^{lk} \nabla_k g^{mc} + g^{mk} \nabla_k g^{lc} - g^{ck} \nabla_k g^{lm} \right) \\ &= -\frac{1}{2} \left(g_{bm} \nabla_a g^{mc} + g_{al} \nabla_b g^{lc} - g_{al} g_{bm} g^{ck} \nabla_k g^{lm} \right). \end{aligned} \tag{2.72}$$

It is then straightforward to show that

$$g^{ab} C^r_{sa} C^s_{rb} = \frac{1}{4} M^{ar}{}_{bcst}(g, \theta = 1) \nabla_a g^{bc} \nabla_r g^{st}. \tag{2.73}$$

This means that we can write the nonlinear action for the special case $\theta = 1$ as

$$\frac{1}{4\lambda^2} \int \mathrm{d}\mathscr{V}_\eta \, M^{ar}{}_{bcst}(g, \theta = 1) \nabla_a g^{bc} \nabla_r g^{st} = \frac{1}{\lambda^2} \int \mathrm{d}\mathscr{V}_\eta \, g^{ab} C^r_{sa} C^s_{rb}. \tag{2.74}$$

What is interesting about this expression is that it permits us to connect with the usual geometrical language of general relativity, with g_{ab} playing the role of the spacetime metric. To see that, let us consider the Einstein-Hilbert action which contains the curvature scalar R of a metric g_{ab}. As we have already discussed in Sec. 2.3.2, if the metric is split as $g^{ab} = \eta^{ab} + \lambda h^{ab}$, the curvature scalar can be written in terms of the covariant derivatives (with respect to the flat reference metric) ∇ of the field C^c_{ab}. Then, we can eliminate a total divergence by just realizing [130] that

$$\begin{aligned} &\frac{2}{\lambda^2} \int \mathrm{d}\mathscr{V}_g \, g^{ab} \left(\nabla_{[c} C^c_{a]b} + C^c_{d[c} C^d_{a]b} \right) \\ &= \frac{2}{\lambda^2} \int \mathrm{d}\mathscr{V}_\eta \nabla_c (\sqrt{-g} \delta^{[c}_a g^{e]b} C^a_{be}) - \frac{2}{\lambda^2} \int \mathrm{d}\mathscr{V}_g \, g^{ab} C^c_{d[c} C^d_{a]b}. \end{aligned} \tag{2.75}$$

The total divergence is given by

$$\frac{2}{\lambda^2} \int \mathrm{d}\mathscr{V}_\eta \nabla_c (\sqrt{-g} \delta^{[c}_a g^{d]b} C^a_{bd}), \tag{2.76}$$

and the remaining action is precisely

$$-\frac{2}{\lambda^2} \int \mathrm{d}\mathscr{V}_g \, g^{ab} C^c_{d[c} C^d_{a]b} = -\frac{1}{\lambda^2} \int \mathrm{d}\mathscr{V}_g \, g^{ab} \left(C^c_{dc} C^d_{ab} - C^c_{da} C^d_{cb} \right). \tag{2.77}$$

[5] The most consistent notation with previous definitions would be $\tilde{C}^c_{ab}$ instead of C^c_{ab}. However, here we have chosen the latter notation which simplifies the appearance of the subsequent equations.

This action was first written by Rosen in the context of a gravitational theory with a preferred flat background [116].

The only thing we need to do to make full contact with our action (2.74) is to impose the condition on the determinant $\sqrt{-g} = \sqrt{-\eta}$. Under this condition, $\mathrm{d}\mathscr{V}_g = \mathrm{d}\mathscr{V}_\eta$ and $C^b_{bc} = 0$ as it can be shown by using a particular Minkowski reference frame:

$$C^b_{bc}\Big|_{\mathscr{M}} = -\frac{1}{2} g_{ab}\partial_c g^{ab} = \frac{1}{\sqrt{-g}}\partial_c\sqrt{-g} = 0. \tag{2.78}$$

Thus the first term in Eq. (2.77) can be dropped and the volume element is replaced by $\mathrm{d}\mathscr{V}_\eta$, making this action completely equivalent to Eq. (2.74). Therefore, we have recovered the general relativity action, subject to the determinant restriction $\sqrt{-g} = \sqrt{-\eta}$. Now it is easy to analyze the gauge symmetries of this theory, the generators of which will be denoted by ξ^a. These symmetries correspond to the transformations of h^{ab} that make the combination $g^{ab} = \eta^{ab} + \lambda h^{ab}$ transform as a transverse diffeomorphism, when keeping η^{ab} constant. Their infinitesimal counterpart is therefore given by the Lie derivative of g^{ab},

$$\delta_\xi h^{ab} = \mathcal{L}_\xi g^{ab} = -\xi^c\nabla_c g^{ab} + g^{ac}\nabla_c\xi^b + g^{bc}\nabla_c\xi^a \ , \tag{2.79}$$

with the additional condition of preserving the Minkowski volume element:

$$\nabla_a\xi^a = 0. \tag{2.80}$$

Here $\mathcal{L}_\xi$ is the Lie derivative operator. Eq. (2.80) is nothing but the first condition in Eq. (2.25). Remember that what singles out the case $\theta = 1$ from the other values from the point of view of the internal symmetry is that the transverse condition [the second condition in Eq. (2.25)] can be dropped even at the level of the free spin-2 theory. This corresponds to the situation analyzed by symmetry arguments in [128], being then our discussion compatible with the content of this work (notice that the transverse condition plays no role in the solution of the iterative equations of the self-coupling problem). On the other hand, when $\theta \neq 1$ we would have additional constraints coming from the transverse condition:

$$\nabla_a g^{ab} = \nabla_a h^{ab} = 0, \tag{2.81}$$

which can be alternatively written as

$$C^a_{bc} g^{bc} = 0. \tag{2.82}$$

This means that we shall also need to impose a deformation of the second constraint over the generators of gauge symmetries in Eq. (2.25). This deformation is given by

$$\nabla_b\delta_\xi h^{ab} = \Box\xi^a + \mathscr{O}(\lambda) = 0. \tag{2.83}$$

The most important difference between these cases is the following. For $\theta = 1$ we have the same number of generators of internal symmetries subjected to the same number of restrictions, both in the linear and nonlinear constructions, so there is no reduction of gauge symmetries, just a deformation. The contrary happens when $\theta \neq 1$. Given that the action in these cases does not correspond to the restriction of the Einstein-Hilbert action to unimodular metrics, the transformations (2.79) would no longer represent symmetries unless an additional condition on the generators (i.e., that the corresponding variation of the action is zero) is imposed, thus implying that the gauge symmetry is reduced.

Lastly, let us consider what happens if the functional space is constrained by the condition $\eta_{ab}h^{ab} = 0$. As in the previous case it is better to start with the particular value $\theta = 1$. The invariance of the traceless condition leads to the deformation

$$\eta_{ab}\delta_\xi h^{ab} = \eta_{ab}g^{ac}\nabla_c\xi^b = \nabla_a\xi^a + \mathscr{O}(\lambda) = 0. \tag{2.84}$$

However, the action (2.74) does no longer correspond to the action of unimodular gravity as the determinant of g^{ab} is unconstrained and Eq. (2.78) does not hold. The only diffeomorphisms that can be a symmetry are those that leave the additional piece in the action invariant, so that the generators ξ^a have to be subjected to the additional condition of preserving the determinant of g^{ab}, given by the deformation

$$\tilde{\nabla}_a\xi^a = \nabla_a\xi^a + \mathscr{O}(\lambda) = 0. \tag{2.85}$$

Note however that Eqs. (2.84) and (2.85) correspond to different deformations, that is, the $\mathscr{O}(\lambda)$ terms are different. Therefore, the gauge symmetry of the original linear spin-2 theory has been deformed but also reduced because the generators ξ^a are subjected to an additional condition. As with the previous case, for $\theta \neq 1$ one would have additional constraints coming from the preservation of the transverse condition at the nonlinear level, but these do not play an important role on the self-coupling problem itself nor in the properties of the resulting action.

This finishes our study of the minimal case, in which we have solved the set of iterative equations from scratch starting from the on-shell picture of gravitons. We have discussed that within the family of solutions there is a privileged solution that corresponds to unimodular gravity. This represents the cleanest situation in which the only constraint on the generators of the gauge symmetries is kept unchanged, but the constraint on the graviton field is deformed from the traceless to the determinant condition. This solution is singled out when the decoupling of the degrees of freedom that fall outside the spin-2 representation of the Poincaré group that is present at the linear order (as we discussed in Sec. 2.1) is extended to all orders in the construction. Nevertheless, there are other solutions as well in which, in principle, the introduction of interactions have broken the original gauge invariance. In the following sections we shall extend these results to the only two possible linear descriptions of gravitons without constraints: Fierz-Pauli theory and Weyl-transverse theory.

2.4 From Fierz-Pauli to Einstein-Hilbert

The previous study of the minimal case has permitted us to understand the structure of the space of solutions of the iterative equations for a given set of initial conditions. However, the very nature of these initial conditions leads to unpleasant features as the consequence of the imposition of constraints at the linear level. We shall now discuss how these features can be circumvented, and what is the result of applying what we have learned before to the standard linear description of gravitons by means of Fierz-Pauli theory. Furthermore, the general character of our discussion will bring us to an appropriate position in order to discuss some of the seemingly contradictory conclusions in the recent literature.

2.4.1 Fierz-Pauli theory

With respect to the on-shell description of gravitons, Fierz-Pauli theory [131] presents an enlargement of the functional space in which the field h^{ab} is defined, with a parallel extension of the gauge symmetries. In this description both the traceless and transverse conditions in Eq. (2.22) are dropped, also extending the gauge symmetry so that the degrees of freedom are kept the same. In other words, the fundamental field is just a symmetric Lorentz tensor h^{ab} (we shall keep using the same notation and name as before for this field) with no constraints on it, and the theory is demanded to be invariant under the gauge transformations

$$\delta h^{ab}_{\xi} = \eta^{ac}\nabla_c \xi^b + \eta^{bc}\nabla_c \xi^a, \tag{2.86}$$

where now the generators ξ^a are unrestricted. In the following we shall call h^{ab} Fierz-Pauli field. Concerning this extension of the gauge symmetry, two comments are in order. The first one is that the transverse and traceless conditions can be imposed on the Fierz-Pauli field only within the space of solutions of the free theory. That is, the so-called transverse-traceless gauge can be applied only for fields h^{ab} verifying the condition

$$\nabla_a \nabla_b h^{ab} = \eta_{ab} \Box h^{ab}, \tag{2.87}$$

which is precisely the trace of the Fierz-Pauli equations [109]. The second comment is that there exists one and only one alternative extension that also reduces on-shell to the minimal description of gravitons. This alternative description leads to Weyl-transverse gravity [111, 108], as we shall discuss in Sec. 2.5.

The action of Fierz-Pauli theory is obtained by demanding the Lorentz-invariant Lagrangian density (2.1) to be also invariant under the transformations (2.86). This condition alone suffices to fix the values of the parameters $\{c_i\}_{i=2,3,4}$ (the details can be found in [29] for instance). The resulting action, including surface terms, is given by

$$\mathscr{F}_0[\eta, \theta] := \frac{1}{4} \int \mathrm{d}\mathscr{V}_\eta \, F^{ar}{}_{bcst}(\eta, \theta) \nabla_a h^{bc} \nabla_r h^{st}, \tag{2.88}$$

with

$$\begin{aligned} F^{ar}{}_{bcst}(\eta,\theta) &:= M^{ar}{}_{bcst}(\eta,\theta) - 2\delta^a_{(b}\delta^r_{c)}\eta_{st} + \eta^{ar}\eta_{bc}\eta_{st} \\ &+ \frac{1-\theta}{2}\Big[\eta_{s(b}\delta^a_t\delta^r_{c)} + \eta_{t(b}\delta^a_s\delta^r_{c)} - \delta^a_{(b}\delta^r_s\eta_{c)t} - \delta^a_{(b}\delta^r_t\eta_{c)s}\Big]. \end{aligned}$$

In the previous discussion, the parameter θ could acquire any value leaving the theory unchanged because the only modification were surface terms, due to the constraints on the field h^{ab}. In this extended setting the value of θ also controls the form of a specific surface term, the form of which is given by

$$\left[\eta_{s(b}\delta^a_t\delta^r_{c)} + \eta_{t(b}\delta^a_s\delta^r_{c)} - \delta^a_{(b}\delta^r_s\eta_{c)t} - \delta^a_{(b}\delta^r_t\eta_{c)s}\right]\nabla_a h^{bc}\nabla_r h^{st} = \left(\delta^a_s\delta^r_b\eta_{ct} - \delta^a_b\delta^r_t\eta_{cs}\right)\nabla_a h^{bc}\nabla_r h^{st}. \tag{2.89}$$

This is why in many places Fierz-Pauli theory is presented as the special case $\theta = 1$. In our discussion we are leaving this term explicit, since we have learned from the previous study that it will be of importance. We are not going to need here the explicit form of the tensor $F^{ar}{}_{bcst}(\eta,\theta)$ though. The only thing we need to keep in mind is that it can be written in terms of $M^{ar}{}_{bcst}(\eta,\theta)$ plus additional terms, which now are not identically zero as the conditions (2.22) no longer hold.

2.4.2 The family of solutions and non-uniqueness

Given that the only difference between the two starting points (on-shell and Fierz-Pauli descriptions of the graviton field) concerns the properties of the tensor field occurring in the linear action, we can try to apply now the same self-interacting scheme, but with $F^{ar}{}_{bcst}(\eta,\theta)$ instead of $M^{ar}{}_{cbst}(\eta,\theta)$, to see whether we are able to obtain general relativity as the outcome. This procedure does not work out so straightforwardly in this case. The first evidence of this is that there does not exist any value of θ for which the analogue of Eq. (2.41) is true, i.e.,

$$\sqrt{-\gamma}\,\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}D^{a\nu\rho}{}_{\mu e(b}F^{\mu r}{}_{c)dst}(\gamma,\theta) \neq -\frac{\delta\sqrt{-\gamma}F^{ar}{}_{pqst}(\gamma,\theta)}{\delta\gamma^{bc}}. \tag{2.90}$$

A way of realizing this is the following: the right-hand side of this equation contains terms that are proportional to γ_{bc}, not contracted with $F^{ar}{}_{pqst}(\gamma,\theta)$, because of the variation of the determinant. However, the left-hand side of this equation does not contain this kind of terms. Independently of the form of $F^{ar}{}_{bcst}(\gamma,\theta)$, for the first term in the left hand side of the previous equation one has e.g.

$$\gamma_{p\nu}\gamma_{q\rho}\gamma^{de}D^{a\nu\rho}{}_{\mu eb} = \gamma_{b(p}\delta^a_\mu\delta^d_{q)} + \delta^a_b\gamma_{\mu(p}\delta^d_{q)} - \gamma_{b(p}\gamma_{q)\mu}\gamma^{ad}. \tag{2.91}$$

The index b never appears in combination with the free index c. The same happens with the second term in Eq. (2.90). This means that one would need to introduce non-minimal couplings even for $\theta = 1$ in order to be able to find solutions to the self-coupling problem.

This was already noticed in the work by Butcher at al. [30]. To do that, these authors performed a reverse engineering exercise that we have already mentioned in Sec. 2.3.2. Let us discuss it briefly here. The Einstein-Hilbert action can be expanded as a series in λ using the decomposition $g^{ab} = \eta^{ab} + \lambda h^{ab}$. As we shall explain below, this series is by construction a solution of the iterative equations (2.21). These authors explicitly show (in the case $\theta = 1$) that, to guarantee that the overall procedure makes sense, one needs to accept the following condition. When writing the lowest order $\mathscr{A}_0[\eta,\theta=1]$ in terms of an auxiliary metric γ_{ab} to obtain $\mathscr{A}_0[\gamma,\theta=1]$, this quantity must contain non-minimal couplings as they are necessary to obtain the stress-energy tensor appearing in the lowest-order iterative equation (this happens also for higher orders). The quadratic action $\mathscr{A}_0[\gamma,\theta=1]$, when particularized to Minkowski space, leads precisely to $\mathscr{F}_0[\eta,\theta=1]$ as defined in Eq. (2.88).

Thus there exists a certain source, obtained through the addition of non-minimal couplings, which permits to recover general relativity as a self-interacting theory of the Fierz-Pauli field. In fact, with the right non-minimal couplings all of the actions (2.88) with an arbitrary value of θ can be uplifted to nonlinear theories that are solutions of the iterative equations. A different issue is the kind of gauge symmetries that these theories could present. In principle, only the value $\theta=1$ leads to a theory with an internal symmetry of the form of the usual diffeomorphism invariance. The final form of these theories is then given by

$$\mathscr{A}[g,\theta]=\frac{1}{4\lambda^2}\int \mathrm{d}\mathscr{V}_g\, F^{ar}{}_{bcst}(g,\theta)\nabla_a g^{bc}\nabla_r g^{st}. \tag{2.92}$$

These correspond to the entire set of solutions of the iterative equations with different initial conditions corresponding to different surface terms at the linear level. Expressions that correspond to surface terms proportional to θ at the linear level are no longer surface terms when completed in a nonlinear fashion. This is the feature that is behind the non-uniqueness of the construction. Note that in all these cases the matter fields, which have been omitted, couple to g^{ab}.

Direct use of the formalism developed in [30] permits to show that the expression (2.92) is a solution of the iterative equations of the self-coupling problem (2.21) for θ arbitrary, leading at the linear order to Fierz-Pauli theory. In terms of the analogy we discussed at the end of Sec. 2.2.2, it is essentially a matter of checking whether a given function is a solution of a certain differential equation. In order to do so, let us write down the functional (2.92) but in terms of $\gamma^{ab}+\lambda h^{ab}$ instead, with γ_{ab} an auxiliary metric. The resulting expression can be expanded in a Taylor series on the deviations λh^{ab} from γ^{ab}, and then evaluated in $\gamma^{ab}=\eta^{ab}$. One can show then that the action is by construction a solution of the iterative equations. The details can be read in [30], but it is easy to see the structure behind this demonstration by using a single-variable function $F(\gamma+\lambda h)$. The Taylor series of this function is

$$F(\gamma+\lambda h)=\sum_{n=0}^{\infty}F_n=\sum_{n=0}^{\infty}\frac{1}{n!}\frac{\partial^n F(\gamma)}{\partial\gamma^n}(\lambda h)^n. \tag{2.93}$$

The elements of the set $\{F_n\}_{n=0}^{\infty}$ verify the relations

$$\frac{\partial F_n}{\partial h}=\frac{\lambda}{(n-1)!}\frac{\partial^n F}{\partial\gamma^n}(\lambda h)^{n-1}=\lambda\frac{\partial F_{n-1}}{\partial\gamma}, \tag{2.94}$$

which are reminiscent of the iterative equations (2.21). However, these are not yet the iterative equations; these should be satisfied by the action

$$A := \int \mathrm{d}^4x\, F(\gamma + \lambda h)(\nabla' h)^2 = \sum_{n=0} \int \mathrm{d}^4x\, F_n(\nabla' h)^2 = \sum_{n=0}^{\infty} A_n. \tag{2.95}$$

Note that the determinant of the field $\gamma^{ab} + \lambda h^{ab}$ that is necessary in order to properly define the integration in the equation above would be included in the term $F(\gamma + \lambda h)$. Also $\nabla' h$ should be symbolically understood as the covariant derivative with respect to γ. Then, discarding some irrelevant boundary terms one has

$$\frac{\partial A_n}{\partial h} = \int \mathrm{d}^4x\, \frac{\partial F_n}{\partial h}(\nabla' h)^2 = \lambda \int \mathrm{d}^4x\, \frac{\partial F_{n-1}}{\partial \gamma}(\nabla' h)^2 = \lambda \frac{\partial A_{n-1}}{\partial \gamma}. \tag{2.96}$$

The boundary terms come from variations of the term $(\nabla' h)^2$ with respect to h and γ. The rigurous proof follows tightly these steps, but including the tensor structure of the graviton field which does not add any essential feature; we refer the reader to [30].

Deser has argued in [21] that the non-uniqueness inherent to the use of Noether currents in the very definition of the self-coupling problem is harmless. His argument is that these identically conserved terms that appear in the definition of the source can be absorbed in a redefinition of the Fierz-Pauli field h^{ab}. If we want to keep us in the linear level, there is only one possible redefinition: shifting h^{ab} by its trace $\eta_{ab}h^{ab}$. This means that one could absorb only certain types of such identically conserved terms. Even if we forget about this, it is difficult to see how this procedure could work as we argue in the following. Let us start with the first-order self-interacting equation

$$O_{abcd}h^{cd} = \lambda T_{ab}(h^{\rho\sigma}) + \lambda \Theta_{ab}(h^{\rho\sigma}), \tag{2.97}$$

where O_{abcd} is a given differential operator (whose form can be obtained from the action (2.88)), $T_{ab}(h^{\rho\sigma})$ is the stress-energy tensor of the graviton field as obtained from the free action, and $\Theta_{ab}(h^{\rho\sigma})$ an identically conserved tensor constructed from h^{ab}. When $\lambda = 0$ we recover the free field equations. Now let us construct a different field

$$h'^{ab} = h^{ab} + \lambda f^{ab}(h^{\rho\sigma}), \tag{2.98}$$

with $f^{ab}(h^{\rho\sigma})$ an arbitrary function of h^{ab} (which, if we want to keep at the linear level, should be proportional to $\eta^{ab}\eta_{cd}h^{cd}$). In [21] it is argued that there always exists a choice of $f^{ab}(h^{\rho\sigma})$ such that the field equations (2.97) can be written as

$$O_{abcd}h'^{cd} = \lambda T_{ab}(h'^{\rho\sigma}), \tag{2.99}$$

thus absorbing the identically conserved term $\Theta_{ab}(h^{\rho\sigma})$. The function $f^{ab}(h^{\rho\sigma})$ is determined by the following equation:

$$O_{abcd}f^{cd}(h^{\rho\sigma}) = -\Theta_{ab}(h^{\rho\sigma}). \tag{2.100}$$

As the operator O_{abcd} satisfies $\nabla^a O_{abcd} = 0$ [10], this equation is well posed and one is tempted to conclude that the identically conserved terms can be shifted away. However, one should not forget about the stress-energy tensor, which is not a mere spectator here but explicitly depends on the Fierz-Pauli field h^{ab}. Thus, at best, one can get instead of (2.99) an equation of the form

$$O_{abcd}h'^{cd} = \lambda T'_{ab}(h'^{\rho\sigma}), \tag{2.101}$$

such that

$$T'_{ab}(h'^{\rho\sigma}) = T_{ab}(h^{\rho\sigma}). \tag{2.102}$$

Notice that one of the strengths of Deser's derivation is that the series is finite so that only one iteration is needed in order to find the solution to the self-coupling problem. This permits to show clearly that the claimed equivalence cannot hold: if the identically conserved terms could be absorbed, the two following sets of equations,

$$O_{abcd}h'^{cd} = \lambda T_{ab}(h'^{\rho\sigma}) \tag{2.103}$$

and

$$O_{abcd}h^{cd} = \lambda T_{ab}(h^{\rho\sigma}) + \lambda\Theta_{ab}(h^{\rho\sigma}), \tag{2.104}$$

must be the same under $h'^{ab} = h^{ab} + \lambda f^{ab}(h^{\rho\sigma})$ with f^{ab} a solution of Eq. (2.100). But this is only possible if

$$T_{ab}(f^{cd}) = 0. \tag{2.105}$$

There is no reason for this constraint on the identically conserved terms to be true, thus implying that the effect of the non-minimal couplings cannot be simply shifted away in general. This is in consonance with the information encoded in the set of nonlinear actions (2.92).

2.4.3 Naturalness of general relativity as a solution

Butcher et al. [30] say that "*general relativity cannot be derived from energy-momentum self-coupling the Fierz-Pauli Lagrangian*". More precisely what they mean is that one cannot use the stress-energy tensor obtained straightforwardly from the Fierz-Pauli Lagrangian density by using a minimal coupling prescription. One has to introduce specific non-minimal couplings in order to consistently solve the iterative equations for $\theta = 1$. From reading this paper and Padmanabhan's one ends up with the impression that to obtain general relativity from self-interaction one needs to know somehow the final result, as one needs to make use of curved-spacetime notions. However, here we have discussed that non-minimal couplings are encompassed by covariant surface terms (though generally different from those controlled by θ), thus forming part of the standard arbitrariness in defining the stress-energy tensor even in flat spacetime. Moreover, the specific non-minimal couplings needed for each value of θ are determined by the iterative equations of the self-coupling problem only. Allowing surface terms one finds a one-parameter family of solutions to the self-coupling problem. From them, general relativity is selected by requiring the final theory to have the largest possible amount of gauge invariance. In other words, $\theta = 1$ is fixed by demanding that the decoupling of some of the degrees of freedom that already takes place at first order, as discussed in Sec. 2.1, holds for all orders. Thus the construction only uses concepts based on Poincaré-invariant field theory and gauge invariance, even when some of the mathematical tools used are geometrical, as the Hilbert prescription to obtain the stress-energy tensor. It is instructive to notice that the necessity of considering the addition of identically conserved terms to the source one would obtain directly from the free action is not exclusive of gravity, but the same thing happens when considering the case of Yang-Mills theory in the second-order formalism, as it is explicitly written (but to some extent ignored) in the original work of Deser [20].

If we consider the non-tensorial general-relativity action

$$\frac{1}{\lambda^2}\int \mathrm{d}^4x\sqrt{-g}\, g^{ab}(\Gamma^c_{da}\Gamma^d_{cb} - \Gamma^c_{ab}\Gamma^d_{cd}), \tag{2.106}$$

and perform an expansion in the coupling constant λ with $g^{ab} = \eta^{ab} + \lambda h^{ab}$, we will see that it precisely exhibits a coupling term of the form $h^{ab}S_{ab}$ at first order in λ. Padmanabhan pointed out the role of this object S_{ab} in any coupling scheme leading to general relativity [29]. He showed for instance that this object S_{ab} can be obtained from the quadratic term (zeroth order in λ) by applying only a half-covariantization scheme which might be regarded at least as unnatural (see Appendix A in [30] for additional comments on this quantity). This means that, somewhat surprisingly, whereas the quadratic action is tensorial, the first order correction should already be non-tensorial. The variation of this new action with respect γ_{ab} might lead in principle to a non-tensorial stress-energy object (though finally this is not the case). Therefore one could argue, as Padmanabhan did [29], that the construction of general relativity from a self-coupling scheme is somewhat unphysical (only at the end of the iterative procedure one would realize the existence of a total divergence that allows the construction of a diffeomorphism-invariant Lagrangian density).

However, in our formulation we always keep track of the flat reference metric. This allows us to construct the tensorial action

$$\frac{1}{\lambda^2}\int \mathrm{d}^4x\sqrt{-g}\, g^{ab}(C^c_{da}C^d_{cb} - C^c_{ab}C^d_{cd}), \tag{2.107}$$

instead of the non-tensorial action (2.106). Performing again the expansion in the coupling constant λ, the cubic term has a form $h^{ab}\bar{S}_{ab}$ where now $\bar{S}_{ab}$ is a proper tensor. Moreover, this object is not and must not be the stress-energy tensor. We have seen that there is a natural definition of $\bar{S}_{ab}$ within the iterative procedure, as the result of the integration of the first-order iterative equation analogue to (2.18). In fact one of the main differences between the work of Padmanabhan [29] and that in here is that we have explicitly performed the integration of the iterative equations. In other words, from the point of view of the self-coupling consistency problem, $\bar{S}_{ab}$ is just a derived quantity and not a fundamental one. One will be led to it by following the equations carefully [recall for example the discussion around (2.37)]. Concerning this last point, Deser makes a similar comment in his reply to Padmanabhan [21]: the only role of $\bar{S}_{ab}$ is to lead to the required source when the variations which respect to h^{ab} are performed, and this is precisely the definition of this quantity.

Concerning boundary terms in the resulting nonlinear action: the Einstein-Hilbert action can be partitioned in a first-derivative action plus a total divergence in several ways. If one does not introduce a fiducial background metric, this partition has to be non-tensorial. Instead, by introducing a flat background metric, one discovers a tensorial partition. In our view what is unnatural from the self-coupling program point of view is precisely to forget about the background metric, making an identification of the invariance under changes of coordinates that is already present in flat spacetime and the invariance under gauge transformations (diffeomorphisms). Once one obtains the action (2.107), which is a scalar, one would not look for complementing this action with additional surface terms to build the scalar curvature. Only when taking the non-trivial conceptual jump of forgetting about the background structure and taking a complete geometrical description in terms of a single metric, one would start worrying about the significance of the surface term and its non-tensorial character. In this stage we agree with Padmanabhan's [29] that the surface term of the Einstein-Hilbert action is not naturally obtainable in the self-coupling problem as it does not affect the equations of motion; one has to add geometrical information.

Most of these comments apply to the classic papers of Deser [20, 21], e.g., the fact that the resulting action will be written in terms of the covariant derivative ∇ with respect to the flat reference metric. The clever choice of independent variables in that work allowed him to lead to completion the iterative procedure in a single step. Precisely, this selection of variables hides the fact that the stress-energy tensor obtained by varying η^{ab} is not the one that one would directly obtain from the minimally-coupled Fierz-Pauli theory. That is, Deser's first order formalism naturally selects the specific non-minimal couplings that lead to the Einstein equations. The reader should not confuse this with the surface term in the Einstein-Hilbert action, which is put by hand (it has nevertheless no impact on the equations of motion). Notice that all the other potential solutions to the self-coupling problem are missing in this particular approach. One could recover them by using additional surface terms (i.e., changing the value of θ) in his direct construction. One cannot simply exclude these possibilities from a logical the perspective, but the resulting iterative series might be infinite. A more general treatment such as the one presented here, in which the use of specific variables is avoided and which permits to handle infinite series, is therefore convenient in order to grasp the nature of these solutions.

In summary, in the Fierz-Pauli case the self-coupling problem naturally leads to a one-parameter set of solutions that includes general relativity. From this point of view general relativity naturally emerges from the self-coupling of a Poincaré-invariant field theory. However, to select general relativity from the other theories one has to require the existence of a maximal gauge symmetry or, equivalently, the decoupling of certain degrees of freedom. There seems to be no alternative guiding principle to directly obtain general relativity. Moreover, what is obtained is closer to the bimetric theory of Rosen [116].[6] Although both theories are observationally equivalent in standard situations, they might suggest different deviations from the classical behavior of gravitational fields, as well as conceptual differences when considering extreme situations such as spacetime singularities.

2.5 Weyl-transverse gravity

One of the main results of our discussion on the minimal picture of gravitons in Sec. 2.3 was obtaining unimodular gravity as the result of the self-coupling problem. Due to some unappealing features of the construction, afterwards we changed the linear description of gravitons to that of Fierz-Pauli theory, obtaining general relativity instead as the natural outcome. The obtention of unimodular gravity thus seems to be an artifact of the constraints imposed to the graviton field from the beginning in the minimal construction. However, in this section we explain how a suitable version of unimodular gravity also arises naturally as a solution to the iterative equations, when some subtleties on the extension from the on-shell linear description to the Fierz-Pauli description are taken into account.

2.5.1 Two extensions of the on-shell picture of gravitons

As reviewed in the introduction of this chapter, the on-shell description of gravitons by means of the corresponding unitary representations of the Poincaré group is naturally expressed in terms of a second-rank, transverse and traceless tensor field h^{ab}. This field satisfies the constraints (2.22) and the generators of the gauge invariance of the linear theory were also constrained by Eq. (2.25). This setup was the starting point for our initial discussion on the self-coupling problem in Sec. 2.3. This on-shell description is the only firm statement one can draw from the assumption that gravity is mediated by a spin-2 graviton only, with no admixture of spin 1 or 0.

Nevertheless, for practical purposes it is useful (as exemplified for instance in our discussion of the previous section) to relax the conditions (2.22) and (2.25), enlarging the gauge symmetry. As explained in Sec. 2.4 this procedure leads to the well-known Fierz-Pauli theory. However, it is remarkable that there exists an alternative extension that also reduces on-shell to the minimal picture of gravitons [111], known as Weyl-transverse theory. As the name suggests, the gauge transformations of the theory are given by

$$h'^{ab} = h^{ab} + \eta^{ac}\nabla_c\xi^b + \eta^{bc}\nabla_c\xi^a + \phi\eta^{ab}, \tag{2.108}$$

with generators satisfying only the first condition in Eq. (2.22), that is, $\nabla_a\xi^a = 0$, and ϕ is an arbitrary scalar function. This is to be compared with Eq. (2.24). The linear action of this theory is given by:

$$\mathscr{A}_0[\eta] := \frac{1}{4}\int \mathrm{d}\mathscr{V}_\eta \left(2\eta_{bs}\delta^a_t\delta^r_c - \eta^{ar}\eta_{bs}\eta_{ct} - \eta_{bc}\delta^a_t\delta^r_s + \frac{3}{8}\eta^{ar}\eta_{bc}\eta_{st}\right)\nabla_a h^{bc}\nabla_r h^{st}. \tag{2.109}$$

[6]The reader should not confuse the resulting theory with what is usually considered a bimetric theory, as here one of the metrics (the flat reference metric) is not a dynamical entity.

Fierz-Pauli theory and Weyl-transverse gravity are the only two linear theories compatible with the on-shell description of free gravitons [108]. This fact alone is interesting enough to explore this last theory to its ultimate consequences. Even if both are by construction equivalent as linear theories, their nonlinear completions could differ.

2.5.2 A different nonlinear theory

From our previous discussion we expect surface terms added to the action (2.109) to be of relevance, leading again to an entire family of solutions when the right non-minimal couplings are also introduced. Nonetheless, from all these solutions one would only preserve the same degrees of freedom or, in other words, the same amount of gauge invariance than the linear theory. While knowing the explicit form of all these solutions would be interesting, in this section we are mainly interested in this particular solution. So it will be enough for our purposes to construct the action of a theory with these features by using symmetry arguments, and then show that it is indeed a solution to the iterative equations (2.21) that leads at the linear level to the very description of gravitons given in the previous section.

Let us therefore start with a nonlinear action motivated by unimodular gravity,

$$\mathscr{A} = \frac{1}{\lambda^2} \int \mathrm{d}\mathscr{V}_\eta \, \mathscr{R}(\hat{g}), \tag{2.110}$$

where $\mathscr{R}(\hat{g})$ has the same functional form as the Ricci scalar of a metric $\hat{g}_{ab}$ whose determinant is constrained by the condition $\det(\hat{g}) = \det(\eta)$. However, for us $\hat{g}_{ab}$ is just a tensor field that lives in a flat background. To make this explicit, let us use the well-known fact [116] that one can express the Ricci scalar in terms of the covariant derivatives associated with the flat metric η_{ab}, denoted by ∇, and integrate by parts to put the action in the form:

$$\mathscr{A} = \frac{1}{4\lambda^2} \int \mathrm{d}\mathscr{V}_\eta \left(2\hat{g}_{bs}\delta^a_t\delta^r_c - \hat{g}^{ar}\hat{g}_{bs}\hat{g}_{ct}\right) \nabla_a \hat{g}^{bc} \nabla_r \hat{g}^{st}. \tag{2.111}$$

The covariant notation we are using makes clear that this theory is invariant under general coordinate transformations. Moreover, by construction this action is invariant under transverse diffeomorphisms, the infinitesimal form of which is

$$\delta_\xi \hat{g}^{ab} = \mathcal{L}_\xi \hat{g}^{ab}, \qquad \nabla_a \xi^a = 0. \tag{2.112}$$

These are the nonlinear version of the first part of the linear symmetry (2.108) with transverse generators. The natural nonlinear deformation of the remaining symmetry of the Weyl-transverse theory are conformal transformations. To include these let us define

$$\hat{g}_{ab} := \kappa^{-1/4} g_{ab}. \tag{2.113}$$

Here $\kappa := \det(g)/\det(\eta)$. This definition makes the action automatically invariant under conformal transformations, whose infinitesimal version is

$$\delta g_{ab} = \delta\omega \, g_{ab}. \tag{2.114}$$

Written in terms of g_{ab} the action is given by:

$$\mathscr{A} = \frac{1}{4\lambda^2}\int \mathrm{d}\mathscr{V}_\eta\, \kappa^{1/4}\Big[(2g_{bs}\delta^a_t\delta^r_c - g^{ar}g_{bs}g_{ct})\nabla_a g^{bc}\nabla_r g^{st} + \frac{1}{\kappa}\delta^a_b\delta^r_c\nabla_a g^{bc}\nabla_r\kappa - \frac{1}{2\kappa}g^{ar}g_{bc}\nabla_a g^{bc}\nabla_r\kappa - \frac{1}{8\kappa^2}g^{ar}\nabla_a\kappa\nabla_r\kappa\Big]. \tag{2.115}$$

One can now see that the transformations (2.112) and (2.114) combine in a way that makes this action invariant under transverse diffeomorphisms (now acting on the field g_{ab}) as well as conformal transformations. The action (2.115) is thus the most general nonlinear covariant action quadratic in the derivatives of the field g_{ab} and satisfying these invariance requisites.

Following our previous discussion it is not difficult to show that this nonlinear theory can be obtained via the self-coupling of gravitons initially described by the action (2.109). To do that, we shall extend the formalism in [30], which was useful to prove the analogue result in the case of general relativity, to consider actions of the type:

$$\frac{1}{4\lambda^2}\int \mathrm{d}\mathscr{V}_\eta\, \left[M^{ar}{}_{bcst}(g,\eta)\nabla_a g^{bc}\nabla_r g^{st} + N^{ar}{}_{bc}(g,\eta)\nabla_a g^{bc}\nabla_r\kappa + O(g,\eta)g^{ar}\nabla_a\kappa\nabla_r\kappa\right]. \tag{2.116}$$

These actions display an additional dependence on the flat metric η_{ab}. We shall comment on the meaning of this feature below. As we discussed in Sec. 2.4, just by performing an expansion $g^{ab} = \gamma^{ab} + \lambda h^{ab}$ one can directly show that the action (2.116) is by construction a solution of the iterative equations. Moreover, the lowest order is given by the first functional derivative of this action which does not vanish in the flat limit $\gamma_{ab} \to \eta_{ab}$. In this case, it is the second order contribution which in the flat limit reads:

$$\frac{1}{4}\int \mathrm{d}\mathscr{V}_\eta\, \left[M^{ar}{}_{bcst}(\eta,\eta)\nabla_a h^{bc}\nabla_r h^{st} - N^{ar}{}_{bc}(\eta,\eta)\eta_{st}\nabla_a h^{bc}\nabla_r h^{st} + O(\eta,\eta)\eta^{ar}\eta_{bc}\eta_{st}\nabla_a h^{bc}\nabla_r h^{st}\right]. \tag{2.117}$$

The expression of this first nontrivial order in terms of a general auxiliary metric γ_{ab} gives the non-minimal couplings which are necessary to the consistency of the formalism and, thus, the source to which the field h^{ab} couples at first order. Higher orders can be directly evaluated from this first order to construct the infinite series of partial actions $\{\mathscr{A}_n\}_{n=1}^\infty$ and the corresponding sources.

The action (2.115) we want to consider is a particular case of (2.116) with

$$\begin{aligned} M^{ar}{}_{bcst}(g,\eta) &:= 2\kappa^{1/4}g_{bs}\delta^a_t\delta^r_c - \kappa^{1/4}g^{ar}g_{bs}g_{ct},\\ N^{ar}{}_{bc}(g,\eta) &:= \frac{1}{\kappa}\kappa^{1/4}\delta^a_b\delta^r_c - \frac{1}{2\kappa}\kappa^{1/4}g^{ar}g_{bc},\\ O(g,\eta) &:= -\frac{1}{8\kappa^2}\kappa^{1/4}. \end{aligned} \tag{2.118}$$

Note that we are ignoring the symmetrization of these quantities, which is not essential due to the fact that in both Eqs. (2.116) and (2.117) these tensors are contracted with quantities that display the relevant symmetries. With these equations at hand, we can check that the leading order (2.117) is exactly (2.109), finishing the proof. Now from these expressions we can see that the additional dependence on the Minkowski metric η_{ab} is only through its determinant. This just reflects that in order to define the action of Weyl-transverse gravity in a coordinate-invariant fashion one has to introduce an auxiliary non-dynamical volume element, which given the nature of the construction is instinctively identified with the flat volume element. Note that this auxiliary volume element remains inert while applying Hilbert's prescription to obtain the source of the nonlinear equations of motion at different orders.

Therefore we conclude that there is no distinction from the perspective of self-coupling between general relativity and Weyl-transverse gravity. Both nonlinear theories lead to the kind of particle one expects to mediate the gravitational interaction in the non-interacting limit and can be explicitly constructed by self-coupling this particle. The only distinction between them is their internal symmetry group. And this very feature explains the results of previous analyses about the uniqueness of general relativity as a solution to the self-coupling problem. Indeed, in these analyses it was always assumed that the gauge symmetry characteristic of gravitons is that of Fierz-Pauli theory. For instance, in [132] the Ward identities associated to the gauge symmetries of Fierz-Pauli theory are an essential part of the demonstration of the uniqueness of general relativity. We completely agree that, in accepting this assumption, general relativity arises as the only consistent nonlinear theory that preserves the original gauge invariance. Nevertheless, assuming Fierz-Pauli theory as the correct description of lineal gravitons is a strong assumption that is not necessary from the perspective of the self-coupling problem; moreover, it is ultimately motivated by the knowledge of the symmetry group of general relativity. This uncovers an additional non-uniqueness at the heart of the self-coupling problem of gravitons.

One could expect that deviating from the usual solution to the self-coupling problem would incur in disagreement with experimental facts. On the contrary, the theory constructed here indeed describes gravity in a way which is compatible with all the known experiments in gravitation, as it is essentially equivalent to unimodular gravity [128, 133]. Matter is naturally coupled to $\hat{g}_{ab}$, possibly with non-minimal couplings. The field equations are, by construction, traceless. In particular, in the gauge $\det(g) = \det(\eta)$ they reduce to the trace-free Einstein equations [134, 135]:

$$\mathscr{R}_{ab} - \frac{1}{4}\mathscr{R} g_{ab} = 2\lambda^2 \left(T_{ab} - \frac{1}{4} T g_{ab} \right). \tag{2.119}$$

These are tensorial with respect to changes of coordinates regardless of the fact that the metric is constrained, as it is a tensorial constraint, $\det(g) = \det(\eta)$. This is different from the strict formulation of unimodular gravity in which active and passive diffeomorphisms are merged and the metric is subjected to a non-tensorial condition $\det(g) = 1$.

As it was shown by Rosen [116], the conservation of the canonical stress-energy tensor of gravity in flat spacetime is equivalent to the covariant conservation of the stress-energy tensor of matter fields. Applying this condition to the field equations (2.119) we recover the Einstein field equations with a phenomenological integration constant, unrelated to zero-point energies of matter, playing the role of a cosmological constant. The resulting field equations include potential energies in the matter sector as, even if gravity is not directly coupled to these terms by construction, they inevitably appear in the definition of the canonical stress-energy tensor [134, 135]. In the next chapter we will study in detail the implications of this theory for the cosmological constant problem.

2.6 Conclusions

In this chapter we have discussed Gupta's original program in detail, concerning the possible theories which arise as self-interacting theories of gravitons propagating in Minkowski spacetime. The discussion applies to quantum theories whose low-energy spectrum contains gravitons interacting with matter in a flat background, as long as one accepts that the long wavelength limit is described by a classical, second-order Lagrangian field theory.

We have explicitly solved the infinite set of iterative equations that appears when using a standard formalism based on the tensor field variable h^{ab} for the graviton, thus complementing previous work on the subject concerning finite series which appear when specific variables are considered. To do that we have constructed a proof by induction and found the formal sum of the resulting series, starting from the on-shell linear picture of gravitons motivated by the irreducible spin-2 representation of the Poincaré group. We have extended and contrasted our approach with previous discussions in the literature that start instead from Fierz-Pauli theory, a linear representation of gravitons with a larger gauge symmetry. The formalism we have used has permitted us to explicitly show the interplay between gauge invariance (a notion which is clearly separated from changes of coordinates) and the self-coupling procedure. Finally, we have considered the only other alternative extension of the gauge symmetry at the very linear level (Weyl-transverse theory), and explored its nonlinear implications for the self-coupling problem.

One of the most important conclusions is that the outcome of the iterative equations of the self-coupling problem is *not* unique. One obtains field equations that are equivalent to those of general relativity (or Weyl-transverse gravity in the alternative description of gravitons) as the result of self-coupling, but one has to demand that the number of generators of gauge symmetries is preserved in this procedure. We have explicitly shown that the construction is completely natural from the perspective of flat spacetime and does not need any information conceptually related to geometric notions: the non-minimal couplings that are necessary in some cases are nothing but natural surface terms that form part of the standard ambiguities in the definition of the stress-energy tensor in flat spacetime. The preservation of gauge invariance is also natural from the perspective of a field theory in flat spacetime. This condition can be understood as the extension to higher orders of the decoupling of the degrees of freedom that is observed when coupling the theory of a Lorentz-invariant theory of gravitons to a conserved current at first order. If one does not require the preservation of gauge invariance, the self-coupling problem exhibits other solutions.

This is specially important for the emergent gravity program. We can conclude that an additional imposition is needed in order to fix general relativity as the relevant low-energy description of the gravitational degrees of freedom. While the occurrence of Fermi points guarantees the emergence at low energies of Lorentz and gauge invariance for an abelian theory as the one considered in the previous chapter, obtaining the specific set of nonlinearities of general relativity requires something more. This could permit to understand the difficulty of obtaining general relativity in condensed-matter-like models, and guide future efforts to the determination of an additional principle or property that selects general relativity as the correct description of the low-energy degrees of freedom for a given universality class. It is interesting to notice that this resonates with the situation in the string theory framework. In the latter formalism, excitations with the properties of gravitons are directly obtained when linearizing around Minkowski spacetime. However, this fact alone would not imply that general relativity is recovered: indeed, the Einstein field equations are shown to be verified by coherent states for gravitons when the additional condition of cancellation of anomalies of local symmetries in these backgrounds is taken into account.

From the perspective of the action principle, the iterative equations of the self-coupling problem determine that the nonlinear action has to be constructed in terms of the combination $g^{ab} = \eta^{ab} + \lambda h^{ab}$ (or $\hat{g}^{ab}$ if starting from Weyl-transverse theory). The Einstein-Hilbert action (or Weyl-transverse gravity action) is therefore one of the possible outcomes, but not the only one. The complete one-parameter family of solutions to the self-coupling problem is obtained by uplifting what at the linear level is a surface term (proportional to a real parameter θ) to a nonlinear term that affects the equations of motion. Only for a specific value of the parameter the resulting theory displays a deformation of the linear gauge invariance, with no reduction of the number of gauge generators.

Even if the resulting theory admits a geometrical interpretation, it is naturally expressed as a bimetric theory à la Rosen [116], and not directly general relativity (or Weyl-transverse gravity), which from this perspective have forgotten the existence of a flat reference metric. Therefore, the structures of general relativity (or Weyl-transverse gravity) do appear naturally without drawing upon curved spacetime notions, but precisely because of this they appear in a form that does not demand a geometrical interpretation in terms of a unique metric. Their form does not demand the supplementation of the action with a total divergence to build the Einstein-Hilbert action either. The geometrical interpretation is certainly appealing as, on the one hand, it provides a natural interpretation of the absence of a local meaning for the gravitational energy (a gauge dependent quantity) and, on the other hand, it makes the theory self-contained, with no externally fixed elements. However, here we adhere to Rosen's comment more than 60 years ago [117]: *"Perhaps this* (flat spacetime interpretation) *may be regarded by some as a step backward. It should be noted, however, that this geometrization referred to has never been extended satisfactorily to other branches of physics, so that gravitation is treated differently from other phenomena. It is therefore not unreasonable to wonder whether it may not be better to give up the geometrical approach to gravitation for the sake of obtaining a more uniform treatment for all the various fields of force that are to be found in nature."*

The most important conclusion to be drawn from this feature is that starting from an effective flat spacetime, as it is generally the case in condensed-matter-like systems, does not represent any drawback to obtain general relativity at low energies. The existence of this flat background is conveniently hidden at the end of the day, being undetectable by means of classical gravitational experiments. However, some of its properties could be revealed when aspects that go beyond classical physics are included, as explained in the second part of this thesis. We shall study the possible low-energy implications that stem from the two independent parts of this background structure: its volume form (Chap. 3) and its conformal, or causal structure (Chap. 4).

As we have remarked several times, there is an additional source of non-uniqueness in the self-coupling problem that has its roots in the very linear representation of gravitons. While the usual representation in terms of Fierz-Pauli theory leads to general relativity when the condition of preservation of gauge invariance is imposed, this is not the only reasonable choice, but one can show that there is only one alternative. This non-uniqueness at the linear level extends to the nonlinear regime, leading to the existence of a different solution to the self-coupling problem that is incarnated by the theory known as Weyl-transverse gravity (intimately related to unimodular gravity). This opens the possibility of dealing with the cosmological constant problem in the framework of emergent gravity, a prospect that is studied in detail in the next chapter.

Appendix A: An algebraic identity

If we evaluate the derivative with respect to the auxiliary metric and forget momentarily about the symmetrization in the pair (s,t), we can write Eq. (2.58) as

$$D^{a\nu\rho}{}_{\mu be}(\gamma_{ps}\gamma_{\nu t}\gamma_{q\rho}\gamma^{de}+\gamma_{p\nu}\gamma_{qs}\gamma_{\rho t}\gamma^{de}-\gamma_{p\nu}\gamma_{q\rho}\delta^d_s\delta^e_t)=\gamma_{p\nu}\gamma_{q\rho}\gamma_{t\beta}\gamma^{d\delta}D^{a\nu\rho}{}_{\theta se}D^{\theta e\beta}{}_{\mu b\delta}. \quad (2.120)$$

Of course, this equation would only be valid when the terms obtained under the exchange $s\leftrightarrow t$ are added. In the following we are going to show that this equation holds. The left-hand side is easier to evaluate; it is composed by three terms:

$$\begin{aligned}&\frac{1}{2}\gamma_{ps}\gamma_{\nu t}\gamma_{q\rho}\gamma^{de}\left(\delta^a_\mu\delta^\nu_b\delta^\rho_e+\delta^a_\mu\delta^\rho_b\delta^\nu_e+\delta^a_b\delta^\nu_\mu\delta^\rho_e+\delta^a_b\delta^\rho_\mu\delta^\nu_e-\delta^a_e\delta^\nu_b\delta^\rho_\mu-\delta^a_e\delta^\rho_b\delta^\nu_\mu\right)\\&+\frac{1}{2}\gamma_{p\nu}\gamma_{qs}\gamma_{\rho t}\gamma^{de}\left(\delta^a_\mu\delta^\nu_b\delta^\rho_e+\delta^a_\mu\delta^\rho_b\delta^\nu_e+\delta^a_b\delta^\nu_\mu\delta^\rho_e+\delta^a_b\delta^\rho_\mu\delta^\nu_e-\delta^a_e\delta^\nu_b\delta^\rho_\mu-\delta^a_e\delta^\rho_b\delta^\nu_\mu\right)\\&-\frac{1}{2}\gamma_{p\nu}\gamma_{q\rho}\delta^d_s\delta^e_t\left(\delta^a_\mu\delta^\nu_b\delta^\rho_e+\delta^a_\mu\delta^\rho_b\delta^\nu_e+\delta^a_b\delta^\nu_\mu\delta^\rho_e+\delta^a_b\delta^\rho_\mu\delta^\nu_e-\delta^a_e\delta^\nu_b\delta^\rho_\mu-\delta^a_e\delta^\rho_b\delta^\nu_\mu\right).\end{aligned} \quad (2.121)$$

The first six terms are:

$$\gamma_{ps}\gamma_{bt}\delta^a_\mu\delta^d_q+\gamma_{ps}\gamma_{qb}\delta^a_\mu\delta^d_t+\gamma_{ps}\gamma_{t\mu}\delta^a_b\delta^d_q+\gamma_{ps}\gamma_{q\mu}\delta^a_b\delta^d_t-\gamma_{ps}\gamma_{q\mu}\gamma_{bt}\gamma^{ad}-\gamma_{ps}\gamma_{qb}\gamma_{t\mu}\gamma^{ad}. \quad (2.122)$$

These are followed by the following six terms:

$$\gamma_{pb}\gamma_{qs}\delta^a_\mu\delta^d_t+\gamma_{qs}\gamma_{bt}\delta^a_\mu\delta^d_p+\gamma_{p\mu}\gamma_{qs}\delta^a_b\delta^d_t+\gamma_{qs}\gamma_{t\mu}\delta^a_b\delta^d_p-\gamma_{p\mu}\gamma_{qs}\gamma_{bt}\gamma^{ad}-\gamma_{pb}\gamma_{qs}\gamma_{t\mu}\gamma^{ad}. \quad (2.123)$$

The last six terms are:

$$-\gamma_{pb}\gamma_{qt}\delta^a_\mu\delta^d_s - \gamma_{pt}\gamma_{qb}\delta^a_\mu\delta^d_s - \gamma_{p\mu}\gamma_{qt}\delta^d_s\delta^a_b - \gamma_{pt}\gamma_{q\mu}\delta^a_b\delta^d_s + \gamma_{p\mu}\gamma_{qb}\delta^d_s\delta^a_t + \gamma_{pb}\gamma_{q\mu}\delta^a_t\delta^d_s. \quad (2.124)$$

So we finish with the following expression symmetric in $p \leftrightarrow q$:

$$\begin{aligned}
&\frac{1}{2}\left(\gamma_{ps}\gamma_{bt}\delta^a_\mu\delta^d_q + \gamma_{qs}\gamma_{bt}\delta^a_\mu\delta^d_p + \gamma_{ps}\gamma_{qb}\delta^a_\mu\delta^d_t + \gamma_{qs}\gamma_{pb}\delta^a_\mu\delta^d_t + \gamma_{ps}\gamma_{t\mu}\delta^a_b\delta^d_q + \gamma_{qs}\gamma_{t\mu}\delta^a_b\delta^d_p\right.\\
&+ \gamma_{ps}\gamma_{q\mu}\delta^a_b\delta^d_t + \gamma_{qs}\gamma_{p\mu}\delta^a_b\delta^d_t - \gamma_{ps}\gamma_{q\mu}\gamma_{bt}\gamma^{ad} - \gamma_{qs}\gamma_{p\mu}\gamma_{bt}\gamma^{ad} - \gamma_{ps}\gamma_{qb}\gamma_{t\mu}\gamma^{ad} - \gamma_{qs}\gamma_{pb}\gamma_{t\mu}\gamma^{ad}\\
&\left.+ \gamma_{p\mu}\gamma_{qb}\delta^d_s\delta^a_t + \gamma_{q\mu}\gamma_{pb}\delta^a_t\delta^d_s - \gamma_{pb}\gamma_{qt}\delta^a_\mu\delta^d_s - \gamma_{pt}\gamma_{qb}\delta^a_\mu\delta^d_s - \gamma_{p\mu}\gamma_{qt}\delta^d_s\delta^a_b - \gamma_{q\mu}\gamma_{pt}\delta^a_b\delta^d_s\right). \quad (2.125)
\end{aligned}$$

This expression must be still symmetrized under $s \leftrightarrow t$. When one does this some of the terms cancel,

$$3 \leftrightarrow 16, \qquad 4 \leftrightarrow 15, \qquad 7 \leftrightarrow 18, \qquad 8 \leftrightarrow 17; \quad (2.126)$$

leaving the simplified result:

$$\begin{aligned}
&\frac{1}{2}\left(\gamma_{ps}\gamma_{bt}\delta^a_\mu\delta^d_q + \gamma_{qs}\gamma_{bt}\delta^a_\mu\delta^d_p + \gamma_{ps}\gamma_{t\mu}\delta^a_b\delta^d_q + \gamma_{qs}\gamma_{t\mu}\delta^a_b\delta^d_p - \gamma_{ps}\gamma_{q\mu}\gamma_{bt}\gamma^{ad}\right.\\
&\left.- \gamma_{qs}\gamma_{p\mu}\gamma_{bt}\gamma^{ad} - \gamma_{ps}\gamma_{qb}\gamma_{t\mu}\gamma^{ad} - \gamma_{qs}\gamma_{pb}\gamma_{t\mu}\gamma^{ad} + \gamma_{p\mu}\gamma_{qb}\delta^d_s\delta^a_t + \gamma_{q\mu}\gamma_{pb}\delta^a_t\delta^d_s\right)_{s\leftrightarrow t}. \quad (2.127)
\end{aligned}$$

This is the equation which we must compare with the right-hand side of Eq. (2.120). In this side, there are 36 terms in total,

$$\begin{aligned}
&\frac{1}{4}\gamma_{p\nu}\gamma_{q\rho}\gamma_{t\beta}\gamma^{d\delta}\left(\delta^a_\theta\delta^\nu_s\delta^\rho_e + \delta^a_\theta\delta^\rho_s\delta^\nu_e + \delta^a_s\delta^\nu_\theta\delta^\rho_e + \delta^a_s\delta^\rho_\theta\delta^\nu_e - \delta^a_e\delta^\rho_s\delta^\nu_\theta - \delta^a_e\delta^\nu_s\delta^\rho_\theta\right)\\
&\times\left(\delta^\theta_\mu\delta^e_b\delta^\beta_\delta + \delta^\theta_\mu\delta^\beta_b\delta^e_\delta + \delta^\theta_b\delta^e_\mu\delta^\beta_\delta + \delta^\theta_b\delta^\beta_\mu\delta^e_\delta - \delta^\theta_\delta\delta^e_\mu\delta^\beta_b - \delta^\theta_\delta\delta^e_b\delta^\beta_\mu\right). \quad (2.128)
\end{aligned}$$

The 36 terms are given by (the following expression is multiplied by 1/4):

$$\begin{aligned}
&\gamma_{ps}\gamma_{qb}\delta^a_\mu\delta^d_t + \gamma_{ps}\gamma_{bt}\delta^a_\mu\delta^d_q + \gamma_{ps}\gamma_{q\mu}\delta^a_b\delta^d_t + \gamma_{ps}\gamma_{t\mu}\delta^a_b\delta^d_q - \gamma_{ps}\gamma_{q\mu}\gamma_{bt}\gamma^{ad} - \gamma_{ps}\gamma_{qb}\gamma_{t\mu}\gamma^{ad}\\
&+ \gamma_{pb}\gamma_{qs}\delta^a_\mu\delta^d_t + \gamma_{qs}\gamma_{bt}\delta^a_\mu\delta^d_p + \gamma_{p\mu}\gamma_{qs}\delta^a_b\delta^d_t + \gamma_{qs}\gamma_{t\mu}\delta^a_b\delta^d_p - \gamma_{p\mu}\gamma_{qs}\gamma_{bt}\gamma^{ad} - \gamma_{pb}\gamma_{qs}\gamma_{t\mu}\gamma^{ad}\\
&+ \gamma_{p\mu}\gamma_{qb}\delta^a_s\delta^d_t + \gamma_{p\mu}\gamma_{bt}\delta^a_s\delta^d_q + \gamma_{pb}\gamma_{q\mu}\delta^a_s\delta^d_t + \gamma_{pb}\gamma_{t\mu}\delta^a_s\delta^d_q - \gamma_{q\mu}\gamma_{tb}\delta^a_s\delta^d_p - \gamma_{qb}\gamma_{t\mu}\delta^a_s\delta^d_p\\
&+ \gamma_{pb}\gamma_{q\mu}\delta^a_s\delta^d_t + \gamma_{q\mu}\gamma_{bt}\delta^a_s\delta^d_p + \gamma_{p\mu}\gamma_{qb}\delta^a_s\delta^d_t + \gamma_{qb}\gamma_{t\mu}\delta^a_s\delta^d_p - \gamma_{p\mu}\gamma_{bt}\delta^a_s\delta^d_q - \gamma_{pb}\gamma_{t\mu}\delta^a_s\delta^d_q\\
&- \gamma_{p\mu}\gamma_{qs}\delta^a_b\delta^d_t - \gamma_{p\mu}\gamma_{qs}\gamma_{bt}\gamma^{ad} - \gamma_{pb}\gamma_{qs}\delta^a_\mu\delta^d_t - \gamma_{pb}\gamma_{qs}\gamma_{t\mu}\gamma^{ad} + \gamma_{qs}\gamma_{bt}\delta^a_\mu\gamma^d_p + \gamma_{qs}\gamma_{t\mu}\delta^a_b\delta^d_p\\
&- \gamma_{ps}\gamma_{q\mu}\delta^a_b\delta^d_t - \gamma_{ps}\gamma_{q\mu}\gamma_{bt}\gamma^{ad} - \gamma_{ps}\gamma_{qb}\delta^a_\mu\delta^d_t - \gamma_{ps}\gamma_{qb}\gamma_{t\mu}\gamma^{ad} + \gamma_{ps}\gamma_{bt}\delta^a_\mu\delta^d_q + \gamma_{ps}\gamma_{t\mu}\delta^a_b\delta^d_q.
\end{aligned} \quad (2.129)$$

There are several terms which cancel:

$$1-33, \qquad 3-31, \qquad 7-27, \qquad 9-25, \qquad 14-23, \qquad 16-24, \qquad 17-20, \qquad 18-22. \quad (2.130)$$

The 20 remaining terms are paired, and they correspond to the 10 terms in Eq. (2.127):

$$\begin{aligned}
&2+35 \sim 1, \qquad 4+36 \sim 3, \qquad 5+32 \sim 5, \qquad 6+34 \sim 7, \qquad 8+29 \sim 2\\
&10+30 \sim 4, \qquad 11+26 \sim 6, \qquad 12+28 \sim 8, \qquad 13+21 \sim 9, \qquad 15+19 \sim 10.
\end{aligned} \quad (2.131)$$

Part II

Applications

Chapter 3

Bypassing the cosmological constant problem

3.1 The quantum vacuum and gravitation

The ultimate goal of the emergent gravity program in the version we are considering in this thesis is the determination of a universality class of condensed-matter-like systems that include, at low energies, all the excitations that are described by the combination of the standard model of particle physics and gravity. When general relativity is taken as the description of the gravitational interaction, this very combination makes sense in principle as an effective field theory below the energy scale set by the Planck energy [136, 137]. This framework even leads to specific predictions concerning genuine quantum corrections on different processes [138, 139], though none of these have been verified up to date due to their smallness. However, this effective field theory displays a famous feature: the so-called cosmological constant problem.

That this is a recurrent problem in contemporary theoretical physics is demonstrated by the number of available reviews about it; see [31, 32, 33, 34] for a small sample. It is not our aim to study all the different aspects, suggestions and ramifications of this problem. On the contrary, we will give a precise (and simple) mathematical meaning to the problem and keep our discussion within this framework.

The root of the problem lies in the gravitating properties of the quantum vacuum. Let us first discuss the concept of the quantum vacuum in the absence of gravity, that is, within the frame of quantum field theory in flat spacetime. The quantum vacuum corresponds to the Poincaré-invariant (whether or not this symmetry is emergent is irrelevant) state of lowest energy. On top of this state, one can define non-vacuum states with a definite number of particles by using the corresponding creation operators, and evaluate the transition amplitudes between different states in a perturbative fashion. These calculations are pictorially represented by Feynman diagrams. From all these perturbative processes, in this chapter we are interested in those that preserve the vacuum state or, in other words, that do not contain physical particles. These correspond in terms of Feynman diagrams to vacuum bubbles: diagrams with no external legs that represent the (perturbative) view of the quantum vacuum as a *sea* of virtual particles [140].

In flat-spacetime quantum field theory, the linked-cluster theorem (see, e.g., [141] for a textbook discussion of the theorem) permits to show explicitly that the contributions of vacuum bubbles cancel out of correlation functions, so that they do not have any physical consequence, and therefore lack any operational meaning. This changes drastically if we include gravity in the discussion by means of general relativity and consider the resulting effective field theory. The decoupling of vacuum bubbles of the matter sector no longer holds as a result of the dependence of the spacetime volume form on the gravitational field: diffeomorphism invariance implies the coupling of gravity to these diagrams. The subsequent effect can be explicitly shown to lead to the renormalization of the cosmological constant [33].

A nontrivial running of the cosmological constant would not directly be worrisome. There are other well-known quantities in physics that are renormalized, such as the electron charge for instance. Indeed, in an effective field theory framework, for any coupling constant there will be a corresponding renormalization group equation that links the value of the constant with the energy scale at which it is measured. The trouble comes then from the specific form of the renormalization group equation for the cosmological constant.

This equation can be evaluated by a number of techniques, all of them giving equivalent results. For instance, one can evaluate the effective action of matter fields, with the introduction of a regulator μ. Using the heat kernel expansion [142] one can easily take into account the necessary counterterms that have to be added in order to renormalize the effective action. Later in this chapter we shall perform a similar procedure step by step. If Λ_0 is the bare cosmological constant, one gets then the equation

$$\Lambda = \Lambda_0 + C_1 \ln\left(\frac{\mu^2}{C_2}\right). \tag{3.1}$$

The occurrence of logarithms is due to the use of dimensional regularization to regulate the divergent integrals; using a hard cutoff would imply the presence of powers of the cutoff μ [99]. In Eq. (3.1) the values of C_1 and C_2 depend on the masses of the matter sector. For a simplified matter sector with only one massive particle with mass m one has, up to irrelevant numerical factors and suitable dimensional constants, $C_1 \sim m^4$ and $C_2 \sim m^2$. For a more involved particle spectrum one gets a repetition of the second term on the right-hand side of Eq. (3.1) for all the different particles.

This sets the stage for the cosmological constant problem. The observed value of the cosmological constant is several order of magnitude smaller than $Gc\,m^4/\hbar^3$ for any of the particles of the standard model [34]. Therefore, a change of order of magnitude on the regulator μ leads to a very large running when compared to the experimental value of the cosmological constant. If one believes Eq. (3.1), it seems difficult to justify the value of the cosmological constant measured in cosmology without invoking severe fine tunings. It is not even necessary to go to cosmological observations to detect this problem. Indeed, it is enough to consider an effective description of the physics of the solar system to demonstrate its existence [33, 143]: solar system experiments constrain the possible cosmological constant to be much smaller than the natural order of magnitude obtained from Eq. (3.1) when the numbers of the standard model of particle physics are used [33]. This tension should not be mixed, as it happens frequently in the literature, with the problem of explaining the observed value of the cosmological constant in cosmological observations (and related issues as, e.g., the cosmic coincidence problem). To explain this value one would have to consider a much wider range of scales, from very short lengths outside the domain of applicability of the effective field theory up to the Hubble length. Let us stress that the need of taking into account so large length scales does *not* make the problem bona fide infrared.

An additional complication arises when the effect of phase transitions on the cosmological constant is taken into account. The standard formalism to describe phase transitions in the cosmological evolution of our universe leads to large shifts of the cosmological constant across these transitions. This is however a phenomenon of different nature than the renormalization group equation of the cosmological constant, therefore requiring a separate study which is out of the scope of this thesis. Note also that, as Weinberg stresses [31], there is no observational evidence that refutes the (calculable) effects of phase transitions on the cosmological evolution of our universe through the corresponding changes on the effective cosmological constant. In other words, there is no evidence that it could not be much bigger in the past; one could say even the contrary, i.e., that this could conform with the nowadays standard inflationary picture. But even solar system observations lead to strong tensions with the renormalization group equation (3.1). Indeed, what prevents to accept that the cosmological constant is a parameter that has to be fixed by observations, as any other fundamental constant in physics as the gravitational constant or the electron charge, is this very same equation (and similar equations that are obtained for higher orders in perturbation theory). This explains our focus on this feature, a perspective that is shared by many reviews; see, for instance, [33, 34].

In the previous chapter we have discussed that, apart from general relativity, there is only one nonlinear theory of gravity that preserves the internal symmetries that are characteristic of the linear description of free gravitons: Weyl-transverse gravity. It is therefore worth exploring, from an emergent perspective, the possibility that the effective low-energy description of the gravitational interaction is given by the latter theory. The determination of the potential differences between these two choices represents an interesting field of study [120]. Even if there are arguments that point to the degeneracy of these two theories at the classical level when certain conditions are met, there exists the possibility that differences are triggered by quantum effects. We may use the following clear analogy: radiative corrections can be understood as perturbations with respect to the tree-level physics; these differences would be equivalent to the (quite common) degeneracy breaking by perturbations in eigenvalue problems. The only known difference, which represents the main result of this chapter, concerns the renormalization group of the gravitational action in the presence of quantum matter fields.[1] We will show explicitly that the renormalization group (3.1) for the cosmological constant is no longer present in the alternative framework, and interpret this result in terms of well-known field-theoretical concepts. For the purposes of this chapter, the term cosmological constant will always refer to the corresponding quantity in the Einstein field equations, and not to any parameter occurring in the gravitational action. The reason for pointing this difference is that, while in general relativity these two notions coincide, this is no longer true in Weyl-transverse gravity.

It is convenient to review briefly the form of the classical field equations in Weyl-transverse gravity to highlight this point. The gravitational symmetries can be exploited in order to fix a gauge in which the field equations take the same form as the traceless Einstein field equations [121, 122]:

$$\mathscr{R}_{ab} - \frac{1}{4}\mathscr{R} g_{ab} = \kappa \left(T_{ab} - \frac{1}{4} T g_{ab} \right). \tag{3.2}$$

These correspond to nine partial differential equations ($\kappa = 2\lambda^2$ where λ is the coupling constant in the previous chapter). As explained in detail in [134, 135], under the condition of the covariant conservation of the source T_{ab} (that makes for a tenth equation) one recovers the full set of Einstein field equations, with $\Lambda := (\mathscr{R} + \kappa T)/4$ an integration constant,

$$\mathscr{R}_{ab} - \frac{1}{2}\mathscr{R} g_{ab} + \Lambda g_{ab} = \kappa T_{ab}. \tag{3.3}$$

The parameter Λ is the quantity that we refer to as the (effective) cosmological constant. In general relativity this quantity is directly linked to a coupling constant in the gravitational action, but in Weyl-transverse gravity there is not such a connection.

3.2 The cosmological constant in Weyl-transverse gravity

Afters years of development of the subject known as quantum field theory in curved spacetimes, the evaluation of the renormalization group for the coupling constants of the (classical) gravitational action in the presence of quantum matter field is a well-known procedure. We shall apply the standard recipes that have been worked out using general relativity as the description of the gravitational interaction, and spotlight the differences that appear in the path.

[1]Note that it has been recently argued that there should be no differences between these theories even with the inclusion of quantum effects [144]. Nevertheless, in the words of these authors their arguments are "less than rigorous". Indeed, the results of this chapter provide a clear counterexample to these incomplete claims.

3.2.1 The semiclassical renormalization group

Let us start by recalling the form of the action of Weyl-transverse gravity. With some minor notational changes with respect to the previous chapter, this action looks like

$$\mathscr{A} := \frac{1}{2\kappa}\int \mathrm{d}^4x\sqrt{|\omega|}\,\mathscr{R}(\hat{g}). \tag{3.4}$$

Here $\mathscr{R}(\hat{g})$ has the same functional form as the Ricci scalar of a metric $\hat{g}_{ab} := |\omega|^{1/4}|g|^{-1/4}g_{ab}$, with g the determinant of g_{ab}. We have introduced an auxiliary, non-dynamical volume form $\boldsymbol{\omega}$ (a nowhere-vanishing differential form of degree $D = 4$), with $\mathrm{d}^4x\sqrt{|\omega|}$ the corresponding volume element in $D = 4$ spacetime dimensions. The contents of this chapter are trivially extendable to arbitrary dimensions but, unless stated explicitly, we shall work in this particular dimensionality. In the discussion of the previous chapter, the volume form $\boldsymbol{\omega}$ was identified with the volume form of the underlying Minkowski spacetime, though this identification is not mandatory in a generic framework. Although reminiscent of unimodular gravity as presented, e.g., in [133, 31], the theory described by the action (3.4) has as field variable an unconstrained second-rank tensor field g_{ab}. Unimodular gravity or, equivalently, the traceless Einstein field equations (3.3), correspond to the specific gauge choice in which the determinant of the gravitational field g_{ab} is fixed to be $|g| = |\omega|$.

As in general relativity, matter is coupled to gravity by following a minimal coupling approach, but replacing η_{ab} with the composite field $\hat{g}_{ab} = |\omega|^{1/4}|g|^{-1/4}g_{ab}$. The resulting gravity-matter action is invariant under gravitational scale transformations, or local scale transformations of the gravitational field, defined as local Weyl transformations that do not affect matter fields. Given a theory of matter with an arbitrary combination of matter fields with different spin (0, 1/2 and 1), minimally coupled to $\hat{g}_{ab} = |\omega|^{1/4}|g|^{-1/4}g_{ab}$, the heat kernel expansion [142] permits to write the regulated effective action so that one can take account of the necessary counterterms and obtain the renormalization group for the gravitational couplings.

Let us work explicitly with a scalar field only, as the generalization to other kinds of matter fields is straightforward. We will follow closely the discussion in [142, 99]. The evolution equations are given by

$$\mathcal{O}_g\phi = \frac{1}{\sqrt{|\omega|}}\partial_a\left(\sqrt{|\omega|}\,\hat{g}^{ab}\partial_b\phi\right) + \left(\frac{mc}{\hbar}\right)^2\phi + \xi\,\mathscr{R}(\hat{g})\phi = 0. \tag{3.5}$$

The parameters m and ξ are real, but otherwise arbitrary. The first part of the differential operator corresponds to the d'Alembert operator associated with $\hat{g}_{ab} = |\omega|^{1/4}|g|^{-1/4}g_{ab}$.

Let us now extract the information encoded in the one-loop effective action $\mathscr{S}_g$, which is essentially the functional determinant of the differential operator $\mathcal{O}_g$. Following the usual practice, when performing manipulations with path integrals some expressions should be understood in the Euclidean sense, in order to guarantee that all is well defined (the reader can read the details on this as well as the standard conventions we follow in, e.g., [145]). First, the classical action leading to the equations of motion (3.5) is given by

$$S[\phi, g] = \frac{1}{2}\langle\phi, \mathcal{O}_g\phi\rangle = \frac{1}{2}\int_{\mathscr{M}}\boldsymbol{\omega}\,\phi\,\mathcal{O}_g\phi, \tag{3.6}$$

where we have used the following notation for the inner product

$$\langle \phi, \phi' \rangle := \int \mathrm{d}^4 x \sqrt{|\hat{g}|}\, \phi(x) \phi'(x) = \int_{\mathscr{M}} \boldsymbol{\omega}\, \phi\, \phi'. \tag{3.7}$$

This inner product is defined for every $\boldsymbol{\omega}$ in a coordinate-free way as the integral of the differential form $\boldsymbol{\omega}\, \phi\, \phi'$ (see, e.g., the corresponding appendix in [146]). Integrating by parts twice we can show that the operator $\mathcal{O}_g$ is symmetric in this inner product, namely

$$\langle \phi, \mathcal{O}_g \phi' \rangle = \langle \mathcal{O}_g \phi, \phi' \rangle. \tag{3.8}$$

We can then perform a decomposition in terms of the eigenfunctions $\{\phi_n\}_{n=1}^{\infty}$ of this operator. This permits to write the action as

$$S[\phi, g] = \frac{1}{2} \langle \phi, \mathcal{O}_g \phi \rangle = \frac{1}{2} \sum_{m=1}^{\infty} \lambda_m c_m^2, \tag{3.9}$$

where $\{c_m\}_{m=1}^{\infty}$ are the coefficients of the expansion in eigenfunctions and $\{\lambda_m\}_{m=1}^{\infty}$ the corresponding eigenvalues. Following the usual normalization conventions, the measure is defined in terms of the $\{c_m\}_{m=1}^{\infty}$ as

$$[\mathscr{D}\phi] := \prod_{n=1}^{\infty} \frac{\mathrm{d}c_n}{\sqrt{2\pi}}, \tag{3.10}$$

so that the path integral is formally given by

$$\begin{aligned} \int [\mathscr{D}\phi] \exp(-S[\phi, g]) &= \int_{-\infty}^{\infty} \left(\prod_{n=1}^{\infty} \frac{\mathrm{d}c_n}{\sqrt{2\pi}} \right) \exp\left(- \sum_{m=1}^{\infty} \lambda_m c_m^2 / 2 \right) \\ = \prod_{n=1}^{\infty} \int_{-\infty}^{\infty} \frac{\mathrm{d}c_n}{\sqrt{2\pi}} \exp\left(-\lambda_n c_n^2/2\right) &= \prod_{n=1}^{\infty} \lambda_n^{-1/2} = \det{}^{-1/2}(\mathcal{O}_g). \end{aligned} \tag{3.11}$$

The one-loop effective action is defined through the relation

$$\exp(-\mathscr{S}_g) := \int [\mathscr{D}\phi] \exp(-S[\phi, g]), \tag{3.12}$$

or, what is equivalent,

$$\mathscr{S}_g = \frac{1}{2} \ln[\det(\mathcal{O}_g)]. \tag{3.13}$$

Now the following integral representation of the real logarithm

$$\ln(x) = - \lim_{\epsilon \to 0} \int_{\epsilon}^{\infty} \frac{\mathrm{d}s}{s} \left[\exp(-xs) - \exp(-s) \right] \tag{3.14}$$

is useful in order to extract the information encoded in the effective action (3.13). Given two real numbers $\alpha, \beta \in \mathbb{R}$ we have

$$\ln(\alpha/\beta) = \lim_{\epsilon\to 0}\int_\epsilon^\infty \frac{\mathrm{d}s}{s}\left[\exp(-s\beta) - \exp(-s\alpha)\right]. \tag{3.15}$$

If we understand these α, β as eigenvalues of the operators $\mathcal{O}_g$, $\mathcal{O}_{g^0}$ corresponding to two different configurations of the gravitational field, and use $\ln\det(\mathcal{O}_g) = \operatorname{Tr}\ln(\mathcal{O}_g)$, we can write then[2]

$$\mathscr{S}_g - \mathscr{S}_{g^0} = \frac{1}{2}\ln\left[\det(\mathcal{O}_g)/\det(\mathcal{O}_{g^0})\right] = \frac{1}{2}\lim_{\epsilon\to 0}\int_\epsilon^\infty \frac{\mathrm{d}s}{s}\operatorname{Tr}\left[\exp(-s\,\mathcal{O}_{g^0}) - \exp(-s\,\mathcal{O}_g)\right]. \tag{3.16}$$

From the way this relation is obtained it is clear that it will be also satisfied for other kinds of matter fields, but in terms of the corresponding differential operators.

While the relation involving the logarithm of two real numbers, Eq. (3.15), is well defined in the limit $\epsilon \to 0$, this is not true for Eq. (3.16) due to the infinite dimensionality of the functional space in which the scalar field ϕ is defined. To isolate these divergences, we can use the heat kernel expansion in the $s \to 0$ limit,

$$\operatorname{Tr}\left[\exp(-s\,\mathcal{O}_g)\right] = \int \mathrm{d}^4x\frac{\sqrt{|\omega|}}{(4\pi s)^2}\left[a_0(\hat{g}) + a_1(\hat{g})s + a_2(\hat{g})s^2 + \mathscr{O}(s^3)\right]. \tag{3.17}$$

We have only written explicitly the terms causing divergences in the $s \to 0$ limit, which are associated with the first Seeley-DeWitt coefficients, $\{a_n\}_{n=0,1,2}$. These coefficients are given (see [99] for instance) by

$$\begin{aligned}
&a_0(\hat{g}) = 1,\\
&a_1(\hat{g}) = k_1\mathscr{R}(\hat{g}) - m^2,\\
&a_2(\hat{g}) = k_2\mathscr{C}_{abcd}(\hat{g})\mathscr{C}^{abcd}(\hat{g}) + k_3\mathscr{R}_{ab}(\hat{g})\mathscr{R}^{ab}(\hat{g}) + k_4\mathscr{R}^2(\hat{g}) + k_5\Box\mathscr{R}(\hat{g}) -\\
&- (mc/\hbar)^2k_1\mathscr{R}(\hat{g}) + \frac{1}{2}(mc/\hbar)^4.
\end{aligned} \tag{3.18}$$

In these expressions, $\{k_i\}_{i=1,2,3,4,5}$ are dimensionless real constants that depend on the kind of field (that is, the differential operator) one is considering [142], and m the mass of the field. All the curvature tensors and invariants, namely the Weyl tensor $\mathscr{C}_{abcd}(\hat{g})$, the Ricci tensor $\mathscr{R}_{ab}(\hat{g})$ and the Ricci scalar $\mathscr{R}(\hat{g})$, are evaluated on $\hat{g}_{ab}$.

The divergent behavior in the effective action is absorbed by means of the renormalization of the gravitational couplings. The necessary counterterms to do so can be read from the following expression, directly obtained from the combination of Eqs. (3.16) and (3.17):

$$\begin{aligned}
\mathscr{S}_g - \mathscr{S}_{g_0} = &\frac{1}{32\pi^2}\left[\lim_{\epsilon\to 0}\int_\epsilon^\infty \frac{\mathrm{d}s}{s^2}\right]\int_{\mathscr{M}}\boldsymbol{\omega}\left[a_1(\hat{g}^0) - a_1(\hat{g})\right]\\
&+ \frac{1}{32\pi^2}\left[\lim_{\epsilon\to 0}\int_\epsilon^\infty \frac{\mathrm{d}s}{s}\right]\int_{\mathscr{M}}\boldsymbol{\omega}\left[a_2(\hat{g}^0) - a_2(\hat{g})\right]
\end{aligned} \tag{3.19}$$

[2]Considering the difference between the effective action for two configurations of the gravitational field makes some of the subsequent expressions to be rigorously defined [142].

The same structure is valid for other kinds of matter fields, just by changing the numeric factors $\{k_i\}_{i=1,2,3,4,5}$ inside of the Seeley-DeWitt coefficients. Once Eq. (3.19) has been obtained, the usual procedure is to regulate the divergent integrals, for instance by introducing a sharp cutoff μ, or by using dimensional regularization.

From this equation we can notice that there is no term corresponding to the $a_0 = 1$ Seeley-DeWitt coefficient, in contrast to what happens in general relativity. This is a consequence of the invariance under gravitational scale transformations, which enforces $\sqrt{|\hat{g}|} = \sqrt{|\hat{g}^0|}$ so that these contributions are independent of the gravitational field. The corresponding piece in general relativity leads to the renormalization of the cosmological constant. In general relativity there are also additional terms that renormalize the cosmological constant, coming from the constant pieces (that is, independent of the gravitational field) of the Seeley-DeWitt coefficients a_1 and a_2 as displayed in Eq. (3.18). These automatically cancel out in Weyl-transverse gravity, as one can read directly from (3.19). Alternatively, the corresponding pieces of the effective action are independent of the physical fields, and are therefore unobservable. This is the main result of this chapter: in Weyl-transverse gravity there is *no* renormalization group equation for the cosmological constant.

The first nonzero contribution in Eq. (3.19), namely the term proportional to $R(\hat{g})$ in $a_1(\hat{g})$ [see Eq. (3.18)], leads to a renormalization of the gravitational coupling constant κ. If we call κ_0 the bare gravitational coupling constant and use dimensional regularization, one can read from the one-loop effective action (3.19) the equation

$$\frac{1}{\kappa} = \frac{1}{\kappa_0} + C_3 \ln\left(\frac{\mu^2}{C_4}\right), \tag{3.20}$$

where C_3 and C_4 are constants with convenient physical dimensions and whose values depend on the particle content of the matter sector [99]. The next contributions involve quadratic expressions in the Ricci tensor $\mathscr{R}_{cd}(\hat{g}_{ab})$ as one can read from the expression of $a_2(\hat{g}_{ab})$ given in Eq. (3.18). These terms, that also appear in general relativity, respect the gravitational symmetries and lead to higher-derivative deviations from the second-order field equations at high energies. These imply the running of the corresponding coefficients in front of these quadratic terms in the Lagrangian density, a feature that is irrelevant for the point we want to make here (the form of these renormalization group equations is the same as in the general relativity case and can be consulted in [99] for instance).

To sum up, the important observation is that the cosmological constant is not renormalized in Weyl-transverse gravity. This parameter, that appears in the equations of motion as an integration constant, is not subjected to Eq. (3.1). This avoids the corresponding radiative instability and therefore the cosmological constant problem.

Let us finish this section with some comments regarding the significance of this result in terms of Feynman diagrams, and in particular vacuum bubbles. As discussed in the introduction, these diagrams are responsible for the renormalization of the cosmological constant in general relativity through their coupling to the gravitational field. In other words, the dependence of the spacetime volume form on the gravitational field that is dictated by diffeomorphism invariance implies the coupling of gravity to these diagrams.

In Weyl-transverse gravity, the internal symmetries imply that the spacetime volume form cannot depend on the gravitational field, and therefore the decoupling of the contributions of vacuum bubbles. The (formally infinite) contributions of these diagrams is as harmless as it is in flat-spacetime quantum field theory, given that no observable quantity is affected by them.

3.2.2 Absence of anomalies in Weyl-transverse gravity

The discussion in the previous section about the distinctive feature of Weyl-transverse gravity that the cosmological constant is unchanged by radiative corrections may benefit of showing a different perspective, in order help building up an intuition on this result. Let us focus our attention on the action of low-energy effective field theory. As long as we are concerned only with the gravitational sector, it is straightforward to check that only requiring the existence of a global symmetry in the form of constant Weyl transformations

$$g_{ab} \to \zeta^2 g_{ab}, \tag{3.21}$$

with ζ a real constant, would suffice to forbid the occurrence in the effective Lagrangian density of any cosmological constant term. However, adding matter suggests that there is no way out of the cosmological constant problem by following this path (see, e.g., the discussion in [34]): first, only massless fields are allowed in order to preserve the symmetry classically; even in this case, in general the symmetry can no longer be maintained when the matter fields are quantized, with the occurrence of what is known as the Weyl or conformal anomaly [147, 148, 149].

Nevertheless, previous analyses leading to these conclusions have been based in general relativity. As we shall discuss next, in Weyl-transverse gravity the local extension of the transformation (3.21) corresponds to a genuine (non-anomalous) symmetry. In this alternative framework one would therefore expect that the symmetry (3.21) protects the cosmological constant sector, rendering it radiatively stable. This is compatible with our discussion in the previous section, in which we have shown that the cosmological constant is not renormalized [in other words, that Eq. (3.1) is not present] in Weyl-transverse gravity. However, to guarantee that this picture is satisfactory, this gravitational scale invariance must survive quantum effects, i.e., be free of anomalies in the presence of quantum matter fields.

Any reader familiar with previous results on scale-invariance anomalies might be skeptical about our statement of protection of the cosmological constant term by means of gravitational scale invariance. While this first reaction is justified, the crucial observation we want to make here is that former results must be revised in the case at hand as one of the previous assumptions has been dropped off: invariance under the entire group of diffeomorphisms. The presence of anomalies can indeed be traced back to this assumption alone, being a feature of the interplay between longitudinal diffeomorphisms and scale transformations. The key resides in the extension of the symmetry (3.21) when we include matter in the game. We encourage the reading of previous work on the subject, namely the corresponding appendix in [150], where some previous comments regarding the issue of anomalies in Weyl-transverse gravity are presented, and [151] for the study of pure gravity and couplings to *conformal* matter. To the best of our knowledge, the definition and properties of the path integral for a general matter content coupled to (classical) Weyl-transverse gravity have not been discussed before.

In the standard case in which it is assumed that the gravitational action is invariant under diffeomorphisms, the couplings between gravitational and matter fields dictated by symmetry considerations generally imply that matter fields should transform nontrivially under scale transformations, as well as being massless, if one wants the preserve the symmetry (3.21) of the matter action. This results in standard conformal invariance (the situation with scale invariance in quantum chromodynamics is completely parallel). On the contrary, in Weyl-transverse gravity the matter-field Lagrangian density is constructed from a general flat-space Lagrangian density by replacing the flat metric η_{ab} by $\hat{g}_{ab} = |\omega|^{1/4}|g|^{-1/4}g_{ab}$ instead of g_{ab}, to guarantee invariance under gravitational scale transformations. Matter fields are inert under these transformations so that no restrictions whatsoever apply. Obviously, this kind of coupling would explicitly break the invariance under longitudinal diffeomorphisms, if present.

To describe anomalies we shall use Fujikawa's approach [152], which is especially suited for making our point in a clean and concise way. A specific anomaly will be given by a (regulated) Jacobian associated with the change of the path integral measure under a given symmetry. Even if gravitational scale transformations do not affect matter fields by construction, this symmetry could be anomalous. The best example of this is a scalar field in two dimensions: in this particular case, the scalar field by itself is unchanged by conformal transformations, but still the symmetry is anomalous as the result of the definition of the path integral measure.

For clarity let us briefly recall the well-known two-dimensional conformal anomaly (see, e.g., [147, 153] and references therein) to compare it with that of gravitational scale invariance (for the purposes of this section only, we will work in an arbitrary dimension D). We will see that the conformal anomaly arises because of the interplay of conformal transformations and longitudinal diffeomorphisms. That is, the classical conformal symmetry is broken by quantum effects if one uses the path integral measure which preserves invariance under arbitrary diffeomorphisms. For the purposes of this example only, the object g_{ab} will momentarily regain its standard geometric interpretation.

To define the path integral measure and, so, the path integral itself, we first define a scalar product

$$(\phi, \phi') := \int \mathrm{d}^2x \sqrt{|g|}\, \phi(x)\phi'(x) = \int_{\mathscr{M}} \boldsymbol{\epsilon}\, \phi\, \phi', \tag{3.22}$$

in which the differential operator occurring in the classical field equations for the scalar field is symmetric ($\boldsymbol{\epsilon}$ is the Levi-Civita tensor [146]), and therefore we can use the corresponding decomposition in terms of the eigenfunctions $\{\bar{\phi}_n\}_{n=1}^{\infty}$ of this operator. The coefficients of the expansion are given by

$$b_n := (\bar{\phi}_n, \phi) = \int_{\mathscr{M}} \boldsymbol{\epsilon}\, \bar{\phi}_n\, \phi. \tag{3.23}$$

The classical action can be written entirely in terms of these coefficients. The path integral measure is defined as the natural measure in the infinite-dimensional space spanned by the $\{b_n\}_{n=1}^{\infty}$, being the path integral formally defined then as a functional determinant.

By construction, the transformation properties of the measure are directly inherited from the transformation properties of the inner product (3.22) and the fields, and can be directly evaluated from (3.23). The general result that we will use is the following [147]: symmetries such that $\delta b_n = 0$ are directly free of anomalies, being thus preserved in the quantum realm, while if $\delta b_n \neq 0$ additional manipulations are needed in order to extract a meaningful finite result.

The path integral measure constructed from the $\{b_n\}_{n=1}^{\infty}$ is invariant under diffeomorphisms by construction. Applying the corresponding transformation laws of the fields to the coefficients b_n as defined in Eq. (3.23), one can check that these are invariant. However, the $\sqrt{|g|}$ factor in the inner product (3.22) enforced by diffeomorphism invariance implies that the coefficients (3.23) are not invariant under conformal transformations. Under an infinitesimal conformal transformation $\delta g_{ab} = \alpha(x) g_{ab}$ one has:

$$\delta b_n = \int_{\mathscr{M}} \boldsymbol{\epsilon}\, \bar{\phi}_n\, \phi\, \alpha. \tag{3.24}$$

It is an algebraic matter to evaluate the Jacobian J associated with the path integral measure for infinitesimal $\alpha(x)$ as:

$$\ln J = \lim_{N\to\infty} \sum_{n=1}^{N} \int_{\mathscr{M}} \boldsymbol{\epsilon}\, \bar{\phi}_n\, \bar{\phi}_n\, \alpha. \tag{3.25}$$

This expression must be regularized in the $N \to \infty$ limit but, after a proper treatment, it leads to the usual results of the conformal anomaly: the trace of the stress-energy tensor of matter fields is no longer zero in general gravitational backgrounds so that conformal invariance is lost [147, 148, 149].

We can apply the same procedure to show that all the local symmetries of Weyl-transverse gravity are free of anomalies in the presence of quantum matter. Again, the first we need to do is to define the path integral measure for the matter fields. Let us detail the scalar field case, to extend it later to a general matter content. In view of our reasoning in previous sections, the inner product for a scalar field in Weyl-transverse gravity in D dimensions will be given by

$$\langle \phi, \phi' \rangle := \int \mathrm{d}^D x \sqrt{|\hat{g}|}\, \phi(x) \phi'(x) = \int_{\mathscr{M}} \boldsymbol{\omega}\, \phi\, \phi'. \tag{3.26}$$

Notice the essential difference with respect to Eq. (3.22), in that this definition displays an auxiliary volume form unrelated to the gravitational field. The inner product is defined for every $\boldsymbol{\omega}$ in a coordinate-free way as the integral of the $D-$form $\boldsymbol{\omega}\, \phi\, \phi'$ [146].

The evolution operator $\mathcal{O}_g$ is symmetric in this inner product. In the following we consider the expansion in terms of its eigenfunctions $\{\phi_n\}_{n=1}^{\infty}$, which permits to construct the path integral from the coefficients,

$$c_n := \langle \phi_n, \phi \rangle = \int_{\mathscr{M}} \boldsymbol{\omega}\, \phi_n\, \phi, \tag{3.27}$$

as the functional determinant given in Eq. (3.11).

By construction, the classical action is invariant under the symmetries of the theory. On the other hand, the transformation properties of the path integral measure can be read off from Eq. (3.27). We can explicitly check that these coefficients and, hence, the corresponding path integral measure, are invariant under both transverse diffeomorphisms and gravitational scale transformations. It is enough to notice that under these transformations $\delta\sqrt{|\hat{g}|} = 0$ (or, equivalently, that $\boldsymbol{\omega}$ is invariant), thus implying

$$\delta c_n = \int \mathrm{d}^D x\, \phi_n(x)\phi(x)\, \delta\sqrt{|\hat{g}|} = 0. \tag{3.28}$$

In contraposition with the former case we see that, as we are not demanding invariance under longitudinal diffeomorphisms, the factor $\sqrt{|g|}$ can be freely tuned to accommodate the remaining gauge symmetries, generated by gravitational scale transformations. This is what we are effectively realizing when inserting $\hat{g}_{ab} = |\omega|^{1/D}|g|^{-1/D}g_{ab}$ instead of g_{ab} in the expressions for the inner products. These considerations can be extended to *any* matter content, with fields of arbitrary spins and masses, as we can always construct an inner product with these invariance properties as long as we leave aside longitudinal diffeomorphisms and make use of the usual recipes but with the gravitational field represented through the combination $|\omega|^{1/D}|g|^{-1/D}g_{ab}$.

A complementary way to realize this is the following. It is a general result that anomalies may occur in the one-loop effective action when the differential operator appearing in the equations of motions for the fields is not *invariant* ($\mathscr{O}'_g = \mathscr{O}_g$) under a symmetry transformation, but only *covariant* ($\mathscr{O}'_g \propto \mathscr{O}_g$) [142]. This description is of course parallel to the previous discussion in terms of the measure (a non-invariant operator would imply a non-invariant measure in the path integral). Diffeomorphism invariance enforces a coupling to the gravitational field g_{ab} which implies that it is not possible to make the differential operator invariant under conformal transformations at the same time. On the contrary, when coupling the matter fields to $\hat{g}_{ab} = |\omega|^{1/D}|g|^{-1/D}g_{ab}$ we are giving up invariance of the field equations under longitudinal diffeomorphisms, while making the corresponding differential operators invariant under gravitational scale transformations. An example is the scalar field differential operator defined in Eq. (3.5), which is clearly invariant under gravitational scale transformations. Let us illustrate how this works with a different kind of matter field, e.g., a fermion field. It is straightforward to consider more complicated matter contents, such as the one in the standard model of particle physics, following the recipes for the coupling to the gravitational field given above. Fermions couple to the composite vierbein field, defined by means of the relation $\hat{g}_{ab} = \hat{e}^c{}_a \hat{e}^d{}_b \eta_{cd}$ or, in terms of the usual vierbein field,

$$\hat{e}^b{}_a = \frac{|\omega|^{1/2D}}{|e|^{1/D}} e^b{}_a, \tag{3.29}$$

with e the determinant of $e^b{}_a$. The Dirac operator in Weyl-transverse gravity will be given then by

$$\mathcal{D}_g = i\gamma_b \hat{e}^b{}_a D^a - Mc/\hbar. \tag{3.30}$$

Here M is the fermion mass and D_a the covariant derivative associated with $\hat{e}^b{}_a$ by means of the corresponding spin connection (see, e.g., [2]), which in our case is given in terms of the Weyl connection,

$$\begin{aligned}\hat{\Gamma}^c_{ab} &= \frac{1}{2}\hat{g}^{cd}\left(\partial_a \hat{g}_{bd} + \partial_b \hat{g}_{ad} - \partial_d \hat{g}_{ab}\right) \\ &= \frac{1}{2} g^{cd}\left(\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}\right) + \frac{1}{2D}\left[\delta^c_b \partial_a + \delta^c_a \partial_b - g_{ab}\partial^c\right] \ln(|\omega|/|g|),\end{aligned} \tag{3.31}$$

as

$$\hat{\omega}^{pq}_a = \hat{e}^p{}_b \partial_a \hat{e}^{q\,b} - \hat{e}^{p\,b} \hat{e}^q{}_c \hat{\Gamma}^c_{ab}. \tag{3.32}$$

While its counterpart in general relativity is just covariant under conformal transformations, the Dirac operator (3.30) is invariant under gravitational scale transformations, as we have claimed.

Up to now we have covered in the discussion particles with spin 0, 1/2 and 1, with general properties and interactions between them. Let us make a remark concerning the spin-2 case which, although not essential for our semiclassical discussion, may be interesting for future developments. When it comes to the quantum properties of gravity itself, the definition of the path integral is subtler but a tentative exploration of the path integral measure following previous works (see, e.g., [154] and references therein) shows that it should be possible to define it in terms of the composite field $\hat{g}_{ab} = |\omega|^{1/D}|g|^{-1/D} g_{ab}$ instead of g_{ab}. This procedure would lead to a non-anomalous path integral with respect to the internal symmetries. Beyond the scope of this thesis, it would be very interesting to carry out this program in detail as well as a study of the possible properties of a quantum theory of gravity with these symmetries. We refer the reader to the recent papers [155, 156] for further advances along these lines.

3.3 Keeping the cosmological constant small at all scales

We have shown that the cosmological constant is stable against radiative corrections in Weyl-transverse gravity, in stark difference with the situation in general relativity. In this section we take a step back to consider a more general discussion concerning the minimal requirements that a theory has to verify in order to avoid the cosmological constant problem, and argue that Weyl-transverse gravity displays all of them. We shall also make some brief comments on a proposal by Volovik to fix the value of the effective cosmological constant that fits naturally in this framework.

3.3.1 Stabilizing the cosmological constant

That Eq. (3.1) does not make appearance in Weyl-transverse gravity can be seen from different, but complementary perspectives. As discussed in Sec. 3.2.2, one can understand this feature as a consequence of the symmetries of the gravitational action. However, the action of Weyl-transverse gravity (plus matter) presents an additional global symmetry, the occurrence of which is intimately related to its local symmetries, that offers a complementary view on the relation between the interplay between the quantum vacuum and the gravitational interaction in this theory.

When considering a field theory on flat spacetime, there is a global symmetry that tell us that only relative energies have physical meaning, namely the shift symmetry

$$\mathscr{L} \longrightarrow \mathscr{L} + C_0, \tag{3.33}$$

where $C_0 \in \mathbb{R}$ is a constant. One can trace back to this symmetry the decoupling of vacuum bubbles from correlation functions, and therefore from physical observables, that we have discussed in the introduction. The shift symmetry (3.33) is broken with the introduction of general relativity. Again, it is the introduction of a spacetime volume form that depends on physical fields the reason behind this feature.

It is therefore clear that maintaining the shift symmetry (3.33) is a necessary condition to deal satisfactorily with the cosmological constant problem (see the related discussions in [157, 158]). In order to guarantee this condition, there must exist a background volume form $\boldsymbol{\omega}$ so that the would-be cosmological constant term in the action is rendered innocuous to the classical dynamics.[3] At the same time, the classical field equations must contain an effective cosmological constant in order to match with cosmological observations. If we assume no deviations from classical physics, which is a well-motivated assumption given the quantum-mechanical nature of the cosmological constant problem, these equations should take the form (3.3). As explained in the introduction of this chapter, this is indeed the case in Weyl-transverse gravity [121, 122].

In terms of the parameters in the field equations (3.3), the fact that the shift transformation (3.33) is a global symmetry of the theory is expressed as

$$T_{ab} \longrightarrow T_{ab} + C_0 g_{ab}, \qquad \Lambda \longrightarrow \Lambda + \kappa C_0. \tag{3.34}$$

Therefore, the combination of the two necessary conditions to deal with the cosmological constant problem, namely that the shift symmetry (3.33) holds and that there exists an effective cosmological constant (or, in other words, that one essentially recovers the Einstein field equations), point to the symmetry (3.34) as a necessary condition that any alternative theory aiming to solve the cosmological constant problem must present. Note that the transformation (3.34) does *not* correspond to a symmetry in general relativity, as in that case Λ corresponds to a coupling constant that cannot be affected by symmetry transformations; in Weyl-transverse gravity Λ is an integration constant that acts as a

[3] It is also possible that ω depends on the physical fields, but its integral is a topological invariant.

label for different solutions and can be therefore shifted by means of genuine symmetry transformations.

Therefore Weyl-transverse gravity presents the symmetry (3.34). However, this theory goes further: following the effective field theory logic, this classical picture would not be enough. A symmetry that forbids the occurrence of $\sqrt{|g|}\,\Lambda$ in the Lagrangian density must be present in order to guarantee that this term is not generated by radiative corrections, which would spoil the shift symmetry (3.33). A natural candidate for the job is the invariance under scale transformations of the gravitational field,

$$g_{ab} \to \zeta^2 g_{ab}, \qquad \zeta \in \mathbb{R}. \tag{3.35}$$

Notice that the invariance under longitudinal diffeomorphisms of general relativity would be broken to guarantee the invariance under (3.35) while keeping us in a second-order field theory. In order to maintain the number of degrees of freedom in the gravitational sector, the symmetry (3.35) is then extended to a gauge symmetry. This discussion presents an alternative route to motivate Weyl-transverse gravity, as the simplest theory that contains the minimal ingredients that permits handling both the classical and quantum aspects of the cosmological constant problem.

From an effective field theory perspective, it is well understood that radiative corrections will generate all the terms in the Lagrangian density which are compatible with the symmetries of a given system. In this framework, the cosmological constant term corresponds to a relevant operator that is, however, not natural [34, 65]. Therefore the value of the cosmological constant is highly sensitive to the ultraviolet details beyond the effective theory, and it has to be fine-tuned in order to match the experiments. On the one hand, this observation is compelling: the value of the cosmological constant is an issue to be treated in a theory which consistently unifies the ultraviolet and infrared details of our universe, that is, a theory of quantum gravity. On the other hand, it challenges the basic working principle according to which the behavior of physics at a given distance scale is insensitive to the fine details of the dynamics at much shorter distances, but for a set of physical constants that are determined by the high-energy physics. It is therefore desirable that an effective field theory rationale exists that permits to work consistently at low energies, leaving the question about the cosmological constant unanswered (as for any other fundamental constant) until we are able to construct a more complete theory. What we have shown is that it is possible to construct a self-consistent low-energy effective field theory, if only replacing general relativity by Weyl-transverse gravity.

At the light of this discussion, it is interesting to consider an alternative proposal that has been recently presented [159]. In this proposal, the modification of just the purely global properties of general relativity permits to overcome the cosmological constant problem without changing the local physics. As discussed by the authors, this is again due to the presence of the symmetry (3.34). In this case, this symmetry is trivially satisfied due to the constraint

$$\Lambda = \frac{\kappa \int \mathrm{d}^4x \sqrt{g}\, T}{4 \int \mathrm{d}^4x \sqrt{g}}. \tag{3.36}$$

It is then clear to see that this proposal "overkills" the cosmological constant problem: recall that the necessary condition to deal with this problem is that the shift symmetry (3.34) holds. This is a statement that concerns only the transformation properties of the effective cosmological constant. But Eq. (3.36) presents a stronger statement, as it imposes a constraint on the effective cosmological constant that is not necessary in order to guarantee that the shift transformation (3.34) is a symmetry. All the bizarre features of this proposal, such as the inclusion of global variables and the corresponding causality violations [160] (note that to evaluate the effective cosmological constant (3.36) in a cosmological context one would need to know the entire evolution of the universe in all its points), can be traced back to the constraint (3.36). What we want to emphasize is that these are additional features that are not needed at all to deal with the cosmological constant problem, as shown by the discussion of this section, and the example of Weyl-transverse gravity.

3.3.2 Setting its value at high energies

Once the cosmological constant (the integration constant that enters the gravitational field equations) is shown to be unaffected by radiative corrections, one may think about the kind of principles that could fix its value. In principle, only a more fundamental theory or principle that goes further than the low-energy effective field theory description (Weyl-transverse gravity plus matter fields) could unveil its nature and set its actual value, which by matching should be the one used at low energies. This explains the title of this brief section: it is natural to expect that the value of the cosmological constant can be obtained in a suitable ultraviolet completion.

In particular, this principle may be related to the vacuum of the high-energy theory. In this regard, Volovik's proposal [161] is one of the most compelling proposals from a physical perspective one can find in the literature. This proposal fits quite naturally in the general theme of this thesis of emergence from condensed-matter-like models, as it indeed arises from the study of quantum liquids. The discussion here follows the original formulation by Volovik [35, 36, 161], with just a subtle (but important) deviation: in the original discussion one gets to the conclusion that the effective field theory must fail dramatically when evaluating some quantities such as the running of the cosmological constant. Taking Weyl-transverse gravity instead as the description of the gravitational interaction permits to overcome this conclusion, thus reconciling the principles of effective field theory and Volovik's arguments in a self-consistent combination. Indeed, one may argue that Weyl-transverse gravity provides the most natural realization of some of the ideas of this author. In our opinion, the following argument is worth further exploring due to its potential to justify the observed value of the cosmological constant.

The overall argument only rests in the assumption that our universe as a whole can be modeled as a quantum liquid, to which one can apply standard thermodynamical considerations [161]. The pressure in the vacuum state p_0 can be evaluated as the variation of the vacuum energy $\langle 0|\hat{H}|0\rangle$, where $|0\rangle$ is the ground state and $\hat{H}$ the Hamiltonian operator, with respect to the change of the volume of the system V. This leads to the equation

$$p_0 = -\frac{d\langle 0|\hat{H}|0\rangle}{dV} = -\frac{d[V\epsilon(N/V)]}{dV}$$
$$= -\epsilon(n) + n\frac{d\epsilon(n)}{dn} = -\frac{1}{V}\langle 0|\hat{H} - \mu\hat{N}|0\rangle. \tag{3.37}$$

In this equation, $\mu := d\epsilon/dn$ is the chemical potential of the system, $\hat{N}$ the number operator with N its mean value in the ground state, $\epsilon := \langle 0|\hat{H}|0\rangle/V$ the mean energy density and $n := N/V$ the mean density of particles. The vacuum pressure is then controlled by the grand-canonical Hamiltonian operator $\hat{H} - \mu\hat{N}$. Recall that this is essentially the same formalism we used in our discussion about emergent electrodynamics in Chap. 1. This leads to the observation that the ground state presents the equation of state $\rho_0 = -p_0/c^2$, typical of the cosmological constant, with

$$\Lambda = \kappa\rho_0 c^2 = \frac{\kappa}{V}\langle 0|\hat{H} - \mu\hat{N}|0\rangle. \tag{3.38}$$

This result is universal: it does not depend on the details of the Hamiltonian operator (regardless of the system being relativistic or not), the statistics of the atoms in the fluid nor the corresponding low-energy effective field theory [161]. Note that only liquid states can exist as isolated systems, that is, with no external pressures being applied.

Then, in the absence of external forces and neglecting surface terms, the pressure p_0 must be identically zero. This implies that the effective cosmological constant (3.38) is also zero. Deviations from the perfect equilibrium would induce a nonzero cosmological constant that could be compared with the observed value in cosmological scenarios, though more precise models have to be constructed in order to make definite assertions in this regard. Note that, following this argument, the effective cosmological constant would be unaffected by phase transitions as well.

This observation must be accompanied by a justification of the mismatch between the results one gets from Eqs. (3.1) and (3.38), when assuming that the low-energy effective field theory contains gravity as described general relativity. Volovik's argument is that the evaluation of the properties of the vacuum (in particular, the cosmological constant) is not responsibility of the low-energy effective field theory. Only the knowledge of the underlying high-energy theory, which in this case is assumed to take the form of a condensed-matter-like model (the specific model is yet to be constructed), can determine the properties of the cosmological constant. This implies a complete failure of the principles of effective field theory in this construction.

However, if the low-energy description of the gravitational interaction is assumed to be given by Weyl-transverse gravity, both high- and low-energy perspectives on the cosmological constant smoothly match. The low-energy effective theory shows unequivocally that the cosmological constant is alien to the effective description, representing a fundamental parameter that is unaffected by radiative corrections. In other words, Weyl-transverse gravity is the natural realization of the observation that the low-energy effective field theory should be oblivious to the properties of the quantum vacuum, but for a set of physical constants that are non-calculable in the framework of the effective low-energy theory. The effective cosmological constant is one of these non-calculable physical constants. Therefore, its value has to be fixed in a wider framework, namely a suitable ultraviolet completion of Weyl-transverse gravity.

3.4 Conclusions

In this chapter we have shown that Weyl-transverse gravity is the first known example in the literature of what we can call "minimal" solution to the cosmological constant problem: under natural assumptions the classical field equations are essentially equivalent to those of general relativity, while the cosmological constant can take arbitrary but radiatively stable values as it is protected by symmetries. Modifications with respect to the classical predictions of general relativity are only triggered by quantum effects, so that tree-level physics is preserved while one-loop and further corrections are changed. It is instructive to observe that the criteria demanded in [34] for a satisfactory solution to the cosmological constant problem are verified. From the perspective of the low-energy physics, the cosmological constant is in this framework as mysterious as (but not more than) any other parameter in physics, such as the gravitational constant or the electron charge. This resonates strongly with previous arguments by Volovik that are based in the potential emergence of the gravitational interaction from condensed-matter-like models that describe our universe as a quantum liquid.

This nonlinear gravitational theory has a strong first-principles justification rooted in local particle-like quantum properties of the gravitational interaction, as explained in the previous chapter of this thesis. This makes it especially suited to describe the infrared limit of a would-be theory of quantum gravity. From a more philosophical perspective, this proposal is inextricably tied up to nontrivial conceptual implications as one needs to accept that some properties of spacetime deviate from the ones associated with a pseudo-Riemannian manifold. This suggests that solving the cosmological constant problem may entail changing our conceptual and mathematical picture of spacetime.

As of future work, from this point different roads can be taken. It would be interesting to explore additional quantum properties of Weyl-transverse gravity as an effective theory, apart from the one which is the main subject of this chapter and which is in accordance with current observations, in order to further distinguish it from general relativity. Promising candidates to display differences may be scattering amplitudes involving the off-shell structure of gravitons, i.e., containing graviton loops. Regarding the cosmological constant, only a more fundamental theory or principle could unveil its nature and set its actual value, which by matching should be the one used in the classical solutions of the effective low-energy theory. The exploration of the nature of such a principle is an interesting issue in itself. Also the possible construction of theories of quantum gravity with non-anomalous gravitational scale invariance and their relation with the more familiar conformal field theories, appears as an attractive problem which could lead to profound implications for our understanding of the gravitational interaction and scale invariance.

Chapter 4

Black- to white-hole transition in gravitational collapse

4.1 Black holes

One of the main goals of any research program dealing with the ultraviolet details of the gravitational interaction is the resolution of the many paradoxes surrounding the formation and evolution of black holes. From an astrophysical perspective, black holes are currently accepted as members of the bestiary of astronomical objects. Their success is explained on the one hand by the simplicity and elegance of their mathematical properties, and on the other hand by the (indirect) observation of the existence of very dense, and intrinsically very dark, distributions of matter in our universe. For instance, observations of the center of our own galaxy point to the existence of an object with such characteristics that the only known theoretical structure in general relativity that can be identified with it is a supermassive black hole [162, 163, 164]. However, it is noteworthy that we lack direct evidence of the most characteristic property of a black hole, namely its horizon (this may actually be impossible [165]; for a short discussion of this issue see Sec. 4.6.1 below), or even of the very dynamical formation of black holes. The latter deficiency could eventually be alleviated in the future with the advent of gravitational wave astronomy [166].

Most importantly, black holes contain odd features that suggest that their theoretical portrait is far from completely understood. The most prominent of these features is the unavoidable presence of a singularity, a region in spacetime in which the known laws of physics break down [37, 38]. It is expected that the successful combination of quantum mechanics and gravity will lead to the regularization of these singularities. In the absence of a full theory of quantum gravity or, in a wider sense, of an ultraviolet completion of general relativity, there has nonetheless emerged, especially after the work of Hawking [38, 167], a consensus picture about the kind of ultraviolet modifications of the classical behavior one should expect. Naturally, the overall picture is far from being completely self-consistent (see for instance the information loss problem [39] or the recent firewall controversy [168]), and surprises may arise as our knowledge about the high-energy properties of the gravitational interaction improves.

The emergent gravity program explored in this thesis strongly suggests that a radically different view on the possible ultraviolet effects in gravitational collapse processes could be realized instead. An extensive discussion is presented in [40], though the emerging picture can be briefly motivated in a natural way within the framework of our discussion in the first part of this thesis. The possibility that the gravitational interaction might just be a collective phenomenon that arises at low energies implies its disappearance (or disconnection) at some point, namely when we go to high enough energy densities. This very phenomenon would take place for the electromagnetic interaction in the model described in Chap. 1. In terms of the formalism introduced in Chap. 2 to deal with the self-interactions of the field describing the gravitational interaction, this phenomenon would be equivalent to the limit $\lambda \to 0$ for the coupling constant λ that controls the nonlinearities of the gravitational sector. Taking into account that the action of a theory resulting from graviton self-interactions must be expressible in terms of a composite field $g_{ab} = \eta_{ab} + \lambda h_{ab}$ (for any nonzero value of λ the field equations of this theory would be equivalent to the Einstein field equations for a metric g_{ab}), in this limit the nonlinearities disappear and the matter fields are effectively decoupled from the graviton field, which evolves separately as a free field. Therefore the causality of the spacetime, given by η_{ab}, is no longer dynamical at high energies: when high-energy phenomena are involved, the underlying causality in the system, which is Minkowskian with no horizons whatsoever, is unveiled. Note that any phenomenon that uncovers this underlying causality would be complementary to our previous discussion about the cosmological constant in Weyl-transverse gravity, which serves to notice the existence of the background volume element.

Let us describe qualitatively what would happen in this situation in a local region around the distribution of matter undergoing gravitational collapse [40]. At some point (when its density is high enough) the collapsing matter will enter the regime in which the local causal structure is Minkowskian and there is no trace of gravity. After a scattering process that takes place in the absence of gravity and which can be idealized as a first approximation as dissipationless, the lump of matter will effectively bounce back, now expanding in time. If we keep following the expanding distribution of matter we will exit the high-energy regime in which the causal structure is not dynamical, and the usual general-relativistic picture with a gravitational field g_{ab} will be restored. This dynamical causal structure is emergent and it will therefore adapt itself to the distribution of matter in spacetime, with the corresponding light cones pointing outwards. This leads to the global geometry that we want to discuss in detail in this chapter. One of the nontrivial results of our discussion in this chapter is that there indeed exists a specific geometry that captures this qualitative picture, representing the transition from a black-hole geometry to a white-hole geometry in a short characteristic time scale (when compared to the estimated lifetime of black holes).

This motivation in terms of a specific conceptualization of the behavior of the gravitational interaction at high energies is by no means necessary, though it represents a compelling possibility. It is possible that ultraviolet completions of different nature than the ones considered in this thesis could permit this transition. Once one has constructed an effective description, one could work within this self-consistent framework without loss of generality; we shall keep our discussion within this restricted setting. From this perspective, geometries describing the black- to white-hole transition would represent the bounce of the matter distribution when reaching Planckian densities, originating a shock wave (described by some effective matter content) that propagates outwards, modifying the near-horizon Schwarzschild geometry. Our approach is heuristic, trying to elucidate the form of the effective geometries describing the black- to white-hole transition and the kind of effects that have the potential of being eventually measurable. In a genuine bottom-up approach, we shall focus in the self-consistency of the picture and the implications that derive from it.

4.2 The standard picture in general relativity

Given the unusual nature of our proposal, it is convenient to present a brief review of the standard picture of black-hole formation and evolution in both classical and quantum gravity, in order to highlight the main differences. In this section we present the purely classical considerations, leaving for the next one the inclusion of ultraviolet effects.

4.2.1 Event horizons

We shall concentrate on the simplest models of gravitational collapse, which are spherically symmetric. Eventually we will consider the effect of introducing small, non-spherical perturbations in this simple picture. As is well known, in the spherically symmetric case and for a star with mass M whose radius is larger than its Schwarzschild radius,

$$r_{\rm s} = \frac{2GM}{c^2}, \tag{4.1}$$

the external metric is unique and is given precisely by the Schwarzschild solution. The Schwarzschild metric has to be glued with the internal metric of the matter distribution along the surface of the star. A stable stellar structure has to be in hydrostatic equilibrium described by means of the Tolman-Oppenheimer-Volkoff equation [169]. On the basis of this equation and known equations of state for the matter inside the star, one concludes that there exists an upper bound to the mass of stable configurations [169, 170, 171]. Accordingly, any star in equilibrium will become unstable if accreting more than a certain amount of matter.

In those extreme situations, the strength of the gravitational interaction surpasses that of any other possible known force in the system, provoking an indefinite contraction of the stellar structure. The resulting trajectory of the surface of the collapsing star can be described by means of a function $R(\tau)$, where τ is the proper time of an observer attached to the surface. If the regime in which the strength of the gravitational interaction surpasses that of any other relevant force in the system is reached, this surface will inevitably cross the Schwarzschild radius in finite time. In terms of equations, for a given initial stellar radius at $\tau = \tau_{\rm i}$ there always exists a finite proper time interval $\Delta\tau$ so that

$$R(\tau_{\rm i} + \Delta\tau) = r_{\rm s}. \tag{4.2}$$

A prime example of this behavior is the Oppenheimer-Snyder model of gravitational collapse [172], in which the perfect fluid representing the matter content is pressureless, so that gravity is strictly the only force present in the model. The Oppenheimer-Snyder model is thus considered to provide a reliable description of the late stages of the gravitational collapse process for massive enough distributions of matter with small asphericities.

Upon crossing, a trapping horizon is formed in the position that corresponds to the Schwarzschild radius. While we leave to the next section the rigorous definition of such a notion, for the moment it is enough to have in mind the intuitive picture that this trapping horizon *locally* forbids that signals originated at the surface of the star reach external observers. This can be seen explicitly if we take the black-hole patch of the Kruskal manifold and use suitable coordinates, for instance Painlevé-Gullstrand coordinates (see [173] and references therein). It is important to keep in mind that this observation by itself does not imply that lumps of matter will never be able to cross outwards the radial position at which the trapping horizon was first formed. The overall dynamical evolution of spacetime around the outgoing distribution of matter should be taken into account in order to consistently conclude so. As we will see, restrictions to the geometry of spacetime in the form of energy conditions are useful to tackle this issue. We leave this discussion for the following section, assuming for the moment the standard picture that horizons are inviolable in classical general relativity.

The resulting object, a so-called black hole, would then possess as defining characteristic an event horizon. They would be absolutely inert objects, the ultimate end point of gravitational collapse, the dead state of stellar physics. No matter content can ever escape from the inside, as the region behind the Schwarzschild radius is causally disconnected from the external observers. Black holes are characterized by just three numbers, mass M, electric charge Q and angular momentum J, for generic initial conditions (even if not spherically symmetric) for the matter that went to form the black hole; see [174] for a thorough technical discussion. This simplicity is an appealing feature for a large number of theorists. In our idealized setting we are demanding spherical symmetry ($J = 0$) and electrical neutrality ($Q = 0$), leaving us with just one relevant parameter.

4.2.2 Singularities

According to classical general relativity, the fate of matter after crossing the trapping horizon is ominous. While the horizon is characterized by a pronounced deformation of the light cones so that (locally) the light that is emitted at the horizon cannot escape, light cones just behind the horizon are deformed even more drastically. The resulting deformation ultimately implies that every light-like trajectory starting inside the horizon must (locally) go inwards, i.e., with the value of the radial coordinate decreasing, until “hitting” the singularity.

Let us be more precise on these notions, which not only will permit us to make sharp statements, but also understand the generality of this picture. In principle one could think (and this was so historically [175]) that the occurrence of the singularity sketched above may be an artifact of the high degree of symmetry of the solution being used. Quite the opposite, these features are completely generic, as was first understood by means of the so-called Penrose singularity theorem. This was the first of a sequence of results, usually known as Hawking-Penrose theorems (see [176] for a recent account of the history of the subject and later developments up to date). These theorems are formulated on spacetime manifolds that possess a non-compact Cauchy hypersurface, i.e., the topology of the manifold is $\mathbb{R} \times \Sigma$ with Σ being a non-compact spacelike three-dimensional manifold. Roughly speaking, this means that the geometry does not incorporate new regions that require additional data for their description, and that there are well-defined notions of future and past.

The first notion we need concerns the "surface" of the black hole, that is, the two-dimensional manifold defined by the equation $r = r_{\rm s}$ at a given moment of time (being the Schwarzschild metric independent of time, this surface is stationary under its flow). We shall use the conventions and notation of [176]. Let us imagine that we are emitting a spherical wavefront of light rays from a given position $r > r_{\rm s}$. Then light rays pointing outwards will move outwards, meaning that the radius of the corresponding sphere of light will grow in time, while light rays pointing inwards will move inwards, with decreasing radius in time. This is no longer true when we enter the Schwarzschild radius: for $r < r_{\rm s}$ both spheres of light move inwards. These surfaces are thus a particular case of the general concept of a closed future-trapped surface: a closed surface whose *future-directed* null geodesics do not flow outwards. A closed past-trapped surface would correspond to the situation in which *past-directed* null geodesics do not flow outwards. In the spherically symmetric case one can consider these surfaces as spherical, filling up the black hole as an onion. In this case the $r = r_{\rm s}$ surface marks the boundary between this trapping behavior and the normal behavior for $r > r_{\rm s}$, and represents the so-called trapping horizon, or marginally future-trapped surface.[1]

The second notion is that of a singularity. This is a slippery concept as, strictly speaking, singularities do not belong to spacetime. We can however characterize the singular behavior by means of particles traveling to or from the singularity or, in geometric terms, of geodesics which finish or start at the singularity. In general, curves are maps from a subset of the real line $\mathbb{R}$ onto the spacetime manifold. When all the geodesics in a manifold are defined for the entire real line $\mathbb{R}$ when parametrized in terms of their affine parameter, the manifold is said to be geodesically complete, being incomplete in the opposite. The usual definition is that a spacetime manifold is singular if and only if it is geodesically incomplete.[2] From a physical perspective this would imply the existence of particles or

[1] In the Schwarzschild solution the event and trapping horizons coincide. This is not generally true for dynamical spacetimes, being the simplest example a small, non-spherical perturbation of the Schwarzschild solution [37].

[2] Most definitions also include the condition of being non-extendable [37], but this technical point is not really relevant for our purposes.

observers that disappear (or materialize) abruptly, which clearly corresponds to an abhorrent behavior. It is important to keep in mind that, for a manifold to be geodesically incomplete, it is enough that there exists just one incomplete geodesic.

The last ingredient is a condition on the curvature of the manifold. This crucial geometrical condition for the theorem is motivated by its translation, through the use of the Einstein field equations, into a condition upon the matter content. The so-called null convergence condition holds by definition if, for every null vector field u^a in the manifold,

$$\mathscr{R}_{ab} u^a u^b \geq 0, \tag{4.3}$$

where $\mathscr{R}_{ab}$ is the Ricci curvature tensor. If we apply the Einstein field equations,

$$G_{ab} = \mathscr{R}_{ab} - \frac{1}{2}\mathscr{R} g_{ab} = \kappa T_{ab}, \tag{4.4}$$

then Eq. (4.3) implies the null energy condition for the matter stress-energy tensor T_{ab}:

$$T_{ab} u^a u^b \geq 0. \tag{4.5}$$

With these ingredients, the Penrose singularity theorem is formulated as follows [177]: given a spacetime containing a non-compact Cauchy hypersurface Σ and a closed future-trapped surface, the convergence condition (4.3) being valid for all the null vectors u^a implies that there exist null geodesics that are incomplete in the future.

In all its generality, this theorem does not provide us with any clue about the behavior of the spacetime near the singularity. To put an example, it does not assert that there is some curvature invariant (a scalar made up from the metric and its derivatives only) which blows up in the singularity, or that an entire set of observers in a given region (e.g., the interior of the future-trapped surface) will certainly experience the singular behavior in the future. Developing new techniques is necessary in order to tackle this thorny problem and understand in fine detail the physical properties of singularities. Among the attempts of doing so in the framework of the classical theory, the so-called BKL conjecture [178, 179], contemporaneous to the singularity theorem stated above, occupies a prominent place. While in ordinary situations gravity is the weakest of all the forces in nature, singularities are the place in which gravity becomes the most prominent actor. Following this idea, in this conjecture it is assumed that matter fields do not play any essential role near the singularity. Moreover, it is postulated that the temporal derivatives of the metric field are the dominant ones, so that spatial derivatives can be ignored near the singularity (which is assumed to be spacelike). Every spatial point of spacetime will then become disconnected from the others, and will evolve following ordinary differential equations that present chaotic behavior. Numerical studies have shown evidence in favor of this behavior; see [180, 181, 182, 183] for modern discussions.

4.3 Ultraviolet effects beyond general relativity

In the previous section we have described the structure of black holes in classical general relativity. Let us now describe how the inclusion of ultraviolet effects could modify the nature of the distinctive elements of a black hole.

4.3.1 Trapping horizons

When trying to build coherent scenarios in which the collapse of matter is affected by the very quantum nature of matter, the classical picture significantly changes. Robust semiclassical calculations tell us that black holes cannot be absolutely stationary, as they evaporate through the emission of Hawking quanta [167, 184]. This is the last piece composing the so-called black-hole thermodynamics: a black hole of mass M has a temperature

$$T_{\rm H} = \frac{\hbar c^3}{8\pi G M k_{\rm B}}. \tag{4.6}$$

Given this result, it is assumed that the classical spacetime representing the collapse of a star to form a black hole should be modified so that the event horizon, instead of setting at a stationary position (once the absorption of surrounding matter and its stabilization has ended up), would shrink up to eventually disappear in a final explosion [185]. It is also clear that the quantum effects responsible for the evaporation of the horizon would entail even more drastic modifications in the region surrounding the singularity. These realizations inevitably trigger the questions: Are these evaporating black holes really black holes in the sense of having an event horizon? Is there some information loss in a complete evaporation process? These questions have been a matter of controversy, and a strong driving force for theoretical development, since the publication of Hawking's paper [38]. However, nowadays even Hawking concedes [186] that the most reasonable solution is that no event horizon would ever form, but only a structure which is locally similar but that lives for a finite, albeit extremely long, amount of time; that is, a trapping horizon as defined in Sec. 4.2.2. Under this view black holes, understood as causally disconnected regions of spacetime, would not strictly exist in nature. However, owing to the similarity of these quantum-corrected objects with classical black holes, they usually keep their name. Here we will generically call them regularly evaporating black holes (REBHs) to distinguish them from their classical cousins.

4.3.2 Preventing singularities

As happens with the event horizon the introduction of quantum mechanics, or any other kind of ultraviolet modifications, is likely to bring important changes to the classical picture concerning the singular behavior of GR. The first indication of these changes is that all kinds of energy conditions, that play a central role in the singularity theorems, are violated by quantum effects [187, 188]. Indeed, from a qualitative perspective it is widely expected that any consistent ultraviolet completion of general relativity will regularize the singular classical behavior. One of the ways to do so is the occurrence of repulsive forces preventing the gravitational collapse process to proceed indefinitely.[3] Given their repulsive character,

[3] It would not be the first known example of this phenomenon: for example, the electron degeneracy pressure is a genuinely quantum effect that stabilizes white dwarfs [189, 190].

these effective forces will generally violate the energy conditions and, eventually, would unlock new options for the fate of gravitational collapse.

These qualitative expectations have been embedded in mathematical frameworks of different nature, essentially since the first explorations in semiclassical gravity [191], and most frequently in the field of cosmology [192]. One of the most popular implementations, probably because of its fundamental flavor, is the one that results from the application of the loop quantum gravity techniques (see [193] for an introduction to the subject). Loop quantum cosmology (see, e.g., [194]) is the result of applying non-perturbative quantization techniques borrowed from the general theory and applied to some highly symmetric cosmological spacetimes. Although the results of this procedure have to be taken with a grain of salt [195, 196], these models are regarded as valuable tools in understanding the implications of the wider quantization program of the canonical structure of general relativity. One of the robust results of this approximation is that near the cosmological singularities, there appear effective forces with a net repulsive effect. These forces are strong enough to overcome the fatal attraction of gravity which would otherwise engender a singularity, provoking the *bounce* of the matter distribution and connecting a classical contracting cosmology with a classical expanding cosmology [197]; the Big Bang event is identified with the moment of the bounce, so that the expanding branch corresponds to our present universe. The bounce generally takes place when a critical matter density $\rho_{\rm c}$, of the order of the Planck density, is reached:

$$\rho_{\rm c} \sim \rho_{\rm P} = \frac{m_{\rm P}}{\ell_{\rm P}^3} = \frac{c^5}{\hbar G^2} \simeq 5 \times 10^{96} \text{ kg m}^{-3}. \tag{4.7}$$

Following this result, quantum effects would be triggered when the radius of the stellar structure is several orders of magnitude greater than the Planck length; for a star of mass M, its minimum radius R_0 will be approximately given by

$$R_0 \sim \ell_{\rm P} \left(\frac{M}{m_{\rm P}}\right)^{1/3}. \tag{4.8}$$

Although this may seem surprising at first, given that quantum gravity effects are usually attached to the Planck length, this very association is only reasonable in the absence of matter, that is, in the pure gravity case. Let us stress that the order of magnitude (4.8) is the natural one on the basis of the Einstein field equations and the expectation that quantum gravity effects should appear when the curvature is of the order of $\ell_{\rm P}^{-2}$ *in the presence of matter*. Actually, for dust matter content of Planck density the trace of the Einstein field equations (4.4) leads to the very same order the magnitude:

$$\mathscr{R} = g^{ab}\mathscr{R}_{ab} = -\kappa g^{ab} T_{ab} = \frac{8\pi G \rho_{\rm P}}{c^2} = \frac{8\pi}{\ell_{\rm P}^2}. \tag{4.9}$$

Thus in the presence of matter the association between Planckian densities and ultraviolet effects is completely straightforward, independently of the specific ultraviolet completion of the gravitational interaction one pursues.

In loop quantum cosmology, the bouncing behavior can be shown in some cases to be successfully captured by an effective Friedmann equation which, in the flat ($k = 0$) case, takes the form [198]

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho\left(1 - \frac{\rho}{\rho_c}\right). \tag{4.10}$$

In this equation it is easy to see that reaching the critical density corresponds to a turning point, $\dot{a} = 0$. The second term in the right-hand side of this equation corresponds to an effective stress-energy tensor that strongly modifies the classical behavior. It will be instructive to have in mind this feature in our later discussion of gravitational collapse. The occurrence of effective repulsive forces will probably be a robust ultraviolet effect, independently of the specific quantization program or ultraviolet completion one choses (see, e.g., [40, 191, 199, 200] for different suggestions), leading eventually to a resolution of the singular behavior of the classical theory. However, while the cosmological situation is quite well understood nowadays, how these considerations would affect the issue of gravitational collapse is a more subtle issue, as we discuss in the following.

4.4 Phenomenological considerations

As explained in the previous section, there is a clear conceptual picture about the kind of modifications of the classical elements characterizing a black hole that could appear as the result of the modification of the gravitational interaction at high energies. Nonetheless, it is seldom remarked that it is hardly possible to combine all these features into a unified theoretical model, while maintaining reasonable prospects for the independent experimental corroboration of its different parts. In this section we highlight this tension, which serves as additional motivation for the discussion of our proposal.

4.4.1 Horizon predominance

The necessity of finding an ultraviolet completion of general relativity has been in the air at least since the first successes of quantum field theory, in particular when applied to electrodynamics. The generalizations of these field-theoretical developments to general relativity were plagued with difficulties, thus forcing to use partial, and therefore necessarily incomplete, approaches. These have been used since the sixties to get insights into the ultraviolet behavior of gravitational collapse processes. Before Hawking's groundbreaking developments, there already were some interesting attempts to see what could result from the incorporation of quantum-mechanical effects to black-hole spacetimes. As early as 1966, Sakharov [201] and Gliner [202] suggested that at the high densities reached close to the singularity formation, the effective matter content might develop a vacuum-like equation of state $\rho = -p$, providing a repulsive force. Soon after, Bardeen showed that it is possible to construct black-hole geometries (here meaning having an event horizon) satisfying the null energy condition but having no singularities ([203], see also [204]). Singularity avoidance (and hence a way out of the singularity theorems) is permitted due the non-existence of an open Cauchy surface: the spacetime develops a topology change from open hypersurfaces to closed ones [205]. Matter crossing the black-hole horizon and falling towards the apparent singularity would end up reappearing from a white-hole horizon into a *different universe*, that is, not in the same asymptotic region. Bardeen's spacetime and others with similar structure (e.g., [206, 207, 208, 209, 210, 211, 212, 213]) are examples of regularizations of the black hole singularity that do not affect the event horizon.

Hawking's idea that event horizons should evaporate, leading eventually to some fundamental loss of information, introduced a new ingredient into the singularity problem. Many different evaporation scenarios have been put forward since then to accommodate solutions to the information and singularity problems. These proposals have generated a large amount of work; see for instance the sample of recent papers [214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232]. What we want to emphasize here is that, at least qualitatively, the resulting effective geometries advocated by most of these scenarios (what we have called REBHs) are essentially equivalent. Similar diagrams describing these geometries can be found for example in [233, 234, 235, 236, 211, 224, 237, 238, 239]. Aside from their fine theoretical details, all of them contemplate REBHs as essentially hollow and almost stationary regions of spacetime; all of them share the substitution of event horizons by *extremely slowly* evaporating trapping horizons as seen by external observers, even if the specific time scales could vary between different models. In this sense these proposals may be regarded as "conservative".

Thus, the prevalent view is that REBHs indeed form in astrophysical scenarios and, ignoring their gravitational interaction with the surrounding matter, they would remain almost inert for very many Hubble times except for a tiny evaporative effect that would eventually make them disappear. The evaporation rate of stellar-mass objects would be so slow, 10^{67} years to halve the size (the so-called Page time) of a Solar mass object, that for all practical purposes they could be considered stationary. REBHs have a ridiculously long lifetime, in whatever measure: the estimated age of the universe is of the order of 10^{10} years. Whereas the distinction between a classical BH and a REBH is of fundamental interest on purely theoretical grounds, it is almost certainly irrelevant for all astrophysical purposes, and arguably for any meaningful operational perspective [240]. This becomes even worse when considering that astrophysical black holes are actually on average growing and not yet evaporating because their Hawking temperature is smaller than the approximately 3K of the cosmic microwave background [241]. This situation could be partially alleviated if primordial black holes were generated in our universe, as it was already proposed by Hawking [185]; see [242] for some recent constraints on their existence. In the present discussion we shall focus on contemporary gravitational collapse processes, which are certainly the most interesting ones from an experimental perspective.

Given the irrelevance of Hawking radiation for astrophysical purposes, any experimental test of the precise way in which the evaporation proceeds, whose result is available to observers outside black holes, would be almost certainly beyond the reach of humankind. Most importantly, the very nature of any kind of process taking place in the interior of the trapping horizon, and in particular the fate of the matter inside it and the ultraviolet regularization of the singularity, will be inevitably hidden for us unless something radical changes this picture. While observers that fall into the black hole and cross the trapping horizon will be able to catch a glimpse of what is going on inside, they will not be able to communicate their experience to the exterior, enforcing the ignorance of external observers.[4]

For the standard REBH perspective the only possible, albeit partial, experimental corroboration of the overall scenario would be the detection of Hawking radiation.[5] In the real world the systems hosting black holes are so complex that there exist multitude of forms of radiation that would eventually frustrate any attempt to measure such a tiny effect.[6] On the other hand, given the ridiculously long times associated with the evaporation process, any event happening inside the trapping horizons formed in gravitational collapse processes is certainly irrelevant for practical purposes. Any attempt to understand the physics behind the horizon, for instance on the lines sketched in Sec. 4.3.2, would be unquestionably pointless: while the fate of matter behind the trapping horizon and the horizon itself could prove an interesting intellectual exercise, it would be impossible to find any experimental corroboration of these developments. The black hole trapping horizon will effectively act as an event horizon, forcing our ignorance about what is behind. This observation is not a matter of improving the sensitivity of experiments, but the issue is a completely different (and for us, quite unpleasant) one.

One could argue that, even accepting that the gravitational collapse process itself could be useless to boost our understanding of the physics near singularities, the Big Bang singularity could be used as a substitute from which to obtain this knowledge. The first observation that comes to mind concerning this assertion is the issue of repeatability: while we expect that numerous processes of gravitational collapse are taking place now around us in the universe, and will be quite surely taking place in the future, in the currently accepted model of the universe there is only one Big Bang event. Also the Big Bang singularity lies in the past so that, while we can hopefully describe its properties in a simple way, it is questionable to what extent we are able to perform experiments in the usual sense of the word, or if the most we can hope to do is constructing some sort of cosmological "archeology". Leaving aside these issues, which are indeed ubiquitous to the field of cosmology (see [249] for instance), there are good physical reasons to believe that the singularities associated with black holes should be different from the initial cosmological singularity. These arguments are based on the second law of thermodynamics and have been repeatedly exposed by Penrose [175]. While one may think that these singularities correspond from a mechanical point of view as the time-reversal of each other, thermodynamical considerations break this apparent time reversal. On the one hand, if the second law of thermodynamics is to be applicable to the universe, the behavior near

[4] We encourage the reader to note the amusing analogy with the existence of life after death.

[5] Even this assertion is debatable, as stellar objects hovering near its Schwarzschild radius could produce a radiation signal very close to the predicted Hawking radiation [243, 244, 245].

[6] A great deal of effort has been put on the detection of the analogue of the Hawking effect in the analogue gravity framework [14, 246, 247]. It is, however, unclear to what extent the analogy is deep enough to consider the measurement of these effects on fluids a detection of the "genuine" gravitational Hawking radiation; some arguments in favor can be read in [248].

the initial singularity is to be associated with a low amount of entropy. On the other hand, the process of gravitational collapse to black holes is generally expected to lead to high entropies. In particular, this should indeed be the case in the standard scenario, in order to fulfill the thermodynamic description of black holes and match the Bekenstein-Hawking entropy formula

$$S = \frac{k_{\mathrm{B}} c^3 A}{4 G \hbar}, \tag{4.11}$$

where A is the area of the horizon. This points to a deep physical distinction between these situations that supersedes the apparent time reversal symmetry.

To end this section let us say a few words about other popular, and somewhat more exotic, approaches to the information and singularity problems. Our comments here will be restricted to their relation with the lifetime problem which is central for our discussion. The complementarity approach of Susskind [250] would share the REBH geometric description but maintaining that it would be only relevant for observers actually crossing the horizon. A complete experimental proof of the overall complementarity scenario would therefore face the same practical problems than more standard REBHs scenarios. The fuzzball proposal of Mathur [251, 252] takes the view that the evaporating horizon marks a real border for the spacetime. Information is accumulated at the gravitational radius, being retrieved only after enormous amounts of time. The same situation occurs in the condensed state scenario of Dvali [253, 254]. Let us recall that the Page time (half-mass evaporation, the assumed time of retrieval of the information in these scenarios) for an astrophysical black hole is as extremely long as the time scale associated with Hawking evaporation. The Giddings remnant scenario [255, 256] adds to the REBHs the presence of massive remnants. Only the last stages of the REBHs picture are expected to be modified so that this scenario also shares the lifetime problem.

Let us note for completeness that there is a recent construction by Haggard and Rovelli [257] that shares some geometric properties with our proposal, as it also describes a black- to white-hole transition with only one asymptotic region. However, it is completely different in essence: by construction, the duration of the overall transition as seen from external observers is extremely long. This is due to the implementation of a specific assumption about the behavior of ultraviolet effects that contrasts with the picture we shall present in Sec. 4.5. This scenario is therefore indistinguishable for an extremely long time from any REBH model, being then not better regarding the lifetime problem. Independently of this, it presents serious internal consistency problems when applied to physically reasonable situations [258].

4.4.2 Preponderance of the singularity regularization

The possibility of measuring the ultraviolet effects that are responsible for singularity avoidance in reasonable time scales would imply that Hawking radiation would not be able to carry away from the object a significant amount of energy in the meanwhile. Indeed, to make these measurements possible the horizon barrier should be broken in some way so that some signals from the inside may reach an external observer. As the evaporation process is extremely slow, the net evaporative effect would be then negligible. Even if we wait the entire age of the universe (specifically, the Hubble time $t_{\rm H}$), the energy loss of a stellar-mass black hole would be of the order of

$$10^{-4}\,\frac{\hbar c^6}{G^2}\frac{t_{\rm H}}{M_\odot^2} \sim 10^{-11}\ {\rm J}. \tag{4.12}$$

This represents a 10^{-49} part of its original rest energy, which is absurdly negligible. Again, from an operational perspective only part of these theoretical developments associated with gravitational collapse could be relevant. We end up with a dilemma: if the physics of horizons as long-lived thermodynamical entities is to be realized in nature, there is very little hope that we will be able to observe any trace of the physics associated with the corresponding (regulated) singularities; while, if the physics of these spacetime singularities is to be observable in reasonable time scales, any dissipation through the emission of Hawking radiation would be physically irrelevant in energetic terms.

An additional tension, of a different kind but related to the previous one, appears when ultraviolet effects inside the collapsing distribution of matter are taken into account. As a first approximation let us consider the homogeneous and isotropic situation inside the star, so that we can use a patch of a cosmological solution to describe the internal geometry. In the pressureless case the surface of the star (as well as any point in its interior) will follow a geodesic in the external metric, i.e., the Schwarzschild metric. Any nonzero pressure will lead to deviations from this behavior. As an extreme example one could imagine a strong enough pressure to halt the collapse before the event horizon is formed; the radius of the star will then reach a fixed value. While this is what happens for structures below the Tolman-Oppenheimer-Volkoff limit of neutron stars [169, 170, 171], for larger masses there is no known force that can compete with the gravitational contraction [259]. It seems unavoidable that, in these situations, the matter distribution crosses its Schwarzschild radius, so that we can even take the idealized Oppenheimer-Snyder model, neglecting any known form of pressure with respect to the gravitational force. Deviations from general relativity would not appear until high enough densities are reached, and certainly *not* at the moment when the star generates its trapping surface (at least if the collapse proceeds in a free-fall manner; departures from this behavior could lead to strong deviations even before of horizon crossing [260]). As we discussed before, if singularities are to be avoided in a suitable ultraviolet completion, a universal effective pressure with a net repulsive effect should appear at some point. This pressure would prevent the formation of singularities, leading to bouncing solutions of similar nature as the ones that were previously described in a cosmological framework. How is this compatible with the geometry outside the star?

The bouncing process of the matter distribution implies that the radius of the star should grow in proper time for an observer in the surface. But this is in contradiction with the external metric: there are no timelike geodesics of non-decreasing radius inside the Schwarzschild radius. In order to avoid a causality violation we should accept a modification of the external metric around the bouncing star. Given the spherical symmetry we are assuming, these geometry deviations should take the form of an effective matter content violating energy conditions. In other words, these general arguments point to the necessary existence of modifications of the geometry even outside the star. These modifications are generally expected to be localized in a small region around the minimum radius of the star R_0 in which the curvature invariants associated with the classical geometry are still large enough.

The situation is, in our opinion, more subtle. Once we have accepted that a bounce of the star should occur near the singularity, both geometric pictures inside and outside the radius of the star seem to be hardly reconcilable. More specifically, gluing these geometries along the wordline (for fixed angular variables) of the surface of the star no longer appears as a viable option, as these present what we can regard as competitive trends. While the external geometry will try to stop the matter distribution to escape, the internal geometry of matter will try to overcome the overwhelming horizon influence. It seems that one should rather face the problem of how these effects compete in order to have a glance to the final outcome. In the most general case the resulting geometry will present a transition layer between these two regimes. How to describe the properties of this layer seems to be currently unknown.

In the standard view of REBHs, the internal geometry is completely dominated by the external geometry, so that the transition layer affects the behavior of the overall internal metric. This would correspond, e.g., to the effective stabilization of the bouncing star at some asymptotic radius *inside* the horizon. Such a situation corresponds to one of the extremes of the entire family of possible interpolating geometries. The situation we want to analyze here is the complementary extreme, with the internal geometry unchanged and the external metric modified in order to guarantee a smooth matching. As we will discuss, in this case the transient may be analogous to a shock wave phenomenon as those occurring in normal fluids.

From a theoretical perspective, the fine knowledge of the resulting geometry will probably have to wait until our understanding of the ultraviolet properties of the gravitational interaction is deep and firm enough to tackle these complex questions. Although ultraviolet effects are expected to be triggered only when curvatures are high enough (i.e., Planckian), we have virtually no knowledge of how these regions of high-curvature, that violate energy conditions, evolve once they are generated, how they propagate and decay in time. While a first-principles study of this issue is highly interesting, we can always revert the logic and try a phenomenological approach. We can study the possible effective geometries and try to extract physical and observational consequences, information that could be eventually used in order to progress in our theoretical understanding.

4.5 Effective bounces, black- to white-hole transitions and shock waves

In plain terms, what we want to analyze is what would happen if the spherically-symmetric collapsing matter, upon reaching Planck density, slowed down and bounced back. In geometric terms we expect the description near the matter crossing the horizon outwards to correspond with the time-reversed geometry to that of a black hole, namely the one associated with a white hole. Birkhoff's theorem is certainly a uniqueness result for the vacuum geometry outside a spherical distribution of matter with mass parameter M and radius greater than its Schwarzschild radius. However, for vacuum geometries inside the Schwarzschild radius this is not true. Birkhoff's theorem asserts that any vacuum patch must be locally isometric to a patch of the maximally extended Kruskal manifold [37]. The vacuum geometry outside a spherical distribution of matter but still inside the Schwarzschild radius could therefore either correspond to the black-hole or the white-hole patch. From this perspective, the bounce can be represented by a transition between these two patches which will necessarily contain features that go beyond general relativity.

It is an essential feature that the flat asymptotic regions corresponding to these geometries are one and the same. The bounce is a physical event taking place entirely in our universe; it does not describe the escape of matter to remote regions that are causally disconnected from the static observers standing outside the stellar structure (as discussed in Sec. 4.4.1, proposals of this kind fall again within the REBH paradigm). As we will see later, the most general metric describing this bounce contains a number of unknown parameters, so that its fine details certainly depend on the underlying ultraviolet completion of general relativity. However, it also displays some robust characteristics that clearly distinguishes it from REBH proposals even from a physical perspective. Let us proceed in a constructive way, and build first a simple toy-model geometry of a time-symmetric bouncing regularization of a collapsing star, that however encapsulates the most relevant features of the process we want to describe. While this geometry will present some singular properties, it can be considered as the distributional limit of well-behaved geometries, thus providing indeed a very good analytical approximation.

We shall describe the bouncing geometry from an explicitly time-symmetric point of view. As stated above, for simplicity we are concentrating on the spherically symmetric collapse, with the collapsing matter being characterized as a pressureless dust cloud. Also we will ignore dissipation effects in a first stance: classically, a spherically-symmetric collapse does not produce any dissipation in the form of gravitational radiation. If we took into account quantum corrections, there would indeed exist some (rather small) dissipation. In the limit of very large masses, this quantum radiation could in principle be made as small as desired. The time-symmetric geometry we are going to present is a reasonable dissipationless approximation to more realistic, dissipative situations, that we explore in Sec. 4.6.

It will prove useful to describe the metric using generalized Painlevé-Gullstrand coordinates [261]. These coordinates are adapted to observers attached to the stellar body. We shall write the metric as

$$ds^2 = -A^2(t,r)dt^2 + \frac{1}{B^2(t,r)}[dr - v(t,r)dt]^2 + r^2 d\Omega_2^2, \tag{4.13}$$

with $d\Omega_2^2$ being the line element of the unit 2-sphere. The three functions $A(t,r)$, $B(t,r)$ and $v(t,r)$ will be given in the following patch by patch. Let us begin considering that the collapse process starts at rest at infinity, so that the line element (4.13) reduces to the standard Painlevé-Gullstrand one, with $A^2(t,r) = c^2$ and $B^2(t,r) = 1$. When familiarized with this situation, we will move on to consider in Sec. 4.5.3 the more general case in which the collapse starts from an initial stellar radius $r_{\rm i}$.

4.5.1 Collapse from infinity and homogeneous thin-layer transition

To describe the bounce, we shall glue two geometries, one corresponding to the Oppenheimer-Snyder collapse of a homogeneous ball of dust, the other one being its time-reversal. Let us consider for the moment the metric outside the star, corresponding to the Schwarzschild solution. The velocity profile $v(t,r)$ presents a flip of sign between the black-hole and white-hole patches of the Kruskal manifold. Thus the metric we are seeking for will be characterized by a velocity profile

$$v(t,r) = v_{\rm s}(r)[1-2\theta(t)] = -c\,[1-2\theta(t)]\sqrt{\frac{r_{\rm s}}{r}}, \qquad R(t) \leq r \leq +\infty. \tag{4.14}$$

Here $v_{\rm s}(r) = c\sqrt{r_{\rm s}/r}$ is the absolute value of the standard velocity profile of the Schwarzschild solution, and the Heaviside function $\theta(t)$ is used to perform an abrupt transition between the different patches. Concerning the function $R(t)$ describing the trajectory of the surface of the star, for the moment the only condition we demand, following our previous discussion, is that it is bounded from below by $R(0) = R_0$ where $t = 0$ corresponds to the moment of the bounce in the coordinates we are using.

This gluing procedure by itself is nothing but a brute force exercise, and the result presents unpleasant features. However, we want to present a constructive procedure in which the details are progressively added on demand, eventually arriving to the complete picture in all its generality. For instance, using the Heaviside function to construct the geometry makes all the $t = 0$ hyperplane singular for $r > R_0$, in the sense that the metric is discontinuous there. Of course, this does not necessarily reflect any physical reality, as we can always introduce a regulator to make the geometry continuous (see Fig. 4.1). Let us introduce the continuous, differentiable but non-analytic function

$$f(t) = \begin{cases} \exp(-1/t) & 0 \leq t, \\ 0 & t \leq 0, \end{cases} \tag{4.15}$$

that permits us to define for $t_{\rm R} \in \mathbb{R}$ the function

$$g_{t_{\rm R}}(t) = \frac{f(1/2 + t/t_{\rm R})}{f(1/2 + t/t_{\rm R}) + f(1/2 - t/t_{\rm R})}. \tag{4.16}$$

This function represents a regulated version of $\theta(t)$ that interpolates between the values $g_{t_{\rm R}} = 0$ (for $t \leq -t_{\rm R}/2$) and $g_{t_{\rm R}} = 1$ (for $t \geq t_{\rm R}/2$). The value of $t_{\rm R}$ thus controls the duration of the interpolation. We can actually define the pointwise limit

$$\theta(t) = \lim_{t_{\rm R} \to 0} g_{t_{\rm R}}(t), \tag{4.17}$$

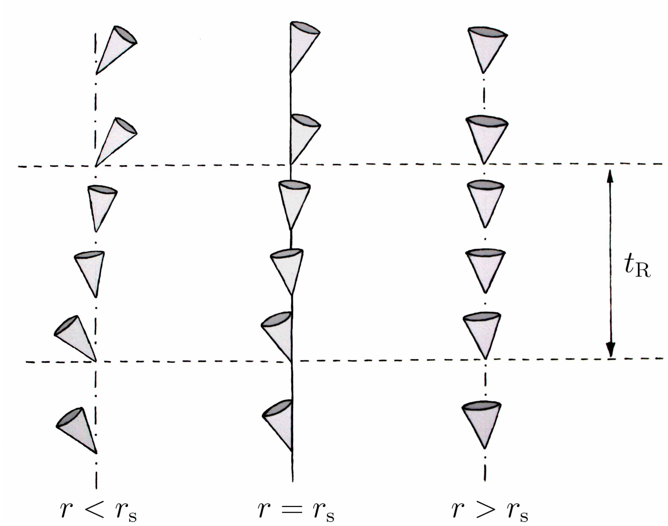

Figure 4.1: Transition from the black-hole patch to the white-hole patch by using a smooth interpolation with characteristic time scale t_{R}. The qualitative behavior of light cones for different values of the radial coordinate is shown.

thus fixing the convention $\theta(0) = 1/2$.

With respect to the field equations, the strict limit $t_{\mathrm{R}} \to 0$ may not be well defined as a solution even in the distributional sense, as its curvature may contain terms that are quadratic in the Dirac delta function, which are ill-defined (in general, the product of two distributions cannot be defined [262]). Let us emphasize again that we always keep in the back of our minds that the relevant physical situation corresponds to t_{R} small, but nonzero. For $t_{\mathrm{R}} > 0$ we can safely insert the metric in the Einstein field equations (4.4) to obtain (using, e.g., Eqs. 3.2 in [261]) the following nonzero components of the effective stress-energy tensor supporting this geometry:

$$\begin{aligned} \mathscr{T}^1{}_0 &= -\frac{c}{2\pi}(1 - 2g_{t_{\mathrm{R}}})\dot{g}_{t_{\mathrm{R}}}\frac{M}{r^2}, \\ \mathscr{T}^1{}_1 &= -\frac{c}{2\pi}\dot{g}_{t_{\mathrm{R}}}\frac{M}{r^2}\left(\frac{r_{\mathrm{s}}}{r}\right)^{-1/2}, \\ \mathscr{T}^2{}_2 &= -\frac{c}{8\pi}\dot{g}_{t_{\mathrm{R}}}\frac{M}{r^2}\left(\frac{r_{\mathrm{s}}}{r}\right)^{-1/2}. \end{aligned} \tag{4.18}$$

We use the notation $\mathscr{T}_{ab}$ in order to make a clear distinction between this effective source originated in the bounce, and the stress-energy tensor T_{ab} of the perfect fluid that forms the stellar structure. It is interesting to note that the effective stress-energy tensor (4.18) only involves first derivatives of the interpolating function $g_{t_{\mathrm{R}}}$, so that the distributional limit is indeed well-defined, corresponding to a $\delta-$function source located at $t = 0$; for $t_{\mathrm{R}} > 0$ this source presents certain width. The effective source (4.18) violates the null energy condition (4.5) in all its radial extension; let us consider for example the hyperplane $t = 0$ and the null vector u^a on it, with components $u^0 = u^1 = 1$, $u^2 = u^3 = 0$. Then

$$\mathscr{T}_{ab}u^a u^b = -\frac{c}{\pi}\frac{M}{r^2}\left(\frac{r_{\rm s}}{r}\right)^{-1/2} \dot{g}_{t_{\rm R}}\big|_{t=0} < 0. \tag{4.19}$$

Taking into account the metric inside the collapsing star, the complete geometry is given by:

$$A^2(t,r) = c^2 B^2(t,r) = c^2, \tag{4.20}$$

and

$$v(t,r) = -c\,[1 - 2g_{t_{\rm R}}(t)] \times \begin{cases} \sqrt{\dfrac{r_{\rm s}}{r}} & R(t) \leq r, \\ \dfrac{r}{R(t)}\sqrt{\dfrac{r_{\rm s}}{R(t)}} & r \leq R(t). \end{cases} \tag{4.21}$$

The function $R(t)$ is constructed from the calculation in general relativity of the Oppenheimer-Snyder collapse starting from infinity [261]:

$$R(t) = \left(\frac{9r_{\rm s}}{4}\right)^{1/3} |c\,t|^{2/3}. \tag{4.22}$$

The point $t = 0$ in this trajectory would correspond to the classical singularity. For $t < 0$ the trajectory goes inwards, while for $t > 0$ it goes outwards. To make the geometry smooth, we shall consider that the function $R(t)$ is modified by means of a smooth interpolation in the interval $t \in [-t_{\rm R}/2, t_{\rm R}/2]$ so that the classical singularity is never formed; as said before, $R(t)$ should be bounded from below by $R(0) = R_0 > 0$. The order of magnitude of R_0 is fixed through model-independent considerations as given by Eq. (4.8).

To sum up, the two main features of the geometry presented in this section are the regularization of the singularity by the bounce of the collapsing matter at $R(0) = R_0$ and the existence of a thin-layer transition region between the black-hole and white-hole patches that violate the null convergence condition (4.3). This thin layer is homogeneous in the sense that the interpolation time $t_{\rm R}$ is independent of the radial coordinate; this dependence will be included in Sec. 4.5.3. Let us discuss for the moment some additional relevant properties of this region of spacetime.

4.5.2 Non-perturbative ultraviolet effects

The curvature scalar corresponding to the effective source (4.18) is given by

$$\mathscr{R} = 6\frac{\dot{g}_{t_{\rm R}}}{c}\sqrt{\frac{r_{\rm s}}{r^3}}. \tag{4.23}$$

This expression permits to make an order of magnitude estimate for $t_{\rm R}$, by the following argument. Let us assume that, when the stellar radius reaches its minimum value, R_0, its density is of the order of the Planck density, corresponding to Planckian curvatures by virtue of Eq. (4.9). These Planckian curvatures generate non-perturbative contributions that prevent the structure from collapsing, triggering the bounce process [recall the form of the modified Friedmann equation (4.10)]. For continuity reasons, it is reasonable to expect that the modifications of the classical geometry outside, but very close to the stellar radius, are at this stage also Planckian. This is guaranteed if, as using Eqs. (4.8), (4.23) and the estimation $\dot{g}_{t_{\rm R}} \sim 1/t_{\rm R}$ reveals,

$$t_{\rm R} \sim \frac{\ell_{\rm P}}{c} = t_{\rm P}, \tag{4.24}$$

with t_P the Planck time. These extremely short times, in conjunction with the mechanism acting as the trigger and the information about the curvature (4.23), suggest the following interpretation of the effective source (4.18) by analogy: it may be understood as a shock wave, originated by the violent bounce of the stellar structure at $r = R_0$, that propagates outwards. The decay rate of this "curvature wave" with the radius goes as $r^{-3/2}$, which is to be compared with, e.g., the decay of the classical value of the Kretschmann scalar $\mathscr{K} = \mathscr{R}_{abcd}\mathscr{R}^{abcd}$ of the Schwarzschild geometry, valid for $|t| > t_{\rm R}/2$,

$$\sqrt{\mathscr{K}|_{|t|\geq t_{\rm R}/2}} = 2\sqrt{3}\,\frac{r_{\rm s}}{r^3}. \tag{4.25}$$

Indeed, to be more consistent we may compare this expression with the ultraviolet corrections to the Kretschmann scalar at $t = 0$ for instance. The evaluation of the leading order of this quantity for short transients shows that it behaves as Eq. (4.23):

$$\sqrt{\Delta\mathscr{K}|_{t=0}} \simeq 6\sqrt{2}\,\frac{\dot{g}_{t_{\rm R}}}{c}\sqrt{\frac{r_{\rm s}}{r^3}}. \tag{4.26}$$

While on R_0 the classical and "ultraviolet" parts of the Kretschmann scalar are of the same order by construction, their decay rates for $r > R_0$ are rather different. A look at Eqs. (4.25) and (4.26) suffices to check that the effects associated with the shock wave decrease more slowly than the curvature of the classical geometry. We could say that the wave preserves its non-perturbative character, with respect to the background, in its way out. As we will discuss in detail in Sec. 4.5.5, this is the mechanism that permits to understand how ultraviolet modifications are able to alter dramatically the geometry on the horizon, even if the curvatures might no longer be Planckian there. Notice that the infinite radial extension of this shock wave is an unphysical feature of this simplified model that will not survive in more elaborated models as we explain below.

4.5.3 Collapse from a finite radius and triangular-shaped transition

Up to now, we have considered that the initial radius of the stellar structure is infinite. Let us look for the description of the more realistic case of gravitational collapse from a finite radius $r_{\rm i}$. The resulting metric will be written again in the form of Eq. (4.13). The collapse of the star begins at rest at the initial radius $r_{\rm i}$ and is represented by the trajectory of the star's surface $R(t)$. The arbitrary zero of time is chosen such that the collapse starts at $t = -t_\mathrm{B}/2$, where

$$t_\mathrm{B} = \pi \frac{r_\mathrm{i}}{c}\sqrt{\frac{r_\mathrm{i}}{r_\mathrm{s}}} \tag{4.27}$$

is twice the classical collapsing time from $r = r_\mathrm{i}$ to $r = 0$ in the Oppenheimer-Snyder model. For $t \leq -t_\mathrm{B}/2$, i.e., before the collapse has started, the metric is exactly Schwarzschild outside r_i, while for $r \leq r_\mathrm{i}$,

$$A^2(t,r) = c^2, \qquad B^2(t,r) = 1 - \frac{r_\mathrm{s} r^2}{r_\mathrm{i}^3}, \qquad v(t,r) = 0. \tag{4.28}$$

This metric corresponds to a star of homogeneous density, maintained static by an appropriate internal pressure. Once the collapse has started, the three coordinate patches we use are $0 \leq r \leq R(t)$, $R(t) \leq r \leq r_\mathrm{i}$, and $r_\mathrm{i} \leq r \leq +\infty$. The functions describing the metric are then:

$$A^2(t,r) = c^2 \times \begin{cases} \dfrac{1 - r_\mathrm{s}/r}{1 - r_\mathrm{s}/r_\mathrm{i}} & r_\mathrm{i} \leq r, \\ 1 & R(t) \leq r \leq r_\mathrm{i}, \\ 1 & r \leq R(t); \end{cases} \tag{4.29}$$

$$B^2(t,r) = \begin{cases} 1 - r_\mathrm{s}/r & r_\mathrm{i} \leq r, \\ 1 - r_\mathrm{s}/r_\mathrm{i} & R(t) \leq r \leq r_\mathrm{i}, \\ 1 - \dfrac{r_\mathrm{s}}{r_\mathrm{i}}\left(\dfrac{r}{R(t)}\right)^2 & r \leq R(t); \end{cases} \tag{4.30}$$

$$v(t,r) = -c\left[1 - 2g_{t_\mathrm{R}(r)}(t)\right] \times \begin{cases} 0 & r_\mathrm{i} \leq r, \\ \sqrt{\dfrac{r_\mathrm{s}}{r} - \dfrac{r_\mathrm{s}}{r_\mathrm{i}}} & R(t) \leq r \leq r_\mathrm{i}, \\ \dfrac{r}{R(t)}\sqrt{\dfrac{r_\mathrm{s}}{R(t)} - \dfrac{r_\mathrm{s}}{r_\mathrm{i}}} & r \leq R(t). \end{cases} \tag{4.31}$$

In this case, the function $R(t)$ is constructed from the calculation in general relativity for the Oppenheimer-Snyder collapse from a finite radius r_i:

$$R(t) = \frac{r_\mathrm{i}}{2}(1 + \cos\eta), \qquad t = \frac{t_\mathrm{B}}{2\pi}(\eta + \sin\eta - \pi), \tag{4.32}$$

with $\eta \in [0, 2\pi]$. The point $\eta = \pi$ ($t = 0$) in this trajectory would correspond to the classical singularity; for $\eta < \pi$ the trajectory goes inwards, while for $\eta > \pi$ it goes outwards. Again, we shall modify the function $R(t)$ so that it is bounded from below by $R(0) = R_0 > 0$.

A new feature in this situation is the necessary introduction of a radial dependence on the parameter t_R that controls the interpolation time, which then becomes a function

$t_{\mathrm{R}}(r)$. This quantity enters through the function $g_{t_{\mathrm{R}}(r)}(t)$ in Eq. (4.31). This is necessary to guarantee that curvature invariants are kept finite in the surroundings of the $r = r_{\mathrm{i}}$ hypersurface. Had we taken this parameter as being constant, the leading order of the Ricci scalar in the limit $r \to r_{\mathrm{i}}$ from below would be given, at $t = 0$ for instance, by

$$\mathscr{R}|_{t=0} \simeq \frac{4r_{\mathrm{s}}(3r_{\mathrm{i}} - 4r)}{r_{\mathrm{i}}\, r^2 \sqrt{\dfrac{r_{\mathrm{s}}}{r} - \dfrac{r_{\mathrm{s}}}{r_{\mathrm{i}}}}} \frac{1}{c\, t_{\mathrm{R}}}. \tag{4.33}$$

On the other hand, the consideration of *decreasing* functions $t_{\mathrm{R}}(r)$ with the radius is clearly motivated by the understanding of the modifications of the near-horizon geometry as the result of the propagation of non-perturbative ultraviolet effects from $r = R_0$ up to $r = r_{\mathrm{i}}$. It is therefore quite remarkable that avoiding that the scalar curvature blows up at $r = r_{\mathrm{i}}$ at $t = 0$ demands that $t_{\mathrm{R}}(r)$ verifies the differential equation

$$(3r_{\mathrm{i}} - 4r)\frac{1}{t_{\mathrm{R}}(r)} + 2r(r_{\mathrm{i}} - r)\frac{d}{dr}\left(\frac{1}{t_{\mathrm{R}}(r)}\right) = 3C\sqrt{r_{\mathrm{i}} - r}, \tag{4.34}$$

where C is an arbitrary constant. In this equation, the first term on the left-hand side corresponds to the contribution reflected in Eq. (4.33), while the second term on the left-hand side corresponds to additional contributions to the Ricci scalar that show up when the radial dependence of the parameter $t_{\mathrm{R}}(r)$ is included. The right-hand side of this equation is the minimal expression which cancels the term that goes to zero in the denominator of Eq. (4.33), thus ensuring a good behavior of the Ricci scalar in the limit $r \to r_{\mathrm{i}}$.

The inhomogeneous solution to this differential equation is given by the decreasing function

$$t_{\mathrm{R}}(r) = \frac{1}{C}\sqrt{r_{\mathrm{i}} - r}. \tag{4.35}$$

The simplest ansatz for our geometry is assuming that Eq. (4.35) holds for the entire interval $R_0 \leq r \leq r_{\mathrm{i}}$. Demanding $t_{\mathrm{R}}(R_0) = t_{\mathrm{P}}$ fixes the constant C, leading to

$$t_{\mathrm{R}}(r) = t_{\mathrm{P}}\sqrt{\frac{r_{\mathrm{i}} - r}{r_{\mathrm{i}} - R_0}}. \tag{4.36}$$

One can show by direct evaluation that this choice regularizes at the same time other curvature invariants such as, e.g., the Kretschmann scalar. Fixing $t_{\mathrm{R}}(r)$ as in Eq. (4.36) specifies completely the geometry, in which the thin-layer region that encloses the non-standard ultraviolet effects that is depicted in Fig. 4.1 will be generally transformed into a (smoothed) triangular-shaped region, with one of the vertices located at $(t, r) = (0, r_{\mathrm{i}})$.

The inevitable introduction of the function $t_{\mathrm{R}}(r)$ is intimately related to the fact that non-perturbative ultraviolet effects are naturally confined in this case into a compact ball of radius r_{i}, so that the form of the Schwarzschild metric is preserved for $r > r_{\mathrm{i}}$. In the most general case the maximum radius reached by ultraviolet effects could be an independent parameter r_{m} such that $r_{\mathrm{m}} > r_{\mathrm{s}}$. The explicit construction of the corresponding geometries is more involved, though there is in principle no obstruction for their construction;

all the necessary ingredients for doing so should be contained in the present discussion. For completeness, the diagrams that would correspond to these most general situations are reproduced in the following. Even if being slightly more restricted, the geometries we have explicitly constructed here have nevertheless the necessary properties to reflect appropriately the main implications of the black- to white-hole transition in short characteristic time scales. We shall therefore take them as the starting point for the exploration of these implications.

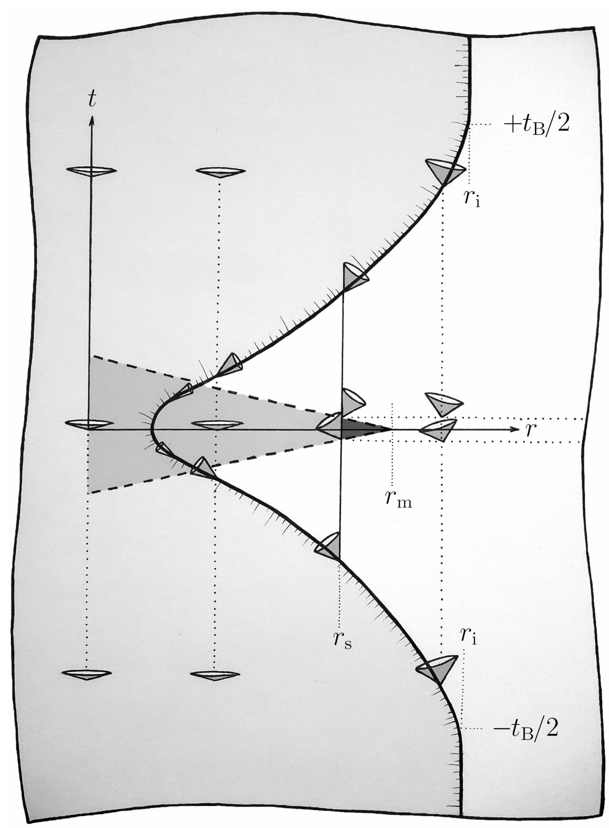

Figure 4.2: The figure represents the collapse and time-symmetric bounce of a stellar object in our proposal (the thick line). The past thick dashed line from $r = 0$ to r_m marks the boundary where the non-standard gravitational effects start to occur. In all the external white region the metric is Schwarzschild. In the region between the two thick dashed lines (which extends outside the stellar matter itself) the metric is not Schwarzschild, including the small dark gray triangle outside the Schwarzschild radius r_s. The drawing tries to capture the general features of any interpolating geometry. The slope of the almost Minkowskian cones close to the origin has been taken larger than the usual 45 degrees to cope with a convenient and explicit time-symmetric drawing. Note that the smoothing of the geometry at $r = r_m$ through the introduction of the function $t_R(r)$ has not been made explicit in the diagram for simplicity.

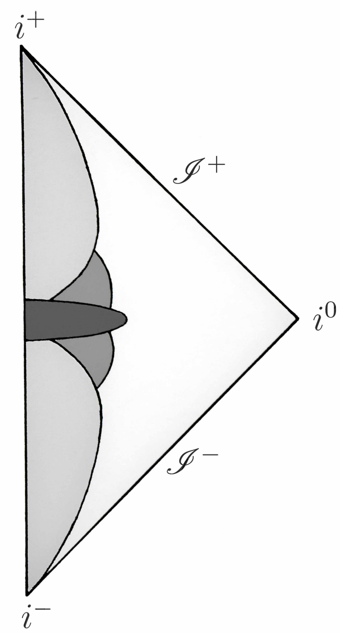

Figure 4.3: The figure represents the Penrose diagram of the proposed geometry. Globally it has the same causality than Minkowski spacetime. Locally it has some peculiarities. The dark gray region represents a non-standard gravitational field, while the up and down gray regions are respectively regions with past- and future-trapped surfaces. The light gray regions on the left-hand side are those filled by matter.

4.5.4 Short-lived trapping horizons

The first of the quantities of interest we can evaluate shows one of the nontrivial features that characterizes the entire family of bouncing geometries we are considering in this chapter. In the first model we have discussed with a thin-layer transition region, calculations that are performed in the limit $t_{\rm R} \to 0$ are exact up to $\mathscr{O}(t_{\rm R}/t_{\rm B})$ terms. This implies that this distributional limit is indeed a very good analytical approximation, on the basis of Eqs. (4.24) and (4.27). This is also true in the second model with a triangular-shaped transition region, again in the limit in which the transition between the black-hole and white-hole geometries is abrupt, namely $t_{\rm R}(R_0) \to 0$ where R_0 is the minimum radius of the stellar structure.

Let us therefore start by considering the distributional limit $t_{\rm R}(R_0) \to 0$ for mathematical simplicity. In this case, the proper time for observers situated at $r = r_{\rm i}$ for the bounce to take place is exactly the same as for observers attached to the surface of the star. Indeed, from the perspective of an observer collapsing with the star, the entire process of collapse and bounce with respect to some reference initial position takes a time $t_{\rm B}$ as defined in Eq. (4.27), that is, twice the free-fall collapsing time. By construction of the coordinate system, the proper time of the observer attached to the stellar structure is simply the Painlevé-Gullstrand time, $d\tau = dt$, as for that observer $dr/dt = v(t,r)$. This same process seen by an observer always at rest at the initial position $r_{\rm i}$ takes *this very same time*. For this observer $dr = 0$, $v(t,r_{\rm i}) = 0$ and $A^2(t,r_{\rm i}) = c^2$, thus implying $d\tau = dt$; its proper time is therefore given by the same temporal coordinate we are using to write down the metric. Seen by observers far outside the collapsing star ($r \gg r_{\rm s}$) the entire process would take this time multiplied by the standard general-relativistic redshift factor, $(1 - r_{\rm s}/r_{\rm i})^{-1/2}$. This redshift factor will be of order unity for stars initially larger than a few times their Schwarzschild radius. The fact that the proper time separation between these two events, the start of the collapse and its return to the initial position, is the same for the surface-attached observer and the one standing still at the initial radial position, is a general property of the geometries we are considering: a finite value of $t_{\rm R}(R_0)$ of the order of Eq. (4.24) leads, taking into account Eq. (4.27), to extremely small $\mathscr{O}[t_{\rm R}(R_0)/t_{\rm B}]$ corrections coming from the finite transient zones. Therefore, these geometries do not exhibit long-lived trapped regions of any sort, but only short-lived trapped regions.

From a different point of view, imagine that one were to monitor with high time resolution this time-symmetric collapse from the reference initial position, that is assumed to be sufficiently far outside the Schwarzschild radius. One sets up two synchronized clocks at this initial position before the collapse. Then one clock is left to follow the collapsing structure and the other is kept at rest in the reference position. By observing with a telescope the ticks of the clock falling with the star, one would see that in the collapsing phase the ticks slow down progressively. However, at some point they start to speed up in such a way that when the two clocks are finally back together they show precisely the same time. This is easily understood if we think in terms of the analogue metric in fluids [14]. When the star is collapsing, the velocity profile is that of a sink, so that signals originated on the star's surface will be emitted at different positions, picking up a delay that depends on the time that light needs to cover up that distance. On the contrary, in the white-hole patch the velocity profile is reminiscent of a source, which effectively shortens the time between different signals for the external observer. Overall, for an external observer, that uses essentially the Schwarzschild time coordinate as its proper time (but for an irrelevant redshift factor depending on his position), the collapsing phase will last longer than the expanding phase, so that he might not realize the time-symmetric character of the process.

In summary, one of the essential features of the process we are considering is that the time lapse associated with the collapse is short, of the order of tenths of a millisecond for neutron-star-like initial configurations. This is equally true both for observers attached to the structure as well as for external stationary observers. Within this quite general family of geometries (arguably a complete set of geometries interpolating between a black-hole geometry with the white-hole geometry), the only way to prescribe geometries that allow for extremely long times of the bounce process as seen by external observers is to introduce extremely slow transients. By using the adjective *extremely slow* we understand a regularization that takes a very long time, measured by observers attached to the surface of the star, in order to overcome the gravitational attraction and start a noticeable expansion of the stellar structure; in mathematical terms this corresponds to very large $t_{\rm R}(R_0)/t_{\rm B}$ quotients. From the perspective of the regularization of the singularity as well as observers *inside* the stellar structure, this quasi-static possibility is arguably quite unnatural. In the framework of our discussion in Sec. 4.4, these solutions would correspond to prioritizing the role of the external metric in the entire process, subjugating the behavior of matter as well as the possible ultraviolet effects to the prevalence of the long-lived horizon.

4.5.5 Short transients and the propagation of non-perturbative ultraviolet effects

All the transients lead to a characteristic imprint as we have already discussed: the deviation of the near-horizon outer geometry, that is, the Schwarzschild geometry beyond $r = r_{\rm s}$. Such a feature is anathema in the orthodox view on the possible relevance of ultraviolet effects on classical geometries. It belongs to the conventional wisdom that appreciable deviations from the classical behavior are to be expected in regions of large curvature, namely of order one when measured in Planck units. The argument is that only then the possible corrections to the Einstein field equations are expected to be non-perturbative. In the Schwarzschild solution the Ricci curvature tensor is zero but the Weyl part of the Riemann curvature tensor leads to the Kretschmann scalar

$$\mathscr{K} = \mathscr{R}_{abcd}\mathscr{R}^{abcd} = \frac{12 r_{\rm s}^2}{r^6}. \tag{4.37}$$

This implies that the geometry around $r = r_{\rm s}$ of stellar-size black holes is to be regarded as a robust feature, not to be affected by any kind of ultraviolet modifications.

While this view certainly makes sense in the study of the static case (i.e., an eternal black hole), the role of these considerations is far from clear for us when talking about dynamical situations. For instance, in our proposal there is no modification at all of the behavior of the geometry near the moment of formation of the trapping horizon in the black-hole patch. An observer who suddenly left the surface of the star to pend close, but outside the horizon, will not notice any deviations from the expected behavior in general relativity in the few initial instants of the process. It is only when matter reaches a dense enough situation that relevant ultraviolet effects altering the geometry are originated. While at first these effects are confined to regions of high curvature, there is nothing that prevents them to propagate outwards, modifying the geometry in their wake. In doing so these effects can reach regions that were characterized by low curvatures, such as the near-horizon region. However, this effect should not be seen as a modification of the behavior of regions with low curvature by unnaturally large ultraviolet effects, but rather as the propagation of a non-perturbative wave of high-curvature through regions of low curvature. Interestingly, this picture is self-consistent only for short transients, with characteristic time scale $t_{\rm R}(R_0)$ of the order of the Planck time.

This is neatly illustrated by taking as working example the geometry considered in Sec. 4.5.1, though the situation is generically the same for all the geometries we have considered. The Kretschmann scalar of the transient in this case is given in Eq. (4.26). As argued above, the short (Planckian) time $t_{\rm R}$ associated with the transient implies that the modifications of the spacetime curvature in the surroundings of the stellar structure, when the latter reaches its minimum radius R_0, are also of the order of the Planck curvature. Thus in geometries with short transients, and only in these cases, the trigger of non-perturbative ultraviolet effects on the metric outside the star is inextricably tied up to spacetime regions with high curvature. This region of high curvature will propagate outwards, getting diluted in this process: already when reaching the horizon, the magnitude of this curvature several orders of magnitude lower, roughly by

$$\left(\frac{R_0}{r_\mathrm{s}}\right)^{3/2} \sim \frac{m_\mathrm{P}}{M}. \tag{4.38}$$

One might be tempted to argue on this basis that ultraviolet effects should not be able to significantly alter the near-horizon geometry, being the corresponding curvature far smaller than Planckian.[7] The situation is, indeed, the contrary. While the classical distribution of the curvature (measured by the Kretschmann scalar) decays with the radius as r^{-3} [Eq. (4.25)], the curvature of the non-standard region associated with ultraviolet effects decays as $r^{-3/2}$ [Eq. (4.26)]. This guarantees that modifications that are non-perturbative at R_0, remain non-perturbative near the horizon. Thus modifications of the near-horizon geometry do not appear because of unnaturally large ultraviolet effects originated there, but rather as a result of the propagation of sudden ultraviolet effects that are originated when the stellar structure undergoes a violent bounce at Planckian densities.

An argument of completely different nature, but also in favor of short transients, is the following. As stated above, the expanding regime can be described by means of the white-hole patch of the Kruskal manifold, excluding a region surrounding the past singularity. It is well known that there exist analyses concluding that white holes are unstable, being rapidly transformed into black holes [263, 264, 265, 266, 267, 268, 269, 270, 271]. Quite remarkably, the characteristic timescale for the different instabilities to develop is of the same order of magnitude as the lifespan of the past-trapped region in our model, roughly r_s/c. A sharp proof of this assertion can be constructed by using suitable null coordinates near the white-hole horizon. Let us take for instance the classical instability against the accretion of matter, though the underlying reason for these instabilities to occur is known to be the same. Then one can show explicitly that, the longer the time for the black- to white-hole transition to unfold completely, the less amount of matter is needed in order to inhibit the transition. The relevant parameter that marks the different regimes is given by

$$V_0 = \exp(-c\Delta t_\mathrm{B}/4r_\mathrm{s}), \tag{4.39}$$

where Δt_B is the additional delay with respect to the perfect bounce situation we have analyzed in detail. The exponential relation between Δt_B and the amount of matter that is needed to inhibit the white hole explosion comes directly from the exponential relation between Kruskal-Szekeres coordinates and Eddington-Finkelstein coordinates. For V_0 of order unity, the white hole explosion is robust against the perturbations that represent the

[7] It is interesting to note that it is in principle possible to construct models in which the consideration of triangular-shaped transition regions could even lead to Planckian curvatures near the horizon, by just making an appropriate ansatz for $t_\mathrm{R}(r)$ from $r = R_0$ up to the near-horizon region.

accreting matter, while it becomes highly unstable for $V_0 \to 0$. Therefore, transitions with long characteristic time scales are highly unlikely to occur, being exponentially suppressed. Only transitions with short characteristic time scales such that Δt_{B} is sufficiently close to zero could take place in astrophysical scenarios.

It is certainly not possible, given our present knowledge of the gravitational interaction, to back up the picture with short characteristic time scales from a first-principles, complete perspective. The work of Hájícek and collaborators [272, 273, 274, 275], represents nevertheless a first step in this direction. In their study of the quantization of collapsing shells, they have shown that the transition from a collapsing branch to an expanding branch occurs, and that a given definition of the crossing time between an external observer and the dynamical shell is short, being essentially of the same order of magnitude as the time t_{B} we are considering, given by Eq. (4.27). It is equally worth remarking their skepticism about this result, as the same authors were expecting very long times to occur, in order for their model to be accommodated within the REBH family. The definition of observables concerning time intervals in these bouncing processes is however far from straightforward [274, 275]. At the light of the developments presented here, this issue certainly deserves further study, both in this particular model and in frameworks that use different quantization techniques, such as [276] for instance. While it is not strictly mandatory that any ultraviolet completion of general relativity has to permit the black- to white-hole transition, as described here, in short characteristic time scales, this is indeed the only chance for this transition to occur. Otherwise, black holes will certainly keep black, with just a tiny evaporative effect due to the Hawking evaporation.

In any case, lacking for the moment a robust fundamental justification for the short transients does not prevent our discussion to be self-consistent at low energies, while presenting new intuitions about the way in which ultraviolet modifications to the behavior of general relativity would behave. On top of this, what we want to stress is the genuine opportunity that these hypothetical processes offer: in contrast with REBH proposals, which are experimentally inert for extremely long times and thus for practical purposes, a process with a short characteristic time scale should lead to clear imprints that could be hopefully detectable. Thus this proposal is audacious, but not without consequences, as it offers prospects of being falsifiable much more easily than any other models that are nowadays present the literature.

4.6 Physical and observational consequences

If there exists a regularization of the classical behavior of the form we have described, the collapse process itself would not constitute the final stage of collapse in stellar physics. One would immediately be impelled to wonder about what would happen after the bounce. The search for new states of equilibrium on the one hand, and the understanding of the transient collapse process itself on the other, become entirely distinct issues.

If our idea is at work in nature it should have many observable effects, though additional work is needed in order to determine the possible signatures of both phases. In the collapse of, say, a neutron star, matter would remain apparently frozen at the Schwarzschild radius for just a few tenths of a millisecond before being expelled again. Even neglecting dissipation, the metric is non-Schwarzschild during an extremely short time interval in a region extending outside the gravitational radius. In realistic situations the bounce will not be completely time-symmetric: part of the matter will go through towards infinity in the form of dissipative winds while the remaining mass will tend to recollapse. In this way one would have a brief and violent transient phase, composed of several bounces, followed by the formation of a new (stable or metastable) equilibrium object with the remaining mass.

4.6.1 Towards new figures of equilibrium

In an ideal situation, perfectly spherically symmetric and without dissipation, the collapsing body would enter a never-ending cycle of contracting and expanding phases. In a realistic situation, though, one expects that the system will dissipate at least quantum-mechanically while searching for new equilibrium configurations. Here the panorama of possibilities is almost unexplored. Let us discuss the possibility we think most plausible. Notice that any development on this direction should properly acknowledge and deal with, for instance, the observations of the existence of extremely dense objects in our own galaxy.

In [260] it was shown that if the velocity of trapping-horizon crossing by the collapsing matter distribution were rather small, then the quantum effects of vacuum polarization would become so powerful that they might even stop the collapse. It is very unlikely that the classically expected almost-free-falling collapse of stellar structures like neutron stars would lead directly to strong vacuum polarization effects. However, in our scenario, when taking into account dissipation, one would expect that each new recollapsing phase would start from a position closer to its Schwarzschild radius than the previous one. In this way, at some stage vacuum polarization effects could start to be relevant and even stop further collapse cycles. This might lead to hypothetical almost-stationary structures hovering extremely close to their gravitational radius.

There might be other mechanisms underlying the stabilization of stellar structures close to their gravitational radius. What is relevant here is that the final metastable object could be small, dark, and heavy, but without black- or white-hole districts (see [277] and references therein). These black stars will not be voids in space, but they will be filled with matter. Since they have no horizons, they will in principle be completely open to astrophysical exploration. Let us stress that black holes as described by classical general relativity might still continue to be very good models for the external gravitational behavior of these black stars.

A natural question in this regard is whether Hawking-like evaporation, being a paradigmatic theoretical feature of black holes, would be also a characteristic of these new objects. When the system stabilizes close to its Schwarzschild radius, it might emit or not and with different spectral properties depending on the specifics of the structure, which at this stage are difficult to envisage. In other words, the Hawking effect *does not* need to be preserved in the black-star scenario. However, we have already mentioned in our previous discussion that there are ways in which these objects could acquire emission and evaporation properties resembling Hawking's scenario. At least two different mechanisms are known. One such structure could emit a Hawking-like flux if it were continuously and asymptotically approaching its Schwarzschild radius (without crossing it) [244], or if it were pulsating in a close to free-fall manner [245]. Should this radiation exist, it would in both cases be Planckian but not strictly thermal, as correlations are maintained by both mechanisms, and the total amount of energy radiated would be in principle negligible given the extremely short time associated with the transients. In relation with Hawking radiation, notice that these scenarios do not invite us to wonder whether information is lost or not, as no singularities and no long-lived trapping horizons are formed in the first place. On the other hand, during the transient phases one would expect quantum dissipation in the form of particle production. This particle production will be in general non-thermal, though at trapping-horizon crossings it would have the form of short bursts of thermal radiation [243, 244, 245].

Once the object has settled down, it would probably be extremely difficult to distinguish it from a standard black hole through astrophysical observations. There have been some proposals to discern whether or not there exists an explorable surface in objects associated with black holes [278, 279]. The absence of Type I X-ray emission in binaries containing a black-hole mimicker have been argued to imply that the candidate did not have an external surface but an event horizon [278]. However, things are clearly not that simple (for some specific criticism see [165]). The reaction of the black hole mimicker when absorbing some matter from its companion would strongly depend on its specific heat capacity. If the black-hole candidate has a heat capacity similar to that of a black hole, which can be expected due the high redshift of its surface, its behavior would be difficult to distinguish in this regard from that of a proper black hole. This kind of observations can play a significant role, though, in putting constraints to specific models of black hole mimickers (see, e.g., [280] for constraints on gravastars). This task might prove even more difficult to accomplish for proposals in which the compact object lacks a hard surface [281].

A different issue is the possibility, in principle, of the hypothetical use of a radar to check the presence of a surface. An *elastic* scattering of a wave in the supposed surface located at $r = r_{\rm star}$ would distinguish in no time whether a hard surface exists or not. The general relativistic delay can be calculated to be

$$T = \frac{2}{c}\int_{r_{\rm star}}^{r_0} \frac{dr}{1 - r_{\rm s}/r} = \frac{2}{c}\left[r_0 - r_{\rm s} + r_{\rm s}\ln\left(\frac{r_0 - r_{\rm s}}{r_{\rm star} - r_{\rm s}}\right)\right], \tag{4.40}$$

with r_0 the observation point. The divergent term is logarithmic so that it can never become too large in realistic situations. For instance, for a Solar mass black star with a radius larger than its Schwarzschild radius (3×10^3 m) by the tiny amount of 10^{-75} m (which is about 10^{-40} times the Planck length), a radar signal sent from a distance of 8 light-minutes would acquire a gravitational delay of only a few milliseconds and would echo back after about 16 minutes plus few milliseconds, in sharp contrast with the *infinite* amount of time necessary in the case of a proper black hole.

4.6.2 Energetics of the transient phase

When considering realistic situations with dissipation, the transient phase might leave some traces, for instance in the physics of gamma-ray bursts (see, e.g., [282]). There is experimental evidence that a subset of these events, the so-called *long* gamma-ray bursts, are associated with the final stages in the life of very massive stars. The most widely accepted theoretical picture is known as the collapsar model [283]. It is natural to expect that a modification of the standard gravitational collapse process to a black hole that is considered here could leave clear imprints associated with a reverberant collapse. In the collapsar model of GRBs the emission zone is supposed to be very far from the collapsed core [282]. This means that the connection between the processes at the core and those at the external wind shells could be very far from direct. However, the general features of the model are enough in order to roughly compare its energetics to those of GRBs. This comparison may be used in order to understand whether or not the bounce process is a reasonable candidate for the mechanism behind these bursts. So let us assume that the picture discussed above is realized in nature: the occurrence of violent bouncing processes, dissipation and final stabilization in the form of a black star. We can estimate the effect of the energy loss in the entire process by using the following argument. Recall that for dust matter content, and in the absence of rotation, the differential equation for the trajectory of the surface of the star is mathematically equivalent to that of a test particle with the overall mass of the star following a radial geodesic in Schwarzschild spacetime. We can then use the conserved quantities associated with the geodesic equations in this spacetime. In particular, we shall use the conserved quantity E that is associated with energy. So let $r_{\rm i}$ be the initial radius and $r_{\rm s}$ the Schwarzschild radius of the star, and consider the Schwarzschild effective potential for radial motion. If the structure was originally at rest, its energy is given by

$$\left(\frac{E}{Mc^2}\right)^2 = 1 - \frac{r_{\rm s}}{r_{\rm i}}. \tag{4.41}$$

This equation has a clear interpretation: the positive term on the right-hand side is the rest energy of the star, while the second term corresponds to the negative gravitational energy of the structure. Then the energy of the resulting compact body, as defined in Eq. (4.41) and taking $r_{\rm i} \simeq r_{\rm s}$, is essentially zero. Energy balance implies that the energy that has to be released in the entire process, for example by means of the emission of a shell of matter that escapes to spatial infinity, is given by:

$$\Delta E = Mc^2 \sqrt{1 - \frac{r_{\rm s}}{r_{\rm i}}}. \tag{4.42}$$

This is a model-independent estimation that tell us that the object has to get rid of a significant portion of its original rest energy in order to reach stabilization. If we take for example a neutron star with $r_{\rm i} = 2\, r_{\rm s}$, then the emitted energy is $Mc^2\sqrt{1/2} \simeq 0.71\, Mc^2$.

How is this compared with the GRBs energetics? Interestingly, Eq. (4.42) is of the same order of magnitude as the energy emission in those events if the emission is considered isotropic (see, e.g., [282, 284]). Indeed, it is of the same order of magnitude as other theoretical estimates that are based, for instance, on the Penrose process [285] or similar mechanisms for energy extraction from charged black holes [286, 287]. A comment that applies to all these estimations is that there is experimental evidence that the emission is collimated, so that the real energy that is emitted is smaller than Eq. (4.42), roughly by a factor of 10^{-2}. Of course, Eq. (4.42) is a crude estimate that does not take into account other effects that would take place near the collapsing object, besides being evaluated in an isotropic model that does not take into account rotation. That with simple, model-independent ingredients we are able to get that close to the observed energetics of GRBs is a strong incentive to consider further developments of the picture. This is a hint in favor of the possibility that the bounce process we are describing might be behind of some GRB events. As we cover in the next section, gravitational waves produced deep inside these violent events may provide a much better observational opportunity, eventually permitting to elucidate the kind of mechanism that is behind them.

4.6.3 Ripples from the transient phase

One of the primary predictions of general relativity that still awaits experimental corroboration is the generation and propagation of local disturbances of spacetime, or gravitational waves. Gravitational-wave astronomy is nowadays a mature branch from a theoretical perspective, while there has been a great deal of experimental effort in order to overcome the difficulties in the detection of these tiny ripples in spacetime. The scientific potential that this new observational window promises is huge; see [166, 288] for instance. The fine theoretical knowledge of the gravitational wave patterns associated with different gravitational phenomena would make possible to unveil information about astrophysical processes that is definitely not possible to obtain from their electromagnetic counterparts. As a prime example, this observational technique is arguably the best tool to finally determine the physical mechanisms that are behind both short and long GRBs. If the process that we are describing in this chapter is realized in nature, the information encoded in gravitational waves may be significantly different, and therefore more surprising than initially expected.

The observation of the gravitational wave pattern associated with the gravitational collapse of a massive star into a black hole, if properly correlated with the electromagnetic counterpart of a long GRB, would be the smoking gun of the collapsar model. The generation and form of these wave patterns are well understood nowadays. In the spherically symmetric case, non-spherical inhomogeneities that are present in the initial stellar structure will generate a gravitational wave signal, that terminates with the relaxation of the perturbed horizon to its stable Schwarzschild form (in the presence of rotation it will be described by the Kerr solution instead [37]). After that, there is complete silence in the gravitational-wave channel. This is a definite characteristic of the classical gravitational collapse process as described in general relativity, that will be shared by *all* the REBH models. The robustness and generality of this result is what makes any departures from it highly interesting. We shall describe in the following why do we expect distinctive departures of this behavior to occur in our model, and how can these be computed. It is noteworthy that this is only proposal in the literature to our knowledge that advocates for these kind of modifications on the gravitational wave patterns of collapsing stars.

Let us start with a very simple analogue, given by the electromagnetic radiation of a pulsating, or bouncing, charged sphere. In the spherically symmetric case there will be no emission of radiation. As in the gravitational case, one has to consider non-spherical distributions of charge in order to trigger the emission of electromagnetic radiation. Now the electromagnetic radiation measured far from the sphere will depend on the given trajectory of the radius of the sphere, $R(t)$, bounded both from below and above. We take this trajectory as the analogue of the trajectory of the star surface in our bouncing geometry. The overall emission of radiation can be obtained in the framework of standard electrodynamics. The emission of electromagnetic radiation in the collapsing branch will be followed by the emission of new radiation pulses coming from the expanding branch, perhaps with a burst corresponding to the bounce event. Notice that the emission of radiation breaks the time symmetry of the trajectory $R(t)$, if present. Additionally, in several bounces we shall get several repetitions of this wave pattern. This is what we expect to occur in the gravitational case, thus leading to a very different gravitational signal when compared with the one obtained in general relativity.

Now when we turn back to gravity, there is a crucial technical difference to take into account when considering this analogy. While the equations of electrodynamics enforce the conservation of charge, the equations of general relativity do the equivalent with energy. In the electromagnetic case the energy that is emitted in the form of radiation is ultimately provided by the mechanism behind the kinematic evolution of the charge distribution. In other words, the non-spherical distributions of charge that cause the emission of radiation do not decay over time. On the contrary, the equations of general relativity automatically take into account the decay of non-spherical perturbations due to the emission of gravitational waves. This makes the problem more involved mathematically but, as we discuss in the following, having at hand the metric describing the bounce of the stellar structure one should be able to obtain a definite answer for the spectrum of gravitational waves that is produced in the process for a given perturbation of the initial configuration.

Let us sketch the necessary steps in order to do so, taking for instance the metric described in Sec. 4.5.3 as a specific representative of the bounce process. This metric satisfies by construction the equations

$$G_{ab} = \kappa \left(T_{ab} + \mathscr{T}_{ab} \right) , \tag{4.43}$$

where G_{ab} is the Einstein tensor, T_{ab} is the dust stress-energy tensor, and $\mathscr{T}_{ab}$ is the non-standard stress-energy tensor, the components of which would be similar to Eq. (4.18), that describes the shock wave produced in the violent bounce. Our previous discussion corresponds to the spherically symmetric situation; in order to produce gravitational waves we need to introduce non-spherical perturbations. These perturbations are introduced in the initial state of the system, namely the star at rest with radius r_{i}, with asymptotic conditions ensuring the absence of radiation at spatial infinity. We can perform a perturbative expansion around the spherically symmetric situation of the form

$$G_{ab} = G_{ab}^{(0)} + G_{ab}^{(1)} + ..., \tag{4.44}$$

where $G_{ab}^{(0)}$ is the Einstein tensor of the spherically symmetric geometry describing the bounce, and $G_{ab}^{(1)}$ will contain the information about the non-spherical perturbations and, in particular, the gravitational wave signal. At first order in the deviations from absolute spherical symmetry, the evolution of these perturbations are then determined by the equations

$$G_{ab}^{(1)} = \frac{8\pi G}{c^4} T_{ab}^{(1)}. \tag{4.45}$$

Notice that $\mathscr{T}_{ab}$ drops off from this equation, as it is constructed to identically cancel the zeroth-order terms. This last equation sets the basis for the study of gravitational wave emission of this process. The study of its implications is currently being carried out, the results of which will be reported elsewhere.

It is probably not necessary to stress the appeal of the possibility, even if it could appear remote at present, of detecting a characteristic gravitational wave signal that deviates from the expected classical template. Such an observation would be the smoking gun of the bounce process as described here, providing a low-energy observational window to genuine ultraviolet effects acting on the gravitational collapse of massive stars.

4.7 Conclusions

The outcome of extreme gravitational collapse processes is one of the great theoretical open problems in gravitational physics. The precise determination of the ultraviolet modifications to the classical behavior encapsulated in general relativity is of course the essential key to unveil its solution. It is not so often stressed that the present understanding of this problem is facing an important dilemma: most of the models in the market largely preserve the semiclassical picture of long-lived trapping horizons, thus obstructing their very experimental verification due to the ridiculously long lifetime of any REBH model. While we are far from denying the theoretical value of these models in understanding the gravitational interaction, in our opinion a serious effort should be made in analyzing the theoretical and observational characteristics of models that exhibit clear testable signatures in the near future.

While the perturbative view of the classical picture that is encoded in the REBH paradigm certainly represents the consensus of the community, we have tried to transmit that there exists an interesting alternative. In this alternative model, ultraviolet effects are no longer perturbative, so that the regularization of the singular behavior of general relativity opens new unexplored avenues for the evolution of the system. The resulting geometries cannot be described as representing essentially the semiclassical perturbation of the trapping horizons that are formed in the collapse, but give preeminence to the regularization of the singularity.

The first nontrivial result is that there exist geometries describing a transition between a black-hole geometry and a white-hole geometry in short characteristic time scales. With the adjective *short* we mean that both observers attached to the structure and distant stationary observers measure a short time interval for the overall bounce process, which is essentially twice the collapsing time evaluated in the Oppenheimer-Snyder model. In order to describe a black- to white-hole transition, it is mandatory that a certain open region outside the Schwarzschild radius deviates from the usual spherically symmetric vacuum solution. Moreover, the study of the curvature invariants of these geometries shows that only for short transients (of the order of the Planck time), the non-perturbative modification of the near-horizon Schwarzschild geometry is justified by the propagation of non-perturbative ultraviolet effects originated at the moment of the bounce. In these situations, the corresponding geometries represent the bounce of the matter distribution when reaching Planckian densities, originating a shock wave that propagates outwards, modifying the near-horizon Schwarzschild geometry. The effective matter content that describes the shock wave violates the standard energy conditions of classical general relativity, which are a fundamental ingredient of singularity theorems.

The rapid bounce of the distribution of matter radically changes the discussion of the possible endpoints of gravitational collapse. The inclusion of dissipative processes makes plausible that the final object is a compact object filled with matter; if this view is indeed realized in nature, black holes could be an idealized approximation to the ultimate stationary objects. That the typical energy scale for the energy that has to be dissipated to reach this equilibrium is roughly coincident to the energetics of GRBs represents a first nontrivial check of this proposal. Most importantly, the short characteristic time scale makes possible to think about the (always exciting) possibility of making contact with the empirical reality, being the most promising detection channel the measurement of the gravitational waves originated in the bounce.

Main conclusions and future directions

The goal of this thesis has been the exploration of both fundamental aspects and applications of the emergent gravity program. The distinctive feature of this research program is that most of the features of the known theories of physics are taken as emergent; for instance, the geometric content of general relativity, or the Lorentz invariance of the standard model of particle physics. In this rationale, the focus is not on the construction of a fundamental theory of nature with very specific features, but rather the study of the robust mechanisms that could lead to the emergence of the relevant low-energy properties within a universality class of systems.

Each of the chapters making up this dissertation presents a final section with a detailed discussion of the conclusions reached. In order to avoid repetition, and make this last section a useful complement to the body of the text, we shall end this dissertation with a description of a selected set of conclusions that have larger scope, and a brief outlook of the possibilities that have been opened in terms of potential future developments.

The discussion in Chap. 1 about the construction of a specific condensed-matter-like model leading to massless electrodynamics at low energies presents a number of remarkable features. It is always instructive to have at hand an explicit example which permits to gain an additional intuition about some arguments or results that arise in more general, but necessarily less detailed discussions. In this regard, one of the most important features to be extracted from this model is the way in which the emergence of an effective Lorentz invariance is ensured. As mentioned in the introduction of the thesis, previous discussions about emergent models have mainly dealt with the evolution of certain fields on top of a given gravitational field, therefore ignoring the very dynamics of the gravitational degrees of freedom. In the simpler framework of electrodynamics, this would correspond to the consideration of background electromagnetic fields. The virtue of the model of emergent electrodynamics presented here is that all the fields are dynamical, and therefore the nature of the electromagnetic excitations is displayed explicitly.

This step forward in the construction of emergent models permits the understanding of some features that were not possible to probe in previous approaches, in particular the mechanisms that could be behind the emergence of Lorentz invariance. As briefly noted in the introduction, and explained in more detail in Chap. 1, the application of standard arguments of effective field theory implies that the introduction of interactions between the effective fields leads to the percolation of Lorentz-violating effects at low energies, leading to unacceptably large effects. In the standard model of particle physics interactions are mediated by bosonic particles which, in the particular case of electrodynamics, correspond to photons. In the condensed-matter-like model we have discussed in detail, there is a coherence length for the effective photons, associated with the very occurrence of Cooper pair formation and their condensation. In practice, this imposes a physical cutoff in loop integrals that forbids the percolation of Lorentz-violating effects at low energies. Therefore, in this particular model, there exists a nontrivial mechanism that guarantees the robustness of Lorentz invariance against radiative corrections. It is natural to conclude that this feature could be extended to more complicated systems that are based in similar mechanisms for the emergence of the low-energy degrees of freedom.

While we have not treated the gravitational case with the same level of detail as the electromagnetic case due to time constraints, we have considered in Chap. 2 the most distinctive properties of this interaction. The results of our in-depth analysis of the self-coupling problem of gravitons differ from the standard lore in this topic. In particular, we have argued that the overall construction is highly non-unique. Once this bumpy road has been traveled, the most clear way to highlight this non-uniqueness is to stress the existence of a family of nonlinear theories of gravity that can be written down explicitly. All the members of this family are shown to be a solution of the equations that define the self-coupling problem, and lead in the non-interacting limit to the free description of gravitons as the irreducible spin-2 representation of the Poincaré group.

Apart from the interest of this result on its own, it has clear implications for the aspirations of the emergent gravity program. In particular, we can conclude that the existence of an excitation with the properties of a graviton in the linear spectrum of an emergent theory of gravity does not directly imply by itself that the nonlinear description of these degrees of freedom has to be given by general relativity. One would need to find an additional condition, the imposition of which guarantees the reproduction of the desired low-energy limit. Following the logic of emergence we have been advocating, this additional condition or principle would be shared by a universality class of systems. It is certainly tempting to draw parallels between this conclusion and the situation in string theory, in which it is the condition of cancellation of anomalies of local symmetries in general backgrounds that permits to recover the form of the Einstein field equations.

The non-uniqueness in the self-coupling problem of gravitons has an additional ramification that affects the discussion of the cosmological constant problem. There exists only one alternative to general relativity as a nonlinear theory of gravity that maintains the degrees of freedom that are present at the linear level. This theory is known as Weyl-transverse gravity, and is nothing but a suitable parametrization of unimodular (or tracefree) gravity.

To what extent unimodular gravity provides a resolution of the cosmological constant problem has been a disputed topic in the literature. In Chap. 3 we have used the embedding of this theory in the formalism of Weyl-transverse gravity to show explicitly that the cosmological constant term in the field equations does not receive radiative corrections, in stark difference to the situation in general relativity. We have linked this result to the symmetries of the gravitational action, showing that radiative corrections to the cosmological constant term are forbidden due to standard symmetry considerations in effective field theory. One of the most important features of this discussion is that this way of avoiding the cosmological constant problem would not have been suggested in a framework in which the geometric structure of general relativity is taken as a fundamental skeleton. Only an approach that is flexible enough to permit this structure to be emergent could reach this possibility, which therefore represents a genuine application of the emergent gravity program.

In Chap. 4 we also elaborate on a suggestion that comes from the conceptualization of the gravitational interaction as an emergent phenomenon in condensed-matter-like systems, but that now affects a different problem: the gravitational collapse of massive stars to black holes. Once again, these considerations would be hardly identified as a viable option in a framework in which general relativity is thought to be fundamental. Our analysis shows that previous qualitative considerations can be condensed in the construction of an effective geometry that describes a transition between a black-hole and a white-hole geometry in a short characteristic time scale. The overall geometry describes the elastic bounce of the collapsing distribution of matter when Planckian densities are reached, with no significant time delay as measured by observers that are kept far away from the matter structure.

This geometric construction gives concreteness to previous partial descriptions of this process, and represents a solid starting point for further studies. In particular, we have discussed how several arguments point towards the self-consistency of the overall geometry as representing a non-perturbative modification of the classical picture that is obtained in general relativity, independently of the particular ultraviolet completion that could permit this phenomenon to occur.

These developments answer a number of questions, keep silence about others, and open new ones. These unresolved issues could only be addressed by further developments. One of the most clear avenue to extend the contents of this thesis is the attempt to construct a specific condensed-matter-like model that replicates general relativity at low energies. This enterprise would probably encounter a number of notable difficulties that reside in the nonlinear character of this theory and its inextricably related geometric features. The hypothetical determination of the condition that selects the correct nonlinear behavior of the gravitational degrees of freedom would permit to establish the relevant universality class of systems that serve for this purpose.

Concerning the applications that have been considered, the possibilities are numerous. An interesting direction of study is the determination of the similarities and dissimilarities between Weyl-transverse gravity and general relativity when the quantum properties of the gravitational interaction are take into account. This study would determine whether or not these theories differ in more aspects than the renormalization properties of the cosmological constant. On the other hand, the geometric model describing a transition between a black-hole and a white-hole geometry represents only a first step that has to be conveniently expanded to become a mature proposal. From fundamental issues concerning its embedding in specific theories that go beyond general relativity, to its phenomenological implications and potential surprises for our conception of astrophysical black holes, it stands as an exciting area for future research.

Bibliography

[1] Craig Callender Nick Huggett. Why quantize gravity (or any other field for that matter)? *Philosophy of Science*, 68(3):S382–S394, 2001.

[2] C. Rovelli. *Quantum Gravity.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2004.

[3] R. Gambini and J. Pullin. *A First Course in Loop Quantum Gravity.* OUP Oxford, 2011.

[4] T. Thiemann. *Modern Canonical Quantum General Relativity.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2007.

[5] T. Jacobson. Thermodynamics of space-time: The Einstein equation of state. *Phys.Rev.Lett.*, 75:1260–1263, 1995.

[6] T. Padmanabhan. A Physical Interpretation of Gravitational Field Equations. *AIP Conf.Proc.*, 1241:93–108, 2010.

[7] T. Padmanabhan. Thermodynamical Aspects of Gravity: New insights. *Rept.Prog.Phys.*, 73:046901, 2010.

[8] G. Volovik. Emergent physics: Fermi-point scenario. *Phil. Trans. R. Soc. A*, 366:2935–2951, August 2008.

[9] T. Konopka, F. Markopoulou, and S. Severini. Quantum Graphity: A Model of emergent locality. *Phys.Rev.*, D77:104029, 2008.

[10] J. Polchinski. *String Theory: Volume 1, An Introduction to the Bosonic String.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1998.

[11] B. Zwiebach. *A First Course in String Theory.* A First Course in String Theory. Cambridge University Press, 2004.

[12] W. G. Unruh. Experimental black hole evaporation. *Phys. Rev. Lett.*, 46:1351–1353, 1981.

[13] M. Visser. Acoustic propagation in fluids: An Unexpected example of Lorentzian geometry. 1993.

[14] C. Barceló, S. Liberati, and M. Visser. Analogue Gravity. *Living Reviews in Relativity*, 8:12, December 2005.

[15] C. Barcelo, M. Visser, and S. Liberati. Einstein gravity as an emergent phenomenon? *Int.J.Mod.Phys.*, D10:799–806, 2001.

[16] J. Collins, A. Perez, D. Sudarsky, L. Urrutia, and H. Vucetich. Lorentz Invariance and Quantum Gravity: An Additional Fine-Tuning Problem? *Physical Review Letters*, 93(19):191301, November 2004.

[17] J. Collins, A. Perez, and D. Sudarsky. Lorentz invariance violation and its role in quantum gravity phenomenology. In *Approaches to quantum gravity, pp. 528-547*. Cambridge University Press, 2006.

[18] G. E. Volovik. *The Universe in a Helium Droplet*. International Series of Monographs on Physics. Clarendon Press, 2003.

[19] C. Barcelo, L. J. Garay, and G. Jannes. Quantum Non-Gravity and Stellar Collapse. *Found.Phys.*, 41:1532–1541, 2011.

[20] S. Deser. Self-interaction and gauge invariance. *General Relativity and Gravitation*, 1:9–18, March 1970.

[21] S. Deser. Gravity from self-interaction redux. *General Relativity and Gravitation*, 42:641–646, March 2010.

[22] S. Carlip. Challenges for Emergent Gravity. *Stud.Hist.Philos.Mod.Phys.*, 46:200–208, 2014.

[23] J. C. Maxwell. On Physical Lines of Force I II. *Philosophical Magazine*, XXI:161–75; 281–91; 338–348, 1861.

[24] J. C. Maxwell. On Physical Lines of Force III IV. *Philosophical Magazine*, XXIII:12–24; 85–95, 1861.

[25] G. E. Volovik. *From Quantum Hydrodynamics to Quantum Gravity*. World Scientific, 2006.

[26] G. E. Volovik. Topology of quantum vacuum. *Lecture Notes in Physics*, 870:343–383, 2013.

[27] R. P. Feynman. *Feynman lectures on gravitation*. 1996.

[28] S. N. Gupta. Gravitation and Electromagnetism. *Physical Review*, 96:1683–1685, December 1954.

[29] T. Padmanabhan. From Gravitons to Gravity: Myths and Reality. *International Journal of Modern Physics D*, 17:367–398, 2008.

[30] L. M. Butcher, M. Hobson, and A. Lasenby. Bootstrapping gravity: A consistent approach to energy-momentum self-coupling. *Physical Review D*, 80(8):084014, October 2009.

[31] S. Weinberg. The cosmological constant problem. *Rev. Mod. Phys.*, 61:1–23, Jan 1989.

[32] J. Polchinski. The Cosmological Constant and the String Landscape. In *The Quantum Structure of Space and Time*, pages 216–236, 2006.

[33] J. Martin. Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask). *Comptes Rendus Physique*, 13:566–665, 2012.

[34] C. P. Burgess. The Cosmological Constant Problem: Why it's hard to get Dark Energy from Micro-physics. In *100e Ecole d'Ete de Physique: Post-Planck Cosmology Les Houches, France, July 8-August 2, 2013*, 2013.

[35] G. E. Volovik. Vacuum energy and cosmological constant: View from condensed matter. 2001.

[36] G. E. Volovik. Cosmological constant and vacuum energy. *Annalen Phys.*, 14:165–176, 2005.

[37] S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1973.

[38] S. W. Hawking. Breakdown of Predictability in Gravitational Collapse. *Phys. Rev.*, D14:2460–2473, 1976.

[39] S. W. Hawking. Information loss in black holes. *Phys. Rev. D*, 72(8):084013, October 2005.

[40] C. Barceló, L. J. Garay, and G. Jannes. Quantum Non-Gravity and Stellar Collapse. *Foundations of Physics*, 41:1532–1541, September 2011.

[41] C.W. Misner, K.S. Thorne, and J.A. Wheeler. *Gravitation*. Number pt. 3 in Gravitation. W. H. Freeman, 1973.

[42] K. P. Sinha, C. Sivaram, and E. C. G. Sudarshan. Aether as a superfluid state of particle-antiparticle pairs. *Foundations of Physics*, 6(1):65–70, 1976.

[43] K. P. Sinha, C. Sivaram, and E. C. G. Sudarshan. The superfluid vacuum state, time-varying cosmological constant, and nonsingular cosmological models. *Foundations of Physics*, 6(6):717–726, 1976.

[44] K. P. Sinha and E. C. G. Sudarshan. The superfluid as a source of all interactions. *Foundations of Physics*, 8(11-12):823–831, 1978.

[45] M. B. Green, J. H. Schwarz, and E. Witten. *Superstring Theory: Volume 1, Introduction*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1988.

[46] S. Longair. *Theoretical Concepts in Physics: An Alternative View of Theoretical Reasoning in Physics*. Cambridge University Press, 2003.

[47] F. Dyson. *Why is Maxwell's theory so hard to understand?* James Clerk Maxwell Foundation, 1999.

[48] T. K. Simpson. *Maxwell on the Electromagnetic Field: A Guided Study*. Masterworks of discovery: guided studies of great texts in science. Rutgers University Press, 1997.

[49] D. F. Moyer. Continuum mechanics and field theory: Thomson and maxwell. *Studies in History and Philosophy of Science Part A*, 9(1):35 – 50, 1978.

[50] J. Cat. On understanding: Maxwell on the methods of illustration and scientific metaphor. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 32(3):395 – 441, 2001.

[51] D. M. Siegel. *Innovation in Maxwell's Electromagnetic Theory*. Cambridge University Press, 1992. Cambridge Books Online.

[52] T. K. Simpson. *Maxwell's Mathematical Rhetoric: Rethinking the Treatise on Electricity and Magnetism*. Green Lion Press, 2010.

[53] G. Rousseaux and E. Guyon. À propos d'une analogie entre la mécanique des fluides et l'élctromagnétisme. *Bull. Un. Phys.*, (96):107–136, 2002.

[54] G. Rousseaux. On the physical meaning of the gauge conditions of classical electromagnetism: the hydrodynamics analogue viewpoint. *Ann. Fond. Louis de Broglie*, (28):261–270, 2003.

[55] R. Brady and R. Anderson. Maxwell's fluid model of magnetism. *ArXiv e-prints*, February 2015.

[56] H. Goldstein, C. P. Poole, and J. L. Safko. *Classical Mechanics*. Addison Wesley, 2002.

[57] A. J. Leggett. A theoretical description of the new phases of liquid ^{3}He. *Reviews of Modern Physics*, 47:331–414, April 1975.

[58] D. Vollhardt and P. Woelfle. *The Superfluid Phases Of Helium 3*. Taylor & Francis, 1990.

[59] A. J. Leggett. *Quantum Liquids: Bose Condensation and Cooper Pairing in Condensed-matter Systems*. Oxford graduate texts in mathematics. OUP Oxford, 2006.

[60] L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskiĭ. *Statistical Physics*. Number pt. 2 in Course of theoretical physics. Butterworth-Heinemann, 1980.

[61] F. Schwabl. *Advanced Quantum Mechanics*. Advanced Texts in Physics Series. Springer, 2008.

[62] L. D. Landau and E. M. Lifshitz. *Mechanics*. Butterworth Heinemann. Butterworth-Heinemann, 1976.

[63] L. D. Landau and E. M. Lifshitz. *Statistical Physics*. Number v. 5. Elsevier Science, 2013.

[64] A. L. Fetter and J. D. Walecka. *Quantum Theory of Many-particle Systems*. Dover Books on Physics. Dover Publications, 2003.

[65] J. Polchinski. Effective field theory and the Fermi surface. In *Theoretical Advanced Study Institute (TASI 92): From Black Holes and Strings to Particles Boulder, Colorado, June 3-28*, 1992.

[66] L. N. Cooper. Bound electron pairs in a degenerate fermi gas. *Phys. Rev.*, 104:1189–1190, Nov 1956.

[67] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Microscopic theory of superconductivity. *Phys. Rev.*, 106:162–164, Apr 1957.

[68] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Theory of superconductivity. *Phys. Rev.*, 108:1175–1204, Dec 1957.

[69] L. D. Landau. *Collected papers of L. D. Landau*. Pergamon Press, 1965.

[70] L. P. Gor'kov. Microscopic derivation of the ginzburg-landau equations in the theory of superconductivity. *Sov. Phys. JETP*, 9:1364–1367, 1959.

[71] M. de Prato, A. Pelissetto, and E. Vicari. Normal-to-planar superfluid transition in ^{3}He. *Phys. Rev. B*, 70(21):214519, December 2004.

[72] L. Gor'kov. On the energy spectrum of superconductors. *Sov. Phys. JETP*, 7:505–508, 1958.

[73] P. Horava. Stability of Fermi Surfaces and K Theory. *Phys. Rev. Lett.*, 95(1):016405, June 2005.

[74] C. Barceló and G. Jannes. A Real Lorentz-FitzGerald Contraction. *Found. Phys.*, 38:191–199, February 2008.

[75] E. J. Mueller, T. L. Ho, M. Ueda, and G. Baym. Fragmentation of bose-einstein condensates. *Phys. Rev. A*, 74:033612, Sep 2006.

[76] B. L. Ioffe. Axial anomaly: The Modern status. *Int.J.Mod.Phys.*, A21:6249–6266, 2006.

[77] N. D. Mermin and T. L. Ho. Circulation and angular momentum in the a phase of superfluid helium-3. *Phys. Rev. Lett.*, 36:594–597, Mar 1976.

[78] T. L. Ho and N. D. Mermin. Equilibrium order parameters and chemical potentials in rotating superfluids. *Phys. Rev. B*, 21:5190–5197, Jun 1980.

[79] J. Dalibard, F. Gerbier, G. Juzeliūnas, and P. Öhberg. Colloquium: Artificial gauge potentials for neutral atoms. *Reviews of Modern Physics*, 83:1523–1543, October 2011.

[80] G. Jannes. *Emergent gravity: the BEC paradigm.* PhD thesis, 2009.

[81] L. Sindoni. Emergent models for gravity: an overview of microscopic models. *SIGMA*, 8:027, 2012.

[82] G. E. Volovik. Axial anomaly in ^{3}He-A: Simulation of Baryogenesis and generation of primordial magnetic fields in Manchester and Helsinki. *Physica B Condensed Matter*, 255:86–107, December 1998.

[83] G. Jannes and G. E. Volovik. The cosmological constant: A lesson from the effective gravity of topological weyl media. *JETP Lett.*, 96:215, 2012.

[84] C. P. Burgess. An Introduction to Effective Field Theory. *Annual Review of Nuclear and Particle Science*, 57:329–362, November 2007.

[85] R. Machleidt and D. R. Entem. Chiral effective field theory and nuclear forces. *Phys. Rep.*, 503:1–75, June 2011.

[86] Y. B. Zel'dovich. Interpretation of Electrodynamics as a Consequence of Quantum Theory. *Soviet Journal of Experimental and Theoretical Physics Letters*, 6:345, November 1967.

[87] C. Itzykson and J. B. Zuber. *Quantum Field Theory.* Dover Books on Physics. Dover Publications, 2012.

[88] C. Barceló, R. Carballo-Rubio, L. J. Garay, and G. Jannes. Electromagnetism as an emergent phenomenon: a step-by-step guide. *New Journal of Physics*, 16(12):123028, December 2014.

[89] S. Liberati, M. Visser, and S. Weinfurtner. Naturalness in an Emergent Analogue Spacetime. *Physical Review Letters*, 96(15):151301, April 2006.

[90] R. Gambini, S. Rastgoo, and J. Pullin. Small Lorentz violations in quantum gravity: do they lead to unacceptably large effects? *Classical and Quantum Gravity*, 28(15):155005, August 2011.

[91] J. Polchinski. Comment on: Small Lorentz violations in quantum gravity: do they lead to unacceptably large effects? *Classical and Quantum Gravity*, 29(8):088001, April 2012.

[92] J. B. Jiménez and A. L. Maroto. The Dark Magnetism of the Universe. *Modern Physics Letters A*, 26:3025–3039, 2011.

[93] X. G. Wen. Origin of gauge bosons from strong quantum correlations. *Phys. Rev. Lett.*, 88:011602, Dec 2001.

[94] L. Maccione, A. M. Taylor, D. M. Mattingly, and S. Liberati. Planck-scale Lorentz violation constrained by Ultra-High-Energy Cosmic Rays. *JCAP*, 4:22, April 2009.

[95] S. Liberati. Tests of Lorentz invariance: a 2013 update. *Classical and Quantum Gravity*, 30(13):133001, July 2013.

[96] M. C. Cross. A generalized Ginzburg-Landau approach to the superfluidity of helium 3. *Journal of Low Temperature Physics*, 21:525–534, December 1975.

[97] J. Dziarmaga. Low-Temperature Effective Electromagnetism in Superfluid ^{3}He-A^1. *Soviet Journal of Experimental and Theoretical Physics Letters*, 75:273–277, March 2002.

[98] A. D. Sakharov. Vacuum Quantum Fluctuations in Curved Space and the Theory of Gravitation. *Soviet Physics Doklady*, 12:1040, May 1968.

[99] M. Visser. Sakharov's Induced Gravity. *Modern Physics Letters A*, 17:977–991, 2002.

[100] A. F. Andreev and M. Y. Kagan. The superhydrodynamics of ^{3}He-A. *Journal of Experimental and Theoretical Physics*, 66:504, September 1987.

[101] J. Iliopoulos, D. V. Nanopoulos, and T. N. Tomaras. Infrared stability or anti-grandunification. *Physics Letters B*, 94:141–144, July 1980.

[102] P. Wölfle. Order-Parameter Collective Modes in ^{3}He-A. *Physical Review Letters*, 37:1279–1282, November 1976.

[103] D. Vollhardt. *Pair Correlations in Superfluid Helium 3*. Springer, June 1997.

[104] S. Weinberg and E. Witten. Limits on massless particles. *Physics Letters B*, 96:59–62, October 1980.

[105] A. S. Goldhaber and M. M. Nieto. Photon and graviton mass limits. *Reviews of Modern Physics*, 82:939–979, January 2010.

[106] E. P. Wigner. On Unitary Representations of the Inhomogeneous Lorentz Group. *Annals Math.*, 40:149–204, 1939. [Reprint: Nucl. Phys. Proc. Suppl.6,9(1989)].

[107] F. Loebbert. The Weinberg-Witten theorem on massless particles: An Essay. *Annalen Phys.*, 17:803–829, 2008.

[108] E. Álvarez, D. Blas, J. Garriga, and E. Verdaguer. Transverse Fierz Pauli symmetry. *Nuclear Physics B*, 756:148–170, November 2006.

[109] T. Ortín. *Gravity and Strings*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2007.

[110] K. A. Meissner and H. Nicolai. Conformal Symmetry and the Standard Model. *Phys. Lett.*, B648:312–317, 2007.

[111] K. Izawa. Derivative Expansion in Quantum Theory of Gravitation. *Progress of Theoretical Physics*, 93:615–619, March 1995.

[112] V. I. Ogievetsky and I. V. Polubarinov. Interacting field of spin 2 and the einstein equations. *Annals of Physics*, 35:167–208, November 1965.

[113] A. Jenkins. *Topics in theoretical particle physics and cosmology beyond the standard model*. PhD thesis, Caltech, 2006.

[114] B. Zwiebach. *A First Course in String Theory*. A First Course in String Theory. Cambridge University Press, 2004.

[115] R. M. Wald. Spin-two fields and general covariance. *Physical Review D*, 33:3613–3625, June 1986.

[116] N. Rosen. General relativity and flat space. i. *Phys. Rev.*, 57:147–150, Jan 1940.

[117] N. Rosen. General relativity and flat space. ii. *Phys. Rev.*, 57:150–153, Jan 1940.

[118] R. H. Kraichnan. Special-relativistic derivation of generally covariant gravitation theory. *Phys. Rev.*, 98:1118–1122, May 1955.

[119] E. R. Huggins. *Quantum Mechanics of the Interaction of Gravity with Electrons: Theory of a Spin-Two Field Coupled to Energy*. PhD thesis, California Institute of Technology, 1962.

[120] E. Alvarez. Can one tell Einstein's unimodular theory from Einstein's general relativity? *Journal of High Energy Physics*, 3:2, March 2005.

[121] E. Alvarez and R. Vidal. Weyl transverse gravity (WTDiff) and the cosmological constant. *Phys.Rev.*, D81:084057, 2010.

[122] E. Alvarez. The Weight of matter. *JCAP*, 1207:002, 2012.

[123] C. Aragone and Stanley Deser. Constraints on gravitationally coupled tensor fields. *Nuovo Cim.*, A3:709–720, 1971.

[124] G. Magnano and L. M. Sokolowski. Symmetry properties under arbitrary field redefinitions of the metric energy-momentum tensor in classical field theories and gravity. *Classical and Quantum Gravity*, 19:223–236, January 2002.

[125] X. Bekaert, N. Boulanger, and P. Sundell. How higher-spin gravity surpasses the spin two barrier: no-go theorems versus yes-go examples. *Rev. Mod. Phys.*, 84:987–1009, 2012.

[126] F. J. Belinfante. On the current and the density of the electric charge, the energy, the linear momentum and the angular momentum of arbitrary fields. *Physica*, 7:449–474, May 1940.

[127] L. Rosenfeld. Sur le tenseur D'Impulsion- Energie. *Mem. Acad. Roy. Belg. Sci.*, 18:1–30, 1940.

[128] J. J. van der Bij, H. van Dam, and Y. J. Ng. The Exchange of Massless Spin Two Particles. *Physica*, 116A:307–320, 1982.

[129] R.M. Wald. *General Relativity*. University of Chicago Press, 2010.

[130] T. Padmanabhan. *Gravitation: Foundations and Frontiers*. Gravitation: Foundations and Frontiers. Cambridge University Press, 2010.

[131] M. Fierz and W. Pauli. On relativistic wave equations for particles of arbitrary spin in an electromagnetic field. *Proc. Roy. Soc. Lond.*, A173:211–232, 1939.

[132] D. G. Boulware and S. Deser. Classical general relativity derived from quantum gravity. *Annals of Physics*, 89:193–240, January 1975.

[133] W. G. Unruh. Unimodular theory of canonical quantum gravity. *Phys. Rev. D*, 40:1048–1052, August 1989.

[134] G. F. R. Ellis, H. van Elst, J. Murugan, and J.-P. Uzan. On the trace-free Einstein equations as a viable alternative to general relativity. *Classical and Quantum Gravity*, 28(22):225007, November 2011.

[135] G. F. R. Ellis. The Trace-Free Einstein Equations and inflation. *Gen.Rel.Grav.*, 46:1619, 2014.

[136] J. F. Donoghue. General relativity as an effective field theory: The leading quantum corrections. *Phys. Rev. D*, 50:3874–3888, September 1994.

[137] C. P. Burgess. Quantum Gravity in Everyday Life: General Relativity as an Effective Field Theory. *Living Reviews in Relativity*, 7:5, April 2004.

[138] N. E. J. Bjerrum-Bohr, J. F. Donoghue, B. R. Holstein, L. Planté, and P. Vanhove. Bending of Light in Quantum Gravity. *Physical Review Letters*, 114(6):061301, February 2015.

[139] J. F. Donoghue and B. R. Holstein. Low energy theorems of quantum gravity from effective field theory. *Journal of Physics G Nuclear Physics*, 42(10):103102, October 2015.

[140] P. W. Milonni. *The Quantum Vacuum: An Introduction to Quantum Electrodynamics*. Academic Press, 1994.

[141] P. Kopietz, L. Bartosch, and F. Schütz. *Introduction to the Functional Renormalization Group*. Introduction to the Functional Renormalization Group. Springer, 2010.

[142] D. V. Vassilevich. Heat kernel expansion: User's manual. *Phys. Rept.*, 388:279–360, 2003.

[143] V. Kagramanova, J. Kunz, and C. Lammerzahl. Solar system effects in Schwarzschild-de Sitter spacetime. *Phys. Lett.*, B634:465–470, 2006.

[144] Antonio Padilla and Ippocratis D. Saltas. A note on classical and quantum unimodular gravity. *Eur. Phys. J.*, C75(11):561, 2015.

[145] R. A. Bertlmann. *Anomalies in Quantum Field Theory*. International Series of Monographs on Physics. Clarendon Press, 2000.

[146] R. M. Wald. *General Relativity*. University of Chicago Press, 2010.

[147] K. Fujikawa and H. Suzuki. *Path Integrals and Quantum Anomalies*. International Series of Monographs on Physics. Clarendon Press, 2004.

[148] D. M. Capper and M. J. Duff. Trace anomalies in dimensional regularization. *Nuovo Cim.*, A23:173–183, 1974.

[149] M. J. Duff. Observations on Conformal Anomalies. *Nucl. Phys.*, B125:334, 1977.

[150] D. Blas. *Aspects of Infrared Modifications of Gravity*. PhD thesis, Barcelona U., 2008.

[151] E. Alvarez and M. Herrero-Valea. No Conformal Anomaly in Unimodular Gravity. *Phys. Rev.*, D87:084054, 2013.

[152] K. Fujikawa. Path Integral for Gauge Theories with Fermions. *Phys. Rev.*, D21:2848, 1980. [Erratum: Phys. Rev.D22,1499(1980)].

[153] V. Mukhanov and S. Winitzki. *Introduction to Quantum Effects in Gravity*. Cambridge University Press, 2007.

[154] E. Mottola. Functional integration over geometries. *J. Math. Phys.*, 36:2470–2511, 1995.

[155] E. lvarez, S. Gonzlez-Martn, M. Herrero-Valea, and C. P. Martn. Unimodular Gravity Redux. *Phys. Rev.*, D92(6):061502, 2015.

[156] E. Álvarez, S. González-Martín, M. Herrero-Valea, and C. P. Martín. Quantum Corrections to Unimodular Gravity. *JHEP*, 08:078, 2015.

[157] L. Smolin. Quantization of unimodular gravity and the cosmological constant problems. *Physical Review D*, 80(8):084003, October 2009.

[158] T. Padmanabhan and H. Padmanabhan. Cosmological Constant from the Emergent Gravity Perspective. *Int. J. Mod. Phys.*, D23(6):1430011, 2014.

[159] N. Kaloper and A. Padilla. Sequestering the Standard Model Vacuum Energy. *Phys. Rev. Lett.*, 112(9):091304, 2014.

[160] A. Padilla. Lectures on the Cosmological Constant Problem. 2015.

[161] G. E. Volovik. *The Universe in a Helium Droplet*. International Series of Monographs on Physics. OUP Oxford, May 2009.

[162] R. Schödel, T. Ott, R. Genzel, R. Hofmann, M. Lehnert, A. Eckart, N. Mouawad, T. Alexander, M. J. Reid, R. Lenzen, M. Hartung, F. Lacombe, D. Rouan, E. Gendron, G. Rousset, A.-M. Lagrange, W. Brandner, N. Ageorges, C. Lidman, A. F. M. Moorwood, J. Spyromilio, N. Hubin, and K. M. Menten. A star in a 15.2-year orbit around the supermassive black hole at the centre of the Milky Way. *Nature*, 419:694–696, October 2002.

[163] R. Schödel, T. Ott, R. Genzel, A. Eckart, N. Mouawad, and T. Alexander. Stellar Dynamics in the Central Arcsecond of Our Galaxy. *Astrophys. J.*, 596:1015–1034, October 2003.

[164] A. M. Ghez, S. Salim, S. D. Hornstein, A. Tanner, J. R. Lu, M. Morris, E. E. Becklin, and G. Duchêne. Stellar Orbits around the Galactic Center Black Hole. *Astrophys. J.*, 620:744–757, February 2005.

[165] M. A. Abramowicz, W. Kluźniak, and J.-P. Lasota. No observational proof of the black-hole event-horizon. *Astron. Astrophys.*, 396:L31–L34, December 2002.

[166] C. L. Fryer and K. C. B. New. Gravitational waves from gravitational collapse. *Living Reviews in Relativity*, 14(1), 2011.

[167] S. W. Hawking. Particle Creation by Black Holes. *Commun. Math. Phys.*, 43:199–220, 1975.

[168] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully. Black holes: complementarity or firewalls? *Journal of High Energy Physics*, 2:62, February 2013.

[169] J. R. Oppenheimer and G. M. Volkoff. On Massive Neutron Cores. *Physical Review*, 55:374–381, February 1939.

[170] C. E. Rhoades and R. Ruffini. Maximum Mass of a Neutron Star. *Physical Review Letters*, 32:324–327, February 1974.

[171] I. Bombaci. The maximum mass of a neutron star. *Astronomy and Astrophysics*, 305:871, January 1996.

[172] J. R. Oppenheimer and H. Snyder. On continued gravitational contraction. *Phys. Rev.*, 56:455–459, 1939.

[173] K. Martel and E. Poisson. Regular coordinate systems for Schwarzschild and other spherical space-times. *Am. J. Phys.*, 69:476–480, 2001.

[174] P. T. Chruściel, J. L. Costa, and M. Heusler. Stationary Black Holes: Uniqueness and Beyond. *Living Reviews in Relativity*, 15:7, May 2012.

[175] R. Penrose. *Cycles of Time: An Extraordinary New View of the Universe*. Bodley Head, 2010.

[176] J. M. M. Senovilla and D. Garfinkle. The 1965 Penrose singularity theorem. 2014.

[177] R. Penrose. Gravitational Collapse and Space-Time Singularities. *Physical Review Letters*, 14:57–59, January 1965.

[178] I. M. Khalatnikov and E. M. Lifshitz. General Cosmological Solution of the Gravitational Equations with a Singularity in Time. *Physical Review Letters*, 24:76–79, January 1970.

[179] V. A. Belinskiĭ, E. M. Lifshitz, and I. M. Khalatnikov. Reviews of Topical Problems: Oscillatory Approach to the Singular Point in Relativistic Cosmology. *Soviet Physics Uspekhi*, 13:745–765, June 1971.

[180] B. K. Berger. Numerical Approaches to Spacetime Singularities. *Living Reviews in Relativity*, 5:1, January 2002.

[181] A. Ashtekar, A. Henderson, and D. Sloan. Hamiltonian formulation of the Belinskii-Khalatnikov-Lifshitz conjecture. *Phys. Rev. D*, 83(8):084024, April 2011.

[182] Jun-Qi Guo, Daoyan Wang, and Andrei V. Frolov. Spherical collapse in $f(\mathrm{R})$ gravity and the Belinskii-Khalatnikov-Lifshitz conjecture. *Phys. Rev.*, D90(2):024017, 2014.

[183] Jun-Qi Guo and Pankaj S. Joshi. Interior dynamics of neutral and charged black holes. *Phys. Rev.*, D92(6):064013, 2015.

[184] W. G. Unruh. Notes on black-hole evaporation. *Phys. Rev. D*, 14:870–892, August 1976.

[185] S. W. Hawking. Black hole explosions. *Nature*, 248:30–31, 1974.

[186] S. W. Hawking. Information Preservation and Weather Forecasting for Black Holes. *ArXiv e-prints*, January 2014.

[187] C. Barceló, M. Visser, and D. V. Ahluwalia. Twilight for the Energy Conditions? *International Journal of Modern Physics D*, 11:1553–1560, 2002.

[188] L.H. Ford. The Classical singularity theorems and their quantum loop holes. *Int.J.Theor.Phys.*, 42:1219–1227, 2003.

[189] R. H. Fowler. On dense matter. *Monthly Not. Roy. Astr. Soc.*, 87:114–122, December 1926.

[190] S. Chandrasekhar. The Maximum Mass of Ideal White Dwarfs. *Astrophys. J.*, 74:81, July 1931.

[191] L. Parker and S. A. Fulling. Quantized Matter Fields and the Avoidance of Singularities in General Relativity. *Phys. Rev. D*, 7:2357–2374, April 1973.

[192] M. Novello and S. E. P. Bergliaffa. Bouncing cosmologies. *Phys. Rep.*, 463:127–213, July 2008.

[193] A. Ashtekar. Gravity and the quantum. *New Journal of Physics*, 7:198, September 2005.

[194] K. Banerjee, G. Calcagni, and M. Martín-Benito. Introduction to Loop Quantum Cosmology. *SIGMA*, 8:16, March 2012.

[195] K. V. Kuchar and M. P. Ryan, Jr. Is minisuperspace quantization valid?: Taub in mixmaster. *Phys. Rev. D*, 40:3982–3996, December 1989.

[196] J. F. Barbero G. and E. J. S. Villaseñor. Quantization of Midisuperspace Models. *Living Reviews in Relativity*, 13:6, October 2010.

[197] M. Bojowald. Absence of a Singularity in Loop Quantum Cosmology. *Physical Review Letters*, 86:5227, June 2001.

[198] V. Taveras. Corrections to the Friedmann equations from loop quantum gravity for a universe with a free scalar field. *Phys. Rev. D*, 78(6):064072, September 2008.

[199] B. Broda. One-Loop Quantum Gravity Repulsion in the Early Universe. *Physical Review Letters*, 106(10):101303, March 2011.

[200] J. Abedi and H. Arfaei. Obstruction of black hole singularity by quantum field theory effects. *ArXiv e-prints*, June 2015.

[201] A. D. Sakharov. The Initial Stage of an Expanding Universe and the Appearance of a Nonuniform Distribution of Matter. *Soviet Journal of Experimental and Theoretical Physics*, 22:241, January 1966.

[202] E. B. Gliner. Algebraic Properties of the Energy-momentum Tensor and Vacuum-like States o$^+$ Matter. *Soviet Journal of Experimental and Theoretical Physics*, 22:378, February 1966.

[203] J. M. Bardeen. Non-singular general-relativistic gravitational collapse. In *Proc. Int. Conf. GR5, Tbilisi*, page 174, 1968.

[204] A. Borde. Open and closed universes, initial singularities and inflation. *Phys. Rev.*, D50:3692–3702, 1994.

[205] A. Borde. Regular black holes and topology change. *Phys. Rev.*, D55:7615–7617, 1997.

[206] I. Dymnikova. Vacuum nonsingular black hole. *Gen. Rel. Grav.*, 24:235–242, 1992.

[207] M. Mars, M. M. Martín-Prats, and J. M. M. Senovilla. Models of regular schwarzschild black holes satisfying weak energy conditions. *Classical and Quantum Gravity*, 13(5):L51, 1996.

[208] E. Ayon-Beato and A. Garcia. Regular black hole in general relativity coupled to nonlinear electrodynamics. *Phys. Rev. Lett.*, 80:5056–5059, 1998.

[209] K. A. Bronnikov. Regular magnetic black holes and monopoles from nonlinear electrodynamics. *Phys. Rev. D*, 63(4):044005, February 2001.

[210] K. A. Bronnikov. Spherically symmetric false vacuum: No-go theorems and global structure. *Phys. Rev. D*, 64(6):064013, September 2001.

[211] S. A. Hayward. Formation and Evaporation of Nonsingular Black Holes. *Physical Review Letters*, 96(3):031103, January 2006.

[212] G. J. Olmo, D. Rubiera-Garcia, and H. Sanchis-Alepuz. Geonic black holes and remnants in Eddington-inspired Born-Infeld gravity. *European Physical Journal C*, 74:2804, March 2014.

[213] G. J. Olmo and D. Rubiera-Garcia. Nonsingular Black Holes in $f(R)$ Theories. *ArXiv e-prints*, September 2015.

[214] A. Ashtekar and M. Bojowald. Quantum geometry and the Schwarzschild singularity. *Class. Quant. Grav.*, 23:391–411, 2006.

[215] A. Ashtekar, V. Taveras, and M. Varadarajan. Information is Not Lost in the Evaporation of 2D Black Holes. *Physical Review Letters*, 100(21):211302, May 2008.

[216] S. Hossenfelder and L. Smolin. Conservative solutions to the black hole information problem. *Phys. Rev. D*, 81(6):064009, March 2010.

[217] H. Kawai, Y. Matsuo, and Y. Yokokura. A Self-consistent Model of the Black Hole Evaporation. *Int.J.Mod.Phys.*, A28:1350050, 2013.

[218] H. Kawai and Y. Yokokura. Phenomenological Description of the Interior of the Schwarzschild Black Hole. 2014.

[219] R. Torres. Singularity-free gravitational collapse and asymptotic safety. *Phys.Lett.*, B733:21–24, 2014.

[220] V. P. Frolov. Do Black Holes Exist? In *18th International Seminar on High Energy Physics (Quarks 2014) Suzdal, Russia, June 2-8, 2014*, 2014.

[221] M. Bojowald. Information loss, made worse by quantum gravity? *Front. Phys.*, 3:33, 2015.

[222] R. Gambini, J. Olmedo, and J. Pullin. Quantum black holes in loop quantum gravity. *Classical and Quantum Gravity*, 31(9):095009, May 2014.

[223] R. Torres and F. Fayos. On the quantum corrected gravitational collapse. 2015.

[224] C. Rovelli and F. Vidotto. Planck stars. *International Journal of Modern Physics D*, 23:42026, December 2014.

[225] A. Barrau and C. Rovelli. Planck star phenomenology. *Phys.Lett.*, B739:405, 2014.

[226] C. Bambi, D. Malafarina, and L. Modesto. Non-singular quantum-inspired gravitational collapse. *Phys. Rev. D*, 88(4):044009, August 2013.

[227] Y. Liu, D. Malafarina, L. Modesto, and C. Bambi. Singularity avoidance in quantum-inspired inhomogeneous dust collapse. *Phys. Rev. D*, 90(4):044040, August 2014.

[228] Y. Zhang, Y. Zhu, L. Modesto, and C. Bambi. Can static regular black holes form from gravitational collapse? *European Physical Journal C*, 75:96, February 2015.

[229] S. K. Modak, L. Ortíz, I. Peña, and D. Sudarsky. Nonparadoxical loss of information in black hole evaporation in a quantum collapse model. *Phys. Rev. D*, 91(12):124009, June 2015.

[230] S. K. Modak, L. Ortíz, I. Peña, and D. Sudarsky. Black hole evaporation: information loss but no paradox. *General Relativity and Gravitation*, 47:120, October 2015.

[231] S. W. Hawking. The Information Paradox for Black Holes. *ArXiv e-prints*, September 2015.

[232] G. 't Hooft. Diagonalizing the Black Hole Information Retrieval Process. *ArXiv e-prints*, September 2015.

[233] V. P. Frolov and G. A. Vilkovisky. Quantum gravity removes classical singularities and shortens the life of black holes. In *Grossmann Mtg.1979:0455*, page 0455, 1979.

[234] V. P. Frolov and G. A. Vilkovisky. Spherically symmetric collapse in quantum gravity. *Physics Letters B*, 106:307–313, November 1981.

[235] T. A. Roman and P. G. Bergmann. Stellar collapse without singularities? *Phys. Rev.*, D28:1265–1277, 1983.

[236] A. Ashtekar and M. Bojowald. Black hole evaporation: A Paradigm. *Class. Quant. Grav.*, 22:3349–3362, 2005.

[237] C. Bambi, D. Malafarina, and L. Modesto. Terminating black holes in asymptotically free quantum gravity. *Eur. Phys. J.*, C74:2767, 2014.

[238] J. M. Bardeen. Black hole evaporation without an event horizon. *ArXiv e-prints*, June 2014.

[239] H. Kawai and Y. Yokokura. Interior of Black Holes and Information Recovery. *ArXiv e-prints*, September 2015.

[240] M. Visser. Physical observability of horizons. *Phys. Rev. D*, 90(12):127502, December 2014.

[241] D. Samtleben, S. Staggs, and B. Winstein. The Cosmic Microwave Background for Pedestrians: A Review for Particle and Nuclear Physicists. *Annual Review of Nuclear and Particle Science*, 57:245–283, November 2007.

[242] B. J. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama. New cosmological constraints on primordial black holes. *Phys. Rev.*, D81:104019, 2010.

[243] C. R. Stephens, G. 't Hooft, and B. F. Whiting. Black hole evaporation without information loss. *Class. Quant. Grav.*, 11:621–648, 1994.

[244] C. Barceló, S. Liberati, S. Sonego, and M. Visser. Hawking-like radiation does not require a trapped region. *Phys. Rev. Lett.*, 97:171301, 2006.

[245] L. C. Barbado, C. Barceló, L. J. Garay, and G. Jannes. The Trans-Planckian problem as a guiding principle. *JHEP*, 1111:112, 2011.

[246] J. Steinhauer. Observation of self-amplifying Hawking radiation in an analogue black-hole laser. *Nature Physics*, 10:864–869, November 2014.

[247] J. Steinhauer. Observation of thermal Hawking radiation and its entanglement in an analogue black hole. *ArXiv e-prints*, October 2015.

[248] W. G. Unruh. Has Hawking Radiation Been Measured? *Foundations of Physics*, 44:532–545, May 2014.

[249] H. F. M. Goenner. What kind of science is cosmology? *Annalen der Physik*, 522:389–418, June 2010.

[250] L. Susskind, L. Thorlacius, and J. Uglum. The Stretched horizon and black hole complementarity. *Phys. Rev.*, D48:3743–3761, 1993.

[251] S. D. Mathur. The Fuzzball proposal for black holes: An Elementary review. *Fortsch. Phys.*, 53:793–827, 2005.

[252] S. D. Mathur. The information paradox: a pedagogical introduction. *Classical and Quantum Gravity*, 26(22):224001, November 2009.

[253] G. Dvali and C. Gomez. Black Holes as Critical Point of Quantum Phase Transition. *Eur. Phys. J.*, C74:2752, 2014.

[254] G. Dvali, C. Gomez, and D. Lüst. Classical Limit of Black Hole Quantum N-Portrait and BMS Symmetry. *ArXiv e-prints*, September 2015.

[255] S. B. Giddings. Black holes and massive remnants. *Phys. Rev.*, D46:1347–1352, 1992.

[256] S. B. Giddings. Comments on information loss and remnants. *Phys. Rev. D*, 49:4078–4088, April 1994.

[257] H. M. Haggard and C. Rovelli. Quantum-gravity effects outside the horizon spark black to white hole tunneling. *Phys. Rev.*, D92(10):104020, 2015.

[258] C. Barceló, R. Carballo-Rubio, and L. J. Garay. Black holes turn white fast, otherwise stay black: no half measures. 2015.

[259] S. E. Woosley, A. Heger, and T. A. Weaver. The evolution and explosion of massive stars. *Reviews of Modern Physics*, 74:1015–1071, November 2002.

[260] C. Barceló, S. Liberati, S. Sonego, and M. Visser. Fate of gravitational collapse in semiclassical gravity. *Phys. Rev.*, D77:044032, 2008.

[261] Y. Kanai, M. Siino, and A. Hosoya. Gravitational collapse in Painlevé-Gullstrand coordinates. *Prog. Theor. Phys.*, 125:1053–1065, 2011.

[262] G. Grubb. *Distributions and Operators*. Graduate Texts in Mathematics. Springer, 2009.

[263] D. M. Eardley. Death of white holes in the early universe. *Phys. Rev. Lett.*, 33:442–444, 1974.

[264] V. Frolov and I. D. Novikov. *Black Hole Physics: Basic Concepts and New Developments*. Fundamental Theories of Physics. Springer Netherlands, 1998.

[265] C. Barrabès, P. R. Brady, and E. Poisson. Death of white holes. *Phys. Rev. D*, 47:2383–2387, 1993.

[266] A. Ori and E. Poisson. Death of cosmological white holes. *Phys. Rev. D*, 50:6150–6157, 1994.

[267] S. K. Blau and A. H. Guth. The stability of the white hole horizon. *Essay written for the Gravity Research Foundation 1989 Awards for Essays on Gravitation*, 1989.

[268] S. K. Blau. Dray-'t Hooft geometries and the death of white holes. *Phys. Rev. D*, 39:2901–2903, 1989.

[269] P. Szekeres. Global Description of Spherical Collapsing and Expanding Dust Clouds. *Nuovo Cim.*, B17:187–195, 1973.

[270] K. Lake and R. C. Roeder. Blue-shift surfaces and the stability of white holes. *Lettere Al Nuovo Cimento*, 16(1):17–21, 1976.

[271] Y. B. Zeldovich, I. D. Novikov, and A. A. Starobinsky. Quantum effects in white holes. *Zh. Eksp. Teor. Fiz.*, 66:1897–1910, 1974.

[272] P. Hájícek and C. Kiefer. Singularity avoidance by collapsing shells in quantum gravity. *Int. J. Mod. Phys.*, D10:775–780, 2001.

[273] P. Hájícek. Quantum Theory of Gravitational Collapse (Lecture Notes on Quantum Conchology). In D. Giulini, C. Kiefer, and C. Laemmerzahl, editors, *Quantum Gravity: From Theory to Experimental Search*, volume 631 of *Lecture Notes in Physics, Berlin Springer Verlag*, pages 255–299, 2003.

[274] M. Ambrus. *How long does it take until a quantum system reemerges after a gravitational collapse?* PhD thesis, Bern U., 2004.

[275] M. Ambrus and P. Hájícek. Quantum superposition principle and gravitational collapse: Scattering times for spherical shells. *Phys. Rev. D*, 72(6):064025, September 2005.

[276] J. Hartle and T. Hertog. Quantum transitions between classical histories. *Phys. Rev. D*, 92(6):063509, September 2015.

[277] M. Visser, C. Barceló, S. Liberati, and S. Sonego. Small, dark, and heavy: But is it a black hole? *ArXiv e-prints*, February 2009.

[278] R. Narayan and J. S. Heyl. On the Lack of Type I X-Ray Bursts in Black Hole X-Ray Binaries: Evidence for the Event Horizon? *Astrophys. J.*, 574:L139–L142, August 2002.

[279] R. Narayan and J. E. McClintock. Advection-dominated accretion and the black hole event horizon. *New Astron. Rev.*, 51:733–751, May 2008.

[280] A. E. Broderick and R. Narayan. Where are all the gravastars? Limits upon the gravastar model from accreting black holes. *Classical and Quantum Gravity*, 24:659–666, February 2007.

[281] F. H. Vincent, Z. Meliani, P. Grandclement, E. Gourgoulhon, and O. Straub. Imaging a boson star at the Galactic center. *ArXiv e-prints*, October 2015.

[282] T. Piran. The physics of gamma-ray bursts. *Reviews of Modern Physics*, 76:1143–1210, October 2004.

[283] A. MacFadyen and S. E. Woosley. Collapsars: Gamma-ray bursts and explosions in 'failed supernovae'. *Astrophys. J.*, 524:262, 1999.

[284] M. Mobberley. *Cataclysmic Cosmic Events and How to Observe Them*. Astronomers' Observing Guides. Springer, 2009.

[285] V. P. Frolov and A. Zelnikov. *Introduction to Black Hole Physics*. OUP Oxford, 2011.

[286] D. Christodoulou and R. Ruffini. Reversible Transformations of a Charged Black Hole. *Phys. Rev. D*, 4:3552–3555, December 1971.

[287] T. Damour and R. Ruffini. Quantum Electrodynamical Effects in Kerr-Newmann Geometries. *Physical Review Letters*, 35:463–466, August 1975.

[288] A. Buonanno and B. S. Sathyaprakash. Sources of Gravitational Waves: Theory and Observations. *ArXiv e-prints*, October 2014.