Imperial College London

Numerical Investigations of the Free Energy for a Scalar Field Theory on Deformed Spheres

Candidate: Emanuele Panella

Supervisor: Toby Wiseman

Submitted in partial fulfilment of the requirements for the degree of Master of Science of Imperial College London.

Department of Physics

September, 2021

Abstract

Casimir energy is object of growing interest in physics, mainly due to recent technological advances in condensed matter theory that might allow the study of quantum field theories in controlled lab conditions, but also for its possible applications in the gravity sector. Recent results seem to suggest that a $(2+1)$-dimensional non-interacting scalar field has its Casimir energy maximised, at least locally, by the round sphere with respect to arbitrarily big axisymmetric deformations. After reviewing the theoretical background surrounding the topic of Casimir energy, we generalise such results, by extending the analysis to a more general family of distortions. We deploy spectral numerical methods to calculate the Casimir energy of a free scalar field living on a $\mathbb{R} \times \Sigma$ static spacetime, where Σ is a spatial manifold with S^2 topology. We investigate the free energy for various geometries Σ, finding that the scalar disfavours the spherical configuration. We further systematically explore the effect that non-minimal coupling to the background geometry has on the free energy, having it received little attention in previous works. We find that this extra contribution manifests itself as a secondary peak in the heat kernel profile, and observe the system to transition to a curvature dominated regime around the $\xi = 1/6$ value for the coupling. Suggestively, this is also the value of the coupling for which the massless scalar becomes conformally invariant. The effects of non-zero temperature and mass on the computed free energy are also considered.

To my family, with deep love and gratitude.

Acknowledgements

This dissertation would have not been possible without the careful supervision and support by Professor Toby Wiseman. I've always looked up to him as a teacher, and it has been inspiring and a pleasure to work with him on this project. He has put me in the right position to make original research, motivating and guiding me whenever I hit a rough patch.

I will never be tired of acknowledging my family. I know the huge emotional sacrifices that my parents had to do to let me achieve my dreams and ambitions, setting aside their own expectations. My father Alessandro's moral fiber and my mother Cristina's determination made me the person I am today. I also want to send a kiss to my little brother Edoardo, who is growing and maturing miles away from me. I know he will continue to prove his worth in the years to come, in his personal and professional life. A heartfelt hug to my grandparents, aunts and uncles. The love of my incredibly close family gave me the strength to live these years away with serenity, and words are not enough to say 'thank you' for all that they have done for me. A special thought goes to nonno Sandro, who nurtured my passion for science and would have loved to see me graduate.

I would also like to thank my friends, all of them, for being the spice that enriches my life. Those at home, for having gifted me with some wonderful teenager years and for making me feel as I have never left Rome. To point fingers, thank you Barto from the bottom of my heart for keeping up with me. Most say we are like a disfunctional married couple, and I love it. Thank you Luca and Gianluca, for completing the famed Riot Club and for having laughed with me our way out of high school. I also want to thank Bianca, as she has been an important presence in my life, gifting me with memories that I hold very dearly. To those in London, thank you for being the most absurdly assorted group, sharing my passions and mindset. These four years have been really something, and you never made me feel alone. To Tancredi, Pietro and Emma: we learnt from each other throughout this degree. Looking back to those four kids sitting at the back of Blackett's lecture theatres, full of hopes and curiosity, I am not surprised that we managed to get where we are today.

Finally, I'd also like to spend two words to thank Dr Mitesh Patel. For the trust he showed in my abilities and for the passion he conveyed for his work, he is probably one of the main reasons I ended up embarking on a PhD after all.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivations and Objectives

It had been long believed that matter in the Universe had to obey some form of a 'positive energy' constraint in the context of general relativity, with many important results assuming the so called null energy condition (NEC). The discovery of the Casimir effect turned the tables, uncovering the fact that quantum effects can actually cause matter field to develop a negative zero point-energy, breaking these commonly assumed energy conditions. The finding put the phenomenological implications that vacuum energy can have under the spotlight, especially in the gravity sector. It has indeed been suggested that a negative Casimir energy could play a role in sourcing exotic solutions to Einstein's equations, such as wormholes [1] [2] [3]. More notably, the vacuum energy of QFTs has been proposed as a natural candidate for Dark Energy, but it has lead to disastrous predictions and as of now there is no consensus that this could in fact be the right approach [4]. On the other hand, the effective description of electrons on 2D materials such as graphene in terms of massless Dirac fermions, together with the technological development of optical lattices, might enable in the foreseeable future the long-awaited opportunity to carry out experiments that allow the study of quantum fields on curved spacetime in controlled conditions [5] [6] [7].

It is the much anticipated role of vacuum energy in the research that will take place in the years to come that motivates our study. In the present analysis, we numerically study the Casimir energy, and its thermal generalisation (the Helmholtz

free energy), for a free minimally and non-minimally coupled scalar field on a (2+1)-dimensional static spacetime via spectral methods. In particular, we consider the spatial geometry Σ to be topologically equivalent to the unit sphere S^2, studying how the free energy changes as Σ is non-axisymmetrically deformed away from the round sphere. In order to perform the computations described, we resort to the heat kernel representation of the free energy. We gather new results that reinforce the conjecture that the round sphere maximises the free energy of quantum field theories, a proposal that sparked from a recent series of papers [8] [9] [10] [11]. We crucially extend the findings in the literature to encompass more general deformations. Moreover, we characterise in detail the effects that non-minimal coupling has on the Casimir energy of the free scalar field, something that did not receive much attention in previous studies.

1.2 Overview

Chapter 2 opens the dissertation with a review of the theory surrounding the topic of Casimir energy. After some mentions on the history of the Casimir effect, we dive into an overview of Casimir's original derivation, discussing the physical significance of the divergences that arise in the calculations along with an outline of dimensional regularisation. We also describe non-zero mass and temperature corrections to the celebrated result. After a brief summary of how to couple the classical Klein Gordon field to gravity, we go over the canonical quantisation of a scalar field in fixed curved background, highlighting the implications that this has for the concept of particle. We then discuss Casimir energy on curved spacetime. In doing so, we introduce the heat kernel method to compute the vacuum energy, and mention the advantages that this approach brings when the problem has to be translated into a computational task. At this point, we review attempts to study the backreaction of the vacuum energy on the spacetime geometry, along with the limitations to the most commonly used semi-classical limit of Einstein's equations. We then present solutions to this semi-classical version of general relativity, discussing in particular the role that Casimir energy is proposed to have in sourcing wormhole solutions. We conclude the first chapter showing how the vacuum energy of a quantum field could cause instability of certain geometrical configurations for 2D layers of graphene-like materials.

The third chapter sets the scene for the numerical analysis. We first introduce the physical setup, describing the Laplacian and the Ricci scalar on a spatial foliation with geometry conformal to the round sphere. We successively show in detail how to obtain the free energy of the quantum field for our problem from the heat kernel, illustrating why taking the difference with respect to a reference configuration can regulate the Helmholtz energy of our theory. We further discuss the heat kernel

asymptotics, and how they guarantee the positivity of the kernel in the small and large kernel time limits. We conclude the chapter by reviewing recent results for free Klein-Gordon and Dirac fields. These have numerically shown the sphere to maximise the free energy with respect to big axisymmetric deformations of Σ and analytically proved that the same is true for small deformations of any form. We also outline a recent paper that proved the free energy to be maximised by S^2 with respect to perturbations of any size and form in the case of a strongly interacting conformal field theory, exploiting the AdS/CFT correspondence.

We present our numerical methods in Chapter 4, where we briefly review the advantages of spectral methods with respect to equispaced finite difference methods. After discussing how to implement the boundary conditions, we describe how to translate our physical problem into a numerical task. We then benchmark our method by analysing the spectrum of the Laplacian on both a 1D interval and on the round sphere. We further study how changing shape and density of the 2D computational grid affects accuracy of the numerically computed spectra.

Chapter 5 contains the novel results obtained in our study. We begin by mentioning the issue of convergence of the numerically evaluated kernel for small kernel time. We then validate the numerical method by showing that the computed kernel curves approach the perturbative predictions for small deformations of the spatial geometry. We successively present the results for the renormalised free energy of both the minimally and non-minimally coupled scalar. We further investigate how the vacuum energy is affected by thermal effects and by adding a mass to the previously massles K-G field. The significance of the results obtained is then highlighted in the last chapter, where we also mention possible interesting extensions of this study.

1.3 Units

Throughout the discussion, we adopt Planck units $G = c = \hbar = k_b = 1$, in order to simplify the expressions.

Chapter 2

Theoretical Background

We now present the theoretical background that surrounds our study. We first review the topic of Casimir energy in flat space, discussing its physical significance and outlining different approaches to compute it. We then cover the subject of QFTs in (fixed) curved spacetime, focusing on the non-interacting minimally and non-minimally coupled scalar. We successively introduce the heat kernel, and how it can be used to compute Casimir energy in curved background geometry. From here, we take a short detour on what happens when we let this vaucum energy backreact on the spacetime by including it as a source in Einstein equations, discussing its proposed role in wormhole solutions. We then introduce some practical applications of our research, reviewing the effective QFT descriptions of electrons on a 2D layer of graphene and how vacuum energy could affect the stability of certain geometrical configurations for graphene-like materials.

2.1 Casimir Energy in Flat Space

2.1.1 A (Brief) History of the Casimir Effect

The realisation that the zero point energy of quantum fields can lead to interesting physical phenomena traces back to the formulation of the famous Casimir effect by Hendrik Casimir in 1948. Casimir first stumbled upon it while he was trying to give a consistent theoretical explanation to the form of the interaction between

non-polar molecules. Thanks to the (at the time) recently developed quantum theory, the German physicist Fritz London had managed to successfully compute the attraction between non-polar molecules by 1930, but experimental results showed that for very large distances London's prediction failed, as the potential energy fell quicker than the theorised R^{-6} behaviour (where R is the intermolecular separation). Some twenty years later, as QED was formulated and refined, Casimir and the Dutch scientist Dirk Polder had all the tools to address the problem. After the long calculations needed to evaluate the forces on an atom in an electromagnetic cavity with perfectly conducting walls using the full QED machinery, they managed to successfully predict the R^{-7} trend for large distances that had been observed in experiments. After reading the paper by Casimir and Polder, Niels Bohr realised that the result was much more profound, and tipped Casimir in the right direction by suggesting that the zero point energy of the Maxwell field could be the origin of such attraction. Inspired by Bohr's remark, Casimir re-derived his results, this time in just 2 pages of calculations, by considering the vacuum energy of the EM field only. He then continued on the same route, and, by the end of 1948, published the famous paper in which he predicted the attraction between two perfectly conducting plates. The Casimir effect was first (accurately) experimentally measured in 1997 [12] [13].

2.1.2 Mode Summation

We now follow Casimir's original mode summation treatment for the calculation of the force between parallel conducting plates [14], outlining alternative approaches that lead to the same result.

Casimir considers a cubic cavity bounded by perfectly conducting walls of side L, with a further conducting square plate of side L a distance a away from the wall given by the $x - y$ plane. The role of the cube is to regulate our space, making it finite. Indeed, for any non zero vacuum energy density, we are bound to have an infinite total zero-point energy if the field lives in an infinite ambient space. He therefore introduces this space cutoff, taking the $L \to \infty$ limit at the end of the calculations and focusing on the vacuum energy per unit area. Casimir further assumes that within the cavity lives a Maxwell field, but, since we'll mostly talk about the scalar field in our discussion, we consider a massless Klein-Gordon field instead (the treatment is virtually identical, recalling that each polarisation mode of the Maxwell field behaves as a massless scalar). Key ingredients when discussing vacuum energy are the conditions that the field ϕ obeys at the boundary, that Casimir assumes to be homogeneous Dirichlet (i.e. $\phi = 0$ on the plates, as they are taken to be perfectly conducting). Recall that the Hamiltonian for a scalar field of

mass M is given by:

$$\hat{H} = \frac{1}{2}\sum_k \omega_k \left(\hat{a}_k^\dagger \hat{a}_k + \hat{a}_k \hat{a}_k^\dagger\right), \tag{2.1}$$

where the sum is formal and over all modes, $\vec{k}$ is the wavevector and

$$\omega_k = \sqrt{\vec{k}^2 + M^2}. \tag{2.2}$$

This highlights the useful physical picture of the scalar field being represented by a system of uncoupled 1D quantum harmonic oscillators. Exploiting the usual commutator relations for the creation/annihilation operators and introducing the number operator $\hat{n}_k = \hat{a}_k^\dagger \hat{a}_k$, we see that we can rewrite Equation 2.1 in the more explicative form:

$$\hat{H} = \sum_k \omega_k \left(\hat{n}_k + \frac{1}{2}\right), \tag{2.3}$$

emphasising that the zero point energy of the field (namely $\langle 0|\hat{H}|0\rangle$, with $|0\rangle$ being the vacuum state) is infinite, and has to be renormalised. In general, two approaches are usually taken to fix this. Either we normal order the operators (effectively moving all the $\hat{a}$ to the right and $\hat{a}^\dagger$ to the left) forcing the vacuum energy to be zero, or we add a constant in our Lagrangian (and, consequently, Hamiltonian). The latter can always be done, and there's no reason not to! This c-number can then be chosen so that it cancels the divergences in the zero-point energy, making it finite and, if one wishes, zero. The reason why we can safely discard the divergent contribution to the energy resides in the fact that, in a Poincaré invariant theory, the physics is never affected by the absolute value of the energy of a physical system. Only differences between, for example, the potential energy of distinct configurations of the same system can have physically measurable effects, meaning that a redefinition of the energy scale should never affect physical predictions. If we want our Hamiltonian to be the generator of time translation (and not some shifted version of $\hat{H}$), its vacuum expectation value should actually vanish, but in the end this is an arbitrary choice and any finite vacuum expectation value (VEV) will do. There is, however, a caveat: gravity. Indeed, the zero-point energy of a physical field contributes to the expectation value of its stress-energy tensor, and should be included (if one were to know how to do it) when evaluating the curvature that the quantum field imprints on the ambient space. Here, we're not free to choose the zero-point energy as we please anymore, since its value has physical consequences. We briefly touch on this issue later, and focus now on how to explicitly evaluate the Casimir force in the original setup.

Let's consider two configurations, one where the moving plate is close (i.e. $a \ll L$) to the $x-y$ plate and one where the two are a significant distance away (e.g. $a = L/2$). The reason why we want to look at two distinct positioning of the plates is

that we take the approach of looking at the difference between the zero-point energies of the two arrangements to normalise the vacuum energy, instead of normal ordering the Hamiltonian or adding to it infinite constants to absorb the divergences. Doing so, the divergent (geometry invariant) contribution is cancelled in the subtraction and we end up (as we shortly see) with a finite physical quantity. This is illustrative since it is exactly how we renormalise the free energy in our numerical investigations. Note, however, that this this does not always happen when talking about Casimir energy: the cancellation is dependent on the geometry of the problem and on the boundary conditions. Conveniently, the infinite contributions happen to drop out from the energy difference both in the standard parallel plates arrangement and in our deformed sphere case [15]. Given that we are ultimately interested in the $L \to \infty$ limit, we take the components of the wavevector k_x and k_y in the x and y direction to be continuous, while k_z has a discrete spectrum due to the confinement caused by the plates, i.e.:

$$k_z = \frac{\pi}{a} n, \tag{2.4}$$

with n an integer. Then, the formal sum when taking the VEV of Equation 2.3 is given by:

$$E = \frac{1}{2}\frac{L^2}{\pi^2} \int_0^\infty \mathrm{d}k_x \int_0^\infty \mathrm{d}k_y \sum_{n=1}^{\infty} \sqrt{n^2 \frac{\pi^2}{a^2} + k_x^2 + k_y^2}. \tag{2.5}$$

It is now convenient to change variables, introducing polar-like coordinates for the $k_x - k_y$ plane. Calling k the magnitude of the component of $\vec{k}$ parallel to the plates and θ the angle with respect to the k_x axis (recall that we are integrating on the positive-positive quadrant), we get:

$$E = \frac{L^2}{2\pi^2} \int_0^{\pi/2} \mathrm{d}\theta \int_0^\infty \mathrm{d}k \sum_{n=1}^{\infty} \sqrt{n^2 \frac{\pi^2}{a^2} + k^2}\ k \tag{2.6}$$

$$= \frac{L^2}{4\pi} \int_0^\infty \mathrm{d}k \sum_{n=1}^{\infty} \sqrt{n^2 \frac{\pi^2}{a^2} + k^2}\ k. \tag{2.7}$$

Now, in the big L limit, the $a = L/2$ configuration can be approximated as having continuous k_z spectrum as well. Taking the difference between the energies of the two setups leads to:

$$\Delta E = \frac{L^2}{4\pi} \int_0^\infty \mathrm{d}k \left(\sum_{n=1}^{\infty} \sqrt{n^2 \frac{\pi^2}{a^2} + k^2} - \int_0^\infty \mathrm{d}k_z \sqrt{k_z^2 + k^2}\ \frac{a}{\pi} \right) k. \tag{2.8}$$

This expression is naively divergent due to the high energy modes. In his treatment, Casimir opts to introduce a high frequency cutoff, motivating the choice with the argument that big $|\vec{k}|$ modes (e.g. X-rays in the case of the photon field) have such a significant penetration depth inside the conducting plates that they actually do not

see them at all (when in experiments the perfectly conducting idealisation does not hold). A more profound rationale for this approach would be the sensible statement that we do not expect our low energy physics to be affected by the UV physics. We then modify our theory in the high energy regime with a regulator such that Equation 2.14 is finite, hoping that the final answer turns out to be non-divergent as well and, crucially, independent from such regulator once we remove it! If this happens, we can be confident to have in our hands a meaningful physical result. A further motivation for this approach is that it would be presumptuous to assume that our theory is valid for all energy scales, so we simply don't! Following Casimir's treatment, we hence introduce a cutoff scale Λ above which the modified energy-momentum relation quickly decays to zero. In particular, we multiply our integrand with a function $f(k)$ which gives unity for $k \ll \Lambda$, but rapidly dies off as $k \to \infty$. This effectively turns Equation 2.14 into:

$$\Delta E = \frac{L^2}{4\pi} \int_0^\infty \mathrm{d}k \left(\sum_{n=1}^{\infty} \sqrt{n^2 \frac{\pi^2}{a^2} + k^2} - \int_0^\infty \sqrt{k_z^2 + k^2}\ \frac{a}{\pi}\ \mathrm{d}k_z \right) k\ f(k). \tag{2.9}$$

The details of this function eventually turn out not to be important, as we hoped, so that our calculations are indeed physically meaningful. Skipping a few steps which involve making use of the Euler-Maclaurin formula to turn the discrete sum into a continuous integral, we find the difference in vacuum energies per unit area between the two plates configurations as $\Lambda \to \infty$ to be given by:

$$\Delta E / L^2 = -\frac{\pi^2}{1440 a^3}, \tag{2.10}$$

which is indeed finite and regulator (both space and momentum cutoff) independent! Furthermore, it has an explicit dependence on the separation between the plates, which manifests itself as an attractive force between the walls that goes as a^{-4} (as one can see by simply differentiating Equation 2.10).

2.1.3 Dimensional Regularisation

Although it gives a clean physical picture of the regularisation procedure, the high momentum cutoff approach is sometimes deprecated due to the fact that it breaks Lorentz invariance by picking a preferred energy scale (until, of course, the regulator is removed). An alternative approach that preserves the Lorentz group is dimensional regularisation, which sacrifices the number of dimensions of our physical reality in the name of finiteness. A detailed derivation is not of interest for the present discussion, but the underlying idea is to deform the number of dimensions d of the physical space away from the 3 spatial dimensions of the cubical cavity. The

integral is then calculated by using the relation:

$$\int \mathrm{d}^d k f(|\vec{k}|^2) = \frac{2\pi^{d/2}}{\Gamma(d/2)} \int \mathrm{d}k |\vec{k}|^{d-1} f(|\vec{k}|^2), \tag{2.11}$$

where f is some function that depends solely on the magnitude of the wavevector and $\Gamma(x)$ is the gamma function. To take this alternative route, instead of subtracting the vacuum energy from some reference configuration, one simply considers the Casimir energy given in Equation 2.14 (which needs to be appropriately modified for a d dimensional cavity, with the conducting walls becoming conducting hyperplanes). One then exploits Equation 2.11 to transform the multidimensional integral into an integral over the magnitude of the wavevector only. Jumping to the result, we obtain:

$$E = \frac{1}{2}\left(\frac{L}{2}\right)^{d-1} \frac{\pi^{(d+1)/2}}{a^d} \frac{\Gamma(-d/2)}{\Gamma(-1/2)} \zeta(-d), \tag{2.12}$$

where $\zeta(x)$ is the Riemann zeta function. While the Γ function is defined for negative non-integers (it has, however, simple poles at negative integers and 0), the ζ function is not! Nonetheless, it can be analytically continued to those values, obtaining:

$$E = -\frac{L^{d-1}}{2a^d} (4\pi)^{-(d+1)/2} \, \Gamma\left(\frac{d+1}{2}\right) \zeta(d+1). \tag{2.13}$$

The last step is, then, to return to our 3-dimensional space, taking the limit $d \to 3$ [16]. In doing so, we realise the result is already finite and agrees with Equation 2.10. But where did all the divergences go? They were dropped when we analytically continued the Riemann zeta function to negative arguments. Indeed, $\zeta(x)$ is defined for $x > 1$ and, having a simple pole at $x = 1$, it is not analytic from $x > 1$ all the way down to $x = -3$! Since only the finite contributions are captured by the analytic continuation, all the divergences are removed in the process.

2.1.4 Mass and Temperature

Generalising Equation 2.10 for massive scalar fields is trivial. Indeed, the only modification one needs to do to the above discussion is not to set $M = 0$! Doing so, Equation 2.14 clearly becomes:

$$\Delta E = \frac{L^2}{4\pi^2} \int_0^\infty \mathrm{d}k \left(\sum_{n=1}^{\infty} \sqrt{n^2 \frac{\pi^2}{a^2} + k^2 + M^2} - \int_0^\infty \mathrm{d}k_z \sqrt{k_z^2 + k^2 + M^2} \, \frac{a}{\pi} \right) k. \tag{2.14}$$

This has to be renormalised as before. In the interest of time, we once more skip a few steps and jump to the result after the regulator is removed, namely:

$$\Delta E = -\frac{L^2}{8\pi^2} \frac{M^2}{a} \sum_{n=1}^{\infty} \frac{1}{n^2} K_2(2aMn). \tag{2.15}$$

Here, K_2 is the second modified Bessel function. To understand what this implies for small masses, we use the expansion

$$K_2(2Man) = \frac{2}{(Man)^2} - \frac{1}{2} + \mathcal{O}(M^2). \tag{2.16}$$

Up to order M^2, we obtain:

$$\Delta E = \Delta E/L^2 = -\frac{\pi^2}{1440a^3} + \frac{L^2M^2}{96a} + ..., \tag{2.17}$$

showing that a small mass gives a positive contribution that opposes the massless negative vacuum energy. Indeed, in the $M \gg a$ limit, we instead obtain:

$$\Delta E = -\frac{L^2}{16\pi^2}\frac{M^2}{a}\left(\frac{\pi}{Ma}\right)^{1/2} e^{-2Ma}, \tag{2.18}$$

showing that as $M \to \infty$ the poistive contribution coming from the mass effects makes the Casimir energy approach zero [16].

The story of thermal corrections is a bit different. Indeed, a thermal states corresponds to real quanta being excited in the field, meaning that the latter will no longer be in a vacuum state. If a large number of states are occupied, then a quantum statistical description is required, and we can borrow concepts from classical thermodynamics to study the system. In the case of a classical system at temperature T, we can extract the free energy (and, really, all the relevant physical quantities) of the ensemble by looking at the partition function:

$$Z_c = \sum_i e^{-E_i/T}, \tag{2.19}$$

where the sum is over all the microstates of the system and E_i is the energy of the ith microstate. Then, the classical Helmholtz free energy of the system is given by:

$$F_c = -T\ln(Z_c). \tag{2.20}$$

For the case of a quantum field, the generalisation to Equation 2.19 is given by the following functional integral:

$$Z = \int \mathcal{D}\phi\ e^{iS}, \tag{2.21}$$

with S being the action of our theory. The quantum Helmholtz free energy is then simply:

$$F = -T\ln(Z). \tag{2.22}$$

If we assume the thermal state of the field to be statistically independent to its vacuum fluctuations, Z factorises as $Z = Z_0Z_T$, with Z_0 and Z_T being the partition functions of the vacuum fluctuations and the excited quanta respectively. Indeed,

assuming such factorisation to occur, we can rewrite:

$$F = -T\ln(Z_0) - T\ln(Z_T) = E + F_T. \tag{2.23}$$

We then interpret the Casimir energy as the Helmholtz free energy in the limit $T \to 0$. To derive the thermal correction for the scalar field in the Casimir setup, we model the field as a series of independent (meaning that the partition function for each mode factorises) oscillators with the usual KG dispersion relation. Then, the partition function is given by:

$$Z = \prod_k \left(\sum_n e^{-\frac{1}{T}\omega_k (n+\frac{1}{2})} \right) = \prod_k \frac{e^{-\frac{1}{2T}\omega_k}}{1 - e^{-\frac{1}{T}\omega_k}}, \tag{2.24}$$

leading to

$$F = \sum_k \left(\frac{\omega_k}{2} + T\ln\left[1 - e^{-\frac{\omega_k}{T}}\right] \right). \tag{2.25}$$

As expected, we immediately recognise the first term in Equation 2.25 as the usual zero-point energy of the scalar field, which we know to diverge. We still need to regulate the free energy as explained before, both with a space regulator and by taking the difference with respect to a reference configuration (i.e. the one with $a \approx L/2$). The second term is what instead gives the thermal corrections. To see what these amount to, we need to evaluate them at both large and small separation of the plates and take the difference of the results. Note that there was in principle no need to regulate the second term, being it manifestly finite. Nonetheless, since we need to do it to get a finite vacuum energy, we need to consider the energy difference also for the thermal corrections. For $a \approx L/2$, we take again the approximation of k_z having a continuous spectrum. Then, making the formal sum over k an integral over $\vec{k}$ we see that:

$$\begin{aligned} F_T &= T\frac{L^2 a}{(2\pi)^3} \int \mathrm{d}^3 k \ln\left(1 - e^{-\frac{\omega(\vec{k})}{T}}\right) \\ &= \frac{T^4 L^2 a}{2\pi^2} \int_0^\infty \mathrm{d}k \ln\left(1 - e^{-x}\right) k^2 \\ &= -\frac{L^2 a}{\pi^2}\zeta(4) T^4 \\ &= -\frac{1}{90}\pi^2 T^4 L^2 a. \end{aligned} \tag{2.26}$$

Repeating the same calculations for the $a \ll L$ limit (i.e. with k_z quantised as in Equation 2.4) and taking the difference between the two configurations, we eventually obtain:

$$\Delta F_T = \frac{L^2 \pi}{2a^2} \left[-\frac{\zeta(3)}{2} T \left(\frac{aT}{\pi}\right)^2 + \sum_{n=1}^{\infty} b(a,T,n) + 2\zeta(4) T \left(\frac{aT}{\pi}\right)^3 \right], \tag{2.27}$$

where we have defined

$$b(a,T,n) = \frac{1}{2}T\int_{n^2}^{\infty} \mathrm{d}y \ln\left[1 - \exp\left(-\frac{\pi\sqrt{y}}{aT}\right)\right]. \tag{2.28}$$

Analytically, we cannot go beyond this since the sum cannot be written in a nice closed form, unless we take the small or big temperature limits. For $aT \ll 1$ we can expand the logarithm in b and resum to obtain:

$$\sum_{n=1}^{\infty} b(a,T,n) = -T\left[\left(\frac{aT}{\pi}\right)^2 + \frac{aT}{\pi}\right] e^{-\frac{\pi}{aT}} + \mathcal{O}\left(e^{-\frac{2\pi}{aT}}\right). \tag{2.29}$$

From this we see that, interestingly, for small temperatures the thermal effects first cause the free energy to become more negative. As T increases, the T^4 term eventually starts dominating, inverting the trend. The high temperature regime is even more suggestive. Indeed, expanding b for $aT \gg 1$, one can show that up to order $\mathcal{O}(e^{-4\pi aT})$:

$$\Delta F_T = L^2\pi^2\left[\frac{1}{1440a^3} - \frac{\zeta(3)T}{16\pi^4a^2} + \frac{T}{8\pi^3}\left(\frac{1}{a^2} + \frac{4\pi T}{a}\right)e^{-4\pi aT}\right]. \tag{2.30}$$

Curiously, the first term *exactly* cancels the negative Casimir energy! It is possible to show that this cancellation is not even unique to the parallel plate setup, but it is a more general result [16].

2.2 QFT in Curved Spacetime

2.2.1 Classical Fields

Before jumping into an overview of the basics of quantum fields in curved spacetime, it's worthwhile to spend some time discussing classical fields living in fixed curved geometry.

The case of the scalar field is quite straightforward, and moving from Minkowski to curved spacetime does not cause any major conceptual difficulty. The main idea is that the actions for the classical fields (Klein Gordon and Dirac fields, for example) in flat space are not diffeomorphism invariant, which is clearly problematic. To deal with this issue we treat the flat space action as the action that an observer sees in a local inertial frame, and follow the usual approach used to move from local to covariant equations in GR. To obtain the full curved space action, we then just need to promote partial derivatives to the covariant derivatives appropriate for the object they're operating on (recall that for scalar fields they are the partial derivatives themselves) and the Minkowski metric $\eta_{\mu\nu}$ to the metric tensor $g_{\mu\nu}$ describing the geometry of spacetime. We also need to replace the integration measure with the usual GR covariant volume element involving the determinant of the metric g.

Therefore, we perform the following substitutions:

$$\partial_\mu \to \nabla_\mu, \qquad \eta_{\mu\nu} \to g_{\mu\nu}, \qquad \mathrm{d}^4x \to \mathrm{d}^4x\sqrt{-g}. \tag{2.31}$$

obtaining for the K-G field:

$$S[\phi] = \frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}(g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - m^2\phi^2). \tag{2.32}$$

The story is, however, is a bit more complex. Indeed, one can also couple ϕ to the scalar curvature, i.e. the Ricci scalar R, obtaining the action:

$$S[\phi] = \frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}(g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \xi R\phi^2 - m^2\phi^2), \tag{2.33}$$

where ξ is just a number that governs the strength of this higher order coupling of the field to the curvature. There is no reason not to include such term, since it is both diffeomorphism invariant and dimensionally consistent. With little thought one sees that both Equation 2.32 and Equation 2.33 reproduce the usual flat space action with the choice $g_{\mu\nu} = \eta_{\mu\nu}$ (as $R = 0$ for Minkowski). What is the correct way to couple the scalar field to gravity, one may then ask? Both, and neither! In the end, there is no a priori answer: our physical world is only locally Minkowski, so the real physics is encapsulated by the action in *curved* spacetime, which cannot be uniquely inferred just by observing phenomena in a locally flat portion of the manifold. This mirrors the problem of the ordering of the operators in quantum mechanics: the real physics is given by the quantum Hamiltonian which has a definite ordering, but as we take the classical limit that information is lost and cannot be recovered just by looking at classical phenomena. All we can do as theorists is to make a choice and let the experiments decide. The simplest one, i.e. the choice made in Equation 2.32, takes the name of *minimal coupling* as ϕ couples to the gravitational field through the lowest possible number of objects (volume element and metric). In our study we'll consider both $\xi = 0$ and non-minimally coupled fields, studying how the Casimir energy changes as we vary ξ. The scalar curvature term becomes particularly interesting in the case of a massless field, as for $\xi = 1/6$ the action (and, hence, the equations of motion) become invariant under:

$$(g_{\mu\nu}, \phi) \to (\Omega^2 g_{\mu\nu}, \phi/\Omega), \tag{2.34}$$

for any non-zero conformal factor $\Omega(x)$. The non-minimally coupled model is of increasing interest in cosmology for various reasons, from Quintessence to its proposed role in inflation [17] [18]. For the non-minimally coupled field, the equation of motion is given by:

$$(\Box + m^2 + \xi R)\phi = 0, \tag{2.35}$$

where the d'Alambertian in curved spacetime acting on a scalar field is

$$\Box\phi = \nabla_\mu \partial^\mu \phi. \tag{2.36}$$

A key object when we introduce backreaction of the Casimir energy on the spacetime geometry is the stress-energy tensor $T^{\mu\nu}$ associated to ϕ. For the classical non-minimally coupled scalar this is given by:

$$T_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi - \frac{1}{2}g_{\mu\nu}((\partial\phi)^2 - m^2\phi^2) + \xi\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + g_{\mu\nu}\Box - \nabla_\mu\nabla_\nu\right)\phi^2. \tag{2.37}$$

As our numerical analysis focuses only on the scalar field, there's no point in spending time talking about fermions. It is worth mentioning, however, that when one couples spinors to curved spacetime, the picture gets a bit more complicated. Indeed, the presence of spinors necessarily implies non-zero torsion, meaning that the connection has to deviate from the Levi-Civita one. This leads to subtleties completely absent for the case of the scalar.

2.2.2 Canonical Quantisation of the Scalar Field

It is now time to try to quantise our scalar field on fixed curved spacetime (no backreaction of the field on the geometry for now) and see if anything interesting happens. Spoiler alert, it does. For our numerical investigations, such effects are not terribly relevant. It is however illustrative to mention them, as they underline the curious consequences that curvature can imprint on quantum fields.

Let's go back to our old friend, the KG field. While canonical quantisation is neither the most elegant nor the most useful approach to quantise a field in curved spacetime, it provides a nice place to start as it highlights some conceptual differences that arise with respsect to the Minkowski case. Before starting the quantisation procedure, let's recall that the conjugate field momentum is given by:

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial(\nabla_t\phi)} = \sqrt{-g}\; g^{\mu t}\nabla_\mu\phi. \tag{2.38}$$

Consider now a folitation of our manifold at time t which gives a spacelike hypersurface Σ_t. There exists a natural inner product between two solutions ϕ_1 and ϕ_2 to Equation 2.35, namely the Klein-Gordon inner product, defined as:

$$(\phi_1, \phi_2) \coloneqq -i\int_{\Sigma_t} \mathrm{d}^3x\sqrt{h}(\phi_1\nabla_\mu\phi_2^* - \phi_2^*\nabla_\mu\phi_1)n^\mu, \tag{2.39}$$

where n^μ is the vector perpendicular to the spacelike hypersurface and h_{ij} is the metric induced on Σ_t by the foliation. The inner product so defined can also be shown to be independent on the foliation chosen. As Equation 2.35 is linear, we can expand any solution in a complete basis of modes $\{f_i\}$, which can be taken to be

orthonormal with respect to the KG inner product:

$$(f_i, f_j) = \delta_{ij}, \qquad (f_i^*, f_j^*) = -\delta_{ij}, \qquad (f_i, f_j^*) = 0. \tag{2.40}$$

Note that the complex conjugate modes have to be included, since they are linearly independent from the other modes and if f solves Equation 2.35 then f^* does as well. In order to build our Fock space, we now need to identify positive and negative frequency solutions, as they allow a proper definition of the vacuum state and of the creation and annihilation operators. If the spacetime admits a global timelike Killing vector K (i.e. it is static), then it is easy to see that we can choose a coordinate system such that the e.o.m. in Equation 2.36 is separable into a pure time derivative term and a term that involves spatial derivatives and functions of space only. We can therefore find separable solutions that obey:

$$\mathscr{L}_K f_i = -i\omega_i f_i, \qquad \mathscr{L}_K f_i^* = i\omega_i f_i^*, \tag{2.41}$$

where $\mathscr{L}_K$ indicates the Lie derivative with respect to K. When this happens, we can assign the label positive/negative energy modes to f_i and f_i^* respectively (note that really $\omega > 0$ in both cases). Obviously, exactly the same happens for the usual flat space solutions, whose basis is given by $f(p) = e^{-ip.x}$ (where the continuous label p is the 4-momentum). However, there is no guarantee that such Killing vector exists, and for most spacetime it doesn't, so let's assume for now that the geometry is not static. Without Equation 2.41, the condition given by Equation 2.40 is not enough to fix the subspace of solutions $\{f_i\}$ uniquely. There are many possible choices of frames (and hence basis) where the separation between positive and negative frequencies can be done, but without a timelike Killing vector none of these is a natural choice. Let's nonetheless let our observer, that we call Alice, choose one of the possible orthonormal sets, and see what an happens. Alice can build the general classical solution as a superposition of the basis modes:

$$\phi(x) = \sum_i \left[a_i f_i(x) + a_i^* f_i^*(x)\right], \tag{2.42}$$

where the sum is formal and over all basis functions (it is an integral if the label i is continuous as in Minkowski) and the a_i are just coefficients (note that we have already imposed the reality condition on ϕ). Plugging Equation 2.42 into Equation 2.38 one can then obtain the form of the conjugate momentum π. It is now time to put the party hats on and start the main event: second quantisation. Analogously to the flat space quantisation, Alice promotes ϕ and π to quantum mechanical

operators, and force them to obey the bosonic equal time commutator relations:

$$\begin{aligned}[\hat{\phi}(t,\vec{x}),\hat{\pi}(t,\vec{y})] &= \frac{i}{\sqrt{-g}}\delta^{(3)}(\vec{x}-\vec{y}),\\ [\hat{\phi}(t,\vec{x}),\hat{\phi}(t,\vec{y})] &= [\hat{\pi}(t,\vec{x}),\hat{\pi}(t,\vec{y})] = 0.\end{aligned} \tag{2.43}$$

What are really being promoted to operators here are the coefficients of the basis functions in Equation 2.42. The field operator is then given by:

$$\hat{\phi}(x) = \sum_i \left[\hat{a}_i f_i(x) + \hat{a_i}^\dagger f_i^*(x)\right], \tag{2.44}$$

and the explicit form of the conjugated momentum follows from Equation 2.38. In order for Equation 2.43 to be satisfied, one finds that $\hat{a_i}$ and $\hat{a_i}^\dagger$ must obey:

$$[\hat{a_i},\hat{a_j}^\dagger] = \delta_{ij}, \qquad [\hat{a_i},\hat{a_j}] = [\hat{a_i}^\dagger,\hat{a_j}^\dagger] = 0, \tag{2.45}$$

reminiscent of the quantum harmonic oscillator commutator relations, from which we can identify $\hat{a}_i$ as the annihilation operator and $\hat{a}_i^\dagger$ as the creation operator. From here Alice can also define the vacuum state with respect to the basis $\{f_i\}$ as the one that satisfies:

$$\hat{a}_i \left|0\right\rangle_f = 0 \qquad \forall i, \tag{2.46}$$

The vacuum can then be used to build n-particles states for the ith eigenmode by acting n times with $\hat{a}_i^\dagger$ (and appropriately normalising). The number operator with respect to the basis $\{f_i\}$ can be finally defined as:

$$\hat{N}_{i|f} \coloneqq \hat{a}_i^\dagger \hat{a}_i. \tag{2.47}$$

We have repeatedly emphasised that the states (and operators) Alice is building are *with respect to a basis*. What does this mean? Shouldn't any correctly normalised basis choice lead to the same answer? In curved spacetime the answer is surprisingly: no! In flat space, the vacuum state and the mentioned operators are invariant under Poincaré transformations. This resides on the fact that neither translations, rotations nor Lorentz boost change the sign of the frequencies of the modes: positive and negative frequencies separated with respect to the global timelike Killing vector (i.e. the Minkowksi time coordinate) are well defined. However, there is no such thing as Poincaré symmetry in curved spacetime (outside a local inertial frame, that is), and positive and negative frequencies have no covariant meaning in general relativity [19]. Consider a second observer, Bob, that lives on the curved geometry but might be using a different coordinate system. As there is no natural time direction, he can pick an alternative basis $\{g_a\}$ and repeat the same procedure. He will have his own Fock space, with a (possibly) different vacuum state $\left|0\right\rangle_g$ that is

annhilated by the new annihilation operators $\hat{b}_a$. Bob's number operator is $\hat{N}_{a|g} = \hat{b}_a^\dagger \hat{b}_a$. However, the two observer are describing the same *scalar* field, which is itself invariant under general coordinate transformations! This gives a route to uncover the linear relation that exists between the two complete basis (the KG equation is still linear, after all):

$$g_a = \sum_i \left[\alpha_{ai} f_i + \beta_{a_i} f_i^*\right], \qquad f_i = \sum_a [\alpha_{ai}^* g_a - \beta_{a_i} g_a^*]. \tag{2.48}$$

Again, the sums here are formal and α_{ai} and β_{ai}, called Bogoliubov coefficients, are just formal matrix elements. These relations take the name of Bogoliubov transformations and the coefficients are obtained through the KG inner product:

$$\alpha_{ai} = (g_a, f_i), \qquad \beta_{ai} = -(g_a, f_i^*), \tag{2.49}$$

which can be seen using orthonormality of the basis functions. They also satisfy the following normalisation conditions:

$$\sum_k (\alpha_{ak}\alpha_{ik}^* - \beta_{ak}\beta_{ik}^*) = \delta_{ai}, \qquad \sum_k (\alpha_{ak}\beta_{ik} - \beta_{ak}\alpha_{ik}) = 0. \tag{2.50}$$

From these expressions we can relate the two sets of ladder operators via:

$$\hat{a}_i = \sum_a [\alpha_{ai}\hat{b}_a + \beta_{a_i}^* \hat{b}_a^\dagger], \qquad \hat{b}_a = \sum_i [\alpha_{ai}^* \hat{a}_i - \beta_{a_i}^* \hat{a}_i^\dagger], \tag{2.51}$$

and, hence, the states of one observer to the ones of the other. A question naturally arises: how does the observer in the basis $\{g_a\}$ see the vacuum state $|0\rangle_f$? The answer is easily found by computing the expectation value of $\hat{N}_{a|g}$ for such a state, which turns out to be:

$${}_f\langle 0|\, \hat{N}_{a|g}\, |0\rangle_f = \sum_i |\beta_{ai}|^2. \tag{2.52}$$

It is (unless $\beta_{ai} = 0$) indeed filled with particles [20]! Curvature strikes again: after energy and momentum, also particles lose their privileged place in the universe and become (at best) observer-dependent physical objects! With no natural time direction to use in order to separate between positive and negative frequencies, each observer can choose its own set of basis modes and obtain a different vacuum (and, hence, Fock space). The process is all too arbitrary and artificial, so there is little reason to talk about particles at all. Of course, in a region of spacetime in which the curvature does not vary significantly (i.e. in a LIF), it still makes sense to talk about particles (otherwise what is CERN doing, one might ask), but the key point is that we can do so only *locally*. For the case of static spacetimes, all observers have a recipe to divide the basis modes between positive and negative frequencies, so they can meaningful compare their Hilbert spaces. This does not mean, however,

that they will all see the same particles: as said, the sign of the frequencies is not covariant, so two observers related by a general coordinate transformation will not usually agree with each other's separation. This strikingly results in the Unruh effect: an accelerated observer in flat space (i.e. and observer in Rindler space) sees the ordinary Minkowski vacuum as a thermal state filled with particle! Clearly, the spacetime is the same for both observers, so they both enjoy the same timelike Killing vector. Nonetheless, the change of coordinates causes a different identification of positive and negative frequencies with respect to K, resulting in a different vacuum for the two observers [21]. Other than in small portions of spacetime and in static spacetimes, one can also talk about particles in a space that is asymptotically flat or with only a compact region (both in time and space) of non-vanishing curvature. In this cases, we also have a natural time direction (i.e. the asymptotic timelike Killing vector), which give to all observers a recipe to separate between positive and negative frequencies when in the asymptotically flat portion of spacetime. The latter situation is considered by Wald [21], who describes a toy model of particle pairs (and, in the case of complex fields, particle-antiparticle pairs) production by gravity. The main idea behind the process is that, while in the curved region the concept of particle breaks down, we can still define positive and negative frequency modes in the asymptotic past and future, where the spacetime is isometric to Minkowski. Then, evolving in time the field and comparing the asymptotic in and out states (much like in a scattering problem) one can count the number of particle produced (the key concept here is that the localised curvature changes the natural basis $\{f_i\}$ so that, as before, the vacuum state and the observed number of particles change as well). The mathematical reason is still the same: if (originally) positive and negative frequencies get mixed by the change of basis, then particles are produced, even if we started with a vacuum state in the first place. Virtually all time-dependent gravitational fields induce such mixing, hence the proposal by cosmologists that such process could be the reason why we ended up with particle after the Big Bang, during inflation. This process of particle creation by a gravitational field is also what causes the theorised Hawking radiation from black holes [19].

2.3 Casimir Energy in Fixed Curved Spacetime

When talking about Casimir energy in curved spacetime, it is natural to start from its thermal generalisation: Helmholtz free energy. Indeed, as discussed before, we can see the Helmholtz free energy for a non-interacting quantum field as composed by the Casimir zero-point energy plus thermal correction. As usual, the free energy of a field is given by:

$$F = -T \ln Z, \tag{2.53}$$

where recall we have set the Boltzmann constant to one, T is the temperature and Z is the partition function of the physical system. For a free scalar field this is given by the functional integral

$$Z = \int \mathcal{D}\phi \; e^{iS} = \int \mathcal{D}\phi \; e^{i\int \mathrm{d}^4x\sqrt{-g}\frac{1}{2}\phi(\Box - m^2)\phi} = \Big[\det(\Box - m^2 - \xi R)\Big]^{-\frac{1}{2}}, \quad (2.54)$$

where in the last step we have simply performed the standard path-integral Gaussian integration. We immediately see the expected divergences starting to appear: the determinant of the differential operator (as for the determinant of a matrix) is given by the product of all the eigenvalues which, in the case of Equation 2.54, are infinite! Z is UV divergent and we will have to regulate it. Let's move on, for now, and reformulate the partition function in what will turn out to be a much more useful form. For any elliptic (suitably linear) differential operator $\mathcal{D}$, we can define the related ζ-function as:

$$\zeta_{\mathcal{D}}(s) = \sum_i \lambda_i^{-s} = \mathrm{Tr}\mathcal{D}^{-s}, \quad (2.55)$$

where the sum is formal and is intended to be over all eigenvalues λ_i of $\mathcal{D}$. The familiar Riemann ζ-function arises (up to a multiplicative factor of two) when we choose $\mathcal{D} = -\partial_\theta^2$, i.e. for the Laplacian on the circle S^1 [22]. It is then straightforward to see that

$$\left.\frac{\mathrm{d}\zeta_{\mathcal{D}}}{\mathrm{d}s}\right|_{s=0} = -\sum_i \ln(\lambda_i)\lambda_i^{-s}|_{s=0} = -\sum_i \ln\lambda_i = \ln\left(\prod_i \lambda_i\right) = -\ln\det(\mathcal{D}), \quad (2.56)$$

which means that the Helmholtz free energy for a non-minimally coupled scalar can be expressed as [23]:

$$F = -\frac{T}{2}\left.\frac{\mathrm{d}}{\mathrm{d}s}\zeta_{\Box - m^2 - \xi R}\right|_{s=0}. \quad (2.57)$$

2.3.1 Heat Kernel Method

We then see that all we need to do to evaluate the free energy is to compute the ζ-function related to its differential operator $\mathcal{D}$. Let ψ_λ be the eigenfunction of $\mathcal{D}$ associated to the eigenvalue λ. We introduce the heat kernel $\mathcal{K}_{\mathcal{D}}$, namely the function satisfying the following heat equation:

$$\mathcal{D}\mathcal{K}_{\mathcal{D}}(x, x'; t_k) = -\frac{\partial}{\partial t_k}\mathcal{K}_{\mathcal{D}}(x, x'; t_k), \quad (2.58)$$

subject to the initial condition

$$\mathcal{K}_{\mathcal{D}}(x, x'; 0) = \delta(x - x'). \quad (2.59)$$

Here, we have added a fifth auxiliary spacetime coordinate t_k that we'll refer to as *kernel time*. It can be proved that such kernel exists and it's unique on a Riemannian manifold without boundary, such as the ones we consider later in the study [24]. It is readily seen that this heat equation is satisfied by:

$$\mathcal{K}_{\mathcal{D}}(x, x'; t_k) = \sum_i \psi_{\lambda_i}(x)\psi^*_{\lambda_i}(x')e^{-\lambda_i t_k}. \tag{2.60}$$

The heat kernel encodes much our QFT. For example, it follows that $\mathcal{K}_{\mathcal{D}}$ provides an integral representation for the propagator D:

$$D^{-1}(x, x') = \int_0^\infty \mathrm{d}t_k \mathcal{K}_D(x, x'; t_k). \tag{2.61}$$

More to the point with regards to our discussion, it is possible to show with little work that $\zeta_{\mathcal{D}}(s)$ can be expressed in terms of the heat kernel as

$$\begin{aligned}\zeta_{\mathcal{D}}(s) &= \frac{1}{\Gamma(s)} \int_0^\infty \mathrm{d}t_k t_k^{s-1} \int \mathrm{d}^4x\sqrt{-g}\mathcal{K}_{\mathcal{D}}(x, x; t_k) \\ &= \frac{1}{\Gamma(s)} \int_0^\infty \mathrm{d}t_k t_k^{s-1} K_{\mathcal{D}}(t_k),\end{aligned} \tag{2.62}$$

where in the last step we have defined:

$$K_{\mathcal{D}}(t_k) \coloneqq \int \mathrm{d}^4x\sqrt{-g}\mathcal{K}_{\mathcal{D}}(x, x; t_k) = Tr\left(e^{-t_k D}\right) \tag{2.63}$$

We have then reduced the issue of finding the free energy of a quantum field to the one of evaluating the heat kernel of the associated differential operator [23]. To make further progress, it is necessary to additionally specify the physical problem. This is done in the next chapter, where we discuss the heat kernel for the case of a scalar field on a (2+1)-dimensional spacetime.

2.4 Backreaction Effects: Wormholes

Ever since its discovery, the Casimir effect attracted much attention also for its potential applications in the gravity sector. Indeed, many results in general relativity assume the so-called null energy condition (NEC), namely that any physically meaningful stress-energy tensor $T_{\mu\nu}$ should obey:

$$T_{\mu\nu}n^\mu n^\nu \geq 0, \tag{2.64}$$

where n^μ is any null vector. On its most basic essence, the NEC is just the general relativity statement that corresponds to positivity of the energy for the matter content of the Universe. The (likely) most important result that assumes the NEC is the Penrose singularity theorem, which states that for a spacetime obeying the

NEC with a closed trapped null surface $\mathcal{T}$ and a non-compact connected Cauchy surface, then either the spacetime is geodesically incomplete or there exist closed timelike curves. The theorem is essentially the proof that singularities and black holes are robust predictions of general relativity [25]. It had been long believed that the NEC would not be violated by any form of matter content in the Universe, but the stress-energy tensor of the famous Casimir setup showed otherwise. Indeed, its expectation value is given by (for the case of the massless scalar field at zero temperature):

$$\langle \hat{T}_{\mu\nu} \rangle = \text{diag}(-1, 1, 1, -3) \frac{\pi^2}{1440 a^4}, \tag{2.65}$$

which clearly violates the null energy condition. However, it does satisfy the achronal averaged null energy condition, stating that for every *achronal* integral curve $C(\lambda)$ (with λ parametrising the curve) of a null vector field k, the stress-energy tensor obeys:

$$\int_C \mathrm{d}\lambda \; T_{\mu\nu} k^\mu k^\nu \geq 0, \tag{2.66}$$

where by achronal we mean that no two points on the null curve are timelike separated. As it turns out, this averaged condition is satisfied by the stress energy tensor of all the usual QFTs, and it is expected to hold for any matter field in the Universe [26].

Exotic matter (i.e. matter content that violates the NEC) gained some traction in recent years for its possible role in the construction of suggestive solutions of Einstein's equations, such as wormholes and warp drives. It is however important to emphasise that Casimir energy is not always negative! As we have discussed, the vacuum energy is dependent on the geometry and topology of the spacetime where the quantum field lives. Famously, it has been shown that for a scalar field confined in a spherical conducting shell (rather than by parallel planar plates), the Casimir energy is positive [16]. Nonetheless, the fundamental point is that for certain configurations, the vacuum energy of the quantum field *can* be negative, opening the doors to interesting physical phenomena.

2.4.1 The Semi-classical Approach to GR and its Limitations

Before discussing the effects that a negative vacuum energy can have as a source for gravity, we need to first to understand how to couple a quantum field to a classical gravitational field in a sensible physical way. Indeed, the classical Einstein's equations

$$G_{\mu\nu} = 8\pi T_{\mu\nu} \tag{2.67}$$

do not make much sense once we quantise the matter content of the universe, with the left hand side being a classical tensor (i.e. a c-number) and the right and side being an operator. One might expect this issue not to be fundamental (if a quantum

theory of gravity is indeed what we will eventually find as a deeper description of reality), but even from a practical point of view, it would be useful to have a semi-classical approximation to Equation 2.67. This hybrid theory could be useful in order to describe the physics at energy scales where the spacetime is effectively classical but the quantum effects in the matter fields are non neglegible. Indeed, the backreaction of the vacuum energy of quantum fields can be the solution to many unsolved problems in cosmology. A famous example is its candidacy as the origin of Dark Energy. Indeed, since the vacuum energy is expected to gravitate and anything that adds to the vacuum energy density of the universe behaves exactly as a cosmological constant, it is a natural guess. Precisely because the actual value of the energy density matters in general relativity, we cannot lightheartedly renormalise it. If we decide not to do so, we then need to add all the zero point energies of our QFT up to a (big) cutoff energy scale Λ_c, where we believe our theory to break down. Back of the envelope QFT calculations give a discrepancy as high as 120 orders of magnitudes between the experimental results and theoretical predictions, when taking the Planck scale as Λ_c [27]. This is one of the biggest failed predictions in the history of physics, and it is known as *Cosmological Constant Problem*. Of course, a solution would be to just renormalise the vacuum energy density, treating it as a parameter that has to be experimentally measured and not theoretically derived. This is, however, unsatisfactory to many, who claim such fine-tuning to be too artificial [4].

There are currently some proposed ways to generalise Equation 2.67 in order to make the QFT backreact on the spacetime. In the naive approach, the most widely used in the literature, we take the semi-classical equation to be just Equation 2.67 but with the stress-energy tensor being replaced by its expectation value with respect to the quantum state of matter $|\psi\rangle$. This amounts to the replacement $T_{\mu\nu} \to \langle \hat{T}_{\mu\nu} \rangle$, leading to:

$$G_{\mu\nu} = 8\pi \langle \psi | \hat{T}_{\mu\nu} | \psi \rangle , \tag{2.68}$$

and it is what we implicitly assumed in the previous subsection when discussing the NEC for the scalar field confined between parallel plates. While the simplicity of this approach might seem appealing, its drawbacks are non-negligible. Indeed, the limitations have been evident since it was first proposed as a way to avoid quantising gravity altogether, with a very influential paper [28] highlighting the contradictions and unphysical prediction that arise when trying to evaluate the backreaction of a quantum system in a superposition (or even a classical mixture) of states. In particular, through a thought experiment involving shooting classical gravitational waves of very short wavelength to a localised quantum particle, it has been shown that Equation 2.68 allows for either superluminal signalling or violation of the Heisenberg uncertainty principle. Even though there is some criticism [29] about both the argument and the conclusions of [28] which present this as strong

evidence for the necessity of a quantum theory of gravity, over the years other issues with this approach have been uncovered. For example, it has been proved that if one assumes collapse of the state upon measurement (as in standard QM), then the right hand side of Equation 2.68 would not be covariantly conserved, contrary to the LHS which is automatically so thanks to the Bianchi identities [30]. The effect that matter has on the spacetime geometry when the superposition is achieved for states of significantly massive particles displaced one from the other is further issue of this semi-classical limit. In such a situation, Equation 2.68 would cause the metric to be influcenced as if there were two bodies of half the mass of the particle in both locations, which is clearly unphysical. While alternative methods have been recently put forward, like approaches that propose to couple classical and quantum degrees of freedom stochastically [31] , it is yet unclear what the best way to couple quantum fields to a classical background is. Hence, for the present discussion, we assume Equation 2.68 to be hold in order to present recent results that take this semiclassical limit to be valid. This is justified by the fact that we consider our field to be in a definite state (i.e. the vacuum), avoiding the complications that we briefly mentioned above.

2.4.2 Features of Traversable Wormholes

As exotic as they might sound, wormholes are present in many of the most familiar solutions to Einstein's equations. Indeed, they even appear in the very first solution to Equation 2.67 ever found: the Schwarzschild spacetime. When maximally extended through Kruskal coordinates as shown in Figure 2.1, Schwarzschild spacetime can be seen as composed by 4 regions. Regions I and III are isometric asymptotically flat portions of the manifold, while region II corresponds to the

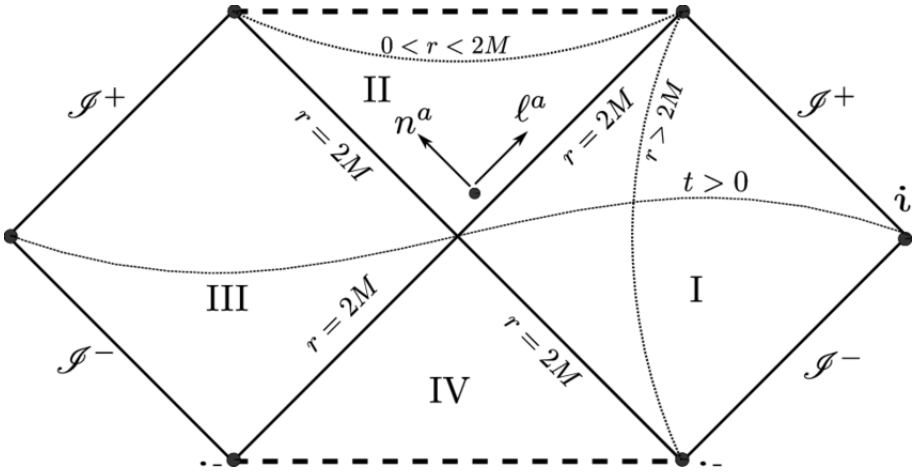


Figure 2.1: Penrose diagram for the maximally extended Schwarzschild spacetime. Reproduced from [32].

interior of a Schwarzschild (non-rotating, neutral) black hole and region IV represents the interior of a Schwarzschild white hole. One can see the two asymptotically flat regions as effectively two distinct Minkowski universes, that, for constant time hypersurfaces, are connected by a region of space corresponding to a Riemannian 3-manifold with topology $\mathbb{R} \times S^2$. We refer to this throat as an Einstein-Rosen bridge. While it might seem at first sight an appealing wormhole to travel between two otherwise disconnected universes, it has to be emphasized that it is a spacelike surface, so it cannot be casually traversed by any observer (nor light)! Indeed, to traverse it, one would have to go through the $r = 2M$ hyperplane which, in Schwarzschild spacetime, corresponds to the event horizon of the black hole: definitely not FDA approved, contrary to Pfizer's COVID vaccine. Furthermore, regions III and IV are not really relevant when discussing black holes originated via gravitational collapse, since they disappear from the solution once the non-zero stress-energy tensor of the interior of the star is taken into consideration. We therefore do not expect such wormholes to be at all present in the Universe. A similar throat appears in sub-extremal magnetically charged black holes that, in natural units, have the magnitude of their charge P smaller than their mass M, i.e. $|P| < M$. As the black hole approaches extremality (i.e. $|P| \to M$) the throat becomes infinitely long. Again, this is still not causally not traversable.

The main issue with the aforementioned examples of wormholes is that they are non-traversable, and, consequently, of little practical interest. When one restricts its attention to traversable wormholes, Casimir energy enters the scene. To simplify calculations, traversable wormholes are tipically assumed to be spherically symmetric and static. The line element of the solution describing such objects can then generally be written as:

$$\mathrm{d}s^2 = -e^{2\Phi(r)}\mathrm{d}t^2 + \frac{1}{1 - b(r)/r}\mathrm{d}r^2 + r^2\mathrm{d}\Omega, \tag{2.69}$$

where $b(r)$ and $\Phi(r)$ are two arbitrary functions, which determine respectively the shape of the wormhole and the gravitational redshift. For this reason, we call $b(r)$ the *shape function* and $\Phi(r)$ the *redshift function*. By a traversable wormhole, we intend an object described by the metric in Equation 2.69 and that guarantees the possibility of causal trips between two asymptotically flat regions (both ways). The form of $\Phi(r)$ and $b(r)$ is then constrained by requiring that there are no horizons (which would prevent two-ways trips), that the trip is safe for human travel (reasonable tidal forces and crossing time) and that the spacetime geometry is sourced by a physically meaningful stress-energy tensor. A careful analysis shows that the first two requirements constraint the shape and redshift functions so much that the stress-energy tensor that sources the wormhole solution is bound to have negative energy density (at least as measured by obeservers travelling at speeds close to c,

but one can choose to make it so for *any* observer) [1]. A natural candidate to provide this negative energy density source is obviously the vacuum state of quantum fields. A possible option would be to source the wormhole by concentrating most of the exotic matter through the throat, but with its density falling off rapidly away from the mouths. Solutions of this type give the so-called Casimir wormholes, and generalisations of the stress energy tensor of a scalar confined by planar plates have been shown to be able to source in principle traversable throats [3]. If one adds the achronal averaged null energy condition into the mix (since, as mentioned earlier, all quantum matter fields are expected to obey it), one further finds the restriction that the throat has to be 'long', meaning that it takes longer to go through the thorat rather than through the ambient space, protecting causality. We now review a very interesting solution recently published, which provides an appealing traversable wormhole sourced by the vacuum energy of Standard Model fields.

2.4.3 A Traversable Wormhole allowed by the Standard Model

The wormhole we now consider, presented by Maldacena, Milekhin and Popov in a recent paper [2], can be seen as given by two extremal magnetically oppositely charged black holes, whose throats, with geometry $AdS_2 \times S^2$ are joined together. This configuration has no event horizon, as required, and is indeed 'long' preserving causality. To build such solution, we exploit the fact that a massless charged fermion gets its cyclotron orbits quantised in the so called discrete Landau levels when immersed in a magnetic field. Being the wormhole traversed by magnetic field lines, this splitting happens inside the throat as well. One can show that the lowest

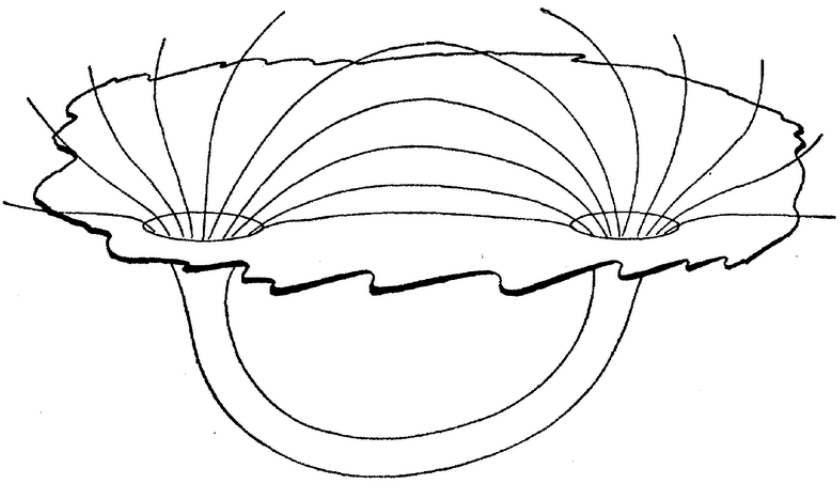


Figure 2.2: Schematic of the magnetic field lines in a wormhole constructed from two oppositely magnetically charged extremal black holes. Reproduced from [33]

Landau level has zero energy on the S^2 direction, while it gives rise to q effective massless 2D fields in the time and radial directions, provided that the integer q (specifying the quantised magnetic flux through the sphere) is big. This large number of fields is what provides the negative energy required to source the wormhole and to make it stable.

We can picture each of these fields as localised along one of the q magnetic field lines that go through the wormhole, constraining the QFT to live on S^1 in the spatial direction, as shown by Figure 2.2. This is crucial, since for a (1+1)-dimensional field on a circular spatial topology the Casimir energy we obtain is indeed negative. In particular, for q complex fermions living on a circle of length L with both right and left moving modes, the vacuum energy is given by:

$$E = -\frac{q}{12}\frac{2\pi}{L}. \tag{2.70}$$

However, our field is not really living on the flat $\mathbb{R} \times S^1$, but rather in AdS_2. We then need to account for the two dimensional conformal anomaly, which turns out to be:

$$E_{\text{anomaly}} = \frac{q}{24}\frac{\pi}{L}. \tag{2.71}$$

We further assume that the two mouths of the throat are very close together with respect to the length of the pipe itself. Doing so, we ignore the contribution coming from the field lines in the region of the ambient space between the two mouths, taking $L = L_{\text{throat}}$. One can show this to be given by:

$$L_{\text{throat}} = \pi l \tag{2.72}$$

where $l = t/\tau$, with t and τ being the time coordinates of respectively the asymptotic Minkowski portion of the manifold and the wormhole region. Putting all together we obtain:

$$E_{\text{wormhole}} = E_{\text{anomaly}} + E = -\frac{q}{8l}, \tag{2.73}$$

which turns out to be negative non-withstanding the positive contribution from the conformal anomaly. For reference, the expectation value of the stress-energy tensor that sources the solution can be proved to be:

$$\langle T_{tt} \rangle = \langle T_{xx} \rangle = -\frac{q}{32\pi^2 l^2 r_e^2}, \tag{2.74}$$

which has already been corrected for the conformal anomaly. Here we have defined:

$$r_e \coloneqq \frac{\pi^2 q^2}{g^2}, \tag{2.75}$$

which can be interpreted as the radius of the entrance of the wormhole, and where

g is the coupling constant of the $U(1)$ gauge field. In order to work out l for the equilibrium configuration of the wormhole, one could directly solve the semiclassical Einstein's equation with the so specified matter content. A faster route, is just to minimise the overall energy of the system given by $E_{\text{tot}} = E_{\text{classical}} + E_{\text{wormhole}}$, where the classical contribution is the energy of the two unconnected near extremal black holes with temperature $T = 1/(2\pi l)$ (since at this temperature the $t = 0$ hypersurface of a near extremal black hole looks similar to the $t = 0$ slice of our wormhole geometry). Adding the classical and quantum energies together we obtain:

$$E_{\text{tot}} = \frac{r_e^3}{l^2} - \frac{q}{8l}, \tag{2.76}$$

which turns out to be minimised by

$$l = \frac{16 r_e^3}{q} = \frac{16\pi^{3/2} q^2}{g^3}. \tag{2.77}$$

Being this the minimum energy configuration we expect the wormhole to eventually settle down to this geometry, no matter in which arrangement it first comes into existence. One might ask how stable this wormhole is against positive energy intakes. The energy gap in the throat can be shown to be $E_{\text{gap}} \propto 1/l \propto 1/q^2$. Furthermore, the binding energy of the throat is of order $1/q$. We then see that quantum fluctuations are not expected to spoil our solution. However, external object can (and most likely will). Indeed, an object that has a Compton wavelength of the order r_e outside the throat has already the required energy to break the two mouths apart, restoring them to the two unconnected black holes. As human travellers would need a r_e much larger than their Compton wavelength to fit through the mouths, we see that the solution really describes a wormhole that can be crossed only by low energy objects and waves without it collapsing.

As the two ends have opposite magnetic charge, they would naturally attract each other, destroying the wormhole. A straightforward way to keep them apart would be to make them orbit each other with some frequency Ω. Including both gravitational and magnetic attraction, one can use the classical Kepler thrid law to show this to be given by:

$$\Omega = 2\sqrt{\frac{r_e}{d^3}}, \tag{2.78}$$

with d being the distance between the two mouths. In order for the wormhole not to break apart, we however need the rotational energy provided by the orbit to be smaller than the binding energy of the throat, meaning $\Omega l \ll 1$. While this can be easily achieved, there is a further effect that might cause the solution to collapse. Indeed, as the mouths rotate, the fermions will feel extra forces, such as the Coriolis force and the quantum Hall force. These forces might push a particle to neighbouring field lines as it goes round, meaning that it could potentially take

more than a cycle to return to the original starting point as the fermion jumps from line to line. This effectively lengthens the loop on which the particles live, lowering the negative Casimir energy contribution and making the wormhole more fragile. Furthermore, because of the rotation of the two mouths (which can be approximated as extremal black holes for all purposes), radiation is bound to play a role. Indeed, both electromagnetic and gravitational radiation lowers the total energy of the system with a rate of:

$$\frac{\mathrm{d}E_{\mathrm{g}}}{\mathrm{d}t} = \frac{2r_e^2 d^4 \Omega^6}{15} \propto \frac{q^5}{d^5}, \qquad \frac{\mathrm{d}E_{\mathrm{EM}}}{\mathrm{d}t} = \frac{2\pi}{3g^2} q^2 d^2 \Omega^4 \propto \frac{q^4}{d^4}. \tag{2.79}$$

Considering only the dominating EM radiation, one gets a finite lifetime for the system of:

$$T_{\mathrm{lifetime}} \propto \frac{d^3}{q^2}. \tag{2.80}$$

This is much larger than the expected time required to go through the wormhole, which, as measured in terms of the asymptotically flat spacetime coordinate t is given by l (hence meeting the traversability condition). One might also worry whether some of this radiation can fall into the mouths and break the configuration, but for the regime of validity of the solution the energy emitted by the orbit is much smaller than the binding energy of the wormhole. While, as we have discussed, the described wormhole is not safe for human travellers, its main achievement is to give a practical example of how the Standard Model fields themselves could source these solutions. Although it is unclear whether two black holes could dinamically merge into such wormhole or not, the reviewed proposal could highlight some of the features that we could encounter if wormholes were to exist in the Universe. This is in contrast with most of the other solutions in the literature, where the source is represented by some artificial Casimir-like cavity [2].

2.5 The Case for Graphene

As mentioned, the study has potentially practical applications beyond the pure theoretical exploration. As an example, we make a case for graphene, the famous 2D material, consisting of a single layer of graphite, that has been the object of extensive research ever since its discovery in 2004 due to its unique properties [34]. Interestingly, its electronic band structure might prove to be the perfect arena to experimentally study the free energy of QFTs on a (2+1)-dimensional space.

2.5.1 Effective Dirac Fermions

The simplest approach to extract the electronic band structure of graphene is, as always, the tight binding model. Here, we treat the electron as bound to a specific

atom (hence the name), but with some probability (parametrised by the so called hopping integral t) to being transferred to the nearest neighbours only. Because of the honeycomb structure of graphene, we see that we can actually interpret the hexagonal lattice as being composed by two sublattices, that we denote A and B, as shown by Figure 2.3. The three nearest neighbours to an atom belonging to the A-sublattice at $\vec{x}$ are given by $\vec{x} + \vec{\delta}_i$ (for $i = 1, 2, 3$) with:

$$\vec{\delta}_1 = \frac{a}{2}(1, \sqrt{3}), \qquad \vec{\delta}_2 = \frac{a}{2}(1, -\sqrt{3}), \qquad \vec{\delta}_3 = -a(1, 0), \tag{2.81}$$

where $a \approx 1.42$Å is the lattice C-C spacing. For the case of an atom belonging to the B-sublattice, the nearest neighbours would instead be positioned at $\vec{x} - \vec{\delta}_i$. With this in mind, the general tight binding Hamiltonian can be written as:

$$\hat{H} = -t \sum_{\vec{x},n,\sigma} \hat{a}^{\dagger}_{\vec{x},\sigma} \hat{b}_{\vec{x}+\vec{\delta}_n,\sigma} + h.c., \qquad n = 1, 2, 3 \tag{2.82}$$

where $\hat{a}_{\vec{x},\sigma}$ and $\hat{b}_{\vec{x},\sigma}$ are the annihilation operators of an electron with spin σ around the atom at position x on respectively the A and B sublattices. It is, however, more convenient to work with the Hamiltonian in k-space. Expanding in Bloch waves we have:

$$\hat{a}_{\vec{x}} = \sum_{\vec{k} \in \mathrm{BZ}} e^{i\vec{k}.\vec{x}} \hat{a}_{\vec{k}}, \qquad \hat{b}_{\vec{x}} = \sum_{\vec{k} \in \mathrm{BZ}} e^{i\vec{k}.\vec{x}} \hat{b}_{\vec{k}}, \tag{2.83}$$

where the sum is over all wavevectors $\vec{k}$ belonging to the 1$^{\mathrm{st}}$ Brillouin zone of the lattice. Recall that the expansion in Bloch waves is allowed because of the (discrete) translational symmetry of the lattice. Plugging those in and representing a state with a 2D vector where in the top slot we have the amplitude on the A sublattice while on the bottom one the amplitude on the B sublattice (basically we treat the full Hilbert space of the electrons on graphene as a tensor product between the Hilbert space on the A-sublattice and the one on the B-sublattice), it can be shown

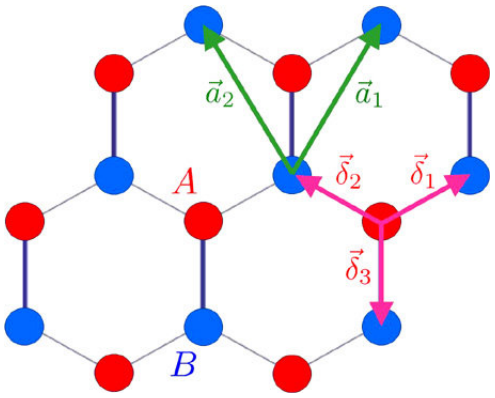


Figure 2.3: Graphene honeycomb lattice, reprinted from [35]

that the Hamiltonian for graphene in k-space is given by:

$$\hat{H}_{\vec{k}} = \begin{pmatrix} 0 & \Delta_{\vec{k}} \\ \Delta^*_{\vec{k}} & 0 \end{pmatrix}, \tag{2.84}$$

with

$$\Delta_{\vec{k}} = -t\left[1 + 4\cos\left(\frac{3k_x a}{2}\right)\cos\left(\frac{\sqrt{3}k_y a}{2}\right) + 4\cos^2\left(\frac{\sqrt{3}k_y a}{2}\right)\right]^{-\frac{1}{2}}. \tag{2.85}$$

The energy bands are then trivially given by the dispersion relation $E_{\vec{k}} = \pm|\Delta_k|$. A very important feature of the bands is that the positive and negative levels meet at the so called *Dirac points*, i.e. where $\Delta_{\vec{k}} = 0$. This happens at the 6 corners of the hexagonal first Brillouin zone, which fall into two inequivalent groups of 3 (much like the 2 sublattices). We take as representatives of the two groups the points in momentum space:

$$\vec{K} = \frac{2\pi}{3a}\left(1, \frac{1}{\sqrt{3}}\right), \qquad \vec{K}' = \frac{2\pi}{3a}\left(1, -\frac{1}{\sqrt{3}}\right). \tag{2.86}$$

To understand the reason behind the name 'Dirac points' for $\vec{K}$ and $\vec{K}'$, let's expand Equation 2.85 near $\vec{q} \coloneqq \vec{k} - \vec{K} = 0$ to obtain:

$$\Delta_{\vec{q}} \approx \frac{\sqrt{3}}{2}|\vec{q}|e^{-i(\pi/2+\theta)}, \tag{2.87}$$

with θ being the angle that $\vec{q}$ makes with the k_x-axis in the $k_x - k_y$ plane. We then see that near the Dirac points the disperion relation is linear in $|\vec{q}|$, with the bands inheriting a conical shape in this region. We consequently call this portion of the bands *Dirac cone*. More interestingly, plugging Equation 2.87 into Equation 2.84 we see:

$$\hat{H}_{\vec{q}} \approx -\frac{\sqrt{3}}{2}at\begin{pmatrix} 0 & q_x - iq_y \\ q_x + iq_y & 0 \end{pmatrix} = v_F\ \vec{\sigma}.\vec{q}, \tag{2.88}$$

with $\vec{\sigma}$ being the formal vector of the usual Pauli matrices (note that we are working in 2D here, so $\underline{\underline{\sigma_z}}$ is not included) and v_F is the Fermi velocity in graphene. Experimentally, it has been found that v_F is not really a constant as this treatment suggests, but by introducing space-dependent strain into the model it is possible to obtain a variable Fermi velocity that matches the experimental observations. What is really curious, however, is that Equation 2.88 is just the Dirac equation for a massless fermion living in a (2+1)-dimensional spacetime (where $c \to v_f$)! Furthermore, a similar result would have been obtained by expanding the Hamiltonian near $\vec{K}'$ instead, with $\Delta_{\vec{q}|\vec{K}'} = \Delta^*_{\vec{q}|\vec{K}}$. The physical interpretation is that we have two Dirac Hamiltonians, one for each inequivalent set of cones, with the helicity of the

effective particle living near $\vec{K}'$ reversed with respect to the one near $\vec{K}$. We then see that electrons occupying the Dirac cones behave for all purposes as massless Dirac fermions (at each inequivalent Dirac point we obviously have particles and antiparticles, corresponding respectively to electrons and holes). It is also interesting to look at the eigenfunctions of the Hamiltonian near the Dirac points, which are given by:

$$\psi^{\pm}_{\vec{q}|\vec{K}} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\theta/2} \\ \pm e^{-i\theta/2} \end{pmatrix}, \tag{2.89}$$

where the + state corresponds to a mode with positive energy (that we interpret as the electron) while the − state corresponds to a mode with negative energy (that we interpret as the hole). Of course, upon second quantisation, we would lose the disturbing negativity of the energy for the antiparticle. Staring at Equation 2.89 we notice two compelling features. First, as expected for fermions, the wavefunctions exhibit a 4π period on θ, meaning that they pick up a minus sign after a 2π rotation in q-space. Secondly, we see that it is a different phase between the modes on the A and B sublattices that leads us to interpret the particle as hole or electron [36].

It would be natural to ask how close to reality this model really is. Would the effective Dirac Hamiltonian still arise as we make the description more complex and realistic? As it turns out [36], the result does not change if we include, for example, second nearest neighbour hopping. Furthermore, including space dependent strain does not alter the Dirac-like behaviour of the electron near the Dirac points [37]. Indeed, Dirac points in graphene have been experimentally verified [38], demonstrating the high level of stability that these features possess. While we have mentioned only graphene so far, other 2D materials such as silicene, germanene, boron and MoS2 could theoretically exhibit Dirac cones, with compelling evidence in silicon found in recent works [39].

2.5.2 Stability of the Membrane

The study of the free energy of a quantum field living on a (2+1)-spacetime can then be used to elucidate the stability of certain 2D membranes, such as a layer of graphene. Indeed, a geometrical arrangement of a lattice is stable if such configuration minimises the free energy of the material. For the case of graphene near a Dirac cone, F has a two-fold origin: a classical contribution F_c, given by the Landau free energy of the geometry, and a quantum contribution F_q which originates from the effective massless Dirac fermions living on the lattice.

We've already discussed the free energy of a quantum field (a scalar, actually, but the story for fermions is quite similar modulo some subtleties about how they couple to the curved spacetime). Let's then give a brief overview on the nature of the Landau energy, and what this looks like for the case of graphene. Landau theory is an

effective theory that describes what happens to a system near a critical point. In our case, the critical point we are interested in represents the transition point between a 'flat' and crumpled phase of the lattice. It was formulated by Lev Landau, who built on the experimental evidence that different systems exhibit similar features near phase transitions, suggesting that a universal description could be achieved. The simplest and most illustrative example of the efficacy of Landau theory is given by the description of the phase transition in the Ising model through the mean field approximation. Landau's approach can be summarised by the assumption that, near phase transition, the free energy of the system is analytic, meaning that it can be expressed as a Taylor expansion in terms of a parameter governing the physics, called the *order parameter*. The only other restriction placed on the form of the free energy is that it must follow the symmetries of the Hamiltonian of the system.

It is interesting to see how the classical free energy changes as we deform a spherical lattice, since this gives the competing energy that might balance the change in the Helmholtz energy of the quantum field living on it. We assume diffeomorphism invariance, even though this is not technically valid for a crystalline material such as graphene due to its discrete symmetries. Moreover, we take as sole contribution to the Landau energy the bending free energy of the material, which we assume to be minimised by the sphere (at least locally, compared to fluctuations around it). Hence, to second order, the difference in bending free energy between a geometry with mean curvature H and bending rigidity κ, and sphere of radius r_0, is given by [40]:

$$\Delta F_c = \kappa \int \mathrm{d}^2x \sqrt{h} \left(H - \frac{2}{r_0} \right)^2 , \tag{2.90}$$

with g_{ij} being the metric with which Σ, the 2D spatial sub-manifold with sphere topology, is endowed. Note that, in some conventions, H is given by the average of the principal curvatures, but here we take it to be just the sum. Consider now a surface slightly deformed from a sphere of radius r_0, parametrised by

$$r = r_0[1 + \epsilon\chi(\theta, \phi)], \tag{2.91}$$

where χ is the deformation function and ϵ is a small parameter. The metric induced by flat space on such Σ is then given by:

$$\mathrm{d}s^2_\Sigma = r_0^2(1 + 2\epsilon\chi)\mathrm{d}\Omega^2 + \mathcal{O}(\epsilon^2), \tag{2.92}$$

where $\mathrm{d}\Omega^2$ is the metric on the unit sphere (we show this later, but one can understand its origin by recalling that any 2D metric can be written in terms of the round sphere metric multiplied by an appropriate Weyl exponential factor). Inserting the determinant of the metric and the mean curvature evaluated for Equation 2.91 in

Equation 2.90, we get to leading order in ϵ:

$$\Delta F_c = \epsilon^2 \kappa \int \mathrm{d}\theta \mathrm{d}\phi \left(2\chi + \overline{\nabla}^2 \chi)\right)^2 + \mathcal{O}(\epsilon^3), \tag{2.93}$$

with $\overline{\nabla}^2$ being the Laplacian on the unit sphere. It is then useful to decompose ξ using the spherical harmonics $Y_{l,m}(\theta, \phi)$:

$$\chi(\theta, \phi) = \sum_{l,m} \chi_{l,m} Y_{l,m}(\theta, \phi), \tag{2.94}$$

where $\chi_{l,m}$ are coefficients and $|m| < l$. Mathematically, the fact that deforming the lattice does not increase its surface area is enforced by setting $\chi_{0,0} = 0$. Then, plugging the expansion back in Equation 2.93, we see that the leading order change on the Landau energy caused by the perturbation is given by:

$$\Delta F_c = \epsilon^2 \kappa \sum_{l,m} |\chi_{l,m}|^2 (l-1)^2 (l+2)^2. \tag{2.95}$$

To understand the balance between classical and quantum free energy in such setup, we consider the example of a CFT living on the lattice. This is somewhat realistic, as the electrons near the Dirac cones do behave as excitations of a CFT if the graphene layer is flat or slightly perturbed away from it. We assume that this description is valid for a spherical arrangement as well (as it makes the discussion that follows easier). It can be shown that the difference in the zero-temperature Helmholtz free energy of the CFT (i.e. the Casimir energy) between the spherical lattice and the perturbed one is [11]:

$$\Delta F_q = -\epsilon^2 \frac{\pi^2 c_T v_F}{48 r_0} \sum_{l,m} |\chi_{l,m}|^2 \frac{(l^2-1)(l+2)}{l} \left(\frac{\Gamma\left(\frac{l+1}{2}\right)}{\Gamma\left(\frac{l}{2}\right)} \right)^2 + \mathcal{O}(\epsilon^3), \tag{2.96}$$

where v_F is the Fermi velocity of the material acting as effective speed of light for the relativistic degrees of freedom living on the lattice and c_T is the central charge of the theory (defined as the coefficient of the two point function of the stress tensor of the CFT). In the case of a Dirac CFT, it takes the value of $c_T = 3/(32\pi^2)$.

At this point, we can see two key differences between Equation 2.95 and Equation 2.96. First of all, in Equation 2.95 the length scale r_0 dropped out, meaning that it is insensitive to the size of the sphere. This is not true, however, for the quantum contribution, which is proportional r_0^{-1}, suggesting that for very small r_0 it might dominate the free energy (note, however, that the analysis is valid only for length scales well above the lattice spacing where the discrete nature of the crystal is masked). Secondly, we note that the classical contribution is proportional to l^4 for big l, while Equation 2.96 grows as l^3 (note that a further l factor comes from

the ratio of the Γ-functions), indicating that for big enough l the positive classical Landau energy is the leading term. Furthermore, the surface area preserving condition imposes $\chi_{00} = 0$, meaning that only $l > 1$ contribute. However, the $l = 1$ term vanishes as well, since those deformations are equivalent to infinitesimal diffeomorphism (under which we have assumed our systems to be invariant). Note that close to $l = 1$ the quantum contribution vanishes linearly in l while the classical Landau energy does so quadratically. This means that there is some region with $l > 1$ where $\Delta F = \Delta F_c + \Delta F_q < 0$ (taking l continuous). Nonetheless, l must be an integer, meaning that to effectively see the free energy go negative it is necessary to have this holding all the way up to $l = 2$. Summing Equation 2.95 and Equation 2.96 and choosing a function with $\chi_{l,m} = 0$ for $l \neq 2$, we then see that we achieve a negative ΔF for:

$$r_0 \leq \frac{3\pi^3 c_T v_F}{1536\kappa}. \tag{2.97}$$

In the case of graphene, we tipically have $\kappa \approx 1$ eV and $v_F \approx 10^6$ m/s. Furthermore, recall that we need to multiply the central charge of the Dirac CFT by 2 since we really have 2 inequivalent Dirac points (hence two Dirac fields living on the lattice). Reinstating the factor of $\hbar$, we see that for the negative CFT contribution to dominate we would need $r_0 \lessapprox 0.01a$, with a being the lattice spacing in graphene. Clearly an unphysical result, highlighting that for a CFT leaving on graphene it is always the case that the sphere is a stable configuration. One could nonetheless conceivably engineer the membrane by modifying its Fermi velocity [37] and bending rigidity [41] in such a way that for a reasonably sized sphere the quantum energy dominates over the classical contribution, making the the spherical geometry unstable.

It is also interesting to note that the quantum contribution seems to indicate that the CFT prefers to crumple rather than to stay on the sphere. Is this behaviour universal for all quantum fields and valid for all deformations? We'll discuss recent analytical results in the next chapter, but all evidence seems to point in that direction. Numerical studies seem to support this tendency also for deformation non treatable with pen and paper, and we come to a similar conclusion with the numerical results we obtained.

Chapter 3

Setup and Previous Results

In this section, we formalise the problem of computing the free energy of a scalar field on a deformed sphere. We first discuss the Laplacian on the sphere, and how this changes as we deform the geometry. We then tackle the computation of the Helmoltz energy analytically, discussing how to regulate the divergences that naturally arise when calculating the zero point energy of the field and presenting the form of the heat kernel for this particular setup. We then review recent findings on the topic, both analytical (for a strongly coupled CFT) and numerical (fermion and scalar field on axysimmetrical deformed spheres).

3.1 Geometry of the Physical Space

We focus on a static (2+1)-dimensional geometry $\mathbb{R} \times \Sigma$, where Σ is a 2D manifold with sphere topology. The general line element in such geometry can then be written as:

$$\mathrm{d}s^2 = \mathrm{d}\tau^2 + g_{ij}(\vec{x})\mathrm{d}x_i\mathrm{d}x_j, \tag{3.1}$$

where $\tau = it$ is the Wick rotated time coordinate, $i, j = 1, 2$ and $g_{ij}(\vec{x})$ is the spatial metric on Σ that depends only on the spatial components of x. In particular, we want to consider Σ to be some kind of deformed sphere, which may (or may not) be embeddable in a 3D ambient space. We make use of the results that any 2D metric is conformally flat and that the n-dimensional round sphere S^n is conformally flat,

to consider, without loss of generality, the metric g on Σ to be:

$$g_{ij} = e^{2f}\bar{g}_{ij}, \tag{3.2}$$

where $\bar{g}_{ij}$ is the metric on S^2 and $f(\theta, \phi)$ is a smooth function that represents our deformation. Recall that for the unit 2-sphere the line element is given by:

$$\mathrm{d}s^2_{S^2} = \mathrm{d}\theta^2 + \sin\theta \mathrm{d}\phi^2. \tag{3.3}$$

3.1.1 Laplacian on a Deformed 2-Sphere

In order to perform the numerical calculations, we need an explicit expression for the Laplacian on Σ. As it turns out, it is straightforward to relate the Laplacian of two conformally equivalent metrics. Recall that the Laplace operator acting on a scalar field in curved space is given by:

$$g^{ij}\nabla_i\partial_j\phi = g^{ij}\left(\partial_i\partial_j - \Gamma^k_{ij}\partial_k\right)\phi, \tag{3.4}$$

with Γ^k_{ij} being the Christoffel's symbols giving components of the Levi-Civita metric connection on Σ. With some algebra, it can be readily shown that for two metrics related by a conformal factor as in Equation 3.2, the components of the Levi-Civita connection are given by:

$$\Gamma^k_{ij} = \bar{\Gamma}^k_{ij} + \delta^k_i\partial_j f + \delta^k_j\partial_i f - \bar{g}_{ij}\bar{\partial}^\alpha f, \tag{3.5}$$

where $\bar{\Gamma}^k_{ij}$ are the components of the connection for $\bar{g}_{ij}$ and $\bar{\partial}^\alpha$ is just the partial derivative with the index raised by $\bar{g}^{ij}$. Plugging Equation 3.5 into Equation 3.4 and using $\delta^i_i = 2$ (since we are working on a 2D manifold), we see that the extra terms cancel and we are left with:

$$\nabla^2\phi = e^{-2f}\bar{g}^{ij}(\partial_i\partial_j - \bar{\Gamma}^k_{ij}\partial_k)\phi = e^{-2f}\overline{\nabla}^2\phi, \tag{3.6}$$

i.e. the two Laplacians are related by the inverse of the conformal factor. Furthermore, one can show that a similar relationship holds for the Ricci scalar [42]:

$$R = e^{-2f}\bar{R} - 2(n-1) - 2f\overline{\nabla}^2 f - (n-2)(n-1)e^{-2f}|\mathrm{d}f|^2, \tag{3.7}$$

with $\mathrm{d}f$ indicating the exterior derivative acting on the deformation function and n the dimensions of the manifold. For the case of $\bar{g}_{ij}$ being the metric on the 2-sphere (and, consequently, $n = 2$) this reduces to:

$$R = 2e^{-2f}(1 - \overline{\nabla}^2 f), \tag{3.8}$$

where we have used the fact that $\bar{R} = 2$. For reference, recall that the Laplacian on the unit sphere is given by:

$$\overline{\nabla}^2 = \partial_\theta^2 + \frac{1}{\sin^2\theta}\partial_\phi^2 + \cot\theta\partial_\theta. \tag{3.9}$$

3.1.2 Embedding of Σ

Not all geometries defined by Equation 3.2 can be embedded in $\mathbb{R}^3$, and can be then physically seen as properly deformed sphere. Nonetheless, we now consider an embeddable 2D surface slightly perturbed away from the unit sphere, illustrating that the metric induced by flat Euclidean space on Σ is indeed of the form shown in Equation 3.2. The motivation behind this excursion is to visualise the physical implications of different f, so that the reader can have a clearer picture of the type of deformations we are considering.

Consider a surface defined by

$$r = R(\theta, \phi) \coloneqq c[\chi, \epsilon](1 + \epsilon\chi(\theta, \phi)), \tag{3.10}$$

where r is the radial coordinate in $\mathbb{R}^3$, ϵ is a numerical parameter that governs the strength of the deformation described by χ and $c[\chi, \epsilon]$ is a functional that acts as a normalisation for the surface area of Σ. In particular, c is chosen such that for any deformation the surface area of the geometry described by $R(\theta, \phi)$ is 4π, i.e. the one of the unit sphere. Note that for Equation 3.10 to describe a valid embedding, we must require $R(\theta, \phi) \geq 0$, meaning that, for each χ, there is a different range of values ϵ can take, i.e. $\epsilon \in [\epsilon_{\min}, \epsilon_{\max}] \subset \mathbb{R}$. Now, recall that an m-dimensional manifold $(\mathcal{M}, g)$ induces a metric $g_\mathcal{N} = h^* g$ on an embedded submanifold $\mathcal{N}$, where the embedding is defined by the smooth map $h : \mathcal{N} \to \mathcal{M}$, via the pull-back h^* (i.e. a map induced by h that allows us to to map the cotangent space on $\mathcal{N}$ to $\mathcal{M}$). Now, taking h to be defined by $x = x(y)$ (where $\{x^\mu\}$ are the coordinates on $\mathcal{M}$ and $\{y^\alpha\}$ are the coordinates on the embedded submanifold), we have that the induced metric on $\mathcal{N}$ is given by:

$$(h^* g)_{\alpha\beta} = \frac{\partial x^\mu}{\partial y^\alpha}\frac{\partial x^\nu}{\partial y^\beta} g_{\mu\nu}. \tag{3.11}$$

For the case of the embedding shown in Equation 3.10 we have $\{x^\mu\} = \{r = R(\theta, \phi), \theta, \phi\}$ and $\{y^\alpha\} = \{\theta, \phi\}$, leading to the line element:

$$\mathrm{d}s_\Sigma^2 = (R'^2 + R^2)\mathrm{d}\theta^2 + (\dot{R}^2 + R^2\sin^2\theta)\mathrm{d}\phi^2 + 2\dot{R}R'\mathrm{d}\theta\mathrm{d}\phi, \tag{3.12}$$

where a dot and a prime indicate respectively partial derivative with respect to ϕ and θ. For illustration purposes it is now easier to work with axisymmetric deformation (which avoids us the trouble of diagonalising the metric). This amounts to setting

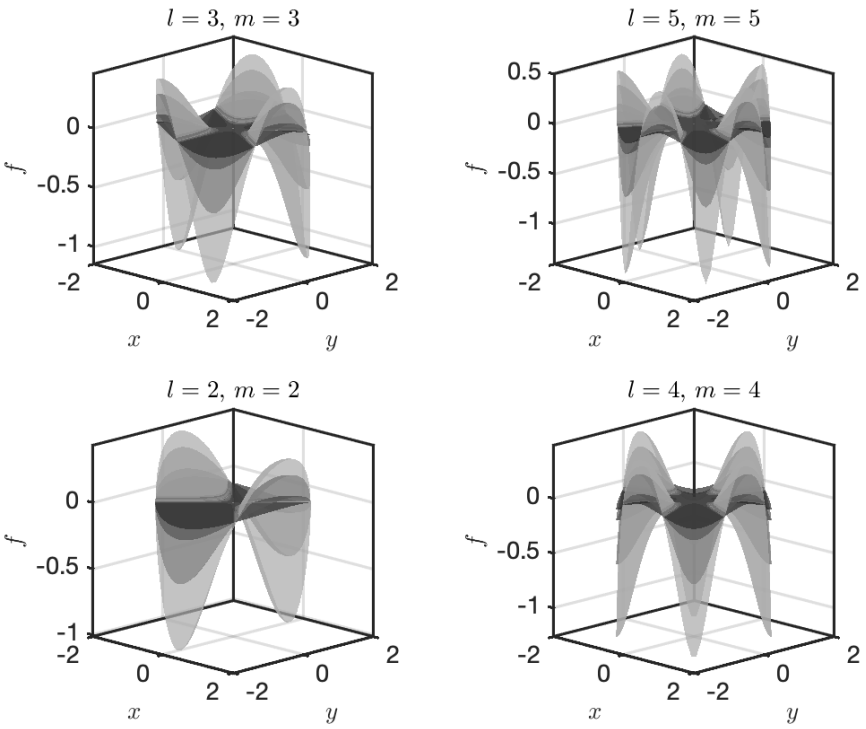


Figure 3.1: Comparison of the deformation function for different ϵ on the upper hemisphere (lighter surface corresponds to bigger ϵ). The x and y Cartesian coordinates correspond to the ones obtained from a polar system with $0 \leq \theta \leq \pi/2$ treated as radial coordinate and $0 \leq \phi \leq 2\pi$ as azimuthal angle.

$\dot{R} = 0$, obtaining:

$$\mathrm{d}s^2_\Sigma = R^2 \left[\left(\frac{R'^2}{R^2} + 1 \right) \mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\phi^2 \right]. \tag{3.13}$$

Finally, we see that keeping only terms up to order ϵ (i.e. considering small deformations), and choosing:

$$f = \ln R = \ln \left[c_{\chi,\epsilon} (1 + \epsilon \chi(\theta, \phi)) \right] \tag{3.14}$$

we do recover the conformal gauge in Equation 3.2 for $\epsilon \ll 1$ (even though higher order terms break the conformal form of the metric). It is then clear what we mean when we say that f defines the deformation of the sphere: with the prescription given in Equation 3.14, we see that $r = e^f$ defines for small ϵ the same embedding of the perturbed sphere given by Equation 3.10.

Later, in our numerical analysis, we focus on non-axisymmetric deformations. We take the function f to be of the form of Equation 3.14, and we choose:

$$\chi^{(l,m)}(\theta,\phi) = \begin{cases} \sqrt{2}\, Y_{l,m} & m = 0 \\ (-1)^m Y_{l,m} + Y_{l,-m}, & |m| > 0 \end{cases} \tag{3.15}$$

With these choice of normalisation factors, we both guarantee reality of χ and obtain deformations of comparable size between axisymmetric and non-axisymmetric distortions (as one can check by looking back at the perturbative results for a CFT in Equation 2.95 and Equation 2.96). Furthermore, as we discuss later, we consider only perturbations that are even with respect to the equatorial line $\theta = \pi/2$. This allows us to solve the problem on half of the sphere only, in order to reduce the computational cost of the numerical calculations. For the deformations to enjoy this symmetry, we exploit the spherical harmonics property that:

$$Y_{l,m}(\theta,\phi) = (-1)^{l+m} Y_{l,m}(\theta - \pi, \phi), \tag{3.16}$$

from which we see that the deformations with the required reflection invariance with

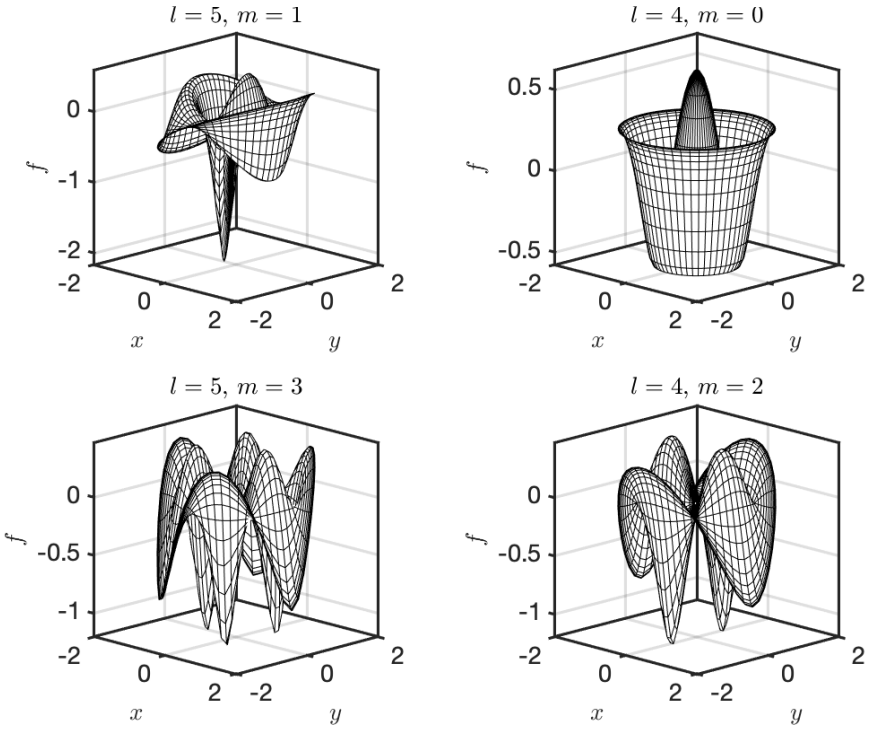


Figure 3.2: Deformation function for different l and m on the upper hemisphere. The x and y Cartesian coordinates correspond to the ones obtained from a polar system with $0 \leq \theta \leq \pi/2$ treated as radial coordinate and $0 \leq \phi \leq 2\pi$ as azimuthal angle.

respect to the $x-y$ plane have the same parity for l and m. Figure 3.1 shows some examples of such non-axisymmetric deformation functions, illustrating how these change as ϵ increases. A comparison between deformations with same l but different m is instead shown in Figure 3.2, where we also see an example of axisymmetric distortion, i.e. the one with $l = 4$ and $m = 0$.

3.2 Scalar Field on a Deformed Sphere

We now consider a scalar field living on our spacetime. We first analyse the UV and IR divergences of the theory, showing how these are manifest in the form of the heat kernel. We then discuss how we can extract finite physical quantities that can be calculated numerically. Successively, we introduce previous analytical and numerical results on the subject, exploring the sign of the difference in free energy between deformed and undeformed sphere in various regimes. We describe that small deformations of the sphere always reduce the free energy for a non-interacting scalar field, and that the same is valid for deformations of any size for a strongly interacting CFT. Finally, we discuss the sign of the heat kernel in different t_k regimes.

3.2.1 UV and IR Divergences

All the physics of our theory is as always encapsulated by the Euclidean (since we have Wick rotated time) action:

$$S_E[\phi] = \frac{1}{2}\int \mathrm{d}\tau \int \mathrm{d}^2x\sqrt{g}\phi(M^2 + \xi R - \Box)\phi, \tag{3.17}$$

where we have used the fact that the metric is static in order to isolate the τ integration. As usual, the free energy is divergent. We can however introduce counterterms in the action to absorb those infinities, regulating our theory using a UV momentum cutoff Λ. On dimensional grounds, we can immediately see that the most general counterterms that are covariant (and that we can therefore add to the action) are:

$$c_1\Lambda^3, \qquad c_2\mu\Lambda^2, \qquad (c_3\mu^2 + c_4 R)\Lambda, \tag{3.18}$$

where μ is the sum of any parameter in Equation 3.17 with mass dimension 1, while the c_i are dimensionless coefficients independent on both Λ and the geometry. With counterterms, the action reads:

$$S_{\text{c.t.}}[\phi] = S_E[\phi] + \int \mathrm{d}\tau \int \mathrm{d}^2x\sqrt{g}\left[c_1\Lambda^3 + c_2\mu\Lambda^2 + (c_3\mu^2 + c_4 R)\Lambda\right]. \tag{3.19}$$

We can now perform the spacetime integration on the counterterms, noting that the first three terms are just constants in x. We make use of the known integrals:

$$\int \mathrm{d}^2x\sqrt{g} = \mathrm{Vol}[g] \tag{3.20}$$

and

$$\int \mathrm{d}^2x\sqrt{g}R = 4\pi\chi_\Sigma, \tag{3.21}$$

where the χ_Σ is the Euler characteristic of Σ, a topological invariant that for geometries with sphere topology is given by $\chi_\Sigma = 2$. This shows that the regulated free energy of the field takes the form (dropping the q subscript on F_q since we only focus on the quantum contribution from now on):

$$F = \mathrm{Vol}[g](c_1\Lambda^3 + c_2\mu\Lambda^2 + c_3\mu^2\Lambda) + 4\pi\chi_\Sigma\Lambda + F_{\text{fin}}. \tag{3.22}$$

Here F_{fin} is the finite contribution to the free energy as $\Lambda \to \infty$. Therefore, the difference in free energy between fields living on distinct Σ, but with same topology and surface area, is actually a finite quantity! Hence, from now on, we renormalise the free energy by considering $\Delta F = F_\Sigma - F_{\Sigma=S^2}$, i.e. taking as control the field living on the unit 2-sphere with standard S^2 metric $\bar{g}$. We also choose Σ such that $\mathrm{Vol}[g] = \mathrm{Vol}[\bar{g}]$ (note that this also conveniently replicates what happens when we deform an actual lattice, as mentioned earlier).

In order to compute the free energy, we resort to the heat kernel method. The reasond behind this choice is two-fold: first of all it makes the UV divergences manifest in the analytical treatment, and secondly it allows for numerical computations (contrary to ζ-function regularisation, where the implementation of the analytic continuation as a numerical task is nontrivial). For a scalar field described by Equation 3.17 we obtain the following heat kernel integral representation for the Helmholtz free energy:

$$F = -\frac{T}{2}\int_0^\infty \mathrm{d}t_k \frac{K_D(t_k)}{t_k}, \tag{3.23}$$

with $D = -\partial_\tau^2 + L + M^2$ being the differential operator acting on the field. Here we have defined $L = -\nabla^2 + \xi R$, with ∇^2 the Laplacian on the spatial Σ geometry. From Equation 3.8 we see also see that L for the non-minimally coupled scalar is explicitly given in terms of the deformation function f by:

$$L = e^{-2f}\left[-\bar{\nabla}^2 + 2\xi\left(1 - \overline{\nabla}^2 f\right)\right]. \tag{3.24}$$

Crucially, we can extract the contributions due to thermal effect and mass of the field, and rewrite F in terms of the heat kernel K_L associated with L instead:

$$F = -\frac{T}{2}\int_0^\infty \mathrm{d}t_k\ e^{-M^2t_k}\Theta_{-1/2}(T^2t_k)\frac{K_L(t_k)}{t_k}, \tag{3.25}$$

where

$$\Theta_\sigma(x) \coloneqq \sum_{n=-\infty}^{\infty} e^{-(2\pi)^2(n-\sigma+1/2)^2 x}. \tag{3.26}$$

From Equation 3.23 we see that there are two natural regimes where the divergent contributions to F can come from: $t_k \ll 1$ and $t_k \to \infty$. Let's briefly analyse the latter. As mentioned earlier, we can formally express the heat kernel for L as:

$$K_L(t_k) = \mathrm{Tr}\left(e^{-Lt_k}\right) = \sum_i e^{-\lambda_i t_k}, \tag{3.27}$$

with λ_i being the ith eigenvalue of L and the sum being over all such eigenvalues. It is evident that if L were to have zero or negative eigenvalues, then the integral in Equation 3.23 would diverge in the $t_k \to \infty$ regime, resulting in tachyonic instabilities if, when the mass term is accounted for, such eigenvalue amounts to a negative effective mass. To be more precise, the presence of the vanishing eigenvalue would only cause divergence for $M = 0$, a typical feature of IR divergences. Nonetheless, we know that for the minimally coupled case we avoid negative eigenvalues due to the conformal form and to the properties of the Laplacian on S^2. $\lambda = 0$, however, is indeed an eigenvalue of the Laplace operator for both the deformed and underformed sphere when $\xi = 0$ since the constant function is killed by both. However, consider the difference between the free energy of Σ and of the unit sphere, which from now on we refer to as *subtracted* free energy. This is trivially given in terms of the respective heat kernels by:

$$\Delta F = T\sigma \int_0^\infty \mathrm{d}t_k \; e^{-M^2 t_k} \Theta_\sigma(T^2 t_k) \frac{\Delta K_L(t_k)}{t_k}, \tag{3.28}$$

where we have introduced $\sigma = -1/2$ to adopt the same notation of [9] and we have defined the subtracted heat kernel

$$\Delta K_L(t_k) \coloneqq K_L(t_k) - \overline{K}_L(t_k) = \sum_i \left(e^{-\lambda_i t_k} - e^{-\bar{\lambda}_i t_k}\right), \tag{3.29}$$

with $\overline{K}_L$ being the heat kernel for the scalar on the control geometry. The integral in Equation 3.28 is now regulated in the big t_k limit, as the contributions from the zero eigenvalues cancel in Equation 3.29. For the non-minimally coupled case, instead, all eigenvalues are positive for vanishing deformations. Indeed, the eigenvalues for the non-minimal Laplacian on the unperturbed sphere are given by:

$$\lambda = l(l+1) + 2\xi, \qquad l = 0, 1, 2, ... \tag{3.30}$$

Nonetheless, inspecting Equation 3.6 we see that the perturbations can potentially cause at some point a sign change for at least the lowest eigenvalue, leading to the catastrophic divergences. We will then have to carefully choose the deformation

function when considering the non-minimally coupled scalar. Even without IR infinities, UV divergences would still force us to consider ΔF rather F itself! To see this more explicitly, consider the small t_k behaviour of K_L, where those divergences are encapsulated. As mentioned earlier, there is a theorem that guarantees the existence (and crucially uniqueness) of the heat kernel on a compact n-dimensional Riemannian manifold $\mathcal{M}$ without boundary (such as Σ) for any elliptic differential operator $\mathcal{D}$. It can also be proved that such heat kernel admits a small t_k asymptotic expansion:

$$\mathcal{K}_{\mathcal{D}}(t_k, x, x) = t_k^{-n/2}(\alpha_0(x) + \alpha_1(x)t_k + \alpha_2(x)t_k^2 + ...), \qquad t \to 0^+, \tag{3.31}$$

where $\alpha_i(x)$ are smooth functions on the manifold $\mathcal{M}$. One can then integrate this over $\mathcal{M}$ to get:

$$K_{\mathcal{D}}(t_k) = t^{-n/2}\sum_{i=0}^{\infty} a_i t^i, \tag{3.32}$$

where

$$a_i = \int_{\mathcal{M}} \alpha_i. \tag{3.33}$$

As it turns out, the coefficients a_i are locally computable in terms of geometric invariants [24]. Precisely for this reason, and because one cannot construct odd-dimensional topological invariants for manifolds without boundaries, we have that the odd-numbered coefficients vanish:

$$a_{2i+1} = 0. \tag{3.34}$$

For the specific case of the scalar field on the 2D geometries under consideration, i.e. for $\mathcal{D} = D$ it can be shown that this reduces to [9]:

$$K_D(t_k) = \frac{1}{\sqrt{4\pi}t^{3/2}T}\sum_{n=0}^{\infty} b_{2n} t^n, \tag{3.35}$$

where the new coefficient b_{2n} can still be expressed in terms of objects depending on the geometry. In particular, we have:

$$b_0 \propto \int \mathrm{d}^2x\sqrt{g} = \mathrm{Vol}[g], \qquad b_2 \propto \int \mathrm{d}^2x\sqrt{g}R = 4\pi\chi_\Sigma, \tag{3.36}$$

which we saw in our dimensional arguments are the terms that control the UV divergences! It is in this sense that we mean that small t_k regime governs the high energy behaviour of our theory, while the big t_k encapsulates the IR physics. The next order coefficients are not equal for all area preserving Σ. To see more explicitly that these two contributions do indeed give UV divergences, note that, when plugged into Equation 3.28, the integral of the heat kernel expanded up to $n = 2$ blows up

in the $t \to 0$ limit. However, the third term (i.e. the one governed by the b_4 coefficient) goes as $t^{-1/2}$, meaning that its integral does give a finite contribution even in the $t_k = 0$ limit. Since the two coefficients in front of the divergent terms are proportional to the surface area of Σ and to a topological invariant, we again see that considering ΔF for area and topology preserving deformations of the sphere leads to a UV finite quantity. Extracting the dependence on temperature and mass, the short time asymptotics of the subtracted heat kernel for the reduced operator L is given by:

$$\Delta K_L(t_k) = t \sum_{n=0}^{\infty} \Delta b_{2n+4} t^n, \tag{3.37}$$

with Δb_i being the difference of the heat kernel coefficients of between the deformed and undeformed sphere (note we used $\Delta b_0 = \Delta b_2 = 0$). Therefore, the leading order behaviour for the subtracted kernel in the small t_k is linear and corresponds to

$$\Delta K_L(t_k) = t\Delta b_4 + \mathcal{O}(t^2), \tag{3.38}$$

where for a non-minimally coupled scalar field we have [9]:

$$\Delta b_4 = \frac{1}{1440\pi}[5(6\xi - 1)^2 + 1] \int \mathrm{d}^2x \left(\sqrt{g}R^2 - \sqrt{\bar{g}}\bar{R}^2\right). \tag{3.39}$$

As we discuss later, the fact that we have an explicit expression for ΔK_L for any deformation at small t_k makes the task of evaluating the subtracted kernel a little easier, being the computational burden required to compute ΔK_L numerically for $t_k \lessapprox 10^{-2}$ significant.

At this point we pause briefly to describe the effects of mass and temperatures that we can extract from Equation 3.25. First of all, the $T \to 0$ limit can be inferred by using the fact that through Poisson resummation one can show:

$$\lim_{T\to 0} T\Theta_\sigma(T^2 t_k) = \frac{1}{\sqrt{4\pi t_k}}. \tag{3.40}$$

For non-zero T, however, we see that $\Theta_{-1/2}$ does not decay to zero as $t_k \to \infty$, but instead approaches 1 (since all terms die off exponentially, except from the $n = -1$ that gives unity for all non-zero times). For both zero and non-zero temperatures, however, as $t_k \to 0$ we see $\Theta_{-1/2} \to \infty$. This is useful to note in order to extract the $T \to \infty$ limit. Indeed, in this case, the $t_k \to 0$ and $T \to \infty$ limits do not commute. However, we know that for all finite temperatures $\Theta_{-1/2}$ diverges as t_k approaches zero, indicating that the correct ordering of the limit is to let first t_k go to zero and to later take $T \to \infty$. Apart from the measure zero point $t_k = 0$, we see that in the infinite temperature limit:

$$\lim_{T\to\infty} \Theta_{-1/2} = 1. \tag{3.41}$$

Hence, we see that the free energy per unit temperature in the 'hot' limit is given by:

$$\lim_{T\to\infty}\frac{\Delta F}{T} = -\frac{1}{2}\int_0^\infty \mathrm{d}t_k\; e^{-M^2 t_k}\frac{\Delta K_L(t_k)}{t_k}. \tag{3.42}$$

Another interesting regime is the long-wavelength limit, obtained when $M\ell \gg 1$, with ℓ being the typical curvature scale of the geometry. The heat kernel coefficients Δb_{2n} scale with such parameter as ℓ^{-2n} for $n \geq 2$, which imply that in the long-wavelength regime we obtain [9]:

$$\Delta F = -\frac{1}{4}\frac{\Delta b_4}{M}\coth\left(\frac{M}{2T}\right) + \dots\ . \tag{3.43}$$

Now, considering ΔF_0 to be the subtracted free energy obtained for a small deformation (as we discuss later we have an analytic expression for such cases) with the same temperature and mass, we obtain:

$$\frac{\Delta F}{\Delta F_0} = \frac{\Delta b_4}{\Delta b_4^{(0)}}, \tag{3.44}$$

where $\Delta b_4^{(0)}$ is the linear coefficient for the case of the small distortion. Also, note that Equation 3.28 shows the same exponential suppression due to finite mass that we observed in the parallel plates configuration.

3.2.2 Sign of the Subtracted Heat Kernel

It is now a good place to stop and look at the information we can extract from the analytic expressions that we have for the subtracted heat kernel. Starting from the small t_k regime, we can trivially see that Δb_4 is a positive coefficient. Indeed, we can use the volume conservation condition, together with Equation 3.20 and Equation 3.21 (and the fact that $\overline{R} = 2$ is a constant) to derive the statement:

$$\begin{aligned}\int \mathrm{d}^2x\left(\sqrt{g}R^2 - \sqrt{\bar{g}}\overline{R}^2\right) &= \int \mathrm{d}^2x\left(R^2 - \overline{R}^2\right)\\ &= \int \mathrm{d}^2x\sqrt{g}\left(R-\overline{R}\right)^2 - 2\int \mathrm{d}^2x\sqrt{g}\overline{R}^2 + 2\int \mathrm{d}^2x\sqrt{g}R\overline{R}\\ &= \int \mathrm{d}^2x\sqrt{g}\left(R-\overline{R}\right)^2 \geq 0.\end{aligned} \tag{3.45}$$

We therefore see that, for small enough t_k, the subtracted heat kernel is guaranteed to be positive.

The large t_k regime is instead governed by the smallest eigenvalue of L, as the other contributions to Equation 2.63 vanish more quickly as $t_k \to \infty$. Now, this is not quite true for the minimally coupled scalar, where we have seen that the constant function is an eigenmode of both the perturbed and unperturbed Laplacian, meaning

that the two contributions exactly cancel in Equation 2.63. To understand the large time asymptotics for the $\xi = 0$ case, it is hence necessary to look at the second smallest eigenvalue for L and $\overline{L}$. In the case of the unperturbed Laplacian, this is given by $\lambda = 2$ (it has degeneracy 3, but this does not matter since it just amounts to a mutliplicative factor which clearly does not change the qualitative large t_k behaviour). The second smallest eigenvalue for the deformed case is known to be bounded by [9]:

$$\lambda_2 \leq \frac{8\pi}{\mathrm{Vol}[g]} = \bar{\lambda}_2, \tag{3.46}$$

i.e. it is guaranteed to be smaller than the unperturbed eigenvalue. Going back to the non-minimally coupled scalar, the lowest eigenvalue is non-zero and different for the deformed and undeformed geometry. It can be shown to be given by the Reyleigh-Ritz formula [9]:

$$\lambda_{\min} = \inf_\phi J[\phi], \qquad J[\phi] \coloneqq \left[\int \mathrm{d}^2x\sqrt{g}\phi^2\right]^{-1} \int \mathrm{d}^2x\sqrt{g}\phi(-\nabla^2 + \xi R)\phi, \tag{3.47}$$

with the infimum intended to be taken over all square-integrable test functions ϕ. We see that we can bound $\lambda_{\min}$ from above with a constant test function, and using the relations in Equation 3.20 and Equation 3.21 one can show such a bound to be:

$$\lambda_{\min} \leq J[\phi = \mathrm{const}] = \frac{4\pi\xi\chi_\Sigma}{\mathrm{Vol}[g]} = \bar{\lambda}_{\min}, \tag{3.48}$$

meaning again that the smallest eigenvalue for the undeformed sphere is bigger than the one for the deformed geometry. This has the implication that for both the minimally and non-minimally coupled scalar, the subtracted heat kernel is greater than zero also in the big t_k regime. The sign of the intermediate t_k region is up to grabs for general deformations, which is the reason why we use numerical calculations to compute it. Nonetheless, it is possible to study it analytically for specific configurations. We shortly review some of such results for the mid t_k range for small deformation and for deformations of any size in the case of a conformal field theory.

The sign of the subtracted heat kernel is important as it gives hints on the sign of the subtracted free energy. Of course it's not necessary for ΔK_D to be positive *everywhere* for ΔF to be negative (recall there is an overall minus sign in Equation 3.23), but showing positivity of ΔK_D for all t_k guarantees the negativity of the energy (and, consequently, the preference of the quantum field to live on the deformed geometry).

3.2.3 Small Deformations

We now summarise analytical results regarding a scalar field living on a geometry only small perturbations away from the unit sphere [9]. We start by considering the

metric in conformal gauge in Equation 3.2 and expanding the deformation function in the small parameter ϵ that governs the strength of the perturbation:

$$f = \epsilon f^{(1)} + \epsilon^2 f^{(2)} + \mathcal{O}(\epsilon^3), \tag{3.49}$$

with $\epsilon = 0$ clearly corresponding to the reference S^2 metric. Now, the area preservation condition amounts to:

$$\int \mathrm{d}^2x\sqrt{\bar{g}}f^{(1)} = 0 \quad \wedge \quad \int \mathrm{d}^2x\sqrt{\bar{g}}\left((f^{(1)})^2 + f^{(2)}\right) = 0. \tag{3.50}$$

Note that keeping the expansion of f to second order is crucial in order not to have $f^{(1)} = 0$ fixed by the volume preservation condition. We can similarly expand the Laplacian L, its eigenvalues λ_I, its eigenfunctions h_I and the subtracted heat kernel in powers of ϵ:

$$\begin{aligned} L &= \bar{L} + \epsilon L^{(1)} + \epsilon^2 L^{(2)} + \mathcal{O}(\epsilon^3), \\ h_I &= \epsilon h_I^{(1)} + \epsilon^2 h_I^{(2)} + \mathcal{O}(\epsilon^3), \\ \lambda_I &= \epsilon \lambda_I^{(1)} + \epsilon^2 \lambda_I^{(2)} + \mathcal{O}(\epsilon^3), \\ \Delta K_L &= \epsilon \Delta K_L^{(1)} + \epsilon^2 \Delta K_L^{(2)} + \mathcal{O}(\epsilon^3). \end{aligned} \tag{3.51}$$

It follows from homogeneity of the round sphere that $\Delta K_L^{(1)}(t_k) = 0$, meaning that the leading contribution to the subtracted heat kernel from the perturbation is quadratic. For a more rigorous treatment, please refer to [9]. However, it can be proved that:

$$\Delta K^{(2)}(t_k) = t\sum_I e^{-\bar{\lambda}_I t_k}\left[\frac{t_k}{2}\left(\lambda_I^{(1)}\right)^2 - \lambda_I^{(2)}\right]. \tag{3.52}$$

We now take L in the form of Equation 3.24, i.e. for a non-minimally coupled scalar field. Recall that on the round sphere (i.e. $f = 0$), the eigenvalues are given by $\bar{\lambda}_l = l(l+1) + 2\xi$ with l a non-negative integer. Then, expanding $f^{(1)}$ in spherical harmonics

$$f^{(1)} = \sum_{l,m} f_{l,m} Y_{l,m}, \tag{3.53}$$

with $f_{l,m}$ being the coefficients of the expansion, one can show that:

$$\Delta K^{(2)}(t_k) = \sum_{l,m} a_l(t_k)|f_{l,m}|^2, \qquad a_l(t_k) \coloneqq t_k \sum_{l'=0}^{\infty} e^{-\bar{\lambda}_{l'} t_k}(\alpha_{l,l'} + \beta_{l,l'} t_k), \tag{3.54}$$

where $\alpha_{l,l'}$ and $\beta_{l,l'}$ are numerical coefficients. For general l, these are given by:

$$\alpha_{l,l'} = \begin{cases} \frac{(2l'+1)\left(\bar{\lambda}_{l'}-\xi C_l\right)^2}{\pi C_l} \frac{\left(\frac{2+l}{2}\right)_{l'}\left(\frac{l}{2}\right)_{-l'}}{\left(\frac{3+l}{2}\right)_{l'}\left(\frac{1+l}{2}\right)_{-l'}}, & l' < \frac{l}{2} \\ \Bigg[\frac{(2l'+1)\left(\bar{\lambda}_{l'}-\xi C_l\right)^2}{2\pi}\left(H_{l'-\frac{l}{2}} - H_{l'+\frac{l}{2}}\right. & \\ \quad \left. + H_{l'+\frac{l+1}{2}} - H_{l'-\frac{l+1}{2}} - \frac{2}{2l'+1}\right) & l' \geq \frac{l}{2} \\ \quad - \frac{(2l'+1)^2\left(\bar{\lambda}_{l'}-\xi C_l\right)}{\pi}\Bigg] \begin{pmatrix} l & l' & l' \\ 0 & 0 & 0 \end{pmatrix}^2 & \end{cases} \tag{3.55}$$

and

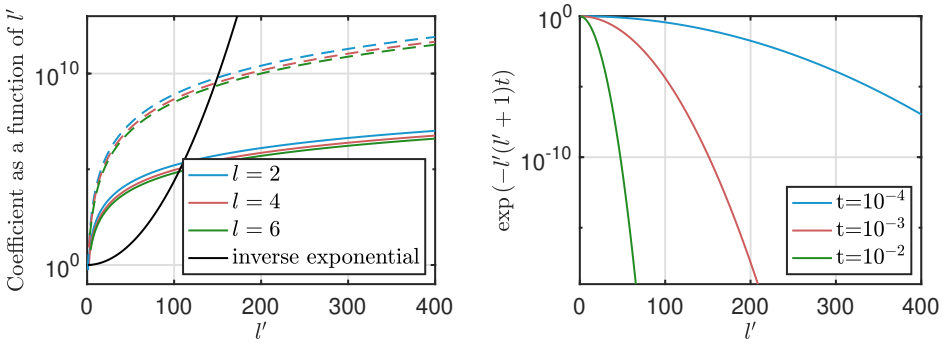
$$\beta_{l,l'} = \frac{(2l'+1)(\bar{\lambda}_{l'}-\xi C_l)^2}{2\pi} \begin{pmatrix} l & l' & l' \\ 0 & 0 & 0 \end{pmatrix}^2, \tag{3.56}$$

where $C_l \coloneqq l(l+1)$, $(x)_n = \Gamma(x+n)/\Gamma(x)$ are the Pochhammer symbols, H_n are the harmonic numbers and

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.57}$$

is the $3j$ symbol. An interesting thing to note from Equation 3.54 is that the perturbative result is insensitive to the value of m from which we build our deformation function, as the result is only dependent on l and on the *coefficients* of the spherical harmonics expansion.

Later, in the presentation of our numerical solutions, we will want to benchmark our numerical results with the analytical predictions. However, Equation 3.54 seems to suggest that we have to numerically evaluate an infinite sum of terms to compute the analytical result. We have to truncate such sum in l', and the exponential terms in Equation 3.54 give us a rough idea of the ideal place to do it. Figure 3.3 shows the magnitude of the α and β coefficients and exponential prefactor for different l, l' and kernel times. Figure 3.3a compares these coefficients at $t = 10^{-3}$ for different l as a function of l', showing that the contributions to Equation 3.2.3 become neglegible for $l' \gtrapprox l'_{\text{cut}} = 200$. Figure 3.3b helps to visualise the fact that in order to compute $\Delta K_L^{(2)}$ for smaller and smaller t_k a significant increase in l'_{cut} is required. Fortunately, we do not necessarily need to compute the subtracted heat kernel this way for small t_k: we can patch it with the small t_k expansion in Equation 3.38, so that it is sufficient to truncate the sum as soon as $\Delta K_L^{(2)}$ approaches such linear behaviour. The situation is even simpler for odd l, however, since the $3j$ symbol vanishes if $l + 2l'$ is an odd integer! Hence, for the case of odd l, the α and β coefficients are



(a) (b)

Figure 3.3: Coefficients for the perturbative expansion of the subtracted heat kernel. (a) Magnitude of the α (solid coloured lines) and β (dashed coloured lines) as a function of l' for different l. The inverse of the (inverse) exponential prefactor is plotted for comparison (black solid line) for $t_k = 10^{-3}$. (b) Magnitude of the inverse exponential prefactor as a function of l' for different heat kernel times.

given by the reduced expressions:

$$\alpha_{l,l'} = \begin{cases} \frac{(2l'+1)\left(\bar{\lambda}_{l'} - \xi C_l\right)^2}{\pi C_l} \frac{\left(\frac{2+l}{2}\right)_{l'} \left(\frac{l}{2}\right)_{-l'}}{\left(\frac{3+l}{2}\right)_{l'} \left(\frac{1+l}{2}\right)_{-l'}}, & l' < \frac{l}{2} \\ 0 & l' \geq \frac{l}{2} \end{cases} \tag{3.58}$$

and

$$\beta_{l,l'} = 0, \tag{3.59}$$

effectively transforming the infinite sum in into a finite one. In the paper, big axisymmetric deformations are then studied numerically for a scalar and a Dirac free field. The results can be shortly summarised by the statement that $\sigma \Delta K$, with $\sigma = -1/2$ for the scalar and $\sigma = 1$ for the fermion, is found to be always negative, showing that the free energy is maximised by S^2 for the deformations considered. Furthermore, it is found that ΔF behaves suggestively similar between the fermionic and scalar field, a feature further discussed in [10].

3.2.4 Strongly Coupled Conformal Field Theories

So far, we have studied a scalar free theory, with no interaction whatsoever. Indeed, one might naively expect that introducing interactions would only complicate the picture, and would not help in the search of general analytical results. While this would be right in most cases, this is not what happens for a strongly coupled (2+1)-dimensional conformal field theory thanks to the incredibly powerful tool of holography. While we do not have much time to go into details, we briefly touch on CFTs and the AdS/CFT correspondence, and then summarise a recent result of interest.

A conformal field theory is a quantum field theory which is invariant under the conformal group. The conformal group is an extension of the Poincare' group that includes dilatations ($x \to \lambda(x)x$) and special conformal transformations (which amount to an inversion, followed by a translation and a further inversion), and it is isomorphic to $SO(1, d+1)$ in an Euclideanised d-dimensional spacetime. Effectively, conformal transformations are coordinate transformations $x^\mu \to \tilde{x}^\mu(x)$ that cause the metric to change as:

$$g_{\mu\nu}(x) \to \Omega^2(x) g_{\mu\nu}(x). \tag{3.60}$$

The statement of invariance under a conformal transformation, is really the statement that the physics is insensible to length scales. CFTs have recently been the centre of much attention due to the formulation of the AdS/CFT correspondence, i.e. the conjectured (but now widely accepted) possibility of translating a physical problem in a $d+1$-dimensional quantum theory of gravity (string theory) in Anti-de Sitter background geometry to one in the conformal quantum field theory without gravity living on its d-dimensional boundary. This is formally stated by equating the generating function of the connected correlation functions for the CFT with the string/gravity partition function on AdS space. The correspondence can be used both ways, i.e. to solve quantum gravity problems by studying the CFT and vice versa. In our case, we make use of the fact that, the bulk $(d+1)$-gravitational theory is effectively classical in the strong coupling limit for the CFT, so that we can use Einstein's equation to analyse the otherwise intractable strongly correlated CFT [43].

A recent paper has studied the Casimir enegry for a CFT living on the same (2+1)-dimensional spacetime that we consider in our analysis [8]. The static spacetime for the boundary CFT is taken to be, without loss of generality, in the ultrastatic frame (as in our case). This is justified by the fact that we can bring the metric in such ultrastatic frame via a conformal transformation, which is shown not to change the Casimir energy of the CFT. Such energy, is given by:

$$E = \int_\Sigma \sqrt{g} T^{\mathrm{CFT}}_{\mu\nu} n^\mu v^\nu, \tag{3.61}$$

where $v = \partial_t$ is the global timelike Killing vector of the static spacetime, n is the unit vector normal to a constant t hypersurface and $T^{\mathrm{CFT}}_{\mu\nu}$ is the stress-energy tensor of the CFT. A thermal state at finite temperature T of the CFT is assumed to be described in the bulk by a gravitational solution that is smooth away from the conformal boundary and that ends only either on such boundary or on smooth killing horizons with a Hawking temperature of T with respect to v. The Casimir energy is then evaluated as the limit of E as $T \to 0$. Without getting into much detail, the AdS/CFT dictionary then dictates the expectation value of the stress-energy tensor of the CFT to be determined by the asymptotic approach of the bulk metric

solution to the conformal boundary. The final result is that the Casimir energy for a strongly coupled CFT is always non-positive, and is maximised by a Σ with constant curvature (which, in the topology we are considering corresponds to the sphere). This means that strongly coupled (2+1)-dimensional CFTs prefer to crumple and disfavour the simpler spherical geometry! Contrary to the results earlier discussed (which were valid only for small perturbations), this is valid for any Σ geometry and, hence, for deformations of any size. It is curious that, conterintuitively, the most general analytical result up to date has been obtained for a strongly coupled field rather than for a (naively simpler to treat) free field. This is an excellent proof of the incredible power of the AdS/CFT conjecture.

Chapter 4

Numerical Methods

In this section we present our approach to the numerical calculations. We first introduce spectral methods, highlighting the advantages that these techniques bring with respect to the standard equispaced finite difference approach. We then examine the discretisation of the physical space, describing Chebyshev nodal collocation and illustrating how to implement the boundary conditions. In order to clarify this as much as possible, an application of the method to a toy model is successively presented, where the eigenvalues and eigenfunctions of the 1D Laplacian are found. We conclude with an analysis on the performance of the numerical methods, looking at how the density and geometry of the computational grid affects the accuracy of the calculations in different regimes.

4.1 Spectral Methods

4.1.1 Runge's Phenomenon

Consider a periodic function $f(x)$ defined on the closed interval $x \in I[a, b]$ on the real line. Suppose now that we discretise such an interval, obtaining a set of $N+1$ nodes $\{x_j\}_{j=0}^{N}$, which may or may not be equidistant. An interesting question would be: how can I reconstruct the function through polynomial interpolation as faithfully as possible using the sampled values $f_j = f(x_j)$? This query, basic as it might seem, opens the door to spectral methods (we'll define them carefully soon) by highlighting how important the collocation of nodes is when approximating a function.

A naive approach would be to just place the lattice points evenly spaced and interpolate them with a polynomial of as high degree as possible. As it turns out, this approach is not only inefficient, but potentially disastrous. It is a known result [44] that the difference at a point $x \in I$ between a function g which is C^{N+1} in the interval I and the interpolating polynomial $P_N(x)$ of degree at most N is given by:

$$e_N(x) \coloneqq g(x) - P_N(x) = \frac{g^{(N+1)}(\eta)}{(N+1)!} \prod_{i=0}^{N} (x - x_i) \tag{4.1}$$

where η is some point in the interval and the x_i is the $(i+1)$th node. To simplify notation, we define:

$$w_N(x) \coloneqq \prod_{i=0}^{N} (x - x_i). \tag{4.2}$$

Equation 4.1 provides a quantitative way to check whether our interpolating polynomial gives a better and better approximation as the number of lattice points (and, consequently, degree of the polynomial) increases. Changing N affects both $w_N(x)$ and the prefactor involving the $(N+1)$th derivative of g. The factorial on the denominator could suggest that a bigger number of nodes achieves a better approximation, but really both the derivative and product terms can step in to reverse that behaviour. This is what's commonly known as *Runge's phenomenon* (similar in looks, but not in origin, to the Gibbs' phenomenon ubiquitous in signal analysis). The effect is named after the mathematician Carl Runge, who first presented the canonical example, that we now illustrate, in which increasing the order of the interpolating polynomial makes the error grow without bounds [45]. Consider:

$$g(x) = \frac{1}{1 + 25x^2}, \qquad x \in [-1, 1]. \tag{4.3}$$

A careful, but lengthy, analysis on $w_N(x)$ and on the derivatives of $g(x)$ shows that for certain x the error diverges as $N \to \infty$ if we take uniform spacing for the nodes [44]. An elegant alternative approach to achieve the same result is instead to analytically continue f to the complex plane. For complex z, the analogous expression to Equation 4.1 is given by the following contour integral:

$$e_N(z) \coloneqq g(z) - P_N(z) = \frac{1}{2\pi i} \int_C \frac{w_N(z)}{w_N(\eta)} \frac{g(\eta)}{\eta - z} \mathrm{d}\eta, \tag{4.4}$$

with C being the boundary of a domain R where f is analytic and where both z and the nodes are contained. It is also useful to define:

$$\sigma_N(z) \coloneqq |w_N(z)|^{1/(N+1)} \tag{4.5}$$

and its limiting value:

$$\sigma(z) \coloneqq \lim_{N \to \infty} \sigma_N(z), \tag{4.6}$$

which in the case of equidistant points can be shown to be [44]:

$$\sigma(z) = \exp\left\{\frac{1}{b-a}\int_a^b \log|z-t|\,\mathrm{d}t\right\}. \tag{4.7}$$

Consider now the family of curves

$$C(\rho) = \{z \in \mathbb{C} | \sigma(z) = \rho\}, \tag{4.8}$$

which for the case of uniform node spacing correspond to smooth curves about the midpoint of the real interval I. In general, we have that for $\eta \in C(\rho)$ and $z \in C(\rho')$

$$\lim_{N\to\infty}\left|\frac{w_N(z)}{w_N(\eta)}\right| = 0 \tag{4.9}$$

if $\rho > \rho'$. This has the key implication that if we can find a curve $C(\rho)$ containing z and that bounds a region of the complex plane where g is analytic, the polynomial interpolation will indeed converge for big enough N. If g is entire, this guarantees convergence on any domain I. This is not true, however, if g has poles. Indeed, let g have a simple pole at z^*, and suppose that there is no curve $C(\rho)$ that encloses z but not the pole. Since Equation 4.4 is valid only if the integration contour bounds a region of the complex plane where g is analytic, we cannot use any member of the family of curve given by Equation 4.8 by itself to estimate the error at z. Nonetheless, we can integrate along $C(\rho)$, if we are careful to make a deviation Γ to avoid z^*. Doing so, we have:

$$e_N(z) = \frac{1}{2\pi i}\int_{C(\rho)} \frac{w_N(z)}{w_N(\eta)}\frac{g(\eta)}{\eta - z}\mathrm{d}\eta + \int_\Gamma \frac{w_N(z)}{w_N(\eta)}\frac{g(\eta)}{\eta - z}\mathrm{d}\eta \tag{4.10}$$

As discussed earlier, the first integral gives a vanishing contribution as $N \to \infty$, while the second one can be trivially evaluated using the residue theorem, leading to:

$$\lim_{N\to\infty} e_N(z) = \frac{w_N(z)}{w_N(z^*)}\frac{\mathrm{Res}_g(z^*)}{z^* - z} = \left(\frac{\sigma_N(z)}{\sigma_N(z^*)}\right)^{N+1}\frac{\mathrm{Res}_g(z^*))}{z^* - z}. \tag{4.11}$$

Here $\mathrm{Res}_g(z^*)$ indicates the residue of the function g at the pole z^*. Taking N big enough, $\sigma_N(z) \to \sigma(z) = \rho$ and, considering $z \in C(\rho')$ and $z^* \in C(\rho^*)$ with $\rho^* < \rho'$, we see that:

$$\lim_{N\to\infty}\left(\frac{\sigma_N(z)}{\sigma_N(z^*)}\right) = \frac{\rho' - \epsilon}{\rho^* + \epsilon} > 1, \tag{4.12}$$

with ϵ small. Consequently, not only convergence is not guaranteed, but actually the error of the interpolating polynomial blows up exponentially with N outside the region bounded by $C(\rho^*)$ (i.e. the curve where the pole lies). This is reflected on our real interval I by the fact that the error grows without bounds outside I_c, i.e. the portion of I enclosed by $C(\rho^*)$. If g has multiple poles, the same applies, with

the error blowing up outside $C(\rho^*_{min})$, where ρ^*_{min} is the minimum ρ such that $C(\rho)$ crosses a pole. Going back to Runge's function in Equation 4.3, we see that it has simple poles at $z^* = \pm i/5$. In this case, one can work out that $\rho^* \approx 0.494$ and, consequently, that the error grows without bounds on the real line for $|x| \gtrapprox 0.727$, i.e. at the edges of the domain (as it is always the case for Runge's distortion)

4.1.2 Chebyshev Nodal Collocation

There are few approaches to mitigate Runge's phenomenon, but the solution we adopt (and, arguably, the most elegant one) is to change the collocation of the nodes. While in theory it would be enough to choose $\{x_j\}$ such that the generated $C(\rho)$ completely enclose the interval I on the real line without containing any poles, this is easier said than done. Instead, we go back to Equation 4.1 and try to envisage a discretisation that minimises the error for a given N, optimising the function $w_N(x)$. As it turns out, this will not only solve the distortion in Runge's example, but also provide a collocation of points that guarantees a much faster convergence for smooth enough functions compared to the uniformly spaced grid in all cases, laying the foundations for spectral methods.

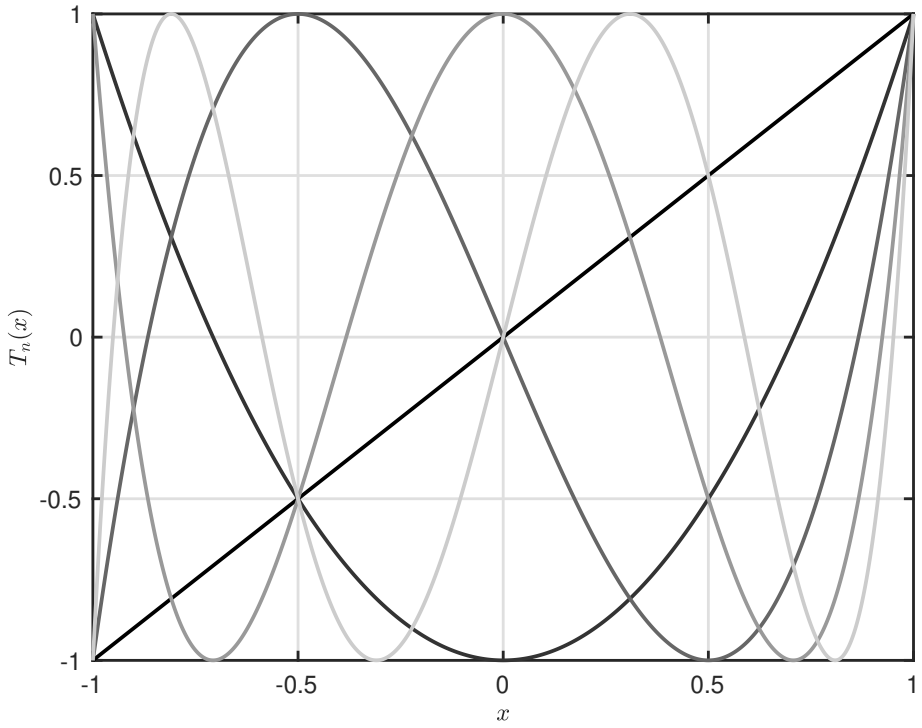


Figure 4.1: The first 5 Chebyshev polynomials (darker corresponds to lower order).

Let's introduce the Chebyshev polynomials of order N as a function of the auxilary variable $\theta = \cos^{-1}(x)$:

$$T_N(x) = \cos(N\theta(x)). \tag{4.13}$$

The first 5 Chebyshev polynomials are show in Figure 4.1. This form is particularly useful, since it elucidates the geometrical interpretation of the so called Chebyshev nodal collocation, where the set of nodes $\{x_j\}$ corresponds to the set of roots of the Nth degree Chebyshev polynomial. Geometrically, the Chebyshev grid on the interval $[-1, 1]$ can be seen to be obtained by projecting on the real line equally spaced points on the unit upper semicircle, as displayed by Figure 4.2. An explicit expression for the Chebyshev nodes for an Nth order polynomial interpolation is given by:

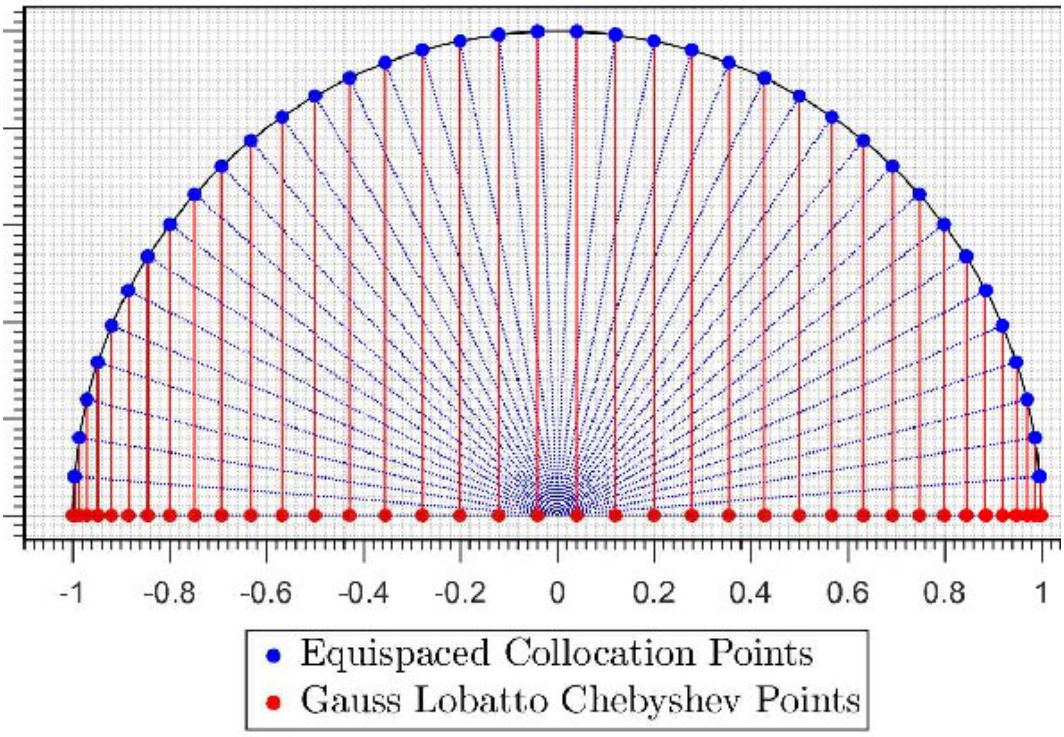


Figure 4.2: Chebyshev collocation.

$$x_j = \cos\left[\frac{(2j+1)\pi}{2N}\right]. \tag{4.14}$$

To help with the visualisation of this alternative collocation methods, Figure 4.3 shows the comparison between the interval $x \in [-1, 1]$ discretised with uniform spacing and the same interval latticised with Chebyshev nodes. The latter has evident clustering of lattice points at the edges of the domain.

Using the identity

$$(x - x_0)(x - x_1)...(x - x_N) = \frac{T_{N+1}(x)}{2^N}, \tag{4.15}$$

where the x_i are the Chebyshev roots, and applying the result that $T_{N+1}/2^N$ is the polynomial of order $N + 1$ and leading coefficient 1 with smallest supremum norm $||\cdot||_\infty$, one can easily see that choosing Chebyshev collocation minimises Equation 4.1

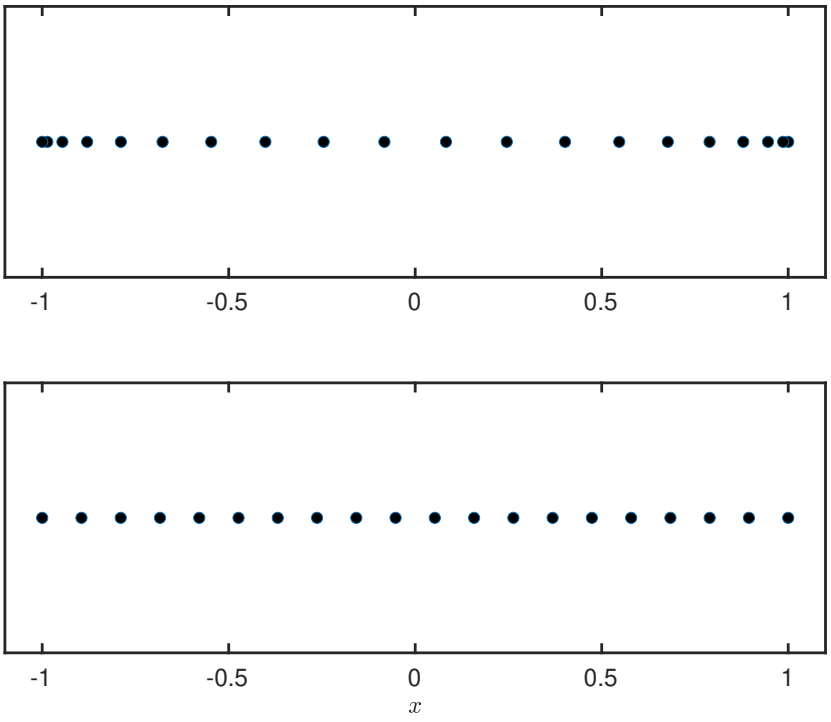


Figure 4.3: $x \in [-1, 1]$ interval discretised with $N = 20$ Chebyshev (top) and equispaced (bottom) nodes.

for fixed g and N [46]. Furthermore, we can see that:

$$\begin{aligned} ||e_N(x)||_\infty &= \left|\left| \frac{g^{(N+1)}(\xi)}{(N+1)!} \prod_{i=0}^{N} (x - x_i) \right|\right|_\infty \\ &= \frac{\left|\left| g^{(N+1)}(\xi) \right|\right|_\infty}{(N+1)!} \frac{||T_{N+1}(x)||_\infty}{2^N} \\ &\leq \frac{\left|\left| g^{(N+1)}(\xi) \right|\right|_\infty}{(N+1)! 2^N}, \end{aligned} \tag{4.16}$$

showing that, for the Chebyshev case, the error decays exponentially with N. This is a significant improvement from the uniformly spaced grid, and it's what goes by the name of *spectral convergence*. The deep reason behind the faster convergence is that the family of curves $C(\rho)$ that one generates with such a choice of nodes is the one of ellipses with foci at $x = \pm 1$ [44]. This allows us to find a contour that avoids the pole but contains the entirety of the domain, hence guaranteeing a bounded (and, crucially, decaying) error as $N \to \infty$. Comparison between the convergence for the interpolation of Runge's function using uniform and Chebyshev grid is shown in Figure 4.4, where the advantage of using Chebyshev nodal collocation is extremely noticeable near the edges of the domain.

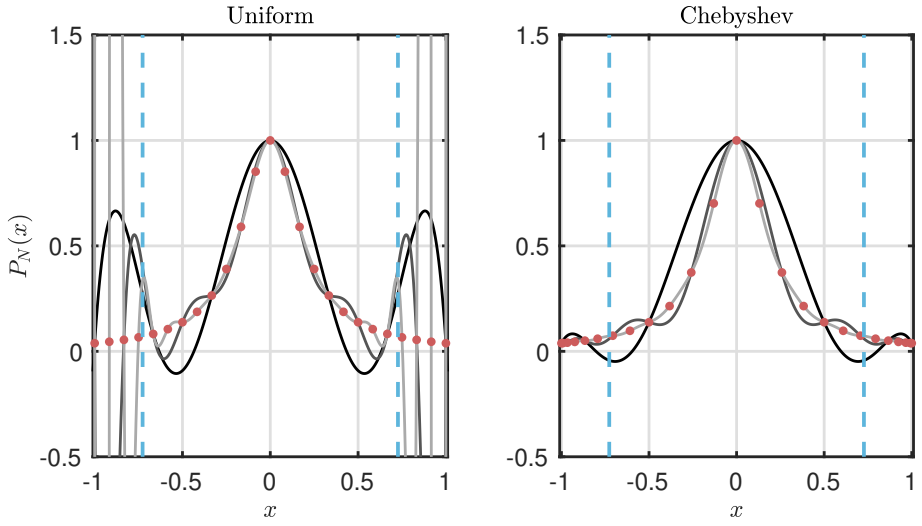


Figure 4.4: Polynomial interpolation of Runge's function (sampled by the red markers) with polynomials of order $N = 7, 13, 25$ (solid lines, darker means smaller N). A uniform collocation of nodes (left) causes a fail to converge to the desired function for $|x| \gtrapprox 0.727$ (blue vertical lines), which is amplified as N increases. The opposite happens for Chebyshev spacing (right), with a better approximation achieved for higher order polynomial.

4.1.3 Differentiation Matrices

The advantage of unevenly spaced grids to interpolate functions on an interval means that an equispaced grid is inefficient even when evaluating the derivative of a function. In particular, the differentiation matrix obtained through a finite difference scheme (uniformly spaced grid) gives an approximation for the derivative that typically converges as $O(N^{-m})$ for some m that depends both on the order of the interpolant from which the matrix is obtained and on the smoothness of the function. On the other hand, differentiation matrices obtained via spectral methods (e.g. Chebyshev collocation) converge at least as $O(N^{-m})$ *for every* m, which is faster than what finite difference schemes give. Moreover, if the function is suitably analytic, one can achieve the even faster convergence of $O(c^N)$ with $0 < c < 1$ (recall that when we say $h(x) = O(k(x))$ we mean that there exists some constant α such that for $x \to \infty$ we have that $|h(x)| < \alpha|k(x)|$) [47]. Spectral methods reside on the idea of appropriately choosing an orthonormal basis to approximate the solution of a differential equation defined on the interval of interest. Different (generally smooth) basis functions can be chosen, from which a variety of differentiation matrices can be extracted. The choice is made taking into consideration the expected properties and symmetries of the solutions. For the case of a periodic function, a Fourier basis is the natural choice, while orthogonal polynomials (among which Chebyshev are the most commonly used) are the best options for non-periodic solutions. For a detailed analysis on how choosing the right basis can significantly improve stability and convergence of the numerical method, please refer to Chapter 4 of Trefethen [47].

As shown earlier, non-periodic functions can be accurately approximated via polynomial interpolation on a Chebyshev grid on the $[-1, 1]$ domain (and, in fact on any domain by appropriate rescaling of the polynomials). Constructing our differentiation matrix from Chebyshev discretisation, we can then exploit the convergence advantages that such approach brings with respect to the standard finite difference techniques. For a grid of $N + 1$ points, the matrix $\underline{\underline{D_N}}$ representing the linear operation that gives the derivative of a function sampled following the Chebyshev prescription has entries given by:

$$\begin{aligned} (\underline{\underline{D_N}})_{00} &= \frac{2N^2+1}{6} = -(\underline{\underline{D_N}})_{NN}, \\ (\underline{\underline{D_N}})_{jj} &= -\frac{x_j}{2(1-x_j^2)} \qquad j = 1, 2, ..., N-1, \\ (\underline{\underline{D_N}})_{ij} &= \frac{c_i}{c_j}\frac{(-1)^{i+j}}{x_i - x_j} \qquad j \neq i, \end{aligned} \tag{4.17}$$

with:

$$c_i = \begin{cases} 1, & i = 0, N \\ 2 & \text{otherwise.} \end{cases} \tag{4.18}$$

To understand the origin of the differentiation matrix, consider a grid function $\vec{u}$ sampled on the Chebyshev nodes (effictively, a vector with $N+1$ entries). Then, to read off the coefficients of the matrix, one would find the unique interpolating polynomial $P_N(x)$ of order at most N that interpolates the data $\{u_j\}$, differentiate it and rewrite the result in a matrix multiplication of the form:

$$\vec{w} = \underline{\underline{D_N}}\vec{u}, \tag{4.19}$$

where we impose that $w_j = P'_N(x_j)$.

From the differentiation matrix for the first derivative we can obtain the one for the nth derivative simply by elevating the former to the nth power. $\underline{\underline{D_N}}$ is then the building block from which we construct any differential operator of interest.

4.1.4 Boundary Conditions

It is at the level of the differentiation matrix that we need to impose the boundary conditions of the physical problem at hand. To better explain how this is done, we solve a simple 1D toy model, and successively move on to the more complex case of the Laplacian on a sphere.

Let's pretend that we want to find the spectrum of the Laplacian on the intevral $x \in [-1, 1]$ on the real line, subject to a mixture of Dirichlet and Neumann homogeneous boundary conditions at the extremal points. Effectively we want to solve the following eigenvalue problem:

$$\frac{\mathrm{d}^2 y_n}{\mathrm{d}x^2} = -\lambda_n y_n, \qquad \phi(y_n(-1), y'_n(-1)) = \chi(y_n(1), y'_n(1)) = 0, \tag{4.20}$$

where y_n and λ_n are respectively the nth eigenfunction and eigenvalue of the 1D Laplacian, while ϕ and χ are some functions of y_n and its derivative at the boundary giving the homogeneous boundary conditions. This example is familiar to any physicist, as it is exactly the first problem that students meet in all the introductory courses on quantum mechanics: the particle in a box (which can be seen by identifying y with the wavefunction and λ with energy). To mirror what we do when looking at the 2D Laplacian on the sphere, we focus on two basic cases:

- homogeneous Dirichlet boundary conditions, i.e. $y(-1) = y(1) = 0$, for which we obtain the following set of non-trivial eigenfunctions (up to an irrelevant

factor A) and eigenvalues:

$$y_n(x) = \begin{cases} A\cos\left(\frac{n\pi x}{2}\right), & n \text{ odd} \\ A\sin\left(\frac{n\pi x}{2}\right), & n \text{ even} \end{cases}, \qquad \lambda_n = \frac{n^2\pi^2}{4}, \qquad n = 1, 2, ... \tag{4.21}$$

- homogeneous Neumann boundary conditions, i.e. $y'(-1) = y'(1) = 0$, that instead result in:

$$y_n(x) = \begin{cases} A\sin\left(\frac{n\pi x}{2}\right), & n \text{ odd} \\ A\cos\left(\frac{n\pi x}{2}\right), & n \text{ even} \end{cases}, \qquad \lambda_n = \frac{n^2\pi^2}{4}, \qquad n = 1, 2, ... \tag{4.22}$$

It is useful to think about any homogeneous boundary condition as a matrix multiplication of the form:

$$\underbrace{\vec{0}}_{m\times 1} = \overbrace{\underline{\underline{B}}}^{m\times M} \underbrace{\vec{y}}_{M\times 1}, \tag{4.23}$$

where the matrix $\underline{\underline{B}}$ acting on $\vec{y}$ gives the equations for such constraints. This can be generalised for the case of non-homogenous conditions substituting the vector of zeros with a vector $\vec{c}$ containing the inhomogeneous part of the constraints. The dimensions highlighted in Equation 4.23 assume M grid points and m constraints, which in our 1D example would translate into $\vec{0}$ being a 2×1 column vector, $\underline{\underline{B}}$ being a $2 \times M$ matrix and the vector of sampled function values $\vec{y}$ a $M \times 1$ vector. To impose the boundary conditions at the level of the differential operator, we rewrite the entries $\vec{y_B}$ of $\vec{y}$ that correspond to the boundary points as a function of the interior points $\vec{y_I}$ [48]. We can rearrange $\underline{\underline{B}}$ into a formal row vector in order to visualise what's going on:

$$\underbrace{\vec{0}}_{m\times 1} = \begin{pmatrix} \overbrace{\underline{\underline{B_B}}}^{m\times m}, & \overbrace{\underline{\underline{B_I}}}^{m\times(M-m)} \end{pmatrix} \begin{pmatrix} \overbrace{\vec{y_B}}^{m\times 1} \\ \underbrace{\vec{y_I}}_{(M-m)\times 1} \end{pmatrix} = \underline{\underline{B_B}}\vec{y_B} + \underline{\underline{B_I}}\vec{y_I}.$$

Here, $\underline{\underline{B_B}}$ and $\underline{\underline{B_I}}$ are respectively a square and a rectangular matrix obtained from $\underline{\underline{B}}$ itself, such that the resulting constraints are left invariant by this formal rearrangement. It is immediately evident why the reorganisation is useful, since we can now isolate $\vec{y_B}$ provided that $\underline{\underline{B_B}}$ is non-singular. Rewriting:

$$\vec{y_B} = -\underline{\underline{B_B}}^{-1}\underline{\underline{B_I}}\vec{y_I} = \underline{\underline{A}}\vec{y_I} \tag{4.24}$$

we can then successfully reduce the degrees of freedom of the problem, finally implementing the constraints.

The next step is to modify the differential operator such that it only acts on the interior points, while still taking into account the conditions on the boundary points. Take the discretised version of the Laplace equation:

$$\underline{\underline{L}}\vec{y} = 0,$$

where $\underline{\underline{L}}$ is the discretised Laplacian (which in the simple example of Equation 4.20 is just the Chebyshev differentiation matrix squared). Let's once more rearrange $\vec{y}$ and, consequently, $\underline{\underline{L}}$ to visualise better the procedure:

$$\begin{pmatrix} \overbrace{\underline{\underline{L_{BB}}}}^{m \times m} & \overbrace{\underline{\underline{L_{BI}}}}^{m \times (M-m)} \\ \underbrace{\underline{\underline{L_{IB}}}}_{(M-m) \times m} & \underbrace{\underline{\underline{L_{II}}}}_{(M-m) \times (M-m)} \end{pmatrix} \begin{pmatrix} \overbrace{\vec{y_B}}^{m \times 1} \\ \underbrace{\vec{y_I}}_{(M-m) \times 1} \end{pmatrix} = \vec{0}.$$

Since we will work only with the interior points, we can discard the top entry of the formal vector that results from this matrix multiplication, and instead focus on the bottom component. Plugging in Equation 4.24 we obtain:

$$0 = \underline{\underline{L_{IB}}}\vec{y_B} + \underline{\underline{L_{II}}}\vec{y_I} = (\underline{\underline{L_{IB}}}\ \underline{\underline{A}} + \underline{\underline{L_{II}}})\vec{y_I} = \underline{\underline{\tilde{L}}}\vec{y_I},$$

where $\underline{\underline{\tilde{L}}}$ is the effective differential operator where the boundary conditions have been implemented. The problem of finding the eigenvalues and eigenfunctions of the differential operator subject to boundary conditions ϕ and χ is then reduced to the one of finding the eigenvalues and eigenvectors of the matrix $\underline{\underline{\tilde{L}}}$ [48]. Of course, we are limited to finding the $M - m$ smallest eigenvalues, but theoretically, as $M \to \infty$, one can capture the full spectrum of the operator.

Let's go back to the example of the Laplacian on a line, considering the case of homogeneous Neumann boundary conditions. Here the matrix $\underline{\underline{B}}$ corresponds just to the two rows of $\underline{\underline{D_{M-1}}}$ that give the derivatives at the boundaries. Explicitly, for a grid with $M = 3$ points the Chebyshev differentiation matrix is given by:

$$\underline{\underline{D_2}} = \begin{pmatrix} \frac{3}{2} & -2 & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 2 & -\frac{3}{2} \end{pmatrix}$$

The first and last rows of $\underline{\underline{D_2}}$ are the ones that give the derivative evaluated at the boundary points x_0 and x_2. Hence, we can extract the required $\underline{\underline{B}}$ matrix and

rewrite the constraints for this example as:

$$\begin{pmatrix} \frac{3}{2} & -2 & \frac{1}{2} \\ -\frac{1}{2} & 2 & -\frac{3}{2} \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

It is helpful to note that to implement homogeneous Dirichlet boundary conditions it is enough to delete from the differential operator columns and rows that respectively act on and give the boundary points. Therefore, for a grid vector of the form $\vec{y} = (y_0, y_1, ..., y_M)$ with y_0 and y_M boundary points, we would take our Laplacian $\underline{\underline{L}}$ and obtain $\tilde{\underline{\underline{L}}}$ by deleting the 1st and Mth columns and rows, i.e. keeping the green portion of the matrix below:

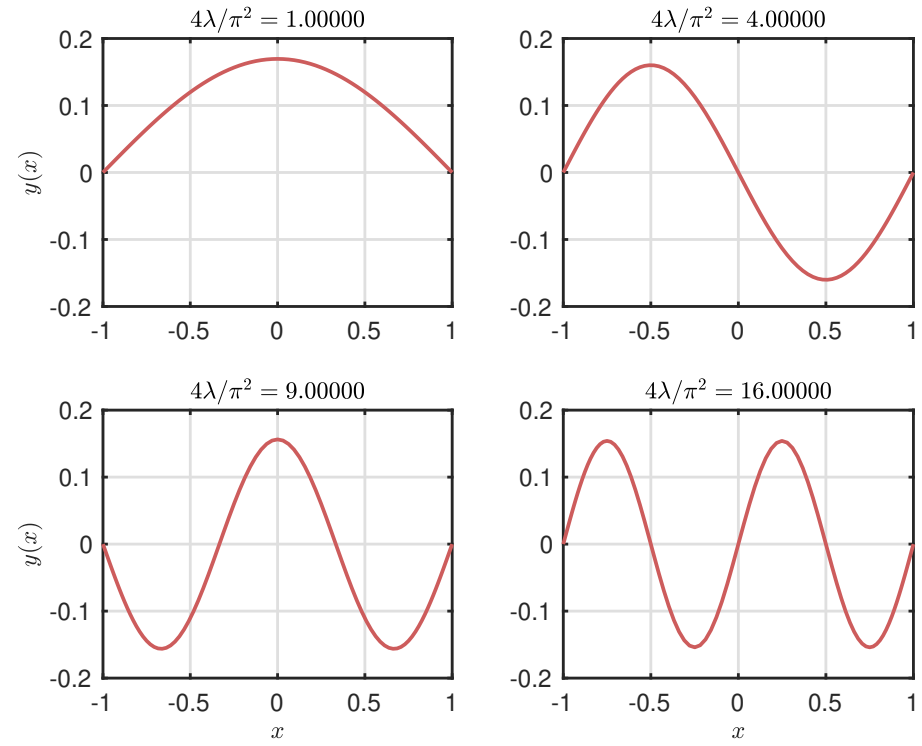
$$\underline{\underline{L}} = \begin{pmatrix} L_{00} & L_{01} & ... & L_{0M} \\ L_{10} & L_{11} & ... & L_{1M} \\ \vdots & \ddots & \ddots & \vdots \\ L_{M0} & L_{M1} & ... & L_{MM} \end{pmatrix}$$

Following the outlined procedure, we can then solve the eigenvalue problems given by Equation 4.21 and Equation 4.22, as shown in Figure 4.5a and Figure 4.5b respectively. Indeed, Figure 4.5 shows the first 4 non-trivial eigenfunctions and eigenvalues, obtained with a grid of $N = 100$ points for both Neumann and Dirichlet boundary conditions. The accuracy of the eigenvalues computed numerically for both sets of boundary conditions is then shown in Figure 4.6, where the relative error for various grid resolutions is displayed. Here we start to see the incredible precision of spectral methods: even with only 10 points, we get the smallest eigenvalues accurate up to a part in 10^{10}. We see that the error stays somewhat constant for most of the spectrum, until a threshold is hit, after which it rapidly increases. This can be explained by the fact that as n grows, the eigenfunctions oscillate faster and faster. At some point, we incur in aliasing, i.e. the nodal spacing are not fine enough to resolve the eigenfunctions (especially in the centre, where the separation between lattice points is larger), leading to a rapidly increasing error.

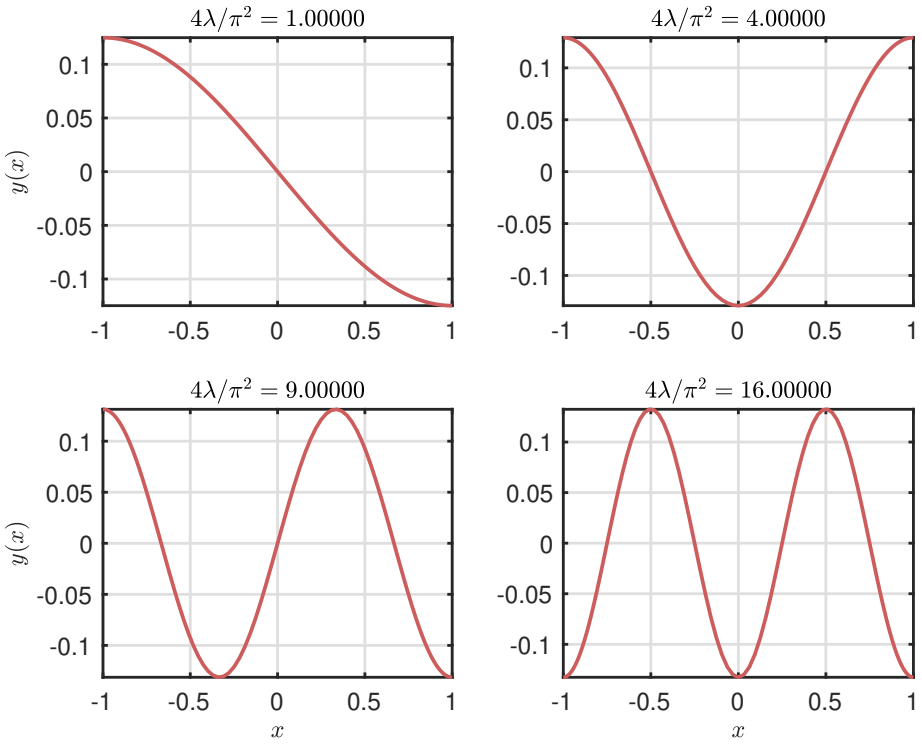
4.2 Computational Set Up

4.2.1 Tensor Product Grid

A crucial difference between the 1D toy model that we just discussed and the case of the Laplacian on a sphere is obviously the dimensionality of the problem. Indeed, the sphere being two dimensional requires a 2D grid, which we naturally obtain by formally performing a tensor product between two 1D grids. Effectively, we can take a 1D grid in one direction and replace all its points with a 1D grid in the other

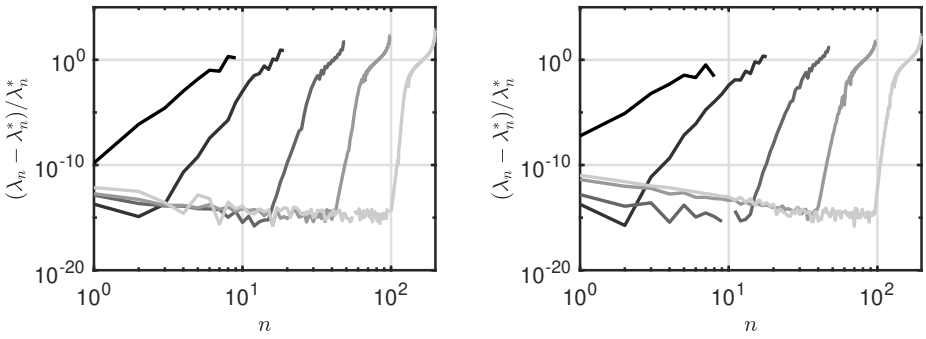


(a)



(b)

Figure 4.5: The 4 smallest eigenvalues and eigenfunctions of the 1D Laplacian on the interval $[-1, 1]$ subject to (a) homogeneous Dirichlet and (b) homogeneous Neumann boundary conditions. Results obtained with $N = 100$ nodes.



(a) Dirichlet BC's (b) Neumann BC's

Figure 4.6: Relative error with respect to the theoretical values λ_n^* for the eigenvalues λ_n of the 1D Laplacian. We compare the results obtained with a grid of $N = 10, 20, 50, 100, 200$ (darker corresponds to smaller N) points for homogeneous (a) Dirichlet (b) Neumann boundary conditions.

direction. First, however, we must pick a coordinate system on the sphere. We make the choice covering S^2 with two charts, each representing one hemisphere. Since we restrict ourselves to deformations which are symmetric about the equator, we can focus our attention to one of the two charts, spanning the region:

$$\theta \in [0, \pi/2], \qquad \phi \in [0, 2\pi].$$

We can see this chart as covering a disk, where the role of the radial coordinate is played by the θ direction, as shown in Figure 4.7. We now follow the standard approach in spectral methods when dealing with 2D polar coordinates on a disk [47], namely we use a Chebyshev nodal collocation in the radial coordinate and a periodic

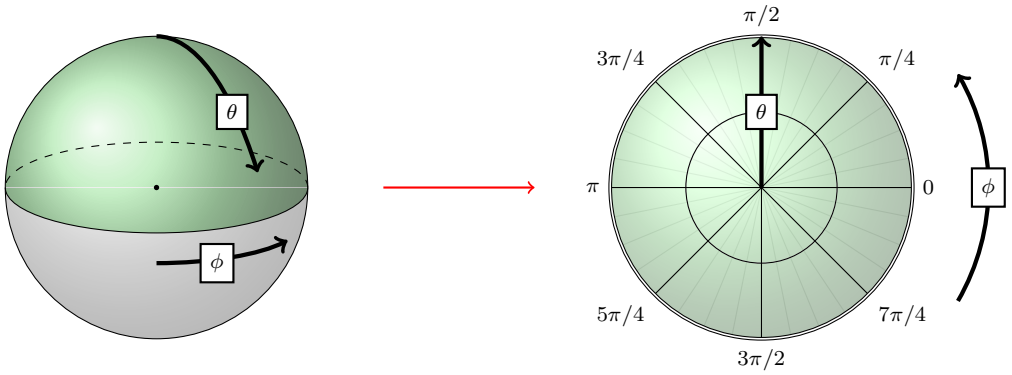


Figure 4.7: Choice of chart covering the sphere. Due to the symmetry of the problem with respect to the equatorial plane, only the top hemisphere is considered and treated as a disk, with the role of radial coordinate played by $\theta \in [0, \pi/2]$

Fourier grid in the angular direction. A Fourier grid is really an equispaced grid, with the difference that the differentiation matrices are evaluated not with polynomial interpolants (which would amount to a standard finite difference method), but in the Fourier basis (hence encapsulating periodicity on the ϕ direction). With the choice made for the θ grid, we end up with a clustering of points both near the equator and near the centre of the disk (representing either our North or South pole). While this would be fine in theory, in practice it is rarely of any use to have such a high density of nodes near the origin, being it a very small region. Indeed, since with spectral methods we always implicitly assume smoothness of our solutions (because the basis functions chosen are smooth), we would be wasting precious computational power, oversampling a not so interesting portion of the domain. A workaround to this is to extend the radial coordinate (in our case θ) to negative values, i.e. considering the coordinate system:

$$\theta \in [-\pi/2, \pi/2], \qquad \phi \in [0, 2\pi].$$

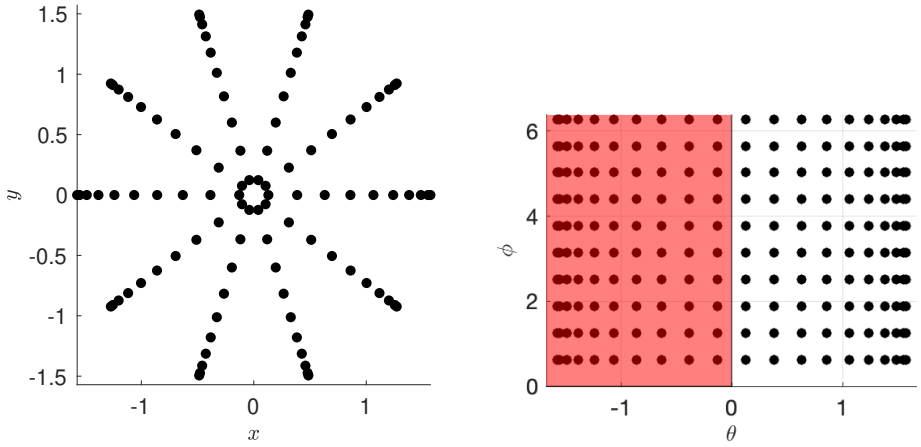
Of course, this means that there is some redundancy in our coordinate choice, since we are mapping each point twice! Indeed, the map from (θ, ϕ) to (x, y) is 2-to-1 (excluding the origin where it really is ∞-to-1, but taking an even number of nodes on the θ direction we'll knowingly avoid that point). To solve this issue, we must impose that our solution has the symmetry:

$$u(\theta, \phi) = u(-\theta, (\phi + \pi)(\mathrm{mod} 2\pi)), \tag{4.25}$$

which forces single-valuedness of the function. Here we are just requiring $u(\theta, \phi) = u(\theta, \phi + 2\pi m)$ for any integer m (i.e periodicity in the ϕ direction, which is what motivates the use of the Fourier grid for that coordinate in the first place) and that the reflection of the radial coordinate $\theta \to -\theta$ really amounts to a rotation by π in the azimuthal direction. We now show how to implement the latter symmetry condition numerically, restricting ourselves once more to 1D. Consider our state vector $\vec{u}$ sampled on $2N$ grid points with Chebyshev nodal collocation, and divide it into two sub-vectors $\vec{u}_+$ and $\vec{u}_-$, representing respectively u sampled at positive and negative radial coordinate. Clearly, the condition given by Equation 4.25 really amounts to $\vec{u}_+ = \vec{u}_-$ (if entries are reordered such that if u_i corresponds to coordinate θ, then u_{i+N} corresponds to coordinate $-\theta$). With this in mind, consider a matrix $\underline{\underline{A}}$ representing any linear operation:

$$\underline{\underline{A}}\vec{u} = \begin{pmatrix} \underline{\underline{A_{++}}} & \underline{\underline{A_{+-}}} \\ \underline{\underline{A_{-+}}} & \underline{\underline{A_{--}}} \end{pmatrix} \begin{pmatrix} \vec{u}_+ \\ \vec{u}_- \end{pmatrix} = \begin{pmatrix} \vec{w}_+ \\ \vec{w}_- \end{pmatrix}. \tag{4.26}$$

The vector $\vec{w}$ so obtained must necessarily possess the same symmetry as $\vec{u}$ with respect to $\theta \to -\theta$. Imposing both the symmetry conditions and Equation 4.26, we



(a) (b)

Figure 4.8: Computational 10×10 grid, shown (a) on the $x - y$ plane and (b) on the $\theta - \phi$ plane. In (b) we show how the grid is formally generated, first collocating 20 Chebyshev nodes on the interval $\theta \in [-\pi/2, \pi/2]$ and discarding then the reduntant $\theta < 0$ region such that the neighborhood of the pole is not oversampled.

obtain:

$$\vec{w}_+ = (\underline{\underline{A_{++}}} + \underline{\underline{A_{+-}}})\vec{u}_+ = (\underline{\underline{A_{-+}}} + \underline{\underline{A_{--}}})\vec{u}_+, \tag{4.27}$$

from which we see that for any $2N \times 2N$ differential operator, the symmetry requirements restrict our matrix to a $N \times N$ one given by the sum of the top two (or, equivalently, bottom two) blocks and acting only on $\vec{u}_+$.

The last piece of the puzzle needed before constructing the Laplacian on the sphere is to understand what the differentiation matrix with respect to only one of the two coordinates is on a tensor product grid. Suppose we have a $2N \times M$ grid with $2N$ lattice points in the θ direction and M nodes for the ϕ coordinate. The required operator is easlily obtained by taking the Kronecker tensor product (namely the usual tensor product between two matrices) between the differentiation matrix with respect to the desired coordinate and the identity matrix defined on the other direction. For example, if we want to construct the operator giving the 1st partial derivative with respect to θ, we would take the $2N \times 2N$ $\underline{\underline{D^{\theta}_{2N-1}}}$ Chebyshev differentiation matrix and tensor it with the $M \times M$ identity matrix $\mathbb{1}_M$, i.e.:

$$\underline{\underline{D_\theta}} = \underline{\underline{D^{\theta}_{2N-1}}} \otimes \mathbb{1}_M,$$

obtaining a $2MN \times 2MN$ matrix acting on the full tensor product grid. This is intuitive, and reminiscent of when, in quantum mechanics, we want to construct an operator that acts on a subspace of our Hilbert space, leaving the rest unaltered. Similarly, the matrix giving the partial derivative with respect to ϕ is obtained

through:

$$\underline{\underline{D_\phi}} = \mathbb{1}_{2N} \otimes \underline{\underline{D_M^\phi}},$$

where $\underline{\underline{D_M^\phi}}$ is the differentiation matrix on the M-point Fourier grid. The careful reader will realise that while the 1D diferentiation matrices are not sparse, the operators acting on the full grid will be! Working with such big matrices can in practice be very expensive with regards to memory, but there are routines optimised for sparse matrices that can minimise the storage costs.

In reality, the story is a bit more complicated. Indeed, we've forgotten about the symmetry condition given by Equation 4.25, which we still need to apply following Equation 4.27! However, there is a further tweak we need to do before imposing it. When we generate the grid points for the θ coordinate they will naturally arrange themselves monotonically with θ, meaning that $\vec{u}_-$ is equal to $\vec{u}_+$ only when its order is reversed. To implement this reversal, we can tensor the $+-$ block of $\underline{\underline{D_M^\theta}}$ with:

$$\begin{pmatrix} 0 & \mathbb{1}_{M/2} \\ \mathbb{1}_{M/2} & 0 \end{pmatrix}.$$

We can consequently see that the actual $N \times M$ matrices that give the partial dervivatives on our tensor product grid are (dropping the dimensionality subscripts on the differentiation matrices for neatness):

$$\underline{\underline{D_\theta}} = \underline{\underline{D_{++}^\theta}} \otimes \mathbb{1}_M + \underline{\underline{D_{+-}^\theta}} \otimes \begin{pmatrix} 0 & \mathbb{1}_{M/2} \\ \mathbb{1}_{M/2} & 0 \end{pmatrix}, \qquad \underline{\underline{D_\phi}} = \mathbb{1}_N \otimes \underline{\underline{D^\phi}}. \tag{4.28}$$

The same can be done for the partial second derivative matrices, which means that we now have all the ingredients to construct the discretised Laplacian on the sphere. Calling $\underline{\underline{\Delta_{2N-1}^\theta}}$ the matrix giving the second derivative on the θ direction, and once again ignoring the dimensionality subscripts for tidiness, the differential operator required is represented by:

$$\underline{\underline{L}} = (\underline{\underline{D_{++}^\theta}} + \underline{\underline{C}}\,\underline{\underline{\Delta_{++}^\theta}}) \otimes \mathbb{1}_M + (\underline{\underline{D_{+-}^\theta}} + \underline{\underline{C}}\,\underline{\underline{\Delta_{+-}^\theta}}) \otimes \begin{pmatrix} 0 & \mathbb{1}_{M/2} \\ \mathbb{1}_{M/2} & 0 \end{pmatrix} + \underline{\underline{S}}^2 \otimes \underline{\underline{D^\phi}}^2, \tag{4.29}$$

where:

$$\underline{\underline{C}} = \mathrm{diag}(\cot\theta_j), \qquad \underline{\underline{S}} = \mathrm{diag}(\csc\theta_j) \qquad j = 1, 2, ..., N.$$

To obtain the Laplacian on the deformed sphere, we can just multiply $\underline{\underline{L}}$ with a diagonal matrix whose diagonal entries are given by the vector of sampled values of the exponential conformal factor. Note that, from now on, when we talk about a $N \times M$ grid, we actually mean the grid *after* the $\theta \to -\theta$ symmetry has been implemented, so N can be indiscriminately even or odd. Figure 4.8 displays the tensor product grid we discussed, with Figure 4.8b in particular illustrating how the

lattice is set up not to incur in oversampling of the pole region.

4.2.2 Round Sphere

Now that we have the discretised Laplacian, the last thing we need to do before extracting the eigenvalues is to enforce the boundary conditions. This is trivially done by extending the procedure discussed in 1D.

For the physical situation at hand, we are interested in enforcing either homogeneous Neumann or homogeneous Dirichlet boundary conditions at the equator. This is motivated by the fact that, focusing only on half of the sphere, we are assuming that the deformation function is symmetric with respect to the equatorial plane. In order to obey the symmetry of the problem, the eigenfunctions will necessarily be either even or odd with respect to such a plane, corresponding respectively to homogeneous Neumann (with respect to the partial derivative in the θ direction) or homogeneous Dirichlet boundary conditions at $\theta = \pi/2$. Indeed, this is also true for the trivial deformation (i.e. $\chi = 0$): the eigenmodes of the Laplacian on a sphere (i.e. the spherical harmonics) are either even or odd with respect to the equator. Enforcing Neumann boundary condition amounts to having the matrix of

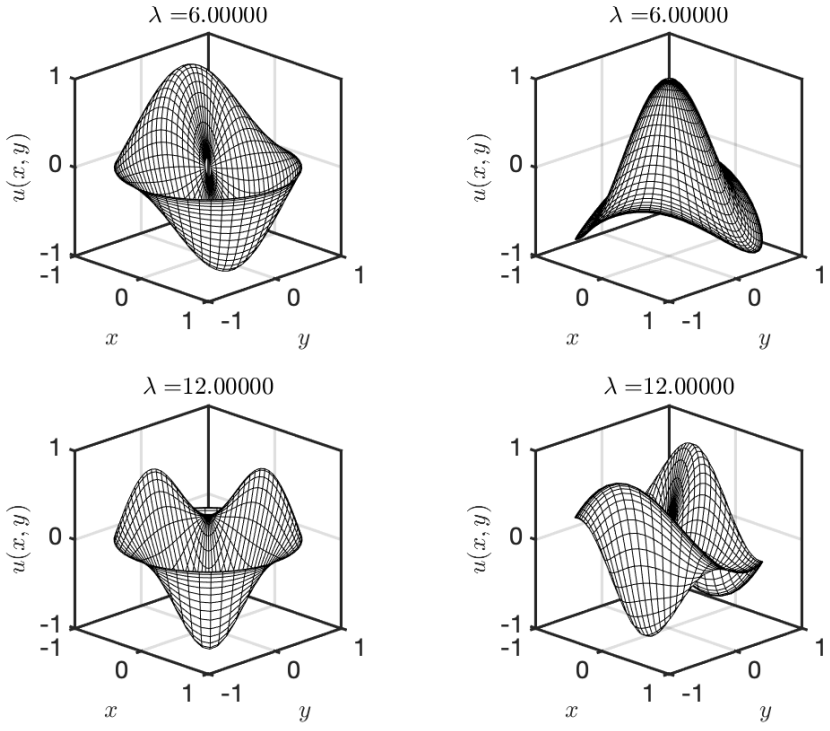


Figure 4.9: 4 eigenmodes of the Laplacian on the unit hemisphere. The left column was obtained imposing homogeneous Dirichlet BC's on the equator, while the right one follows homogeneous Neumann BC's. The computation was performed on a 25×50 grid.

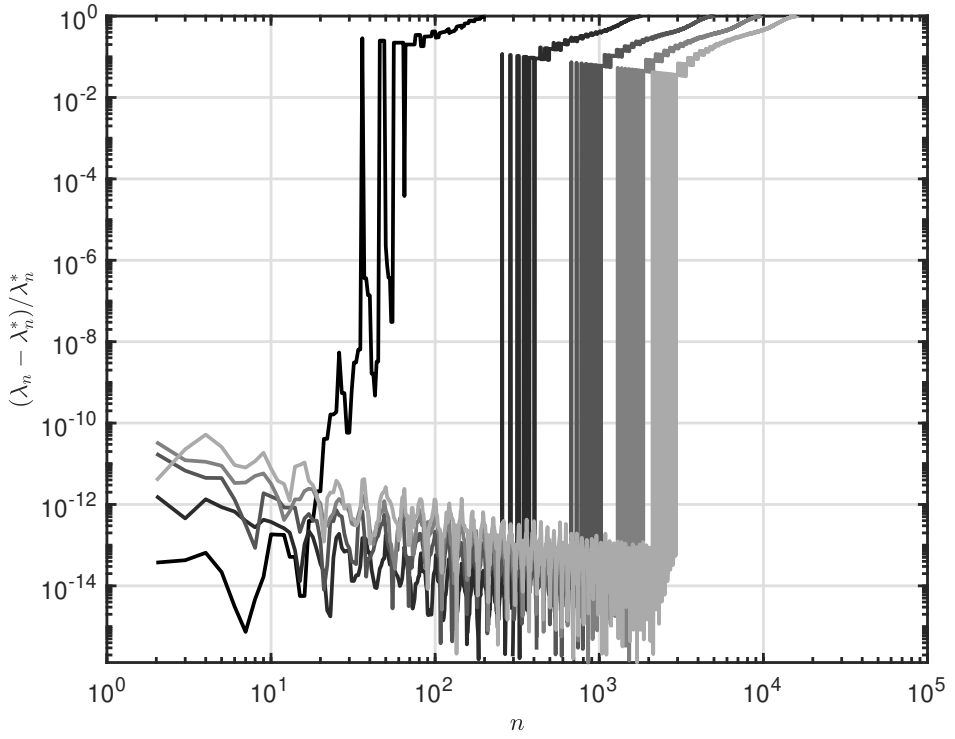


Figure 4.10: Relative error with respect to the theoretical results λ_n^* for the eigenvalues of the Laplacian on the unit hemisphere λ_n evaluated with a $N \times N$ grid with $N = 10, 30, 50, 70, 90$ (darker corresponds to smaller N).

constraints $\underline{\underline{B}}$ composed of the rows of $\underline{\underline{D_\theta}}$ that give the θ partial derivative on the $\theta = \pi/2$ points. On the other hand, deleting the rows and columns of the Laplacian that correspond to the equatorial nodes imposes homogeneous Dirichlet boundary conditions. The periodic conditions on ϕ are already accounted for by the Fourier grid.

To benchmark our method, we numerically evaluated the spectrum of the Laplacian on the undeformed sphere, similarly to what we have done before with the 1D Laplacian. Results are shown in Figure 4.9, where we've plotted 4 of the evaluated eigenfunctions, and in Figure 4.10 where the relative error of the estimated eigenvalues is shown. Here, the precision of spectral methods is once more manifest. Indeed, even with a 50×50 grid we obtain the 200 smallest eigenvalues correct with a part in 10^{13}. The error then increases, albeit still remaining very small, with the 1000th eigenvalues deviating from the theoretical result by only 1% of its magnitude. Again, we see the error steeply increasing once a threshold is hit, suggesting that aliasing on the ϕ direction should be the main cause for the accuracy dropping so quickly. To investigate this hypothesis, we compared the spectra of eigenvalues obtained with rectangular grids of various sizes in Figure 4.11. In Figure 4.11a we

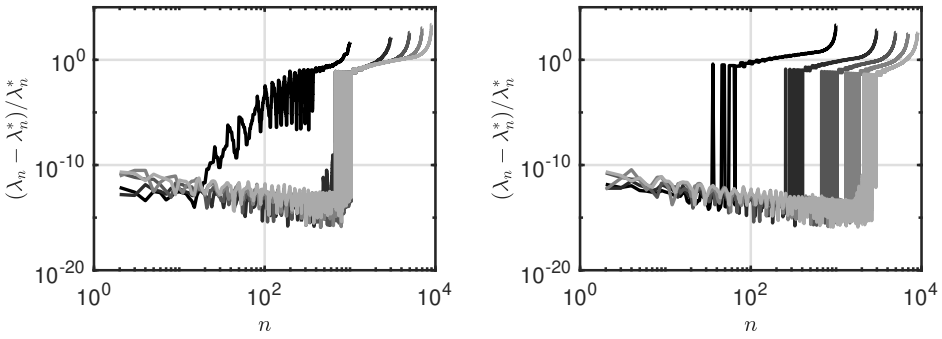


Figure 4.11: Relative error with respect to the theoretical results λ_n^* of the eigenvalues λ_n of the Laplacian on the unit hemisphere estimated with a $N \times 50$ grid for $N = 10, 30, 50, 70, 90$ (darker corresponds to smaller N). The two panels show: (a) fixed number of nodes on the ϕ coordinate and (b) fixed number of nodes on the θ coordinate.

see that increasing the density of nodes on the θ coordinate has mainly an effect on the maximum precision of the estimated eigenvalues, but does not change the threshold after which the error on the eigenvalues grows significantly. On the contrary, Figure 4.11b displays the effect of increasing the number of lattice points in the ϕ direction, which pushes that boundary further and further. This reinforces the idea that the failure to encapsulate the biggest eigenvalues to high precision is mainly due to aliasing, and suggests that the best geometry for the grid (for fixed total number of nodes) is rectangular, with higher density on the ϕ direction. In particular, a $N/2 \times N$ mesh seems to be an ideal compromise and we will adopt it going forward.

Chapter 5

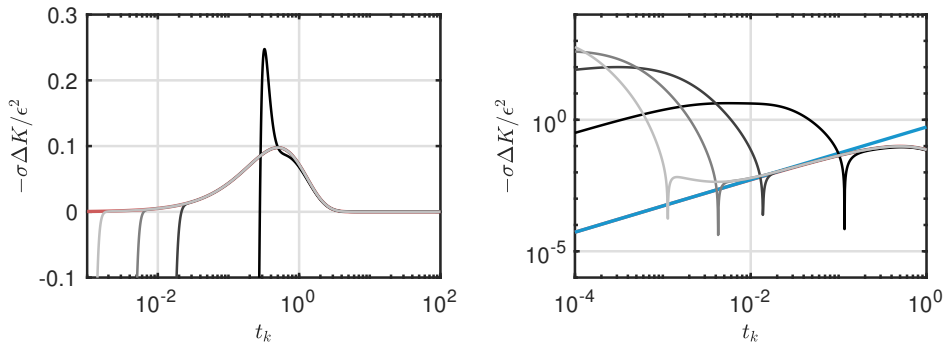
Results

In Chapter 3 we reviewed recent analytical and numerical results that suggest the counterintuitive conclusion that S^2 maximises the free energy of the scalar field (and, possibly, higher spin fields as well), making it an unstable configuration. We now present new findings obtained through the computational methods described in Chapter 4, that extend the above results to a more general class of deformations for the non-interacting scalar. We first mention the difficulties of calculate the subtracted heat kernel for small t_k, discussing how we can go around it by patching the analytical small time asymptotics to the numerically evaluated kernel. We then move on to the novel results, starting with an analysis of the behaviour of the subtracted heat kernel under deformations of increasing strength, followed by a comparison between axisymmetric and non-axisymmetric distortions. We then evaluate the subtracted free energy for such perturbations, further studying the effect that adding a mass term to our field or introducing a non-zero temperature have on the system. We finally repeat most of the above analysis for a non-minimally coupled scalar, finding that the contribution from the coupling to the Ricci scalar appears as a noticeable peak in the profile of the subtracted heat kernel.

5.1 Limitations: the Small t_k Regime

We now turn our attention to the evaluation of the subtracted heat kernel, the key object in our numerical investigations. As shown in Equation 2.63, this can be computed from the eigenvalues of the Laplacian on the deformed and undeformed

sphere. Of course, we do not have access to the full spectra of the differential operators due to the discrete (and finite) nature of the matrices we use to represent them. It's then natural to ask how the evaluation of ΔK (we drop the L subscript from now on) is affected in different regimes. As discussed in Chapter 3, only the smallest (non-trivial) eigenvalue contributes for big enough t_k. Working backwards from $t_k \to \infty$, an ever increasing number of eigenvalues start to give a non-negligible contribution, until for very small time our truncated spectra are not enough to accurately evaluate the kernel. Therefore, we expect that the number of eigenvalues that we have access to inevitably affects our ability to compute the small t_k behaviour. The size of the spectra that we can numerically produce is fundamentally limited by the dimensions of the matrix representing our Laplacian, which are in turn fixed by the dimensions of the computational grid. Fortunately, we know what the small t_k behaviour must be for any deformation, thanks to the kernel expansion given by Equation 3.38. This means that, as long as we have large enough spectra that allows us to compute ΔK until it approaches the linear behaviour, we can patch the numerically evaluated kernel with the small t_k expansion to successfully capture all of the physics. In order to reproduce the linear regime, we can just evaluate the coefficient of the kernel expansion by numerically integrating Equation 3.39. Nonetheless, it is sometimes computationally very expensive to compute a sufficient number of eigenvalues to get to small enough time such that the patching can be done, especially for the big deformation cases.



(a) (b)

Figure 5.1: Comparison between the perturbative analytical results (red line) and the numerical computation (grayscale lines) of $\Delta K(t_k)$ for a small deformation with $l = 3$, $m = 1$ and $\epsilon = 0.001$. Numerical calculations are performed on a $N/2 \times N$ grid with $N = 10, 50, 100, 200$ (darker corresponds to smaller N). (a) Semilog plot of the calculated heat kernel profile. (b) Loglog plot of the small t_k region, with the linear theoretical behaviour in blue. In the right panel, the red line corresponding to the analytical perturbative curve is completely overlaid by the linear approximation for $t \lessapprox 10^{-1}$ and, as it diverges from the small t_k behaviour and the numerical results converge to the theoretical results, by the computational calculations.

Figure 5.1 illustrates the performance of the numerical methods in reproducing the perturbative analytical results for small ϵ deformations. In particular, Figure 5.1 shows the relative error between the numerically computed kernel ΔK and the theoretical perturbative result ΔK_0 for different grid densities in the case of a $l = 3$, $m = 1$ deformation with $\epsilon = 0.001$. While it demonstrates once more the power of spectral methods, which give extremely accurate results for $t_{ker} \gtrapprox 0.1$ with already a 5×10 grid, it also shows that even a 100×200 mesh struggles to reproduce the subtracted kernel for $t_{ker} \lessapprox 10^{-2}$. A closer look at the linear regime is displayed in Figure 5.1b, a loglog plot of the interval $10^{-4} < t_k < 1$. It is first important to notice that the linear small t_k expansion is valid only for $t_k \lessapprox 10^{-1}$, after which the perturbative analytical results deviates from the linear behaviour (recall that in a loglog plot the a straight line corresponds to polynomial function, with the slope being the order of the polynomial and the vertical offset giving the leading coefficient). Further, we see that it becomes progressively harder to compute the kernel for small t_k, with the 100×200 grid only managing to get slightly below $t_k \approx 10^{-2}$,

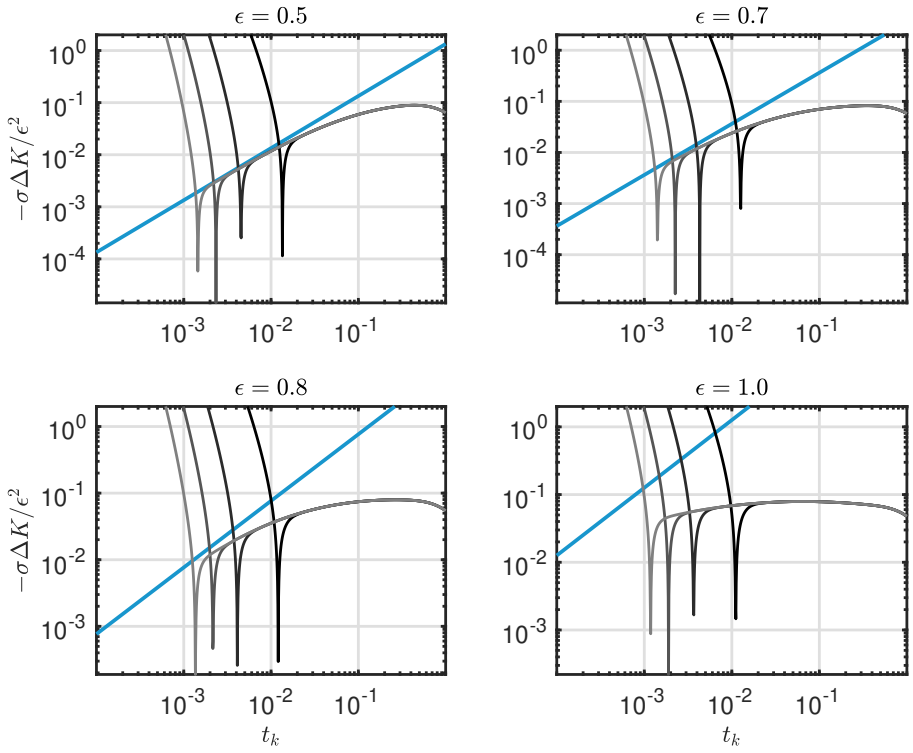


Figure 5.2: Comparison between analytical results and numerical computations of $\Delta K(t_k)$ for small t_k. Each of the 4 panels shows the subtracted heat kernel a $l = 4$, $m = 2$ deformation with a different ϵ. Grayscale lines show the result evaluated with a $N/2 \times N$ grid with $N = 50, 100, 150, 200$ (lighter corresponds to bigger N). The solid blue line gives the analytical small t_k behaviour extracted from the heat kernel expansion.

notwithstanding the high computational burden. Since we expect high ϵ deformation to anticipate significantly the deviation from the linear regime, it is possible that we will not be able to cover adequately the small t_k region in order to patch the numerical result with the linear expansion in those cases. An example of this is given by Figure 5.2, where we compare the numerically calculated ΔK against the linear asymptotics at small t_k in the case of a $l = 4$, $m = 2$ deformation. We see that we struggle to reach the linear behaviour even with the extremely computationally expensive 100×200 mesh already for a $\epsilon = 0.8$ deformation. On the contrary, we observe that also the 50×100 grid manages to achieve such regime comfortably for $\epsilon = 0.5$. In light of this analysis, we restrict our attention to deformations below $\epsilon = 0.5$ when calculating the subtracted free energy, so that we are able to patch our numerical result with the analytical expansion at small t_k with a reasonable computational burden. Nonetheless, we are still allowed to investigate the behaviour of the subtracted heat kernel for bigger deformations, keeping in mind that, when we do so, we are not able to capture the high energy contributions coming from $t_k \lessapprox 10^{-2}$.

5.2 Minimally Coupled Scalar

We begin by studying the simpler case of the minimally coupled scalar field. Recall that we use perturbations of the form given by Equation 3.14 and Equation 3.15, in order to both guarantee reality of $\chi^{(l,m)}$ and at the same time to have equal strength for axisymmetric and non-axisymmetric deformations. Furthermore, due to the fact that we are working with one 'hemisphere' of Σ only, keep in mind that we are forced to consider even m for even l and odd m for odd l, meaning that we are able to take into consideration the axisymmetric $m = 0$ deformations only if l is even. Moreover, we avoid $l = 1$ distortions since, to lowest order, they correspond to diffeomorphisms of the round sphere, giving a vanishing subtracted heat kernel.

5.2.1 Subtracted Heat Kernel

We just discussed that for the tensor product grid we deploy it is more efficient to use a rectangular $N/2 \times N$ lattice, with N being the number of points in the ϕ direction. For the current analysis, we take $N = 200$. Even though such a dense grid requires extensive computational resources, it is illustrative to adopt it for now since it allows for computation of the kernel up to $t_k \approx 10^{-3}$, enabling the possibility to look at more significant deformations. It is however impractical to go much below such limit, contrary to what other studies that restricted their attention to axisymmetric deformations have been able to do. This is due to the fact that, because of the axisymmetry, in those cases it is enough to look at a $\phi =$ constant slices, making the problem 1D and therefore more tractable.

Figure 5.3 shows the profile of the subtracted heat kernel (normalised by the square of the deformation parameter to ease visualisation) for different l while varying ϵ (lighter corresponds to bigger ϵ). Here we took $m = 1$ and $m = 2$ for odd and even l respectively. Recall that we have defined $\sigma = -1/2$. The first thing that jumps to the eye is that ΔK is positive throughout the domain for the deformations considered. While this is harder to state with certainty for the highest ϵ studied, where the numerically evaluated kernel doesn't quite reach the linear behaviour before failing at $t_k \approx 10^{-3}$, it is definitely true for the other distortions. If for the $\epsilon = 1$ case (for example) it does become negative outside the domain of validity of our computational result, it has to do so in a very small region below $t_k \lessapprox 10^{-3}$, before becoming positive again and only then approach the linear behaviour (which we now to have $\Delta K \geq 0$ from the heat kernel expansion). Not only this seems unnatural, but it is even harder that, if it were to happen, upon integration such negative would area more than offset the contribution from the positive region. As expected, the effect that increasing the parameter of the deformation seems to have

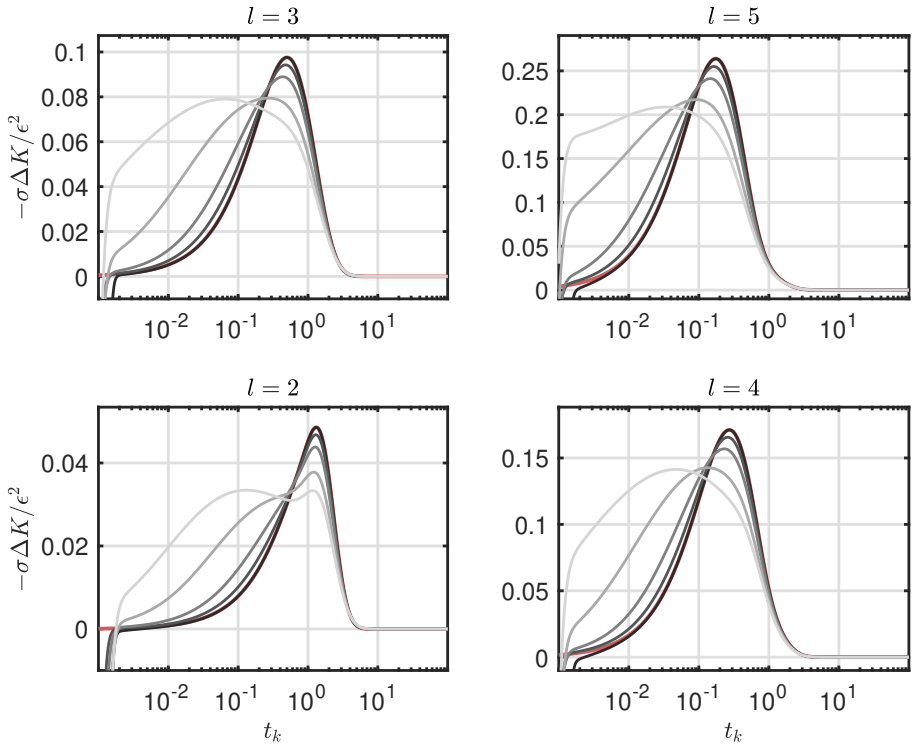


Figure 5.3: subtracted heat kernel for a minimally coupled scalar for $l = 2, 3, 4, 5$, normalised by the square of the deformation parameter. We use $m = 1$ and $m = 2$ for respectively even and odd l. Each panel shows the perturbative result (red line) and the numerical calculations (grayscale lines) for $\epsilon \in [0.01, 0.3, 0.5, 0.8, 1]$ (lighter corresponds to bigger ϵ). Results obtained for a computational grid with $N = 200$.

on the heat kernel is to significantly anticipate the cutoff time after which the linear behaviour fails, by effectively moving the maximum of the curve to earlier times. Furthermore, while the peak becomes relatively smaller (recall that we're normalising by ϵ^2, so the maximum is not actually decreasing), it also gets significantly wider. From this behaviour it is sensible to anticipate that we expect the absolute value of the subtracted free energy to increase as the deformation becomes stronger. Note that we see the curves approach the perturbative expansion for decreasing ϵ (the red solid line in the figure, almost completely overlaid by the small distortion results), as expected. At this stage, no significant deviation is observed between the ΔK computed with different l, apart from the $l = 2$ case. Indeed, we see that contrary to the other perturbations considered, $l = 2$ deformations develop a secondary peak for smaller t_k as ϵ increases. Why this does not happen for higher l requires further analysis, but it might be related to the fact that $l = 2$ is the lowest value of l such that the subtracted heat kernel is non-zero up to non-trivial order. This maybe allows contributions that become negligible for bigger l to be instead noticeable here. Further, we see that increasing l has also the effect of inflating the

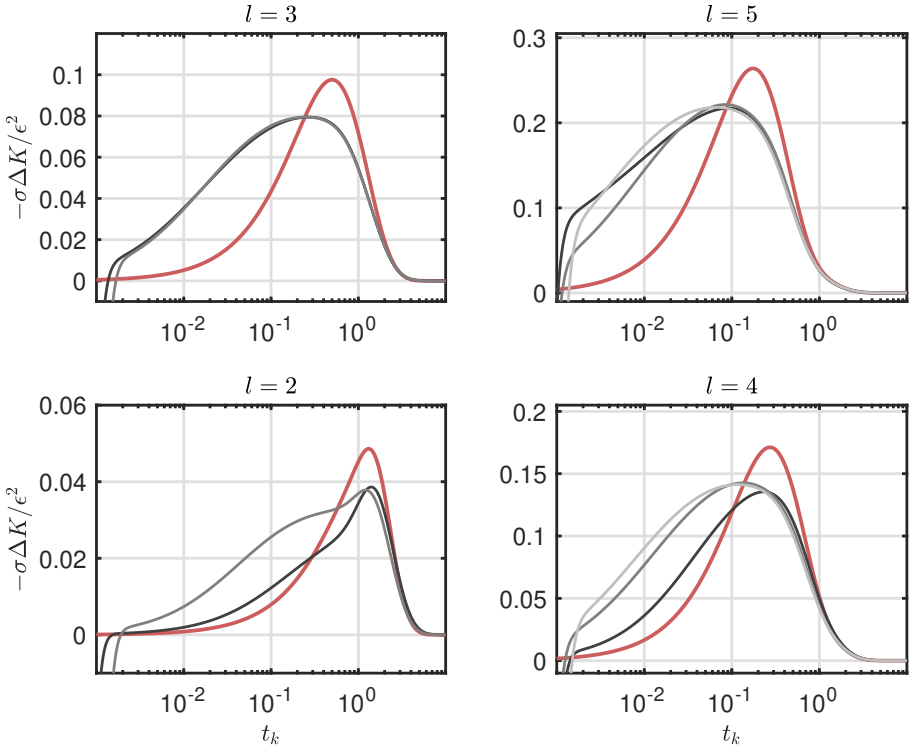


Figure 5.4: subtracted heat kernel for a minimally coupled scalar for $l = 2, 3, 4, 5$, normalised by the square of the deformation parameter and $\epsilon = 0.8$. Each panel shows the perturbative result (red line) and the numerical calculations (grayscale lines) for $0 \leq m \leq l$ (lighter corresponds to bigger m). Recall, that we take only m of the same parity of l. Results obtained for a computational grid with $N = 200$.

overall magnitude of the subtracted kernel, mirroring the naive expectation that higher l correspond to more energetic configurations.

We can see the effect that changing the m parameter has on the kernel in Figure 5.4. Here the subtracted heat kernel for deformations with the same set of l as before and $\epsilon = 0.8$ is displayed, with the grayscale lines represent the numerical solutions for different m (lighter corresponds to increasing m). As discussed earlier, the perturbative solution is really insensitive to m, so we expect that, if there is any difference between ΔK obtained with distinct m, it should be evident only for big enough deformations. Looking at Figure 5.3, we see that $\epsilon = 0.8$ is the highest we can go without throwing away potentially interesting parts of ΔK with the chosen resolution, hence the choice. First of all, looking at the even l curves, we notice a significant difference between the axisymmetric and non-axisymmetric deformations, with the latter subtending a larger area and hence corresponding to a more negative subtracted free energy. Moreover, we see that increasing l while leaving ϵ invariant enhances the differences between curves of distinct m. In the $l = 3$ case, the $m = 1$ and $m = 3$ heat kernels are practically overlaid, while for $l = 5$ we can clearly see three distinct functions. Furthermore, the differences become more evident in the small t_k limit, suggesting that they originate from how high energy modes interact with the geometry. It is counterintuitive that the curves do not change 'monotonically' with m in the $l = 5$ case. It might be interesting to see whether this holds for bigger l as well, i.e. whether it turns out that there is just no hierarchy with increasing m.

5.2.2 Subtracted Free Energy

We now turn our attention to computing the actual subtracted free energy of the scalar field. This is easily achieved by integrating numerically the subtracted heat kernel as shown by Equation 3.28. As mentioned earlier, we patch the computational result with the small t_k behaviour at t_c (i.e. the cutoff time below which we cannot trust our numerical solution), but to do so we need a dense enough grid such that the evaluated ΔK has indeed approached the linear regime at the cutoff. For the present discussion, we opted to focus on deformations up to $\epsilon = 0.5$ and we use a computational grid with $N = 120$, since we have seen before that such choice does manage to reasonably approach the linear behaviour by the chosen cutoff time of $t_c = 6 \times 10^{-3}$. From now on, unless otherwise stated, we stick with the choice $N = 120$. Hence, we integrate the analytic expression for the small t_k expansion up to t_c, and then add the integral of the numerically estimated kernel. The latter integral is truncated at $t_k = 10^3$, after which the contribution is negligible. We start by considering a massless scalar field at zero temperature, recalling that in this 'cold' limit we obtain the Casimir vacuum energy, i.e. $\Delta F \to \Delta E$. Figure 5.5 shows the result as a function of ϵ for deformations of different l and m. As before, darker

markers correspond to smaller m, while the red circles are used to highlight the axisymmetric $m = 0$ case. Solid lines are just splines interpolants drawn to guide the eye. We further normalise the subtracted free energy by both ϵ^2 and the perturbative result $\Delta E_0/\epsilon_0^2$ (with ϵ_0 being the parameter at which ΔE_0 is calculated). Note that the perturbative subtracted energy is negative, meaning that all the calculated data points do indeed correspond to negative ΔE. Qualitatively, we observe exactly what we could have expected from Figure 5.3 and Figure 5.4, namely that the subtracted free energy is more negative for non-axisymmetric perturbations, showing that the field disfavours the more symmetric configuration. It also has to be emphasised that although the axisymmetric curves seem suggest that there is a local minimum, with $\Delta E/\Delta E_0$ first decreasing, this is actually not the case, since recall that we are normalising by ϵ^2/ϵ_0^2. In fact, all curves show a monotonically decreasing ΔE. It is also interesting to note that, for the $l = 5$ deformation, the behaviour of the subtracted free energy is extremely close between $m = 1$ and $m = 5$ in the interval

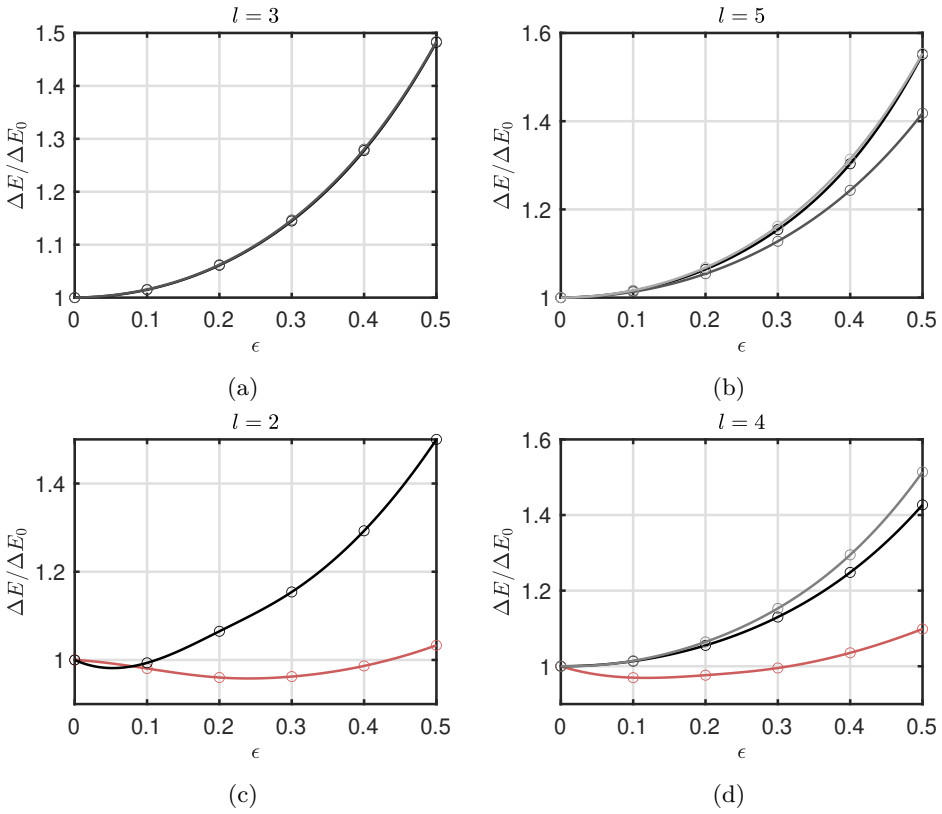
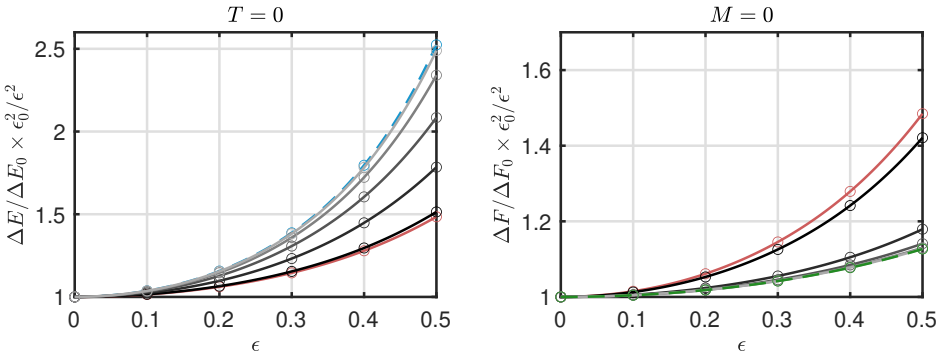


Figure 5.5: subtracted free energy for a massless minimally coupled scalar at zero temerature for $l = 2, 3, 4, 5$ as a function of ϵ. The subtracted free energy is first normalised by the ϵ at which it is calculated, and then by the perturbative result ΔE_0. Each panel shows the numerical results (grayscale markers) for $1 \leq m \leq l$ (lighter corresponds to bigger m). The red markers display the $m = 0$ axisymmetric results. Note that the solid lines are splines interpolators, drawn to guide the eyes.



(a) (b)

Figure 5.6: Subtracted free energy as a function of ϵ for the representative example of $l = 3$ and $m = 1$ with non-zero masses and temperatures. The grayscale markers correspond to fields with (a) $M = 0.5, 1, 2, 4, 8, 16$ and $T = 0$ and (b) $M = 0$ and $T = 0.5, 1, 2, 4, 8, 16$, with smaller value of the changing parameter being represented by a darker circle. The red markers indicate the $M = 0$, $T = 0$ results. The dashed blue and green lines give the long-wavelength and small temperature limits respectively. Note that most of the high T curves in (b) are overlaid in the hot limit.

under consideration, with $m = 3$ significantly deviating. It could be potentially interesting to look at other families of perturbations, and see whether this is just a feature of our particular choice of deformation function or if there is some significance for this result.

Let's now consider the effects of having non-zero mass M and temperature T for our theory by varying them independently as displayed by Figure 5.6 for the representative case of a $l = 3$, $m = 1$ deformation. In both figures, the red markers indicate the $M = 0$, $T = 0$. The blue dashed line gives the long wavelength behaviour, while the green one shows the 'hot' asymptotics. In Figure 5.6a the grayscale markers show the calculated subtracted free energy for a range of non-zero masses (lighter corresponds to bigger mass). We know that negativity of the free energy for a given configuration is bounded by the long wavelength limit given by Equation 3.43, and we indeed observe large M curves approach such bound from below, interpolating between the massless and long wavelength limits. The effect of thermal correction can be instead observed in Figure 5.6b, where we the greyscale markers indicate the calculated energy difference for a range of non-zero temperature (lighter corresponds to higher T). Here, we know that as $T \to \infty$ the curves should approach the limiting value given by Equation 3.42. This is indeed what we obtain, with the non-zero temperatures interpolating between the $M = 0, T = 0$ result and the 'hot' limit.

5.3 Non-minimal Coupling

We now proceed to study the effect of non-minimal coupling for the free scalar field. We repeat a similar analysis as for the minimally coupled case, first looking at how the subtracted heat kernel is modified by the new term and later computing the subtracted free energy as a function of ϵ. However, before jumping into the computation of the kernel, it is important to analyse the spectrum of the modified Laplacian for the non-minimally coupled field. Indeed, as discussed, for big enough deformations we expect the smallest eigenvalue to potentially become negative. This would amount to an effective negative mass which would mean incurring into an ill-defined theory with tachyonic instabilities. Figure 5.7 shows the lowest eigenvalue λ_{min} of the differential operator in Equation 3.24 for $l = 3$, $m = 1$ and $l = 4$, $m = 2$ deformations as a function of ϵ. Grayscale markers show the numerically computed eigenvalue for distinct ξ (lighter corresponds to a bigger coupling), while solid lines are drawn to guide the eye. As expected, for $\xi \to 0$ we obtain $\lambda_{\text{min}} \to 2\xi$. Moreover, we see that for increasing ϵ the eigenvalue monotonically decreases, with bigger rapidity as either the coupling to the Ricci scalar or l increases. Crucially, we

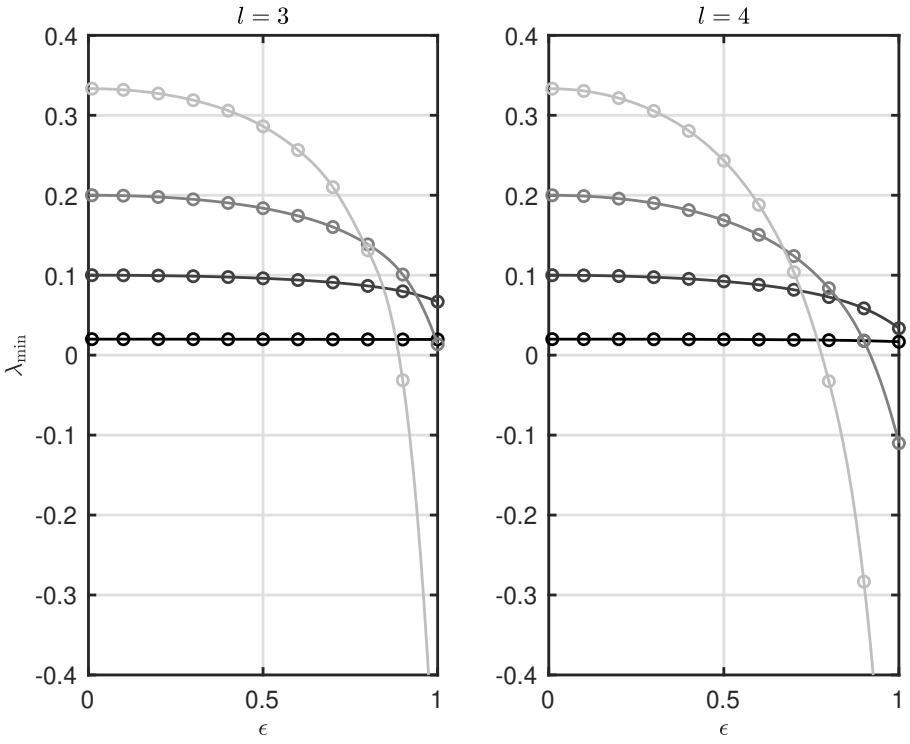


Figure 5.7: Lowest eigenvalue of non-minimally coupled scalar as a function of ϵ for $l = 3$, $m = 1$ and $l = 4$, $m = 2$ deformations. Grayscale markers illustrate the computed eigenvalue for $\xi = 0.01, 0.05, 0.1, 1/6$ (lighter corresponds to greater coupling) and the solid lines are interpolating splines drawn to guide the eye.

indeed observe negative eigenvalues for the $\xi = 1/6$ case even for a deformation with $\epsilon = 0.9$ and $l = 3$, and even sooner for the $l = 4$ perturbation. As later we consider $l = 4$ deformations with the same values of ξ as studied in Figure 5.7, we restrict ourselves to $0 < \epsilon \leq 0.7$, so that we are protected by the tachyonic instabilities and the theory is well behaved.

5.3.1 Subtracted Heat Kernel

We now compute ΔK for different values of the coupling constant ξ, so to understand how the contribution from the extra term shows up in the profile of the subtracted kernel.

The 4 panels of Figure 5.8 show the numerically computed heat kernel for a $l = 4$, $m = 2$ deformation with different ϵ (lighter corresponds to bigger ϵ) for the cases $\xi = 0.01, 0.05, 0.1, 1/6$. As usual, we normalise by ϵ^2 to aid visualisation and the red solid line is the perturbative result. The biggest difference in the profile of ΔK with respect to the minimally coupled case in Figure 5.3 is the presence of a

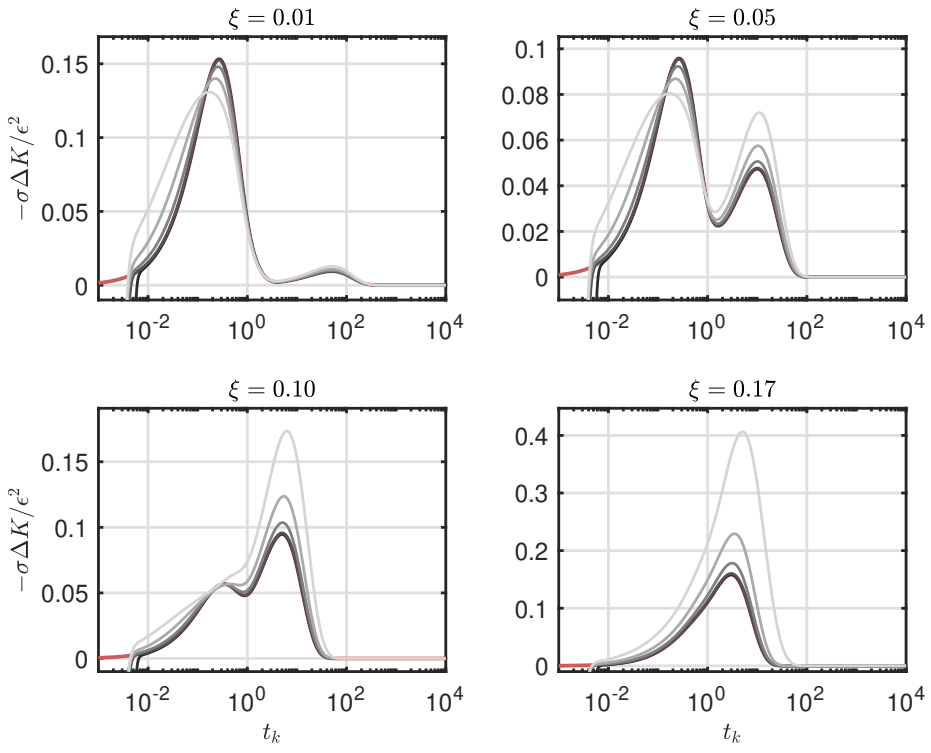


Figure 5.8: Subtracted heat kernel normalised by the square of the deformation parameter for a non-minimally coupled scalar with $\xi = 0.01, 0.05, 0.1, 1/6$ and a $l = 4$, $m = 2$ perturbation. Each panel shows the perturbative result (red line) and the numerical calculations (grayscale lines) for $\epsilon = 0.01, 0.1, 0.3, 0.5, 0.7$ (lighter corresponds to bigger m). Results obtained for a computational grid with $N = 100$.

secondary peak. This crest manifests itself for larger t_k than the original peak, and it is completely absent for $\xi = 0$. As ξ increases, its magnitude as well increases with respect to the other bell, and eventually 'eats' the latter. We can interpret this as a phase transition to a curvature dominated regime. It is suggestive that the transition from one to two crests happens slightly before $\xi = 1/6 \approx 0.17$, which is the value of the coupling for which the massless scalar becomes conformally invariant. Moreover, while bigger ϵ causes the normalised original peak to flatten, in the case of the crest related to the non-minimal coupling we see the opposite behaviour for the deformations under consideration. Furthermore, the variations to the profile of the curvature crest are much more pronounced, suggesting that we expect the subtracted free energy to be more negative for the non-minimally coupled scalar when comparing equivalent perturbations.

We further analyse the effect of m for the $l = 4$ deformations to the non-minimally coupled field in Figure 5.9. Here, the solid red line is again the perturbative result, while the grayscale lines display the computed heat kernel for $m = 0, 2, 4$ (lighter corresponds to bigger m) and $\epsilon = 0.7$. We observe distinct behaviours for

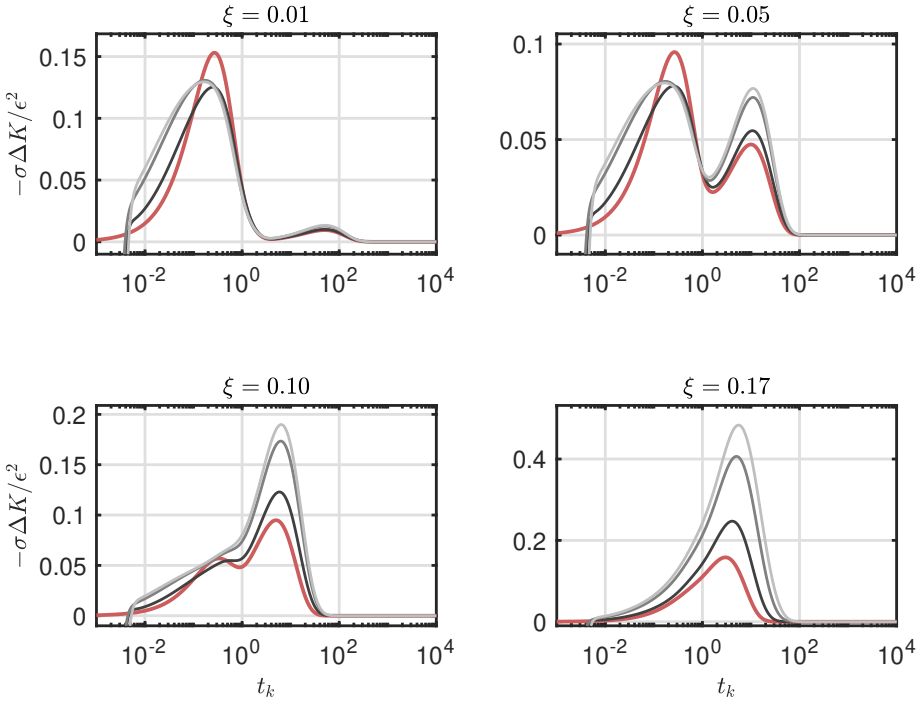


Figure 5.9: Subtracted heat kernel for a non-minimally coupled scalar for $\xi = 0.01, 0.05, 0.1, 1/6$ and $\epsilon = 0.7$, normalised by the square of the deformation parameter. Each panel shows the perturbative result (red line) and the numerical calculations (grayscale lines) for $l = 4$ and $m = 0, 2, 4$ (lighter corresponds to bigger m). Results obtained for a computational grid with $N = 100$.

the two peaks. The leftmost crest, i.e. the one present also in the $\xi = 0$ case, behaves equivalently to what seen in Figure 5.4, with the axisymmetric deformation sottending a much smaller area with respect to the $m = 2$ and $m = 4$ modes. Furthermore, there is little difference between the two non-axisymmetric perturbations. Conversely, the curvature peak discriminates more significantly between deformations of different m. We observe that bigger m corresponds to a significantly higher peak, even though the most pronounced difference is still between axisymmetric and non-axisymmetric deformations.

It is worth highlight again that we see no evidence of the subtracted heat kernel going negative for any of the considered deformation even for the non-minimally coupled case. This reinforces still the idea that the sphere does indeed maximise (at least locally) the Casimir energy of a free scalar field.

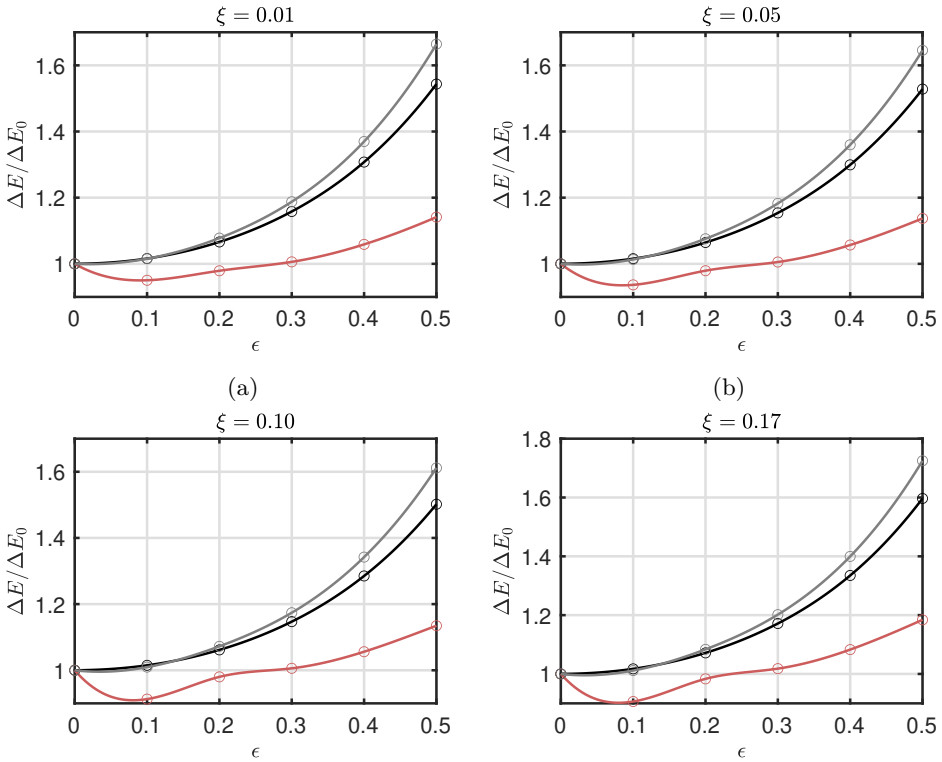


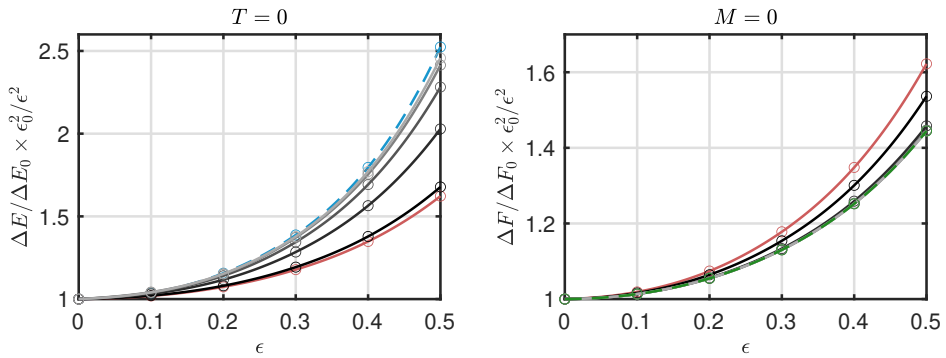
Figure 5.10: Subtracted free energy for a massless non-minimally coupled scalar at zero temerature for $\xi = 0.01, 0.05, 0.1, 1/6$ as a function of ϵ. The subtracted free energy is first normalised by the ϵ^2 at which it is calculated, and then by the perturbative result ΔE_0. Each panel shows the numerical results (grayscale markers) for $m = 2, 4$ (lighter corresponds to bigger m). The red markers display the $m = 0$ axisymmetric results. Note that the solid lines are splines interpolators, drawn to guide the eyes.

5.3.2 Subtracted Free Energy

We finally turn our attention back to the free energy, this time for the non-minimally coupled field. We start by considering the massless field at zero temperature, and then compute the finite mass and temperature corrections.

Figure 5.10 shows the computed Casimir energy for $\xi = 0.01, 0.05, 0.1, 1/6$. We display the results for axisymmetric (red markers) and non axisymmetric (grayscale markers, lighter corresponds to bigger m) deformations as a function of ϵ. As expected from Figure 5.8 and Figure 5.9, we observe the non-axisymmetric perturbations to produce a more negative subtracted energy with respect to the axisymmetric one. It is interesting to note that, even though they have significantly different heat kernel profiles, distinct ξ produce a staggeringly similar behaviour for ΔF once appropriately normalised.

Figure 5.11 shows the subracted free energy when non-zero masses and temperatures for are considered for the representative case of a non-minimally coupled scalar with $\xi = 1/6$ and deformation given by $l = 3$ and $m = 1$. As in the minimally coupled case, for increasing M the numerical results approach the analytical long wavelength limit (blue dashed line), while for big T they reproduce the hot limit (green dashed line). We observe no significant difference between the behaviour of minimally and non-minimally coupled scalar.



(a) (b)

Figure 5.11: Subtracted free energy as a function of ϵ for the representative example of $l = 3$, $m = 1$ and $\xi = 1/6$ with non-zero masses and temperatures. The grayscale markers correspond to fields with (a) $M = 0.5, 1, 2, 4, 8, 16$ and $T = 0$ and (b) $M = 0$ and $T = 0.5, 1, 2, 4, 8, 16$, with smaller value of the changing parameter being represented by a darker circle. The red markers indicate the $M = 0$, $T = 0$ results. The dashed blue and green lines give the long-wavelength and small temperature limits respectively. Solid lines are splines drawn to guide the eye. Note that most of the high T curves in (b) are overlaid in the hot limit.

Chapter 6

Conclusions

In this dissertation, the Helmohltz free energy of a free scalar quantum field on a curved (2+1)-dimensional spacetime was studied. In particular, we placed our K-G QFT on a $\mathbb{R} \times \Sigma$ spacetime, with Σ being a spatial manifold topologically equivalent to S^2, and studied how such free energy changed as Σ was deformed away from the round sphere. We deployed numerical techniques involving spectral methods and the heat kernel representation to compute such energy. We also reviewed recent results that seem to point in the direction that the round sphere maximises the free energy of QFTs. Our study generalises some of these findings, extending the analysis to non-axisymmetric deformations.

6.1 Summary of the Main Results and Future Studies

Before jumping into the physically relevant results, we first benchmarked our computational technique, studying the Laplacian on a 1D interval and on a 2-sphere which have well known analytical solutions. We further investigated the effects that the geometry of the computational grid has on the estimation of the spectra of such operators, finding that the optimal configuration is a rectangular lattice with more density of nodes in the azimuthal direction, which reduces the impact of aliasing.

We then presented our novel results, starting with a minimally coupled scalar under non-axisymmetric deformations defined in terms of spherical harmonics. While similar studies have been recently conducted on the subject, they limited their investigations to axisymmetric perturbations. Our research extends their results to

a more general family of deformations. We first demonstrated the positivity of the subtracted heat kernel for distortions with a wide range of intensities. We further described the effects that different values of the m parameter of the deformation function have on the profile of the kernel. We observe that the difference between modes with distinct m are more evident in the small kernel time regimes, showing that the IR physics is mostly insensitive to it. We then computed the subtracted free energy of the field, with our results confirming that, for all the perturbations considered, the free scalar field prefers to live on a crumpled geometry rather than on the round sphere. We further found that non-axisymmetric deformations produce a significantly more negative subtracted free energy than their axisymmetric counterparts. We concluded the discussion by repeating the same analysis on a field with non-minimal coupling to the Ricci scalar. Again, we found negativity of the subtracted free energy in all cases. We also described the effect that non-minimal coupling has on the profile of the subtracted kernel, showing up as an extra crest in the big kernel time regime that gets amplified as the non-minimal coupling increases, until it completely dominates the other contributions. We also suggestively find this transition to happen silghtly before $\xi = 1/6$, namely the value for which a (2+1)-dimensional massless scalar field becomes conformally invariant. Interestingly, we also observe that, while the profile of the heat kernel varies significantly between $\xi = 0$ and $\xi = 1/6$, the behaviour of the subtracted free energy as a function of ϵ, once normalised by the perturbative results, stays effectively invariant throughout the whole range of the values of ξ considered.

Even though we do not provide a formal proof, the novel results we obtained, along with the recent findings that we reviewed in Chapter 3, give good reasons to believe that the free scalar field (and the statement seems to be generalizable to any quantum field theory) prefers deformed geometries to the round sphere: an extremely counterintuitive conclusions. S^2 seems to (at least locally) maximise the free energy of QFTs. Thanks to recent technological developments in the field of material engineering, it is coinceivable that in the near future it will be possible to devise an experiment in laboratory in which such properties of QFTs can be tested. As discussed, exploiting the flexibility of optical lattices, one can already imagine designing a 2D material with the right properties such that the quantum zero-point energy of the effective fields can have measurable effects. Theorywise, it would be interesting to assess whether the similarity in the behaviour of the subtracted free energy between a free scalar and free fermion field observed in previous works [10] still holds for non-axisymmetric deformations of Σ. Our method can be straightforwardly adapted to study fermionic fields by appropriately changing the Laplacian, so it is indeed a natural extension of this study.

Bibliography

[1] M. S. Morris and K. S. Thorne, "Wormholes in spacetime and their use for interstellar travel: A tool for teaching general relativity," *American Journal of Physics*, vol. 56, no. 5, pp. 395–412, May 1988. DOI: 10.1119/1.15620.

[2] J. Maldacena, A. Milekhin, and F. Popov, "Traversable wormholes in four dimensions," 2020. arXiv: 1807.04726 [hep-th].

[3] R. Garattini, "Casimir wormholes," *The European Physical Journal C*, vol. 79, no. 11, Nov. 2019, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7468-y. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-019-7468-y.

[4] R. J. Adler, B. Casey, and O. C. Jacob, "Vacuum catastrophe: An elementary exposition of the cosmological constant problem," *American Journal of Physics*, vol. 63, no. 7, pp. 620–626, 1995. DOI: 10.1119/1.17850. eprint: https://doi.org/10.1119/1.17850. [Online]. Available: https://doi.org/10.1119/1.17850.

[5] S. Longhi and G. D. Valle, "Realization of interacting quantum field theories in driven tight-binding lattices," *New Journal of Physics*, vol. 14, no. 5, p. 053 026, May 2012. DOI: 10.1088/1367-2630/14/5/053026. [Online]. Available: https://doi.org/10.1088/1367-2630/14/5/053026.

[6] F. Bemani, R. Roknizadeh, and M. Naderi, "Quantum simulation of discrete curved spacetime by the Bose–Hubbard model: From analog acoustic black hole to quantum phase transition," *Annals of Physics*, vol. 388, pp. 186–196, Jan. 2018, ISSN: 0003-4916. DOI: 10.1016/j.aop.2017.11.010. [Online]. Available: http://dx.doi.org/10.1016/j.aop.2017.11.010.

[7] N. Szpak, "Curved spacetimes in the lab," 2014. arXiv: 1410.1567 [quant-ph].

[8] A. Hickling and T. Wiseman, "Vacuum energy is non-positive for (2 + 1)-dimensional holographic CFTs," *Classical and Quantum Gravity*, vol. 33, no. 4, p. 045 009, Jan. 2016, ISSN: 1361-6382. DOI: 10.1088/0264-9381/33/4/045009. [Online]. Available: http://dx.doi.org/10.1088/0264-9381/33/4/045009.

[9] S. Fischetti, L. Wallis, and T. Wiseman, "Does the round sphere maximize the free energy of (2+1)-dimensional QFTs?" *Journal of High Energy Physics*, vol. 2020, no. 10, p. 78, 2020. DOI: 10.1007/JHEP10(2020)078. [Online]. Available: https://doi.org/10.1007/JHEP10(2020)078.

[10] K. Cheamsawat, S. Fischetti, L. Wallis, and T. Wiseman, "A surprising similarity between holographic CFTs and a free fermion in (2 + 1) dimensions," *Journal of High Energy Physics*, vol. 2021, no. 5, p. 246, 2021. DOI: 10.1007/JHEP05(2021)246. [Online]. Available: https://doi.org/10.1007/JHEP05(2021)246.

[11] S. Fischetti and T. Wiseman, "On universality of holographic results for (2 + 1)-dimensional CFTs on curved spacetimes," *Journal of High Energy Physics*, vol. 2017, no. 12, Dec. 2017, ISSN: 1029-8479. DOI: 10.1007/jhep12(2017)133. [Online]. Available: http://dx.doi.org/10.1007/JHEP12(2017)133.

[12] S. K. Lamoreaux, "Casimir forces: Still surprising after 60 years," *Physics Today*, vol. 60, no. 2, pp. 40–45, 2007. DOI: 10.1063/1.2711635. eprint: https://doi.org/10.1063/1.2711635. [Online]. Available: https://doi.org/10.1063/1.2711635.

[13] S. K. Lamoreaux, "Demonstration of the Casimir force in the 0.6 to 6μm range," *Phys. Rev. Lett.*, vol. 78, pp. 5–8, 1 Jan. 1997. DOI: 10.1103/PhysRevLett.78.5. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.78.5.

[14] H. B. G. Casimir, "On the attraction between two perfectly conducting plates," *Indag. Math.*, vol. 10, pp. 261–263, 1948.

[15] M. Visser, "Regularization versus renormalization: Why are Casimir energy differences so often finite?" *Particles*, vol. 2, no. 1, pp. 14–31, 2019, ISSN: 2571-712X. DOI: 10.3390/particles2010002. [Online]. Available: https://www.mdpi.com/2571-712X/2/1/2.

[16] G. Plunien, B. Müller, and W. Greiner, "The Casimir effect," *Physics Reports*, vol. 134, no. 2, pp. 87–193, 1986, ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(86)90020-7. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0370157386900207.

[17] F. Perrotta, C. Baccigalupi, and S. Matarrese, "Extended quintessence," *Physical Review D*, vol. 61, no. 2, Dec. 1999, ISSN: 1089-4918. DOI: 10.1103/physrevd.61.023507. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.61.023507.

[18] V. Faraoni, "Nonminimal coupling of the scalar field and inflation," *Phys. Rev. D*, vol. 53, pp. 6813–6821, 12 Jun. 1996. DOI: 10.1103/PhysRevD.53.6813. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.53.6813.

[19] S. W. Hawking, "Particle creation by black holes," *Communications in Mathematical Physics*, vol. 43, no. 3, pp. 199–220, 1975. DOI: cmp/1103899181. [Online]. Available: https://doi.org/.

[20] B. S. DeWitt, "Quantum field theory in curved spacetime,", vol. 19, no. 6, pp. 295–357, Aug. 1975. DOI: 10.1016/0370-1573(75)90051-4.

[21] R. M. Wald, *General Relativity*. University of Chicago Press, 1984, ISBN: 978-0-226-87033-5. DOI: 10.7208/chicago/9780226870373.001.0001.

[22] S. Moroianu, "Adiabatic limits of eta and zeta functions of elliptic operators," *Mathematische Zeitschrift*, vol. 246, no. 3, pp. 441–471, Mar. 2004, ISSN: 1432-1823. DOI: 10.1007/s00209-003-0578-z. [Online]. Available: http://dx.doi.org/10.1007/s00209-003-0578-z.

[23] K. Bormann and F. Antonsen, "The Casimir effect of curved spacetime (formal developments)," 1996. arXiv: 9608142 [hep-th].

[24] D. Grieser, "Notes on heat kernel asymptotics," 2004.

[25] J. Earman, "The penrose-hawking singularity theorems: History and implications," in *The Expanding Worlds of General Relativity*, H. Goenner, J. Renn, J. Ritter, and T. Sauer, Eds. 1999, p. 235.

[26] N. Graham and K. D. Olum, "Achronal averaged null energy condition," *Physical Review D*, vol. 76, no. 6, Sep. 2007, ISSN: 1550-2368. DOI: 10.1103/physrevd.76.064001. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.76.064001.

[27] S. Weinberg, "The cosmological constant problem," *Rev. Mod. Phys.*, vol. 61, pp. 1–23, 1 Jan. 1989. DOI: 10.1103/RevModPhys.61.1. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.61.1.

[28] K. Eppley and E. Hannah, "The necessity of quantizing the gravitational field," *Foundations of Physics*, vol. 7, no. 1, pp. 51–68, 1977. DOI: 10.1007/BF00715241. [Online]. Available: https://doi.org/10.1007/BF00715241.

[29] A. Kent, "Simple refutation of the Eppley–Hannah argument," *Classical and Quantum Gravity*, vol. 35, no. 24, p. 245 008, Nov. 2018, ISSN: 1361-6382. DOI: 10.1088/1361-6382/aaea20. [Online]. Available: http://dx.doi.org/10.1088/1361-6382/aaea20.

[30] D. N. Page and C. D. Geilker, "Indirect evidence for quantum gravity," *Phys. Rev. Lett.*, vol. 47, pp. 979–982, 14 Oct. 1981. DOI: 10.1103/PhysRevLett.47.979. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.47.979.

[31] J. Oppenheim and Z. Weller-Davies, "The constraints of post-quantum classical gravity," 2020. arXiv: 2011.15112 [hep-th].

[32] B. Krishnan, "Quasi-local black hole horizons," *Springer Handbook of Spacetime*, Mar. 2013. DOI: 10.1007/978-3-642-41992-8_25.

[33] J. A. Wheeler, "Geons," *Phys. Rev.*, vol. 97, pp. 511–536, 2 Jan. 1955. DOI: 10.1103/PhysRev.97.511. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.97.511.

[34] S. K. Tiwari, S. Sahoo, N. Wang, and A. Huczko, "Graphene research and their outputs: Status and prospect," *Journal of Science: Advanced Materials and Devices*, vol. 5, no. 1, pp. 10–29, 2020, ISSN: 2468-2179. DOI: https://doi.org/10.1016/j.jsamd.2020.01.006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S246821792030006X.

[35] G. Montambaux, "Artificial graphenes: Dirac matter beyond condensed matter," *Comptes Rendus Physique*, vol. 19, Nov. 2018. DOI: 10.1016/j.crhy.2018.10.010.

[36] A. J. Leggett, "Selected topics in condensed matter physics, chapter 5," 2010.

[37] C. Hwang, D. A. Siegel, S.-K. Mo, W. Regan, A. Ismach, Y. Zhang, A. Zettl, and A. Lanzara, "Fermi velocity engineering in graphene by substrate modification," *Scientific Reports*, vol. 2, no. 1, Aug. 2012, ISSN: 2045-2322. DOI: 10.1038/srep00590. [Online]. Available: http://dx.doi.org/10.1038/srep00590.

[38] S. Y. Zhou, G. .-H. Gweon, J. Graf, A. V. Fedorov, C. D. Spataru, R. D. Diehl, Y. Kopelevich, D. .-H. Lee, S. G. Louie, and A. Lanzara, "First direct observation of Dirac fermions in graphite," *Nature Physics*, vol. 2, no. 9, pp. 595–599, 2006. DOI: 10.1038/nphys393. [Online]. Available: https://doi.org/10.1038/nphys393.

[39] S. Sadeddine, H. Enriquez, A. Bendounan, P. Kumar Das, I. Vobornik, A. Kara, A. J. Mayne, F. Sirotti, G. Dujardin, and H. Oughaddou, "Compelling experimental evidence of a Dirac cone in the electronic structure of a 2d silicon layer," *Scientific Reports*, vol. 7, no. 1, p. 44 400, 2017. DOI: 10.1038/srep44400. [Online]. Available: https://doi.org/10.1038/srep44400.

[40] D. Nelson and L. Peliti, "Fluctuations in membranes with crystalline and hexatic order," *Journal de Physique*, vol. 48, no. 7, pp. 1085–1092, 1987. DOI: 10.1051/jphys:019870048070108500. [Online]. Available: https://hal.archives-ouvertes.fr/jpa-00210530.

[41] X. Shi, B. Peng, N. M. Pugno, and H. Gao, "Stretch-induced softening of bending rigidity in graphene," *Applied Physics Letters*, vol. 100, no. 19, p. 191 913, 2012. DOI: 10.1063/1.4716024. eprint: https://doi.org/10.1063/1.4716024. [Online]. Available: https://doi.org/10.1063/1.4716024.

[42] J. T. Wheeler, "Weyl geometry," *General Relativity and Gravitation*, vol. 50, no. 7, p. 80, 2018. DOI: 10.1007/s10714-018-2401-5. [Online]. Available: https://doi.org/10.1007/s10714-018-2401-5.

[43] J. McGreevy, "Holographic duality with a view toward many-body physics," *Advances in High Energy Physics*, vol. 2010, pp. 1–54, 2010, ISSN: 1687-7365. DOI: 10.1155/2010/723105. [Online]. Available: http://dx.doi.org/10.1155/2010/723105.

[44] J. F. Epperson, "On the Runge example," *The American Mathematical Monthly*, vol. 94, no. 4, pp. 329–341, 1987. DOI: 10.1080/00029890.1987.12000642.

[45] C. Runge, "Über empirische funktionen und die interpolation zwischen äquidistanten ordinaten," *The Scandinavian Journal of Economics*, vol. 46, pp. 224–243, 1901.

[46] A. Pierce, "Lecture notes on variational and approximate methods in applied mathematics: Lecture 3," 2017.

[47] L. N. Trefethen, *Spectral Methods in MATLAB*. SIAM, 2000, ISBN: 978-0-89871-465-4.

[48] J. Hoepffner, "Implementation of boundary conditions," 2007.