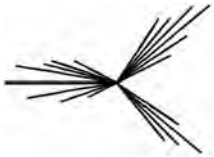

IFT

Instituto de Física Teórica
Universidade Estadual Paulista

MASTER'S THESIS IFT-D.010/11

Can entanglement explain black hole entropy?

Katja Ried

Supervisor

George Emanuel Avraam Matsas

September 2011

To boldly go...

Acknowledgements

This work was supported by scholarships from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, process number 136114/2009-1) and Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp, process number 2009/10774-8).

On a personal note, I am grateful to the people who have supported me in the many endeavours that contributed, directly or indirectly, to this work. I hope the interactions were as stimulating for you as they were for me.

Abstract

When seeking inspiration for a future theory of quantum gravity, studying black holes is a promising ansatz, since they present us with several puzzles at the intersection of quantum theory and gravity. Among these is their entropy: although there are compelling arguments for its existence, its origin and statistical meaning remain a mystery. Previous work showed that at least some aspects of this phenomenon can be accounted for by the entanglement of quantum fields across the horizon: If a field is globally in a pure state, yet part of it is hidden behind the event horizon, then the reduced state of the remainder possesses non-zero entropy. This is the possibility we explore in the present work, in the simplest of settings: a ground-state scalar field, defined in three-dimensional, flat or uniformly curved space. We reduce the problem to that of coupled harmonic oscillators by discretizing space, and derive an expression for the entropy, which is evaluated numerically. The results show that the entropy scales with the boundary area of the inaccessible region, a key feature of black hole entropy known as the area law. Furthermore, we conclude that the dominant contribution to the entropy is due to short-range interactions, and discuss some physical implications of this insight for the puzzle of black hole entropy.

Keywords: black hole, entanglement entropy, area law, quantum field theory in curved spacetime

Contents

Conventions and notation

Throughout this work, unless stated otherwise, we use natural units, such that the speed of light in vacuum, and the Newton, Boltzmann and Planck constants are all equal to unity:

$$c = G = k = \hbar = 1.$$

The unit length in this system is the Planck length, $L_{Pl} \equiv \sqrt{\hbar G/c^3} \approx 10^{-35} m$, and the unit mass is $M_{Pl} =\equiv \sqrt{\hbar c/G} \approx 10^{-8} kg$, the Planck mass.

Throughout this work, the indices l and m are used mostly for partial waves. When the context makes it clear that this is the case, their range, $l = 0, 1, ...$ and $m = -l, -l+1, ..., l$, is suppressed for the sake of notational simplicity.

In the sections involving general relativity, we use the metric signature $(-, +, +, +)$.

The components of a matrix (or vector) A are denoted by A_{ij} (or A_i), and its transpose is A^T. We use A^{-1} for the inverse of a square matrix, if it exists. If A can be diagonalized by a transformation U, let $A_D \equiv UAU^T$ denote the resulting diagonal matrix. We use $\mathbb{I}$ for the identity matrix.

The asterisk (*) denotes complex conjugation, and $\imath \equiv \sqrt{-1}$.

1 Introduction

The two pillars on which our current understanding of physics is built, quantum theory and general relativity, are at the same time indispensable and incompatible. The conceptual tensions between them run so deep that their unification, leading to some larger theory of quantum gravity, has been declared the "Holy Grail" of modern physics. The approaches to the construction of this new theory can be roughly divided into two classes: on the one hand, "top-down" approaches attempt to deduce the entire structure of the new theory from formal principles, before making predictions about specific phenomena. However, we find it more promising to begin by studying the phenomena governed by such a theory, seeking conceptual insights and fundamental physical principles to guide the search for a more abstract description, in the spirit of a "bottom-up" approach. One might argue that this is the way in which quantum mechanics was first discovered, starting with the puzzle of black body radiation and experiments about interfering matter waves. This, not formal considerations about Hilbert spaces, led to the change of paradigm from classical to quantum. It stands to reason that the same approach will inspire the next, probably even more revolutionary leap in our view of the world.

The challenge lies in finding such phenomena, that is, systems in which quantum mechanics and gravity play equally important roles. This is the reason for studying black holes: They readily provide us with several puzzles at the interface of quantum theory, in particular as it pertains to information, and gravity. One major issue is known as the *information loss paradox*: Hawking [1] showed that black holes evaporate, losing mass while emitting thermal, featureless radiation. Thus, any information they absorbed throughout their existence must either be returned to the outer universe, by some as yet unknown mechanism, or be destroyed. The latter possibility is particularly troubling in the context of quantum mechanics, which does not allow for information loss, since that would imply non-unitary evolution. However, Wald [2], for example, argues that the information passes into a region of spacetime to which the asymptotic observers have no access, thereby avoiding the paradox. However, we defer this discussion to the literature on the topic[1], since this phenomenon is not the object of the present work.

We focus instead on another unanswered question, which arises from the fact that black holes act as thermodynamic objects. This idea, too, builds on the discovery that black holes emit thermal radiation, hence possessing a temperature. It was also suggested by Bekenstein's argument that black holes could violate the second law of

[1]Besides the references cited in the text, Preskill's review [3] and Hawking's seminal paper on the problem [4] may be of interest.

thermodynamics unless one assigns an entropy to them as well [5].[2] The microscopic origin and statistical meaning of this entropy, however, remain unclear.

The question has been addressed in a variety of contexts, and several approaches suggest that entanglement may be in some form responsible, although the details are still being debated. For this reason, we base this work on the least complex mechanism, as it was originally proposed by Bombelli, Koul, Lee and Sorkin [6]: they attribute the entropy to quantum fields living in the black hole spacetime, which are entangled across the event horizon and hence possess entanglement entropy. We stress that this explanation requires no additional hypotheses beyond comparatively well-understood, noncontroversial ideas from quantum theory and general relativity, yet it can account for the appearance of entropy. Indeed, a similar approach, of equal conceptual simplicity, has already successfully explained another thermodynamic property of black holes, in the context of the Hawking effect (see section 4.2).

The model can be simplified even further by transposing it to flat space instead of a black hole background, introducing an arbitrary ad hoc boundary to replace the event horizon. Srednicki [7] showed that this setup still generates entropy in the case of spherical regions, and reproduces the key features one expects from the entropy of a black hole (as detailed in section 4.3). This raises the question of how well entanglement alone, without explicit reference to gravitational aspects, can account for black hole entropy. It is this question that we propose to address here, by extending Srednicki's work to regions of different geometries in flat space, as well as space with uniform curvature, and massive fields. We hope that this will provide valuable insights into the mechanism that endows black holes with entropy.

To this end, we begin by reviewing concepts and results from quantum theory, general relativity and quantum field theory (sections 2 through 4) that are relevant to our approach to black hole entropy. Section 5 builds on this basis in deriving analytical expressions for the entanglement entropy of coupled harmonic oscillators. These are then generalized to quantum fields, and the details of evaluating them numerically for different setups are discussed. The results regarding the behaviour of the entanglement entropy (for instance how it depends on factors such as the geometry of the traced-out region and the mass of the field, and in particular the area law) are presented in section 6. Finally, section 7 is devoted to the conclusions we draw from these results. Lengthier derivations of mathematical identities that are used throughout the work are gathered in appendix A, while appendix B contains the Mathematica® code used in the numerical calculations.

[2]The subject of black hole thermodynamics is introduced in more detail in section 4.3.

2 Quantum theory

This section covers specific topics from quantum theory that are relevant to the present work. Although part of the content may be familiar to the reader from undergraduate courses, it has been included not only for the sake of completeness, but to clarify the conceptual foundations on which the later chapters are built.

2.1 Density matrices, pure and mixed states

In this work, by a *pure (quantum) state* we mean a state that can be represented by a single state vector, which is a linear combination of elements of a basis of Hilbert space. This basis can consist of eigenstates of an observable, for instance $\{|+_z\rangle, |-_z\rangle\}$, the eigenstates of the spin component S_z, to describe the spin degrees of freedom of a single (localized) electron. Examples of pure states are $|+_z\rangle$ and $\frac{1}{\sqrt{2}}(|+_z\rangle + |-_z\rangle) = |+_x\rangle$, which are eigenstates of S_z and S_x, respectively. In the latter case, while a measurement of S_x will certainly yield $+\hbar/2$ (spin "pointing" in the $+\hat{x}$ direction) , a measurement of S_z has a 50% probability of finding the spin up ($+\hat{z}$) and 50% down ($-\hat{z}$).

The definition of a pure state is best illustrated by contrasting it with that of a *mixed state*: the latter is described not by a single linear combination of basis states, but several such pure states, which can be found with certain probabilities. Consider again the electron, now in a state that is 50% $|+_z\rangle$ and 50% $|-_z\rangle$. A measurement of S_z can, of course, detect the spin pointing either up or down with equal probabilities. However, measuring S_x can also yield either $+\hat{x}$ or $-\hat{x}$, setting this state apart from the pure $|+_x\rangle$ considered above.

In order to track down the difference between pure and mixed states, suppose we wish to compute the probability of finding the spin pointing in the $+\hat{x}$ direction. Since the corresponding projection operator is

$$\text{Proj} = |+_x\rangle\langle+_x| = \frac{1}{\sqrt{2}}(|+_z\rangle + |-_z\rangle)\,\frac{1}{\sqrt{2}}(\langle+_z| + \langle-_z|), \tag{2.1}$$

the probability of finding this result in a measurement on the pure state $|\varphi_p\rangle = \frac{1}{\sqrt{2}}(|+_z\rangle + |-_z\rangle)$ is

$$\begin{aligned} P_p &= |\langle+_x|\varphi_p\rangle|^2 = \langle\varphi_p|\text{Proj}|\varphi_p\rangle \\ &= \frac{1}{\sqrt{2}}(\langle+_z| + \langle-_z|)\left[\frac{1}{\sqrt{2}}(|+_z\rangle + |-_z\rangle)\,\frac{1}{\sqrt{2}}(\langle+_z| + \langle-_z|)\right]\frac{1}{\sqrt{2}}(|+_z\rangle + |-_z\rangle) \\ &= \frac{1}{2}(1+1)\,\frac{1}{2}(1+1) = 1. \end{aligned} \tag{2.2}$$

On the other hand, for the mixed state with 50% $|+_z\rangle$ and 50% $|-_z\rangle$, this probability is

$$P_m = \frac{1}{2}\langle +_z|\mathrm{Proj}|+_z\rangle + \frac{1}{2}\langle -_z|\mathrm{Proj}|-_z\rangle = \frac{1}{2}\left(\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\right) + \frac{1}{2}\left(\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\right) = \frac{1}{2}. \quad (2.3)$$

Both expressions can be written as single matrix elements: for the mixed state,

$$P_m = \langle +_x|\left[\frac{1}{2}|+_z\rangle\langle +_z| + \frac{1}{2}|-_z\rangle\langle -_z|\right]|+_x\rangle = \frac{1}{2}, \quad (2.4)$$

and for the pure state

$$P_p = \langle +_x|\left[\frac{1}{2}|+_z\rangle\langle +_z| + \frac{1}{2}|+_z\rangle\langle -_z| + \frac{1}{2}|-_z\rangle\langle +_z| + \frac{1}{2}|-_z\rangle\langle -_z|\right]|+_x\rangle = 1. \quad (2.5)$$

The difference between the pure and mixed states becomes evident; in this case, it lies in the cross terms, of the form $|+_z\rangle\langle -_z|$.

The example shows that vectors are not sufficient to describe mixed states. However, it already suggests the necessary generalization: the *density operator* (or density matrix) ρ. For a normalized pure state $|\varphi\rangle$, it is simply $\rho \equiv |\varphi\rangle\langle\varphi|$, or, in the position representation

$$\rho\left(x, x'\right) = \varphi^*\left(x\right)\varphi\left(x'\right), \quad (2.6)$$

where $\varphi\left(x\right) \equiv \langle x|\varphi\rangle$ is the wave function. When several such states $|\varphi_k\rangle$ are mixed together, ρ is the sum of the corresponding density operators, multiplied by the respective probabilities q_k:

$$\rho \equiv \sum_k q_k|\varphi_k\rangle\langle\varphi_k|. \quad (2.7)$$

The probability of finding a given eigenvalue x_n in a measurement of an observable X is then the matrix element of the density operator between the corresponding eigenvectors $|x_n\rangle$:

$$P_n = \langle x_n|\rho|x_n\rangle = \sum_k q_k\left|\langle x_n|\varphi_k\rangle\right|^2 \geq 0. \quad (2.8)$$

The expectation value of X becomes

$$\langle X\rangle = \sum_n x_n P_n = \sum_{m,n}\langle x_m|\left[\rho|x_n\rangle x_n\langle x_n|\right]|x_m\rangle = \mathrm{Tr}\left(\rho X\right), \quad (2.9)$$

and taking X to be the identity operator, we find the normalization condition for ρ:

$$\mathrm{Tr}\left(\rho\right) = 1. \quad (2.10)$$

We furthermore note that ρ in equation (2.7) is Hermitian, so its eigenvectors $|\psi_k\rangle$ span the Hilbert space. In this basis, ρ is diagonal,

$$\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|; \qquad \langle\psi_k|\psi_l\rangle = \delta_{kl}, \tag{2.11}$$

and its eigenvalues are

$$p_k = \langle\psi_k|\rho|\psi_k\rangle = P_k. \tag{2.12}$$

These are the probabilities of finding the basis state $|\psi_k\rangle$ in the mixture.

2.2 Bipartite systems, partial trace and entanglement

We will see that mixed states and consequently entropy appear naturally in bipartite (and, by extension, multipartite) systems. For this reason, we begin by introducing the formalism that will be used to treat them. Let AB be a system that can be divided into A and B, let $\mathcal{H}_A$ be the Hilbert space and $\{|a_i\rangle\}$ an arbitrary orthonormal basis of the subsystem A (and similarly for B). Then the state space of the composite system is given by the Kronecker product

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B, \tag{2.13}$$

and a basis for it is

$$\{|a_i b_j\rangle\} \equiv \{|a_i\rangle \otimes |b_j\rangle\}. \tag{2.14}$$

Hence a generic state can be expanded as

$$|\varphi\rangle = \sum_{i,j} C_{ij} |a_i b_j\rangle. \tag{2.15}$$

Likewise, operators on $\mathcal{H}_{AB}$ are given by products of operators on $\mathcal{H}_A$ and $\mathcal{H}_B$,

$$O_{AB} = O_A \otimes O_B, \tag{2.16}$$

or combinations thereof. The action of such operators is distributive,

$$O_{AB}|\varphi\rangle = \sum_{i,j} C_{ij} \left(O_A|a_i\rangle\right) \otimes \left(O_B|b_j\rangle\right), \tag{2.17}$$

as are scalar products:

$$\left[\sum_{i,j} C_{ij}\langle a_i b_j|\right]\left[\sum_{k,l} C'_{kl}|a_k b_l\rangle\right] = \sum_{i,j,k,l} C_{ij}C'_{kl}\langle a_i|a_k\rangle\langle b_j|b_l\rangle = \sum_{i,j} C_{ij}C'_{ij}. \tag{2.18}$$

In particular, suppose the two subsystems do not interact: each is governed by a Hamiltonian H_A (H_B) that acts only on the space $\mathcal{H}_A$ ($\mathcal{H}_B$). Then the total Hamiltonian of the system is

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathbb{I}_B + \mathbb{I}_A \otimes \mathcal{H}_B, \tag{2.19}$$

and the ground state is

$$|0_{AB}\rangle = |0_A\rangle \otimes |0_B\rangle, \tag{2.20}$$

the product of the ground states of A and B.

We can now describe how mixed states "appear" in a bipartite system: when the whole, isolated system is in a pure state, but one ignores a part of this system (A), then the remaining subsystem (B) is not necessarily left in a single, well-defined, pure state, but sometimes in a mixed one.

As an example, consider two electrons, localized in different positions so that they are distinguishable. If this system is in the (pure) state

$$|\varphi_p\rangle = \frac{1}{\sqrt{2}}\left(|+_{zA}+_{zB}\rangle + |+_{zA}-_{zB}\rangle\right) = |+_{zA}\rangle \otimes \frac{1}{\sqrt{2}}\left(|+_{zB}\rangle + |-_{zB}\rangle\right), \tag{2.21}$$

then any measurement bearing on B will give the same results as if one had taken a single, isolated particle in the (also pure) state $\frac{1}{\sqrt{2}}\left(|+_z\rangle + |-_z\rangle\right)$. $|\varphi_p\rangle$ is an example of a *product state*: one that can be written as a product of pure states of the subsystems. In terms of the expansions coefficients in equation (2.15), the necessary and sufficient condition is that they can be factored as

$$C_{ij} = c_{Ai}c_{Bj} \qquad \forall i,j, \tag{2.22}$$

or, equivalently,

$$|\varphi_p\rangle = \sum_{i,j} c_{Ai}|a_i\rangle \otimes c_{Bj}|b_j\rangle = |\varphi_A\rangle \otimes |\varphi_B\rangle. \tag{2.23}$$

Another example of product states are the eigenstates of products of operators. Consider, for instance, the position operators X_A and X_B, and their eigenstates, $\{|x_{A,B}\rangle\}$. The product states $\{|x_A x_B\rangle\}$ are eigenstates of $X_{AB} \equiv X_A \otimes X_B$, as can be verified by substituting in equation (2.17). This implies that the wave function of a product state in the position representation is

$$\varphi_p\left(x_A, x_B\right) \equiv \langle x_A x_B|\varphi_p\rangle = \langle x_A|\varphi_A\rangle\langle x_B|\varphi_B\rangle \equiv \varphi_A\left(x_A\right)\varphi_B\left(x_B\right), \tag{2.24}$$

the product of the wave functions of the factors.

If, on the other hand, the state of the two-electron system is

$$|\varphi_e\rangle \equiv \frac{1}{\sqrt{2}}\left(|+_{zA}+_{zB}\rangle - |-_{zA}-_{zB}\rangle\right), \tag{2.25}$$

then no such factorization into states of the subsystems A and B is possible. However, there still is a single-particle state that describes B, in the sense that it correctly predicts the probabilities of all possible outcomes in measurements bearing only on that subsystem. This is the *reduced state* of the subsystem B. To find this state, it is not sufficient to simply ignore the state of A in $|\varphi_e\rangle$ and add the state vectors of B:

$$\begin{aligned} |\varphi_e\rangle &= \tfrac{1}{\sqrt{2}}\left(|+_{zA}+_{zB}\rangle - |-_{zA}-_{zB}\rangle\right) \\ \nRightarrow \quad |\varphi_B\rangle &= \tfrac{1}{\sqrt{2}}\left(|+_{zB}\rangle - |-_{zB}\rangle\right). \end{aligned} \tag{2.26}$$

The problem is that global states (of the whole system) in which A is in different states (such as $|+_{zA}\pm_{zB}\rangle$ and $|-_{zA}\pm_{zB}\rangle$) are still orthogonal, even if one chooses to ignore this subsystem. Therefore, in order to calculate the probability of finding B in a given state, one must add the probability amplitudes corresponding to such states, not the vectors. Thus, $|\varphi_e\rangle$ leads to a reduced state that is 50% $|+_{zB}\rangle$ and 50% $|-_{zB}\rangle$, that is, a mixed state.

The fact that we may be dealing with mixed states suggests that we use density operators in the general procedure for finding the reduced state. Given ρ_{AB}, one takes the matrix elements between the ket and the bra of a certain state of A – that is, the probability amplitude – and sums (or integrates, for continuous variables) over all these states. Since this amounts to taking a trace[3] over part of the system, it is called a *partial trace*, or "tracing out" subsystem A. The result of a partial trace over A is the reduced state

$$\rho_B = \sum_n \langle a_{nA}|\rho_{AB}|a_{nA}\rangle, \tag{2.27}$$

where $\{|a_{nA}\rangle\}$ is a basis of the state space of A. In particular, if A corresponds to one region of space, for instance consisting of the field degrees of freedom located in that region, whereas the subsystem B corresponds to a different region, one can also say that the region A has been traced out.

[3]The trace of a matrix ρ is the sum of its diagonal elements on any basis $\{|n\rangle\}$,

$$Tr\,(\rho) \equiv \sum \langle n|\rho|n\rangle.$$

As an example, let us calculate the reduced state of electron B when the global state is $|\varphi_e\rangle$, given by expression (2.25) above: the density matrix of the whole system is

$$\rho_e \equiv |\varphi_e\rangle\langle\varphi_e| = \frac{1}{2}\left[|+_{zA}+_{zB}\rangle\langle+_{zA}+_{zB}| - |-_{zA}-_{zB}\rangle\langle+_{zA}+_{zB}| + \ldots\right]. \quad (2.28)$$

When A is traced out, the first term survives, since the factors that refer to particle A are equal in the ket and the bra[4]. The second, however, is $|-_{zA}\pm_{zB}\rangle\langle+_{zA}\pm_{zB}|$, so it does not count toward the reduced state. After applying this procedure to every term, we find

$$\rho_B = \frac{1}{2}\left[|+_{zB}\rangle\langle+_{zB}| + |-_{zB}\rangle\langle-_{zB}|\right], \quad (2.29)$$

a mixed state. It appears despite the fact that, if we had access to the complete state $|\varphi_e\rangle$, and found A in the $|+_{zA}\rangle$ ($|-_{zA}\rangle$) state, then B would certainly be in $|+_{zB}\rangle$ ($|-_{zB}\rangle$). The mixture stems from our ignorance about A, either chosen deliberately (for example, by not performing measurements on that subsystem) or imposed by fundamental physical restrictions, such as the fact that the interior of a black hole is inaccessible to asymptotic observers.

This phenomenon, that the reduced state is mixed even though the global state is pure, arises because the subsystems A and B are *entangled*. In the case of a pure global state, such as $|\varphi_e\rangle$ in expression (2.25), this means simply that the system is not in a product state. For a generic state ρ_e, which can be pure or mixed, the subsystems are said to be entangled if ρ_e can not be written as a convex combination of product states [8],

$$\rho_e \neq \sum_i p_i \rho_A^i \otimes \rho_B^i. \quad (2.30)$$

2.3 Entanglement entropy

According to statistical mechanics, a classical system that can occupy different microstates with a probability distribution $\{p_k\}$ possesses an entropy[5]

$$S_{stat}(\rho) = -\sum_k p_k \ln p_k. \quad (2.31)$$

Similarly, a mixed quantum state ρ can be assigned a statistical entropy. In that case, the $\{p_k\}$ in equation (2.31) are given by the eigenvalues of ρ, which represent the

[4]For convenience, we perform the derivation in the $\{|+_z\rangle, |-_z\rangle\}$ basis, but the results are independent of that choice.

[5]We note that k_B, the Boltzmann constant, does not appear explicitly in any expression for the entropy because we use natural units.

probabilities of finding the system in each of the orthonormal $|\psi_k\rangle$ that diagonalize the density matrix (cf equation (2.11)). This gives the *von Neumann entropy* of the state. It can also be expressed directly in terms of ρ as[6]

$$S_{vN}(\rho) \equiv -k_B \mathrm{Tr}(\rho \ln \rho). \tag{2.32}$$

Like every entropy, $S_{vN}(\rho)$ is zero if ρ is a pure state and positive otherwise. Note, however, that it is possible for a system AB to have zero von Neumann entropy (corresponding to a pure state) even when the subsystems A and B both have positive entropy (mixed states), because S_{vN} is not necessarily additive over entangled subsystems. For a product state[7], on the other hand, additivity can be proven as follows: let the state of the system be

$$\rho_{AB} = \rho_A \otimes \rho_B. \tag{2.33}$$

In the bases that diagonalize ρ_A and ρ_B,

$$\rho_{AB} = \sum_{n,m} p_{nA} p_{mB} |\psi_{nA}\psi_{mB}\rangle\langle\psi_{nA}\psi_{mB}|, \tag{2.34}$$

subject to the normalization condition for ρ_A and ρ_B:

$$\sum_n p_{nA} = \sum_m p_{mB} = 1. \tag{2.35}$$

Since the density matrix is diagonal, its logarithm is simply

$$\ln \rho_{AB} = \sum_{n,m} \ln(p_{nA} p_{mB}) |\psi_{nA}\psi_{mB}\rangle\langle\psi_{nA}\psi_{mB}|. \tag{2.36}$$

The entropy of ρ_{AB} is therefore

$$S_{vN}(\rho_{AB}) = \sum_{n,m} p_{nA} p_{mB} \ln(p_{nA} p_{mB}). \tag{2.37}$$

[6]The logarithm of a matrix can be evaluated using the identity

$$\ln \rho = \lim_{n \to 0} \frac{\rho^n - \mathbb{I}}{n},$$

which arises in the context of the *replica trick*. Another possibility is to express the von Neumann entropy directly in terms of the eigenvalues of ρ, as per equation (2.31).

[7]The most general non-entangled (that is, separable) state is given by a convex combination of product states. However, one can only derive equalities for the entropy of such a state in special cases, and for the purposes of this work, considering a single product state is sufficient.

Considering the normalization of ρ_A and ρ_B, the von Neumann entropy of a product state can then be written as

$$\begin{aligned} S_{vN}\left(\rho_A \otimes \rho_B\right) &= \left(\sum_n p_{nA} \ln p_{nA}\right)\left(\sum_m p_{mB}\right) + \left(\sum_n p_{nA}\right)\left(\sum_m p_{mB} \ln p_{mB}\right) \\ &= \sum_n p_{nA} \ln p_{nA} + \sum_m p_{mB} \ln p_{mB} \\ &= S_{vN}\left(\rho_A\right) + S_{vN}\left(\rho_B\right), \end{aligned} \tag{2.38}$$

the sum of the entropies of ρ_A and ρ_B.

In particular, we are interested in the entropy that arises because part of a system is inaccessible: suppose a bipartite system AB is in a pure state ρ_{AB}, but one only has access to one of the subsystems. If its reduced state, ρ_{red}, is mixed, it possesses non-zero entropy. Since this is only the case if A and B are entangled, the quantity is known as *entanglement entropy*:

$$S_e\left(\rho_{AB}, AB\right) \equiv S_{vN}\left(\rho_{red}\right). \tag{2.39}$$

In most of this work, when the meaning is clear from context, both the von Neumann and the entanglement entropy will be denoted simply by S. The purpose of the more explicit notation above is to highlight a fact which, if overlooked, could cause some confusion: The entanglement entropy is a global property of the whole system, depending both on the global state ρ_{AB} and on the division into the subsystems A and B. Nevertheless, it can be defined and calculated in terms of the reduced state ρ_{red} of only one of the subsystems.

Notably, if the global state ρ_{AB} is pure, then it does not matter which of the subsystems one uses to calculate the entanglement entropy: whether one traces over A and takes $-Tr\left[\rho_B \ln \rho_B\right]$ or vice versa, the entropy is the same. This is known as the *symmetry theorem*:

$$S_{vN}\left(\rho_A\right) = S_{vN}\left(\rho_B\right) = S. \tag{2.40}$$

Proof: the density matrix of a pure state can be written as $\rho_{AB} = |\varphi\rangle\langle\varphi|$, for some $|\varphi\rangle = \sum C_{ij}|a_i b_j\rangle$, where $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$ are bases of the Hilbert spaces of A and B. Treating the coefficients C_{ij} as matrices, the matrix elements of the reduced density operators on the respective bases are $\left(\rho_A\right)_{ij} = \left(CC^\dagger\right)_{ij}$ and $\left(\rho_B\right)_{ij} = \left(C^T C^*\right)_{ij} = \left(C^\dagger C\right)_{ji}$. Now, the trace of a product of matrices is unchanged under a cyclic permutation of the factors, so that $Tr\left(\rho_A^k\right) = Tr\left(\rho_B^k\right)$ for all exponents

$k \in \mathbb{N}$. In terms of the eigenvalues p_{nA} of ρ_A and p_{mB} of ρ_B, we have

$$\sum_n p_{nA}^k = \sum_m p_{mB}^k \qquad \forall k \in \mathbb{N}, \tag{2.41}$$

which can be true only if ρ_A and ρ_B have the same eigenvalues, apart from extra zeroes. Considering equation (2.31), the entropies of the subsystems must therefore be equal.

The symmetry theorem is important for the following argument, proposed by Srednicki [7]: suppose A and B are two spatial regions of a larger whole, whose total volume is fixed. When one changes the division into A and B, the reduced states ρ_A and ρ_B also change, and their entropy can either increase or decrease, but it must do the same for both A and B. This implies in particular that it can not scale with the volume, since when A grows, B must shrink. Instead, S must be given by some expression that is symmetric under an exchange of A and B. In particular, it can be a function of some feature that both regions have in common, such as the surface area or the geometry of their boundary. We note that this reasoning does not guarantee that S will be a linear function of the area, therefore, if our data do indeed reproduce the Bekenstein-Hawking entropy (equation (4.21)), we have a non-trivial result. The symmetry theorem does, however, rule out extensivity (scaling with the volume). This sets entanglement entropy apart from the entropy of ordinary thermodynamic systems, and qualifies it as a promising candidate for black hole entropy.

2.4 The quantum simple harmonic oscillator

This section gathers some basic facts about the quantum simple harmonic oscillator, mainly for ease of reference: since the same formalism applies to the modes of (quantum) fields, later sections draw on a number of concepts and results presented here.

Consider a quantum mechanical system governed by the Hamiltonian

$$H = \frac{1}{2m}P^2 + \frac{1}{2}m\omega^2 X^2, \tag{2.42}$$

where P and X are the momentum and position observables, respectively, satisfying the commutation relation

$$[X, P] = \imath \tag{2.43}$$

in natural units. This describes a one-dimensional simple harmonic oscillator with mass m and frequency ω. We use the real number x to denote the eigenvalues of X,

and also as an index for the corresponding eigenstates:

$$X|x\rangle = x|x\rangle. \tag{2.44}$$

In the position representation, that is, on the basis formed by these states, $\{|x\rangle\}$, the "expansion coefficients" of a generic ket $|\varphi\rangle$ constitute its wave function,

$$\varphi(x) \equiv \langle x|\varphi\rangle. \tag{2.45}$$

In this sense, a function (be it a wave function or, by extension, a density operator) of a variable $x \in \mathbb{R}$ represents a state "of the oscillator x", that is to say, of a system of the kind described above, whose position eigenvalues are x.

Introducing the destruction operator,

$$a \equiv \frac{1}{\sqrt{2}}\left[\sqrt{m\omega}X + \imath\frac{1}{\sqrt{m\omega}}P\right], \tag{2.46}$$

and its adjunct, the creation operator $a^\dagger$, the Hamiltonian can be written as

$$H = \omega\left(a^\dagger a + \frac{1}{2}\right). \tag{2.47}$$

Its eigenstates, denoted by $|n\rangle$, possess well-defined energy

$$E_n = \omega\left(n + \frac{1}{2}\right) \tag{2.48}$$

and are connected by the creation and destruction operators,

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \qquad a|n\rangle = \sqrt{n}|n-1\rangle. \tag{2.49}$$

In particular, the ground state, $|0\rangle$, is annihilated by a. In the position representation, this fact is expressed by the differential equation

$$\frac{1}{\sqrt{2m\omega}}\left[m\omega x + \frac{d}{dx}\right]\langle x|0\rangle = 0. \tag{2.50}$$

The normalized ground-state wave-function is therefore

$$\varphi_0(x) \equiv \langle x|0\rangle = \left(\frac{m\omega}{\pi}\right)^{1/4}\exp\left(-\frac{1}{2}m\omega x^2\right). \tag{2.51}$$

Starting from $|0\rangle$, the excited states can be found by repeated application of the creation operator. Once again in the position representation, this gives the wave

functions

$$\begin{aligned}\varphi_n(x) &\equiv \langle x|n\rangle = \frac{1}{\sqrt{n!}}\langle x|\left(a^\dagger\right)^n|0\rangle \\ &= \frac{1}{\sqrt{n!\,(2m\omega)^n}}\left[m\omega x - \frac{d}{dx}\right]^n \varphi_0(x) \\ &= \left(\frac{m\omega}{\pi}\right)^{1/4}\exp\left(-\frac{1}{2}m\omega x^2\right)\frac{1}{\sqrt{n!2^n}}H_n\left(\sqrt{m\omega}x\right), \end{aligned} \tag{2.52}$$

where H_n is an n-th degree Hermite polynomial.

2.5 Thermal states

When the oscillator described above is immersed in a heat bath at a temperature T, it will assume a mixed state: there is a probability

$$p_n = Z^{-1}\exp\left(-E_n/T\right) \tag{2.53}$$

of finding it in the (non-degenerate) stationary state $|n\rangle$ with energy E_n. In that expression, the partition function

$$Z \equiv \sum_{n=0}^{\infty}\exp\left(-E_n/T\right) = \frac{\exp\left(-\omega/2T\right)}{1-\exp\left(-\omega/T\right)} \tag{2.54}$$

ensures the normalization ($\sum_n p_n = 1$). Introducing the variable

$$\xi \equiv \exp\left(-\omega/T\right), \tag{2.55}$$

the probabilities p_n, as explicit functions of temperature, frequency and energy level, can be written simply as

$$p_n = [1-\xi]\,\xi^n. \tag{2.56}$$

Substituting this expression into equation (2.31), one finds the von Neumann entropy of the thermal state:

$$\begin{aligned} S &= -(1-\xi)\sum_{n=0}^{\infty}\xi^n\left[\ln(1-\xi) + n\ln\xi\right] \\ &= -\ln(1-\xi) - \frac{\xi}{1-\xi}\ln\xi. \end{aligned} \tag{2.57}$$

If we intend to use this result in our calculations, we must ensure that the state whose entropy we wish to compute is indeed a thermal state. That is, its density

operator must be formally like that of the thermal state, which is given by

$$\rho_T = Z^{-1} \sum_{n=0}^{\infty} \exp\left(-E_n/T\right) |n\rangle\langle n| = Z^{-1} \exp\left(-H/T\right). \tag{2.58}$$

In order to make this comparison, it is more convenient to express ρ_T in the position representation,

$$\rho_T\left(x, x'\right) \equiv \langle x|\rho_T|x'\rangle, \tag{2.59}$$

where $|x\rangle$, $|x'\rangle$ are eigenvectors of the position operator X. Combining equations (2.52) and (2.58), we have

$$\begin{aligned} \rho_T\left(x, x'\right) = & \frac{\sqrt{m\omega/\pi}}{1 - \exp\left(-\omega/T\right)} \exp\left(-\frac{1}{2} m\omega x^2 - \frac{1}{2} m\omega x'^2\right) \\ & \times \sum_{n=0}^{\infty} \frac{1}{n! 2^n} \exp\left(-n\omega/T\right) H_n\left(\sqrt{m\omega}x\right) H_n\left(\sqrt{m\omega}x'\right). \end{aligned} \tag{2.60}$$

It is, however, more convenient to write this as

$$\rho_T\left(x, x'\right) = \sqrt{\frac{\gamma - \beta}{\pi}} \exp\left[-\frac{\gamma}{2}\left(x^2 + x'^2\right) + \beta x x'\right]. \tag{2.61}$$

The equality is proven in appendix A.2 and the coefficients are related to the physical parameters by

$$m = 1 \qquad \omega = \sqrt{\gamma^2 - \beta^2} \qquad T = \omega / \ln\left(\frac{\gamma + \omega}{\beta}\right). \tag{2.62}$$

Substituting these in equation (2.57) gives the entropy associated with the density matrix (2.61).

Finally, we generalize the above results to a system of N non-interacting oscillators: consider the state

$$\begin{aligned} \rho_T\left(x, x'\right) &= c \prod_{i=1}^{N} \exp\left[-\frac{1}{2}\left(x_i \bar{\gamma}_i x_i + x'_i \bar{\gamma}_i x'_i\right) + x_i \bar{\beta}_i x'_i\right] \\ &= c \exp\left[-\frac{1}{2} \sum_{i,j=1}^{N} \left(x_i \bar{\gamma}_i \delta_{ij} x_j + x'_i \bar{\gamma}_i \delta_{ij} x'_j\right) + \sum_{i,j=1}^{N} x_i \beta_i \bar{\delta}_{ij} x'_j\right] \\ &= c \exp\left[-\frac{1}{2}\left(x^T \bar{\gamma} x + x'^T \bar{\gamma} x'\right) + x^T \bar{\beta} x'\right], \end{aligned} \tag{2.63}$$

where $\bar{\gamma}$ and $\bar{\beta}$ are diagonal matrices, and c is the normalization constant. Since this is a product of density matrices like (2.61), each x_i corresponds to an oscillator

in a thermal state, whose frequency ω_i and temperature[8] T_i are encoded in the pair of coefficients $\left(\bar{\beta}_i, \bar{\gamma}_i\right)$, according to equation (2.62). It follows that each oscillator contributes an entropy S_i, given by expression (2.57). Furthermore, since the system is in a product state, according to the demonstration leading up to equation (2.38), the entropy is additive:

$$S = \sum_{i=1}^{N} S_i. \tag{2.64}$$

[8]Note that the oscillators must not necessarily be at the same temperature, since they do not interact.

3 General relativity

In section 2.2, we have seen how quantum theory predicts that entropy arises from pure states through the "ignorance" of an observer, but we have not explored the origin of that lack of information. Quantum theory would most likely attribute it to a choice the observer made, to ignore measurements performed on a part of the system. The associated entropy would therefore be highly subjective, and not particularly meaningful. General relativity, on the other hand, predicts situations in which fundamental physical principles render regions of spacetime inaccessible for observation. We present two of them below: the most prominent example, black holes, and the Rindler universe, where a similar effect occurs in flat spacetime.

Moreover, at least in the case of black holes, relativistic considerations lead one to postulate the existence of an entropy, yet without providing a clear picture of its microscopic origin. The details of this prediction, which must be accounted for by any theory claiming to explain black hole entropy, will also be discussed below. First, however, we introduce the formalism that general relativity uses to describe what a given observer can or cannot see. [9]

3.1 Causal structure

General relativity describes spacetime as a 3+1-dimensional pseudo-Riemannian manifold M, characterized by a metric $g_{\mu\nu}$ (for which we adopt the sign convention $(-,+,+,+)$) and covered by coordinates x^μ. Particles traverse this manifold on parametrized curves $x^\mu(\lambda)$, known as *world lines*. Precluding the possibility of superluminal travel, which would lead to paradoxes, the tangent to these curves must be either *timelike*,

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} < 0, \tag{3.1}$$

or *lightlike*,

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = 0. \tag{3.2}$$

Lightlike intervals (also called null, since the norm is zero) can only be traversed by massless particles, such as photons. For this reason, null geodesics are often identified with light rays. Massive particles, on the other hand, follow timelike world lines.

Only such particles (be they massive or massless) can carry information and thereby affect a system, for instance triggering a detector. Hence, an event A in spacetime can only be causally related to another event A' if they can be connected by a curve that is time- or lightlike at all points. For a given A, the set of all such

[9] For a comprehensive introduction to general relativity, as well as in-depth details, we refer the reader to [9], in particular chapters 6 and 8.

A' can be divided into the *causal past*, $J^-(A)$, and future, $J^+(A)$, and its boundary is the *light cone*[10] of A (see figure 3.1(left)). Thus, even in a nontrivial spacetime, the orientation of the light cones at different points throughout the manifold allows one to visualize its causal structure.

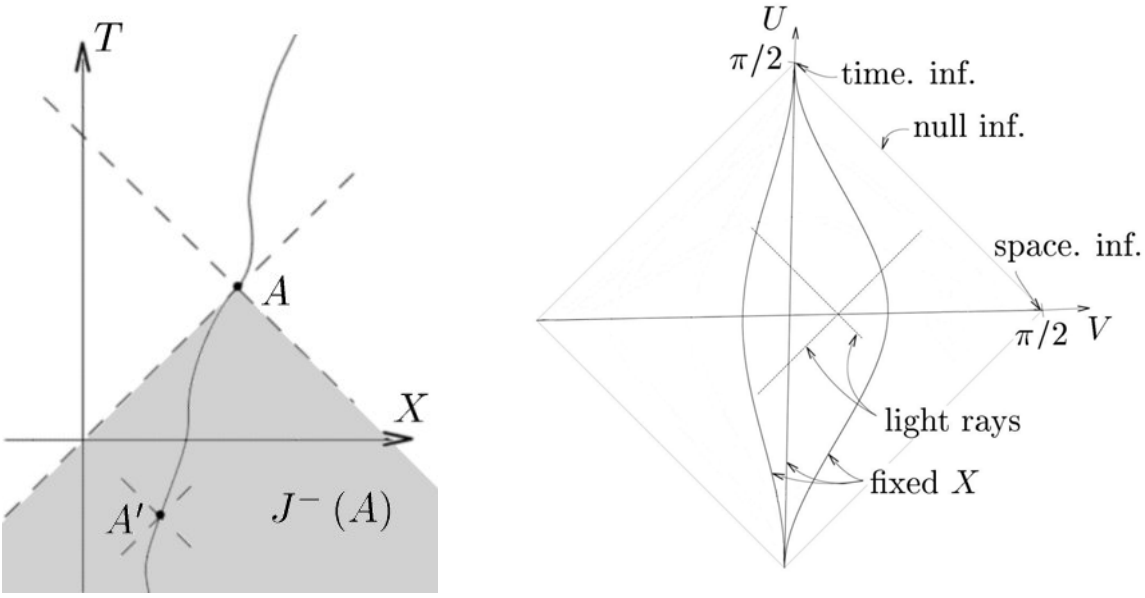

Figure 3.1: (left) In the 1+1-dimensional Minkowski universe, light rays (dashed lines) are at 45° angles with the Cartesian coordinate axes X and T at all points, so that all light cones have the same orientation. The world line of a massive particle (solid curve) lies locally inside the light cone at each point, since the particle can not reach or exceed the speed of light. Therefore, any two events on the world line, such as A and A', are timelike related: A' lies in the causal past (J^-, shaded) of A, so the former can causally affect the latter.
(right) In the Penrose diagram of Minkowski spacetime, all infinities are brought to a finite distance, yet light rays remain at 45° angles. Conversely, the curves of constant Cartesian coordinates X and T, such as the world lines of particles that remain stationary at fixed X, are no longer represented by straight lines.

This fact is exploited by Penrose diagrams: representations of a spacetime constructed such that lightlike curves are at 45° angles with the axes at all points. Furthermore, these diagrams are compact: the spacetime is mapped to a conformally related one[11] such that all infinities – spacelike, timelike and null – are brought to a finite distance. This is particularly useful for illustrating the definition of a black hole (see section 3.2.1 below), which refers to the *future null infinity* $\mathcal{I}^+$, generated by extending null geodesics from the whole manifold M into the future.

[10]It follows from this definition that the light cone of an event A is generated by the light rays that pass through that event.

[11]For the purpose of illustrating causal structures in the present work, it is sufficient to bring the infinities to a finite coordinate distance, without resorting to a different metric.

As an example, consider the Penrose diagram of two-dimensional Minkowski spacetime, in figure 3.1(right): light rays are already at 45° with the Cartesian coordinate axes X and T, so the transformation to the new coordinates

$$\begin{aligned} U &\equiv \tfrac{\arctan(T+X)+\arctan(T-X)}{2} \\ V &\equiv \tfrac{\arctan(T+X)-\arctan(T-X)}{2} \end{aligned} \tag{3.3}$$

serves solely to bring the infinities to a finite coordinate distance.

Based on these concepts, we can now define black holes and make a rigorous statement about how they prevent certain observers from attaining information about part of their universe.

3.2 Curved spacetime: black holes

3.2.1 Defining properties of black holes

Black holes are essentially regions of spacetime from which nothing, in particular no information, can escape. A more rigorous definition is given by Wald, in terms of the manifold M (which represents the entire spacetime) and the causal past[12] J^- of the future null infinity $\mathcal{I}^+$, as defined in the previous section:

> A strongly asymptotically predictable spacetime is said to contain a black hole if M is not contained in $J^-(\mathcal{I}^+)$. The black hole region, B, of such a spacetime is define to be $B = [M - J^-(\mathcal{I}^+)]$, and the boundary of B in M, $H = \dot{J}^-(\mathcal{I}^+) \cap M$, is called the event horizon. (Reference [9], p 300)

This definition is illustrated by a Penrose diagram of the Schwarzschild spacetime (figure 3.2 below). It shows in particular that no massive objects (following timelike world lines) nor massless carriers of energy and momentum (lightlike world lines) can return from beyond the event horizon and reach an asymptotic observer (in the future null infinity). Therefore, said observer can not obtain any information regarding the state of the interior of the black hole. On the other hand, an observer inside the event horizon does see parts of the interior, which suggests that they may have a different perspective on the entropy of the system. We will, however, leave the issue of observer-dependent entropy to other works, and restrict ourselves to the point of view of asymptotic observers where black holes are concerned.

As an aside, we note that the above definition makes no reference to singularities. Roughly speaking, these are "points" in spacetime where the curvature becomes infinite, although defining the concept rigorously is far from trivial.[13] The fact

[12] We furthermore use $\dot{J}^-$ to denote the boundary of the causal past.

[13] For a discussion, see reference [9], chapter 9.

that one can study at least some aspects of black holes without dealing with the singularity allows us to separate the object of our interest, quantum effects due to the event horizon, from the problem of physics in extremely curved space. In fact, if the black hole is large enough, the curvature at its horizon becomes negligible, and one can use flat space as an approximation. This is exploited in a number of works on the subject, including this one.

3.2.2 The Schwarzschild solution

For definiteness, we review some characteristics of an actual black hole spacetime. The simplest example is due to Schwarzschild: a static[14], spherically symmetric vacuum solution of the Einstein equations. The most general such metric can be written as

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 \left[d\theta^2 + \sin^2\theta d\phi^2\right], \tag{3.4}$$

where t is the temporal coordinate, θ and ϕ are angular and r is a radial coordinate, defined such that the area of the two-sphere at constant r is $4\pi r^2$. The parameter M gives the mass that asymptotic observers associate with the black hole.

This can be seen by considering stationary test bodies: in order for a body to remain at a fixed radial coordinate r in this spacetime, its proper acceleration must be

$$a = M/\left[r^2\sqrt{(1 - 2M/r)}\right]. \tag{3.5}$$

Thus, one finds that the Schwarzschild metric is asymptotically flat (when $r \gg M$), and in the weak field limit it reproduces the Newtonian gravity of a body of mass M. The other extreme is $r_S = 2M$, the Schwarzschild radius, where bodies must have infinite proper acceleration to remain stationary and are therefore unable to escape in the direction of increasing r. Indeed, a global analysis shows that nothing can reach an asymptotic observer from inside the surface at $r = r_S$, which therefore constitutes an event horizon. Given the definition of the radial coordinate, its area is

$$A_S = 16\pi M^2. \tag{3.6}$$

It can be shown that the metric diverges at this point only due to an unfortunate choice of coordinates, whereas at $r = 0$ lies a true, physical singularity.

Figure 3.2(right) shows a Penrose diagram of a Schwarzschild spacetime. The angular dimensions are suppressed, since they simply generate S_2 hypersurfaces. For the remaining degrees of freedom, one introduces the Kruskal coordinates, which are

[14]Since the Schwarzschild spacetime is static, it describes an eternal black hole, and not, rigorously, the remains of an actual collapsed star. It is, however, a useful approximation.

related to the radial and temporal coordinates used above by

$$X^2 - T^2 = \left(\frac{r}{2M} - 1\right) \exp\left(\frac{r}{2M}\right), \tag{3.7}$$

$$\operatorname{arctanh}\left(\frac{T}{X}\right) = \frac{1}{2}\frac{t}{2M}, \tag{3.8}$$

in order to put all null geodesics at 45°. Then the diagram is compactified using Penrose coordinates, which are given in terms of X and T by the same expressions as for the Minkowski spacetime, equation (3.3).

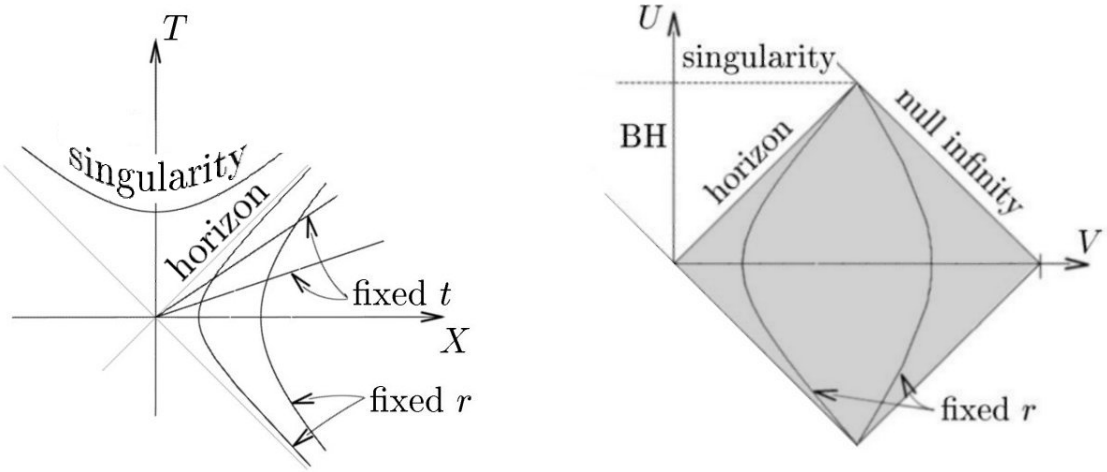

Figure 3.2: (left) Schwarzschild spacetime in Kruskal coordinates, X and T, defined such that all null geodesics lie at 45°. They illustrate the impossibility of returning from beyond the event horizon ($r = 2M$, $t \to \infty$): since it, too, lies at 45°, escaping would require faster-than-light travel. The coordinates r and t, on the other hand, are associated with observers that remain at a fixed proper distance from the central singularity ($r = 0$).
(right) The Penrose diagram of an eternal Schwarzschild black hole spacetime, in which the future null infinity, $\mathcal{I}^+$, is reduced to a finite region, shows that its causal past, $J^-(\mathcal{I}^+)$ (shaded), does not cover the whole spacetime. The complement of the causal past is the black hole (BH). (The curves $r/r_S = 1.032$ (closer to the horizon) and $r/r_S = 1.72$ (farther away) were plotted as examples of the world lines of observers that lie at rest outside the hole.)

There are more general solutions, containing black holes with charge and angular momentum, but since we only intended to give a concrete example of a black hole for the purpose of illustration, the Schwarzschild solution will suffice.

3.3 Flat spacetime: the Rindler wedge

The Minkowski universe is arguably the simplest spacetime conceivable: it is a vacuum solution, with identically zero curvature tensor, symmetric under translations,

boosts and rotations, and the causal structure is likewise trivial (see figure 3.1, on page 3.1). In particular, there are no event horizons like those found around black holes. However, this trivial spacetime can model a surprising number of features usually associated with black holes, when it is described from the point of view of a different family of observers.

Since an observer that wants to remain stationary outside a black hole must be constantly accelerated, let us consider a family of observers that traverse the Minkowski universe with constant proper acceleration. These are known as Rindler observers. For simplicity, we again restrict ourselves to 1+1 dimensions. Let X and T be the usual coordinates of Minkowski spacetime, relative to which a given family of inertial observers is at rest. We then introduce new coordinates, x and t, such that each Rindler observer remains stationary at a fixed x, while ∂_t is tangent to their world line. A convenient choice[15] is given implicitly by

$$X^2 - T^2 = \exp(2x) \qquad \operatorname{arctanh}\left(\frac{T}{X}\right) = t. \tag{3.9}$$

With x and t assuming all real values, this does not cover the whole Minkowski universe, but only the region $X > |T|$, known as the right Rindler wedge. Analogous sets of coordinates can be introduced to cover the remainder of the Minkowski universe.

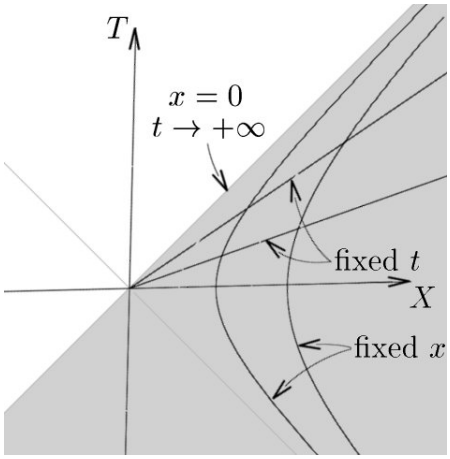

Figure 3.3: In the Minkowski universe, the Cartesian coordinates, X and T, put light rays at 45°. The coordinates x and t are associated with the Rindler observers, who have constant proper acceleration and experience a horizon (at $x = 0$, $t \to +\infty$), since their causal past only covers the region $X > T$ (shaded).

The world line of the Rindler observer at a given x is

$$X = \exp x \cosh t, \qquad T = \exp x \sinh t, \tag{3.10}$$

[15]There are, in fact, infinitely many families of Rindler observers, generated by letting $X^2 - T^2 = \alpha^{-1} \exp(2\alpha x)$ and $\operatorname{arctanh}(T/X) = \alpha t$ with different $\alpha \in \mathbb{R}$, but for our purposes, one will suffice.

shown in figure 3.3. The corresponding proper acceleration is $a = \exp(-x)$, which goes to zero for large x and diverges as $x \to -\infty$. Thus, there is no escape (in the direction of increasing x) from the region $X < |T|$, and in particular no information from it can reach the Rindler observers. They experience a horizon at $X = T$, which is the boundary of the causal past of their world lines. The parallels with the Schwarzschild spacetime as seen by observers at fixed r are illustrated by a comparison of figures 3.3 and 3.2. These structural similarities suggest that there may also be analogies between phenomena occurring in the two spacetimes. Of particular interest to the present work are two effects predicted by quantum field theory, which we will now explore.

4 Quantum field theory in curved spacetime

Since a complete theory of quantum gravity, involving the quantization of the gravitational field, is not yet available, we resort to the next best alternative: the quantization of matter fields in a background determined by (classical) general relativity. This is a semiclassical theory of gravity, known as quantum field theory in curved spacetime, and it is a reasonable approximation as long as the curvature is not at the Planck scale.[16]

Notably, this approach leads to novel phenomena even before one considers actually curved spacetime, but simply due to the use of general relativistic formalism and concepts, in particular to describe accelerated observers. This is the case of the Unruh effect, which occurs in the Rindler universe, but the same ideas also explain Hawking radiation, emitted by black holes. The two effects are discussed below, in view of their significance for the present work: both predict radiation with a thermal spectrum. In a search for the origin of a black hole's entropy, a mechanism that assigns a temperature to it is a promising start. Moreover, this mechanism is semiclassical and also works in flat space. We consider this evidence in favour of beginning our search for black hole entropy in flat spacetime as well.

4.1 The Unruh effect

Although this effect actually occurs in flat spacetime, it illustrates how different observers can assign different particle contents to a given state of a quantum field. We only give a brief outline of its derivation here, and refer the reader to [12] for a more comprehensive review.

Let φ be a scalar field of mass M, subject to the Klein-Gordon equation

$$\left(\Box - M^2\right)\varphi = 0. \tag{4.1}$$

We use a Minkowski background, which we reduce to two dimensions for simplicity:

$$ds^2 = -dT^2 + dX^2. \tag{4.2}$$

For stationary observers, T is the proper time and X is the spatial coordinate, since ∂_T is tangent to their world lines and ∂_X is orthogonal to it. Therefore, the eigenfunctions of the operators ∂/∂_T and ∂/∂_X are the energy-momentum eigenfunctions

[16] For a general introduction to the theory, we recommend the book by Birrell and Davies [10]. More details on the topics that we introduce here can also be found in Jacobson's lecture notes [11].

for these observers: they are given by

$$u_k\left(T,X\right) \propto \exp\left[-\imath\omega T + \imath k X\right], \tag{4.3}$$

(omitting a normalization constant), with $\omega = \sqrt{M^2+k^2}$. These modes, with $\omega > 0$, are positive-frequency eigenstates of the field Hamiltonian, and correspond to particles. Antiparticles can also be written in terms of positive ω, as

$$u_k^*\left(T,X\right) \propto \exp\left[-\imath\left(-\omega\right)T + \imath\left(-k\right)X\right]. \tag{4.4}$$

Together these modes form a basis, on which one can expand the state of the field:

$$\varphi\left(T,X\right) = \int dk \left[a_k u_k\left(T,X\right) + a_k^\dagger u_k^*\left(T,X\right)\right]. \tag{4.5}$$

When the field is treated as an operator, the coefficients $a_k^\dagger$ and a_k acquire the same status, and they are identified with the creation and destruction of quanta of momentum k, respectively. They allow us to write the number of excitations, or particles, in the mode k as

$$N_k \equiv a_k^\dagger a_k \tag{4.6}$$

The state with zero particles in all modes, that is, the one that is annihilated by all destruction operators a_k, is the Minkowski vacuum $|0_M\rangle$:

$$a_k|0_M\rangle = 0 \;\forall k. \tag{4.7}$$

Now we turn to Rindler observers, whose natural choice of coordinates are x and t, introduced in section 3.3. The states to which these observers assign well-defined energy and momentum are given by eigenfunctions of ∂/∂_t and ∂/∂_x, namely

$$v_k\left(t,x\right) \propto \exp\left[-\imath\omega t + \imath k x\right] \tag{4.8}$$

with $\omega = \sqrt{M^2+k^2}$ for particles, and v_k^* for antiparticles. Let $b_k^\dagger$ and b_k be the creation and destruction operators for these modes, so that a given field can also be expanded as

$$\varphi\left(t,x\right) = \int dk \left[b_k v_k\left(t,x\right) + b_k^\dagger v_k^*\left(t,x\right)\right]. \tag{4.9}$$

We are interested in the relation between the two sets of modes, u_k and v_k, which is conveniently expressed by the Bogoliubov transformations between the operators associated with them:

$$a_k = \int d^3k' \left[b_{k'}\alpha_{k'k} - b_{k'}^\dagger \beta_{k'k}^*\right], \tag{4.10}$$

and its conjugate for $a_k^\dagger$. The details of the calculation of the coefficients, $\alpha_{k'k}$ and $\beta_{k'k}$, exceed the scope of this introduction, but they can be found in [13]. The crucial point is that the $\beta_{k'k}$ are not identically zero, so that the annihilation operators of one set of modes are partially transformed into creation operators of the other set. Therefore, the number of particles that the Rindler observers detect in a mode k (given by the observable $b_k^\dagger b_k$) when traversing the Minkowski vacuum (which is annihilated only by the a_k, as expressed in equation (4.7)) is

$$\langle 0_M | b_k^\dagger b_k | 0_M \rangle \neq 0. \tag{4.11}$$

In fact, it can be shown that the Rindler observer with proper acceleration a experiences a thermal bath at temperature

$$T_U = a/2\pi \tag{4.12}$$

in the Minkowski vacuum, whereas inertial observers detect no particles in this state. This effect does not only illustrate the observer-dependence of the particle content in general, but in fact bears a striking resemblance to the Hawking effect, to which we turn next.

4.2 The Hawking effect

The Hawking effect is the emission of thermal radiation from black holes, in close analogy with the Unruh effect, but occurring in the background spacetime of a black hole. It first appeared in the context of cosmological particle creation, whose derivation according to Birrell and Davies [10] is as follows:

Consider a quantum field in an asymptotically flat spacetime. In the past and future infinity, the field can be expanded in modes that are energy-momentum eigenstates of inertial observers. The number of excitations in these modes, $\langle \psi | N | \psi \rangle$, is the particle content that the observers associate with the field in a state $|\psi\rangle$. In particular, the state $|0\rangle$ such that $\langle 0|N|0\rangle = 0$ is the vacuum for them. However, if there is a perturbation of the spacetime between the past (*in*) and the future (*out*), the operators N_{in} and N_{out} are not necessarily the same. Thus, a state $|0_{in}\rangle$, that was initially vacuum, $\langle 0_{in}|N_{in}|0_{in}\rangle = 0$, can contain particles afterwards:

$$\langle 0_{in} | N_{out} | 0_{in} \rangle \neq 0. \tag{4.13}$$

Hawking applied this analysis to the spacetime of a star collapsing to form a black hole, and he found that particle creation does take place, that is, the black hole emits

radiation (known as Hawking radiation) during its formation. More surprisingly, the spectrum of this radiation is thermal.

Such radiation is also found to emanate from static black holes, though the derivation is slightly different: the state of the field must now be $|0_{HH}\rangle$, the Hartle-Hawking vacuum, since this is the only one that is regular (free of divergent observables). Furthermore, the observers that see the radiation in this case are those that are stationary outside the black hole: for them

$$\langle 0_{HH}|N_{stat}|0_{HH}\rangle \neq 0, \tag{4.14}$$

and the spectrum can again be shown to be thermal.

Based on the emission of thermal radiation, one attributes a temperature to the black hole. For observers close to the horizon, the perceived temperature depends on a redshift factor and can be expressed in terms of their proper acceleration a as

$$T_H^{obs} = a/2\pi.$$

In this limit, the temperature of the Hawking radiation is exactly like that of Unruh radiation (equation (4.12)). Measured at infinity, on the other hand, it depends solely on the surface gravity κ: a property of a stationary black hole, given by the force required to hold a unit mass at its horizon[17]. This is known as the Hawking temperature,

$$T_H = \kappa/2\pi, \tag{4.15}$$

and it plays an important role in our understanding of the thermodynamics of black holes.

4.3 Black hole thermodynamics

The formalism of thermodynamics is known for its applicability to a wide variety of physical systems, and indeed it can also describe black holes: If one makes the right identifications of thermodynamic parameters with properties of the black hole, one can recover the laws that govern them, as predicted by general relativity. However, there are additional insights to be gained from this approach, notably the fact that black holes must have entropy.

One simple yet compelling argument in favour of this idea is due to Bekenstein [5]: suppose an object with non-zero entropy falls into a black hole, effectively vanishing from the universe as seen by an observer outside the hole. The only way to prevent a decrease of the total entropy, which would violate the second law of

[17]This force is also measured at infinity, since, for an observer next to the mass, the force diverges.

thermodynamics, consists in postulating a corresponding increase in the entropy associated with some other system, and the black hole itself is the obvious choice.

A more rigorous statement follows from treating black holes as thermodynamic systems. This begins with the *no hair theorem* [14]: any black hole, when in equilibrium, is completely characterized by its mass M, its electric charge Q and its angular momentum $\vec{J}$. The area A of its event horizon, for example, can be expressed as a function of M, Q and $\vec{J}$. These variables therefore correspond to macroscopic parameters of ordinary thermodynamics[18]. The generalization of the first law makes this identification more precise: Bardeen, Carter e Hawking [15] proved that, for infinitesimal differences between two static black hole configurations,

$$dM = \frac{\kappa}{8\pi} dA + \vec{\Omega} \cdot d\vec{J} + \Phi dQ, \tag{4.16}$$

where κ is the surface gravity, $\vec{\Omega}$ is the angular velocity and Φ, the electric potential at the horizon. We can identify

$$dM = dE \tag{4.17}$$

as the change in the energy stored in the black hole and

$$\vec{\Omega} \cdot d\vec{J} + \Phi dQ = dW \tag{4.18}$$

is the work done on it. Analogy with the first law suggests

$$TdS = \frac{\kappa}{8\pi} dA. \tag{4.19}$$

The identification of the entropy S with A is supported[19] by the generalization of the second law: according to the area theorem [17, 9], the area of the event horizon never decreases in classical processes[20], and in fact increases in most of them:

$$dA \geq 0. \tag{4.20}$$

[18]It seems natural to identify M, Q, $\vec{J}$ and A as extensive parameters of the system, since one intuitively expects them to be additive over subsystems, and therefore scale with the size of the total system. However, the concepts involved do not simply carry over to black hole thermodynamics. For instance, we will see in equation (4.21) that A does not scale with the "size" of the black hole, if we assume that the latter is measured by its "volume" or mass. In fact, even properties like the additivity of mass over subsystems can not be taken for granted: one can not simply divide one black hole of mass M into two halves of $M/2$ each. These are but two examples of the puzzles that arise when one attempts to extend ordinary thermodynamics to black holes.

[19]Bekenstein, for example, makes this fact the basis for a thermodynamic treatment of black holes in reference [16].

[20]Quantum mechanics does allow for a decrease of the horizon area, for example in the process of evaporation by emission of Hawking radiation. The difference lies in the existence of vacuum energy, which violates one of the assumptions that the area theorem is based on.

This suggests that the entropy of a black hole should be a monotonically increasing function of its horizon area.

We note that, so far, we have only established a formal analogy between black hole dynamics, as governed by general relativity, and the laws of thermodynamics. It leads us to assume that there is a temperature T associated with an equilibrium black hole, and that it must be proportional to its surface gravity κ.[21] However, in order to find the value of the proportionality constant and to clarify the physical meaning of this temperature, one must turn to quantum field theory in curved spacetime. It predicts the Hawking temperature (equation (4.15)), and thereby allows us to isolate an expression for the entropy of the black hole:

$$S_{BH} = \frac{A}{4L_{Pl}^2}. \tag{4.21}$$

This is the Bekenstein-Hawking entropy.

In natural units, it reads simply $S_{BH} = A/4$, but we have reintroduced the Planck length, $L_{Pl} \equiv \sqrt{\hbar G/c^3}$, for later reference. We also wish to stress the fact that it is required for dimensional reasons: since the entropy is dimensionless, the area of the horizon must be divided by some other quantity that also has the dimension of area. However, the spacetime of the most general black hole is characterized by only one such quantity, the horizon area itself, so the second parameter must be an intrinsic constant of the theory that is used to describe the situation, such as the square of its length scale. We will return to this point in section 5.3.4.

Unlike the entropy of usual thermodynamic systems, S_{BH} is not extensive (scaling with the volume of the region), but proportional to the area of its boundary. This so-called *area law* is arguably the most prominent feature that any explanation of black hole entropy must account for.

We stress that the above derivation of black hole entropy does not draw on the observers' ignorance about the region hidden behind the event horizon, nor does it account for the microscopic origin of the entropy. However, given the case of the Hawking effect, it seems only natural to believe that this gap in the explanation can be bridged by quantum theory. The details and consequences of this idea are explored in the following sections.

[21] Indeed, one can prove that κ is constant over the horizon of an equilibrium black hole, in accordance with a generalized zeroth law.

5 Calculation of the entanglement entropy

We intend to explore whether the entanglement entropy that a ground-state quantum field acquires when a region of spacetime is inaccessible to observation can account for black hole entropy. The following section derives a technique for calculating this entropy, beginning with the simpler cases of coupled oscillators.

5.1 Two classical oscillators

Although our goal is a quantum-mechanical description, that encompasses phenomena like energy quantization and entanglement, we will first study the system in the framework of classical mechanics. This provides a concrete picture of the physical situation and illustrates the concept of normal modes, as well as the usefulness of the matrix formalism. As the simplest example, we turn to just two coupled harmonic oscillators.

Consider two equal point masses m, whose positions along a single common axis we denote by x_i, with indices $i = 1, 2$. Let each one be subject to a potential that allows harmonic oscillation with frequency ω_0 about an equilibrium position $x_i = 0$. This can be implemented, for instance, by attaching each mass to a wall[22] using a spring with constant $k_0 = m\omega_0^2$. The masses are also coupled to each other by a third spring, whose constant is $k_c = m\omega_c^2/2$ and whose undistended length equals the distance between the equilibrium positions of the masses. The setup is illustrated in figure 5.1.

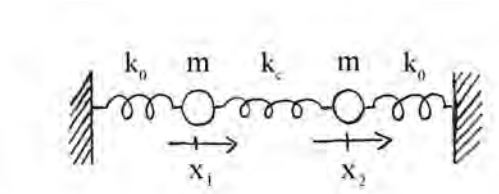

Figure 5.1: The two equal masses m are coupled by springs with different constants (k_0, k_c) such that their positions, x_1 and x_2, obey the equations for two coupled classical harmonic oscillators.

The potential energy of the oscillators is

$$E = \frac{1}{2}k_0x_1^2 + \frac{1}{2}k_0x_2^2 + \frac{1}{2}k_c\left(x_1 - x_2\right)^2. \tag{5.1}$$

[22]The idea that there are walls to which the springs are attached is not to be taken too literally: the potential in which the masses move is only that of an harmonic oscillator, and the range of the positions x_i is $]-\infty, +\infty[$.

This can be conveniently expresses in matrix notation: let $x \equiv (x_1, x_2)^T$ be a vector, and K be a *coupling matrix*, so that we can write

$$E = \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} k_0 + k_c & -k_c \\ -k_c & k_0 + k_c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \equiv \frac{1}{2} x^T K x. \tag{5.2}$$

K is always real, and we will choose it to be symmetric for convenience.

In terms of components, the equations of motion are

$$m \frac{d^2}{dt^2} x_i = -\frac{d}{dx_i} \frac{1}{2} \sum_{j,l=1}^{2} x_j K_{jl} x_l = -\frac{1}{2} \left(\sum_{l=1}^{2} K_{il} x_l + \sum_{j=1}^{2} x_j K_{ji} \right), \tag{5.3}$$

and since K is symmetric,

$$\frac{d^2}{dt^2} x_i = -\frac{1}{m} \sum_{j=1}^{2} K_{ij} x_j. \tag{5.4}$$

This is a set of coupled differential equations, because of the non-diagonal elements of K. However, since K is real and symmetric, there exists a matrix U that makes

$$K_D \equiv U K U^T \tag{5.5}$$

diagonal. The rows of U are given by the orthonormalized[23] eigenvectors of K, therefore the matrix is orthogonal: all its elements are real and

$$U U^T = U^T U = \mathbb{I}. \tag{5.6}$$

Furthermore, U does not depend on time, so we can introduce the vector

$$\overline{x} \equiv U x \tag{5.7}$$

in order to uncouple the equations of motion:

$$\frac{d^2}{dt^2} \bar{x}_i = -\frac{1}{m} \left(U K U^T U x \right)_i = -\frac{1}{m} \sum_{j=1}^{2} \left(K_D \right)_{ij} \bar{x}_j = -\frac{(K_D)_{ii}}{m} \bar{x}_i. \tag{5.8}$$

These equations describe another system of masses m, also confined to move in one dimension, but not coupled to each other. Their positions, given by the $\bar{x}_i$, are simple harmonic functions of time, with frequencies $\omega_i = \sqrt{(K_D)_{ii}/m}$. It is thus easier to solve for the $\bar{x}$ as a function of time, but the solution is equivalent

[23] One must only deliberately choose the eigenvectors to be orthogonal if K has degenerate eigenvalues. The eigenvectors of different eigenvalues are already orthogonal, since K is Hermitian.

to finding x, since $x = U^T \bar{x}$. The degrees of freedom described by these more convenient variables are the normal modes of the system.

In the case of two oscillators, there is straightforward interpretation of the normal modes. The transformation that diagonalizes the coupling matrix K is

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \tag{5.9}$$

and its eigenvalues are found to be

$$(K_D)_{11} = k_0 + 2k_c \qquad (K_D)_{22} = k_0. \tag{5.10}$$

Therefore

$$\begin{cases} \frac{d^2}{dt^2}\bar{x}_1 = -\frac{k_0+2k_c}{m}\bar{x}_1 \\ \frac{d^2}{dt^2}\bar{x}_2 = -\frac{k_0}{m}\bar{x}_2 \end{cases} \tag{5.11}$$

are the equations of motion for the normal mode variables

$$\begin{cases} \bar{x}_1 = \frac{1}{\sqrt{2}}(x_1 - x_2) \\ \bar{x}_2 = \frac{1}{\sqrt{2}}(x_1 + x_2). \end{cases} \tag{5.12}$$

This describes simply the uncoupled motion of the centre of mass $(x_1 + x_2)$ and the relative motion $(x_1 - x_2)$ of the particles.

5.2 N quantum oscillators

5.2.1 Ground state

The matrix formalism carries over readily to a system of N coupled one-dimensional oscillators, indexed by $i = 1, \ldots, N$. Let their masses be equal and unitary in natural units, that is, $m = M_{Pl} = 1$. Most importantly, we will now deal with quantum oscillators, which are described by wave-functions and density operators instead of single points in phase space. Therefore, the vector $X = (X_1, \ldots, X_N)^T$ consists of position operators, and we introduce the corresponding momentum operators $P = (P_1, \ldots, P_N)^T$. They satisfy the canonical commutation relations,

$$[X_i, X_j] = 0 = [P_i, P_j] \tag{5.13}$$

and

$$[X_i, P_j] = \imath\delta_{ij}. \tag{5.14}$$

Using matrix notation, the general Hamiltonian of such a system can be written as

$$H = \frac{1}{2}P^T P + \frac{1}{2}X^T K X, \tag{5.15}$$

with a generic coupling matrix K. In particular, since we are developing a general formalism, we do not assume that K couples only next neighbours (that is, $K_{ij} = 0$ if $|i - j| > 1$), which would restrict us to one-dimensional chains. In fact, we will see in later sections that this Hamiltonian can just as well describe networks of coupled oscillators in higher dimensions and with different geometries. Two assumptions we can make without loss of generality are that K is real and symmetric.

Again we introduce the orthogonal transformation U that diagonalizes K, as in equation (5.5), and the new variables

$$\bar{X} \equiv UX \qquad \bar{P} \equiv UP. \tag{5.16}$$

The commutation relations become

$$\left[\bar{X}_i, \bar{P}_j\right] = \sum_{k,l=1}^{N} U_{ik} U_{jl} \left[X_k, P_l\right] = \imath \sum_{k=1}^{N} U_{ik} \left(U^T\right)_{kj} = \imath \left(UU^T\right)_{ij} = \imath \delta_{ij}, \tag{5.17}$$

and similarly

$$\left[\bar{X}_i, \bar{X}_j\right] = 0 = \left[\bar{P}_i, \bar{P}_j\right]. \tag{5.18}$$

Since U is orthogonal (equation (5.6)), the momentum term is not altered,

$$P^T P = \bar{P}^T \bar{P}, \tag{5.19}$$

and the Hamiltonian becomes

$$\begin{aligned} H &= \frac{1}{2}\bar{P}^T \bar{P} + \frac{1}{2}\bar{X}^T K_D \bar{X} \\ &= \frac{1}{2}\sum_{i=1}^{N} \bar{P}_i^2 + \frac{1}{2}\sum_{i=1}^{N} K_{Dii} \bar{X}_i^2 \\ &\equiv \sum_{i=1}^{N} H_i\left(\bar{P}_i, \bar{X}_i\right). \end{aligned} \tag{5.20}$$

This is the Hamiltonian of a multipartite system whose subsystems do not interact, since there are no terms containing variables with different indices i. Just as in the classical treatment of two oscillators, each of the uncoupled subsystems is a simple harmonic oscillator of mass $m = 1$ and frequency $\omega_i = \sqrt{(K_D)_{ii}/m}$:

$$H_i = \frac{1}{2}\bar{P}_i^2 + \frac{1}{2}K_{Dii}\bar{X}_i^2, \tag{5.21}$$

whose ground-state wave-function, according to equation (2.51), is

$$\langle \bar{x}_i | 0_i \rangle = \varphi_0 (\bar{x}_i) = \left(\frac{\omega_i}{\pi} \right)^{1/4} \exp \left(-\frac{1}{2} \omega_i \bar{x}_i^2 \right). \tag{5.22}$$

According to equations (2.20) and (2.24), the ground state of the system is then described by the product

$$\varphi_0 (\bar{x}) = \pi^{-N/4} \left(\prod_{i=1}^{N} \omega_i \right)^{1/4} \exp \left(-\frac{1}{2} \sum_{i=1}^{N} \bar{x}_i \omega_i \bar{x}_i \right). \tag{5.23}$$

We introduce the matrix Ω_D, defined by

$$(\Omega_D)_{ij} \equiv \delta_{ij} \omega_i = \delta_{ij} \sqrt{(K_D)_{ii}}, \tag{5.24}$$

and

$$\Omega \equiv U^T \Omega_D U, \tag{5.25}$$

so that Ω is also real, symmetric and positive-definite. This allows us to write the wave function simply as

$$\begin{aligned} \varphi_0 (\bar{x}) &= \pi^{-N/4} (\det \Omega_D)^{1/4} \exp \left(-\tfrac{1}{2} \bar{x}^T \Omega_D \bar{x} \right) \\ &= \pi^{-N/4} (\det \Omega)^{1/4} \exp \left(-\tfrac{1}{2} x^T \Omega x \right) = \varphi_0 (x), \end{aligned} \tag{5.26}$$

and the ground-state density operator becomes

$$\rho_0 (x, x') = \varphi_0^* (x') \varphi_0 (x) = c \exp \left(-\frac{1}{2} x^T \Omega x - \frac{1}{2} x'^T \Omega x' \right). \tag{5.27}$$

For simplicity, we will not keep track of the different values of the normalization constant (c, c') over the course of the following calculation, and only normalize the density operator when computing its eigenvalues.

5.2.2 Reduced state

The next step is to take the partial trace over some of the oscillators and find the reduced state of the rest of the system, in order to calculate its entropy.

Let $j = 1, \ldots, n$ be the oscillators that are traced out. Note that they do not have to be physically adjacent; assigning them sequential indices is merely a matter of algebraic convenience. Also recall that, since the ground state is pure, the entanglement entropy depends only on the division into subsystems, but not on which one is traced out and which remains (see equation (2.40)). Thus, tracing out $j = n+1, \ldots, N$ gives the same entropy derived in the following.

The reduced density matrix is given by

$$\begin{aligned}&\rho_{red}\left(x_{n+1},\ldots,x_N;x'_{n+1},\ldots,x'_N\right)\\&\equiv\int\left[\prod_{j=1}^{n}dx_j\right]\rho_0\left(x_1,\ldots,x_N;x_1,\ldots,x_n,x'_{n+1},\ldots,x'_N\right).\end{aligned}\tag{5.28}$$

The technique used to calculate such integrals is described in Appendix A.1.3: one decomposes Ω into submatrices,

$$\Omega_{N\times N}=\begin{pmatrix}A_{n\times n} & B_{n\times(N-n)}\\ B^T_{(N-n)\times n} & C_{(N-n)\times(N-n)}\end{pmatrix},\tag{5.29}$$

in terms of which one defines the $(N-n)\times(N-n)$ matrices

$$\beta=\frac{1}{2}B^TA^{-1}B\qquad\gamma=C-\beta.\tag{5.30}$$

From this point on, we will only work in the space of $j=n+1,\ldots,N$, and for the sake of notational simplicity we relabel the x_j, introducing the $(N-n)$-component vector $(\tilde{x}_1,\ldots,\tilde{x}_{N-n})\equiv(x_{n+1},\ldots,x_N)$. Since Ω is real, symmetric and positive-definite (cf equation (5.25)), the identity (A.26) holds, and

$$\rho_{red}\left(\tilde{x},\tilde{x}'\right)=c'\exp\left[-\frac{1}{2}\left(\tilde{x}^T\gamma\tilde{x}+\tilde{x}'^T\gamma\tilde{x}'\right)+\tilde{x}^T\beta\tilde{x}'\right].\tag{5.31}$$

This expression is similar to (2.63), but we can only apply the results from that section if the coefficient matrices are diagonal. To that end, we introduce the variables $y\equiv\gamma_D^{1/2}V\tilde{x}$, where V is the orthogonal matrix that makes $\gamma_D\equiv V\gamma V^T$ diagonal. The reduced density matrix becomes

$$\rho_{red}\left(y,y'\right)=c'\exp\left[-\frac{1}{2}\left(y^Ty+y'^Ty'\right)+y^T\beta'y'\right],\tag{5.32}$$

with

$$\beta'\equiv\gamma_D^{-1/2}V\beta V^T\gamma_D^{-1/2}.\tag{5.33}$$

This definition has the following useful property: suppose the entire coupling matrix K is multiplied by a factor ν. Then so are Ω, and consequently γ and β, yet β' remains unchanged. Since all the following results, in particular the entanglement entropy, depend only on the latter, they will not be affected by a global factor in K.

Now let $z\equiv Wy$, with W such that $\beta'_D\equiv W\beta'W^T$ is diagonal. Then

$$\rho_{red}\left(z,z'\right)=c'\exp\left[-\frac{1}{2}\left(z^Tz+z'^Tz'\right)+z^T\beta'_Dz'\right].\tag{5.34}$$

This matrix does describe a system of uncoupled oscillators z_i (in the sense defined after equation (2.45)) in thermal states, as in equation (2.61), with the identification $\bar{\gamma} = \mathbb{I}$ and $\bar{\beta} = \beta'_D$. Denoting by β'_i the diagonal elements of β'_D, which are given by the eigenvalues of β', the frequencies and temperatures of the uncoupled oscillators z_i become, as per equation (2.62),

$$\omega_i = \sqrt{1^2 - (\beta'_i)^2} \qquad T_i = \omega_i / \ln\left(\frac{1+\omega_i}{(\beta'_i)}\right). \tag{5.35}$$

In terms of the variables

$$\xi_i \equiv \exp\left(-\frac{\omega_i}{T_i}\right) = \frac{\beta'_i}{1+\sqrt{1-(\beta'_i)^2}}, \tag{5.36}$$

the entropy contributions are

$$S_i = -\ln(1-\xi_i) - \frac{\xi_i}{1-\xi_i}\ln\xi_i, \tag{5.37}$$

Since we are dealing with a product state, according to equation (2.38), the total entropy of the reduced state is then the sum

$$S = \sum_{i=1}^{N-n} S_i. \tag{5.38}$$

5.2.3 Thermal state

We note in equation (5.35) that the temperatures T_i of the equivalent uncoupled oscillators z_i will not necessarily all be equal. However, that would have to be the case if we were to identify the reduced state as a thermal state: all field degrees of freedom in the region to which the observer has access would have to be in thermal equilibrium, at a single temperature (as is the case for Unruh and Hawking radiation). Thus, the condition for a thermal state is that all the β'_i be the same, that is,

$$\beta' = (\beta'_i)\,\mathbb{I}. \tag{5.39}$$

Substituting in equation (5.33), this is equivalent to

$$(\beta'_i)\,\gamma_D = V\beta V^T \quad \Leftrightarrow \quad \beta = (\beta'_i)\,\gamma \quad \Leftrightarrow \quad [\beta,\gamma] = 0. \tag{5.40}$$

It might be interesting to determine what conditions the coupling matrix K, and, by extension, the Hamiltonian must meet to guarantee that the reduced state be thermal, but, to our knowledge, this has not been investigated so far.

5.3 Quantum field

5.3.1 Choice of the background and traced-out region

Although we wish to understand entanglement entropy in the context of black holes eventually, we begin studying it in a flat background instead of an actual black hole spacetime. We note that this takes us beyond the usual approach, which merely approximates the spacetime near the horizon of a large black hole as being flat. Going to exactly flat space may seem like an oversimplification, but phenomena like the Unruh effect suggest that it can contain effects that are of interest for our investigation, and the results of our calculations bear this out. In fact, the discussion in section 6.2.2 suggests that flat space may be entirely equivalent to the actual curved background for the purpose of calculating entanglement entropy.[24]

The significant difference, of course, is that flat space does not contain an objective event horizon, as can be found in the Schwarzschild spacetime, nor will we take the point of view of a specific family of observers who experience a horizon, such as the Rindler observers. Instead, the region Ω, over which the partial trace is taken, is chosen arbitrarily. And even though the ignorance is deliberate, the boundary H does separate Ω, a region we know nothing about, from its complement $\bar{\Omega}$, whose reduced state and entropy are calculated. In that sense, H can be considered a horizon.

5.3.2 Choice of the spin and state of the field

In this work, we consider a real scalar field with mass M, in its ground state.

Again, simplicity is the main reason for this choice, since we intend a first foray into the subject of entanglement entropy. However, the scalar field should not be viewed solely as a toy model for higher spin fields: it is also relevant in its own right, for treating small perturbations to the metric.

Assuming the field to be in the ground state is not as restrictive as it might appear at a first glance. Das et al. [18], for example, argue that a simple coordinate transformation can reduce generalized coherent and squeezed states to the ground state, at least for the purposes of our calculation. Thus, the results extend to that case.

Finally, non-zero mass can be incorporated into the calculations with only minor modifications, allowing us to track the changes in entropy that it causes.

[24] It is also possible to reduce the curved-space problem to flat space rigorously under certain circumstances (such as in [18], appendix B), but we will not take that approach here.

5.3.3 Hamiltonian

In 3+1-dimensional flat space, the Lagrangian density of a real, massive scalar field φ is

$$\mathcal{L} = \frac{1}{2}\left(-\partial^{\alpha}\varphi\partial_{\alpha}\varphi - M^{2}\varphi^{2}\right), \tag{5.41}$$

which leads to the Klein-Gordon field equation,

$$\left(\Box - M^{2}\right)\varphi = 0. \tag{5.42}$$

Let π be the canonical momentum,

$$\pi \equiv \frac{\partial\mathcal{L}}{\partial\left(\partial_{t}\varphi\right)} = \partial_{t}\varphi, \tag{5.43}$$

subject to the equal-time (both evaluated at time t) commutation relation

$$\left[\varphi\left(\vec{x}\right), \pi\left(\vec{x}'\right)\right] = \imath\delta^{3}\left(\vec{x} - \vec{x}'\right), \tag{5.44}$$

where we use $\vec{x}$ for the spatial three-vectors. The Hamiltonian is

$$H = \frac{1}{2}\int\left[\pi^{2}\left(\vec{x}\right) + \left|\nabla\varphi\left(\vec{x}\right)\right|^{2} + M^{2}\varphi^{2}\left(\vec{x}\right)\right]d^{3}x. \tag{5.45}$$

5.3.4 Discretization and cutoff

The problem of the entropy of a field (which is defined at every point in continuous space) can be reduced to that of coupled oscillators, which we solved above, by approximating space as a grid of discrete points, and taking the value of the field at each one of them as the amplitude of an oscillator. Thus, taking the partial trace over some of the oscillators amounts to tracing out the field degrees of freedom at certain points, or in entire regions of space.

A straightforward way to implement this consists in discretizing all three dimensions, giving rise to a cubic lattice with uniform spacing a. This length naturally introduces an ultraviolet cutoff, since it is the smallest wavelength that can still be represented in the discretized system. For instance, when describing phonons in condensed matter, a would be the intermolecular spacing, because wavelengths shorter than this do not make sense. When dealing with quantum fields in spacetime, there is, as of now, no evidence of such a natural lower bound on the validity of the continuous treatment. There are, however, known limitations to the validity of today's theories, namely the Planck scale, which for length is of the order of $L_{Pl} \equiv \sqrt{\hbar G/c^{3}} \approx 10^{-35}m$. Regardless of whether waves shorter than L_{Pl} actually

exist, we are not yet able to make meaningful predictions about them, and should therefore exclude them from our model.

Besides the inherent limitations of our theories, there is also a simple technical reason for introducing an ultraviolet cutoff: the area law we expect to find requires one for dimensional reasons, as discussed in section 4.3. Without it, that is, in the limit where the cutoff goes to zero, the entropy diverges. Thus, one cannot have a physically meaningful entropy without introducing an ultraviolet cutoff.[25]

In order to reduce the problem to a finite number of grid points, one must also introduce an infrared cutoff. Letting L_x be the length (in the x dimension) of the region in which the field is defined, we introduce the total number of discrete points along that axis,

$$x_{tot} \equiv L_x/a. \tag{5.46}$$

Then the discretization can be summed up by the prescription

$$x \to ak \; k \in \mathbb{N}; \quad \int dx \to a \sum_{k=1}^{xtot}; \quad \frac{\partial f(x)}{\partial x} \to \frac{f_{k+1} - f_k}{a}, \tag{5.47}$$

and similarly for y and z, so that

$$\varphi(\vec{x}) \to \varphi_{klm}\,; \quad \pi(\vec{x}) \to \frac{1}{a}\pi_{klm}. \tag{5.48}$$

The commutation relation (5.44) becomes

$$[\varphi_{klm}, \pi_{k'l'm'}] = \imath\delta_{kk'}\delta_{ll'}\delta_{mm'}, \tag{5.49}$$

and the Hamiltonian (5.45) can be written as

$$\begin{aligned} H \; &= \frac{a}{2}\sum_{k=1}^{xtot}\sum_{l=1}^{ytot}\sum_{m=1}^{ztot}\Big[\pi_{klm}^2 + \left(\varphi_{(k+1)lm} - \varphi_{klm}\right)^2 + \left(\varphi_{k(l+1)m} - \varphi_{klm}\right)^2 \\ &\quad + \left(\varphi_{kl(m+1)} - \varphi_{klm}\right)^2 + a^2M^2\varphi_{klm}^2\Big]. \end{aligned} \tag{5.50}$$

Finally, we introduce a single index $j = j(k,l,m) = 1, \ldots, x_{tot}y_{tot}z_{tot}$ to label all lattice points sequentially. This formally reduces the problem to that of $x_{tot}y_{tot}z_{tot} \equiv N$ coupled harmonic oscillators, and the entropy after tracing out those with indices $j = 1, \ldots, n$ can be computed as described above. As mentioned in section 5.2.2, the spatial positions in the lattice can be labelled in an arbitrary order, so that one

[25] An alternative approach to the problem is to implement the cutoff directly in momentum space, or to assume the field to be continuous only up to a certain distance from the horizon. (The latter is implemented, for example, by Bombelli, Lee, Koul and Sorkin [6].) However, despite minor technical differences, the results are equivalent, and lead to the same conclusions.

has complete freedom to choose the geometry of the traced-out region, by deciding which sites are assigned $j \leq n$.

5.3.5 Spherical symmetry

Spherical symmetry is present in all static black hole spacetimes as well as in flat spacetime, and it can be exploited to reduce the computational cost of the calculation, besides revealing new physical features.

Partial waves To that end, we introduce spherical coordinates, (r, θ, ϕ), and decompose the field into partial waves[26]:

$$\varphi_{lm}(r) \equiv r \int d\Omega Z_{lm}(\theta, \phi)\, \varphi(\overrightarrow{x}) \tag{5.51}$$

$$\pi_{lm}(r) \equiv r \int d\Omega Z_{lm}(\theta, \phi)\, \pi(\overrightarrow{x}), \tag{5.52}$$

that is,

$$\varphi(\vec{x}) = \sum_{l,m} \frac{\varphi_{lm}(r)}{r} Z_{lm}(\theta, \phi) \tag{5.53}$$

$$\pi(\vec{x}) = \sum_{l,m} \frac{\pi_{lm}(r)}{r} Z_{lm}(\theta, \phi). \tag{5.54}$$

The factor r is introduced for dimensional reasons, so that the final expression for the Hamiltonian will be of the same form as in equation (5.15). Since the field is real, and in order to make φ_{lm} and π_{lm} real as well, we use the real spherical harmonics $Z_{lm}(\theta, \phi)$. In terms of the ordinary spherical harmonics, they are given by

$$Z_{lm} = \begin{cases} \sqrt{2} Re\left[Y_{lm}\right] & (m > 0) \\ Y_{l0} & (m = 0) \\ \sqrt{2} Im\left[Y_{lm}\right] & (m < 0) \end{cases}. \tag{5.55}$$

The factors $\sqrt{2}$ in this definition ensure orthonormality,

$$\int \left[Z_{lm} Z_{l'm'}\right] d\Omega = \delta_{ll'}\delta_{mm'}, \tag{5.56}$$

and that

$$\int \left[\left(\frac{\partial Z_{lm}}{\partial \theta}\right)\left(\frac{\partial Z_{l'm'}}{\partial \theta}\right) + \frac{1}{\sin^2\theta}\left(\frac{\partial Z_{lm}}{\partial \phi}\right)\left(\frac{\partial Z_{l'm'}}{\partial \phi}\right)\right] d\Omega = l(l+1)\,\delta_{ll'}\delta_{mm'}. \tag{5.57}$$

[26] When it is clear from the context that the indices l and m refer to partial waves, their range, $l = 0, 1, ...$ and $m = -l, -l+1, ..., l$, is suppressed for the sake of notational simplicity.

Considering that in spherical coordinates

$$|\nabla\varphi|^2 = \left(\frac{\partial\varphi}{\partial r}\right)^2 + \frac{1}{r^2}\left[\left(\frac{\partial\varphi}{\partial\theta}\right)^2 + \frac{1}{\sin^2\theta}\left(\frac{\partial\varphi}{\partial\phi}\right)^2\right], \tag{5.58}$$

the terms in the Hamiltonian (5.45) become, in terms of the Z_{lm},

$$\int\left[\pi^2\right]d\Omega r^2 dr = \sum_{l,m}\left[\int\left(\frac{\pi_{lm}(r)}{r}\right)^2 r^2 dr\right], \tag{5.59}$$

$$\int\left[\varphi^2\right]d\Omega r^2 dr = \sum_{l,m}\left[\int\left(\frac{\varphi_{lm}(r)}{r}\right)^2 r^2 dr\right], \tag{5.60}$$

$$\int\left[\left(\frac{\partial\varphi}{\partial r}\right)^2\right]d\Omega r^2 dr = \sum_{l,m}\left[\int\left(\frac{\partial}{\partial r}\frac{\varphi_{lm}(r)}{r}\right)^2 r^2 dr\right], \tag{5.61}$$

and

$$\begin{aligned}&\int\frac{1}{r^2}\left[\left(\frac{\partial\varphi}{\partial\theta}\right)^2 + \frac{1}{\sin^2\theta}\left(\frac{\partial\varphi}{\partial\phi}\right)^2\right]d\Omega r^2 dr\\ &= \sum_{l,m}\left[\int\frac{l(l+1)}{r^2}\left(\frac{\varphi_{lm}(r)}{r}\right)^2 r^2 dr\right].\end{aligned} \tag{5.62}$$

Furthermore, given the properties

$$\left[\varphi\left(\overrightarrow{x}\right),\pi\left(\overrightarrow{x}'\right)\right] = \imath\frac{1}{r^2}\delta(r-r')\frac{1}{\sin\theta}\delta(\theta-\theta')\delta(\phi-\phi') \tag{5.63}$$

(equation (5.44)) and

$$\int d\Omega'\frac{1}{\sin\theta}\delta(\theta-\theta')\delta(\phi-\phi')f(\theta',\phi') = f(\theta,\phi), \tag{5.64}$$

the orthonormality of the modes ensures the equal-time commutation relation

$$\begin{aligned}\left[\varphi_{lm}(r),\pi_{l'm'}(r')\right] &= rr'\int d\Omega d\Omega' Z_{lm}(\theta,\phi)Z_{l'm'}(\theta',\phi')\left[\varphi\left(\overrightarrow{x}\right),\pi\left(\overrightarrow{x}'\right)\right]\\ &= \imath\frac{rr'}{r^2}\delta(r-r')\int d\Omega Z_{lm}(\theta,\phi)Z_{l'm'}(\theta,\phi)\\ &= \imath\delta(r-r')\delta_{ll'}\delta_{mm'}.\end{aligned} \tag{5.65}$$

Thus, the Hamiltonian of the field can be decomposed into a sum,

$$H = \sum_{l,m} H_{lm}, \tag{5.66}$$

of the Hamiltonians H_{lm} of each partial wave,

$$H_{lm} = \frac{1}{2}\int_0^\infty \left[\pi_{lm}^2(r) + r^2\left(\frac{\partial}{\partial r}\frac{\varphi_{lm}(r)}{r}\right)^2 + \frac{l(l+1)}{r^2}\varphi_{lm}^2(r) + M^2\varphi_{lm}^2(r)\right]dr. \tag{5.67}$$

Discretization As before, we apply a discretization scheme, but only to the one remaining continuous variable, the radial coordinate. This implies that the number of grid points involved in the calculation increases linearly with the ratio of infrared to ultraviolet cutoff, allowing us to study situations with grid sizes of the order of 10^2, whereas with cubic symmetry it grows as $(L/a)^3$, making grids as small as $20 \times 20 \times 20$ intractable. Conversely, a calculation using spherical symmetry also involves a sum over partial waves, which we discuss below.

Letting a be the ultraviolet cutoff in the radial direction and $L \equiv n_{tot}a$ the infrared, we have the prescription

$$\begin{aligned} r \to aj; \quad \int dr \to a\sum_{j=1}^{ntot}; \quad g(r)\frac{\partial f(r)}{\partial r} \to g_{j+1/2}\frac{f_{j+1}-f_j}{a}; \\ \varphi_{lm}(r) \to \varphi_{lmj}; \quad \pi_{lm}(r) \to \tfrac{1}{a}\pi_{lmj}. \end{aligned} \tag{5.68}$$

Hence the Hamiltonian becomes

$$\begin{aligned} H_{lm} &= \frac{1}{2a}\sum_{j=1}^{ntot}\left[\pi_{lmj}^2 + \left(j+\frac{1}{2}\right)^2\left(\frac{\varphi_{lm(j+1)}}{j+1} - \frac{\varphi_{lmj}}{j}\right)^2\right. \\ &\quad \left. + \left(\frac{l(l+1)}{j^2} + a^2M^2\right)\varphi_{lmj}^2\right] \\ &= \frac{1}{a}\left[\frac{1}{2}\sum_{j=1}^{ntot}\pi_{lmj}^2 + \frac{1}{2}\sum_{i,j=1}^{ntot}\varphi_{lmi}\left(K_{lm}\right)_{ij}\varphi_{lmj}\right], \end{aligned} \tag{5.69}$$

and the commutation relations (based on equation (5.65)) read

$$[\varphi_{lmj}, \pi_{lmj'}] = \imath\delta_{jj'}. \tag{5.70}$$

Thus, we have formally reduced each mode lm to a system of coupled oscillators with a coupling matrix

$$\begin{aligned} \left(K_{lm}\right)_{ij} &= \delta_{ij}\left[2 + \frac{1}{2j^2} + \frac{l(l+1)}{j^2} + a^2M^2\right] \\ &\quad -\delta_{i(j+1)}\frac{(j+1/2)^2}{j(j+1)} - \delta_{(i+1)j}\frac{(i+1/2)^2}{i(i+1)}. \end{aligned} \tag{5.71}$$

When calculating the entanglement entropy of a solid sphere of radius R, we trace out the sites $j = 1, \ldots, n_{out}$, keeping $j = n_{out} + 1, \ldots, n_{tot}$ (or vice versa), with

$$n_{out} = R/a. \tag{5.72}$$

For simplicity, we will also refer to n_{out} as the radius of this sphere, though one should bear in mind that this is in units of the ultraviolet cutoff. It is also possible to trace out only a shell, with $j = n_{in} + 1, \ldots, n_{out}$, or several concentric shells. However, one is always restricted to spherically symmetric regions, whereas a cubic grid allows one to either trace out or keep each site independently.

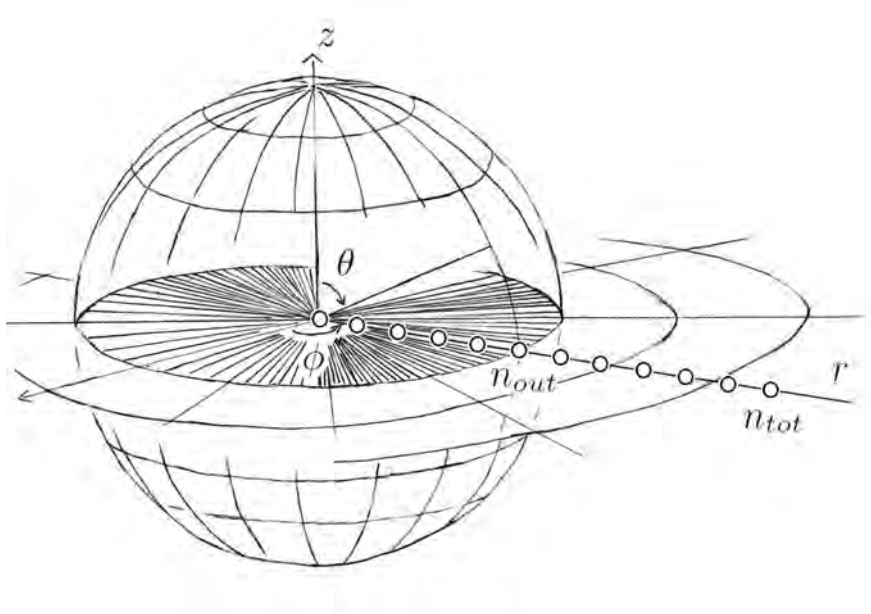

Figure 5.2: Illustration of the traced-out region in flat space with spherical symmetry: when the sites (represented by points) $j = 1, ...n_{out}$ in the radial dimension are traced out for every partial wave l, m, this amounts to eliminating all degrees of freedom inside a sphere (generated by the coordinates θ and ϕ). The plane of $\theta = \pi/2$, generated by $\phi \in [0, 2\pi]$ (angular coordinate) and $r \in [0, \infty[$ (radial), including the portion of it that is traced out (shaded), is highlighted, for comparison with the corresponding hypersurface in the Einstein universe (see figure 5.3).

Sum over partial waves Applying the technique from section 5.2 to the coupling matrix (5.71), we find the entropy due to each partial wave, S_{lm}. In order to ensure that these contributions are additive, we must prove that the reduced state of the whole field is a product of reduced states of the partial waves.

We know this to be the case for the ground state: the density matrix of the entire field is a product of the ground states of the partial waves,

$$\rho_0 = \prod_{l,m} (\rho_0)_{lm} \,. \tag{5.73}$$

Each $(\rho_0)_{lm}$ is a function of the values φ_{lmj} of a certain partial wave l, m at all the sites $j = 1, \ldots, n_{tot}$:

$$(\rho_0)_{lm} = (\rho_0)_{lm} \left(\varphi_{lm1}, ...\varphi_{lmn_{tot}}, \varphi'_{lm1}, ...\varphi'_{lmn_{tot}}\right) . \tag{5.74}$$

Taking the partial trace over an interval of the radial coordinate, that is, over certain values of j, implies the integration

$$\rho_{red} = \int \left[\prod_{l,m} \prod_{j} d\varphi_{lmj}\right] \rho_0. \tag{5.75}$$

However, each of the φ_{lmj} over which we integrate appears only in the arguments of the corresponding $(\rho_0)_{lm}$, so that we can write

$$\rho_{red} = \prod_{l,m} \left[\int \prod_{j} d\varphi_{lmj} \, (\rho_0)_{lm}\right] , \tag{5.76}$$

and the reduced state is indeed also a product:

$$\rho_{red} = \prod_{l,m} (\rho_{red})_{lm} \,. \tag{5.77}$$

The S_{lm} are therefore additive:

$$S = \sum_{l,m} S_{lm}. \tag{5.78}$$

In fact, the index m does not appear in the coupling matrix (5.71), so the entropy of a partial wave does not depend on it:

$$S_{lm} = S_{l0} \quad \forall m, \tag{5.79}$$

and the total entanglement entropy becomes

$$S = \sum_{l=0}^{\infty} (2l+1) \, S_{l0} \equiv \sum_{l=0}^{\infty} S_l. \tag{5.80}$$

5.3.6 Curved background

Since we intend to study entanglement entropy in the spacetime of black holes eventually, it is natural to introduce curvature to the background. However, there

is also another reason: As we will discuss in section 6.2, the results from cubic and spherical grids in flat space show that the dominant contribution to the entropy is due to interactions whose range is of the order of the cutoff. The effect is therefore insensitive to the large-scale geometry of the traced-out region, at least to differences in its shape. What we intend to verify is whether the same holds for intrinsic curvature.

For this purpose, one does not need an acutal black hole spacetime. Instead, we choose a simple generalization of Minkowski spacetime, in which only the spatial geometry is different: the Einstein universe[27], whose metric is given by

$$ds^2 = -dt^2 + R^2 \left[d\alpha^2 + \sin^2 \alpha \left(d\theta^2 + \sin^2 \theta d\phi^2 \right) \right], \tag{5.81}$$

where

$$\alpha \in [0, \pi] \; ; \; \theta \in [0, \pi] \; ; \; \phi \in [0, 2\pi[, \tag{5.82}$$

and R is a constant. The angular variables θ and ϕ generate a two-sphere, as they do in flat space, so that the decomposition into partial waves is still applicable. The difference lies in the surface area of those spheres: instead of $4\pi (R\alpha)^2$, the surface with fixed radius $r = R\alpha$ has an area $A(r) = 4\pi R^2 \sin^2 \alpha$ in this spacetime. If an area law continues to hold, we expect the entropy to be proportional to the latter (divided by the square of the ultraviolet cutoff).

The discretization in the radial dimension is also altered: in this spacetime, the greatest possible distance from the origin is finite, given by $R\pi$. Let this interval be covered by n_{tot} discrete points, implying a spacing (ultraviolet cutoff) $a = R\pi/n_{tot}$. The traced-out region, on the other hand, extends from the origin to a site n_{out}, so its radius is $r_{out} = R\pi (n_{out}/n_{tot})$ and the area, $A_{out} = 4\pi R^2 \sin^2 [\pi (n_{out}/n_{tot})]$. Thus, the variable we expect to find in an area law is

$$\frac{A_{out}}{a^2} = \frac{4}{\pi} n_{tot}^2 \sin^2 \left[\pi \frac{n_{out}}{n_{tot}} \right]. \tag{5.83}$$

Since this spacetime differs from the Minkowski universe only in the spatial geometry, the statements leading up to equation (5.45) require only minor modifications. The Hamiltonian for a scalar field in the Einstein universe becomes

$$\begin{aligned} H \; &= \frac{1}{2} \int R^3 \sin^2 \alpha d\alpha d\Omega \left[\pi^2 (\overrightarrow{x}) + \frac{1}{R^2} \left(\frac{\partial \varphi}{\partial \alpha} \right)^2 \right. \\ &\left. + \frac{1}{R^2 \sin^2 \alpha} \left[\left(\frac{\partial \varphi}{\partial \theta} \right)^2 + \frac{1}{\sin^2 \theta} \left(\frac{\partial \varphi}{\partial \phi} \right)^2 \right] + M^2 \varphi^2 (\vec{x}) \right]. \end{aligned} \tag{5.84}$$

[27] This metric is a special case of Friedman-Robertson-Walker solutions, with constant scale factor and Gaussian curvature $k = +1$.

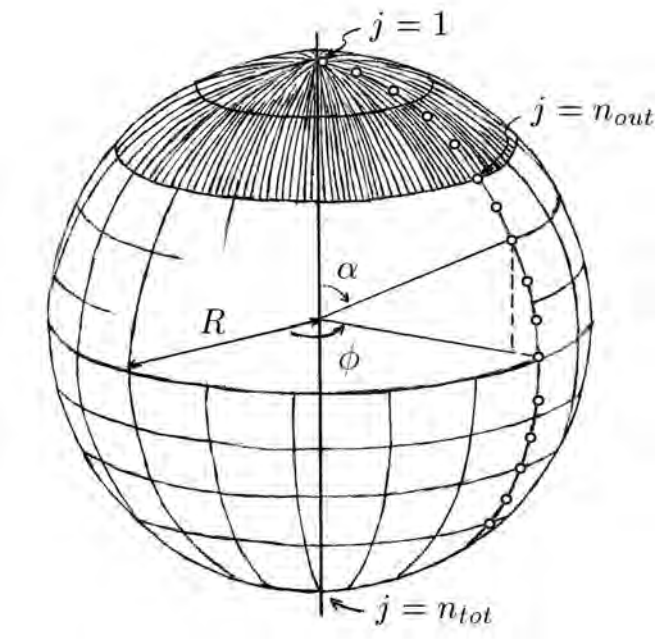

Figure 5.3: In the Einstein universe, the hypersurface with fixed angular coordinate $\theta = \pi/2$ is a two-sphere, generated by the coordinates $\phi \in [0, 2\pi]$ (angular) and $r = R\alpha \in [0, R\pi]$ (radial), instead of the infinite plane one finds in flat space (see figure 5.2). The finite interval spanned by the radial coordinate is divided into n_{tot} discrete points. The first n_{out} sites (shaded region) are then traced out, in order to calculate the entanglement entropy.

Partial waves Here, let

$$\varphi_{lm}(\alpha) \equiv R \sin\alpha \int d\Omega Z_{lm}(\theta, \phi)\, \varphi(\overrightarrow{x}) \tag{5.85}$$

$$\pi_{lm}(\alpha) \equiv R \sin\alpha \int d\Omega Z_{lm}(\theta, \phi)\, \pi(\overrightarrow{x}), \tag{5.86}$$

so that

$$\int \left[\pi^2\right] d\Omega R^3 \sin^2\alpha d\alpha = \sum_{l,m} \left[\int \pi_{lm}^2(\alpha)\, R d\alpha\right], \tag{5.87}$$

$$\int \left[\varphi^2\right] d\Omega R^3 \sin^2\alpha d\alpha = \sum_{l,m} \left[\int \varphi_{lm}^2(\alpha)\, R d\alpha\right], \tag{5.88}$$

$$\int \left[\frac{1}{R^2}\left(\frac{\partial\varphi}{\partial\alpha}\right)^2\right] d\Omega R^3 \sin^2\alpha d\alpha = \sum_{l,m} \left[\int \sin^2\alpha \left(\frac{\partial}{\partial\alpha}\frac{\varphi_{lm}(\alpha)}{R\sin\alpha}\right)^2 R d\alpha\right] \tag{5.89}$$

and

$$\begin{aligned} &\int \tfrac{1}{R^2\sin^2\alpha}\left[\left(\tfrac{\partial\varphi}{\partial\theta}\right)^2 + \tfrac{1}{\sin^2\theta}\left(\tfrac{\partial\varphi}{\partial\phi}\right)^2\right] d\Omega R^3 \sin^2\alpha d\alpha \\ &\quad = \textstyle\sum_{l,m} l(l+1)\left[\int \left(\frac{\varphi_{lm}(\alpha)}{R\sin\alpha}\right)^2 R d\alpha\right]. \end{aligned} \tag{5.90}$$

As in flat space, the Hamiltonian can be written as a sum,

$$H = \sum_{l,m} H_{lm}, \tag{5.91}$$

of the H_{lm} that govern each mode,

$$H_{lm} = \frac{R}{2} \int \left[\pi_{lm}^2(\alpha) + \sin^2\alpha \left(\frac{\partial}{\partial\alpha} \frac{\varphi_{lm}(\alpha)}{R \sin\alpha} \right)^2 + \left(\frac{l(l+1)}{R^2 \sin^2\alpha} + M^2 \right) \varphi_{lm}^2(\alpha) \right] d\alpha. \tag{5.92}$$

Discretization The appropriate prescription for the Einstein universe is

$$\begin{aligned} R \to a\tfrac{n_{tot}}{\pi} \; ; \quad \alpha \to \tfrac{aj}{R} = \tfrac{\pi}{n_{tot}} j \; ; \quad \textstyle\int d\alpha \to \tfrac{\pi}{n_{tot}} \sum_{j=1}^{ntot} \; ; \\ g(\alpha) \tfrac{\partial f(\alpha)}{\partial \alpha} \to g_{j+1/2} \tfrac{n_{tot}}{\pi} (f_{j+1} - f_j) \; ; \\ \varphi_{lm}(\alpha) \to \varphi_{lmj} \; ; \quad \pi_{lm}(\alpha) \to \tfrac{1}{a} \pi_{lmj}, \end{aligned} \tag{5.93}$$

which gives

$$\begin{aligned} H_{lm} \; = \frac{1}{2a} \sum_{j=1}^{ntot} \Bigg[& \pi_{lmj}^2 + \left(\left(\frac{\pi}{n_{tot}} \right)^2 \frac{l(l+1)}{\sin^2\left(\frac{\pi}{n_{tot}} j\right)} + a^2 M^2 \right) \varphi_{lmj}^2 \\ & + \sin^2\left(\frac{\pi}{n_{tot}} (j+1/2) \right) \left(\frac{\varphi_{lm(j+1)}}{\sin\left(\frac{\pi}{n_{tot}}(j+1)\right)} - \frac{\varphi_{lmj}}{\sin\left(\frac{\pi}{n_{tot}} j\right)} \right)^2 \Bigg]. \end{aligned} \tag{5.94}$$

We have again reduced the problem to that of coupled harmonic oscillators with coupling matrix

$$\begin{aligned} (K_{lm})_{ij} \; = \delta_{ij} \Bigg[& \left(\frac{\pi}{n_{tot}} \right)^2 \frac{l(l+1)}{\sin^2\left(\frac{\pi}{n_{tot}} j\right)} + a^2 M^2 \\ & + \left[\sin^2\left(\frac{\pi}{n_{tot}} (j+1/2) \right) + \sin^2\left(\frac{\pi}{n_{tot}} (j-1/2) \right) \right] \frac{1}{\sin^2\left(\frac{\pi}{n_{tot}} j\right)} \Bigg] \\ & - \delta_{i(j+1)} \frac{\sin^2\left(\frac{\pi}{n_{tot}}(j+1/2)\right)}{\sin\left(\frac{\pi}{n_{tot}}(j+1)\right) \sin\left(\frac{\pi}{n_{tot}} j\right)} \\ & - \delta_{j(i+1)} \frac{\sin^2\left(\frac{\pi}{n_{tot}}(i+1/2)\right)}{\sin\left(\frac{\pi}{n_{tot}}(i+1)\right) \sin\left(\frac{\pi}{n_{tot}} i\right)}. \end{aligned} \tag{5.95}$$

The difference between this K and the corresponding matrix in flat space (equation (5.71)) lies only in the factors $\sin\left(\frac{\pi}{n_{tot}}j\right)$: replacing them by j, one recovers the previous expressions. Since $\sin\left(\frac{\pi}{n_{tot}}n_{tot}\right) = 0$, the elements of K with i or $j = n_{tot}$ must be left out in the calculation. However, we verified that small changes, such as including or excluding a single grid point, do not affect the results of the calculations significantly.

5.4 Numerical calculation

The calculations detailed in the previous section were implemented[28] in Mathematica®, to compute the entanglement entropy numerically in a number of situations. In the following, we describe these setups, in particular introducing the variables that characterize them, and discuss some points that have to be addressed before any meaningful results about the entropy itself can be obtained.

The first test concerns the robustness of our calculation against small changes in the discretization prescription, such as replacing x with $a(j+1)$ instead of aj, and shows that the consequent changes are negligible ($<1\%$). For each of the setups, we also verify the symmetry theorem from section 2.3, which states that (a) tracing out a region Ω and calculating the entropy of the reduced state of the remainder, $\bar{\Omega}$, and (b) keeping Ω and tracing out $\bar{\Omega}$ give the same value for the entanglement entropy. In particular, in the Einstein universe, this fact allows us to restrict ourselves to the range $n_{out} \leq n_{tot}/2$, since a given n_{out} and another $\overline{n}_{out} = n_{tot} - n_{out}$ represent the same division of the system, therefore resulting in the same entropy.

Once we obtain numerical values for the entropy and verify that it obeys an area law, we also estimate the accuracy of the value found for the coefficient κ of that law, by comparing the results found for different values of the parameters. Although this analysis is not sufficient to determine the error with any precision, it does show that κ is at least an order of magnitude larger than its error, ensuring meaningful results.

5.4.1 Flat space with cubic symmetry

We begin with the setup illustrated in figure 5.4: a finite region of flat three-dimensional space that is covered by a cubic grid, consisting of x_{tot} by y_{tot} by z_{tot} sites with spacing a in each dimension. The traced-out region Ω is a smaller rectangular parallelepiped that encompasses x_{ext} by y_{ext} by z_{ext} sites, with one vertex at $(x_{min}, y_{min}, z_{min})$. We then introduce a hole in Ω, with sides x_{iext}, y_{iext} and z_{iext} and one vertex at $(x_{imin}, y_{imin}, z_{imin})$, so that the trace is only taken over a hollow shell, while keeping both its interior and the exterior. As remarked at the end

[28] Two examples of the code used in these simulations can be found in appendix B.

of section 5.3.4, a cubic grid gives us complete freedom to choose which sites are traced out. This allows us to test the robustness of our results under changes of the geometry of the traced-out region, for instance by adding a tunnel through the shell, which connects the inside and the outside.

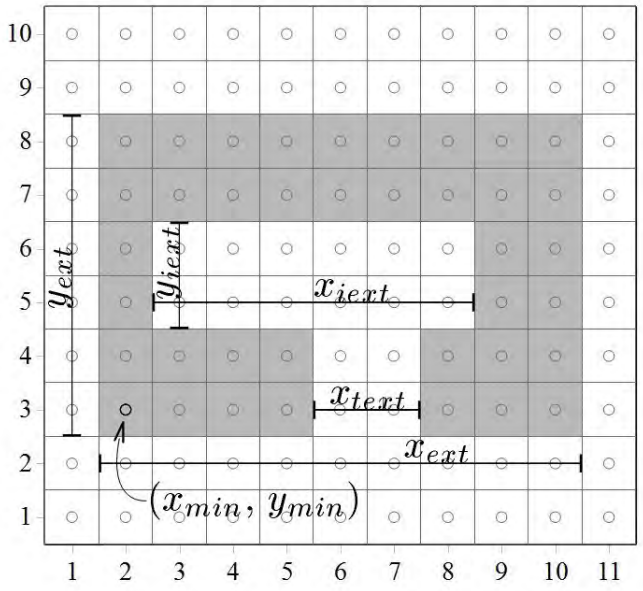

Figure 5.4: Illustration of the parameters that describe the traced-out region on the cubic lattice. (For simplicity, the z dimension is suppressed.) The grid consists of $x_{tot} = 11$ by $y_{tot} = 10$ sites (represented by circles). The traced-out region (shaded) is an x_{ext} by y_{ext} rectangle, located with one corner at x_{min}, y_{min}, in which there is a hole (x_{iext} by y_{iext}, aligned at x_{imin}, y_{imin}). Furthermore, a section of width x_{text} is removed from the fixed-y wall, so that a tunnel connects the inside and the outside.

The first parameter to be determined when planning the setup is the size of the grid, $x_{tot}y_{tot}z_{tot}$: because of its impact on the computational cost[29], it should be as small as it can be without affecting the result. Fortunately, the data (for example figure 6.4(left)) show that the entropy of a fixed region Ω only changes perceptibly ($>0.2\%$) when these parameters become so small that the horizon coincides exactly with the boundary of the grid. This effect and its implications are discussed in more detail in section 6.2, but for the purpose of determining the parameters, it is sufficient to note that we must only choose the size of the grid and the position of Ω such that there is always at least one site between their surfaces.

[29]The main concern is the duration of the calculation, but it is not the only one: the absolute limit in this simulation was a $20 \times 20 \times 20$ grid, which was found to exceed the memory of the software.

5.4.2 Spherical symmetry

In section 5.3.5, we introduce the parameters n_{tot}, n_{out} and n_{in}, which characterize the grid and the spherical shell which is traced out (see figure 5.2), and derive a technique for calculating the contribution S_{lm} $(= S_{l0} \ \forall m)$ of each partial wave l, m to the total entropy,

$$S = \sum_{l=0}^{\infty} (2l+1)\, S_{l0} \equiv \sum_{l=0}^{\infty} S_l. \tag{5.96}$$

Again, the first variable to be chosen is the number of grid points (in this case only in the radial dimension), n_{tot}. Comparing the S_l obtained for fixed l, n_{in} and n_{out}, but with different n_{tot} (ranging from $n_{tot} = n_{out}$, of the order of 10, to several hundred), one finds again that the results are mostly independent of n_{tot} (variations of less than 1%). In this case, however, boundary effects appear already when the distance between the outer surface of the traced-out region, n_{out}, and the boundary of the grid, n_{tot}, decreases to approximately three grid points. This, too, is discussed further in section 6.2. Until that point, we take $n_{tot} = 1.2 n_{out}$ in order to avoid n_{tot}-dependence without incurring excessive computational cost.

However, the cost is also increased by another fact: in principle, there are infinitely many contributions S_l to be calculated. In practice, the series must be truncated at some point, which is why we now turn to the behaviour of the S_l as a function of the index l, to determine whether the series will converge. Figure 5.5 shows that the S_l increase with growing l at first, up to some dominant contribution S_{ldom} at l_{dom}. From that point on, S_l falls off monotonically, reaching half its maximum value at roughly $3.5 l_{dom}$.

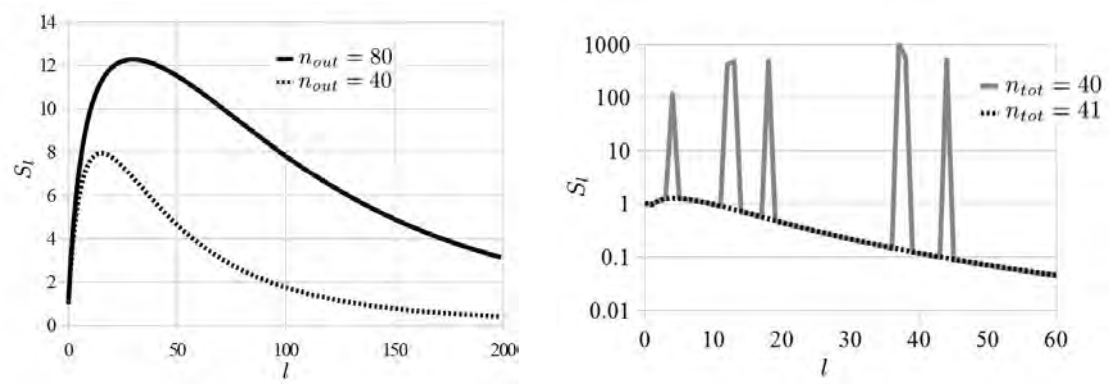

Figure 5.5: Entropy contributions S_l due to partial waves as a function of the index l: (left) in flat space, for spherical shells with $n_{in} = 20$ and different outer radii n_{out}, and (right) in curved space, for a solid sphere with $n_{out} = 10$, but different n_{tot}. In the Einstein universe, there are considerable deviations from the smooth curve, but they disappear when one uses odd values of n_{tot}.

This general behaviour does not change for different values of the parameters; only the values l_{dom} and S_{ldom} increase with n_{in} and n_{out}. However, we have not conducted further inquiries into this phenomenon, since we consider other facets of the problem more promising in terms of physical insights.

The fact that the S_l go to zero for large l suggests that the series may converge, and figure 5.6 supports this. We choose to truncate the series when it drops to a certain fraction of its maximum value: at the first $l > l_{dom}$ such that $S_l < t \cdot S_{ldom}$. When the parameter is $t = 0.03$, the entropy found in this way lies consistently about 5% below S^{exact}, the value extrapolated for $t \to 0$. In general, S, the partial sum of the series when truncated with a given t, reaches (approximately) the same fraction of S^{exact} for all values of n_{out} and n_{in}. This is the reason for choosing this truncation criterion: it allows us to track how the entropy changes when other parameters are altered without having to compute it to high precision in every case.

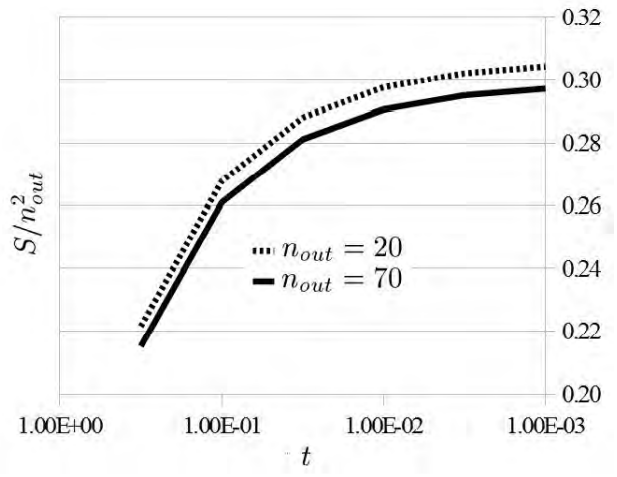

Figure 5.6: Convergence of the value found for S as a function of the truncation parameter t, for two solid spheres with different n_{out}. (The entropy is divided by n_{out}^2 to allow for an easier comparison.)

In a curved background, despite the additional factors in the coupling matrix (5.95), the calculations lead to the same spectra of S_l as a function of l, and the same truncation criterion applies. A notable difference arises for certain even numbers of grid points, n_{tot}: at several values of l, the contributions S_l suddenly spike several orders of magnitude above the rest of the curve, as shown in figure 5.5(right). However, the phenomenon disappears when n_{tot} is changed to the next odd value. This suggests that the deviations are most likely artifacts, and allows us to circumvent the problem entirely by simply avoiding even values of n_{tot}.

6 Behaviour of the entropy

6.1 Area law

The most important result of our calculations is that the entanglement entropy indeed obeys an area law,

$$S = \kappa \frac{A}{a^2}, \tag{6.1}$$

where A is the boundary area of the region that is being traced out, divided by the square of the ultraviolet cutoff for dimensional reasons. We have also found the numerical value of the proportionality constant κ, as detailed in the following.

6.1.1 Numerical results

Solid sphere in flat space The classic setup is flat space with spherical symmetry, permeated by a massless scalar field. A solid sphere of radius $R = n_{out}a$ and, consequently, surface area $A/a^2 = 4\pi n_{out}^2$ is traced out, as shown in figure 5.2. This situation was studied numerically by Srednicki [7] and he found an area law with the coefficient κ given by[30]

$$\kappa_{sph} \approx 0.024. \tag{6.2}$$

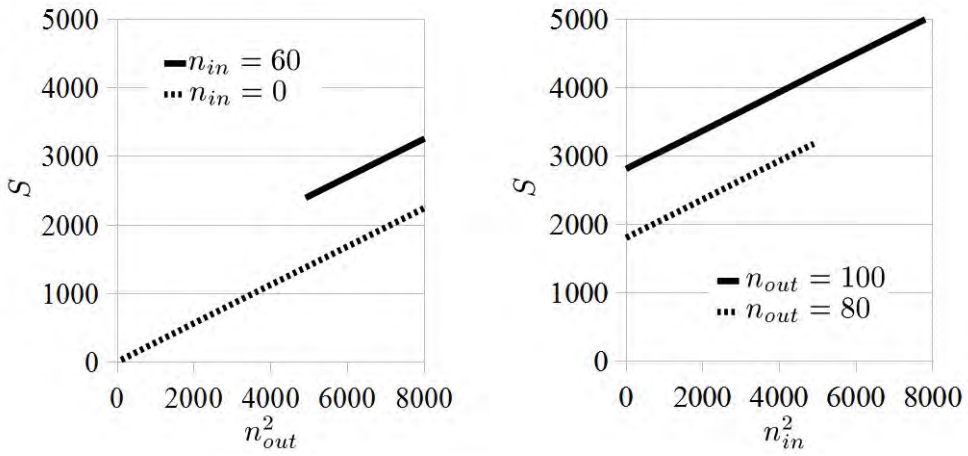

Figure 6.1: Entropy of a spherical shell; (left) as a function of the outer radius squared, for fixed inner radius, and (right) vice versa.

Spherical shell in flat space As a first step, we repeated that calculation, then traced over a hollow spherical shell of finite thickness, characterized by inner and outer radii n_{in} and n_{out}, respectively. Figure 6.1 shows that S is a linear function of n_{out}^2 and n_{in}^2, and the coefficient in both cases is the same. In terms of $A_{out}/a^2 =$

[30] The author does not provide an estimate of the error of this value.

$4\pi n_{out}^2$ and $A_{in}/a^2 = 4\pi n_{in}^2$, we recover an area law,

$$S_{shell} = \kappa_{shell} \frac{A_{out} + A_{in}}{a^2}, \tag{6.3}$$

with[31]

$$\kappa_{shell} \approx 0.022. \tag{6.4}$$

This κ also holds for the solid sphere, which is the limiting case of $n_{in} = 0$, and our results are in good agreement with those of Srednicki.

Sphere in curved space Using the coupling matrix (5.95) in similar calculations, one finds the entropy of a spherical region in curved space. Although the area of the boundary is now given by expression (5.83), the entropy is still proportional to it, as evidenced by figure 6.2. One can isolate the coefficient

$$\kappa_{curv} \approx 0.0231. \tag{6.5}$$

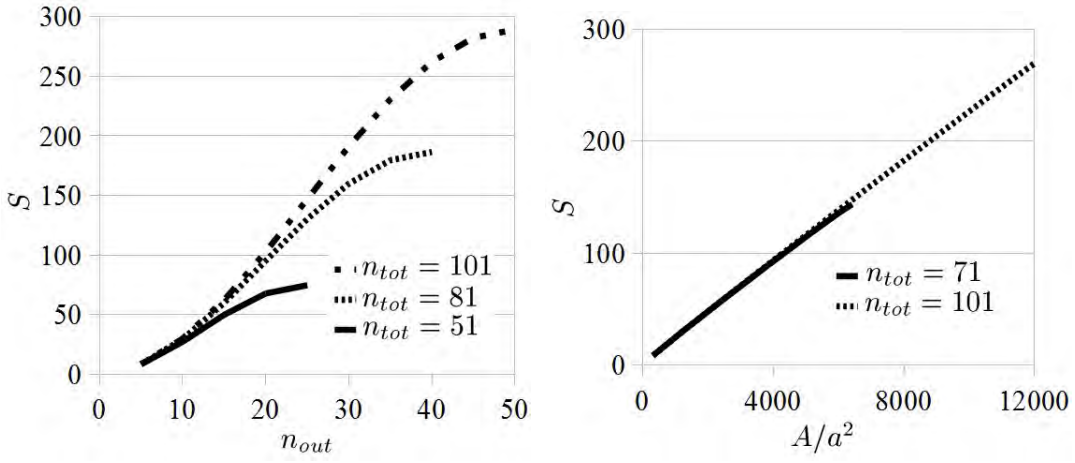

Figure 6.2: Entanglement entropy of a spherical region in Einstein universes with different curvature radii n_{tot}; as a function of (left) the radius n_{out} of the inaccessible region, and (right) its boundary area, given by equation (5.83). We stress that a given boundary area always corresponds to the same entropy, regardless of the curvature of the background.

We stress that the entropy depends solely on the boundary area of the region that we ignore, but not on its interior. Indeed, the latter can change depending on the curvature of the background space, yet the entropy of spheres with the same surface area is always the same. This is especially interesting in the context of the dimensionality argument presented in section 4.3: although the Einstein universe is

[31] As noted in section 5.4, the error of this coefficient for different setups could only be estimated very roughly, but it was found to be always at least an order of magnitude smaller than κ.

characterized by a third constant with the dimension of length, namely its curvature radius R, the entropy of the black hole still depends only on the combination A/a^2, its horizon area divided by the square of the ultraviolet cutoff of the theory.

Cubic grid When one traces out a solid parallelepiped on a cubic grid, the entropy grows linearly when one increases the length of its side in the x direction, x_{ext}. The same holds for y_{ext} and z_{ext}, as should be expected based on the symmetry under an exchange of x, y and z. The most general expression for the entropy that is linear in each parameter separately is

$$\begin{aligned} S = \quad & S_0 + b\,(x_{ext} + y_{ext} + z_{ext}) \\ & + c\,(x_{ext}y_{ext} + x_{ext}z_{ext} + y_{ext}z_{ext}) + d\,(x_{ext}y_{ext}z_{ext})\,. \end{aligned} \tag{6.6}$$

The entropy is zero when none of the sites are traced out, that is, $S_0 = 0$. For the remaining coefficients, the data indicate $c \approx 0.031$, $b \approx -0.007$ and $d \approx 0.002$. Since the surface area of the solid parallelepiped is $A_{out} = 2a^2\,(x_{ext}y_{ext} + x_{ext}z_{ext} + y_{ext}z_{ext})$, this reflects, to good approximation, an area law, with

$$\kappa_{paral} \approx 0.016. \tag{6.7}$$

When one adds a $x_{iext} \times y_{iext} \times z_{iext}$ hole in the interior of a parallelepiped (as illustrated in figure 5.4), while keeping its outer dimensions fixed, there arises a second interface between the traced-out region, Ω, and its complement, $\bar{\Omega}$. This is reflected by a second contribution to the entropy, which increases linearly with the dimensions of the hole. Thus, this S is also of the form (6.6), in terms of x_{iext}, y_{iext} and z_{iext}. The corresponding S_0 is given by the entropy when the size of the hole is zero, that is, of the solid parallelepiped. Then the data indicate $c \approx 0.032$, while $b < 10^{-2}$ and $d < 10^{-3}$. Besides being much smaller than c, the values of b and d vary considerably between setups (for instance, when one changes the size of the parallelepiped in which the hole is placed), even alternating between positive to negative. For this reason, we consider them artifacts, and assume $b = d = 0$.

This leads to a more general area law, contemplating both internal and external surfaces,

$$A = A_{out} + A_{in}, \tag{6.8}$$

and both with the same coefficient:

$$\kappa_{cub} \approx 0.016. \tag{6.9}$$

The area law holds for the entire interface between Ω and $\bar{\Omega}$, regardless its shape. This remains true when one adds a tunnel connecting the exterior and the interior of the hollow shell, as illustrated in figure 5.4. Once one takes into account all the changes in surface area that this modification causes, one finds that their sum is proportional to the change in entropy. The coefficient is ≈ 0.015, compatible with the value found above for the area law on a cubic grid.

6.1.2 Value of κ

Before discussing physical implications of the coefficients found above, we note that the results for cubic and spherical symmetry are slightly, yet perceptibly different. This can be attributed to the difference between the discretization schemes, as illustrated by the following scenario (see figure 6.3):

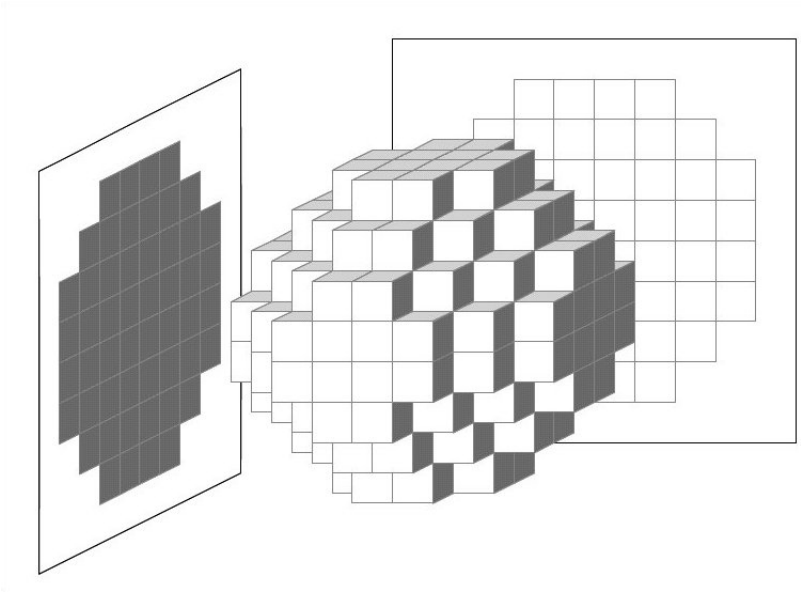

Figure 6.3: When an (approximately) spherical region is constructed out of cubic blocks, its effective surface area is larger than the equivalent smooth surface. Projecting all the unit squares with the same orientation onto a single plane shows that, in terms of the radius r, one has $6\pi r^2$ (six circles) instead of $4\pi r^2$ (sphere).

Suppose one were to trace out all the unit blocks on a cubic grid that lie inside a sphere of radius r. Since the faces of these blocks lie in the xy, yz or xz planes, instead of being parallel to the surface of the sphere, the actual surface area of the traced-out region would be approximately $6\pi r^2$. On the other hand, if one were to take the arguably more realistic approach of spherically symmetric discretization, the area would be only $4\pi r^2$. Yet the entropy of a given region should be independent of the discretization scheme, therefore the coefficient for cubic symmetry must be correspondingly smaller. The expected numerical factor of roughly $2/3$ is in good agreement with the data presented above.

That said, we can establish from our results that, if A is the area of a surface that is smooth on a length scale set by a, then κ in equation (6.1) is ≈ 0.02. However, expression (4.21), for the Bekenstein-Hawking entropy as predicted by black hole thermodynamics, appears to suggest that it should be $1/4$, almost an order of magnitude larger.

There have been attempts to explain this discrepancy as being related to the number of fields found in nature: if each non-interacting field (neglecting, for now, its spin) adds a contribution of $\kappa A/a^2$ to the entropy of a black hole, then one could reach the desired $1/4$ by summing over about a dozen fields. This point of view also suggests other possibilities, such as restricting the number of fields that can exist – even if they are not "accessible" via interactions with the known fields – by demanding that the total entropy added up in this manner not exceed the Bekenstein-Hawking value. In general, the implications of the existence of different fields (and different "species" of particles associated with them) for the problem of black hole entropy are studied under the name of *species puzzle*.

However, even if one considers only one field, the discrepancy between the κ found here and the expected $1/4$ can be accounted for: as discussed in section 5.3.4, the ultraviolet cutoff a of our model, representing the shortest wavelength it can describe, must only be of the order of the Planck length, whereas the factor $1/L_{Pl}^2$ in the Bekenstein-Hawking formula arises from an exact derivation. Therefore, we can simply choose a such that $\kappa/a^2 = 1/4L_{Pl}^2$, so that a single field generates the Bekenstein-Hawking entropy. If N fields are taken into account, the requirement becomes

$$N\frac{\kappa}{a^2} = \frac{1}{4}\frac{1}{L_{Pl}^2}, \tag{6.10}$$

which is perfectly compatible with our results, as long as $a/L_{Pl}\sqrt{N}$ has the right value. However, a better understanding of why the ultraviolet cutoff has this value may well require full quantum gravity.

6.1.3 Corrections to the area law

For the sake of completeness, we must mention at least two classes of situations in which corrections to the area law do arise: higher spin fields and excited states. However, neither of these fall within the scope of this work, and we refer the interested reader to other publications about the precise nature and possible interpretations of these corrections, such as [18] and [19].

6.2 Boundary effects

In section 5.4, we note that the entropy is mostly independent of the size of the grid on which the traced-out region is defined. Beyond its relevance for fixing the parameters for the numerical calculation, this fact also has physical implications: it shows that the entropy does not depend on the infrared cutoff of the model, the longest wavelength that can be represented on the grid. Srednicki [7] had already obtained some results in this direction for a solid sphere, but he restricted himself to sets of values such that $n_{tot} > 2n_{out}$. His reason for not using larger n_{out} was that the area law must be violated sooner or later as $n_{tot} \to n_{out}$, since, when the two are equal, the entire field is traced out, and the entropy drops to zero.

In our own calculations, we explore this transition. The traced-out region is maintained unchanged, so that the entropy should be constant. Then the size of the grid is decreased, eliminating one point at a time, until the frontier of the grid reaches the horizon. The resulting changes in the entropy are tracked, because one expects a drop in S upon eliminating degrees of freedom that are responsible for the entropy. Thus, studying the boundary effects should provide insights into the localization of the entanglement entropy.

6.2.1 Numerical results

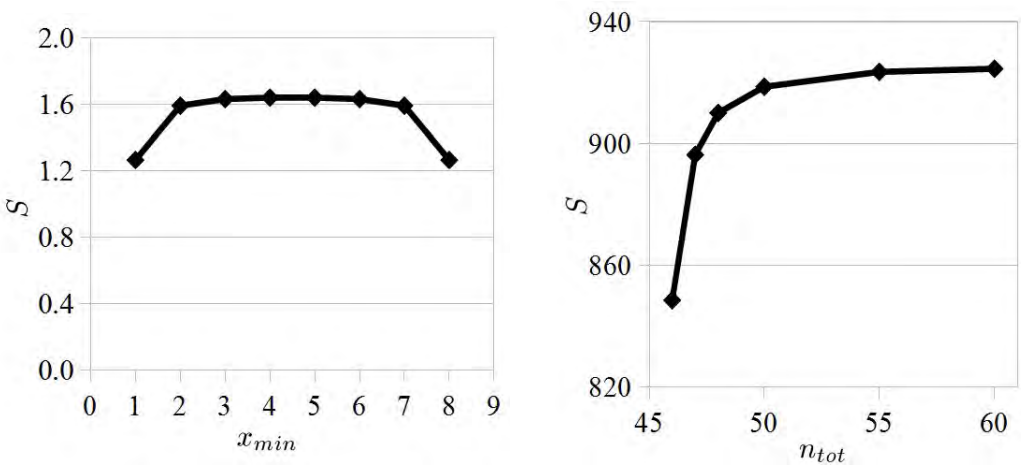

Figure 6.4: Boundary effects (left) on a cubic grid, in the x dimension: a solid parallelepiped with extension $x_{ext} = 3$ is moved to different positions x_{min} on a grid with $x_{tot} = 10$ sites, until it touches the borders. (right) with spherical symmetry: decreasing the grid size (n_{tot}) around a sphere of fixed outer radius $n_{out} = 45$ and inner radius $n_{in} = 35$ causes a drop in entropy when n_{tot} approaches n_{out}.

The results are most striking on a cubic grid: when a parallelepiped of fixed extension x_{ext}[32] is moved to different positions x_{min}, its entropy does not change even when

[32]The statements are made for the x dimension for the sake of simplicity, but they carry over naturally to y and z.

the boundaries of the traced-out region and the grid are adjacent ($x_{min} - 1 = 1$ or $x_{tot} - (x_{min} + x_{ext} - 1) = 1$). Only when the two coincide is there a drop in the entropy, as shown in figure 6.4(left).

This can be explained simply in terms of the area law: since one face of the parallelepiped no longer lies inside the grid (see figure 6.5), it does not contribute to the entropy. Indeed, the drop in entropy is proportional to the decrease in area, with the coefficient given exactly by the κ_{cub} of the area law found above.

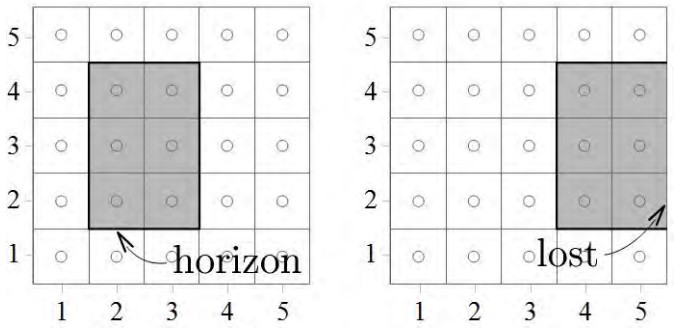

Figure 6.5: When the traced-out region is placed such that its surface, the horizon, lies at the boundary of the grid , one of its sides is unaccounted for, and the entropy decreases accordingly.

With spherical symmetry, both in flat and in curved space, the transition occurs gradually over approximately three sites (figure 6.4). A similar result was found by Das et al. [18], although they investigated the phenomenon using a slightly different approach: still for a ground-state scalar field decomposed into spherical waves, they modified the coupling matrix (5.71), manually "turning off" the interactions outside a given window. When this happens within approximately 3 grid points on either side of the horizon, Das et al. found a marked drop in the entropy. It is not clear why the boundary effect is more "washed out" in spherical symmetry, but one should expect the transition to become sharper as more terms of the S_l series are taken into account, since in the limit of summing over all l, both the cubic grid and the partial waves approach must lead to the same results.

One might suspect that this boundary effect is merely an artifact, due to our choice to approximate the field by an array of coupled discrete oscillators. Indeed, the prescription (5.47) for derivatives is the reason why the discretized Hamiltonian (5.50) contains only next-neighbour couplings[33], which in turn seems like a satisfactory explanation the occurrence of a boundary effect on a length scale that is given by the distance between these neighbours. However, the same phenomenon is also observed when the entropy is calculated in continuous space, where the ultraviolet

[33]The same holds for the corresponding equations for spherical symmetry.

cutoff appears as the limit of integration [6]. This additional evidence makes it more likely that this is a legitimate physical effect.

6.2.2 Localization of the entropy and small wavelength dominance

The above results clearly indicate that the entropy is localized near the horizon. There are several equivalent ways of stating this fact: for one, that short-distance correlations (specifically entanglement), between degrees of freedom just inside and just outside the horizon, are responsible for the appearance of the entropy. These correlations are due to short-range interactions, mediated by small wavelengths, which in turn correspond to high-energy modes. Any of these points of view may be useful in interpreting the results.

There are also a number of interesting implications to this insight. One of them concerns the area-dependence of S: if the entropy is due to short-distance interactions, at wavelengths much smaller than the length scale of the entire horizon, then each element of area dA must give the same contribution $dS = \kappa dA$ to entropy, without any effects of the large-scale geometry. This can account for the fact that the total entropy grows linearly with the area. The same effect also explains why the infrared cutoff and the curved background of the Einstein universe do not affect S: the entropy is due to interactions whose length scale is much smaller than the total size of the grid and the radius of curvature, so that they are effectively "blind" to it.

According to this reasoning, short-distance interactions are a sufficient condition to ensure the validity of an area law, regardless of possible background curvature. In particular, this justifies our choice of restricting ourselves to flat space when studying the entanglement entropy of black holes – at least as long as the ultraviolet cutoff of the model is indeed much larger than the radius of spacetime curvature at the horizon. In the limit of small black holes, we expect our description to break down, and full quantum gravity to be required.

6.2.3 Contribution from longer-range interactions

In light of this importance of short-wave dominance, we should also note when it breaks down. Das et al. [18] found one such instance: when applying the same procedure described above, of eliminating interactions outside a certain "window", to excited states, they found some contribution from longer-range interactions. This may be related to the fact that the area law does not hold for ordinary thermodynamical systems. That discussion, however, goes well beyond the scope of this work.

6.3 Massive field

Unlike excited states and non-zero spin, giving the field a non-zero mass M is a step towards a more realistic situation that requires only minor modifications to the calculation. Perhaps because of this, most of the results carry over: the independence of the size of the grid, the boundary effects, the spectra of $S_l \times l$ in spherical symmetry, and the area law itself. Only the coefficient κ changes: massive fields account for less entanglement entropy than a massless field in the same situation, as shown in figure 6.6.

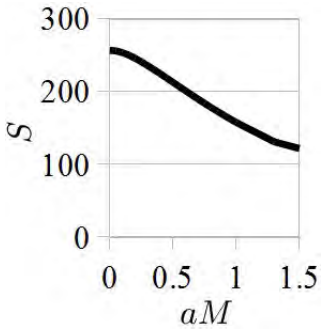

Figure 6.6: Entanglement entropy of a solid sphere with $n_{out} = 30$ in flat space, as a function of the field mass M, multiplied by the discretization constant a for dimensional reasons.

The picture of a discrete grid of oscillators, whose components are entangled with each other across the horizon, suggests a simple explanation for this effect. It is based on the concept of *correlation length*, the characteristic distance over which two points in a field are typically correlated, for instance via entanglement. This length is infinite in the case of massless fields, and decreases as $1/M$ as the mass grows. When the field is modelled by a discrete grid, if the correlation length becomes smaller than the distance between nearest neighbours, there can be no more correlations, no entanglement, ergo no entropy. The trend we observe can be understood as the transition from zero mass to this situation, and its scale is set once again by the ultraviolet cutoff.

Indeed, the quantity that multiplies the mass term in the Hamiltonian (5.50) (and the corresponding ones for spherical symmetry), and which therefore determines the decrease of S, is the product aM (that is, in the context of the above explanation, the ratio of the grid spacing to the correlation length, $a/(1/M)$). Since the cutoff a is of the order of the Planck length, whereas M is of the order of $10^{-17} M_{Pl}$ for the top quark, and even smaller for other known fields, we always have $aM \ll 1$ (in natural units), which implies that the change of S is negligible.

The data appear to indicate that even a field of one Planck mass would still generate roughly half as much entropy as a massless one. However, one should bear in mind that our model is based on theories that break down at the Planck scale, and the numbers found for large masses are not reliable. We can only extrapolate from tendencies observed in the low mass limit, which suggest that there will be some decrease in the entropy due to mass.

7 Conclusion

Our analysis of entanglement entropy shows that it provides a favourable cost-benefit ratio, as explanations go: even in an exceedingly simple setup - a ground-state scalar field, in flat or uniformly curved space, with a boundary that is not imposed by causality, but chosen at will - and without any extravagant concepts or formalisms beyond those already familiar from general relativity and quantum field theory, it generates entropy that obeys an area law, regardless of the shape of the traced-out region and possible background curvature. The latter, in particular, will be relevant in a future generalization to an actual black hole background. More importantly, we conclude from our results that all these features, which one must demand in any potential explanation for black hole entropy, arise because the effect is dominated by short-range interactions. Thus, any phenomenon that can generate entropy based on short-distance correlations has the potential to provide a mechanism for black hole entropy.

However, the dominance of short-range interactions naturally requires a constant to define what is "short", and the physical role this constant plays is somewhat obscure. The Planck scale is the natural choice, since it provides a length scale determined by constants of nature and marks the limit beyond which one must employ full quantum gravity, but knowing the value of the constant does not necessarily bring us any closer to understanding its meaning. One can try to gain some insight into that matter from the results presented here. They show that the cutoff, implemented in our model as the spacing a of a discrete grid, plays a twofold role in the behaviour of entanglement entropy: for one, it determines the size of the unit area that contributes one bit of entanglement entropy. Moreover, it sets the scale for the localization of the entropy, in the sense that the dominant contribution comes from degrees of freedom located within no more than $\sim 3a$ of the horizon. Indirectly, this is also responsible for the mass dependence: if the mass of a field is so large that interactions over distances greater than a are suppressed, then it is unlikely to entangle even nearest neighbours on the grid, without which there is no entropy. However, we feel this understanding of the role of the cutoff is far from complete, and may warrant further investigation.

The above evidence seems to suggest that physics below the Planck scale may actually be sufficient to explain the origin of a black hole's entropy, just as it can account for its temperature[34]. This is only unsatisfactory in that it deprives us of

[34]In the simplified situations studied in this work, the reduced state is not, in fact, thermal, as discussed in section 5.2.3. Further research into why our model fails to reproduce this particular feature would hopefully lead to a better understanding of this approach to black hole thermodynamics.

what we hoped would be an opportunity to observe phenomena pertaining to the quantum gravity regime.

However, if this is indeed the mechanism that is responsible for black holes behaving as if they possessed temperature and entropy, we feel it important to stress that neither of those properties actually pertain to the hole itself, in the sense of counting its microstates or measuring the change of the energy stored in them. Rather, the mixed quantum state with a thermal distribution describes the fields around the hole. It merely arises because of the unusual causal structure that constitutes the black hole. Hopefully, this insight, that the physical manifestation of the branch of physics known as "black hole thermodynamics" is not restricted to the gravitational phenomenon alone, will prove one step towards a deeper understanding of this theory.

A Mathematical identities

A.1 Gaussian integrals

A.1.1 Gaussian integral in one dimension

In its prototypical form, this identity is

$$\int_{-\infty}^{+\infty} dx \exp\left(-a\left(x-x_0\right)^2 - c\right) = \sqrt{\frac{\pi}{a}} \exp\left(-c\right), \tag{A.1}$$

where $a > 0$, c and x_0 are constants.

A.1.2 Gaussian integral in n dimensions

Let the argument of the integrand be a second-order polynomial of n variables, $x_1, \ldots, x_n$, over which the integration is performed:

$$\left[\prod_{j=1}^{n} \int_{-\infty}^{+\infty} dx_j\right] \exp\left(-\sum_{j,l=1}^{n} x_j A_{jl} x_l - 2\sum_{j=1}^{n} x_j \bar{B}_j - \bar{C}\right), \tag{A.2}$$

where $A_{n\times n}$ is real, symmetric and positive-definite (all eigenvalues are strictly positive), and $\bar{B}_j$ and $\bar{C}$ are constants.

Given that A is real and symmetric, there exists an orthogonal matrix S,

$$SS^T = S^T S = \mathbb{I}, \tag{A.3}$$

that diagonalizes A,

$$A_D = SAS^T. \tag{A.4}$$

Denoting the eigenvalues of A by a_j, the diagonal matrix is

$$\left(A_D\right)_{jl} \equiv a_j \delta_{jl}, \tag{A.5}$$

and

$$\det A = \prod_{j=1}^{n} a_j. \tag{A.6}$$

We are furthermore assuming that all eigenvalues of A are strictly positive (non-zero), so that

$$\frac{1}{a_j}\delta_{jl} = \left(A_D^{-1}\right)_{jl}, \tag{A.7}$$

in terms of which the inverse of A is simply

$$A^{-1} = S^T \left(A_D\right)^{-1} S. \tag{A.8}$$

Let

$$S^T y \equiv x \qquad S\bar{B} \equiv b, \tag{A.9}$$

then

$$\sum_{j,l=1}^{n} x_j A_{jl} x_l = x^T A x = y^T S \left(S^T A_D S\right) S^T y = y^T A_D y = \sum_{j=1}^{n} y_j^2 a_j, \tag{A.10}$$

$$\sum_{j=1}^{n} x_j \bar{B}_j = x^T \bar{B} = y^T S \bar{B} = \sum_{j=1}^{n} y_j b_j \tag{A.11}$$

and

$$\prod_{j=1}^{n} dx_j = |\det S^T| \prod_{j=1}^{n} dy_j = \prod_{j=1}^{n} dy_j, \tag{A.12}$$

so that

$$\begin{aligned} &- \textstyle\sum_{j,l=1}^{n} x_j A_{jl} x_l - 2 \sum_{j=1}^{n} x_j \bar{B}_j - \bar{C} \\ &\quad = \textstyle\sum_{j=1}^{n} \left(-a_j y_j^2 - 2 b_j y_j\right) - \bar{C} \\ &\quad = \textstyle\sum_{j=1}^{n} -a_j \left(y_j + \frac{b_j}{a_j}\right)^2 + \frac{b_j^2}{a_j} - \bar{C}. \end{aligned} \tag{A.13}$$

The second to last term can be expressed as

$$\sum_{j=1}^{n} \frac{b_j^2}{a_j} = \sum_{j,l=1}^{n} b_j \frac{1}{a_j} \delta_{jl} b_l = b^T S A^{-1} S^T b = \bar{B}^T A^{-1} \bar{B}. \tag{A.14}$$

The variables y thus allow us to write the integral in n-dimensional space as a product of n one-dimensional Gaussian integrals:

$$\begin{aligned} &\left[\textstyle\prod_{j=1}^{n} \int_{-\infty}^{+\infty} dx_j\right] \exp\left(- \textstyle\sum_{j,l=1}^{n} x_j A_{jl} x_l - 2 \sum_{j=1}^{n} x_j \bar{B}_j - \bar{C}\right) \\ &= \textstyle\prod_{j=1}^{n} \left[\int_{-\infty}^{+\infty} dy_j \exp\left(-a_j \left(y_j + \frac{b_j}{a_j}\right)^2\right)\right] \exp\left(+\bar{B}^T A^{-1} \bar{B} - \bar{C}\right). \end{aligned} \tag{A.15}$$

Since all eigenvalues a_j are positive, we can substitute equation (A.1), finding

$$\begin{aligned}\left[\prod_{j=1}^{n}\int_{-\infty}^{+\infty}dx_j\right]\exp\left(-\sum_{j,l=1}^{n}x_jA_{jl}x_l-2\sum_{j=1}^{n}x_j\bar{B}_j-\bar{C}\right)\\ =\sqrt{\frac{\pi^n}{\det A}}\exp\left(\sum_{j,l=1}^{n}\bar{B}_j\left(A^{-1}\right)_{jl}\bar{B}_l-\bar{C}\right).\end{aligned}\tag{A.16}$$

A.1.3 Gaussian integral in n of N dimensions

We now consider an exponential whose argument is a product of two N-component vectors, but which is integrated only over the first n variables:

$$\left[\prod_{j=1}^{n}\int_{-\infty}^{+\infty}dx_j\right]\exp\left(-\frac{1}{2}x^T\Omega x-\frac{1}{2}x'^T\Omega x'\right),\tag{A.17}$$

with $x_j=x'_j$ for $j=1,\ldots,n$. We assume that Ω is real, symmetric and positive-definite.

We begin by decomposing Ω into submatrices,

$$\Omega_{N\times N}=\begin{pmatrix}A_{n\times n} & B_{n\times(N-n)}\\ B^T_{(N-n)\times n} & C_{(N-n)\times(N-n)}\end{pmatrix},\tag{A.18}$$

so that

$$x^T\Omega x=\sum_{j,l=1}^{n}x_jA_{jl}x_l+2\sum_{j=1}^{n}\sum_{l=n+1}^{N}x_jB_{jl}x_l+\sum_{j,l=n+1}^{N}x_jC_{jl}x_l.\tag{A.19}$$

Then the argument of the exponential takes the same form as in equation (A.16):

$$-\frac{1}{2}x^T\Omega x-\frac{1}{2}x'^T\Omega x'=-\sum_{j,l=1}^{n}x_jA_{jl}x_l-2\sum_{j=1}^{n}x_j\bar{B}_j-\bar{C},\tag{A.20}$$

with the n-component vector $\bar{B}$ given by

$$\bar{B}_j=\frac{1}{2}\sum_{l=n+1}^{N}B_{jl}\left(x_l+x'_l\right),\tag{A.21}$$

and

$$\bar{C}=\frac{1}{2}\sum_{j,l=n+1}^{N}\left(x_jC_{jl}x_l+x'_jC_{jl}x'_l\right).\tag{A.22}$$

Furthermore, A is positive-definite and invertible (see proof below), so that the identity (A.16) can be used.

The first term in the argument of the exponential is

$$\sum_{j,l=1}^{n} \bar{B}_j \left(A^{-1}\right)_{jl} \bar{B}_l = \frac{1}{4} \sum_{j,l=n+1}^{N} \left(x_j + x'_j\right) \left(B^T A^{-1} B\right)_{jl} \left(x_l + x'_l\right). \tag{A.23}$$

For notational convenience, we introduce the $(N-n) \times (N-n)$ matrices

$$\beta = \frac{1}{2} B^T A^{-1} B \qquad \gamma = C - \beta, \tag{A.24}$$

and note that x now represents an $(N-n)$-component vector, $(x_{n+1}, \ldots, x_N)$. Thus,

$$\begin{aligned} \textstyle\sum_{j,l=1}^{n} \bar{B}_j \left(A^{-1}\right)_{jl} \bar{B}_l - \bar{C} &= \frac{1}{2} \left(x + x'\right)^T \beta \left(x + x'\right) - \frac{1}{2} x^T C x - \frac{1}{2} x'^T C x' \\ &= -\frac{1}{2} \left(x^T \gamma x + x'^T \gamma x'\right) + x^T \beta x', \end{aligned} \tag{A.25}$$

using the fact that β is symmetric as well. Therefore, finally,

$$\begin{aligned} &\left[\textstyle\prod_{j=1}^{n} \int_{-\infty}^{+\infty} dx_j\right] \exp\left(-\tfrac{1}{2} x^T \Omega x - \tfrac{1}{2} x'^T \Omega x'\right) \\ &= \sqrt{\tfrac{\pi^n}{\det A}} \exp\left(-\tfrac{1}{2} \left[x^T \gamma x + x'^T \gamma x'\right] + x^T \beta x'\right), \end{aligned} \tag{A.26}$$

under the conditions and using the definitions given in the derivation.

Proof that A is positive-definite and invertible We assume that $\Omega_{N\times N}$ is real, symmetric and positive-definite. This implies that there exists a matrix U that makes

$$\Omega_D \equiv U \Omega U^T \tag{A.27}$$

diagonal, and that the eigenvalues of Ω are strictly positive:

$$\omega_k > 0 \;\; \forall k = 1, \ldots, N. \tag{A.28}$$

The rows of U are given by the eigenvectors of Ω, which we will denote by

$$u_i^{(k)} \equiv U_{ki} \;\; \forall i, k = 1, \ldots, N. \tag{A.29}$$

They form a complete set, which spans the space of N-dimensional vectors: any N-component vector w can be expanded as

$$w_i = \sum_{k=1}^{N} \left[\sum_{j=1}^{N} w_j u_j^{(k)}\right] u_i^{(k)} \equiv \sum_{k=1}^{N} c^k u_i^{(k)}. \tag{A.30}$$

In particular, if only the first n components of w are non-zero, its expansion coefficients can also be written as

$$c^k = \sum_{j=1}^{n} w_j u_j^{(k)}. \tag{A.31}$$

Note that the only vector whose c^k are all zero is the null vector.

Now, we define the $n \times n$ submatrix A, given by

$$A_{ij} = \Omega_{ij} = \sum_{k=1}^{N} \omega_k u_i^{(k)} u_j^{(k)} \quad \forall i, j = 1, \ldots, n. \tag{A.32}$$

Like Ω, it is real and symmetric, and consequently diagonalizable. Let $s^{(l)}$ be its eigenvectors and a_l the corresponding eigenvalues:

$$\sum_{j=1}^{n} A_{ij} s_j^{(l)} = a_l s_i^{(l)} \quad \forall i = 1, \ldots, n. \tag{A.33}$$

We can assume without loss of generality that the $s^{(l)}$ are orthonormal,

$$\sum_{i=1}^{n} s_i^{(k)} s_i^{(l)} = \delta_{kl} \quad \forall k, l = 1, \ldots, n, \tag{A.34}$$

so that we can isolate

$$\begin{aligned} a_l &= \sum_{i,j=1}^{n} s_i^{(l)} A_{ij} s_j^{(l)} \\ &= \sum_{k=1}^{N} \omega_k \left[\sum_{i=1}^{n} s_i^{(l)} u_i^{(k)} \right]^2 \\ &= \sum_{k=1}^{N} \omega_k \left(c_l^k \right)^2 . \end{aligned} \tag{A.35}$$

The c_l^k are expansion coefficients of the form introduced in equation (A.31); they are real, so that $\left(c_l^k\right)^2 \geq 0 \ \ \forall k = 1, \ldots, N$. Furthermore, since the vectors $s^{(l)}$ form a basis, and can therefore not be identically zero, at least one of the factors $\left(c_l^k\right)^2$ for any given l must be strictly positive. Finally, equation (A.28) establishes that $\omega_k > 0 \ \ \forall k = 1, \ldots, N$. Consequently,

$$a_l > 0 \quad \forall l = 1, \ldots, n, \tag{A.36}$$

all eigenvalues of A are strictly positive. Considering also that A is diagonalizable, it follows that it is invertible.

A.2 One-oscillator thermal density operator in the position representation

We will show that the normalized density matrix

$$\rho_{\gamma,\beta}(x,x') = \sqrt{\frac{\gamma-\beta}{\pi}} \exp\left[-\frac{\gamma}{2}\left(x^2+x'^2\right)+\beta x x'\right] = \langle x|\rho_{\gamma,\beta}|x'\rangle \tag{A.37}$$

describes a simple harmonic oscillator with mass $m=1$ and frequency

$$\omega = \sqrt{\gamma^2-\beta^2}, \tag{A.38}$$

in a mixed state corresponding to a temperature

$$T = \omega/\ln\left(\frac{\gamma+\omega}{\beta}\right). \tag{A.39}$$

In order to prove this, one could begin with the known thermal density matrix in terms of energy eigenstates, given by equation (2.58):

$$\rho_T = Z^{-1}\sum_{n=0}^{\infty} \exp\left(-E_n/T\right)|n\rangle\langle n|. \tag{A.40}$$

In the position representation, this becomes

$$\begin{aligned}\rho_T(x,x') = \; & \frac{\sqrt{m\omega/\pi}}{1-\exp\left(-\omega/T\right)} \exp\left(-\frac{1}{2}m\omega x^2-\frac{1}{2}m\omega x'^2\right) \\ & \times\sum_{n=0}^{\infty}\frac{1}{n!2^n}\exp\left(-n\omega/T\right)H_n\left(\sqrt{m\omega}x\right)H_n\left(\sqrt{m\omega}x'\right).\end{aligned} \tag{A.41}$$

However, transforming this expression into $\rho_{\gamma,\beta}$ requires extensive calculations. Instead, we begin with expression (A.37) and prove that the wave functions $\varphi_n(x) \equiv \langle x|n\rangle$ of stationary states are eigenfunctions of the density operator, and that the eigenvalues p_n are the probabilities of finding $|n\rangle$ in a mixture at temperature T, if ω and T are related to γ and β by the expressions given above. That is, we will first prove that

$$\int dx'\langle x|\rho_{\gamma,\beta}|x'\rangle\langle x'|n\rangle = p_n\langle x|n\rangle. \tag{A.42}$$

According to equation (2.52), the wave functions are

$$\varphi_n(x) \equiv \langle x|n\rangle = \frac{1}{\sqrt{n!2^n}}\left(\frac{\omega}{\pi}\right)^{1/4}\exp\left(-\frac{1}{2}\omega x^2\right)H_n\left(\sqrt{\omega}x\right), \tag{A.43}$$

so the integral on the left-hand side of equation (A.42) is

$$I_n(x) \equiv \int dx' \sqrt{\frac{\gamma-\beta}{\pi}} \exp\left[-\frac{\gamma}{2}\left(x^2+x'^2\right)+\beta x x'\right] \times \frac{1}{\sqrt{n!2^n}}\left(\frac{\omega}{\pi}\right)^{1/4} \exp\left(-\frac{1}{2}\omega x'^2\right) H_n\left(\sqrt{\omega}x'\right). \tag{A.44}$$

Then

$$I_n(x) = \frac{1}{\sqrt{n!2^n}}\left(\frac{\omega}{\pi}\right)^{1/4} \sqrt{\frac{\gamma-\beta}{\pi}} \exp\left[-\frac{\gamma}{2}x^2\right] J_n(x), \tag{A.45}$$

with

$$J_n(x) \equiv \int dx' \exp\left[-\frac{\gamma+\omega}{2}x'^2+\beta x x'\right] H_n\left(\sqrt{\omega}x'\right). \tag{A.46}$$

We use the generating function of the Hermite polynomials,

$$\exp\left[-\lambda^2+2\lambda z\right] = \sum_{n=0}^{\infty}\frac{\lambda^n}{n!}H_n(z), \tag{A.47}$$

to write

$$\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}J_n(x) = \int dx' \exp\left[-\frac{\gamma+\omega}{2}x'^2+\beta x x'\right] \exp\left[-\lambda^2+2\lambda\sqrt{\omega}x'\right]. \tag{A.48}$$

The argument of the exponential can be rearranged as

$$-\frac{\gamma+\omega}{2}\left(x'-\frac{\beta x+2\lambda\sqrt{\omega}}{\gamma+\omega}\right)^2+\frac{\left(\beta x+2\lambda\sqrt{\omega}\right)^2}{2(\gamma+\omega)}-\lambda^2, \tag{A.49}$$

and we can apply formula (A.1) to evaluate the integral:

$$\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}J_n(x) = \sqrt{\frac{2\pi}{\gamma+\omega}} \exp\left[-\lambda^2+\frac{4\lambda^2\omega+4\lambda\sqrt{\omega}\beta x+\beta^2x^2}{2(\gamma+\omega)}\right]. \tag{A.50}$$

Differentiating n times with respect to λ and setting $\lambda = 0$, we isolate

$$J_n(x) = \sqrt{\frac{2\pi}{\gamma+\omega}}\left[\frac{d^n}{d\lambda^n}\exp\left[-\lambda^2+\frac{4\lambda^2\omega+4\lambda\sqrt{\omega}\beta x+\beta^2x^2}{2(\gamma+\omega)}\right]\right]_{\lambda=0}. \tag{A.51}$$

The argument of this exponential can be written as

$$-\frac{\gamma-\omega}{\gamma+\omega}\left(\lambda-\frac{\sqrt{\omega}\beta x}{\gamma-\omega}\right)^2+\frac{\gamma+\omega}{2}x^2, \tag{A.52}$$

and we define

$$\lambda' \equiv \sqrt{\frac{\gamma-\omega}{\gamma+\omega}}\left(\lambda - \frac{\sqrt{\omega}\beta x}{\gamma-\omega}\right) \qquad \frac{d}{d\lambda} = \sqrt{\frac{\gamma-\omega}{\gamma+\omega}}\frac{d}{d\lambda'}, \tag{A.53}$$

so that

$$\begin{aligned} J_n(x) = \; & \sqrt{\frac{2\pi}{\gamma+\omega}}\left(\frac{\gamma-\omega}{\gamma+\omega}\right)^{n/2} \exp\left(\frac{\gamma+\omega}{2}x^2\right) \\ & \cdot \times \left[\frac{d^n}{d\lambda'^n}\exp\left(-\lambda'^2\right)\right]_{\lambda'=-\sqrt{\omega}\beta x/\sqrt{\gamma^2-\omega^2}}. \end{aligned} \tag{A.54}$$

Another useful property of the Hermite polynomials is that

$$H_n(z) = (-1)^n \exp\left(z^2\right)\frac{d^n}{dz^n}\exp\left(-z^2\right), \tag{A.55}$$

from which it follows that

$$\frac{d^n}{d\lambda'^m}\exp\left(-\lambda'^2\right) = \exp\left(-\lambda'^2\right)(-1)^n H_n(\lambda') = \exp\left(-\lambda'^2\right) H_n(-\lambda'). \tag{A.56}$$

Substituting,

$$J_n(x) = \sqrt{\frac{2\pi}{\gamma+\omega}}\left(\frac{\gamma-\omega}{\gamma+\omega}\right)^{n/2} \exp\left[\left(\frac{\gamma+\omega}{2} - \frac{\omega\beta^2}{\gamma^2-\omega^2}\right)x^2\right] H_n\left(\frac{\sqrt{\omega}\beta x}{\sqrt{\gamma^2-\omega^2}}\right), \tag{A.57}$$

and using this expression in (A.45), we obtain

$$\begin{aligned} I_n(x) = \; & \frac{1}{\sqrt{n!2^n}}\left(\frac{\omega}{\pi}\right)^{1/4}\sqrt{\frac{\gamma-\beta}{\pi}}\sqrt{\frac{2\pi}{\gamma+\omega}}\left(\frac{\gamma-\omega}{\gamma+\omega}\right)^{n/2} \\ & \times \exp\left[\left(\frac{\omega}{2} - \frac{\omega\beta^2}{\gamma^2-\omega^2}\right)x^2\right] H_n\left(\frac{\sqrt{\omega}\beta x}{\sqrt{\gamma^2-\omega^2}}\right). \end{aligned} \tag{A.58}$$

The integral is proportional to $\varphi_n(x)$, that is, the wave function of a stationary state is an eigenfunction of $\rho_{\gamma,\beta}$, if

$$\beta = \sqrt{\gamma^2-\omega^2}. \tag{A.59}$$

Then

$$\left(\frac{\gamma-\omega}{\gamma+\omega}\right)^{1/2} = \frac{\sqrt{\gamma^2-\omega^2}}{\gamma+\omega} = \frac{\beta}{\gamma+\omega} \equiv \xi \tag{A.60}$$

and

$$2\frac{\gamma-\beta}{\gamma+\omega} = 1 + \frac{\gamma-\omega}{\gamma+\omega} - 2\frac{\beta}{\gamma+\omega} = 1 + \xi^2 - 2\xi = (1-\xi)^2, \tag{A.61}$$

so that

$$I_n(x) = (1-\xi)\,\xi^n\varphi_n(x) = p_n\varphi_n(x). \tag{A.62}$$

The eigenvalues are of the same form as the probabilities of finding energy eigenstates in a thermal mixture, given by equation (2.56). A comparison of the expressions implies that the temperature is

$$T = \omega / \ln\left(\frac{\gamma+\omega}{\beta}\right). \tag{A.63}$$

Thus, we have

$$\int dx' \langle x|\rho_{\gamma,\beta}|x'\rangle\langle x'|n\rangle = p_n\langle x|n\rangle, \tag{A.64}$$

and, since

$$\int dx' |x'\rangle\langle x'| = \mathbb{I} = \sum_{m=0}^{\infty} |m\rangle\langle m|, \tag{A.65}$$

this implies that

$$p_n\langle x|n\rangle = \langle x|\rho_{\gamma,\beta}|n\rangle = \sum_{m=0}^{\infty}\langle x|m\rangle\langle m|\rho_{\gamma,\beta}|n\rangle = \sum_{m=0}^{\infty}\langle x|m\rangle\delta_{mn}p_n. \tag{A.66}$$

Since this holds for all $|n\rangle$, which form a basis, we can isolate

$$\langle m|\rho_{\gamma,\beta}|n\rangle = \delta_{mn}p_n, \tag{A.67}$$

and consequently

$$\rho_{\gamma,\beta} = \sum_{n=0}^{\infty} p_n|n\rangle\langle n| = \rho_T, \tag{A.68}$$

the two density matrices describe the same state.

B Code for numerical calculations

This section contains examples of the code used to implement the calculations developed in section 5, designed for the command-line interface of the Mathematica® standalone kernel (version 7.0). We restrict ourselves to two examples, which represent slightly different algorithms, but they can be customized (for instance to include different sets of independent variables or to produce different outputs) with a basic understanding of the underlying programming language. Most of the variables used in the code are defined in subsections 5.2 and 5.4, and additional explanations have been inserted as comments (formatted as (*comment*)) where necessary. To simplify the input in the command line, one first defines a function, `ent`, such that it returns the entropy of a certain setup (for instance a spherical shell in flat space). Then `ent` can be called repeatedly, with the parameters (such as n_{in} and n_{out}) of different situations as input.

B.1 Cubic grid

The following code defines the function ent for a solid parallelepiped on a cubic grid. The input variables in this version are xmin and xext, which determine the position and the extension in the x dimension of the traced-out region.

```
ent:=Function[{xmin,xext},
(*part 1: set variables and define functions*)
m=0; (*massless field*)
tot=10; xtot=tot; ytot=tot; ztot=tot; (*number of grid points in each dimension*)
min=3; ymin=min; zmin=min; (*position of traced-out region in y and z dimensions*)
ext=4; yext=ext; zext=ext; (*extension of traced-out region in y and z dimensions*)
si=Simplify[ If[Abs[#]<(10^-10), 0, -Log[1-(#/(1+Sqrt[1-#^2]))] -Log[#/(1+Sqrt[1-#^2])]
(#/(1+Sqrt[1-#^2]))/ (1-(#/(1+Sqrt[1-#^2])))]&];
(*implements equation (5.37), giving the entropy contribution of each eigenvalue of β′*)
(*part 2: generate coupling matrix*)
trace=ConstantArray[0,{xtot,ytot,ztot}];
indtrac=0; indkeep=0; ntot=xtot*ytot*ztot;
(*The trace array assigns a single sequential index to each grid point, as suggested at the end
of section 5.3.4. Sites with indices from 1 through indtrac=ntot-indkeep are traced out, while
indtrac+1 through ntot are kept.*)
For[x=1,x<=xtot,x++,
For[y=1,y<=ytot,y++,
For[z=1,z<=ztot,z++,
If[((xmin<=x<xmin+xext)&&(ymin<=y<ymin+yext)&&(zmin<=z<zmin+zext))(*At this point,
one can insert a second condition to generate a hole in the traced-out region, of the form
&&!(condition for a point to lie in the hole).*),
indtrac++; trace[[x,y,z]]=indtrac,
trace[[x,y,z]]=ntot-indkeep; indkeep++] ] ]];
k=ConstantArray[0,{ntot,ntot}]; (*coupling matrix*)
For[x=1,x<=xtot,x++,
For[y=1,y<=ytot,y++,
For[z=1,z<=ztot,z++,
k[[ trace[[x,y,z]],trace[[x,y,z]] ]]=N[6+m^2];
If[x>1,k[[ trace[[x,y,z]],trace[[x-1,y,z]] ]]=N[-1]];
If[x<xtot,k[[ trace[[x,y,z]],trace[[x+1,y,z]] ]]=N[-1]];
If[y>1,k[[ trace[[x,y,z]],trace[[x,y-1,z]] ]]=N[-1]];
If[y<ytot,k[[ trace[[x,y,z]],trace[[x,y+1,z]] ]]=N[-1]];
If[z>1,k[[ trace[[x,y,z]],trace[[x,y,z-1]] ]]=N[-1]];
If[z<ztot,k[[ trace[[x,y,z]],trace[[x,y,z+1]] ]]=N[-1]]; ]]];
(*part 3: compute β′, which characterizes the reduced state*)
{eivalk,u}=Eigensystem[k]; (*eigenvalues and -vectors of the coupling matrix*)
```

```
kdiag=DiagonalMatrix[eivalk];
omega= Transpose[u].MatrixPower[kdiag,1/2].u;
a=omega[[1;;indtrac,1;;indtrac]];
b=omega[[indtrac+1;;ntot,1;;indtrac]];
c=omega[[indtrac+1;;ntot,indtrac+1;;ntot]];
beta=(1/2)Transpose[b].Inverse[a].b;
gamma=c-beta;
{eivalgamma,v}=Eigensystem[gamma]; (*eigenvalues and -vectors of gamma*)
gammadiag=DiagonalMatrix[eivalgamma];
gammasqrt=ConstantArray[0,{indkeep,indkeep}];
For[i=1,i<=indkeep,i++,gammasqrt[[i,i]]=(gammadiag[[i,i]])^(-1/2)];
betaprime= gammasqrt.v.beta.Transpose[v].gammasqrt ;
(*part 4: compute entropy*)
listsi=Map[si,Eigenvalues[betaprime]];
s=Re[Sum[Part[listsi,i],{i,indkeep}]] (*this variable is returned as the output of the function
ent*)]
```

Then the entropy for $x_{min} = 3$ and $x_{ext} = 4$, for example, can be computed by calling

```
In[2]:= ent[3,4]
Out[2]:= 1.976
```

B.2 Spherical symmetry

The algorithm in this case is similar to that for a cubic grid; the main difference being the iteration over the partial wave index l (until some truncation criterion is met) to compute the series of contributions S_l (equation (5.80)). In order to allow for plots like figure 5.5, the program prints out this series (listsl) along with the total value of the entropy.

A more technical difference is that we do not relabel the sites (as we did in the cubic case), but simply use the radial index to designate the rows and columns of Ω. Hence, when describing a hollow shell, the sites with indices $i = n_{in} + 1, ...n_{out}$ must be traced out, while both $1, ...n_{in}$ and $n_{out} + 1, ...n_{tot}$ are kept, which requires some rearranging of the matrices (see below). This version of the code also contains the option traceshell, which determines whether the shell (radial indices $i = n_{in}+1, ...n_{out}$) is traced out while the remainder is kept (traceshell=1) or vice versa (0). Comparing the results in both cases allows one to verify that the entropy is unaffected by the choice, as predicted by the symmetry theorem.

```
ent:=Function[{nout,nin,traceshell},
(*part 1: set variables and define functions*)
m=0; (*massless field*)
t=0.03; (*truncation parameter*)
lmax=3000; (*override truncation to prevent infinite calculations in case of bad parameters*)
ntot=IntegerPart[1.2*nout]+1; (*total grid points in radial direction*)
indkeep=ntot-nout+nin;
indtrac=nout-nin;
entropy=0; (*initial value for summation*)
sl=1; (*initial value is set non-zero, so that the comparison "sl>t*Max[listasl]" gives "true"*)
listsl={}; (*stores all S_l*)
If[(indkeep==0)||(indtrac==0),Goto[empty]]; (*if nothing is traced out*)
si=Simplify[If[Abs[#]<(10^-10),0,-Log[1-(#/(1+Sqrt[1-#^2]))]-Log[#/(1+Sqrt[1-#^2])](#/(1+Sqrt[1-
#^2]))/(1-(#/(1+Sqrt[1-#^2])))]&];
(*implements equation (5.37), giving the entropy contribution of each eigenvalue of β'*)
For[l=0,(l<lmax)&&(sl>t*Max[listasl]),l++, (*iteration over l*)
(*part 2: generate coupling matrix*)
k= N[Table[KroneckerDelta[i,j](2+1/(2j^2)+(l+1)*l/j^2+m^2)
-KroneckerDelta[i,j+1](j+1/2)^2/(j^2+j)
-KroneckerDelta[j,i+1](i+1/2)^2/(i^2+i),{i,ntot},{j,ntot}]];
(*part 3: compute β', which characterizes the reduced state*)
{eivalk,u}=Eigensystem[k]; (*eigenvalues and -vectors of the coupling matrix*)
kdiag=DiagonalMatrix[eivalk];
omega= Transpose[u].MatrixPower[kdiag,1/2].u;
(*rearrange matrices*)
c=ConstantArray[0,{indkeep,indkeep}];
c[[1;;nin,1;;nin]]=omega[[1;;nin,1;;nin]];
c[[1;;nin,nin+1;;indkeep]]=omega[[1;;nin,nout+1;;ntot]];
c[[nin+1;;indkeep,1;;nin]]=omega[[nout+1;;ntot,1;;nin]];
c[[nin+1;;indkeep,nin+1;;indkeep]]=omega[[nout+1;;ntot,nout+1;;ntot]];
b=ConstantArray[0,{indtrac,indkeep}];
b[[1;;indtrac,1;;nin]]=omega[[nin+1;;nout,1;;nin]];
b[[1;;indtrac,nin+1;;indkeep]]=omega[[nin+1;;nout,nout+1;;ntot]];
a=omega[[nin+1;;nout,nin+1;;nout]];
If[traceshell==1,
beta=(1/2)Transpose[b].Inverse[a].b; gamma=c-beta; dim=indkeep,
beta=(1/2)b.Inverse[c].Transpose[b]; gamma=a-beta; dim=indtrac];
v=Map[Normalize,Eigenvectors[gamma]];
vinv=Transpose[v];
gammadiag=Re[v.gamma.vinv];
For[i=1,i<=dim,i++,
```

```
If[gammadiag[[i,i]]<0,Print["Gamma error at l=",l," i=",i," eigenvalue: ",gammadiag[[i,i]]]];
(*checks for negative eigenvalues*)
gammasqrt=MatrixPower[gammadiag,-1/2];
betaprime= gammasqrt.v.beta.vinv.gammasqrt;
(*part 4: compute entropy*)
listsi=Map[si,Eigenvalues[betaprime]];
sl=(2l+1)*Re[Sum[Part[listsi,i],{i,dim}]];
listsl=Append[listsl,sl];
entropy+=sl];
If[l==lmax-1,Print["lmax exceeded"]];
Label[empty];
Print["Partial wave contributions (S_l): ",listsl];
Print["Entropy: ", entropy]]
```

The above code contains the coupling matrix for spherical symmetry in flat space (expression (5.71)), but the generalization to a curved background is straightforward: n_{tot} is no longer set to $1.2n_{out}$, but an independent variable, and the K_{ij} are given by expression (5.95), excluding the indices $i = 0$ and n_{tot} to avoid divergences. When studying even values of n_{tot}, there is also the problem of deviant S_l (see figure 5.5(right)): since they tend to trigger the truncation criterium $S_l < t \cdot S_{ldom}$ before the series has actually begun to converge, one should either set t to a large value or remove the condition altogether, using instead l<lmax alone.

References

[1] S. W. Hawking, Nature **248**, 30 (1974).

[2] R. M. Wald, *Quantum field theory in curved spacetime and black hole thermodynamics* (The University of Chicago Press, 1994).

[3] J. Preskill, *Do black holes destroy information?*, URL `hep-th/9209058`.

[4] S. W. Hawking, Phys. Rev. D **14**, 2460 (1976).

[5] J. D. Bekenstein, Lettere Al Nuovo Cimento **4**, 737 (1972).

[6] L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin, Phys. Rev. D **34**, 373 (1986).

[7] M. Srednicki, Phys. Rev. Lett. **71**, 666 (1993).

[8] R. F. Werner, Phys. Rev. A **40**, 4277 (1989), URL `http://link.aps.org/doi/10.1103/PhysRevA.40.4277`.

[9] R. M. Wald, *General Relativity* (The University of Chicago Press, 1984).

[10] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, 1982).

[11] T. Jacobson, *Introductory lectures on black hole thermodynamics* (1996), URL `www.physics.umd.edu/grt/taj/776b/lectures.pdf`.

[12] L. C. B. Crispino, A. Higuchi, and G. E. A. Matsas, Rev. Mod. Phys. **80** (2008).

[13] P. Longhi and R. Soldati, Phys. Rev. D **83**, 107701 (2011).

[14] B. Carter, *General Relativity* (Cambridge University Press, 1979).

[15] J. M. Bardeen, B. Carter, and S. W. Hawking, Commun. Math. Phys. **31**, 161 (1973).

[16] J. D. Bekenstein, Phys. Rev. D **7**, 2333 (1973).

[17] S. W. Hawking, Phys. Rev. Lett. **26**, 1344 (1971).

[18] S. Das, S. Shankaranarayanan, and S. Sur, *Black hole entropy from entanglement: a review*, URL `arXiv:0806.0402 [gr-qc]`.

[19] D. Gioev and I. Klich, Phys. Rev. Lett. **96**, 100503 (2006).