Universitat de Barcelona

Master in Astrophysics, Particle Physics and Cosmology

Master Thesis

Spacetime thermodynamics and entanglement entropy: the Einstein field equations

Author: Martí Berenguer Mimó

NIUB: 17527646

Advisor: Roberto Emparan

September 2019

Ackonwledgements

First of all, I would like to thank my tutor and advisor, Roberto Emparan, for all the help and dedication on supervising my thesis for the last months. He has always been there to assist me when needed. Thank you.

I would also like to thank my family, for encouraging me during all these years of study, and to who I have stolen a lot of time.

Finally, I would like to thank my master classmates. In particular, Oscar and Joseba, for all those sleepless nights working on the assignments, and with whom I have spent an amazing year.

Abstract

In this master thesis, the posibility of a connection between spacetime dynamics (driven by the Einstein equations) and thermodynamics is discussed. Some known results, like the Raychaudhuri equation or the Unruh effect are reviewed in order to make the presentation self-contained.

The Einstein equations are derived in two different ways from thermodynamic arguments. The first one (Section (3)) uses the thermodynamic relation $\delta Q = TdS$, together with the proporionality of entropy and horizon area. In the second derivation (Section (4)), the Einstein equations are derived from an hypothesis about entanglement entropy in a maximally symmetric spacetime.

Some questions regarding the implications of this thermodynamic interpretation of spacetime are discussed as a conclusion of the thesis.

Contents

1 Introduction

One of the most surprising results of Albert Einstein's general theory of relativity was the existence of black holes, regions of spacetime were gravity is so strong, that nothing, even light, could escape. The interior of a black hole is separated from the rest of the universe by an event horizon. This means that any particle, massive or massless, that is located inside the black hole, will never be able to escape, and is doomed to reach a singularity in its future, where its proper time suddenly ends and the known theories of physics stop to work.

When Stephen Hawking studied black holes from a more mathematical point of view, he found an interesting result: the area of the event horizon never decreases with time and, in general, it will increase. This implies that, if two black holes collide and merge, the area of the final black hole will be larger than the sum of areas of the colliding black holes.

This behavior is analogous to the behavior of entropy in thermodynamic systems, where the Second Law of Thermodynamics says that the entropy of a system can never decrease, and that the total entropy of a system is larger than the entropy of its subsystems:

Second Law of Black Hole Mechanics:

$$\delta A \geq 0$$

Second Law of Thermodynamics:

$$\delta S \geq 0$$

This analogy is more evident with the *First Law of Black Hole Mechanics*, which relates the change in mass of a black hole with the change in area of the horizon and the change in angular momentum and electric charge. From here one can see that if the area of the event horizon is analogous to the entropy, then the surface gravity κ is analogous to the temperature:

First Law of Black Hole Mechanics:

$$\delta E = \frac{\kappa}{8\pi}\delta A + \Omega\delta J + \Phi\delta Q$$

First Law of Thermodynamics:

$$\delta E = T\delta S + P\delta V$$

There is even a *Zeroth Law of Black Hole Mechanics*:

Zeroth Law of Black Hole Mechanics:

κ is the same along the horizon in a time-independent black hole.

Zeroth Law of Thermodynamics:

Temperature is the same in all points in a system in thermal equilibrium.

Because of these similarities, Bekenstein proposed that the entropy of a black hole should be proportional to the area of the event horizon, and proposed the Generalized Second Law [1]: the sum of the entropy of the black hole and the entropy of the matter outside the black hole never decreases.

The fact that black holes have temperature and entropy implies that they should radiate, but this was completely against the classical picture of black holes. One of the most important results in theoretical physics during the last century was the one obtained by Stephen Hawking in 1974, when he found that, when quantum effects around a black hole are considered, it radiates thermally at the so-called *Hawking temperature* [2]:

$$T_H = \frac{1}{8\pi M} \tag{1.1}$$

Another important result is that this thermal behavior is not exclusive of black hole horizons. In 1976, W.G. Unruh demonstrated the following: the vacuum state, defined by inertial observers, has a thermal character for uniformly accelerated observers with proper acceleration a (Rindler observers) at the Unruh temperature [3, 4]:

$$T = \frac{\hbar a}{2\pi} \tag{1.2}$$

This means that, around any event, in any spacetime, there is a class of observers that will perceive the spacetime as hot. This thermal character of spacetime (not only for black hole horizons) will be of great importance for the following sections.

Moreover, this thermodynamic interpretation of spacetime invokes some questions about the structure of spacetime at smallest scales. From standard thermodynamics it is known that a macroscopic system like, for example, a gas, can be described with some thermodynamic variables, like the temperature or the entropy, but for a long time, the real meaning of these variables was unknown. It was Boltzmann who gave an explanation to these variables, essentially saying "if you can heat it, it has microscopic degrees of freedom". Before that, it was considered that matter was continuous even at the smallest scales, and the concepts of heat and temperature were added “by hand”. Boltzmann used the discrete interpretation of matter and found that the thermodynamic fenomena were related with the averages of the properties of these microscopic degrees of freedom.

This is profound. It tells that the existence of microscopic degrees of freedom leaves a signature at macroscopic scales, in the form of temperature and heat. Then, if spacetime is seen as hot by some observers, what are the microscopic degrees of freedom that give raise to the temperature and the entropy? There are many approaches that try to give an interpretation to these microscopic degrees of freedom [5, 6], but there is not a clear answer yet. What seems reasonable is that, if spacetime is, at its deepest level, a thermodynamic entity, we should be able to derive the equations that drive its evolution (the Einstein equations) from a purely thermodynamic point of view.

This is what we will do in Sections (3) and (4). In Section (3), the Einstein equations are derived from the thermodynamic relation $\delta Q = TdS$ and the proportionality of the entropy and horizon area, working from the point of view of a Rindler observer in the neighbourhood of the causal horizon of the Rindler space.

In Section (4), an alternative derivation of the Einstein equations will be given, based in the assumption that the entanglement entropy in a geodesic ball is maximal when the geometry and the quantum fields are varied from maximal symmetry.

Section (2) includes some of the conceptual ideas and equations that will be necessary for the two derivations.

In Section (5), some comments about the derivations and some questions about the implications of them are considered, while Section (6) contains the main conclusions of the thesis.

2 Basic previous topics

This section contains some topics that will be necessary to have in mind during the two derivations of the Einstein equations in the following sections. Here we will briefly talk about the Raychaudhuri equation, Rindler space, the Unruh effect, and entanglement entropy.

2.1 Raychaudhuri equation

The Raychaudhuri equation is an evolution equation for what is called the *expansion* of a congruence of geodesics. In order to understand the meaning of the expansion (and two more quantities that appear in the equation, the shear and rotation), it is useful to think first about the kinematics of a deformable medium.

Suppose, in a purely Newtonian context, a two-dimensional medium, with some internal motion whose dynamics are not of our interest. From a purely kinematic point of view, we can always write that, for an infnitesimal displacement ξ^a from a reference point O,

$$\frac{d\xi^a}{dt} = B^a{}_b(t)\xi^b + \mathcal{O}(\xi^2) \tag{2.1}$$

for some tensor $B^a{}_b$, which depends on the internal dynamics of the medium. For short intervals of time,

$$\xi^a(t_1) = \xi^a(t_0) + \Delta\xi^a(t_0) \tag{2.2}$$

where

$$\Delta\xi^a(t_0) = B^a{}_b(t_0)\xi^b(t_0)\Delta t + \mathcal{O}(\Delta t^2) \tag{2.3}$$

and $\Delta t = t_1 - t_0$. To describe the action of $B^a{}_b$ we will consider the situation that $\xi^a(t_0) = r_0(\cos\phi, \sin\phi)$; that is, a circle of radius r_0 in the two-dimensional medium.

Expansion

Suppose that $B^a{}_b$ is a pure-trace matrix, i.e., proportional to the identity, with the form

$$B^a{}_b = \begin{pmatrix} \frac{1}{2}\theta & 0 \\ 0 & \frac{1}{2}\theta \end{pmatrix}$$

In this case, $\Delta\xi^a = \frac{1}{2}\theta r_0 \Delta t(\cos\phi, \sin\phi)$, which corresponds to a change in the circle's radius by an amount $\frac{1}{2}\theta r_0\Delta t$. The correspoonding change in area is given by

$$\Delta A = A_1 - A_0 = \pi r_0^2 \theta \Delta t \tag{2.4}$$

This means that

$$\theta = \frac{1}{A_0}\frac{\Delta A}{\Delta t} \tag{2.5}$$

θ measures the fractional change of area per unit time, and is called the *expansion parameter*.

Shear

Suppose now that $B^a{}_b$ is symmetric and trace-free:

$$B^a{}_b = \begin{pmatrix} \sigma_+ & \sigma_\times \\ \sigma_\times & -\sigma_+ \end{pmatrix}$$

In this case, $\Delta\xi^a = r_0 \Delta t(\sigma_+ \cos\phi + \sigma_\times \sin\phi, -\sigma_+ \sin\phi + \sigma_\times \cos\phi)$. If $\sigma_\times = 0$, we have an ellipse with the major axis oriented in the $\phi = 0$ direction. If $\sigma_+ = 0$, what we have is an ellipse oriented in the $\phi = \pi/4$ direction. The general situation is an ellipse oriented along an arbitrary direction. The area of the figure is not affected by the action of $B^a{}_b$. What we have is a shearing of the figure, and the parameters $\sigma_\times$ and σ_+ are called the *shear parameters*.

Rotation

Finally, if $B^a{}_b$ is antisymmetric,

$$B^a{}_b = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}$$

we have that $\Delta\xi^a = r_0 \omega \Delta t(\sin\phi, -\cos\phi)$, and $\xi^a(t_1) = r_0(\cos\phi', \sin\phi')$, with $\phi' = \phi - \omega\Delta t$. This corresponds to an overall rotation of the original figure, keeping the area fixed. ω is called the *rotation parameter*.

The most general decomposition of this tensor into algebraically irreducible components under rotations is

$$B^a{}_b = \begin{pmatrix} \frac{1}{2}\theta & 0 \\ 0 & \frac{1}{2}\theta \end{pmatrix} + \begin{pmatrix} \sigma_+ & \sigma_\times \\ \sigma_\times & -\sigma_+ \end{pmatrix} + \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}$$

which can also be expressed as

$$B_{ab} = \frac{1}{2}\theta\delta_{ab} + \sigma_{ab} + \omega_{ab} \tag{2.6}$$

where $\theta = B^a{}_a$ (the expansion scalar) is the trace part of B_{ab}, $\sigma_{ab} = B_{(a,b)} - \frac{1}{2}\theta\delta_{ab}$ (the shear tensor) is the symmetric-tracefree part of B_{ab}, and $\omega_{ab} = B_{[a,b]}$ (the rotation tensor) is the antisymmetric part of B_{ab}. For a three-dimensional medium, the decomposition is the same, but with a prefactor of 1/3 instead of 1/2 in the trace term, and the interpretation of the expansion, shear and rotation are the same, but changing the area by the volume.

Once the classical 2-dimensional medium has been introduced, we can move now to the study of congruences of (for now, timelike) geodesics.

Let $\mathcal{O}$ be an open region of spacetime. A congruence of geodesics in $\mathcal{O}$ is a family of geodesics such that through each point in $\mathcal{O}$ passes one and only one geodesic from this family. We will assume that the geodesics are timelike. We are interested in the evolution of the deviation vector ξ^a between two neighbouring geodesics in the congruence as a function of the proper time τ (see Figure (1)).

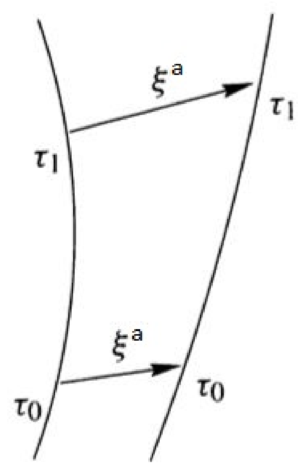

Figure 1: Two neighboring geodesics, with a deviation vector ξ^a as a function of τ.

Let u^a be the (timelike) tangent vector to the geodesics. Then, the spacetime metric g_{ab} can be decomposed in a longitudinal part $-u_a u_b$ and a transverse part h_{ab},

$$h_{ab} = g_{ab} + u_a u_b \tag{2.7}$$

The transverse metric h_{ab} is purely spatial, in the sense that it is orthogonal to u^a. We introduce now the tensor field

$$B_{ab} = \nabla_b u_a \tag{2.8}$$

This tensor determines the evolution of the deviation vector ξ^a. To see this, note that from $u^b\nabla_b\xi^a = \xi^b\nabla_b u^a$ we obtain

$$u^b\nabla_b\xi^a = B^a{}_b\xi^b \tag{2.9}$$

That is, $B^a{}_b$ measures the failure of ξ^a to be parallel transported along the congruence. Equation (2.9) is analogous to (2.1), and therefore we can decompose the tensor $B^a{}_b$ in the same way as before, with the same interpretation for the diffrent terms that appear:

$$B_{ab} = \frac{1}{3}\theta h_{ab} + \sigma_{ab} + \omega_{ab} \tag{2.10}$$

In order to find the evolution equation for the expansion θ, we can start by finding an evolution equation for B_{ab}:

$$\begin{aligned} u^c\nabla_c B_{ab} &= u^c\nabla_c\nabla_b u_a \\ &= u^c\left(\nabla_b\nabla_a u_a - R_{adbc}u^d\right) \\ &= u^c\nabla_b\nabla_c u_a - R_{adbc}u^c u^d \\ &= \nabla_b\left(u^c\nabla_c u_a\right) - \left(\nabla_b u^c\right)\left(\nabla_c u_a\right) - R_{adbc}u^c u^d \\ &= -B^c{}_b B_{ac} - R_{adbc}u^c u^d \end{aligned} \tag{2.11}$$

Taking the trace of this equation, we obtain

$$\frac{d\theta}{d\tau} = -B^{ab}B_{ba} - R_{ab}u^a u^b \tag{2.12}$$

Now, from the definition of B_{ab}, we find that $B_{ab}B^{ab} = \frac{1}{3}\theta^2 + \sigma^{ab}\sigma_{ab} - \omega^{ab}\omega_{ab}$, so (2.12) becomes

$$\frac{d\theta}{d\tau} = -\frac{1}{3}\theta^2 - \sigma^{ab}\sigma_{ab} + \omega^{ab}\omega_{ab} - R_{ab}u^a u^b \tag{2.13}$$

which is kown as the Raychaudhuri equation, and gives the evolution of the expansion parameter θ for a congruence of timelike geodesics. For the case of null geodesics, which is the one that will be of interest in the following sections, the line of argument is the same as for timelike geodesics, but the calculation is a bit more tedious because of the difficulty to precisely define the transverse spacetime. In the above case it was simply the spatial components, but in the case of null geodesics, if k^a is the (null) tangent vector to the geodesics, the orthogonal space to k^a includes k^a because it is orthogonal to itself. Once this technical part is solved, the logic of the derivation and the result are very similar. The Raychaudhuri equation for a congruence of null geodesics reads:

$$\boxed{\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \sigma^{ab}\sigma_{ab} + \omega^{ab}\omega_{ab} - R_{ab}k^a k^b} \tag{2.14}$$

2.2 Rindler space

The Rindler space is introduced when one is interested in the motion of an accelerated observer in flat spacetime. This will be necessary in Section (3), where the whole argumentation line will be centered from the perspective of an accelerated observer in an approximately flat region of spacetime.

For simplicity, let's consider the 2-dimensional Minkowski space, whose metric, in the usual (t, x) coordinates is

$$ds^2 = -dt^2 + dx^2 \tag{2.15}$$

An observer moving at a uniform acceleration α will follow the trajectory $x^\mu(\tau)$ given by

$$t(\tau) = \frac{1}{\alpha}\sinh(\alpha\tau) \tag{2.16}$$

$$x(\tau) = \frac{1}{\alpha}\cosh(\alpha\tau) \tag{2.17}$$

This can be checked taking into account that the components of the 4-acceleration

$$a^\mu = \frac{D^2 x^\mu}{d\tau^2} = \frac{d^2 x^\mu}{d\tau^2} \tag{2.18}$$

where the covariant derivative is equal to the ordinary derivative because the Christoffel symbols vanish in these coordinates, are given by

$$a^t = \alpha\sinh(\alpha\tau) \tag{2.19}$$

$$a^x = \alpha\cosh(\alpha\tau) \tag{2.20}$$

In this way, the magnitude of the acceleration is

$$\sqrt{a_\mu a^\mu} = \sqrt{-\alpha^2\sinh^2(\alpha\tau) + \alpha^2\cosh^2(\alpha\tau)} = \alpha \tag{2.21}$$

Thus, this trajectory corresponds to a uniformly accelerated observer. The trajectory of this observer obeys

$$x^2(\tau) = t^2(\tau) + \frac{1}{\alpha^2} \tag{2.22}$$

which is an hyperboloid asymptoting to null paths $x = -t$ in the past and $x = t$ in the future (see Figure (2)).

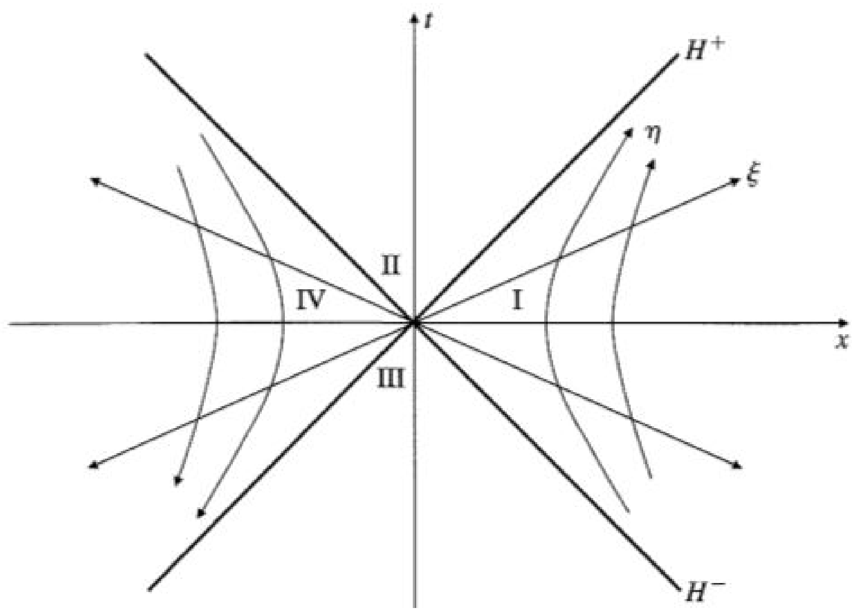

Figure 2: Minkowski spacetime in Rindler coordinates. An observer with constant acceleration in the $+x$ direction follows the hyperbolic trajectories dranw in region I. The patches H^+ and H^- act as horizons for this class of observers.

Notice from (2.22) that the larger the acceleration α, the closer the trajectory is to the patches $x = -t$ and $x = t$. This fact will be important in Section (3).

We can define new coordinates (η, ξ) in the following way:

$$t = \frac{1}{a} e^{a\xi} \sinh(a\eta), \qquad x = \frac{1}{a} e^{a\xi} \cosh(a\eta), \qquad (x > |t|) \tag{2.23}$$

which cover the wedge $x > |t|$ (region I in Figure (2)). Although these are not the usual Rindler coordinates, they are the most appropiate for the derivation of the Unruh effect, which is the purpose of the next section. Notice that an accelerated observer with acceleration $\alpha = a$ follows a world line that is given by $\xi = const$ and $\eta = \tau$.

In these coordinates, the metric is given by

$$ds^2 = e^{2a\xi}\left(-d\eta^2 + d\xi^2\right) \tag{2.24}$$

Region I, with these coordinates, is known as Rindler space (although it is only a part of Minkowski space). A Rindler observer is an observer moving along a constant acceleration path (in the diagram, this corresponds to the hyperbolic trajectories).

Because the metric components are independent of η, the vector ∂_η is a Killing vector. In the (t, x) coordinates, this Killing vector is

$$\begin{aligned} \partial_\eta &= \frac{\partial t}{\partial \eta}\partial_t + \frac{\partial x}{\partial \eta}\partial_x \\ &= e^{a\xi}\left[\cosh(a\eta)\partial_t + \sinh(a\eta)\partial_x\right] \\ &= a\left(x\partial_t + t\partial_x\right) \end{aligned} \tag{2.25}$$

Notice that the patches $x = -t$ and $x = t$ (H^- and H^+) act as Killing horizons for this vector field, because its norm vanishes (only) there:

$$V = a\left(x\frac{\partial}{\partial t} + t\frac{\partial}{\partial x}\right) \quad \Rightarrow \quad V_\mu V^\mu = a^2\left(t^2 - x^2\right) = a(t+x)(t-x) \tag{2.26}$$

The surface gravity of this Killing horizon is

$$\kappa = \sqrt{-\frac{1}{2}\nabla^\mu \xi^\nu \nabla_\mu \xi_\nu} = a \tag{2.27}$$

Although there is no gravitational field (we are in flat spacetime), the surface gravity characterizes the acceleration of the Rindler observers.

It will be convenient for the Unruh effect to define coordinates (η, ξ) for the region IV, by flipping the signs of those defined in region I:

$$t = -\frac{1}{a}e^{a\xi}\sinh(a\eta), \qquad x = -\frac{1}{a}e^{a\xi}\cosh(a\eta), \qquad (x < |t|) \tag{2.28}$$

2.3 The Unruh effect

The basic statement of the Unruh effect is that an accelerating observer in flat space will observe the Minkowski vacuum as a thermal spectrum of particles. The basic idea of this result is the fact that observers with different notions of positive and negative frequency modes will disagree on the particle content of a given state.

In flat spacetime, this problem does not arise for non-accelerated (inertial) observers. For inertial observers, we introduce a set of positive and negative frequency modes, and the fields are expressed as a combination of these modes, interpreting the operator coefficients as creation and annihilation operators. In flat spacetime we can choose a natural set of modes by demanding that they are positive-frequency modes with respect to the time coordinate. Obviously, the time coordinate is not unique, because we can perform Lorentz transformations, but the vacuum state and the number operators are invariant under these transformations.

In curved spacetime (or accelerated observers) we can find a set of modes, but we can find many other sets that are equally good, and the notion of vacuum and number operators will be very sensitive to the set we choose.

We can always find a set of orthonormal modes f_i, and expand the fields in terms of these modes:

$$\phi = \sum_i \left(\hat{a}_i f_i + \hat{a}_i^\dagger f_i^* \right) \tag{2.29}$$

where the operators $\hat{a}_i$ and $\hat{a}_i^\dagger$ obey the usual commutation relations. We can define a vacuum state $|0_f\rangle$, which will be annihilated by all the annihilation operators,

$$\hat{a}_i \left|0_f\right\rangle = 0, \qquad \forall i \tag{2.30}$$

From this vacuum we can define an entire Fock basis, defining the excitations as the states created by the action of $a_i^\dagger$. The number operator can be defined too,

$$\hat{n}_{f_i} = \hat{a}_i^\dagger \hat{a}_i \tag{2.31}$$

where the subscript f makes reference to the fact that this operator is defined with respect to the set of modes f_i. But we can find another complete basis with respect to which expand the fields,

$$\phi = \sum_i \left(\hat{b}_i g_i + \hat{b}_i^\dagger g_i^* \right) \tag{2.32}$$

where, again, $\hat{b}_i$ and $\hat{b}_i^\dagger$ obey the usual commutation relations. The vacuum state, the Fock basis, and the number operator for the $\hat{b}_i$ operators are defined in the same way as for the operators of the f_i modes:

$$\hat{b}_i \left|0_g\right\rangle = 0 \ \ \forall i; \qquad \hat{n}_{g_i} = \hat{b}_i^\dagger \hat{b}_i \tag{2.33}$$

If one observer defines particles with respect to the set of modes f_i and a different observer defines particles with respect to the set of modes g_i, in general they will disagree on the number of particles they observe. To see this, we can expand each set of modes in terms of the other:

$$g_i = \sum_j \left(\alpha_{ij} f_i + \beta_{ij} f_j^*\right) \tag{2.34}$$

$$f_i = \sum_j \left(\alpha_{ji}^* g_j - \beta_{ji} g_j^*\right) \tag{2.35}$$

The transformation that allows to write one set of modes in terms of the other is called *Bogoliubov transformation*, and the coefficients α_{ij}, β_{ij} are called *Bogoliubov coefficients*, which satisfy the normalization conditions

$$\sum_j \left(\alpha_{ij}\alpha_{jk}^* - \beta_{ik}\beta_{jk}^*\right) = \delta_{ij} \tag{2.36}$$

$$\sum_j \left(\alpha_{ik}\beta_{jk} - \beta_{ik}\alpha_{jk}\right) = 0 \tag{2.37}$$

and can be used to relate not only the modes, but also the operators:

$$\hat{a}_i = \sum_j \left(\alpha_{ij}\hat{b}_j + \beta_{ji}^*\hat{b}_j^\dagger\right) \tag{2.38}$$

$$\hat{b}_i = \sum_j \left(\alpha_{ij}^*\hat{a}_j - \beta_{ij}^*\hat{a}_j^\dagger\right) \tag{2.39}$$

The discrepancy on the number of particles can be seen from the following calculation: imagine that the system is in the f-vacuum (in which the observer using the f_i modes would not see any particle). We want to know the number of particles that an observer using the g-modes will observe. Then, we compute the expectation value of the g number operator in the f-vacuum:

$$
\begin{aligned}
\langle 0_f | \hat{n}_{g_i} | 0_f \rangle &= \langle 0_f | \hat{b}_i^\dagger \hat{b}_i | 0_f \rangle \\
&= \langle 0_f | \sum_{jk} \left(\alpha_{ij} \hat{a}_j^\dagger - \beta_{ij} \hat{a}_j \right) \left(\alpha_{ik} \hat{a}_k - \beta_{ik}^* \hat{a}_k^\dagger \right) | 0_f \rangle \\
&= \sum_{jk} \beta_{ij} \beta_{ik}^* \langle 0_f | \hat{a}_j \hat{a}_k^\dagger | 0_f \rangle \\
&= \sum_{jk} \beta_{ij} \beta_{ik}^* \langle 0_f | \left(\hat{a}_k^\dagger \hat{a}_j + \delta_{jk} \right) | 0_f \rangle \\
&= \sum_{jk} \beta_{ij} \beta_{ik}^* \delta_{jk} \langle 0_f | 0_f \rangle \\
&= \sum_{j} \beta_{ij} \beta_{ij}^* \\
&= \sum_{j} |\beta_{ij}|^2
\end{aligned}
\tag{2.40}
$$

$$
\Rightarrow \qquad \boxed{\langle 0_f | \hat{n}_{g_i} | 0_f \rangle = \sum_j |\beta_{ij}|^2} \tag{2.41}
$$

In general, this coefficient does not vanish: an observer that defines particles with respect to the g-modes will detect particles where the observer that defines particles with respect to the f-modes will see the vacuum.

This can be applied to the case of an accelerated observer in flat spacetime (Rindler observer). For simplicity, we will consider a massless Klein-Gordon field in 2 dimensions. The Klein-Gordon equation in Rindler coordinates takes the form

$$
\Box \phi = e^{-2a\xi} \left(-\partial_\eta^2 + \partial_\xi^2 \right) \phi = 0 \tag{2.42}
$$

which admits as solutions normalized plane waves of the form

$$
g_k = \frac{1}{\sqrt{4\pi\omega}} e^{-i\omega\eta + ik\xi} \tag{2.43}
$$

with $\omega = |k|$. Because the choice of coordinates for regions I and IV needed a difference of sign between them, we need to define two sets of modes, one for each of the two regions [7]:

$$
g_k^{(1)} = \begin{cases} \frac{1}{\sqrt{4\pi\omega}} e^{-i\omega\eta + ik\xi} & \text{I} \\ 0 & \text{IV} \end{cases}
$$

$$g_k^{(2)} = \begin{cases} 0 & \text{I} \\ \frac{1}{\sqrt{4\pi\omega}} e^{+i\omega\eta + ik\xi} & \text{IV} \end{cases}$$

In this way, each set of modes is positive-frequency with respect to the corresponding future-directed timelike Killing vector,

$$\partial_\eta g_k^{(1)} = -i\omega g_k^{(1)} \tag{2.44}$$

$$\partial_{(-\eta)} g_k^{(2)} = -i\omega g_k^{(2)} \tag{2.45}$$

Introducing the corresponding creation and annihilation operators for each region, the field can be expressed as

$$\phi = \int dk \left(\hat{b}_k^{(1)} g_k^{(1)} + \hat{b}_k^{(1)\dagger} g_k^{(1)*} + \hat{b}_k^{(2)} g_k^{(2)} + \hat{b}_k^{(2)\dagger} g_k^{(2)*} \right) \tag{2.46}$$

The modes to which we will compare them will be the usual Minkowski modes, which expand the field as

$$\phi = \int dk \left(\hat{a}_k f_k + \hat{a}_k^\dagger f_k^* \right) \tag{2.47}$$

The Minkowski vacuum state $|0_M\rangle$ and the Rindler vacuum state $|0_R\rangle$ are defined as usual:

$$\hat{a}_k |0_M\rangle = 0 \tag{2.48}$$

$$\hat{b}_k^{(1)} |0_R\rangle = \hat{b}_k^{(2)} |0_R\rangle = 0 \tag{2.49}$$

The next step now is to compute the Bogoliubov coefficients relating both sets of modes, and compute the expectation value of the Rindler number operator in the Minkowski vacuum. This is a bit tedious, and the usual procedure is the following: instead of using the previously defined Rindler modes, we will take a set of modes $h_k^{(1)}$, $h_k^{(2)}$ that share the same vacuum state as the Minkowski modes (but the excited states are different). The way to do this is to start with the Rindler modes, analytically extend them to the entire spacetime, and express them in terms of the orignal Rindler modes.

Then, the field will be expanded as

$$\phi = \int dk \left(\hat{c}_k^{(1)} h_k^{(1)} + \hat{c}_k^{(1)\dagger} h_k^{(1)*} + \hat{c}_k^{(2)} h_k^{(2)} + \hat{c}_k^{(2)\dagger} h_k^{(2)*} \right) \tag{2.50}$$

The properly normalized version of these modes is [7]

$$h_k^{(1)} = \frac{1}{\sqrt{2\sinh\left(\frac{\pi\omega}{a}\right)}}\left(e^{\pi\omega/2a}g_k^{(1)} + e^{-\pi\omega/2a}g_{-k}^{(2)*}\right) \tag{2.51}$$

$$h_k^{(2)} = \frac{1}{\sqrt{2\sinh\left(\frac{\pi\omega}{a}\right)}}\left(e^{\pi\omega/2a}g_k^{(2)} + e^{-\pi\omega/2a}g_{-k}^{(1)*}\right) \tag{2.52}$$

Just like before, the Bogoliubov coefficients allow to relate also the creation and annihilation operators:

$$\hat{b}_k^{(1)} = \frac{1}{\sqrt{2\sinh\left(\frac{\pi\omega}{a}\right)}}\left(e^{\pi\omega/2a}\hat{c}_k^{(1)} + e^{-\pi\omega/2a}\hat{c}_{-k}^{(2)\dagger}\right) \tag{2.53}$$

$$\hat{b}_k^{(2)} = \frac{1}{\sqrt{2\sinh\left(\frac{\pi\omega}{a}\right)}}\left(e^{\pi\omega/2a}\hat{c}_k^{(2)} + e^{-\pi\omega/2a}\hat{c}_{-k}^{(1)\dagger}\right) \tag{2.54}$$

In this way, the Rindler number operator in region I,

$$\hat{n}_R^{(1)}(k) = \hat{b}_k^{(1)\dagger}\hat{b}_k^{(1)} \tag{2.55}$$

can be expressed in terms of the new operators $\hat{c}_k^{(1,2)}$, and because they share the same vacuum state as the Minkowski modes, we have that

$$\hat{c}_k^{(1)}\left|0_M\right\rangle = \hat{c}_k^{(2)}\left|0_M\right\rangle = 0 \tag{2.56}$$

The fact that the excited states do not coincide is not a problem, because we are only interested in what the Rindler observer sees when the state is the Minkowski vacuum. For a Rindler observer in region I, the expectation value of the number operator will be

$$\begin{aligned}
\left\langle 0_M\right|\hat{n}_R^{(1)}(k)\left|0_M\right\rangle &= \left\langle 0_M\right|\hat{b}_k^{(1)\dagger}\hat{b}_k^{(1)}\left|0_M\right\rangle \\
&= \frac{1}{2\sinh\left(\frac{\pi\omega}{a}\right)}\left\langle 0_M\right|e^{-\pi\omega/a}\hat{c}_{-k}^{(1)}\hat{c}_{-k}^{(1)\dagger}\left|0_M\right\rangle \\
&= \frac{e^{-\pi\omega/a}}{2\sinh\left(\frac{\pi\omega}{a}\right)}\left\langle 0_M\right|\hat{c}_{-k}^{(1)}\hat{c}_{-k}^{(1)\dagger}\left|0_M\right\rangle \\
&= \frac{1}{e^{2\pi\omega/a}-1}
\end{aligned} \tag{2.57}$$

This result corresponds to a Planck spectrum with temperature

$$T = \frac{a}{2\pi} \tag{2.58}$$

Thus, a uniformly accelerated observer through the Minkowski vacuum will detect a thermal flux of particles.

2.4 Entanglement entropy

In order to conceptually understand entanglement entropy, it is useful to first take a look to the following discrete problem [8]: imagine a lattice model, with discrete degrees of freedom located at the lattice sites, which are separated a distance ϵ (see Figure (3)). At each site (labeled by α) we have a finite-dimensional Hilbert space $\mathcal{H}_\alpha$ (for instance, a qubit per site). A pure quantum state of the system can be written as

$$|\Psi\rangle \in \otimes_\alpha \mathcal{H}_\alpha \tag{2.59}$$

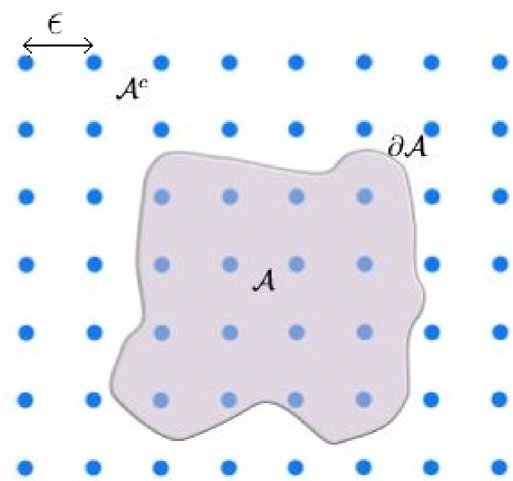

Figure 3: Discrete lattice system, with a Hilbert space at each place. The grey region is called $\mathcal{A}$, while $\mathcal{A}^c$ is its complementary, separted by the boundary $\partial\mathcal{A}$. The distance between places is ϵ.

We can divide the lattice system into two complementary subsystems, namely $\mathcal{A}$ and $\mathcal{A}^c$, separated by the boundary $\partial\mathcal{A}$, which we shall call the *entangling surface*, as can be seen in Figure (3). The Hilbert space of the total system has been split into the direct product of two Hilbert spaces,

$$\otimes_\alpha \mathcal{H}_\alpha = \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_{\mathcal{A}^c} \tag{2.60}$$

Now, one can construct the *reduced density matrix* of the subsystem $\mathcal{A}$, which is constructed by tracing out the degrees of freedom of $\mathcal{A}^c$:

$$\rho_\mathcal{A} = \mathrm{Tr}_{\mathcal{A}^c}(|\Psi\rangle \langle\Psi|) \tag{2.61}$$

If the state $|\Psi\rangle$ is factorized when the system is split, then we will have a pure state in $\mathcal{H}_\mathcal{A}$. However, if the state can not be written as a direct product of states from the two subsystems, the state is entangled and the density matrix gives the probabilities for the ocurrence of the states in $\mathcal{H}_\mathcal{A}$. The amount of entanglement that exists in $|\Psi\rangle$ when the system is split is quantified by the Von-Neumann entropy of the reduced density matrix, or *entanglement entropy*, which is given by

$$S_\mathcal{A} = -\,\mathrm{Tr}_\mathcal{A}(\rho_\mathcal{A} \log \rho_\mathcal{A}) \tag{2.62}$$

In a discrete system, this can be computed diagonalizing the density matrix and obtaining its eigenvalues λ_i. Then, the entanglement entropy is simply

$$S_\mathcal{A} = -\sum_i \lambda_i \log \lambda_i \tag{2.63}$$

Because $|\Psi\rangle$ is a pure state, it can be decomposed via the Schmidt decomposition, $|\Psi\rangle = \sum_i \lambda_i |\alpha_i\rangle_\mathcal{A} |\beta_i\rangle_{\mathcal{A}^c}$. This tells us that non-trivial eigenvalues of $\rho_\mathcal{A}$ are the same as those of $\mathcal{A}^c$. Then, the traces are the same, and the entanglement entropies are also the same:

$$S_\mathcal{A} = S_{\mathcal{A}^c} \tag{2.64}$$

The fact that the entropy is the same for both regions means that it can not depend on the size of each region, but only on the degrees of freedom shared by the two regions. That is, it must be proportional to the area of the boundary $\partial\mathcal{A}$ instead of being proportional to the volume of the regions, as it would be expected in classical thermodynamic systems.

The continuum limit of this system can be defined as taking the limit $\epsilon \to 0$. When this is done, the result for the entropy is sensitive to the ultra-violet (UV) physics, as we should expect.

For a d-dimensional free field theory, the entropy is a UV-divergent quantity, with the leading term being proportional to the area of the entangling surface [8]:

$$S_{\mathcal{A}} = \gamma \frac{\text{Area}(\partial \mathcal{A})}{\epsilon^{d-2}} + ... \tag{2.65}$$

where γ is a constant that depends on the model used. This quantity is divergent when $\epsilon \to 0$ unless there is some physical UV cutoff (presumably, of the order of the Planck scale), with which the entropy would be finite and proportional to A/L_p^2, matching with the Bekenstein-Hawking entropy for black holes [9–12]. Thus we shall assume in all cases that due to the UV physics, the entanglement entropy is finite in small regions, with a leading term given by S=ηA.

Another important quantity that will be useful is the *relative entropy*. Given two density matrices ρ and σ we can define the relative entropy,

$$S(\rho||\sigma) = \text{Tr}(\rho \log \rho) - \text{Tr}(\rho \log \sigma) \tag{2.66}$$

which gives information about the distinguishability between the two density matrices. An important property of the relative entropy is that it is always positive or equal to zero, being equal to zero only when the two density matrices are the same. We can define the *modular hamiltonian* as

$$K_\rho = -\log \rho \tag{2.67}$$

and rewrite the relative entropy as

$$\begin{aligned} S(\rho||\sigma) &= \text{Tr}(\rho \log \rho) - \text{Tr}(\rho \log \sigma) + \text{Tr}(\sigma \log \sigma) - \text{Tr}(\sigma \log \sigma) \\ &= -S(\rho) + \text{Tr}(\rho K_\sigma) - \text{Tr}(\sigma K_\sigma) + S(\sigma) \\ &= \Delta \langle K \rangle - \Delta S \end{aligned} \tag{2.68}$$

where $\Delta S = S(\rho) - S(\sigma)$ is the entropy difference between the states, and $\Delta\langle K\rangle = \text{Tr}(\rho K_\sigma) - \text{Tr}(\sigma K_\sigma)$ is the difference in the expectation values of the modular hamiltonian K_σ for ρ and σ.

If we consider σ to be a reference state $\sigma = \rho_0$, and ρ a state close to it, we can expand the latter in a power series in a parameter λ, $\rho(\lambda) = \rho_0 + \lambda\rho_1 + \lambda^2\rho_2 + ...$ in such a way that $\rho(0) = \rho_0 = \sigma$. The relative entropy can be expanded as

$$S(\rho(\lambda)||\sigma) = S(\rho(0)||\sigma) + \frac{d}{d\lambda} S(\rho(\lambda)||\sigma)\big|_{\lambda=0} \lambda + \mathcal{O}(\lambda^2) \tag{2.69}$$

The first term is zero because of the definition of the relative entropy. The term of order λ is also zero, because the relative entropy is a monotonically increasing function around σ. Thus, the relative entropy is at least quadratic in the deviation parameter. This means that, for first-order variations, we have that

$$\delta S = \delta \langle H_\sigma \rangle \tag{2.70}$$

This is known as the *first law of entanglement entropy*.

3 First derivation. Equilibrium thermodynamics in Rindler space

In this approach, Einstein's equations are derived from the proportionality of entropy and horizon area, together with the thermodynamic relation $\delta Q = TdS$, relating heat, temperature and entropy (and area, due to the relation entropy $\sim$ area).

In standard thermodynamics, heat is defined as energy that flows from, or to a thermodynamic system. Here, we shall define heat as energy that flows across a causal horizon (not necessarily a black hole horizon).

In the relation $\delta Q = TdS$, we associate δQ with an energy flux across the horizon, and we shall use that the entropy is proportional to the area of this horizon. It remains to identify the temperature T. Using Unruh's results, we can take T to be the Unruh temperature if we consider that the observer is in accelerated motion. Then, for consistency, the heat flow must be defined as the energy flux that this observer measures. In order to apply local equilibrium thermodynamics, two conditions must be imposed in the construction of our system:

- We need the observer to be as near as possible to the horizon. In the limit that the accelerated worldine approaches the horizon, the acceleration diverges, and so do the temperature and the energy flux, but their ratio remains finite.

- In general, the horizon will be expanding, contracting or shearing. In order to impose equilibrium, we need the expansion, shear and rotation to be zero at first order in a neighbourhood of the horizon.

The introduction of an accelerated observer gives as a natural choice for the horizon the Rindler horizon associated to this accelerated observer.

The key idea to be shown can be expressed as [13]:

"In order to satisfy the thermodynamic equilibrium relation $\delta Q = TdS$, intrepreted in terms of the energy flux and area of local Rindler horizons, the gravitational lensing by matter energy must distort the causal structure of spacetime in a way that the Einstein equation holds."

The next step is to define precisely this local causal horizon. It can be done as follows:

By means of the equivalence principle, the neighbourhood of any point p can be thought as a piece of flat spacetime. Around p we consider a 2-dimensional surface $\mathcal{P}$. As usual, this 2-surface will be represented as a point in the conformal diagram. The boundary of the past of $\mathcal{P}$ has two components, each of which is a null surface generated by a congruence of null generators k^a orthogonal to $\mathcal{P}$. The local causal horizon is defined as one of these two components. We take λ as the affine parameter for k^a, in such a way that λ vanishes at $\mathcal{P}$ and is negative to the past of $\mathcal{P}$ (see Figure (4)).

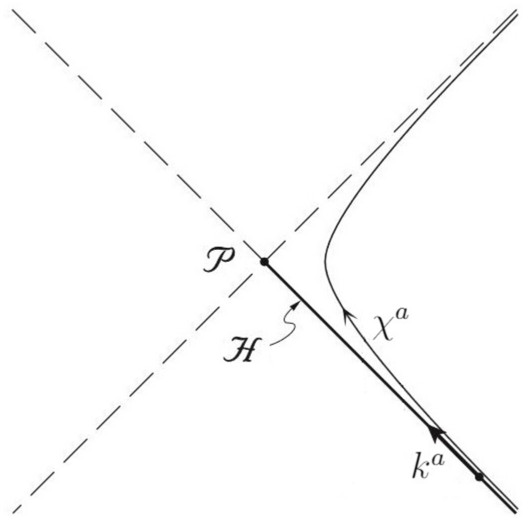

Figure 4: Rindler horizon $\mathcal{H}$ of a 2-sphere $\mathcal{P}$. The accelerated observer follows the trajectory of the Killing vector χ^a. k^a is the generator of the horizon.

In order to define the temperature and the heat, note that in the approximately flat region around p the usual Poincaré symmetries hold. In particular, there is an approximate Killing field χ^a generating boosts orthogonal to $\mathcal{P}$ and vanishing at $\mathcal{P}$. Because we are at very short distances, the Minkowski vacuum state (or any other state) is a thermal state with temperature $T = \hbar a/2\pi$ with respect to the boost hamiltonian, where a is the acceleration of this orbit. The heat flow is then defined through the boost-energy current of matter, $T_{ab}\chi^a$, where T_{ab} is the stress-energy tensor.

The Killing field defining the orbits of Rindler observers coincides at the null surface with the generators for sufficiently accelerated observers. Then, in the limit that the observer is sufficiently close to the horizon, the Killing field χ^a is parallel to the horizon generator k^a, and, at first order, we have that $\chi^a = -\kappa\lambda k^a$ and $d\Sigma^a = k^a d\lambda d\mathcal{A}$, where $d\mathcal{A}$ is the area element on a cross section of the horizon [6, 13, 14].

Then, the heat flux is given by

$$\delta Q = \int_{\mathcal{H}} T_{ab}\chi^a d\Sigma^b = -\kappa \int_{\mathcal{H}} \lambda T_{ab} k^a k^b d\lambda d\mathcal{A} \tag{3.1}$$

Assuming that the entropy is proportional to the area, we have that $dS = \eta\delta\mathcal{A}$, where $\delta\mathcal{A}$ is the area variation of a cross section of a pencil of generators of $\mathcal{H}$. For now, the constant η is left undetermined. The area variation is given by

$$\delta A = \int_{\mathcal{H}} \theta d\lambda d\mathcal{A} \tag{3.2}$$

where θ is the expansion of the horizon generators.

The expression $\delta Q = TdS \propto \delta A$ is telling that the presence of the energy flux is associated with a focussing of the horizon generators. Then, the Raychaudhuri equation (2.14) enters in the game, because it tells precisely the rate of focussing of the generators. The stationarity conditions imposed above imply that, at p, both the expansion and the shear vanish[1], and the Raychaudhuri equation simplifies to

$$\frac{d\theta}{d\lambda} = -R_{ab}k^a k^b \tag{3.3}$$

where the θ^2 and σ^2 are higher-order contributions that can be neglected when integrating to find θ around $\mathcal{P}$. For a small interval of λ, this integration is simply $\theta = -\lambda R_{ab}k^a k^b$. Then,

$$\delta A = -\int_{\mathcal{H}} \lambda R_{ab}k^a k^b d\lambda d\mathcal{A} \tag{3.4}$$

Now, from (3.4) and (3.1), we see that $\delta Q = TdS = \frac{\hbar\kappa}{2\pi}\eta\delta A$ is valid if

$$T_{ab} k^a k^b = \frac{\hbar\eta}{2\pi} R_{ab}k^a k^b \tag{3.5}$$

is valid for all null vectors k^a. This is equivalent to the tensorial equation

$$\frac{2\pi}{\hbar\eta} T_{ab} = R_{ab} + f g_{ab} \tag{3.6}$$

for some undetermined function f. The stress-energy momentum is divergence-free, which means that the rhs of (3.6) must also be divergence-free. This gives the constraint $f = -\frac{R}{2} + \Lambda$ for some undetermined constant Λ. Then, we find:

[1] It is always possible to find a 2-surface $\mathcal{P}$ so that both the expansion and shear vanish in a first order neighbourhood of p.

$$R_{ab} - \frac{1}{2}Rg_{ab} + \Lambda g_{ab} = \frac{2\pi}{\hbar\eta}T_{ab} \tag{3.7}$$

If $\eta = \frac{1}{4G\hbar}$, as the Bekenstein-Hawking entropy formula tells, we recover the Einstein equation:

$$\boxed{R_{ab} - \frac{1}{2}Rg_{ab} + \Lambda g_{ab} = 8\pi G T_{ab}} \tag{3.8}$$

Thus, the Einstein equation appears from the relation $\delta Q = TdS$, from a purely thermodynamic point of view.

4 Second derivation. Entanglement entropy

In this derivation, the Einstein equation appears as a consequence of a maximal vacuum entanglement hypothesis in a small region of spacetime. The main hypothesis can be expressed as [15]:

"When the geometry and quantum fields are simultaneously varied from maximal symmetry, the entanglement entropy in a small geodesic ball is maximal at fixed volume."

The system to consider now can be defined as follows:

Consider any point o of a spacetime of dimension d. If we choose a timelike unit vector u^a, we can generate a $(d-1)-$dimensional spacelike ball Σ of radius l if we consider all the geodesics of length l that leave p in all directions orthogonal to u^a. The point p is located at the center of the ball, and we call the surface of the ball $\partial\Sigma$. The region causally connected to the sphere Σ is called the causal diamond (see Figure (5)). We will consider that the radius l of the ball is much smaller than the characteristic radius of curvature of the spacetime in that region: $l \ll L_{curv}$.

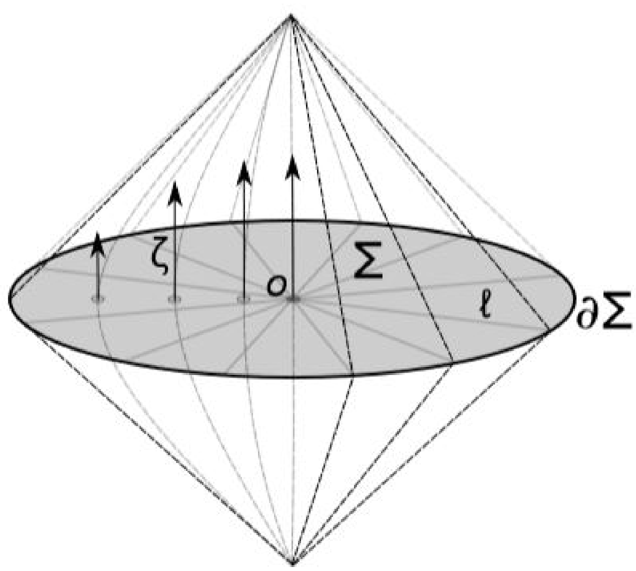

Figure 5: Causal diamond associated to a geodesic ball centered at o and geodesic radius l.

It is known that, at sufficiently short distances, all the fields look like the vacuum state. Moreover, if the condition $l \ll L_{curv}$ is satisfied, the spacetime around p can be treated as flat. Then, when we perform the variations with respect to the geometry and the quantum states, we will perform them with respect to flat space, and to the vacuum state.

A way of interpreting geometrically the Einstein equation is the following: in classical vacuum (without any matter source), any small geodesic ball of given volume has the same area as in flat spacetime. However, when there is a source of matter or energy (given by some expectation value of the stress-energy tensor), curvature causes a spatial ball of given volume to have a smaller surface area than it would have in flat spacetime. This area deficit can be computed at fixed geodesic radius, or at fixed volume. The expressions are given by [15]

$$\delta A\big|_l = -\frac{\Omega_{d-2} l^d}{6(d-1)}\mathcal{R} \tag{4.1}$$

$$\delta A\big|_V = -\frac{\Omega_{d-2} l^d}{2(d^2-1)}\mathcal{R} \tag{4.2}$$

where $\mathcal{R} = R_{ik}R^{ik}$ is the spatial Ricci scalar at p. Note that

$$\delta A\big|_l = \frac{d+1}{3}\delta A\big|_V \tag{4.3}$$

For convenience, as will be seen at the end, we will take the variations to be at fixed volume. To connect this expression with the Einstein equation, note that the spatial Ricci scalar can be related to the 00-component of the Einstein tensor as follows:

$$G_{00} = R_{00} - \frac{1}{2}Rg_{00} = \frac{1}{2}\left(R - 2R_0{}^0\right) = \frac{1}{2}R_{ik}{}^{ik} = \frac{1}{2}\mathcal{R} \tag{4.4}$$

Then, we can write

$$\delta A\big|_V = -\frac{\Omega_{d-2} l^d}{d^2-1}G_{00} \tag{4.5}$$

and, by virtue of the Einstein equation,

$$\delta A\big|_V = -\frac{8\pi G\Omega_{d-2} l^d}{d^2-1}T_{00} \tag{4.6}$$

Under a simultaneous variation of the geometry and the quantum fields, the variation of the entanglement entropy will have two contributions: a UV-contribution δS_{UV} from the area change when the metric is varied with respect to flat spacetime (δg_{ab}), and an IR-contribution δS_{IR} due to the variation of the fields ($\delta \left|\psi\right>$), so we can write

$$\delta S = \delta S_{UV} + \delta S_{IR} \tag{4.7}$$

We shall assume that the UV-part of the entanglement entropy is finite at leading order, and is proportional to the area variation computed above. That is, $\delta S_{UV} = \eta \delta A$. As in the previous derivation, the constant η is left undetermined until the end.

In order to compute δS_{IR}, we take into account that the vacuum state of any QFT, when restricted to the diamond, can be written as a thermal density matrix,

$$\rho = \frac{1}{Z} e^{-K/T} \tag{4.8}$$

where $T = \hbar/2\pi$, and K is the modular hamiltonian. From this thermal density matrix, the entropy variation can be computed and it is given by $\delta S_{IR} = \delta\langle K\rangle$. In general, K is not a local operator, and there is not a general expression for it. However, in the case of the vacuum of a conformal field theory (CFT), the situation is different. The diamond has a conformal boost Killing vector generating it (see Figure (5)), given by

$$\zeta = \frac{1}{2l}\left[\left(l^2 - u^2\right)\partial_u + \left(l^2 - v^2\right)\partial_v\right] \tag{4.9}$$

in null coordinates u, v, or

$$\zeta = \frac{1}{2l}\left[\left(l^2 - r^2 - t^2\right)\partial_t - 2rt\partial_r\right] \tag{4.10}$$

in the usual t, r coordinates. For the vacuum of a CFT, there is a conformal transformation relating the diamond to Rindler space and, in this case, K is equal to H_ζ, the Hamiltonian generating the flow of the above Killing vector [9, 10, 16], which means that

$$H_\zeta = \frac{2\pi}{\hbar}\int_\Sigma T^{ab}\zeta_b d\Sigma_a \tag{4.11}$$

With the previous Killing vector on the $t = 0$ surface, we obtain [10]

$$\delta\langle K\rangle = \frac{2\pi}{\hbar}\int \delta\langle T_{ab}\rangle \zeta^a d\Sigma^b = \frac{2\pi}{\hbar}\int d^{d-1}x \frac{l^2 - r^2}{2l}\delta\langle T_{00}\rangle \tag{4.12}$$

If we consider that $\delta\langle T_{00}\rangle$ is constant within the ball, it can be taken out of the integral, and we have

$$\begin{aligned}
\delta\langle K\rangle &= \frac{2\pi}{\hbar}\delta\langle T_{00}\rangle \int d^{d-1}x \frac{l^2-r^2}{2l} \\
&= \frac{2\pi}{\hbar}\delta\langle T_{00}\rangle \int d\Omega_{d-2}\int_0^l r^{d-2}\frac{l^2-r^2}{2l}dr \\
&= \frac{2\pi}{\hbar}\frac{\Omega_{d-2}l^d}{d^2-1}\delta\langle T_{00}\rangle
\end{aligned} \tag{4.13}$$

This result, together with the one providing the area variation at fixed volume, gives

$$\begin{aligned}
\delta S\big|_V &= \eta\delta A + \delta\langle K\rangle \\
&= \frac{\Omega_{d-2}l^d}{d^2-1}\left[-\eta G_{00} + \frac{2\pi}{\hbar}\delta\langle T_{00}\rangle\right]
\end{aligned} \tag{4.14}$$

Now, imposing the assumption that the entanglement entropy is maximal at fixed volume (that is, $\delta S\big|_V=0$), we obtain the relation

$$G_{00} = \frac{2\pi}{\hbar\eta}\delta\langle T_{00}\rangle \tag{4.15}$$

If we require this variation to vanish at all points and with all timelike unit vectors, we obtain a tensor equation,

$$G_{ab} = \frac{2\pi}{\hbar\eta}\delta\langle T_{ab}\rangle \tag{4.16}$$

This is the Einstein equation, provided we define the constant η to be $\eta = \frac{1}{4}\frac{1}{\hbar G}$, which is the precise value required by the Bekenstein-Hawking entropy formula.

For the non-CFT case, K is not given by (4.11), and some assumptions must be made in order to find an expression for $\delta\langle K\rangle$. The main conjecture [15] is to consider that $\delta\langle K\rangle$ is given by

$$\delta\langle K\rangle = \frac{\Omega_{d-2}l^d}{d^2-1}\left(\delta\langle T_{00}\rangle + \delta X\right) \tag{4.17}$$

where δX is a spacetime scalar, maybe related to the trace of T_{ab}. Calculations [17, 18] support this assumption, although it is still being investigated.

When the maximization of entropy is considered, one obtains

$$G_{ab} = \frac{2\pi}{\hbar\eta} \left(\delta\langle T_{ab}\rangle - \delta\langle X\rangle g_{ab}\right) \tag{4.18}$$

This result has a problem, because from the Bianchi identity, the lhs of (4.18) is divergence-free, and so is the term $\delta\langle T_{ab}\rangle$ because of energy-momentum conservation. This implies that $\nabla_a \delta\langle X\rangle = 0$ and, if it is related to the trace of T_{ab}, it is a too strong constraint.

This problem can be solved if, instead of comparing it to the Minkowski vacuum, the variations are compared to some other maximally symmetric spacetime (MSS), because any MSS seems as good candidate for the vacuum as flat spacetime. The Einstein tensor in a MSS of curvature scale λ is given by $G_{ab}^{MSS} = -\lambda g_{ab}$. When the area variation is compared to this MSS, the area variation at fixed volume is given by the same expression as before, but replacing G_{00} by $G_{00} - G_{00}^{MSS}$. The variation of entropy reads now

$$\begin{aligned} \delta S\big|_V &= \eta \delta A\big|_V + \frac{2\pi}{\hbar}\delta\langle K\rangle \\ &= \frac{\Omega_{d-2} l^d}{d^2-1}\left[-\eta\left(G_{00} + \lambda g_{00}\right) + \frac{2\pi}{\hbar}\left(\delta\langle T_{00}\rangle + \delta X\right)\right] \end{aligned} \tag{4.19}$$

Again, when we consider that the variation vanishes at all points and with all timelike unit vectors, the equation becomes a tensorial equation,

$$G_{ab} + \lambda g_{ab} = \frac{2\pi}{\hbar\eta}\left(\delta\langle T_{ab}\rangle - \delta X g_{ab}\right) \tag{4.20}$$

Taking the divergence of this equation, the term of the Einstein tensor and the term of the stress-energy tensor vanish, because of the Bianchi identity and the conservation of energy-momentum, respectively. Then, we obtain a constraint between λ and δX:

$$\Lambda = \frac{2\pi}{\hbar\eta}\delta X + \lambda \tag{4.21}$$

where Λ is a spacetime constant. When this relation is plugged into (4.20), we obtain

$$G_{ab} + \Lambda g_{ab} = \frac{2\pi}{\hbar\eta}\delta\langle T_{ab}\rangle \tag{4.22}$$

This is the Einstein equation with a cosmological constant Λ, provided that, again, $\eta = \frac{1}{4}\frac{1}{\hbar G}$, in agreement with the Bekenstein-Hawking entropy. Thus, the Einstein equations have been derived from an entanglement entropy hypothesis.

5 Comments and discussion

This section contains a discussion about some issues related to the derivations, together with some of the main questions that arise from this new interpretation of spacetime, and the possible answers (more or less satisfactory) that can be given with the current knowledge of physics.

- **Why are the variations taken at fixed volume instead of at fixed geodesic radius?**

In the derivation based on entanglement entropy, we have taken the variations at fixed volume "for convenience". We argue here why this has been done.

First of all, notice that we have obtained the desired result because the geometric term $\Omega_{d-2}l^d/(d^2-1)$ that appears as a prefactor in both variations is the same for δS_{UV} and δS_{IR}, and it can be factorized. Had we taken the variations at fixed geodesic radius instead of fixed volume, the terms would have not been the same.

But there are other arguments to take the variations at fixed volume. The first law of causal diamonds is a variational identity, analogous to the first law of black hole mechanics, which relates variations, away from flat spacetime, of the area, volume, and cosmological constant inside the diamond. The first law reads [19]

$$-\frac{\kappa k}{8\pi G}\delta V+\frac{V_\zeta}{8\pi G}\delta\Lambda=T\delta S_{gen} \tag{5.1}$$

where κ is the surface gravity of the Killing horizon, k is the trace of the outward extrinsic curvature of the boundary $\partial\Sigma$ when embedded in Σ, and $V_\zeta\equiv\int_\Sigma|\zeta|dV$. Λ is the cosmological constant, T is minus the Hawking temperature, and S_{gen} is defined as the sum of the horizon entropy and the entanglement entropy of matter. At fixed volume and fixed cosmological constant, the first law of causal diamonds implies that the entropy is stationary when varied away from the vacuum, as it has been considered in the derivation.

- **What is the best way to proceed if we want to find a quantum theory of gravity?**

The fact that spacetime dynamics can be derived from thermodynamic arguments suggests the possibilty that gravity is not a fundamental force, but a macroscopic result of some microscopic degrees of freedom of spacetime [20–22]. These degrees of freedom have been called by some authors as "Atoms of Spacetime", in analogy to the standard relation between thermodynamics and statistical mechanics.

If this is the case, it explains why the quantization of General Relativity has shown to be much more problematic than for other microscopic forces.

In [21, 22], some properties that these atoms of spacetime should have are discussed, and with a particular model of atoms of spacetime for the geometric part of the action, the Einstein's equations are recovered from a purely thermodynamic argument. Other works [23, 24] have used particular models of microstructure to recover the Hawking temperature and entropy for black holes.

- **Could we obtain higher-curvature corrections to the Einstein's equations with the thermodynamic interpretation?**

The classical Einstein's equations,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{5.2}$$

are derived from the Einstein-Hilbert action,

$$S = \frac{1}{16\pi G}\int \sqrt{-g}R\, d^4x \tag{5.3}$$

However, the Einstein-Hilbert gravity can be treated as a low-energy effective theory, so we should expect to have corrections to this action, of the form

$$S = \frac{1}{16\pi G}\int \sqrt{-g}\left(R + \alpha_1\Lambda + \alpha_2 R^2 + \alpha_3 R_{\mu\nu}R^{\mu\nu} + \alpha_4 R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} + ...\right) d^4x \tag{5.4}$$

and the field equations arising from this action would contain higher-curvature terms. These terms include higher derivatives of the metric, which correspond to terms with higher and higher curvature (and a lower and lower associated curvature radius). At some point, this curvature radius is of the order of the Planck length. Thus, in a theory of quantum gravity, we expect these terms to be important. The question that arises now is: we have obtained the classical Einstein's equations from a thermodynamic point of view. If the spacetime is really a thermodynamic entity, should we be able to obtain these higher-curvature terms in the field equations with a similar argument?

There is not a clear answer to this question. In the derivation of Section (4), in the expression for the area deficit we have neglected terms of order l/L_{curv}, while the next-higher-curvature correction to the field equations might be of order $(l_1/L_{curv})^2$, with l_1 a length scale appearing in the corresponding term in the action. To obtain this next-order term in the field equations, we need $l/L_{curv} < (l_1/L_{curv})^2 \;\Rightarrow\; l/l_1 <$

l_1/L_{curv}. The rhs must be smaller than 1 (otherwise, the higher-order terms would dominate), which means that we need $l < l_1$. That is, the diamond must be smaller than l_1. If l_1 is, for instance, the Planck length, the diamond should be smaller than the Planck length, and the classical geometry and quantum field theory used in the derivation would not work in that regime.

There have been some attempts to find the field equations when these corrections are considered [25–27], but because of the presence of these terms, some technical difficulties appear and it is required to use non-equilibrium thermodynamics.

However, an interesting result has been found in [28]. There, they show that, for spherically symmetric systems with a horizon, the Einstein equations arising from the Einstein-Hilbert action can be put in the form of the relation $TdS = dE + PdV$, matching the entropy S and the energy E with the already know expressions. They go one step beyond, and do the same for the first correction to the Einstein-Hilbert action, the so-called Gauss-Bonnet correction:

$$S = \frac{1}{16\pi G}\int \sqrt{-g}\,(R + \alpha\mathcal{L}_{GB})\,d^4x, \quad \mathcal{L}_{GB} = R^2 - 4R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} \tag{5.5}$$

finding again that, once the field equations are written in the form $TdS = dE + PdV$, the entropy and the energy match with the expressions obtained by other authors. Finally, they generalize this result to the complete Lanczos-Lovelock action in D dimensions, matching again the results for S and E with independent calculations.

These results suggest that the thermodynamic route to obtain the field equations also works for higher-curvature theories of gravity, and the quantum corrections to the Einstein-Hilbert action appear as quantum corrections for the entropy and the energy [11]. However, the microscopic structure beyond this thermodynamics remains mysterious.

- **Is it appropiate to consider the entanglement entropy to be finite at UV scales?**

As we have seen, for a d-dimenional free field theory the entanglement entropy is a UV-divergent quantity, with the leading term being proportional to the area of the entangling surface, which we rewrite here for convenience:

$$S_{\mathcal{A}} = \gamma\frac{\text{Area}(\partial\mathcal{A})}{\epsilon^{d-2}} + ... \tag{5.6}$$

where ϵ is the cutoff length that is sent to 0 in the continuum limit and originates the divergences. But, because of the fact that the fields that contribute more to the entanglement entropy are those of high energy, we expect to have modifications to the background geometry. This backreaction of spacetime may lead to a finite entropy. Susskind and Uglum show in [29] that when these are considered, the divergences that appear are the same ones that appear in the renormalization of the gravitational constant G. When this renormalization process is done, one obtains a finite result for the entanglement entropy, with the leading term coinciding with the Bekenstein-Hawking entropy [30, 31]:

$$S_{\mathcal{A}} = \frac{1}{4} \frac{\text{Area}(\partial \mathcal{A})}{G_R \hbar} + ... \tag{5.7}$$

where G_R is the renormalized gravitational constant. Again, in order to understand the microscopic origin of the entanglement entropy and its divergences, a microscopic understanding of the theory is needed.

The great similarity between the entanglement entropy and the Bekenstein-Hawking entropy suggests that, maybe, black hole entropy can be originated purely from entanglement [5, 11]. This identification, however, can only be done once the divergences of entanglement entropy are properly solved, because the Bekenstein-Hawking entropy has no divergences.

6 Conclusions

After the realization of the thesis, some conclusions can be extracted:

- The analogy between thermodynamics and black holes found by Bekenstein and Hawking 50 years ago can be extended to spacetime itself. That is, spacetime is a thermodynamic entity.

- Given this thermodynamic nature of spacetime, one should be able to derive the Einstein equations from thermodynamic arguments. This has been done in two different ways: in Section (3), the Einstein equations have been derived from the thermodynamic relation $\delta Q = TdS$ near a Rindler horizon, and in Section (4) they have been derived from an hypothesis about entanglement entropy.

- This thermodynamics suggests some kind of microstructure of spacetime at smallest scales. Although there is not fully satisfactory explanation of what this microstructure is, it seems reasonable that it will be important when trying to find a complete quantum theory of gravity.

References

[1] J.D. Bekenstein. *Generalized second law of thermodynamics in black-hole physics*. Phys. Rev. **D 9** 3292 (1974).

[2] S.W. Hawking. *Particle creation by black holes*. Commun. Math. Phys. **43** 199 (1975).

[3] W.G. Unruh. *Notes on black-hole evaporation*. Phys. Rev. **D 14**, 870 (1976).

[4] P.C.W. Davies. *Scalar Production in Schwarzschild and Rindler Metrics*. J. Phys. A.: Math. Gen. **8** 609 (1975).

[5] J.D. Bekenstein. *Do we understand black hole entropy?* (1994). ArXiv: gr-qc/9409015.

[6] T. Jacobson; R. Parentani. *Horizon entropy*. Found. Phys. **33** 323-348 (2003). ArXiv: gr-qc/0302099.

[7] S. Carroll. *Spacetime and geometry. An introduction to General Relativity*. Addison-Wesley (2004).

[8] T. Takayanagi M. Rangamani. *Holographic entanglement entropy*. (2017) ArXiv: hep-th/1609.01287.

[9] R. Myers D. Blanco; H. Casini; L. Hung. *Relative entropy and holography*. JHEP **2013** 60 (2013). ArXiv: hep-th/1305.3182.

[10] H. Casini; M. Huerta; R. Myers. *Towards a derivation of holographic entanglement entropy*. JHEP **1105** 036 (2011). ArXiv: hep-th/1102.0440.

[11] S. Solodukhin. *Entanglement entropy of black holes*. Living Rev. Rel. **14**, 8 (2011). ArXiv: hep-th/1104.3712.

[12] M. Srednicki. *Entropy and Area*. Phys. Rev. Lett. **71** 666-669 (1993). ArXiv: hep-th/9303048.

[13] T. Jacobson. *Thermodynamics of Spacetime: the Einstein Equation of State*. Phys. Rev. Lett. **75** 1260 (1995). ArXiv: gr-qc/9504004.

[14] D. Kothawala. *The thermodynamic structure of Einstein tensor*. Phys. Rev. **D 83**, 024026 (2011). ArXiv: gr-qc/1010.2207.

[15] T. Jacobson. *Entanglement Equilibrium and the Einstein Equation*. Phys. Rev. Lett. **116** 201101 (2016). ArXiv: gr-qc/1505.04753.

[16] D. Blanco; H. Casini; M. Leston; F. Rosso. *Modular energy inequalities from relative entropy*. JHEP **01** 154 (2018) ArXiv: hep-th/1711.04816.

[17] A.J. Speranza. *Entanglement entropy of excited states in conformal perturbation theory and the Einstein equation*. JHEP **04** 105 (2016). ArXiv: hep-th/1602.01380.

[18] H. Casini; D.A. Galante; R.C. Myers. *Comments on Jacobson's Entanglement equilibrium and the Einstein equation*. JHEP **03** 194 (2016). ArXiv: hep-th/1601.00528.

[19] T. Jacobson; M. Visser. *Gravitational Thermodynamics of Causal Diamonds in (A)dS*. (2019). ArXiv: hep-th/1812.01596.

[20] A. Granger. *Thermodynamic Gravity and the emergence of space with geometry*. Imperial College London (2010).

[21] T. Padmanabhan. *Atoms of Spacetime and the Nature of Gravity*. J. Phys. **701** 012018 (2016).

[22] T. Padmanabhan. *Atoms of Spacetime and the Nature of Gravity*. Entropy **17** 7420 (2015). ArXiv: gr-qc/1508.06286.

[23] J. Mäkelä. *Partition Function of the Schwarzschild Black Hole*. Entropy **2011**, 13, 1324-1354.

[24] R. Raitio. *A Statistical Model of Spacetime, Black Holes and Matter*. Open Access Library Journal, **3**: e2487.

[25] C. Elling; R. Guedens; T. Jacobson. *Non-equilibrium Thermodynamics of Spacetime*. Phys. Rev. Lett. **96** 121301 (2006).

[26] R. Guedens; T. Jacobson; S. Sarkar. *Horizon entropy and higher curvature equations of state*. Phys. Rev. **D 85** 064017 (2011). ArXiv: gr-qc/1112.6215.

[27] T. Jacobson. *Gravitation and vacuum entanglement entropy*. Int. J. Mod. Phys. **D21** 1242006 (2012). ArXiv: gr-qc/1204.6349.

[28] A. Paranjape; S. Sarkar; T. Jacobson. *Thermodynamic route to field equations in Lanczos-Lovelock Gravity*. Phys. Rev. **D 74** 104015 (2006). ArXiv: hep-th/0607240.

[29] L. Susskind; J. Uglum. *Black hole entropy in canonical quantum gravity and superstring theory*. Phys. Rev. **D 50** 2700-2711 (1994). ArXiv: hep-th/9401070.

[30] J.H. Cooperman; M.A. Luty. *Renormalization of entanglement entropy and the gravitational effective action*. J. High Energy Phys. **2014** 45 (2014). ArXiv: hep-th/1302.1878.

[31] F. Larsen; F. Wilczek. *Renormalization of black hole entropy and of the gravitational coupling constant*. Nucl. Phys. **B 458** 249 (1996). ArXiv: hep-th/9506066.