FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

DOCTORAL THESIS

Marek Liška

Thermodynamics of spacetime: corrections from the quantum realm

Institute of Theoretical Physics

Supervisor of the doctoral thesis: Ana Alonso-Serrano, Ph.D.
Study programme: Theoretical Physics, Astronomy and Astrophysics

Prague 2024

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date ....

Author's signature

Let me be truthful, just once, but nonetheless
In this short piece that does not quite belong
With the rest. All that follows is worthless
It doesn't really matter if it's right or wrong

What's important are only those few moments
In which, after weeks of confused meandering,
I glimpsed, in lines of math, a few fragments
Of something unearned, unlooked for, something

That went beyond the drudgery of equations
And, if not true, was truly beautiful
In those moments my thoughts were revelations
So wondrously childlike, simple and cheerful

And my mind, so exhausted and exposed
Danced in exhilaration, all unease gone
I knew then of truths never before proposed
And I knew I'm close, but never could go on

This is all, I have nothing more to tell
For my mind fails to describe where it went
So, stop reading, before you break the spell
And just find for yourself such a moment.

I thank Luis J. Garay and Antonio Vicente-Becerril for their very significant contributions to the research presented in this thesis. I also thank Lucía Menéndez-Pidal, Gerardo García-Moreno, Julio Arrechea Rodríguez, Carlos Barceló Serón, Valentin Boyanov, Javier Olmedo and Robie A. Hennigar for helpful discussions. I gratefully acknowledge the support of the Charles University Grant Agency projects No. GAUK 297721 and No. GAUK 90123. Furthermore, I acknowledge the hospitality of the Max Planck Institut für Gravitationsphysik, Instituto de Astrofísica de Andalucía, Universidad Complutense de Madrid, and Institut de Ciències del Cosmos, Universitat de Barcelona during various phases of the preparation of the thesis.

On a more personal note, I would like to thank the department's secretary Lenka Knotková for the resourcefulness and nearly limitless patience she displayed whenever I turned to her for help. I further thank my thesis advisor Oldřich Semerák for his continuous support and, especially, for introducing me to general relativity in one of the few truly flawless classes I ever frequented. I am also grateful to other teachers who have helped forming my way of thinking about physics; Jiří Podolský, Pavel Krtouš, Jiří Novotný, David Kubizňák and, especially, the ones are unfortunately no longer with us; Martin Scholtz, Jiří Langer and Jiří Bičák.

Of course, my deepest gratitude goes to Ana Alonso-Serrano who has been my supervisor for nearly seven years, co-operating with me on three theses and (so far) ten peer-reviewed publications. While I deeply appreciate Ana as a creative yet rigorous scientist, my admiration for her abilities as a mentor is greater still. I hope I will one day be able to at least approach her level of excellence in this role with my own students.

Lastly, I am grateful to all my friends, especially Yuliia, Eliška, Lenka, Petr, David, Dominika, Yelizaveta and Ondra. I probably do not make my appreciation of you as clear as I ought to, but I could not survive the last four years without your friendship.

Title: Thermodynamics of spacetime: corrections from the quantum realm

Author: Marek Liška

Institute: Institute of Theoretical Physics

Supervisor: Ana Alonso-Serrano, Ph.D., Humboldt University of Berlin and Max Planck Institute for Gravitational Physics

Advisor: Doc. RNDr. Oldřich Semerák, DSc., Institute of Theoretical Physics

Abstract: This thesis explores the deep relation between gravitational dynamics and thermodynamics. It focuses on the notion that the gravitational dynamics is encoded in the thermodynamic equilibrium conditions applied to locally constructed, observer-dependent causal horizons. At the level of the (semi)classical gravity, we show that this approach leads to Weyl transverse gravity. The classical solutions of this gravitational theory are the same as those of general relativity, but it has different local symmetries, which leads to a different origin of the cosmological constant. As a useful computational tool for the (semi)classical part of the thesis, we also develop covariant phase space formalism for Weyl transverse gravity and identify the corresponding Wald entropy. We also generalise the formalism to arbitrary gravitational theories with the same symmetry group. Moreover, we explore the implications of the equivalence between gravitational dynamics and local equilibrium conditions for the low energy quantum gravitational dynamics. At the linearised level, we find a result consistent with quadratic gravity. However, beyond the linearised regime the dynamics becomes significantly different from quadratic gravity. We derive the equations for gravitational dynamics that capture this difference and study their physical implications on the example of the homogeneous, isotropic cosmologies. The thesis is conceived as a self-contained introductory text into the relationship between thermodynamics and Weyl transverse gravity, as well as into our program to search for the implications of thermodynamics for quantum gravitational dynamics.

Keywords: thermodynamics of spacetime; causal diamonds; quantum gravity; unimodular gravity; Weyl transverse gravity; covariant phase space; Wald entropy; equivalence principle

Contents

Introduction

The main subject of this thesis is the recovery of the equations governing the gravitational dynamics by thermodynamic reasoning. This area of research arose in the mid nineties Jacobson [1995]. However, it is based on the systematic exploration of phenomena lying at the intersection of gravity, quantum physics and thermodynamics dating back to the early seventies. Gravity indeed shows a deep interconnection with thermodynamics, which is not yet fully understood and remains an active subject of exploration. It first emerged with the realisation that entropy of any matter falling inside a black hole apparently vanishes, breaking the second law of thermodynamics Wheeler and Ford [1998]. Moreover, a proposal to use a black hole in a construction of a *perpetuum mobile* was put forward Geroch [1971], Bekenstein [1973], leading to another challenge to the second law, which forbids a cyclic process operating with no energy loss. Such potential conflicts with the fundamental laws of thermodynamics were addressed by the idea that black holes possess entropy proportional to the area of a spatial cross-section of their event horizon Bekenstein [1973]. Then, the increase of black hole entropy due to absorption of matter restores the validity of the second law, as the total entropy of the universe does not decrease. A more detailed analysis also showed that assigning entropy to a black hole makes the construction of a *perpetuum mobile* impossible[1]. Thus, black holes are compatible with the second law of thermodynamics. Along the proposal of black hole entropy, an equation governing a small, stationary perturbation of a black holes spacetime was derived, relating the change in the black hole's energy to the change of its entropy Bekenstein [1973], Bardeen et al. [1973]. For a Schwarzschild black hole, it takes the form $\Delta M = \Theta \Delta S$, where M denotes energy, S entropy, and Θ is a function constant on the horizon and proportional to its surface gravity Bekenstein [1973]. If the function Θ could be interpreted as the black hole temperature, this equation would become a genuine first law of thermodynamics relating the change of the total energy to the heat transfer term $T\Delta S$. This interpretation was confirmed soon afterwards, with the realisation that, when one takes into account the quantum effects, black holes emit a black body radiation corresponding precisely to the required temperature Hawking [1975]. Therefore, a black hole can be seen as a genuine thermodynamic object that possesses a well defined entropy and temperature and follows the standard laws of thermodynamics. Notably, quantum physics plays a crucial role in completing the thermodynamic description of the black hole's horizon, making it a phenomenon at the intersection of quantum and gravitational physics. Exploiting this unique position of gravitational thermodynamics is one of the principal aims of the present text.

Thermodynamics of the black hole horizons thus provides a connection between gravitational dynamics, which determines the form of the first law, and quantum physics, which provides a notion of black hole temperature. Follow-

[1] Both statements are somewhat subtle and depend on assuming some reasonable properties of the matter (essentially the null energy condition and the Bekenstein entropy bound). See, e.g. Wald [2001] for a short review and further references.

ing the seminal papers on the subject, the thermodynamic description has been extended to other types of horizons, e.g. the de Sitter cosmological horizon Gibbons and Hawking [1977] and even locally constructed, observer-dependent horizons Unruh [1976], Bombelli et al. [1986], Jacobson and Parentani [2003], Barbado and Visser [2012], Baccetti and Visser [2014], Chakraborty et al. [2016], Jacobson and Visser [2019a, 2023a,b].

Extending the thermodynamic description to local horizons opened up a different perspective on the connection between gravity and thermodynamics. In particular, two decades after the inception of gravitational thermodynamics, it was realised that thermodynamic equilibrium conditions applied to local, approximate Rindler (acceleration) horizons constructed in every regular spacetime point imply the Einstein equations of gravitational dynamics Jacobson [1995]. In other words, thermodynamics actually encodes the information necessary to reconstruct gravitational dynamics. The idea of the thermodynamic reconstruction of the gravitational dynamics lies at the heart of this thesis. Hence, we find it worthwhile to recount the principal ideas and assumptions involved, emphasising the conceptual issues rather than the mathematical details.

To begin our exposition, we note that an approximate Rindler horizon can indeed be constructed in an arbitrary regular spacetime point, P, provided that we only consider a part of the wedge much smaller than the local curvature length scale (an inverse of the square root of the largest eigenvalue of the Riemann tensor). Then, constructing an approximately flat coordinate system in the vicinity of P (e.g. Riemann normal coordinates) allows us to identify approximate worldlines of uniformly accelerating observers and the corresponding Rindler horizon. We show the details of the construction of a local Rindler horizon in figure 1. The equations governing the gravitational dynamics are then encoded in the Clausius equilibrium relation, which equates a small change in entropy of the Rindler horizon to the heat flux divided by the temperature of the system.

Thence, to discuss the thermodynamics of local Rindler horizons, we first require a suitable notion of temperature. Due to the Unruh effect, an eternally uniformly accelerating observer with acceleration a, who perceives the Rindler horizon, sees the Minkowski vacuum as a thermal bath of particles at a temperature $T_{\text{U}} = a/2\pi$. While the observers associated with an approximate Rindler wedge do not accelerate precisely uniformly and can be well defined only in finite regions, the Unruh effect still applies under certain conditions Chirco and Liberati [2010], Barbado and Visser [2012], Baccetti and Visser [2014], Fewster et al. [2016], Shevchenko [2017], Rick Perche [2021, 2022]. First, the local approximate Minkowski vacuum state must be well defined. The sufficient condition is the Einstein equivalence principle, which states that all the non-gravitational test physics behaves locally as it would in the absence of gravity di Casola et al. [2015]. Then, the Minkowski vacuum can be locally defined in the usual way, as the state invariant under the (approximate) Poincaré group associated with the local inertial system, we chose Chirco and Liberati [2010]. Second, while the observers we study do not have an exactly uniform acceleration and the approximation we consider breaks down in finite proper time, they still may perceive the Minkowski vacuum as a thermal bath of particles at the Unruh temperature. The necessary condition is that the acceleration magnitude is much greater than both the variations in the acceleration and the inverse of the finite proper time interval. Only

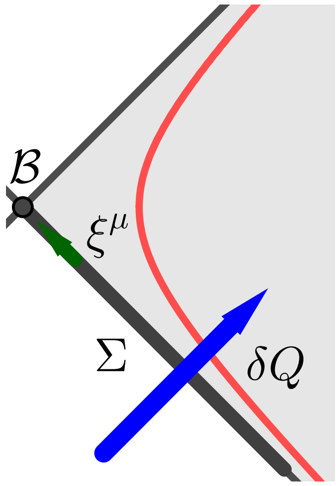

Figure 1: The construction of the local approximate Rindler wedge. Its centre is the bifurcate spacelike $(n-2)$-surface $\mathcal{B}$ denoted by a small black circle. The red line shows a sample trajectory of a uniformly accelerating observer who perceives the horizon. The thick black line represents a part of one branch of the past causal horizon Σ. This part must be chosen small enough to be viewed as an approximate Rindler horizon. The choice of the past boundary of Σ is arbitrary. The blue arrow shows the physical direction of the heat flux across the horizon, whereas the green arrow is the null normal to Σ.

the second requirement matters in our setup, since the acceleration variations occur only due to spacetime curvature and we must anyway choose our Rindler wedge much smaller than the local curvature length scale.

To complete the right hand side of the Clausius equilibrium relation, we need a definition of the heat flux crossing the horizon. For a timelike $(n-1)$-surface $\mathcal{S}$ (n denotes the dimension of the spacetime), the heat flux measured by a uniformly accelerating observer with velocity v^μ tangent to $\mathcal{S}$ is simply defined as an integral of the energy-momentum tensor

$$\delta Q = -\int_{\mathcal{S}} T^{\mu\nu} v_\mu \mathrm{d}\mathcal{S}_\nu. \tag{1}$$

For a null surface, i.e., in the limit of v^μ becoming light-like, the heat flux diverges. However, the acceleration a and, therefore, the Unruh temperature measured by the observer with velocity v^μ also becomes divergent in this limit. It turns out one may rigorously compute the null limit of the ratio $\delta Q/T_{\mathrm{U}}$, obtaining a finite result Baccetti and Visser [2014]. Incidentally, the limit of infinite acceleration also ensures that all the approximations required to invoke the Unruh effect can always be done, as they rely on the acceleration being sufficiently large.

For the left hand side of the Clausius relation, we require an expression for the change of entropy of the horizon due to the heat flux across it. Several standard methods for computing entropy associated with the presence of a causal horizon have been developed, most notably the Wald entropy formula Wald [1993], Iyer and Wald [1994], the Euclidean canonical ensemble construction Gibbons and Hawking [1977], Jacobson and Visser [2023a,b] or the vacuum entanglement entropy paradigm Bombelli et al. [1986], Srednicki [1993], Solodukhin [2011]. Each method prescribes the same entropy to any causal horizon, regardless of whether it is the event horizon of a black hole or a locally constructed, observer-dependent horizon. The key point is that the entropy computation is performed in the spacetime region causally accessible to the given observer, which is naturally bounded by the horizon(s). The correct microscopic interpretation of this entropy associated with a causal horizon remains an open problem, which we discuss in chapter 2. The upshot of this discussion is that any causal horizon possesses entropy proportional, to the leading order, to the area of its spatial cross-section, i.e., $S = \eta\mathcal{A}$. The proportionality constant η cannot be determined without specifying the gravitational dynamics.

At this point, we have all the ingredients necessary to reconstruct the equations governing the gravitational dynamics. If the local approximate Rindler wedge is in thermodynamic equilibrium, the Clausius equilibrium relation implies $\eta\delta\mathcal{A} = \delta Q/T_{\mathrm{U}}$, where the heat flux is given by the integral of the energy-momentum tensor over the horizon. The change in the horizon area $\mathcal{A}$ is encoded in the expansion of the null congruence forming it. Using the Raychaudhuri equation Raychaudhuri [1955] to compute the expansion we find it to be proportional to the Ricci tensor Jacobson [1995]. The Clausius equilibrium relation is then equivalent to the traceless equations governing the dynamics of the spacetime metric, which are valid locally in the point P[2]

$$R_{\mu\nu}(P) - \frac{1}{n}R(P)\,g_{\mu\nu}(P) = \frac{2\pi}{\hbar\eta}\left(T_{\mu\nu}(P) - \frac{1}{n}T(P)\,g_{\mu\nu}(P)\right). \tag{2}$$

The equivalence principle guarantees that the same equations can be derived in this way at any regular spacetime point and, therefore, hold throughout the spacetime. The Newtonian limit of these equations allows us to identify the Newton gravitational constant G as $G = 1/(4\hbar\eta)$. Crucially, G is *defined* in terms of the entropy proportionality constant η and the Planck constant $\hbar$. Notice that, the entropy density of any causal horizon has a universal constant value $\eta = 1/4l_{\mathrm{P}}^2$, which agrees with the Bekenstein entropy formula Bekenstein [1973], Hawking [1975], Wald [2001] valid for black hole horizons in general relativity.

Finally, if we *postulate* that the energy-momentum tensor is divergenceless, we recover the Einstein equations

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{3}$$

with Λ appearing as an arbitrary integration constant. Notice that, while $\hbar$ disappears from the final gravitational equations, it is a necessary ingredient in the derivation. The reason is the role of the quantum Unruh effect in defining the Clausius entropy flux.

All the key physical insights of the original thermodynamic derivation Jacobson [1995] have been confirmed by later works, although some more technical points required further refinement Chirco and Liberati [2010], Baccetti and Visser [2014]. The derivation has even been extended beyond general relativity to a class of local, purely metric theories of gravity Eling et al. [2006], Padmanabhan [2010], Guedens et al. [2012], Bueno et al. [2017], Parikh and Svesko [2018], Svesko [2019]. The derivation then relies on identifying the entropy of local causal horizons with Wald entropy. However, the details are rather subtle and we reserve their discussion for chapter 2.

[2]Throughout the thesis, we set $c = 1$ for simplicity.

Of course, we do not need to make any reference to the Unruh effect or local equilibrium conditions to derive the Einstein equations (for a slightly dated, but nevertheless impressive list of the various approaches to their derivation, see Misner et al. [2017]). Then, why should we be interested in the thermodynamic approach which relies on extra assumptions? Is it anything more than a curiosity? The answer proposed in the seminal paper and extended in a number of follow-up works is that gravity is emergent Jacobson [1995], Padmanabhan [2010], Verlinde [2011], Svesko [2019]. In this paradigm, the gravitational interaction emerges in a suitable thermodynamic limit from the behaviour of some fundamental quantum spacetime degrees of freedom, in principle unrelated to the metric. Then, the Einstein equations have the status of equations of state and gravitons acquire a similar role as phonons, being merely a convenient way to describe excitations of the underlying quantum spacetime structure. This emergence interpretation of gravitational dynamics circumvents the issue of having to quantise the Einstein-Hilbert action and has proven to be stimulating for new ideas regarding the (quantum) nature of the spacetime Padmanabhan [2010], Verlinde [2011], Kothawala [2013], Kothawala and Padmanabhan [2014, 2015], Padmanabhan [2020], Padmanabhan and Chakraborty [2022]. However, the fact that the Einstein equations can be derived from thermodynamics in no way implies that gravity is emergent. To even formulate the local equilibrium conditions, one already requires a classical curved spacetime, which would presumably only arise in the thermodynamic limit applied to the quantum degrees of freedom making up the spacetime. Then, the derivation is not really in a position to tell us anything about the existence of such a limit, much less about whether the Einstein equations can emerge from it. Moreover, it has been shown that the thermodynamic derivations are consistent with loop quantum gravity, which is a manifestly non-emergent approach to quantum gravitational dynamics Chirco et al. [2014].

In this thesis, we show that the local equilibrium conditions allow us to learn a lot about gravity even without embracing the emergence scenario. Instead, we shall proceed with a more modest requirement that the local equilibrium conditions *encode* the gravitational dynamics, or at least its most important features. In other words, we consider them to be the consistency conditions that any candidate theory of quantum gravity should ultimately recover. This way of thinking is similar in spirit to the expectation that quantum gravity ought to reproduce Bekenstein entropy of black holes in a suitable limit. However, the recovery of Bekenstein entropy only allows us to test the already proposed candidate theories. In contrast, as we will show, the local equilibrium conditions make it possible to develop new predictions relevant independently of the microscopic nature of gravity.

In particular, we report on two main branches of our research into the relationship between gravitational dynamics and local equilibrium conditions. One branch consists of re-examining the implications of thermodynamics of spacetime for the (semi)classical gravity, the other one in employing it to gain insights into the quantum gravitational dynamics.

In regards to the (semi)classical gravitational dynamics, we note a somewhat awkward point occurring in the majority of the thermodynamic derivations Jacobson [1995], Padmanabhan [2010], Chirco and Liberati [2010], Jacobson

[2015], Parikh and Svesko [2018]; the need to introduce a divergenceless energy-momentum tensor as an extra assumption. We argue that one should instead treat the traceless equations (2) as the final description of the gravitational dynamics implied by the local equilibrium conditions (together with the equivalence principle). These equations are actually equivalent to the equations of motion of Weyl transverse gravity Unruh [1989], Finkelstein et al. [2001], Tiwari [2006], Álvarez et al. [2006], Barceló et al. [2014], Carballo-Rubio [2015], Álvarez et al. [2016], Barceló et al. [2018], Carballo-Rubio et al. [2022]. In other words, if one assumes that the local equilibrium conditions and the (strong) equivalence principle encode all the information about the classical gravitational dynamics, the resulting theory is Weyl transverse gravity rather than general relativity Alonso-Serrano and Liška [2020a, 2022], Alonso-Serrano et al. [2024].

Weyl transverse gravity has the same classical solutions as general relativity, but its local symmetries are the spacetime volume preserving diffeomorphisms and Weyl rescaling of the metric. It has been suggested that this difference in the symmetry group partially addresses the so called cosmological constant problem which arises in general relativity Unruh [1989], Carballo-Rubio [2015], Barceló et al. [2018] (the key point is that vacuum energy does not couple to gravity). To develop a full understanding and control over the relationship between the local equilibrium conditions and Weyl transverse gravity, we found it necessary to first study the thermodynamics of this theory. We chose to apply the covariant phase space formalism, which allows a rigorous derivation of the first law of horizon thermodynamics and a (heuristic) identification of black hole entropy, known as Wald entropy Lee and Wald [1990], Wald [1993], Iyer and Wald [1994], Iyer [1997], Wald and Zoupas [2000]. Since the covariant phase space formalism in its original form only applies to fully diffeomorphism invariant theories, we had to develop it for Weyl transverse gravity from the scratch. We actually obtained formalism applicable to a wide class of arbitrary, local gravitational theories which are invariant under volume preserving diffeomorphism and Weyl transformations Alonso-Serrano et al. [2023a, 2022]. Aside from applying the covariant phase formalism in the context of local causal horizons, we also studied other physical relevant settings, in particular stationary black holes and the de Sitter horizon.

The second principal direction of our research is the application of the local equilibrium conditions to study low-energy quantum gravitational effects. In particular, we introduce a correction term to entropy of the local horizons which is logarithmic in the horizon area. Remarkably, this form of the leading order correction to entropy is predicted nearly universally by all the major candidate theories of quantum gravity Kaul and Majumdar [2000], Banerjee et al. [2011], Faulkner et al. [2013], entanglement entropy calculations Solodukhin [1995] as well as various phenomenological approaches Adler et al. [2001], Gour and Medved [2003], Hod [2004]. Therefore, the predictions for gravitational dynamics implied by the presence of the logarithmic term in the entropy of local causal horizon are in principle theory-independent. While the completely general analysis of the relevant local equilibrium conditions is very complicated, we were able to treat two different simplified cases. First, by linearising in the curvature tensors, we derive a result equivalent to the linearised equations of motion of quadratic gravity Alonso-Serrano and Liška [2023a]. Second, a different set of simplifying assumptions allows us to explore the nonlinear regime, leading to relatively simple equations with correction terms quadratic in the Ricci tensor Alonso-Serrano and Liška [2020b, 2023b]. Notably, in the cosmological setting these equations reproduce the low energy dynamics of loop quantum cosmology Alonso-Serrano et al. [2023b,c]. Whereas the (semi)classical part of the thesis can be considered essentially complete, the quantum phenomenological gravitational dynamics encoded in the local equilibrium conditions remains an active area of research. On the one side, we are currently studying its physical implications in more complicated settings, e.g. anisotropic cosmologies and gravitational collapse. On the other side, we are working to complete the derivation of the equations without any simplifying assumptions.

The thesis is organised as follows. In chapter 1, we first recall the main features of Weyl transverse gravity. We also include a novel discussion of the status of the various formulations of the equivalence principle in it. Then, we introduce a version of the covariant phase space formalism we obtained for Weyl transverse gravity and more general theories with the same symmetry group Alonso-Serrano et al. [2023a, 2022]. Chapter 2 shows that the local equilibrium conditions are fully consistent with Weyl transverse gravity. We argue this point in several distinct setups, providing a compelling and complete overall picture. Chapter 3 sums up our results in regards to the quantum gravitational corrections implied by the logarithmic term in entropy. Lastly, chapter 4 discusses the physical implications of these corrections in the context of a homogeneous, isotropic cosmological spacetime.

1. Weyl transverse gravity and the covariant phase space formalism

In this chapter, we explore some interesting features of Weyl transverse gravity, an alternative theory to general relativity. This theory reproduces all the known solutions of general relativity Álvarez et al. [2006], Barceló et al. [2014, 2018], Carballo-Rubio et al. [2022]. It also offers a different perspective on the nature of the cosmological constant, partially addressing the problems related to its value. Moreover, as we show in chapter 2, from the thermodynamic perspective, Weyl transverse gravity appears to provide a more natural description of gravity than general relativity. Given these attractive features of Weyl transverse gravity, we decided to develop the covariant phase space formalism for it. This formalism has the potential to address several of the most interesting aspects of Weyl transverse gravity. First, it provides a way to define the conserved quantities and to derive the first law of horizon thermodynamics for the theory. Second, it allows us to check how the different status of the cosmological constant (as compared to general relativity) affects both the thermodynamics and the phase space structure in general. Lastly, it opens a way to compare Weyl transverse gravity and general relativity on the level of their phase space formulations. The last possibility is not explored here in depth, but we plan to study it in a future work.

The chapter is organised as follows. Section 1.1 introduces Weyl transverse gravity as well as the more general class of local, WTDiff-invariant gravitational theories. While we mostly recount the well established results (summed up in the recent excellent review Carballo-Rubio et al. [2022]), we also include a novel discussion of the equivalence principle. The remarkable (though unsurprising) outcome is that Weyl transverse gravity, much like general relativity, obeys the strong equivalence principle, i.e., even a strongly self-gravitating test body (e.g., a sufficiently small black hole) locally behaves as it would in the absence of external gravitational fields. Hence, general relativity and Weyl transverse gravity are the only two theories in four spacetime dimensions known to respect the strong equivalence principle[1]. Section 1.2 recalls the main features of the covariant phase space formalism, both in full generality and as a useful method to study the thermodynamics of local, Diff-invariant theories of gravity. The last two sections then contain our new results. Namely, in section 1.3, we present the construction of the covariant phase space formalism for Weyl transverse gravity and its application to horizon thermodynamics. Section 1.4 then generalises this construction to arbitrary local, WTDiff-invariant theories of gravity.

[1] Nordström gravity also obeys the weak gravitational equivalence principle that applies to strongly self gravitating bodies. However, it is not invariant under transverse diffeomorphisms and, therefore, does not incorporate the strong equivalence principle di Casola et al. [2014, 2015].

1.1 Weyl transverse gravity

Weyl transverse gravity[2] is a gravitational theory offering a classical alternative to general relativity. Unlike many other alternative theories of gravity, its equations of motion have the same classical solutions as those of general relativity Carballo-Rubio et al. [2022]. The difference between both theories lies in the symmetry group of Weyl transverse gravity, which no longer consists of arbitrary diffeomorphisms. Instead, it contains the subgroup of spacetime volume preserving diffeomorphisms together with Weyl transformations. As we will show, this difference in local symmetries leads to a different behaviour of the cosmological constant. Consequently, Weyl transverse gravity partially addresses the cosmological constant problems which plague general relativity Carballo-Rubio [2015], Barceló et al. [2018].

From the field theory viewpoint, gravity can be described as theory of interacting spin 2 massless particles, the gravitons. If we want gravitons to be described by an object invariant under Poincaré transformations, we must choose a symmetric rank 2 tensor $h_{\mu\nu}$ with $n\left(n+1\right)/2$ independent components (n being the spacetime dimension) Deser [1970], Padmanabhan [2008], Barceló et al. [2014, 2018]. The most general action for linearised theory of massless spin 2 particles described by $h_{\mu\nu}$ reads Álvarez et al. [2006]

$$I = \int_{\mathcal{V}} \left(\sum_{k=1}^{4} M_{(k)}^{\alpha\tau\beta\gamma\rho\sigma} \nabla_\alpha h_{\beta\gamma} \nabla_\tau h_{\rho\sigma} \right) \omega \mathrm{d}^n x \tag{1.1}$$

where $\mathcal{V}$ is our spacetime manifold, $\omega = \sqrt{-\eta}$ stands for the determinant of the flat Minkowski metric $\eta_{\mu\nu}$ (in arbitrary coordinates, so $\omega \neq 1$ in general), and matrices $M_{(k)}^{\alpha\tau\beta\gamma\rho\sigma}$ are given as

$$M_{(1)}^{\alpha\tau\beta\gamma\rho\sigma} = -\frac{1}{4}\eta^{\alpha\tau}\eta^{\beta(\rho}\eta^{\sigma)\gamma}, \tag{1.2}$$

$$M_{(2)}^{\alpha\tau\beta\gamma\rho\sigma} = \frac{1}{4}\left(1+w_1\right)\left(\eta^{\gamma\tau}\eta^{\beta(\rho}\eta^{\sigma)\alpha} + \eta^{\beta\tau}\eta^{\gamma(\rho}\eta^{\sigma)\alpha}\right), \tag{1.3}$$

$$M_{(3)}^{\alpha\tau\beta\gamma\rho\sigma} = -\frac{1}{4}\left(1+w_2\right)\left(\eta^{\rho\sigma}\eta^{\alpha(\beta}\eta^{\gamma)\tau} + \eta^{\beta\gamma}\eta^{\tau(\rho}\eta^{\sigma)\alpha}\right), \tag{1.4}$$

$$M_{(3)}^{\alpha\tau\beta\gamma\rho\sigma} = \frac{1}{4}\left(1+w_3\right)\eta^{\alpha\tau}\eta^{\beta\gamma}\eta^{\rho\sigma}, \tag{1.5}$$

with w_1, w_2, w_3 being arbitrary numbers. While $h_{\mu\nu}$ contains $n\left(n+1\right)/2$ independent functions, a massless particle can only have 2 physical polarisations. To ensure that no other physical degrees of freedom appear in $h_{\mu\nu}$, we must introduce $n\left(n+1\right)/2$ constraints and/or gauge symmetries for it. Various choices of numbers w_1, w_2, w_3 require different combinations of gauge symmetries and constraints. However, only two sets of the parameters allow maximum gauge freedom, leaving $h_{\mu\nu}$ as an arbitrary symmetric rank 2 tensor with $n\left(n+1\right)/2$ gauge

[2] A brief note on the terminology is in order: names Weyl transverse gravity and unimodular gravity are often used interchangeably. However, the latter also refers to other theories distinct from Weyl transverse gravity Henneaux and Teitelboim [1989], Padilla and Saltas [2014], Bufalo et al. [2015]. For the sake of clarity, we stick to Weyl transverse gravity in the present text, although many notable recent works prefer to call the same theory unimodular gravity Carballo-Rubio et al. [2022], Álvarez et al. [2023].

symmetries. First, for $w_1 = w_2 = w_3$, action (1.1) is invariant under infinitesimal diffeomorphisms (Diff invariant), i.e., under any transformation of the form

$$h'_{\mu\nu} = h_{\mu\nu} + 2\nabla_{(\mu}\xi_{\nu)}, \tag{1.6}$$

where ξ^μ can be arbitrary sufficiently smooth vector field. This choice leads to linearised general relativity (and including graviton self-interactions allows one to iteratively construct the full non-linear theory). However, there is a second option. By setting $w_1 = 0$, $w_2 = -1/2$, $w_3 = -5/8$ we obtain an action invariant under the following transformations

$$h'_{\mu\nu} = h_{\mu\nu} + 2\sigma\eta_{\mu\nu} + 2\nabla_{(\mu}\xi_{\nu)}, \tag{1.7}$$

with σ being any non-negative function. Vector field ξ^μ cannot be arbitrary in this case, but must obey the divergence-free condition $\nabla_\mu\xi^\mu = 0$. The term proportional to σ corresponds to a Weyl rescaling of the metric, and we refer to the transformation generated by ξ^μ as a transverse (or volume preserving) diffeomorphism. Therefore, we call the resulting theory Weyl transverse gravity and its symmetry group WTDiff.

To conclude this introduction let us briefly compare Diff and WTDiff groups in the linearised setting we considered so far. Clearly, both groups contain as a proper subgroup the transverse diffeomorphisms. In fact, WTDiff group has a structure of a semi-direct product of the group of transverse diffeomorphisms and the group of Weyl transformations Álvarez et al. [2016]. Apart from transverse diffeomorphisms, the only other kind of transformations in the intersection of Diff and WTDiff groups are those generated by the conformal Killing vectors ζ^μ defined by the conformal Killing condition $\pounds_\zeta\eta_{\mu\nu} = \sigma\eta_{\mu\nu}$, where σ is again an arbitrary function. In flat spacetime, the conformal Killing vectors can be given explicitly

$$\zeta^\mu = a^\mu + bx^\mu + x_\nu x^\nu c^\mu + 2c_\nu x^\nu x^\mu, \tag{1.8}$$

where a^μ, b, c^μ are arbitrary constants. It is easy to check that both Diff and WTDiff groups contain transformations that cannot be written as a linear combination of a transverse diffeomorphism and a conformal Killing transformation Álvarez et al. [2006]. Therefore, both groups overlap but neither of them is a subset of the other one. Since Diff and WTDiff groups differ, Weyl transverse gravity represents a gravitational theory distinct from general relativity or any gauge fixed version thereof.

1.1.1 Non-linear vacuum theory

Upon discussing the field theoretical motivation for Weyl transverse gravity, let us introduce the full, non-linear theory in a geometric language (a field theoretical viewpoint is also possible for the non-linear theory Barceló et al. [2014, 2018], but it is not well suited for the purposes of this thesis). To write a non-linear WTDiff-invariant gravitational action, it turns out to be necessary to introduce some non-dynamical background structure. One possibility is suggested by the field theoretical viewpoint, in which one keeps track of the background metric $\eta_{\mu\nu}$ while constructing the theory of self-interacting gravitons Barceló et al. [2014]. In the Diff-invariant case (general relativity), this metric can be safely discarded upon completing the construction. However, with the WTDiff-invariance, one eventually obtains the following action

$$I_{\mathrm{WTG}} = \frac{1}{16\pi}\int_{\mathcal{V}}\left[R + \frac{(n-1)(n-2)}{n^2}g^{\alpha\beta}\partial_\alpha \ln\frac{\sqrt{-\mathfrak{g}}}{\omega}\partial_\beta \ln\frac{\sqrt{-\mathfrak{g}}}{\omega}\right]\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}}\omega \mathrm{d}^n x$$
$$-\frac{1}{8\pi}\int_{\mathcal{V}}\lambda\omega \mathrm{d}^n x + \frac{1}{8\pi}\frac{n-1}{n}\oint_{\partial\mathcal{V}}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{1}{n}}\partial_\mu \ln\frac{\sqrt{-\mathfrak{g}}}{\omega}n^\mu \omega \mathrm{d}^{n-1}x, \qquad (1.9)$$

where $\mathfrak{g}$ denotes the determinant of the dynamical metric $g_{\mu\nu}$, $\omega = \sqrt{-\eta}$ is the determinant of the background metric, and λ is a constant. The boundary integral does not affect the equations of motion, but ensures that the Weyl invariance of the theory is exact. The dependence of the action on ω cannot be removed. To work with ω, we may keep track of the full background metric $\eta_{\mu\nu}$, obtaining a Rosen-style bi-metric theory Rosen [1940, 1973]. However, while interesting in its own right (e.g. for its relation with analogue gravity Barceló et al. [2022]), this choice is a bit of an overkill. The minimal amount of the non-dynamical background information we need to preserve is encoded in the volume n-form associated to the background metric $\boldsymbol{\omega} = \omega(x)\,\mathrm{d}x^0 \wedge \mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^{n-1}$, with ω being a strictly positive function. We can then discard the rest of the information about the background metric. In the present work, we choose this route.

As we study gravity in the geometric paradigm, we do not even need to introduce the background metric at any point. We simply consider a manifold $\mathcal{V}$ endowed with metric $g_{\mu\nu}$ and n-form $\boldsymbol{\omega}$. On this manifold, we wish to write the most general theory of gravity invariant under WTDiff transformations and linear in the spacetime curvature. The unique solution is the action (1.9) Álvarez and Herrero-Valea [2013a] (much like general relativity is the unique metric, Diff-invariant theory linear in the curvature). The term $-(1/8\pi)\int_{\mathcal{V}}\lambda\omega \mathrm{d}^n x$ does not depend on the metric (the only dynamical variable in vacuum) in any way. Therefore, it does not contribute to the equations of motion or even to higher order effective field theory calculations Carballo-Rubio [2015] and it can be removed without any loss of generality. Therefore, unless specified otherwise, we always set $\lambda = 0$ in the following.

The derivatives of the metric determinant in action (1.9) are somewhat awkward to handle. We can write it in a much more elegant form by switching to suitable auxiliary variables that respect the WTDiff symmetry (keeping the metric $g_{\mu\nu}$ as the actual dynamical variable). First, we define the WTDiff-invariant (but not Diff-invariant) auxiliary metric

$$\tilde{g}_{\mu\nu} = \left(\sqrt{-\mathfrak{g}}/\omega\right)^{-2/n} g_{\mu\nu}. \qquad (1.10)$$

We further introduce an auxiliary connection which is Levi-Civita with respect to $\tilde{g}_{\mu\nu}$, i.e.,

$$\tilde{\Gamma}^\mu{}_{\nu\rho} = \Gamma^\mu{}_{\nu\rho} - \frac{1}{n}\left(\delta^\mu_\nu\delta^\alpha_\rho + \delta^\mu_\rho\delta^\alpha_\nu - g_{\nu\rho}g^{\mu\alpha}\right)\partial_\alpha \ln\frac{\sqrt{-\mathfrak{g}}}{\omega}. \qquad (1.11)$$

Finally, we get the auxiliary Riemann tensor by the standard expression

$$\tilde{R}^\mu{}_{\nu\rho\sigma} = 2\tilde{\Gamma}^\mu{}_{\nu[\sigma,\rho]} + 2\tilde{\Gamma}^\mu{}_{\lambda[\rho}\tilde{\Gamma}^\lambda{}_{\sigma]\nu}. \qquad (1.12)$$

These WTDiff-invariant objects allow us to rewrite the action (1.9) in a way more reminiscent of the Einstein-Hilbert action of general relativity

$$I_{\mathrm{WTG}} = \frac{1}{16\pi}\int_{\mathcal{V}} \tilde{R}\omega \mathrm{d}^n x, \tag{1.13}$$

where we defined the auxiliary scalar curvature $\tilde{R} = \tilde{g}^{\mu\nu}\tilde{R}^{\lambda}{}_{\mu\lambda\nu}$. We stress that the actions (1.9) and (1.13) are identical, including the boundary terms (the only difference being that we have already set $\lambda = 0$ in the action (1.13)).

Before proceeding, let us stress a subtle, but very important point. The auxiliary metric $\tilde{g}_{\mu\nu}$ corresponds to the dynamical metric $g_{\mu\nu}$ restricted to the unimodular gauge, in which $\sqrt{-\mathfrak{g}} = \omega$. However, we do not impose any such gauge restrictions. Instead, we treat $\tilde{g}_{\mu\nu}$ as a mere notational device. We keep $g_{\mu\nu}$ as our dynamical variable and also use it to raise and lower indices (with the obvious exception of the contravariant auxiliary metric $\tilde{g}^{\mu\nu}$ which we define as an inverse of $\tilde{g}_{\mu\nu}$).

By construction, the action I_{WTG} is invariant under Weyl transformations

$$\delta g_{\mu\nu} = e^{2\sigma} g_{\mu\nu}, \tag{1.14}$$

and transverse diffeomorphisms,

$$\delta g_{\mu\nu} = 2\nabla_{(\nu}\xi_{\mu)}, \tag{1.15}$$

$$\tilde{\nabla}_\mu \xi^\mu = 0 \qquad \Longleftrightarrow \qquad \nabla_\mu \xi^\mu = \xi^\mu \partial_\mu \ln\frac{\sqrt{-\mathfrak{g}}}{\omega}, \tag{1.16}$$

where the second equation represents a WTDiff-invariant generalisation of the flat background transversality condition, $\nabla_\mu \xi^\mu = 0$, valid for the full, non-linear theory. The transversality condition does not depend on the connection. Hence, we can separate the transverse diffeomorphism subgroup from the Diff group in a general background. However, the remaining vectors lying in the intersection of the WTDiff and the Diff groups, i.e., the conformal Killing vectors, are defined by a connection-dependent condition, $\pounds_\zeta g_{\mu\nu} = \sigma g_{\mu\nu}$. Thence, for a generic metric, we are unable to specify the intersection of the WTDiff and the Diff groups as we did for the flat background in the previous subsection.

To find the vacuum equations of motion for Weyl transverse gravity, we vary I_{WTG} with respect to $g^{\mu\nu}$. Upon some straightforward manipulations, we get

$$\tilde{R}_{\mu\nu} - \frac{1}{n}\tilde{R}\tilde{g}_{\mu\nu} = 0. \tag{1.17}$$

These equations are explicitly WTDiff-invariant and traceless. We may also restate them in a divergenceless form which is closer to the Einstein equations. To that end, we use the contracted Bianchi identities

$$2\tilde{g}^{\nu\rho}\tilde{\nabla}_\nu \tilde{R}_{\mu\rho} = \tilde{\nabla}_\mu \tilde{R}, \tag{1.18}$$

to show that, on shell,

$$\left(n-2\right)/\left(2n\right)\tilde{R}\tilde{g}_{\mu\nu} = \Lambda\tilde{g}_{\mu\nu}, \tag{1.19}$$

where Λ is an arbitrary integration constant. Adding this expression to the traceless equations of motion for Weyl transverse gravity yields

$$\tilde{R}_{\mu\nu} - \frac{1}{2}\tilde{R}\tilde{g}_{\mu\nu} + \Lambda\tilde{g}_{\mu\nu} = 0. \tag{1.20}$$

By comparison to the Einstein equations, it can be seen that Λ plays the role of the cosmological constant. Crucially, unlike in general relativity, Λ has no connection to the constant parameter λ in the Lagrangian. Furthermore, Λ is only well defined on shell, i.e., for a particular solution of the equations of motion, as it only appears in the process of integrating them. Then, Λ naturally takes different values for different solutions of the theory. In other words, Λ represents an additional, global degree of freedom of Weyl transverse gravity.

Equations (1.20) reduce to the Einstein equations (aside from the origin of Λ) of general relativity in the unimodular gauge, $\sqrt{-\mathfrak{g}} = \omega$. Hence, Weyl transverse gravity has the same classical solution as general relativity. Moreover, it can be shown that linearised quantisations of both theories are equivalent, and both theories allow for equivalent formulations of the Euclidean path integral quantisation and embedding in the string theory Carballo-Rubio et al. [2022], Garay and García-Moreno [2023]. The only known physical difference between them is thus the behaviour of Λ, which is a fixed Lagrangian parameter in general relativity, but a global degree of freedom in Weyl transverse gravity. Given the similarity of general relativity and Weyl transverse gravity, one may wonder whether there exists a parent theory invariant under both arbitrary diffeomorphisms and Weyl transformations. General relativity and Weyl transverse gravity would then just correspond to different gauge fixing in this theory. A scalar-tensor theory with this property has indeed been found Oda [2017], but it tacitly assumes the restriction $\Lambda = 0$. If one keeps the value of Λ free, a recent result shows that no such theory exists García-Moreno and Jiménez Cano. The reason is that a Diff-invariant metric theory of gravity cannot accommodate the cosmological constant as a global degree of freedom (although a Diff-invariant action with this property has been recently proposed in a first-order formalism Montesinos and Gonzalez [2023]).

1.1.2 Coupling to matter fields

Upon introducing the vacuum theory, we now turn to matter fields minimally coupled to Weyl transverse gravity. We start by choosing an appropriate WTDiff-invariant action. It reads

$$I_\psi = \int_{\mathcal{V}} \left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} L_\psi \omega \mathrm{d}^n x, \tag{1.21}$$

where $\mathcal{V}$ is again the spacetime, L_ψ denotes a function of the matter variables ψ, their partial derivatives, and k contravariant metric tensors $g^{\mu\nu}$. If more fields are present, I_ψ simply becomes a sum of several contributions of this form, possibly including interaction terms. The factor $(\sqrt{-\mathfrak{g}}/\omega)^{2k/n}$ compensates the behaviour of $g^{\mu\nu}$ under Weyl transformations and is added to ensure that the action is WTDiff-invariant. We stress that the Weyl transformations by definition do not affect the matter variables and only act on the gravitational sector of the theory.

A simultaneous variation of the matter action I_ψ and the gravitational action I_{WTG} with respect to the dynamical metric $g^{\mu\nu}$ yields the full gravitational equations of motion

$$\tilde{R}_{\mu\nu} - \frac{1}{n}\tilde{R}\tilde{g}_{\mu\nu} = 8\pi \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k-1}{n}} \left(T_{\mu\nu} - \frac{1}{n}T g_{\mu\nu}\right). \tag{1.22}$$

The energy momentum tensor is defined via the Hilbert prescription

$$T_{\mu\nu} = -2\frac{\partial L_\psi}{\partial g^{\mu\nu}} + L_\psi g_{\mu\nu}, \tag{1.23}$$

and $T = T_{\mu\nu}g^{\mu\nu}$ is its trace. Under Weyl transformations, the energy-momentum tensor behaves as

$$T'_{\mu\nu} = e^{-2(k-1)\sigma}T_{\mu\nu}, \tag{1.24}$$

making the right hand side of the equations of motion (1.22) WTDiff-invariant.

Taking a WTDiff-invariant divergence of the equations of motion (1.22), we find that the contracted Bianchi identities do not enforce a divergenceless energy-momentum tensor. Instead, they only imply Álvarez and Herrero-Valea [2013a]

$$8\pi\tilde{\nabla}_\nu\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} T_\mu^{\ \nu}\right] = \tilde{\nabla}_\mu\mathcal{J}, \tag{1.25}$$

where $\mathcal{J}$ is some function. The Einstein-like divergenceless equations of motion then read

$$\tilde{R}_{\mu\nu} - \frac{1}{2}\tilde{R}\tilde{g}_{\mu\nu} + \Lambda\tilde{g}_{\mu\nu} = 8\pi \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k-1}{n}} T_{\mu\nu} - \mathcal{J}\tilde{g}_{\mu\nu}. \tag{1.26}$$

where Λ is again an arbitrary integration constant. The vanishing divergence of the energy-momentum tensor is equivalent to the local energy-momentum conservation. Conversely, if $\mathcal{J}$ does not equal zero, it represents a measure of the local energy non-conservation.

To conclude this subsection, we provide a brief discussion of an illustrative example of a scalar field ϕ minimally coupled to Weyl transverse gravity. The complete WTDiff-invariant action reads

$$I = \frac{1}{16\pi}\int_{\mathcal{V}}\tilde{R}\omega \mathrm{d}^n x + \int_{\mathcal{V}}\left[\frac{1}{2}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n} g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi + V(\phi)\right]\omega \mathrm{d}^n x, \tag{1.27}$$

where $V(\phi)$ denotes the potential (possibly including a mass term). The gravitational equations of motion are

$$\tilde{R}_{\mu\nu} - \frac{1}{n}\tilde{R}\tilde{g}_{\mu\nu} = 8\pi\left(\partial_\mu\phi\partial_\nu\phi - \frac{1}{n}\tilde{g}^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi\tilde{g}_{\mu\nu}\right), \tag{1.28}$$

whereas the equation of motion for the scalar field reads

$$\tilde{g}^{\alpha\beta}\tilde{\nabla}_\alpha\tilde{\nabla}_\beta\phi = V'(\phi). \tag{1.29}$$

Taking a divergence of the gravitational equations of motion and invoking the Bianchi identities yields

$$\begin{aligned}\frac{n-2}{2n}\tilde{\nabla}_\mu\tilde{R} =& 8\pi\left(\frac{n-2}{n}\tilde{g}^{\alpha\beta}\tilde{\nabla}_\mu\tilde{\nabla}_\alpha\phi\tilde{\nabla}_\beta\phi + \tilde{g}^{\alpha\beta}\tilde{\nabla}_\alpha\tilde{\nabla}_\beta\phi\tilde{\nabla}_\mu\phi\right)\\ =& 8\pi\tilde{\nabla}_\mu\left(\frac{n-2}{2n}\tilde{g}^{\alpha\beta}\tilde{\nabla}_\alpha\phi\tilde{\nabla}_\beta\phi + V(\phi)\right),\end{aligned} \tag{1.30}$$

where we used the equation of motion for the scalar field to get the second equality. Integrating and subtracting the result from the traceless gravitational equations of motion finally leads to

$$\tilde{R}_{\mu\nu} - \frac{1}{2}\tilde{R}\tilde{g}_{\mu\nu} + \Lambda\tilde{g}_{\mu\nu} = 8\pi\left(\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}\tilde{g}^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi\tilde{g}_{\mu\nu} - V\left(\phi\right)\tilde{g}_{\mu\nu}\right), \quad (1.31)$$

where Λ again appears as an integration constant. The right hand side is just the Hilbert energy-momentum tensor of the scalar field. Hence, in this case $\mathcal{J} = 0$ and the equations of motion for the scalar field ensure the local energy conservation, even though the gravitational equations of motion in principle allow its violation. We further discuss this observation in subsection 1.1.5.

1.1.3 The cosmological constant problem

Observational evidence shows that the expansion of the universe is accelerating. Its acceleration agrees with the effect of a small non-zero cosmological constant Λ, present in the Einstein equations, with the value $\Lambda \approx 10^{-122} l_{\mathrm{P}}^{-2}$ Barrow and Shaw [2011]. However, there exists no fully convincing microscopic interpretation of Λ Padmanabhan [2008]. Even if one were to accept Λ as a fundamental constant of nature, another problem occurs. In semiclassical gravity, which treats matter fields as quantum and gravity as classical, the vacuum energy density ρ_0 contributes to the Einstein equations by a term of the form $\rho_0 g_{\mu\nu}$. This contribution has the same form as the cosmological constant term $\Lambda g_{\mu\nu}$. However, it has been proposed that the natural scale of the vacuum energy density is the Planck length, i.e., $\rho_0 \approx l_{\mathrm{P}}^{-2}$ Weinberg [1989], Burgess [2013]. Thus, there is a huge discrepancy between the expected value of ρ_0 and the cosmological constant required to explain the observed acceleration of the universe (or even the largest cosmological constant not ruled out by solar system experiments Kagramanova et al. [2006]). Then, there would need to be a very precise cancellation between Λ and ρ_0 almost to zero. Even more seriously, in the effective field theory approach to gravity, each higher loop contribution to ρ_0 again apparently leads to a value much larger than $10^{-122} l_{\mathrm{P}}^{-2}$ Weinberg [1989], Carballo-Rubio [2015], requiring an infinite amount of fine-tuning. In other words, the value of the cosmological constant is radiatively unstable Weinberg [1989]. Then, either the effective field theory approach to gravity fails (and the value of Λ can only be explained by the full quantum gravity), or some mechanism cancelling out the huge contributions to ρ_0 must be found. While such mechanisms have been put forward even in the context of general relativity Mooij and Shaposhnikov [2021a,b], Donoghue [2021], de Brito et al. [2021], Hossenfelder [2021], Mottola [2022], Weyl transverse gravity offers a particularly elegant and robust solution. Since the equations of motion are traceless, they are invariant under adding a term of the form $\rho_0 g_{\mu\nu}$ (with ρ_0 being a constant) to the energy-momentum tensor. Then, the vacuum energy simply does not gravitate. Moreover, it can be shown that the value of the cosmological constant is manifestly radiatively stable in Weyl transverse gravity Carballo-Rubio [2015]. This result is possible thanks to the WTDiff invariance of the theory and to the absence of any quantum anomaly corresponding to local Weyl symmetry Álvarez and Herrero-Valea [2013b]. While Weyl transverse gravity, by itself, still does not explain the observed value of the cosmological constant, it allows us to apply the effective field theory approach without any extra assumptions regarding the vacuum energy.

1.1.4 General WTDiff-invariant theories

Weyl transverse gravity represents only the simplest example of WTDiff-invariant gravity. The most action describing such a local theory reads

$$\begin{aligned} I = \int_{\mathcal{V}} L\Big(&\tilde{g}_{\mu\nu}, \tilde{R}^{\mu}{}_{\nu\rho\sigma}, \tilde{\nabla}_{\alpha_1}\tilde{R}^{\mu}{}_{\nu\rho\sigma}, \ldots, \tilde{\nabla}_{(\alpha_1}\ldots\tilde{\nabla}_{\alpha_p)}\tilde{R}^{\mu}{}_{\nu\rho\sigma}, \\ &\psi, \tilde{\nabla}_{\alpha_1}\psi, \ldots, \tilde{\nabla}_{(\alpha_1}\ldots\tilde{\nabla}_{\alpha_p)}\psi\Big)\omega \mathrm{d}^n x, \end{aligned} \tag{1.32}$$

where p is a natural number, and ψ denotes some collection of matter variables (with suppressed spacetime and gauge indices for simplicity of notation). We are free to only use fully symmetrised covariant derivatives, as we can employ the definition of the auxiliary Riemann tensor to rewrite any other ordering of derivatives in terms of a fully symmetric one and some lower derivative terms. By construction, action 1.32 is WTDiff-invariant. To obtain the gravitational equations of motion, we vary action (1.32) with respect to $g_{\mu\nu}$, obtaining

$$16\pi\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}}\frac{\delta S}{\delta g_{\mu\nu}} = \mathring{A}^{\mu\nu} = -8\pi\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k+1}{n}}\left(T^{\mu\nu} - \frac{1}{n}Tg^{\mu\nu}\right), \tag{1.33}$$

where, from now on, symbol $\circ$ signifies traceless, WTDiff-invariant tensors. We define the energy momentum tensor via the Hilbert prescription (1.23) applied to the part of the Lagrangian that does not depend on the auxiliary connection. We include the other, non-minimally coupled terms depending on the matter fields ψ in the "gravitational part" of the equations of motion $\mathring{A}^{\mu\nu}$. To conform with the literature on the covariant phase space formalism, we vary the action with respect to $g_{\mu\nu}$ rather than $g^{\mu\nu}$. Consequently, the overall sign of the equations of motion is opposite to the one usually preferred in different contexts (which comes from variations with respect to the contravariant metric $g^{\mu\nu}$).

Next, as with Weyl transverse gravity, we would like to rewrite the equations of motion so that they are divergenceless. To this end, we use that action (1.32) is invariant under transverse diffeomorphisms, which implies

$$\begin{aligned} 0 =& \frac{\delta S}{\delta g_{\mu\nu}}\nabla_{(\nu}\xi_{\mu)} = \frac{1}{16\pi}\int_V \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{-\frac{2}{n}}\mathring{A}^{\mu\nu}\nabla_{(\nu}\xi_{\mu)}\omega \mathrm{d}^n x \\ =& \frac{1}{16\pi}\int_V \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{-\frac{2}{n}}\xi_\mu\tilde{\nabla}_\nu\mathring{A}^{\mu\nu}\omega \mathrm{d}^n x. \end{aligned} \tag{1.34}$$

We tacitly imposed the appropriate Neumann boundary conditions to ensure that all the boundary integrals vanish[3]. Equation (1.34) holds for an arbitrary generator ξ^μ of an infinitesimal transverse diffeomorphism. Therefore, we must have

$$\tilde{\nabla}_\nu \mathring{A}^{\mu\nu} = \tilde{\nabla}_\mu \Phi, \tag{1.35}$$

where Φ is a function. In principle, we can fully specify Φ, up to a constant, by taking a WTDiff-invariant divergence of $\mathring{A}^{\mu\nu}$ and rewriting it as a gradient of a scalar. For the divergence of the energy-momentum tensor, we can use equation (1.25). In total, we have on shell

$$\Phi = -\Lambda + \mathcal{J} - \frac{1}{n}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k}{n}} T, \tag{1.36}$$

where Λ is an arbitrary integration constant. Hence, the divergenceless equations of motion are

$$\mathring{A}^{\mu\nu} - (\Phi + \Lambda)\, \tilde{g}^{\mu\nu} = -8\pi \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k+1}{n}} T^{\mu\nu} + \mathcal{J}\tilde{g}^{\mu\nu}. \tag{1.37}$$

In the unimodular gauge, $\sqrt{-\mathfrak{g}} = \omega$, and with $\mathcal{J} = 0$, these equations reduce to the equations of motion of some Diff-invariant theory with the same classical solutions Carballo-Rubio et al. [2022]. In this way, one can show that for every local WTDiff-invariant theory with a conserved energy-momentum tensor, there exists a corresponding local, Diff-invariant theory with the same classical solutions (see Carballo-Rubio et al. [2022] for a more detailed version of this argument).

1.1.5 Local energy (non-)conservation

One might wonder how realistic are the scenarios of local energy non-conservation which any WTDiff-invariant gravitational equations of motion in principle allow. To be as general as possible, we tackle this issue for any collection of non-minimally coupled matter fields ψ. Since matter variables are unaffected by Weyl transformations, they are automatically WTDiff-invariant. Hence, to construct

[3] The use of the Gauss theorem for a WTDiff-invariant divergence deserves a brief explanation. For an integral of a WTDiff-invariant divergence of some WTDiff-invariant vector W^μ it holds

$$\int_V \tilde{\nabla}_\mu W^\mu \omega_{\alpha_1\ldots\alpha_n} = \int_V (\omega \partial_\mu W^\mu + W^\mu \partial_\mu \omega)\, \epsilon_{\alpha_1\ldots\alpha_n} = \int_V \partial_\mu (\omega W^\mu)\, \epsilon_{\alpha_1\ldots\alpha_n},$$

where $\epsilon_{\alpha_1\ldots\alpha_n}$ stands for the n-dimensional anti-symmetrisation symbol and $\omega_{\alpha_1\ldots\alpha_n} = \omega\epsilon_{\alpha_1\ldots\alpha_n}$ is the non-dynamical volume element. The Gauss theorem now gives us

$$\int_V \partial_\mu (\omega W^\mu)\, \mathrm{d}^n x = \int_{\partial V} W^\mu n_\mu n^{\alpha_1} \omega_{\alpha_1\ldots\alpha_n},$$

where n^μ is a unit normal to ∂V. Under Weyl transformations, it changes as $n'^\mu = e^{-\sigma} n^\mu$. For greater clarity, suppose we choose the coordinate system so that n^μ becomes a coordinate vector. Then, we have

$$\int_{\partial V} W^\mu n_\mu n^{\alpha_1} \omega_{\alpha_1\ldots\alpha_n} = \int_{\partial V} \left(\sqrt{-\mathfrak{g}}/\omega\right)^{-1/n} W^\mu n_\mu \mathrm{d}^{n-1} x.$$

the matter action, we can only include WTDiff-invariant objects constructed from the metric, i.e., the auxiliary metric $\tilde{g}_{\mu\nu}$, the WTDiff-invariant derivative $\tilde{\nabla}_\mu$, and the auxiliary Riemann tensor $\tilde{R}^{\mu}{}_{\nu\rho\sigma}$. The matter Lagrangian is then a local functional $L_\psi\left[\psi, \tilde{g}_{\mu\nu}, \tilde{\nabla}_\mu, \tilde{R}^{\mu}{}_{\nu\rho\sigma}\right]$. Taking the corresponding Diff-invariant Lagrangian $L_\psi\left[\psi, g_{\mu\nu}, \nabla_\mu, R^{\mu}{}_{\nu\rho\sigma}\right]$ (its existence has been shown in the previous subsection), we easily find that $\nabla_\nu T_\mu{}^\nu = 0$ for it. Then, we also have (just using the definition of the WTDiff-invariant derivative)

$$\tilde{\nabla}_\nu \left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} T_\mu{}^\nu\right] = 0. \tag{1.38}$$

The right hand side of equations (1.26) then has $\mathcal{J} = 0$ and is divergenceless. In total, for any WTDiff-invariant Lagrangian $L_\psi\left[\psi, \tilde{g}_{\mu\nu}, \tilde{\nabla}_\mu, \tilde{R}^{\mu}{}_{\nu\rho\sigma}\right]$ (with any matter fields present and including non-minimal coupling) the local conservation of energy holds.

We are aware of three possible ways to bypass this restriction. First, resigning on having a Lagrangian description for the matter sector would certainly allow us more possibilities Josset et al. [2017]. However, it is unclear how to interpret an energy-momentum tensor that cannot be obtained using a variational principle. Second, we may consider matter variables with non-trivial behaviour under Weyl transformations. These would allow us to construct a Lagrangian for locally non-conserved matter, but the idea is rather artificial. Third, we can use that the background n-volume measure ω does not change under Weyl transformations and add terms constructed from the matter variables and derivatives of ω (e.g. terms of the form $S_{\text{nc}} = \int_\omega \tilde{g}^{\alpha\beta}\left(\omega_{,\alpha}/\omega\right)\phi_{,\beta}\omega \mathrm{d}^n x$). However, derivatives of ω then explicitly appear in the equations of motion for the matter fields, allowing us in principle to experimentally probe the non-dynamical volume top form $\boldsymbol{\omega}$. This possibility does mark a clear departure from the usual framework of WTDiff-invariant gravity in which $\boldsymbol{\omega}$ remains experimentally inaccessible Barceló et al. [2018]. Then, the time derivative of ω can serve as a non-dynamical measure of time, breaking the classical correspondence between Weyl transverse gravity and general relativity. In any case, aside from these three exotic and/or pathological scenarios, it appears that WTDiff-invariance suffices to enforce the local energy conservation. Nevertheless, to be as general as possible, we consider the possibility of the energy non-conservation in the following.

Finally, we remark that the constraint on local energy conservation we discussed only works for the WTDiff-invariant Lagrangians. If we restrict our symmetry group to just transverse diffeomorphisms (without Weyl transformations) as is common in some approaches to unimodular gravity Bufalo et al. [2015], Josset et al. [2017], Perez et al. [2018], we are free to use any powers and derivatives of the metric determinant in our matter Lagrangian. Then, the local energy conservation need not be respected Alonso-López et al. [2023], Jaramillo-Garrido et al. [2024].

1.1.6 Equivalence principle(s)

The equivalence principle played an important role in guiding the development of the general relativity Einstein [1911], and its validity is experimentally tested with a high precision Overduin et al. [2009], Pacilio and Liberati [2017]. From the theoretical perspective, the main interest in the various formulations of the equivalence principle nowadays lies in studying their validity in alternative theories of gravity di Casola et al. [2014] and in situations in which quantum effects become relevant Giacomini and Brukner, Wagner et al. [2023], Balsells and Bojowald [2023], Zhang [2024]. Furthermore, the Einstein and the strong equivalence principles play a key role in studying the gravitational dynamics from the perspective of thermodynamics of spacetime. Thence, it is of interest to us to see whether and how all the different equivalence principles are respected by Weyl transverse gravity and WTDiff-invariant theories of gravity in general. The discussion we present here is novel and its conclusions complement the state of the art in the understanding of the (non)equivalence of Weyl transverse gravity and general relativity Carballo-Rubio et al. [2022].

We proceed by checking the validity of the relevant formulations of the equivalence principle one by one. The classification of the equivalence principles we adapt follows di Casola et al. [2015]. The weakest form of the equivalence principle is known as the Newton equivalence principle. It states "In the Newtonian limit, the inertial and gravitational masses of a body are equal" di Casola et al. [2015]. Since this formulation concerns only the Newtonian limit, it is naturally obeyed by Weyl transverse gravity.

The weak gravitational principle asserts that "Test particles with negligible self-gravity behave, in a gravitational field, independently of their properties"di Casola et al. [2015]. A suitable test particle is one whose back-reaction at its environment can be disregarded. By a negligible self gravity, we mean that the ratio of the object's mass and size is much smaller than unity in the natural units. The weak equivalence principle is obeyed if the effect of gravity on a test particle can be (locally, i.e., disregarding geodesic deviation and similar effects relevant on length scales comparable with the curvature length scale) fully captured by the connection. To discuss this requirement in the context of WTDiff-invariant gravity, we must first properly define the motion in a gravitational field. The Diff-invariant geodesic equation reads

$$v^{\nu}\nabla_{\nu}v^{\mu} = Pv^{\mu}, \tag{1.39}$$

where v^{μ} denotes a unit vector tangent to the geodesic and P is an arbitrary function (for affine parametrisation $P = 0$). However, this equation changes under Weyl transformations. Therefore, force-free trajectories in one Weyl frame would be subjected to forces in a different frame. This behaviour is clearly in conflict with Weyl invariance of physics required in WTDiff-invariant gravity. Therefore, we must adapt a different, WTDiff-invariant definition of a geodesic[4]. The obvious solution lies in replacing the derivative by the WTDiff-invariant one

$$\tilde{v}^{\nu}\tilde{\nabla}_{\nu}\tilde{v}^{\mu} = P\tilde{v}^{\mu}. \tag{1.40}$$

[4]This form of the geodesic equation is equivalent to the condition (1.25) on the WTDiff-invariant divergence of the energy-momentum tensor Alonso-Serrano et al. [2024].

To keep the tangent vector $\tilde{v}^\mu$ properly normalised in any Weyl frame, we define it in a WTDiff-invariant manner, i.e.,

$$\tilde{v}^\mu = \left(\sqrt{-\mathfrak{g}}/\omega\right)^{-1/n} v^\mu. \tag{1.41}$$

The geodesics found as solutions of equation (1.40) correspond to true WTDiff-invariant force-free trajectories. With this definition of a geodesic, the weak equivalence principle is incorporated in Weyl transverse gravity as well as in any WTDiff-invariant theory of gravity given by a Lagrangian of the form (1.32).

Before continuing, one should appreciate the significance of the WTDiff-invariant geodesic equation. It directly implies that, while the dynamical metric $g_{\mu\nu}$ remains the dynamical variable describing gravity, the geometry of the spacetime is actually described by the auxiliary metric $\tilde{g}_{\mu\nu}$. Since the difference between both metrics lies only in the, experimentally inaccessible, measure of the spacetime volume, there is no (known) way in which this use of two different metrics allows to distinguish WTDiff-invariant and Diff-invariant theories of gravity Barceló et al. [2018], Carballo-Rubio et al. [2022]. Nevertheless, the difference is conceptually significant and important for the understanding of Weyl transverse gravity.

A stronger formulation of the weak equivalence principle is the Einstein equivalence principle, which states "Fundamental non-gravitational test physics is not affected, locally and at any point of spacetime, by the presence of a gravitational field" di Casola et al. [2015]. It extends the principle of equivalence just from the motion of particles to all non-gravitational physics. While the status of the equivalence principle for matter fields is an interesting subject in its own right di Casola et al. [2015], WTDiff-invariant theories of gravity do not change the non-gravitational physics (in particular, it is unaffected by Weyl transformations). Therefore, the principle applies to WTDiff-invariant gravity in the same way it does to Diff-invariant theories and with the same caveats.

A different way to strengthen the weak equivalence principle lies in including self-gravitating bodies in it, resulting in the gravitational weak equivalence principle "Test particles behave, in a gravitational field and in vacuum, independently of their properties" di Casola et al. [2015]. The principle is not a straightforward generalisation of the weak equivalence principle as it needs to be restricted to vacuum. Otherwise, the intrinsic gravitational field of the test particle would affect nearby particles, breaking the universality. A simple criterion for the validity of the gravitational weak equivalence principle has been proposed di Casola et al. [2014], based on a generalisation of the Geroch-Jang theorem Geroch and Jang [1975]. The key condition for the validity of the gravitational weak equivalence principle is $\bar{\nabla}_\nu \mathcal{E}_\mu{}^\nu = 0$, where $\mathcal{E}_\mu{}^\nu = 0$ are the linearised (divergenceless) vacuum equations of motion of the theory and $\bar{\nabla}_\nu$ denotes the background covariant derivative. The idea is that the linearised $\mathcal{E}_\mu{}^\nu$ corresponds to the gravitational perturbation generated by the test particle. Then, the condition $\mathcal{E}_\mu{}^\nu = 0$ together with the Geroch-Jang theorem allows us to find a geodesic for the motion of the body in the unperturbed background spacetime. In this way, it encodes the universality of the geodesic motion for self-gravitating bodies and, thus, the gravitational weak equivalence principle. If we take $\mathcal{E}_\mu{}^\nu = 0$ to be a perturbation of the divergenceless equations (1.20) for vacuum Weyl transverse gravity and consider $\bar{\nabla}_\nu$ to be the Weyl invariant covariant divergence with respect to the background (see the discussion above regarding the geodesics), we indeed find $\bar{\nabla}_\nu \mathcal{E}_\mu{}^\nu = 0$ (we show the calculation in appendix A.1). Therefore, just like general relativity, Weyl transverse gravity obeys the gravitational weak equivalence principle. As for the general WTDiff-invariant theories, only the Lovelock ones (purely metric theories with second order equations of motion) incorporate the gravitational weak equivalence principle. We can show that by a simple modification of the arguments presented for Diff-invariant gravity di Casola et al. [2014]. Therefore, Weyl transverse gravity and general relativity appear to be the only two metric gravitational theories compatible with the gravitational weak equivalence principle.

Finally, the strong equivalence principle generalises the Einstein equivalence principle to test gravitational physics "All test fundamental physics (including gravitational physics) is not affected, locally, by the presence of a gravitational field" di Casola et al. [2015]. It essentially has the same relation to the Einstein equivalence principle as the generalised weak equivalence principle to the weak equivalence principle. As such, we expect it to be obeyed by Weyl transverse gravity and WTDiff-invariant versions of the other Lovelock theories, but not by any other (known) WTDiff-invariant gravitational theory[5].

1.2 Covariant phase space formalism

One of the crucial results presented in this thesis is the construction of the covariant phase space formalism for arbitrary local, WTDiff-invariant theories of gravity. This formalism offers a method to construct the phase space as a symplectic manifold without the need to break the general covariance. The covariant phase space formalism also provides powerful tools for computing conserved quantities Wald [1993], Iyer and Wald [1994], Wald and Zoupas [2000]. Moreover, it allows for a straightforward check of (non)-equivalence of different theories Margalef-Bentabol and Villaseñor [2021], Fernando Barbero G. et al. [2021], Barbero G. et al. [2022]. Herein, we recall the basics of the covariant phase space formalism in general and, in particular, its applications to Diff-invariant theories. We first introduce the construction of the symplectic structure. Then, we show how to express the symplectic form in terms of the conserved currents and charges. Lastly, we apply the general formalism to the particular case of Killing horizons in arbitrary diffeomorphism invariant theories of gravity.

1.2.1 General formalism

Consider a manifold $\mathcal{V}$ equipped with a volume form $\boldsymbol{\varepsilon}$. We define a covariant derivative on this manifold by condition $\nabla_\mu \boldsymbol{\varepsilon} = 0$ (obviously, ∇_μ is non-unique). We introduce a Lagrangian L as a local functional of a collection of dynamical

[5]While a detailed proof of such a statement would be quite involved, it essentially follows from our arguments for the Einstein and the gravitational weak equivalence principles. Indeed, it has been conjectured that the simultaneous validity of both principles together already implies the strong equivalence principle di Casola et al. [2015]. Nevertheless, one should keep in mind that the validity of the strong (and even Einstein) equivalence principle represents a very subtle issue even in the context of general relativity.

variables ϕ; non-dynamical variables γ, and their covariant derivatives. Non-dynamical variables are fixed a priori, whereas the dynamical variables are governed by the equations of motion. Under their small, arbitrary variation $\delta_1\phi$, the Lagrangian changes as

$$\delta_1 L = A_\phi \delta_1 \phi + \nabla_\mu \theta^\mu \left[\delta_1\right], \tag{1.42}$$

where $A_\phi = 0$ are the equations of motion for the fields ϕ. The Gauss theorem ensures that $\nabla_\mu \theta^\mu \left[\delta_1\right]$ contributes only with a boundary term to the variation of the action

$$I = \int_{\mathcal{V}} L \boldsymbol{\varepsilon}. \tag{1.43}$$

We call θ^μ the symplectic potential current density (or, for brevity, symplectic potential) for reasons that will become apparent soon Lee and Wald [1990].

Now we consider a second arbitrary variation of the dynamical variables $\delta_2\phi$. The commutator of both variations acting on $\delta_1 L$ gives

$$\left(\delta_1\delta_2 - \delta_2\delta_1\right) L = \delta_1 A_\phi \delta_2 \phi - \delta_2 A_\phi \delta_1 \phi + \nabla_\mu \Omega^\mu \left[\delta_1, \delta_2\right], \tag{1.44}$$

where

$$\Omega^\mu \left[\delta_1, \delta_2\right] = \delta_1 \theta^\mu \left[\delta_2\right] - \delta_2 \theta^\mu \left[\delta_1\right], \tag{1.45}$$

is called the symplectic current density (or symplectic current). Defining $\Omega\left[\delta_1, \delta_2\right]$ as an integral of $\Omega^\mu\left[\delta_1, \delta_2\right]$ over a suitable initial data surface $\mathcal{C}$, i.e.,

$$\Omega\left[\delta_1, \delta_2\right] = \int_{\mathcal{C}} \Omega^\mu \left[\delta_1, \delta_2\right] \mathrm{d}\mathcal{C}_\mu, \tag{1.46}$$

we can introduce a 2-form

$$\boldsymbol{\Omega} = \Omega_{AB} \delta^A \phi \delta^B \phi. \tag{1.47}$$

Defining similarly a 1-form

$$\boldsymbol{\theta} = \theta_A \delta^A \phi = \int_{\mathcal{C}} \theta^\mu \left[\delta\right] \mathrm{d}\mathcal{C}_\mu, \tag{1.48}$$

we have $\boldsymbol{\Omega} = \mathrm{d}\boldsymbol{\theta}$. It is then easy to see that $\boldsymbol{\Omega}$ is antisymmetric and closed. Hence, it satisfies two of the three requirements for a symplectic form. The third requirement, its non-degeneracy, is not in general satisfied. To ensure it, we must restrict $\boldsymbol{\Omega}$ to the subspace of field configurations ϕ satisfying all the constraints of the theory. This subspace together with the symplectic form $\boldsymbol{\Omega}$ then forms a symplectic manifold corresponding to the covariant phase space. For a more detailed exposition of this procedure, see e.g. Lee and Wald [1990]. Unless specified otherwise, we always assume that such a restriction of $\boldsymbol{\Omega}$ has been performed.

Hamiltonian

In the special case when one of the field variations corresponds to a Lie derivative with respect to the vector field ξ^μ, i.e., $\delta_1\phi = \pounds_\xi \phi$ (for simplicity, we also write $\delta_2\phi = \delta\phi$), the Hamilton equations of motion imply

$$\delta H_\xi = \Omega\left[\pounds_\xi, \delta\right], \tag{1.49}$$

provided that the Hamiltonian H_ξ corresponding to the evolution along ξ^μ exists. Thus, the symplectic form directly offers a prescription for a small variation of the Hamiltonian. If ξ^μ generates a global symmetry of all ϕ, it follows that H_ξ is conserved and, therefore, $\delta H_\xi = 0$. Equation (1.49) then represents a useful constraint on small changes of fields ϕ, as we explore in the following.

Let us now discuss when is the Hamiltonian H_ξ well defined. The necessary condition turns out to be Wald and Zoupas [2000]

$$\int_{\partial\mathcal{C}} \Omega^\mu \left[\pounds_\xi, \delta\right] \xi^\nu \mathrm{d}\mathcal{C}_{\mu\nu} = 0, \tag{1.50}$$

where $\partial\mathcal{C}$ denotes the boundary of the initial data surface $\mathcal{C}$. To see it, consider some solution of the equations of motion for the fields ϕ. For this solution, select a vector field ξ^μ, such that the Hamiltonian H_ξ exists. For two independent variations of ϕ's, we have $\left(\delta_1\delta_2 - \delta_2\delta_1\right) H_\xi = 0$. If we expand this expression using equations (1.49), (1.45) and (1.48), we obtain[6]

$$0 = \left(\delta_1\delta_2 - \delta_2\delta_1\right) H_\xi = \delta_1 \pounds_\xi \boldsymbol{\theta}\left[\delta_2\right] - \delta_2 \pounds_\xi \boldsymbol{\theta}\left[\delta_1\right]. \tag{1.51}$$

The Cartan magic formula $\pounds_\xi \boldsymbol{\theta} = \xi \cdot \mathrm{d}\boldsymbol{\theta} + \mathrm{d}\left(\xi \cdot \boldsymbol{\theta}\right)$ then implies

$$\mathrm{d}\left(\xi \cdot \delta_1 \boldsymbol{\theta}\left[\delta_2\right] - \xi \cdot \delta_2 \boldsymbol{\theta}\left[\delta_2\right]\right) = 2\int_{\mathcal{C}} \nabla_\nu \left(\xi^{[\nu}\Omega^{\mu]}\left[\pounds_\xi, \delta\right]\right) \mathrm{d}\mathcal{C}_\mu. \tag{1.52}$$

Finally, applying the Gauss theorem yields condition (1.50).

Noether charges

Noether theorems assert that to every symmetry of a Lagrangian L corresponds a conserved quantity. To construct these quantities, we consider a variation $\hat{\delta}\phi$ corresponding to a local symmetry of fields ϕ. Then, the change in the Lagrangian is a total divergence, $\hat{\delta} L = \nabla_\mu \alpha^\mu \left[\hat{\delta}\right]$. Defining a vector

$$j^\mu\left[\hat{\delta}\right] = \theta^\mu\left[\hat{\delta}\right] - \alpha^\mu\left[\hat{\delta}\right], \tag{1.53}$$

we can easily check that its covariant divergence is $\nabla_\mu j^\mu\left[\hat{\delta}\right] = -A_\phi \hat{\delta}\phi$. Therefore, the vector $j^\mu\left[\hat{\delta}\right]$ is divergence-free on shell. We call $j^\mu\left[\hat{\delta}\right]$ the Noether current corresponding to local symmetry $\hat{\delta}$. The conserved Noether charge associated with j^μ is just its integral over the initial data surface $\mathcal{C}$

$$Q = \int_{\mathcal{C}} j^\mu \mathrm{d}\mathcal{C}_\mu. \tag{1.54}$$

As we will see, in the case of Diff-invariant theories, the Noether charges can be used to express the symplectic form. Equation $\delta H_\xi = \Omega\left[\pounds_\xi, \delta\right]$ for a global symmetry then provides a relation between small changes of the various conserved charges.

[6] We stress that $\delta\xi^\mu = 0$ by definition.

1.2.2 Application to diffeomorphism invariant theories of gravity

Upon introducing the covariant phase space formalism, we review its application to arbitrary local, Diff-invariant theories of gravity. We also use this formalism to find the first law of black hole mechanics and a (heuristic) prescription for black hole entropy in modified theories of gravity.

Covariant phase space formalism for diffeomorphism invariant theories of gravity

We consider the most general local, Diff-invariant gravitational action in n spacetime dimensions

$$\begin{aligned} I =& \frac{1}{16\pi}\int_{\mathcal{V}}\Big[L\Big(g_{\mu\nu}, R^{\mu}{}_{\nu\rho\sigma}, \nabla_{\alpha_1}R^{\mu}{}_{\nu\rho\sigma}, ..., \nabla_{(\alpha_1}...\nabla_{\alpha_p)}R^{\mu}{}_{\nu\rho\sigma}, \\ & \psi, \nabla_{\alpha_1}\psi, ..., \nabla_{(\alpha_1}...\nabla_{\alpha_p)}\psi\Big) - 2\Lambda\Big]\sqrt{-\mathfrak{g}}\mathrm{d}^n x, \end{aligned} \tag{1.55}$$

where $\mathcal{V}$ denotes the spacetime manifold, Λ is the cosmological constant, and ψ are the matter fields (for simplicity of notation, we suppress their spacetime indices). We stress that any matter fields with gauge symmetries require separate treatment Prabhu [2017], Elgood et al. [2020]. The indices of covariant derivatives can be fully symmetrised without any loss of generality, since any other combinations of derivatives can be removed using the Bianchi identities.

Varying the action with respect to the metric and the matter fields yields

$$\delta I = \int_{\mathcal{V}}\left[\frac{\sqrt{-\mathfrak{g}}}{16\pi}\left(A^{\mu\nu} - \Lambda g^{\mu\nu}\right)\delta g_{\mu\nu} + A_{\psi}\delta\psi + \nabla_{\mu}\theta^{\mu}\left[\delta\right]\right]\mathrm{d}^n x. \tag{1.56}$$

The first term gives the equations of motion for the metric[7]. The second term corresponds to the equations of motion for the matter fields. The last term contributes only with a boundary integral, with $\theta^{\mu}\left[\delta\right]$ being the symplectic potential of the theory. The symplectic current and symplectic form are then defined by equations (1.45) and (1.46). Hence, we now have everything we need to work out the explicit form of a variation of the Hamiltonian H_{ξ} corresponding to the evolution along some vector field ξ^{μ}. While this direct method for finding δH_{ξ} in principle works, we can also use that our theory is diffeomorphism invariant, and conveniently express δH_{ξ} in terms of the Noether charges corresponding to the diffeomorphism generated by ξ^{μ}.

The Noether current associated with an infinitesimal diffeomorphism generated by ξ^{μ} reads

$$j^{\mu}_{\xi} = \theta^{\mu}\left[\pounds_{\xi}\right] - L\xi^{\mu}, \tag{1.57}$$

Since its divergence vanishes on shell, it follows that j^{μ}_{ξ} is given by the sum of terms proportional to the equations of motion and a divergence of an antisymmetric rank 2 tensor density Wald [1990]. An explicit calculation shows that

$$j^{\mu}_{\xi} = -\frac{\sqrt{-\mathfrak{g}}}{8\pi}\left(A_{\nu}{}^{\mu} - \Lambda\delta^{\mu}_{\nu}\right)\xi^{\nu} - \left(A_{\psi}\cdot\psi\cdot\xi\right)^{\mu} + \nabla_{\nu}Q^{\nu\mu}_{\xi}. \tag{1.58}$$

[7] Following the seminal papers Wald [1993], Iyer and Wald [1994], we vary the Lagrangian with respect to $g_{\mu\nu}$ rather than $g^{\mu\nu}$. Consequently, the equations of motion of general relativity are $-G_{\mu\nu} - \Lambda g_{\mu\nu} + 8\pi T_{\mu\nu} = 0$ in our convention.

where $A_{\mu\nu} - \Lambda g_{\mu\nu} = 0$ are the equations of motion for the metric, and $A_\psi = 0$ the equations of motion for the matter fields[8]. The antisymmetric tensor density $Q_\xi^{\nu\mu}$ corresponds to the Noether charge. It equals

$$Q_\xi^{\nu\mu} = 2E^{\nu\mu\rho\sigma}\nabla_\rho\xi_\sigma + W_\rho{}^{\nu\mu}\xi^\rho, \tag{1.59}$$

where

$$E_\mu{}^{\nu\rho\sigma} = \sum_{i=0}^{p} (-1)^i \nabla_{\alpha_1}...\nabla_{\alpha_i} \left(\frac{\partial L}{\partial \nabla_{(\alpha_1}...\nabla_{\alpha_i)} R^\mu{}_{\nu\rho\sigma}}\right), \tag{1.60}$$

and $W_\rho{}^{\nu\mu} = W_\rho{}^{[\nu\mu]}$ is a rank 3 tensor density whose precise form does not matter for our purposes. Notably, while $W_\rho{}^{\nu\mu}$ depends both on gravitational and matter fields, the tensor density $E^{\nu\mu\rho\sigma}$ is fully determined by the vacuum gravitational Lagrangian Iyer and Wald [1994], Iyer [1997].

To relate Noether charges $Q_\xi^{\nu\mu}$ with the symplectic form, we first study a small perturbation of the Noether current, δj_ξ^μ. If we assume that the equations of motion are satisfied both before and after the perturbation, we have, starting from the definition of j_ξ^μ (1.57)

$$\delta j_\xi^\mu = \delta\theta^\mu \left[\pounds_\xi\right] - \xi^\mu \nabla_\nu \theta^\nu \left[\delta\right]. \tag{1.61}$$

At the same time, it holds $\delta j_\xi^\mu = \nabla_\nu \delta Q_\xi^{\nu\mu}$. Putting both expressions for δj_ξ^μ together, we have after some straightforward manipulations,

$$\nabla_\nu \delta Q_\xi^{\nu\mu} = \Omega^\mu \left[\delta, \pounds_\xi\right] + 2\nabla_\nu \left(\xi^{[\nu}\theta^{\mu]}\left[\delta\right]\right), \tag{1.62}$$

where the symplectic current $\Omega^\mu\left[\delta, \pounds_\xi\right]$ is given by equation (1.45). Integrating this relation over some Cauchy surface $\mathcal{C}$ then yields the symplectic form which is equal to the perturbation of the Hamiltonian H_ξ, i.e.,

$$\delta H_\xi = \Omega\left[\delta, \pounds_\xi\right] = \int_{\mathcal{C}} \Omega^\mu \left[\delta, \pounds_\xi\right] \mathrm{d}\mathcal{C}_\mu = \int_{\partial\mathcal{C}} \left(\delta Q_\xi^{\nu\mu} - 2\xi^{[\nu}\theta^{\mu]}\left[\delta\right]\right) \mathrm{d}\mathcal{C}_{\mu\nu}, \tag{1.63}$$

where we used the Gauss theorem to convert the right hand side to an integral over the boundary $\partial\mathcal{C}$ of the Cauchy surface $\mathcal{C}$.

From equation (1.63) for δH_ξ, we can infer the expression for the full Hamiltonian H_ξ

$$H_\xi = \int_{\partial\mathcal{C}} \left(Q_\xi^{\nu\mu} - 2\xi^{[\nu}B^{\mu]}\right) \mathrm{d}\mathcal{C}_{\mu\nu}, \tag{1.64}$$

where B^μ obeys $\theta^\mu\left[\delta\right] = \delta B^\mu$ (B^μ need not be covariant). This establishes an important result: if the Hamiltonian corresponding to the evolution along the vector field ξ^μ exists for a Diff-invariant theory, it can always be given as an integral over the boundary of the Cauchy surface. In the next section we show that this result no longer holds for WTDiff-invariant theories.

[8]For simplicity of notation, we suppress the index structure of the matter fields. The expression $(\psi \cdot A_\psi \cdot \xi)^\mu$ should be understood as $\psi A_\psi \xi^\mu$ for scalar fields and $2\psi_\nu A_\psi^{(\mu}\xi^{\nu)}$ for vector fields. More general tensorial fields need to be treated separately.

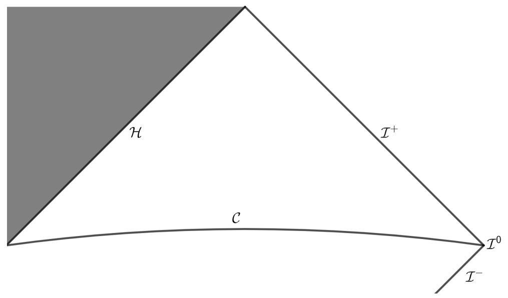

Figure 1.1: Relevant part of the Penrose diagram of the stationary black hole spacetime we analyse. $\mathcal{H}$ denotes the horizon, $\mathcal{I}^+$ and $\mathcal{I}^-$ are the future and past null infinity, and $\mathcal{I}^0$ the spatial infinity. We draw the Cauchy surface $\mathcal{C}$ for the exterior region as the oblique line extended from $\mathcal{H}$ to $\mathcal{I}^+$. The grey region represents the interior of the black hole, whose structure we do not specify.

First law of black hole mechanics and Wald entropy

Expression (1.63) for δH_ξ becomes especially useful if ξ^μ is a Killing vector, i.e., $\pounds_\xi g_{\mu\nu} = 0$. We show its usefulness on the physically interesting case of a stationary, asymptotically flat spacetime containing a single black hole in four spacetime dimensions. We display the corresponding Penrose diagram in figure 1.1. The spacetime possesses a Killing vector $\xi^\mu = t^\mu + \Omega_\mathcal{H}\varphi^\mu$, where t^μ denotes the time translational Killing vector, φ^μ the rotational Killing vector, and $\Omega_\mathcal{H}$ is the constant angular velocity of the black hole event horizon. The vector ξ^μ has timelike norm outside the horizon, null norm on the horizon, and spacelike inside it. Hence, the event horizon coincides with the Killing horizon. We evaluate equation (1.63) on a spacelike surface $\mathcal{C}$ orthogonal to ξ^μ and extended from the spatial infinity $\mathcal{I}^0$ to the horizon $\mathcal{H}$. If the null energy condition holds (notably, this implies absence of Hawking radiation), $\mathcal{C}$ represents a Cauchy surface for the region outside the horizon. Then, the details of the internal structure of the black hole, e.g. the presence of a Cauchy horizon or a singularity, are irrelevant for our purposes.

Let us now expand equation (1.63) in the above described setting. Since ξ^μ is a Killing vector and $\pounds_\xi g_{\mu\nu} = 0$, it follows that $\delta H_\xi = \Omega\left[\pounds_\xi, \delta\right] = 0$. To evaluate the right hand side, we split the boundary in two pieces, the intersection of the Cauchy surface with the asymptotic infinity $\partial\mathcal{C}_\infty$, and with the horizon $\partial\mathcal{C}_\mathcal{H}$. In total, equation (1.63) becomes

$$\begin{aligned}&\int_{\partial\mathcal{C}_\infty}\left(\delta Q_t^{\nu\mu} - 2t^{[\nu}\theta^{\mu]}\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu} + \Omega_\mathcal{H}\int_{\partial\mathcal{C}_\infty}\delta Q_\varphi^{\nu\mu}\mathrm{d}\mathcal{C}_{\mu\nu}\\&-\int_{\partial\mathcal{C}_\mathcal{H}}\left(\delta Q_\xi^{\nu\mu} - 2\xi^{[\nu}\theta^{\mu]}\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu} = 0,\end{aligned} \tag{1.65}$$

where we split the integral over $\partial\mathcal{C}_\infty$ into the time translational and rotational contributions and used that $\varphi^\nu\mathrm{d}\mathcal{C}_{\mu\nu} = 0$ there. The first two terms define the variations of the canonical energy

$$\delta M = \int_{\partial\mathcal{C}_\infty}\left(\delta Q_t^{\nu\mu} - 2t^{[\nu}\theta^{\mu]}\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu}, \tag{1.66}$$

and the canonical angular momentum

$$\delta\ J = -\int_{\partial\mathcal{C}_\infty}\delta Q_\varphi^{\nu\mu}\mathrm{d}\mathcal{C}_{\mu\nu}, \tag{1.67}$$

of the spacetime. For general relativity, both expressions reduce to the corresponding ADM formulas Iyer and Wald [1994]. For the integral over the horizon we use the explicit form (1.59) of the Noether charge, obtaining

$$-\int_{\partial\mathcal{C}_{\mathcal{H}}}\left(\delta Q_{\xi}^{\nu\mu}-2\xi^{[\nu}\theta^{\mu]}\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu}=-\int_{\partial\mathcal{C}_{\mathcal{H}}}\Big[\delta\Big(2E^{\nu\mu\rho\sigma}\nabla_{\rho}\xi_{\sigma}+W_{\rho}{}^{\nu\mu}\xi^{\rho}\Big)-2\xi^{[\nu}\theta^{\mu]}\left[\delta\right]\Big]\mathrm{d}\mathcal{C}_{\mu\nu}. \tag{1.68}$$

Note that the horizon $\mathcal{H}$ can be analytically extended to contain a bifurcation surface at which ξ^{μ} vanishes and $\nabla_{\rho}\xi_{\sigma}=\kappa\epsilon_{\rho\sigma}$, where $\kappa=\sqrt{-\nabla_{\rho}\xi_{\sigma}\nabla^{\rho}\xi^{\sigma}/4}|_{\mathcal{H}}$ is the surface gravity of the black hole, and $\epsilon_{\rho\sigma}$ denotes the bi-normal to the horizon. Since ξ^{μ} is a Killing vector, expression (1.68) has the same value at any cross-section of the horizon orthogonal to ξ^{μ}. Thus, we are free to evaluate it at the bifurcation surface Jacobson et al. [1994], where the last two terms disappear[9] and the first term simplifies to $-\int_{\partial\mathcal{C}_{\mathcal{H}}}\delta\left(2\kappa E^{\nu\mu\rho\sigma}\epsilon_{\rho\sigma}\right)\epsilon_{\mu\nu}\mathrm{d}^{n-2}\mathcal{A}$, where $\mathrm{d}^{n-2}\mathcal{A}$ denotes the area element on $\partial\mathcal{C}_{\mathcal{H}}$. Assuming $\delta\kappa=0$, we have in total

$$\delta M-\Omega_{\mathcal{H}}\delta J-\frac{\kappa}{2\pi}\int_{\partial\mathcal{C}_{\mathcal{H}}}\delta\left(4\pi E^{\nu\mu\rho\sigma}\epsilon_{\rho\sigma}\right)\epsilon_{\mu\nu}\mathrm{d}^{n-2}\mathcal{A}=0. \tag{1.69}$$

This equation is the first law of black hole mechanics for a stationary, asymptotically flat black hole spacetime in an arbitrary local, Diff-invariant theory of gravity Wald [1993], Iyer [1997]. For general relativity, it reduces to the expression derived in the seminal papers concerning the laws of black hole mechanics Bekenstein [1973], Bardeen et al. [1973].

The analysis in curved spacetime quantum field theory shows that black holes radiate at temperature $T_{\mathrm{H}}=\kappa/2\pi$ known as the Hawking temperature Hawking [1975], Visser [2003]. If we identify T_{H} in the first law of black hole mechanics (1.69), we may rewrite it as a genuine first law of thermodynamics

$$\delta M=T_{\mathrm{H}}\delta S_{\mathrm{W}}+\Omega_{\mathcal{H}}\delta J, \tag{1.70}$$

where δM plays the role of internal energy (or enthalpy Kubiznak and Mann [2015]), $\Omega_{\mathcal{H}}\delta J$ represents a work term appearing due to the change in black hole rotation, and

$$\delta S_{\mathrm{W}}=4\pi\int_{\partial\mathcal{C}_{\mathcal{H}}}E^{\nu\mu\rho\sigma}\epsilon_{\rho\sigma}\epsilon_{\mu\nu}\mathrm{d}^{n-2}\mathcal{A}, \tag{1.71}$$

is the Wald entropy of the black hole event horizon. In general relativity, Wald entropy is proportional to the area $\mathcal{A}$ of the horizon's cross-section, i.e., $S_{\mathrm{W}}=\mathcal{A}/4$, which is just the famous Bekenstein entropy Bekenstein [1973]. It has been argued that Wald entropy represents the appropriate generalisation of Bekenstein entropy for modified theories of gravity Wald [1993]. It indeed satisfies the second law of thermodynamics, although the entropy prescription needs to be refined to account for dynamically changing black hole horizons Wall [2012], Dong [2014], Wall [2015], Hollands et al. [2024].

The covariant phase space formalism yields not only the first law, but also the Smarr formula, a mathematical identity relating the parameters of a black

[9] Unless $W_{\rho}{}^{\nu\mu}$ or θ^{μ} diverge at the bifurcation surface Jacobson et al. [1994], Sarkar and Wall [2011].

hole. One just needs to integrate the on-shell relation $j^\mu_\xi = \nabla_\nu Q^{\nu\mu}_\xi$ over a Cauchy surface $\mathcal{C}$. The integration proceeds along the same lines as for the first law and we find the following Smarr formula

$$\mathcal{M} = 2T_{\mathrm{H}}S_{\mathrm{W}} + 2\Omega_{\mathcal{H}}\mathcal{J}. \tag{1.72}$$

This identity can be easily verified by plugging in the parameters of the Kerr metric (the most general vacuum stationary, asymptotically flat black hole spacetime in four dimensional general relativity).

Covariant phase space formalism for WTDiff-invariant gravity

Given the status of Weyl transverse gravity as a classical alternative to general relativity, it is of interest to also apply the covariant phase space formalism to it and to derive expressions for the conserved charges as well as the first law of black hole mechanics.

Unfortunately, the standard treatment of the covariant phase space formalism works only for local, fully Diff-invariant theories of gravity without any non-dynamical structures Iyer and Wald [1994]. An extension of the formalism has also been applied to Einstein-aether gravity which breaks the local Lorentz invariance by introducing a preferred direction of time Foster [2006], Berglund et al. [2012], Pacilio and Liberati [2017], Ho et al. [2018], Ding and Zhai [2020], Kucukakca and Akbarieh [2020], Chan et al. [2022]. However, since the time direction is introduced as a dynamical vector (or scalar) field and the action is written in a covariant form, the loss of the local Lorentz symmetry does not really represent any technical difficulties for the covariant phase space approach. The situation in WTDiff-invariant gravity is somewhat more subtle. On the one side, the presence of a non-dynamical volume n-form $\boldsymbol{\omega}$ requires careful treatment. On the other side, the cosmological constant appears as a global degree of freedom and is only meaningfully defined on shell. We were able to address both challenges, obtaining the symplectic structure for general local, WTDiff-invariant theories of gravity. The results are mostly physically equivalent to the ones found in the Diff-invariant case. However, the different behaviour of the cosmological constant leads to a volume term being present in the Hamiltonian. This marks a departure from Diff-invariant theories, where the Hamiltonian always reduces to a surface integral.

Herein, we discuss the construction of the WTDiff-invariant covariant phase space formalism in detail. Furthermore, we apply it to derive the first law of black hole mechanics, an expression for Wald entropy, and to study thermodynamics of causal diamonds. First, to introduce our approach, we treat the special case of Weyl transverse gravity. Then, we move on to the fully general local, WTDiff-invariant theories.

1.3 Covariant phase space formalism for Weyl transverse gravity

Before going to the most general local, WTDiff-invariant gravitational action, we introduce the covariant phase space formalism for vacuum Weyl transverse gravity. We begin by varying the Lagrangian (1.13) with respect to the dynamical metric $g_{\mu\nu}$,

$$\delta L = -\frac{1}{16\pi}\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\tilde{R}^{\mu\nu} - \frac{1}{n}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{-2/n}\tilde{R}\tilde{g}^{\mu\nu}\right]\delta g_{\mu\nu} + \tilde{\nabla}_\mu\theta^\mu. \quad (1.73)$$

This expression contracted with $\delta g_{\mu\nu}$ corresponds to the vacuum equations of motion. The second term yields the symplectic potential

$$\theta^{\mu}[\delta] = \frac{1}{16\pi}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{4}{n}} \left(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\sigma}g^{\nu\rho}\right)\tilde{\nabla}_{\sigma}\delta\tilde{g}_{\nu\rho}, \tag{1.74}$$

where the variation of the auxiliary metric $\delta\tilde{g}_{\nu\rho}$ reads

$$\delta\tilde{g}_{\nu\rho} = \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{-\frac{2}{n}} \left(\delta g_{\nu\rho} - \frac{2}{n}g_{\nu\rho}\delta\ln\frac{\sqrt{-\mathfrak{g}}}{\omega}\right). \tag{1.75}$$

One can easily verify that $\theta^{\mu}[\delta]$ is WTDiff-invariant. The symplectic current $\Omega^{\mu}[\delta_1, \delta_2]$ and the symplectic form $\Omega[\delta_1, \delta_2]$ are then simply given by the general expressions (1.45) and (1.46), respectively.

As in the Diff-invariant case, we want to relate the symplectic form with the Noether currents and charges associated with the local symmetries of the theory, i.e., Weyl transformations and transverse diffeomorphisms. We start with the Noether current corresponding to an infinitesimal Weyl transformation, $\delta_{\mathrm{W}}g_{\mu\nu} = 2\sigma g_{\mu\nu}$. We recall the general definition of the Noether current

$$j^{\mu}\Big[\delta_{\mathrm{W}}\Big] = \theta^{\mu}\Big[\delta_{\mathrm{W}}\Big] - \alpha^{\mu}\Big[\delta_{\mathrm{W}}\Big], \tag{1.76}$$

where $\alpha^{\mu}[\delta_{\mathrm{W}}]$ obeys $\tilde{\nabla}_{\mu}\alpha^{\mu}[\delta_{\mathrm{W}}] = \delta_{\mathrm{W}}L$, with L being the Lagrangian. Since action (1.13) is exactly Weyl invariant, $\delta_{\mathrm{W}}L = 0$, and $\alpha^{\mu}[\delta_{\mathrm{W}}]$ vanishes. The symplectic potential $\theta^{\mu}[\delta_{\mathrm{W}}]$ is proportional to the variation of the auxiliary metric, $\delta_{\mathrm{W}}\tilde{g}_{\mu\nu}$. However, the auxiliary metric does not change under Weyl transformation. Thence, $\theta^{\mu}[\delta_{\mathrm{W}}] = 0$ and, in total, the Noether current $j^{\mu}[\delta_{\mathrm{W}}]$ vanishes identically. This result has been obtained in different ways by several authors, leading them to call Weyl invariance "a fake symmetry" Jackiw and Pi [2015], Oda [2017, 2022]. It has also been argued that the vanishing Noether current plays a key role in the absence of a quantum anomaly for this local Weyl symmetry Oda [2017]. The absence of the anomaly is then crucial for the explicit radiative stability of the cosmological constant in Weyl transverse gravity Carballo-Rubio [2015].

For transverse diffeomorphisms, we have $\pounds_{\xi}L = \xi^{\mu}\tilde{\nabla}_{\mu}L$, implying $\alpha^{\mu}_{\xi} = L\xi^{\mu}$ (where we used the transversality condition $\tilde{\nabla}_{\mu}\xi^{\mu} = 0$). Putting this together with the symplectic potential $\theta^{\mu}[\pounds_{\xi}]$ we obtain, after some straightforward manipulations

$$j^{\mu}_{\xi} = \frac{1}{8\pi}\left(\tilde{g}^{\mu\rho}\tilde{R}_{\rho\nu} - \frac{1}{2}\tilde{R}\delta^{\mu}_{\nu}\right)\xi^{\nu} + \tilde{\nabla}_{\nu}\left[\frac{1}{8\pi}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\tilde{\nabla}^{[\nu}\xi^{\mu]}\right], \tag{1.77}$$

where the second term corresponds to the divergence of the Noether charge antisymmetric tensor

$$Q^{\nu\mu}_{\xi} = \frac{1}{8\pi}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}}\tilde{\nabla}^{[\nu}\xi^{\mu]}, \tag{1.78}$$

which is explicitly WTDiff-invariant.

On shell, we have the following expression for the Noether current

$$j^{\mu}_{\xi} = -\frac{1}{8\pi}\Lambda\xi^{\mu} + \tilde{\nabla}_{\nu}\left[\frac{1}{8\pi}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\tilde{\nabla}^{[\nu}\xi^{\mu]}\right]. \tag{1.79}$$

It is easy to verify that the WTDiff-invariant divergence of the Noether current vanishes.

Let us now compare our result for the Noether current in Weyl transverse gravity with general relativity. We can do so by fixing the unimodular gauge, $\sqrt{-\mathfrak{g}} = \omega$, for the metric, in which the equations of motion of Weyl transverse gravity reduce to the traceless part of the Einstein equations. In this gauge, the Noether current for Weyl transverse gravity simplifies to

$$j^{\mu}_{\xi} = \frac{1}{8\pi}\sqrt{-\mathfrak{g}}\left(R_{\nu}{}^{\mu} - \frac{1}{2}R\delta^{\mu}_{\nu}\right)\xi^{\rho} + \frac{1}{8\pi}\sqrt{-\mathfrak{g}}\nabla_{\nu}\nabla^{[\nu}\xi^{\mu]}, \tag{1.80}$$

whereas the Noether current for transverse diffeomorphisms in general relativity reads

$$j^{\mu}_{\mathrm{GR},\xi} = \frac{1}{8\pi}\sqrt{-\mathfrak{g}}\left(R_{\nu}{}^{\mu} - \frac{1}{2}R\delta^{\mu}_{\nu} + \Lambda\delta^{\mu}_{\nu}\right)\xi^{\nu} + \frac{1}{8\pi}\sqrt{-\mathfrak{g}}\nabla_{\nu}\nabla^{[\nu}\xi^{\mu]}. \tag{1.81}$$

Comparing both expressions, the Noether current for Weyl transverse gravity lacks the term $\sqrt{-\mathfrak{g}}\Lambda\xi^{\mu}/8\pi$. In general relativity, this contribution enters the Noether current through the Lagrangian term $-L\xi^{\mu}$ which contains Λ as a fixed parameter. By contrast, the Lagrangian of Weyl transverse gravity can include a fixed parameter λ, but, due to WTDiff-invariance, it does not affect the equations of motion. Instead, Λ appears as an arbitrary integration constant. If we do not set $\lambda = 0$, the Noether current for transverse diffeomorphisms in Weyl transverse gravity is

$$j^{\mu}_{\xi} = \frac{1}{8\pi}\left(\tilde{g}^{\mu\rho}\tilde{R}_{\rho\nu} - \frac{1}{2}\tilde{R}\delta^{\mu}_{\nu} + \lambda\delta^{\mu}_{\nu}\right)\xi^{\nu} + \tilde{\nabla}_{\nu}\left[\frac{1}{8\pi}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\tilde{\nabla}^{[\nu}\xi^{\mu]}\right]. \tag{1.82}$$

It might appear that setting $\lambda = \Lambda$ allows us to recover the Noether current of general relativity in the unimodular gauge. However, since Λ in Weyl transverse gravity arises as an integration constant in the process of solving the equations of motion, it is only defined on shell. Moreover, Λ takes a different value for every solution. We obviously cannot set the fixed off shell parameter λ to match all the possible values of Λ. Therefore, the difference between the Noether currents in Weyl transverse gravity and general relativity is genuine and cannot be removed. In the following, we again keep the irrelevant parameter λ equal to zero, unless specified otherwise.

Let us note that the symplectic potential, the symplectic current, the Noether current, and the Noether charge are not completely unambiguous. First, adding a total divergence $\tilde{\nabla}_{\mu}\gamma^{\mu}$ to the Lagrangian leads to

$$\theta^{\mu}\left[\delta\right] \quad\rightarrow\quad \theta^{\mu}\left[\delta\right] + \delta\gamma^{\mu}, \tag{1.83}$$

$$\Omega^{\mu}\left[\delta_1,\delta_2\right] \quad\rightarrow\quad \Omega^{\mu}\left[\delta_1,\delta_2\right], \tag{1.84}$$

$$j^{\mu}_{\xi} \quad\rightarrow\quad j^{\mu}_{\xi} - 2\tilde{\nabla}_{\nu}\left(\xi^{[\nu}\gamma^{\mu]}\right), \tag{1.85}$$

$$Q^{\nu\mu}_{\xi} \quad\rightarrow\quad Q^{\nu\mu}_{\xi} - 2\xi^{[\nu}\gamma^{\mu]}. \tag{1.86}$$

Furthermore, we can shift the symplectic potential by a WTDiff-invariant divergence of an arbitrary antisymmetric rank 2 tensor density $\tilde{\nabla}_{\nu}Y^{\nu\mu}\left[\delta\right]$. We are also free to shift the Noether charge tensor by a divergence of any fully antisymmetric rank 3 tensor $\tilde{\nabla}_\lambda Z_\xi^{\lambda\nu\mu}$ Jacobson et al. [1994], Iyer and Wald [1994]. Altogether, the possible ambiguities are (same as in the Diff-invariant case)

$$\theta^\mu[\delta] \quad\rightarrow\quad \theta^\mu[\delta] + \delta\gamma^\mu + \tilde{\nabla}_\nu Y^{\nu\mu}[\delta], \tag{1.87}$$

$$\Omega^\mu[\delta_1,\delta_2] \quad\rightarrow\quad \Omega^\mu[\delta_1,\delta_2] + \tilde{\nabla}_\nu\left(\delta_1 Y^{\nu\mu}[\delta_2] - \delta_2 Y^{\nu\mu}[\delta_1]\right), \tag{1.88}$$

$$j_\xi^\mu \quad\rightarrow\quad j_\xi^\mu + 2\tilde{\nabla}_\nu\left(\xi^{[\nu}\gamma^{\mu]}\right) + \tilde{\nabla}_\nu Y^{\nu\mu}[\pounds_\xi], \tag{1.89}$$

$$Q_\xi^{\nu\mu} \quad\rightarrow\quad Q_\xi^{\nu\mu} + 2\xi^{[\nu}\gamma^{\mu]} + Y^{\nu\mu}[\pounds_\xi] + \tilde{\nabla}_\lambda Z_\xi^{\lambda\nu\mu}. \tag{1.90}$$

We discuss the potential physical implications of these ambiguities in the next subsection.

1.3.1 Hamiltonian for transverse diffeomorphisms

Upon defining the Noether currents and charges for transverse diffeomorphisms, we show how to relate them with the Hamiltonian corresponding to the evolution along the transverse diffeomorphism generator ξ^μ. To find this relation, we study a small perturbation of the Nother current j_ξ^μ. We assume that both the background spacetime and the perturbation solve the equations of motion. Then, we can vary the on shell expression for j_ξ^μ (1.79), obtaining (we stress that $\delta\xi^\mu = 0$)

$$\delta j_\xi^\mu = -\frac{1}{8\pi}\xi^\mu\delta\Lambda + \tilde{\nabla}_\nu\delta\left[\frac{1}{8\pi}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\tilde{\nabla}^{[\nu}\xi^{\mu]}\right]. \tag{1.91}$$

At the same time, we can also work with the general definition of the Noether current, $j_\xi^\mu = \theta[\pounds_\xi] - L\xi^\mu$. By varying it, after some manipulations and using the definition of the symplectic current (1.45), we find

$$\delta j_\xi^\mu = \Omega^\mu[\pounds_\xi,\delta] + 2\tilde{\nabla}_\nu\left(\xi^{[\nu}\theta^{\mu]}[\delta]\right). \tag{1.92}$$

Comparing expressions (1.91) and (1.92) for δj_ξ^μ, we get an equation for the symplectic current

$$\Omega^\mu[\pounds_\xi,\delta] = \tilde{\nabla}_\nu\left\{\delta\left[\frac{1}{8\pi}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\tilde{\nabla}^{[\nu}\xi^{\mu]}\right] - 2\xi^{[\nu}\theta^{\mu]}[\delta]\right\} - \frac{1}{8\pi}\xi^\mu\delta\Lambda. \tag{1.93}$$

Now, choose a Cauchy surface $\mathcal{C}$ in the unperturbed spacetime. We can define a WTDiff-invariant volume element on $\mathcal{C}$ by $\mathrm{d}\mathcal{C}_\mu = \left(\sqrt{-\mathfrak{g}}/\omega\right)^{-1/n} n_\mu\omega\mathrm{d}^{n-1}x$, with n_μ being a unit normal to $\mathcal{C}$ ($g_{\mu\nu}n^\mu n^\nu = \pm 1$) and $\mathrm{d}^{n-1}x$ a coordinate volume element on $\mathcal{C}$. Integrating equation (1.93) over the Cauchy surface $\mathcal{C}$ yields the symplectic form

$$\Omega[\pounds_\xi,\delta] = \int_{\partial\mathcal{C}}\left(\delta Q_\xi^{\nu\mu} - 2\xi^\nu\theta^\mu[\delta]\right)\mathrm{d}\mathcal{C}_{\mu\nu} - \int_{\mathcal{C}}\frac{1}{8\pi}\delta\Lambda\xi^\mu\mathrm{d}\mathcal{C}_\mu, \tag{1.94}$$

where we used the Gauss theorem to convert the first term into a boundary integral over $\partial\mathcal{C}$ and introduced a WTDiff-invariant surface element on $\partial\mathcal{C}$, $\mathrm{d}\mathcal{C}_{\mu\nu} = \left(\sqrt{-\mathfrak{g}}/\omega\right)^{-2/n} n_{[\mu}m_{\nu]}\omega\mathrm{d}^{n-2}x$, where m_μ denotes the unit normal to $\partial\mathcal{C}$ with respect to its embedding in $\mathcal{C}$ and $\mathrm{d}^{n-2}x$ is the coordinate area element on $\partial\mathcal{C}$.

If condition (1.50) is fulfilled, there exist a Hamiltonian H_ξ for the evolution along the vector field ξ^μ. By the Hamilton equations of motion, its perturbation equals the symplectic form $\Omega\left[\pounds_\xi, \delta\right]$, i.e.,

$$\delta H_\xi = \int_{\partial\mathcal{C}} \left(\delta Q_\xi^{\nu\mu} - 2\xi^\nu\theta^\mu\left[\delta\right]\right) \mathrm{d}\mathcal{C}_{\mu\nu} - \int_{\mathcal{C}} \frac{1}{8\pi}\delta\Lambda\xi^\mu \mathrm{d}\mathcal{C}_\mu. \tag{1.95}$$

In fully Diff-invariant theories, the variation of the Hamiltonian can always be written as a surface integral Wald [1993], Iyer and Wald [1994]. By contrast, in Weyl transverse gravity, we also have a volume integral proportional to $\delta\Lambda$. Since the volume integral clearly gives an infinite contribution and variations of Λ occur generically, we must suitably regularise the Hamiltonian (except for the special case $\Lambda = \delta\Lambda = 0$). We discuss the regularisation on some examples in subsection 1.3.3.

To conclude the general discussion of the variation of the Hamiltonian, we check whether it is affected by the ambiguities in $\theta^\mu\left[\delta\right]$, $\Omega^\mu\left[\delta_1, \delta_2\right]$, j_ξ^μ, and $Q_\xi^{\nu\mu}$ mentioned in the previous subsection. In the definition of the symplectic form $\Omega\left[\pounds_\xi, \delta\right]$, we have only the following ambiguity

$$\Delta\Omega\left[\pounds_\xi, \delta\right] = \int_{\mathcal{C}} \Delta\Omega^\mu\left[\pounds_\xi, \delta\right] \mathrm{d}\mathcal{C}_\mu = \int_{\partial\mathcal{C}} \left(\pounds_\xi Y^{\nu\mu}\left[\delta\right] - \delta Y^{\nu\mu}\left[\pounds_\xi\right]\right) \mathrm{d}\mathcal{C}_{\mu\nu}. \tag{1.96}$$

This term may affect the symplectic form in general Jacobson et al. [1994], Iyer and Wald [1994], Compère and Fiorucci [2018]. However, in the following, we will be mostly interested in the case of Killing vector fields ξ^μ. Then, by definition, $\pounds_\xi \tilde{g}_{\mu\nu} = 0$ and we have $\Delta\Omega\left[\pounds_\xi, \delta\right] = 0$ (the same reasoning applies even in the presence of matter fields). Hence, we can treat the Hamiltonian corresponding to a Killing vector field as unambiguous. In general, the ambiguity can be cured by applying a more sophisticated method known as the relative bi-complex framework, which goes beyond the scope of the present work Margalef-Bentabol and Villaseñor [2021, 2022].

1.3.2 The first law in vacuum and Wald entropy

We now have an expression for a variation of the Hamiltonian generating the evolution along a transverse diffeomorphism generator ξ^μ. As in subsection 1.2.2, we can apply this expression to Killing vectors in stationary black hole spacetimes. In this subsection, we focus on asymptotically flat black hole spacetimes with $\Lambda = \delta\Lambda = 0$. Since Weyl transverse gravity and general relativity lead to equivalent classical physics aside from the behaviour of Λ, we expect to recover the first law of black hole mechanics obtained in general relativity.

Any stationary, asymptotically flat black hole spacetime of dimension n possesses a Killing vector $\xi^\mu = t^\mu + \sum_{i=1}^{(n-1)/2} \Omega_{\mathcal{H},(i)}\varphi_{(i)}^\mu$, where t^μ denotes the time translational Killing vector, $\varphi_{(i)}^\mu$ are the rotational Killing vectors, and $\Omega_{\mathcal{H},(i)}$ denote the angular velocities of the black hole Killing horizon in various directions (the rigidity theorems guaranteeing constancy of $\Omega_\mathcal{H}$ translates directly from general relativity to Weyl transverse gravity). We stress that the Killing vectors are defined in a WTDiff-invariant manner, i.e., the Killing equations reads

$\pounds_\xi \tilde{g}_{\mu\nu} = 0$[10]. Likewise, the horizons (and, indeed, the entire causal structure) are defined with respect to the WTDiff-invariant geodesic equation (1.41).

We introduce the Cauchy surface $\mathcal{C}$ orthogonal to ξ^μ for the exterior of the horizon as in figure 1.1. Let us evaluate the Hamiltonian perturbation (1.95) for this Cauchy surface. Since ξ^μ is a Killing vector, the symplectic form $\Omega\left[\pounds_\xi, \delta\right]$ vanishes and so does δH_ξ. Splitting the boundary of $\mathcal{C}$ into its intersection with the asymptotic infinity $\mathcal{C}_\infty$, and with the horizon $\mathcal{C}_\mathcal{H}$, we have

$$\int_{\partial\mathcal{C}_\infty} \left(\delta Q_\xi^{\nu\mu} - 2\xi^\nu\theta^\mu\left[\delta\right]\right) \mathrm{d}\mathcal{C}_{\mu\nu} - \int_{\partial\mathcal{C}_\mathrm{H}} \left(\delta Q_\xi^{\nu\mu} - 2\xi^\nu\theta^\mu\left[\delta\right]\right) \mathrm{d}\mathcal{C}_{\mu\nu} = 0. \tag{1.97}$$

The first term yields perturbations of quantities measured at the asymptotic infinity. Namely, the contribution corresponding to time translations defines the total canonical energy of the spacetime

$$\delta M = \int_{\partial\mathcal{C}_\infty} \left(\delta Q_t^{\mu\nu} - 2t^\nu\theta^\mu\left[\delta\right]\right) \mathrm{d}\mathcal{C}_{\mu\nu}, \tag{1.98}$$

whereas the rotational contributions quantify the total angular momentum for the given axis of rotation

$$\delta J_{(i)} = -\int_{\partial\mathcal{C}_\infty} \delta Q^{\mu\nu}_{\varphi_{(i)}} \mathrm{d}\mathcal{C}_{\mu\nu}, \tag{1.99}$$

where the overall minus sign ensures positive $J_{(i)}$. As $\varphi^\mu_{(i)}$ are all orthogonal to $\mathrm{d}\mathcal{C}_{\mu\nu}$, we have no contribution proportional to $\varphi^\nu_{(i)}\theta^\mu$. The expressions for both mass and angular momenta are Weyl invariant. In the unimodular gauge and for perturbations that do not change the metric determinant, $\delta\mathfrak{g} = 0$, δM and $\delta J_{(i)}$ reduce to the expressions for variations of the ADM energy and angular momenta valid in general relativity.

Next, we analyse the integral over the Killing horizon. By the same arguments as in Diff-invariant gravity Jacobson et al. [1994], any terms proportional to ξ^μ vanish on the horizon (barring those singular on the bifurcation surface), and a covariant derivative $\tilde{g}_{\lambda\mu}\tilde{\nabla}_\nu\xi^\lambda$ contributes as $\kappa\epsilon_{\nu\mu}$, with $\epsilon_{\nu\mu}$ being the bi-normal to the horizon, and κ the (WTDiff-invariant) surface gravity

$$\kappa = \sqrt{g_{\mu\nu}g^{\rho\sigma}\tilde{\nabla}_\rho\xi^\mu\tilde{\nabla}_\sigma\xi^\nu}\bigg|_\mathcal{H}. \tag{1.100}$$

In total, equation (1.97) then yields

$$\delta M - \sum_{i=1}^{(n-1)/2} \Omega_{\mathcal{H},(i)}\delta J_{(i)} - \frac{1}{8\pi}\kappa\int_{\partial\mathcal{C}_\mathcal{H}} \delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\epsilon^{\nu\mu}\right] \mathrm{d}\mathcal{C}_{\mu\nu} = 0. \tag{1.101}$$

This is the first law of black hole mechanics in Weyl transverse gravity. In the unimodular gauge and for perturbations which leave the metric determinant invariant, it reduces to the familiar first law of general relativity Bekenstein [1973], Bardeen et al. [1973], Wald [1993]

$$\delta M - \sum_{i=1}^{(n-1)/2} \Omega_{\mathcal{H},(i)}\delta J_{(i)} - \frac{1}{8\pi}\kappa\delta\mathcal{A} = 0, \tag{1.102}$$

[10]While the auxiliary metric does not change along Killing vectors, this statement depends on the specific form of the metric. Thence, the Noether currents and charges derived from the general action do not vanish identically for Killing vectors, although the ones derived from a reduced action respecting the Killing symmetries would indeed be zero.

where $\mathcal{A}$ stands for the area of the horizon's cross-section $\partial\mathcal{C}_{\mathcal{H}}$ computed with respect to the dynamical metric $g_{\mu\nu}$. Hence, the physical content of the first law of black hole mechanics in Weyl transverse gravity and in general relativity is the same, as long as we set $\Lambda = \delta\Lambda = 0$.

So far, we have strictly considered classical gravitational physics, and, accordingly, found the first law of black hole mechanics which looks similar to that of thermodynamics, but lacks a $T\delta S$ term. To identify the black hole temperature, we need to invoke the Hawking effect which requires insights from quantum field theory in a curved background Hawking [1975], Visser [2003]. Since Hawking radiation is a kinematic effect and results from fluctuations of matter fields which are unaffected by Weyl transformations (and there are no quantum anomalies associated with local Weyl symmetry Álvarez and Herrero-Valea [2013b], Carballo-Rubio [2015]), the Hawking radiation calculation works the same way for WTDiff-invariant and Diff-invariant setting. The only difference is that the Hawking temperature $T_{\mathrm{H}} = \kappa/2\pi$ contains WTDiff-invariant surface gravity (1.100). If we heuristically take into account the Hawking effect in our otherwise fully classical calculation, we can identify the last term on the left hand side of the first law (1.101) as $T_{\mathrm{H}}\delta S$. Then, equation (1.101) becomes a genuine first law of black hole thermodynamics,

$$\delta M - \sum_{i=1}^{n-3} \Omega_{\mathcal{H},(i)}\delta J_{(i)} - T_{\mathrm{H}}\delta S = 0. \tag{1.103}$$

Wald entropy of the black hole Killing horizon appearing in equation (1.103) reads

$$S = \frac{1}{4}\int_{\partial\mathcal{C}_{\mathcal{H}}} \left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n} \epsilon^{\nu\mu}\mathrm{d}\mathcal{C}_{\mu\nu} = \frac{1}{4}\int_{\partial\mathcal{C}_{\mathcal{H}}} \left(\sqrt{-\mathfrak{g}}/\omega\right)^{(2-n)/n} \sqrt{\mathfrak{h}}\mathrm{d}^{n-2}x, \tag{1.104}$$

where $\mathfrak{h}$ denotes the determinant of the $(n-2)$-dimensional reduced metric on $\partial\mathcal{C}_{\mathcal{H}}$ and $\mathrm{d}^{n-2}x$ the corresponding coordinate area element. Wald entropy of Weyl transverse gravity is by construction WTDiff-invariant and, in the unimodular gauge, it coincides with Bekenstein entropy of general relativity, $S_{\mathrm{B}} = \mathcal{A}/4$.

1.3.3 First law with non-zero cosmological constant

We have pointed out that the only known physical difference between Weyl transverse gravity and general relativity lies in the nature of the cosmological constant, which appears as a global degree of freedom in Weyl transverse gravity. Moreover, this difference also manifests in the Hamiltonian (1.95). Hence, it is of interest to apply our framework to the cases with $\Lambda \neq 0$ and/or $\delta\Lambda \neq 0$. In particular, we choose two important yet tractable examples, a Schwarzschild-anti-de Sitter black hole and de Sitter spacetime in four spacetime dimensions. Results for more general cases can be straightforwardly inferred from our discussion here and from the calculations performed in the context of general relativity.

Schwarzschild-anti-de Sitter spacetime

We begin by deriving the first law of black hole mechanics in Schwarzschild-anti-de Sitter spacetime with $\Lambda < 0$. The spacetime is static and thus possesses a time translational Killing vector field, t^μ, which is timelike everywhere in the black hole exterior and becomes null on the horizon. In this case, equation (1.79) for the on-shell Noether current j_t^μ integrated over the Cauchy surface $\mathcal{C}$ orthogonal to t^μ yields

$$\int_{\mathcal{C}} j_t^\mu \mathrm{d}\mathcal{C}_\mu = -\frac{1}{8\pi}\Lambda \int_{\mathcal{C}} t^\mu \mathrm{d}\mathcal{C}_\mu + \int_{\partial\mathcal{C}_\infty} Q_t^{\nu\mu} \mathrm{d}\mathcal{C}_{\mu\nu} - \int_{\partial\mathcal{C}_\mathcal{H}} Q_t^{\nu\mu} \mathrm{d}\mathcal{C}_{\mu\nu}. \tag{1.105}$$

We first evaluate the Noether current from its general definition (1.57). Since $\pounds_t \tilde{g}_{\mu\nu} = 0$, the symplectic potential term vanishes (see equation (1.74)). For the special case $\Lambda = 0$, we have $\tilde{R} = 0$ and the Lagrangian vanishes. Then, the integral of $Q_t^{\nu\mu}$ over $\mathcal{C}_\infty$ corresponds to $M/2$, where M is in this case the Komar mass Iyer and Wald [1994] (the Komar, ADM and canonical definitions of mass of course coincide in this case). The integral over $\mathcal{C}_\mathcal{H}$ then gives $T_\mathrm{H} S$ in the thermodynamic interpretation and we have the well-known Smarr formula

$$\frac{M}{2} - T_\mathrm{H} S = 0, \tag{1.106}$$

for a Schwarzschild black hole. However, for $\Lambda < 0$, both sides of equation (1.105) diverge. Since these infinities are of the same nature as in the pure anti-de Sitter spacetime, we can choose it as our reference background, and impose that the Noether current and charge vanish there. Stated differently, we define the physical Noether charge and current as the difference of their value in our Schwarzschild-anti de Sitter spacetime and the pure anti-de Sitter spacetime characterised by the same value of Λ[11], obtaining

$$Q_{t,\mathrm{phys}}^{\nu\mu} = Q_{t,\text{S-AdS}}^{\nu\mu} - Q_{t,\mathrm{AdS}}^{\nu\mu}, \tag{1.107}$$
$$j_{t,\mathrm{phys}}^{\mu} = j_{t,\text{S-AdS}}^{\mu} - j_{t,\mathrm{AdS}}^{\mu} = \tilde{\nabla}_\nu Q_{t,\mathrm{phys}}^{\nu\mu}. \tag{1.108}$$

Using these physical variables, the Smarr formula (1.105) becomes finite and reads

$$\begin{aligned}\int_{\mathcal{C}} j_{t,\mathrm{phys}}^\mu \mathrm{d}\mathcal{C}_\mu &= \int_{\partial\mathcal{C}_\infty} Q_{t,\mathrm{phys}}^{\nu\mu} \mathrm{d}\mathcal{C}_{\mu\nu} - \int_{\partial\mathcal{C}_\mathcal{H}} Q_{t,\mathrm{phys}}^{\nu\mu} \mathrm{d}\mathcal{C}_{\mu\nu}, \\ 0 &= \frac{1}{2}M - T_\mathrm{H} S + \frac{1}{3}\Lambda r_\mathcal{H}^3,\end{aligned} \tag{1.109}$$

where $r_\mathcal{H}$ is the horizon radius. This is simply the WTDiff-invariant version of the Smarr formula valid for a Schwarzschild-anti-de Sitter black hole in general relativity Kastor et al. [2009, 2010].

Upon, deriving the Smarr formula we turn to the first law. Thus, we consider a perturbation δH_t of the Hamiltonian generating evolution along t^μ between two Schwarzschild-anti-de Sitter solutions of the equations of motion of Weyl transverse gravity that are related by a small perturbation. As we discussed previously, the cosmological constant generically varies between different solutions in Weyl transverse gravity. Hence, to subtract the divergent terms in δH_t, we must first define the physical Hamiltonians in both Schwarzschild-anti-de Sitter spacetimes by subtracting the corresponding anti-de Sitter backgrounds (one with

[11]Since Λ is only defined on shell, we cannot perform the subtraction at the level of the action as it is usually done in general relativity Hawking and Horowitz [1996].

Λ and the other with $\Lambda + \delta\Lambda$). Only then can we compute δH_t as a difference of these physical Hamiltonians. The result is

$$\delta H_{t,\text{phys}} = \delta H_{t,\text{S-AdS}} - \delta H_{t,\text{AdS}} = \int_{\partial\mathcal{C}} \left(\delta Q^{\nu\mu}_{t,\text{phys}} - 2t^\nu \theta^\mu_{\text{phys}}\left[\delta\right]\right) \mathrm{d}\mathcal{C}_{\mu\nu}, \qquad (1.110)$$

where

$$\theta^\mu_{\text{phys}} = \theta^\mu_{\text{S-AdS}} - \theta^\mu_{\text{AdS}}. \qquad (1.111)$$

Evaluating the integrals, we obtain the first law of thermodynamics valid for a Schwarzschild-anti-de Sitter black hole in Weyl transverse gravity

$$M - T_\text{H}\delta S + \frac{4\pi}{3} r^3_{\mathcal{H}} \frac{\delta\Lambda}{8\pi} = 0. \qquad (1.112)$$

A varying negative cosmological constant has been extensively studied in the literature, in the context of the so-called black hole chemistry Kastor et al. [2009], Kubizňák and Mann [2014]. There, it has been argued that $-\Lambda/8\pi$ effectively acts as a pressure in the black hole thermodynamic system. This observation led to a definition of the thermodynamic volume of a Schwarzschild-anti-de Sitter black hole as $4\pi r^3_{\mathcal{H}}/3$ (this quantity has no relation to the black hole's geometric volume, which is infinite). As the first law (1.112) contains a term proportional to a variation of pressure, the canonical mass M of the spacetime plays the role of enthalpy. This picture allows one to apply many insights from standard thermodynamics to asymptotically anti-de Sitter black holes, including, e.g. the concept of phase transitions Hawking and Page [1996], Kubizňák and Mann [2014], Kubiznak and Mann [2015], Kubizňák et al. [2017]. While perturbations of the cosmological constant $\delta\Lambda$ must be added somewhat ad hoc in general relativity, they appear naturally in Weyl transverse gravity already in the fully classical setting.

De Sitter spacetime

Asymptotically de Sitter spacetimes generically contain both an event and a cosmological horizon (both Killing horizons). The presence of two accessible horizons makes their thermodynamics somewhat complicated Aneesh et al. [2019]. Herein, we thus limit ourselves to a cosmological horizon in a pure de Sitter spacetime. It is a Killing horizon with respect to the time translational Killing vector t^μ which is timelike inside the horizon and spacelike outside of it.

We study a small variation of the metric in de Sitter spacetime that satisfies the vacuum equations of motion of Weyl transverse gravity. The corresponding variation of the Hamiltonian generating the evolution along t^μ vanishes, implying

$$\delta H_t = \sqrt{\frac{\Lambda}{3}} \int_{\partial\mathcal{C}} \delta \left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{1/2} \epsilon^{\nu\mu}\right] \mathrm{d}\mathcal{C}_{\mu\nu} + \frac{1}{8\pi}\delta\Lambda \int_{\mathcal{C}} t^\mu \mathrm{d}\mathcal{C}_\mu = 0. \qquad (1.113)$$

where we again generically have to consider a variation of the cosmological constant, in contrast to the situation in general relativity. We can explicitly compute the integrals, obtaining the first law of the de Sitter cosmological horizon

$$\sqrt{\frac{\Lambda}{3}} \left(\delta\mathcal{A}_{\partial\mathcal{C}} - \frac{1}{2}\mathcal{A}_{\partial\mathcal{C}}\delta\frac{\sqrt{-\mathfrak{g}}}{\omega}\right) + \frac{1}{8\pi}\mathcal{V}_{\mathcal{C}}\delta\Lambda = 0, \qquad (1.114)$$

where $\mathcal{A}_{\partial\mathcal{C}} = 12\pi/\Lambda$ denotes the area of $\partial\mathcal{C}$ and $\mathcal{V}_{\mathcal{C}} = 4\sqrt{3}\pi/\Lambda^{3/2}$ the volume of $\mathcal{C}$. If we consider quantum field theory on curved backgrounds to define the Hawking temperature of the de Sitter horizon, $T_{\rm dS} = (1/2\pi)\sqrt{\Lambda/3}$, we can identify the entropy of de Sitter horizon. It reads

$$S = 3\pi/\Lambda, \tag{1.115}$$

in agreement with the result in general relativity.

We may also derive the Smarr formula relating the volume and the area of $\mathcal{C}$. Integrating the on-shell relation for j_t^μ (1.79) over $\mathcal{C}$, we find

$$\int_{\mathcal{C}} j_t^\mu \mathrm{d}\mathcal{C}_\mu = -\frac{1}{8\pi}\Lambda \int_{\mathcal{C}} t^\mu \mathrm{d}\mathcal{C}_\mu + \int_{\partial\mathcal{C}} Q_t^{\nu\mu} \mathrm{d}\mathcal{C}_{\mu\nu}. \tag{1.116}$$

Since $j_t^\mu = -Lt^\mu = -\Lambda t^\mu/4\pi$ (the symplectic potential contribution to the Noether current vanishes for any Killing vector), the final result reads

$$\frac{\Lambda\mathcal{V}_{\mathcal{C}}}{8\pi} = \frac{1}{2\pi}\sqrt{\frac{\Lambda}{3}}\frac{\mathcal{A}_{\partial\mathcal{C}}}{4}, \tag{1.117}$$

The validity of this Smarr formula can be easily checked by plugging in the expressions for $\mathcal{V}_{\mathcal{C}}$ and $\mathcal{A}_{\partial\mathcal{C}}$.

It has been argued that Euclidean path integral calculations of de Sitter entropy in Weyl transverse gravity (or any model of unimodular gravity) yield a different result than in general relativity Fiol and Garriga [2010]. This would clearly lead to a contradiction with our result for the de Sitter entropy (1.115). Thence, we briefly discuss the Euclidean path integral result and address this apparent contradiction. The approach is based on approximating the partition function Z of the de Sitter spacetime canonical ensemble by the classical action Gibbons and Hawking [1977]. A standard thermodynamics argument then implies (assuming no pressure and chemical potentials are present)

$$\ln Z = -E/T + S, \tag{1.118}$$

where T denotes the de Sitter temperature, E the total energy, and S the entropy. The Euclidean action of general relativity in the region inside the de Sitter horizon (a 4-sphere of radius $\sqrt{3/\Lambda}$) equals

$$I = -\frac{3\pi}{\Lambda}. \tag{1.119}$$

Since $E = 0$, we then straightforwardly have that entropy obeys equation (1.115). In Weyl transverse gravity, the same procedure yields (keeping the parameter λ in the action nonzero)

$$I = -\frac{3\pi}{\Lambda}\left(2 - \frac{\lambda}{\Lambda}\right). \tag{1.120}$$

If we assume $E = 0$ as before, we obtain entropy which differs from the general relativistic result and from Wald entropy we obtained (1.115), unless we choose $\lambda = \Lambda$ Fiol and Garriga [2010]. While this resolution was previously suggested, we have argued that it cannot work, since λ is just a single parameter, whereas Λ is only defined on shell and takes different values for different solutions.

To reconcile the apparent discrepancy in entropy, we instead propose the following. As we pointed out above, it has been argued in the context of general relativity that varying negative cosmological constant can be identified with pressure, $p_\Lambda = -\Lambda/8\pi$ Kubiznak and Mann [2015]. Adopting the same interpretation for the varying positive cosmological in de Sitter (we note that variations of cosmological constant appears in Weyl transverse gravity without any further assumptions), we have

$$I = -S + p_\Lambda \mathcal{V}_{\text{dS}} = -S - \frac{3\pi}{\Lambda}, \tag{1.121}$$

where $\mathcal{V}_{\text{dS}}$ denotes the volume of the Euclidean 4-sphere enclosed by the de Sitter horizon. Entropy thus obeys

$$S = \frac{3\pi}{\Lambda}\left(1 - \frac{\lambda}{\Lambda}\right). \tag{1.122}$$

By setting $\lambda = 0$ one then recovers the expected entropy (1.115). The λ-proportional term has a simple interpretation in Weyl transverse gravity. Its contribution to the action (in any spacetime) equals $\mathcal{V}^{(\omega)}\lambda$, where $\mathcal{V}^{(\omega)}$ denotes the spacetime volume of the integration domain $\mathcal{V}$ evaluated with respect to the non-dynamical volume measure ω. Then, $\mathcal{V}^{(\omega)}\lambda$ is a universal constant which quantifies our freedom to shift the value of entropy by a constant. In conclusion, we can see the equivalence of the Euclidean canonical ensemble and the covariant phase space formalism calculations of entropy (since the only possible obstruction to equivalence is the behaviour of Λ, it clearly holds in general).

1.3.4 First law of black hole mechanics in the presence of matter

So far, we have considered vacuum Weyl transverse gravity. However, the first law of black hole mechanics in general relativity was originally derived for a stationary, asymptotically flat black hole spacetime with a perfect fluid Bardeen et al. [1973]. The covariant phase space formalism then allowed to generalise this derivation to arbitrary local, Diff-invariant theory of gravity Iyer [1997]. For the sake of comparison, we would like to obtain the same result in Weyl transverse gravity. In order to do so, we first discuss the first law in the presence of general matter content. Then, we introduce a suitable WTDiff-invariant Lagrangian description of a perfect fluid. Lastly, we specialise the first law of black hole mechanics in the presence of matter for this model of the perfect fluid. For simplicity, we set $\Lambda = \delta\Lambda = 0$ throughout, but the generalisation to nontrivial Λ is fairly straightforward.

First law in the presence of matter

To obtain the symplectic potential θ^μ_ψ and symplectic current Ω^μ_ψ for matter fields, we can simply apply the general equations (1.42) and (1.45), respectively, to the WTDiff-invariant matter Lagrangian (1.21). The derivation of the expressions for the matter Noether currents is more subtle. Since Weyl transformations do not act on matter fields, the corresponding Noether current trivially vanishes. Regarding transverse diffeomorphisms generated by some vector field ξ^μ, we can straightforwardly apply the general equation (1.53) with $\alpha^\mu_{\psi,\xi} = \left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} L_\psi \xi^\mu$. The result is

$$j^\mu_{\psi,\xi} = \theta^\mu_\psi \left[\pounds_\xi\right] - L_\psi \xi^\mu. \tag{1.123}$$

To learn something about the matter Noether current in general, rather than for specific matter fields, we evaluate its WTDiff-invariant divergence, obtaining

$$\begin{aligned}\tilde{\nabla}_\mu j^\mu_{\psi,\xi} =& \tilde{\nabla}_\mu \theta^\mu_{\psi,\xi} - \pounds_\xi \left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} L_\psi\right] \\ =& - A_\psi \pounds_\xi \psi - \left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} \left(T^{\mu\nu} - \frac{1}{n} T g^{\mu\nu}\right) 2\nabla_{(\mu}\xi_{\nu)},\end{aligned} \tag{1.124}$$

where we used the definition of the symplectic potential (1.42) and denoted the matter equations of motion by $A_\psi = 0$. By a series of straightforward manipulations, we get Iyer [1997]

$$\tilde{\nabla}_\mu j^\mu_{\psi,\xi} = \tilde{\nabla}_\mu \left[-\left(\psi \cdot A_\psi \cdot \xi\right)^\mu - \left(\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} T_\nu^{\ \mu} - \mathcal{J}\delta^\mu_\nu\right)\xi^\nu\right], \tag{1.125}$$

where $\mathcal{J}$ corresponds to the potential local energy non-conservation defined by equation (1.25). Since we can express $\tilde{\nabla}_\mu j^\mu_{\psi,\xi}$ by equation (1.125), we must have for the Noether current Wald [1990], Iyer [1997]

$$j^\mu_{\psi,\xi} = -\left(\psi \cdot A_\psi \cdot \xi\right)^\mu - \left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} T_\nu^{\ \mu} - \mathcal{J}\delta^\mu_\nu\right]\xi^\nu + \tilde{\nabla}_\nu Q^{\nu\mu}_{\psi,\xi}, \tag{1.126}$$

where $Q^{\nu\mu}_{\psi,\xi}$ denotes the matter Noether charge. For any given matter Lagrangian, we can easily find the Noether charge by comparing expressions for $j^\mu_{\psi,\xi}$, (1.123) and (1.126). The second term in equation (1.126) is just twice the right hand side of the divergence-free equations (1.26) of Weyl transverse gravity. Therefore, the sum of the matter (1.126) and the gravitational (1.77) Noether currents on shell becomes the divergence of the Noether charge and a cosmological constant contribution, just like in vacuum Weyl transverse gravity (see equation (1.79)).

We stress that the matter Noether charge is proportional to the transverse diffeomorphism generator ξ^μ. To see this, we note that a Lagrangian for minimally coupled matter fields can contain at most first derivatives of the matter variables. Then, the symplectic potential can only contain variations of the matter variables and not their derivatives. If these variations correspond to transverse diffeomorphisms, they are simply given by Lie derivatives along ξ^μ, which depend at most on first derivatives of ξ^μ. The Noether current $j^\mu_{\psi,\xi}$ and, hence, also the divergence of the Noether charge $Q^{\nu\mu}_{\psi,\xi}$ can also contain at most first derivatives of ξ^μ. This implies that $Q^{\nu\mu}_{\psi,\xi}$ does not contain any derivatives of ξ^μ and there exists a WTDiff-invariant antisymmetric tensor $W_\rho^{\ \nu\mu} = W_\rho^{\ [\nu\mu]}$, such that $Q^{\nu\mu}_{\psi,\xi} = \xi^\rho W_\rho^{\ \nu\mu}$. Since we have shown that terms of the form $W_\rho^{\ \nu\mu} = W_\rho^{\ [\nu\mu]}$ do not affect Wald entropy, minimally coupled matter fields do not contribute to it.

To study the Hamiltonian, we need the matter symplectic current $\Omega^\mu_\psi \left[\pounds_\xi, \delta\right]$ corresponding to a transverse diffeomorphism and an arbitrary small perturbation of the metric and the matter fields. Starting from the general definition (1.45), and following the same strategy as for the gravitational symplectic current, we eventually obtain

$$\begin{aligned}\Omega^{\mu}_{\psi}\left[\pounds_{\xi},\delta\right]=&\tilde{\nabla}_{\nu}\left(\delta Q^{\nu\mu}_{\psi,\xi}-2\xi^{[\nu}\theta^{\mu]}_{\psi}\left[\delta\right]\right)-\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}T_{\nu}{}^{\mu}-\mathcal{J}\delta^{\mu}_{\nu}\right]\xi^{\nu}\\&+\frac{1}{2}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}\xi^{\mu}\left(T^{\alpha\beta}-\frac{1}{n}Tg^{\alpha\beta}\right)\delta g_{\alpha\beta},\end{aligned}\tag{1.127}$$

where we assumed that both the original spacetime and the perturbed one satisfy the equations of motion. An integral of $\Omega^{\mu}_{\psi}\left[\pounds_{\xi},\delta\right]$ over a suitable Cauchy surface $\mathcal{C}$ then yields the matter symplectic form $\Omega_{\psi}\left[\pounds_{\xi},\delta\right]$ (see equation (1.46)). If the matter Hamiltonian exists (this is guaranteed by condition (1.50)), the Hamilton equations of motion imply $\delta H_{\psi,\xi}=\Omega_{\psi}\left[\pounds_{\xi},\delta\right]$. Thence, we have

$$\begin{aligned}\delta H_{\psi,\xi}=&\int_{\partial\mathcal{C}}\left(\delta Q^{\nu\mu}_{\psi,\xi}-2\xi^{\nu}\theta^{\mu}_{\psi}\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu}-\int_{\mathcal{C}}\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}T_{\nu}{}^{\mu}-\mathcal{J}\delta^{\mu}_{\nu}\right]\xi^{\nu}\mathrm{d}\mathcal{C}_{\mu}\\&+\frac{1}{2}\int_{\mathcal{C}}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}\left(T^{\alpha\beta}-\frac{1}{n}Tg^{\alpha\beta}\right)\delta g_{\alpha\beta}\xi^{\mu}\mathrm{d}\mathcal{C}_{\mu}.\end{aligned}\tag{1.128}$$

The complete Hamiltonian perturbation consists of a sum of the matter $\delta H_{\psi,\xi}$ and the gravitational $\delta H_{\mathrm{g},\xi}$ parts (we introduce subscript g to set $\delta H_{\mathrm{g},\xi}$ apart from the perturbation of the total Hamiltonian δH_{ξ}), so that

$$\delta H_{\xi}=\delta H_{\mathrm{g},\xi}+\delta H_{\psi,\xi}.\tag{1.129}$$

For the perturbation of the total on shell Hamiltonian, we obtain Alonso-Serrano et al. [2023a] (following the same steps as in the vacuum case)

$$\begin{aligned}\delta H_{\xi}=&\int_{\mathcal{C}}\Omega^{\mu}\left[\pounds_{\xi},\delta\right]\mathrm{d}\mathcal{C}_{\mu}=\int_{\partial\mathcal{C}}\left(\delta Q^{\nu\mu}_{\mathrm{g},\xi}+\delta Q^{\nu\mu}_{\psi,\xi}-2\xi^{\nu}\theta^{\mu}_{\mathrm{g}}\left[\delta\right]-2\xi^{\nu}\theta^{\mu}_{\psi}\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu}\\&-\int_{\mathcal{C}}\frac{1}{8\pi}\delta\Lambda\xi^{\mu}\mathrm{d}\mathcal{C}_{\mu},\end{aligned}\tag{1.130}$$

where the total symplectic current is again simply a sum of the matter and the gravitational contributions, i.e., $\Omega^{\mu}=\Omega^{\mu}_{\mathrm{g}}+\Omega^{\mu}_{\psi}$. As before, the total Hamiltonian can be expressed in terms of a surface integral, except for the volume contribution proportional to the cosmological constant. Combining equations (1.129) and (1.130) for δH_{ξ} together with the expression (1.128) for $\delta H_{\psi,\xi}$, we find for the perturbation of the gravitational Hamiltonian

$$\begin{aligned}\delta H_{\mathrm{g},\xi}=&\int_{\partial\mathcal{C}}\left(\delta Q^{\nu\mu}_{\mathrm{g},\xi}-2\xi^{\nu}\theta^{\mu}_{\mathrm{g}}\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu}-\int_{\mathcal{C}}\frac{1}{8\pi}\delta\Lambda\xi^{\mu}\mathrm{d}\mathcal{C}_{\mu}\\&-\int_{\mathcal{C}}\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}T_{\nu}{}^{\mu}-\mathcal{J}\delta^{\mu}_{\nu}\right]\xi^{\nu}\mathrm{d}\mathcal{C}_{\mu}\\&+\frac{1}{2}\int_{\mathcal{C}}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}\left(T^{\alpha\beta}-\frac{1}{n}Tg^{\alpha\beta}\right)\delta g_{\alpha\beta}\xi^{\mu}\mathrm{d}\mathcal{C}_{\mu}.\end{aligned}\tag{1.131}$$

In contrast with the vacuum case, the perturbation of the gravitational Hamiltonian now contains volume integrals given by the matter content of the theory. Equation (1.131) represents the starting point for a straightforward derivation of the first law of black hole mechanics, as we show in the following.

Consider an n-dimensional, stationary, asymptotically flat spacetime with arbitrary minimally coupled matter fields present. We have, just like in the vacuum case, a time translational Killing vector field , t^{μ}, and $(n-1)/2$ rotational

Killing vectors, $\varphi^\mu_{(i)}$. Their combination $\xi^\mu = t^\mu + \sum_{i=1}^{n-3} \Omega_{\mathcal{H},(i)}\varphi^\mu_{(i)}$, with $\Omega_{\mathcal{H},(i)}$ being the constant angular velocities of the horizon, is again a Killing vector. The black hole event horizon is then a Killing horizon with respect to ξ^μ. Our task is now to evaluate equation (1.131) for the Cauchy surface $\mathcal{C}$ orthogonal to ξ^μ. Its boundary consists of the intersection with the spatial infinity, $\mathcal{C}_\infty$, and with the horizon, $\partial\mathcal{C}_\mathcal{H}$.

Since ξ^μ is a Killing vector field, $\delta H_{\mathrm{g},\xi}$ vanishes. If $\delta H_{\psi,\xi}$ was zero as well, we would have $\delta H_\xi = 0$ and equation (1.130) would then give us the first law of black hole mechanics in terms of boundary integrals (since we set $\delta\Lambda = 0$). However, the matter fields do not in general possess the same symmetries as the spacetime (although the energy-momentum tensor does) and their Lie derivatives with respect to a Killing vector do not necessarily vanish Iyer [1997]. This occurs, e.g., in the physically relevant case of perfect fluids. Their Lagrangian must include Lagrange multipliers which in general do not share the spacetime symmetries, although fluid's entropy, temperature, velocity, particle density, energy density and pressure do. To sum up, $\delta H_{\psi,\xi}$ does not generically equal zero and we have to use the more complicated equation (1.131) to derive the first law of black hole mechanics.

Evaluating the perturbation of the gravitational Hamiltonian (1.131) in our stationary, asymptotically flat setup yields

$$\int_{\partial\mathcal{C}} \left(\delta Q^{\nu\mu}_{\mathrm{g},\xi} - 2\xi^\nu\theta^\mu_{\mathrm{g}}\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu} - \int_{\mathcal{C}} \delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} T_\nu{}^\mu - \mathcal{J}\delta^\mu_\nu\right]\xi^\nu\mathrm{d}\mathcal{C}_\mu$$
$$+\frac{1}{2}\int_{\mathcal{C}}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}\left(T^{\alpha\beta} - \frac{1}{n}Tg^{\alpha\beta}\right)\delta g_{\alpha\beta}\xi^\mu\mathrm{d}\mathcal{C}_\mu = 0. \tag{1.132}$$

The first term comes from the gravitational degrees of freedom and has the same interpretation as in vacuum

$$\int_{\partial\mathcal{C}} \left(\delta Q^{\nu\mu}_{\mathrm{g},\xi} - 2\xi^\nu\theta^\mu_{\mathrm{g}}\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu} = \delta E - \sum_{i=1}^{(n-1)/2}\Omega^{(i)}_\mathcal{H}\delta J_{(i)}$$
$$-\frac{1}{8\pi}\kappa\int_{\partial\mathcal{C}_\mathcal{H}}\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\epsilon^{\nu\mu}\right]\mathrm{d}\mathcal{C}_{\mu\nu}, \tag{1.133}$$

where perturbations of the canonical energy, δE, and the angular momenta, $\delta J_{(i)}$, include contributions from both the black hole and the matter fields. Let us now look at the matter field contributions. There, we further assume that a vector field co-moving with the matter takes the form $U^\mu = t^\mu + \sum_{i=1}^{(n-1)/2}\Omega_{(i)}\varphi^\mu_{(i)}$, where $\Omega_{(i)}$ denotes the matter angular velocities (not necessarily constant) in the various rotational directions. Then, we can conveniently rewrite the first volume integral in equation (1.132) as

$$-\int_{\mathcal{C}}\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}T_\nu{}^\mu - \mathcal{J}\delta^\mu_\nu\right]\xi^\nu\mathrm{d}\mathcal{C}_\mu = -\int_{\mathcal{C}}\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}T_\nu{}^\mu\right]U^\nu\mathrm{d}\mathcal{C}_\mu$$
$$-\int_{\mathcal{C}}\delta\mathcal{J}\xi^\mu\mathrm{d}\mathcal{C}_\mu + \int_{\mathcal{C}}\sum_{i=1}^{(n-1)/2}\left(\Omega^{(i)} - \Omega^{(i)}_\mathcal{H}\right)\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}T_\nu{}^\mu\right]\varphi^\nu_{(i)}\mathrm{d}\mathcal{C}_\mu, \tag{1.134}$$

where the first term is proportional to the perturbation of the energy-momentum tensor contracted with the vector field U^ν co-moving with the matter. The second term quantifies the possible local energy non-conservation (keep all the caveats of such a possibility we discussed in subsection 1.1.5 in mind). The last contribution yields perturbations of the WTDiff-invariant angular momenta of the matter $\delta\tilde{J}^{\mu}_{(i)}$, defined by the standard prescription

$$\delta\tilde{J}^{\mu}_{(i)} = \delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} T_{\nu}{}^{\mu}\right]\varphi^{\nu}_{(i)}. \tag{1.135}$$

The terms $\Omega_{\mathcal{H},(i)}\delta\tilde{J}^{\mu}_{(i)}$ can be combined with the total angular momenta perturbations present in equation (1.133), leaving only the black hole contribution $J_{\mathcal{H},(i)}$. In total, the first law for a stationary, asymptotically flat black hole spacetime in the present of matter fields read

$$\begin{aligned}
&\delta E - \sum_{i=1}^{(n-1)/2}\Omega^{(i)}_{\mathcal{H}}\delta J^{(i)}_{\mathcal{H}} - \frac{1}{8\pi}\kappa\int_{\partial\mathcal{C}_{\mathcal{H}}}\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\epsilon^{\nu\mu}\right]\mathrm{d}\mathcal{C}_{\mu\nu} \\
&- \int_{\mathcal{C}}\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} T_{\nu}{}^{\mu}\right]U^{\nu}\mathrm{d}\mathcal{C}_{\mu} \\
&+ \frac{1}{2}\int_{\mathcal{C}}\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n}\left(T^{\alpha\beta} - \frac{1}{n}Tg^{\alpha\beta}\right)\delta g_{\alpha\beta}\xi^{\mu}\mathrm{d}\mathcal{C}_{\mu} \\
&+ \int_{\mathcal{C}}\sum_{i=1}^{(n-1)/2}\Omega^{(i)}\delta\tilde{J}^{\mu}_{(i)}\mathrm{d}\mathcal{C}_{\mu} + \int_{\mathcal{C}}\delta\mathcal{J}\xi^{\mu}\mathrm{d}\mathcal{C}_{\mu} = 0.
\end{aligned} \tag{1.136}$$

Further analysis of the first law requires specifying the matter content. We choose as an example a black hole surrounded by a perfect fluid, a setup treated already in the seminal paper on black hole mechanics Bardeen et al. [1973]. However, we first need to introduce a Lagrangian description for a WTDiff-invariant perfect fluid.

WTDiff-invariant perfect fluids

To apply the covariant phase space formalism to a perfect fluid, we require a Lagrangian description for it[12]. We choose to introduce a WTDiff-invariant (and somewhat simplified) variant of the formalism introduced in Brown [1993]. Hence, we take the entropy per particle s, and the particle density ν, as the configurational variables for the fluid. The fluid equation of state then expresses the energy density as a function of these variables, i.e., $\rho = \rho(s,\nu)$. We also introduce the velocity of the fluid u^{μ}, normalised to $u_{\mu}u^{\mu} = -1$. To make the normalisation Weyl-invariant, u^{μ} must change under Weyl transformations as $u'^{\mu} = e^{-\sigma}u^{\mu}$ (stated differently, $g_{\mu\nu} + u_{\mu}u_{\nu}$ remains a projector to the subspace orthogonal to u^{μ} in every Weyl gauge). The particle number density flux then reads

$$I^{\mu} = \left(\sqrt{-\mathfrak{g}}/\omega\right)^{1/n}\nu u^{\mu}, \tag{1.137}$$

[12] The form of the first law (1.136) in the presence of matter we derived, works even without knowing the Lagrangian, since the matter Noether charges do not enter it (and it can be derived by different methods just from the equations of motion and boundary conditions Bardeen et al. [1973]). Then, we may in principle apply it even to non-Lagrangian matter fields which can violate the local energy conservation and have $\delta\mathcal{J} = 0$. However, the Lagrangian description of the fluid is still needed to make the covariant phase space derivation self-consistent, as well as to generalise it to arbitrary local, WTDiff-invariant theories of gravity, which we do in the next section.

where the included power of the metric determinant ensures its overall WTDiff invariance.

The main part of our fluid Lagrangian can be chosen to be simply the energy density, $\rho\left(\nu, s\right)$ Brown [1993]. To have a perfect fluid, we must further ensure that the fluxes of the particle number density and entropy are conserved along the flow lines. We add these conditions using Lagrange multipliers, obtaining the following full Lagrangian[13]

$$L_{\mathrm{f}} = -\rho\left(s, \nu\right) + I^{\mu}\left(\tilde{\nabla}_{\mu}\eta + s\tilde{\nabla}_{\mu}\tau\right). \tag{1.138}$$

Here, the functions η and τ are the Lagrange multipliers. As we will see, their physical interpretation is encoded in the equations of motion. Varying the action with respect to the dynamical metric gives us the traceless part of the energy-momentum tensor

$$T_{\mu\nu} - \frac{1}{n}Tg_{\mu\nu} = \left(\rho + p\right)\left(u_{\mu}u_{\nu} + \frac{1}{n}g_{\mu\nu}\right), \tag{1.139}$$

where we identify the pressure as

$$p = \nu\frac{\partial\rho}{\partial\nu} - \rho, \tag{1.140}$$

using a comparison our result with the standard form of the perfect fluid energy-momentum tensor. By design, variations with respect to the Lagrange multipliers η and τ lead to the conservation laws for the particle number density flux and the entropy flux, respectively,

$$\tilde{\nabla}_{\mu}I^{\mu} = 0, \qquad \tilde{\nabla}_{\mu}\left(sI^{\mu}\right) = 0. \tag{1.141}$$

Varying the action with respect to the entropy per particle s yields

$$-\frac{\partial\rho}{\partial s} + I^{\mu}\tilde{\nabla}_{\mu}\tau = 0. \tag{1.142}$$

If we define the Weyl-invariant fluid temperature as

$$\mathcal{T} = \left(\sqrt{-\mathfrak{g}}/\omega\right)^{1/n} u^{\mu}\tilde{\nabla}_{\mu}\tau, \tag{1.143}$$

equation (1.142) becomes

$$\mathcal{T} = \frac{1}{\nu}\frac{\partial\rho}{\partial s}, \tag{1.144}$$

which corresponds to the first law of thermodynamics for the fluid Brown [1993]. Lastly, varying the action with respect to the particle number density ν gives

$$\frac{\partial\rho}{\partial\nu} - \left(\sqrt{-\mathfrak{g}}/\omega\right)^{1/n} u^{\mu}\left(\tilde{\nabla}_{\mu}\eta + s\tilde{\nabla}_{\mu}\tau\right) = 0. \tag{1.145}$$

[13]To specify the flow lines and fix their form at the boundaries, we should also include their description in terms of Lagrange coordinates Brown [1993]. However, in our case of a stationary, asymptotically flat spacetime, the fluid is not present at the boundaries and no such terms are necessary.

The last term is simply $-\mathcal{T}s$, whereas the first one corresponds to $(\rho+p)/\nu$ (see equation (1.140)). If we invoke the Gibbs-Duhem relation of the standard thermodynamics, we can define the chemical potential μ

$$\mu = \frac{\rho+p}{\nu} - \mathcal{T}s = \frac{\partial \rho}{\partial \nu} - s\left(\sqrt{-\mathfrak{g}}/\omega\right)^{1/n} u^{\mu}\tilde{\nabla}_{\mu}\tau. \tag{1.146}$$

Equation (1.145) then represents a relation between the chemical potential and the Lagrange multiplier η

$$\mu = \left(\sqrt{-\mathfrak{g}}/\omega\right)^{1/n} u^{\mu}\tilde{\nabla}_{\mu}\eta. \tag{1.147}$$

Varying the action with respect to the fluid velocity u^{μ} gives an equation governing the behaviour of the Lagrange multipliers η and τ on the surfaces orthogonal to the flow lines, which is irrelevant for our purposes.

One can easily check that the fluid equations of motion imply that the energy-momentum tensor has a vanishing divergence, i.e., $\tilde{\nabla}_{\nu}T_{\mu}{}^{\nu}=0$, and, therefore, $\mathcal{J}=0$. As expected based on our discussion in subsection 1.1.5, the local energy-momentum conservation is directly built into the Lagrangian description of the fluid.

To illustrate the general covariant phase space formalism for matter fields, let us apply it to our perfect fluid Lagrangian. The expressions for the symplectic potential and the Noether current corresponding to transverse diffeomorphisms can be straightforwardly derived from the general definitions. The results read

$$\theta^{\mu}_{\mathrm{f}} = I^{\mu}\left(\delta\eta + s\delta\tau\right), \qquad j^{\mu}_{\mathrm{f},\xi} = -T_{\nu}{}^{\mu}\xi^{\nu}. \tag{1.148}$$

Thence, the Noether charge $Q^{\nu\mu}_{\mathrm{f},\xi}$ vanishes identically. Finally, the on-shell perfect fluid symplectic current $\Omega^{\mu}_{\mathrm{f}}\left[\pounds_{\xi},\delta\right]$, corresponding to a transverse diffeomorphism and a small perturbation of the metric and the fluid variables, equals

$$\Omega^{\mu}_{\mathrm{f}}\left[\pounds_{\xi},\delta\right] = -2\tilde{\nabla}_{\nu}\left(\xi^{[\nu}\theta^{\mu]}_{\mathrm{f}}\right) - \xi^{\nu}\delta T_{\nu}{}^{\mu} + \frac{1}{2}\xi^{\mu}\left(T^{\alpha\beta} - \frac{1}{n}Tg^{\alpha\beta}\right)\delta g_{\alpha\beta}. \tag{1.149}$$

By integrating this expression over a Cauchy surface, we directly obtain a perturbation of the fluid Hamiltonian $\delta H_{\xi,\mathrm{f}}$ (see equation (1.128)).

First of black hole mechanics with a perfect fluid

We now have a first law of mechanics for stationary, asymptotically flat black hole spacetimes in the presence of arbitrary, minimally coupled matter fields, and a WTDiff-invariant Lagrangian description of a perfect fluid. All that remains is combining these two results. We aim to both illustrate how the matter field contributions to the Hamiltonian perturbation affect the first law and to show the physical equivalence of the final expression with the result in general relativity. Starting from the general equation (1.136), we specialise it to find the first law of black hole mechanics in the presence of a perfect fluid. For the perturbation of the fluid energy-momentum tensor we obtain Alonso-Serrano et al. [2023a]

$$\begin{aligned} U^{\nu}\delta T_{\nu}{}^{\mu} =& \frac{1}{2}U^{\nu}\left(\rho+p\right)\left(u_{\nu}u^{\mu} + \frac{1}{n}\delta^{\nu}_{\mu}\right) + \left(\sqrt{-\mathfrak{g}}/\omega\right)^{-1/n}|U|\mu\delta I^{\mu} \\ &+ \left(\sqrt{-\mathfrak{g}}/\omega\right)^{-1/n}|U|\mathcal{T}\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{1/n}\nu s u^{\mu}\right]. \end{aligned} \tag{1.150}$$

The first term is simply the traceless part of the energy-momentum tensor, which precisely cancels out with the last integral in equation (1.136). Then, the first law of black hole mechanics for a stationary, asymptotically flat black hole spacetime filled with a perfect fluid becomes

$$\delta E - \Omega_{\mathcal{H}}^{(i)} \delta J_{(i)} - \frac{1}{8\pi}\kappa \int_{\partial \mathcal{C}_{\mathcal{H}}} \delta \left[\left(\sqrt{-\mathfrak{g}}/\omega \right)^{2/n} \epsilon^{\nu\mu} \right] \mathrm{d}\mathcal{C}_{\mu\nu}$$
$$- \int_{\mathcal{C}} \left(\sqrt{-\mathfrak{g}}/\omega \right)^{-1/n} |U| \mu \delta I^{\mu} \mathrm{d}\mathcal{C}_{\mu} - \int_{\mathcal{C}} \left(\sqrt{-\mathfrak{g}}/\omega \right)^{-1/n} |U| \mathcal{T} \delta \tilde{S}^{\mu} \mathrm{d}\mathcal{C}_{\mu}$$
$$- \int_{\mathcal{C}} \sum_{i=1}^{n-3} \Omega^{(i)} \delta \tilde{J}^{\mu}_{(i)} \mathrm{d}\mathcal{C}_{\mu} = 0. \tag{1.151}$$

In the unimodular gauge and for perturbations that leave the metric determinant unchanged, $\delta\mathfrak{g} = 0$, this formula reduces to the form of the first law in general relativity Bardeen et al. [1973], Iyer [1997], and it has the same physical meaning. Let us briefly focus on the content of equation (1.151). In the first line we find the familiar terms given by the surface integrals over the asymptotic infinity and the horizon. The first integral on the second line quantifies the change in the fluid's energy as a result of an absorption of particles by the black hole. It is given by the perturbation of the particle density current I^{μ}, multiplied by the chemical potential μ, and a red-shift factor $(\sqrt{-\mathfrak{g}}/\omega)^{-1/n} |U|$. The latter accounts for the red-shift between the event horizon and the asymptotic infinity. The last term on the second line quantifies the heat flow from the fluid into the black hole, with $\tilde{S}^{\mu} = (\sqrt{-\mathfrak{g}}/\omega)^{1/n} s u^{\mu}$ being the WTDiff-invariant entropy flux, and $(\sqrt{-\mathfrak{g}}/\omega)^{-1/n} |U| \mathcal{T}$ the red-shifted fluid temperature. This term leads to the decrease of the fluid entropy. Therefore, unless the black hole also possesses entropy which correspondingly increases, it violates the second law of thermodynamics. This observation is one of the main arguments for identifying the term $-(1/8\pi)\,\kappa \int_{\partial\mathcal{C}_{\mathcal{H}}} \delta \left[(\sqrt{-\mathfrak{g}}/\omega)^{2/n} \epsilon^{\nu\mu} \right] \mathrm{d}\mathcal{C}_{\mu\nu}$ with the heat term for the black hole, and, hence, for the notion of Wald entropy. It actually quantifies the famous John Wheeler's question about pouring a cup of tea into a black hole, which helped to shape the field of black hole thermodynamics Wheeler and Ford [1998].

1.3.5 The first law of causal diamonds

The Diff-invariant covariant phase space formalism applies straightforwardly not only to Killing vectors, but also to conformal Killing vectors generating an infinitesimal Weyl transformation of the metric,

$$\delta_{\zeta} g_{\mu\nu} = \pounds_{\zeta} g_{\mu\nu} = 2\nabla_{(\mu}\zeta_{\nu)} = \frac{1}{n}\nabla_{\rho}\zeta^{\rho} g_{\mu\nu}. \tag{1.152}$$

Since ζ^{μ}, like any vector field, generates an infinitesimal diffeomorphism, it corresponds to an infinitesimal symmetry of any Diff-invariant theory. Hence, we can directly apply the machinery of Hamiltonian perturbations expressed in terms of the Noether charge. In Weyl transverse gravity, the situation is different. Since $\tilde{\nabla}_{\mu}\zeta^{\mu} \neq 0$, ζ^{μ} is not a transverse diffeomorphism transformation. While it does generate a Weyl transformation, this statement does not hold in an arbitrary spacetime (unlike the transversality condition), but depends on a particular background metric. Therefore, we cannot define a background-independent conserved Noether current corresponding to ζ^μ. Nevertheless, we can still compute the symplectic potential, symplectic current and, eventually, the Hamiltonian perturbation by brute force, although we lose its relation with background independent Noether charges.

At this point, the difference between δ_ζ and $\pounds_\zeta$ becomes important. While any variation of a non-dynamical volume measure vanishes by definition, i.e., $\delta_\zeta \boldsymbol{\omega} = 0$, for its Lie derivative we have $\pounds_\zeta \boldsymbol{\omega} = \boldsymbol{\omega} a \tilde{\nabla}_\mu \zeta^\mu \neq 0$. Consequently, it can be shown that a conformal Killing vector leaves the auxiliary metric invariant, $\delta_\zeta \tilde{g}_{\mu\nu} = 0$.

Herein, we first derive the Hamiltonian perturbation for a completely arbitrary vector field ζ^μ. Then, we apply it to the particular case of a causal diamond in flat background, which possesses a conformal Killing symmetry. There, we derive the first law of causal diamonds in Weyl transverse gravity and show its physical equivalence with the one valid in general relativity. This result (as well as its generalisation to arbitrary local, WTDiff-invariant theories we derive in the next section) plays an important role in our discussion of thermodynamics of spacetime in chapter 2.

Hamiltonian

The symplectic current corresponding to a diffeomorphism generated by an arbitrary vector field ζ^μ and a small spacetime perturbation reads

$$\Omega^\mu \left[\delta_\zeta, \delta\right] = \delta\theta^\mu \left[\delta_\zeta\right] - \delta_\zeta \theta^\mu \left[\delta\right]. \tag{1.153}$$

We are interested in comparing two solutions of Weyl transverse gravity related by a small perturbation. Therefore, in the following, we assume that both the original and the perturbed spacetime satisfy the vacuum equations of motion. The first term is a perturbation of the symplectic potential $\theta^\mu \left[\delta_\zeta\right]$, which we express from its general definition (1.74)

$$\theta^\mu \left[\delta_\zeta\right] = \frac{1}{16\pi} \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{4}{n}} \left(g^{\mu\nu} g^{\rho\sigma} - g^{\mu\sigma} g^{\nu\rho}\right) \tilde{\nabla}_\sigma \delta_\zeta \tilde{g}_{\nu\rho}, \tag{1.154}$$

with

$$\tilde{\nabla}_\sigma \delta_\zeta \tilde{g}_{\nu\rho} = 2 \tilde{\nabla}_\sigma \tilde{\nabla}_{(\nu} \left(\tilde{g}_{\rho)\lambda} \zeta^\lambda\right) - \frac{2}{n} \tilde{g}_{\nu\rho} \tilde{\nabla}_\sigma \tilde{\nabla}_\lambda \zeta^\lambda. \tag{1.155}$$

After some straightforward calculations, we obtain

$$\theta^\mu \left[\delta_\zeta\right] = \frac{1}{8\pi} \tilde{g}^{\mu\rho} \tilde{R}_{\rho\nu} \zeta^\nu + \tilde{\nabla}_\nu \left(\frac{1}{8\pi} \left(\sqrt{-g}/\omega\right)^{2/n} \tilde{\nabla}^{[\nu} \zeta^{\mu]}\right) + \Pi^\mu_\zeta, \tag{1.156}$$

where we introduced

$$\Pi^\mu_\zeta = \frac{1}{8\pi} \frac{n-1}{n} \tilde{g}^{\mu\nu} \tilde{\nabla}_\nu \tilde{\nabla}_\rho \zeta^\rho. \tag{1.157}$$

Term Π^μ_ζ gives the only contribution which does not appear in the previously analysed case of transverse diffeomorphisms. One may notice that the first term on the right hand side of equation (1.156) corresponds to the Noether charge, in the case when $\tilde{\nabla}_\mu \zeta^\mu = 0$. However, a general ζ^μ does not generate a local symmetry of Weyl transverse gravity and we cannot assign a conserved Noether charge to it (at least not for a general metric). For the perturbation of the symplectic potential (1.156) $\delta\theta^\mu\left[\delta_\zeta\right]$, we have, using the vacuum equations of motion,

$$\delta\theta^\mu\left[\delta_\zeta\right]=\frac{1}{8\pi}\frac{1}{n}\zeta^\mu\delta\tilde{R}+\frac{1}{8\pi}\tilde{\nabla}_\nu\left[\delta\left(\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\tilde{\nabla}^{[\nu}\zeta^{\mu]}\right)\right]+\delta\Pi^\mu_\zeta. \tag{1.158}$$

The second term in equation (1.153) obeys

$$\delta_\zeta\theta^\mu\left[\delta\right]=\zeta^\nu\tilde{\nabla}_\nu\theta^\mu\left[\delta\right]-\theta^\nu\left[\delta\right]\tilde{\nabla}_\nu\zeta^\mu+\frac{2}{n}\theta^\mu\left[\delta\right]\tilde{\nabla}_\nu\zeta^\nu+\frac{1}{16\pi}\tilde{\nabla}_\nu\tilde{\nabla}_\rho\zeta^\rho\delta\tilde{g}^{\mu\nu}. \tag{1.159}$$

Plugging expressions (1.158) and (1.159) into equation (1.153) we find, after some work (see our paper Alonso-Serrano et al. [2023a] for details),

$$\begin{aligned}\Omega^\mu\left[\delta_\zeta,\delta\right]=&-\frac{1}{8\pi}\zeta^\mu\delta\Lambda+\frac{1}{8\pi}\tilde{\nabla}_\nu\left[\delta\left(\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\tilde{\nabla}^{[\nu}\zeta^{\mu]}\right)\right]-\tilde{\nabla}_\nu\left(\zeta^{[\nu}\theta^{\mu]}\left[\delta\right]\right)\\&-\zeta^\mu\tilde{\nabla}_\nu\theta^\nu\left[\delta\right]+\frac{n-2}{n}\theta^\mu\left[\delta\right]\tilde{\nabla}_\nu\zeta^\nu+\frac{1}{16\pi}\frac{n-2}{n}\tilde{\nabla}_\nu\tilde{\nabla}_\rho\zeta^\rho\delta\tilde{g}^{\mu\nu}.\end{aligned} \tag{1.160}$$

Integrating the symplectic current over a Cauchy surface $\mathcal{C}$ yields the symplectic form $\Omega\left[\delta_\zeta,\delta\right]$ and, thus, the Hamiltonian perturbation (assuming that the Hamiltonian exists)

$$\begin{aligned}\delta H_\zeta=&\int_{\partial\mathcal{C}}\left\{\frac{1}{8\pi}\delta\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n}\tilde{\nabla}^{[\nu}\zeta^{\mu]}\right]-2\zeta^\nu\theta^\mu\left[\delta\right]\right\}\mathrm{d}\mathcal{C}_{\mu\nu}\\&-\int_{\mathcal{C}}\frac{1}{8\pi}\delta\Lambda\zeta^\mu\mathrm{d}\mathcal{C}_\mu+\frac{n-2}{n}\int_{\mathcal{C}}\left(\frac{1}{16\pi}\tilde{\nabla}_\nu\tilde{\nabla}_\rho\zeta^\rho\delta\tilde{g}^{\mu\nu}+\theta^\mu\left[\delta\right]\tilde{\nabla}_\nu\zeta^\nu\right)\mathrm{d}\mathcal{C}_\mu.\end{aligned} \tag{1.161}$$

Notably, even for $\delta\Lambda=0$, the perturbation of the Hamiltonian depends on a volume integral. We clarify its interpretation on the example of a causal diamond in the following.

Causal diamonds

Causal diamonds in flat spacetime are defined as the intersection of past and future light cones of a spacelike $(n-1)$-dimensional ball. Being very simple compact objects with a null boundary which encode the causal structure of the spacetime, they are extensively studied in the field of thermodynamics of spacetime Jacobson [2015], Jacobson and Visser [2019a,b, 2023a,b], Svesko [2019], Alonso-Serrano and Liška [2020a,b, 2022, 2023a,b]. Hence, it is worthwhile to also analyse their thermodynamics in Weyl transverse gravity.

In a flat spacetime, we specify the causal diamond by choosing an arbitrary point P, a length scale l and a unit timelike vector n^μ. Then, we construct the $(n-1)$-dimensional spacelike ball Σ_0 of radius l, centred in P and orthogonal to n^μ. The causal diamond is defined as the union of the internal regions of the past and future light cones of Σ_0. We show this construction in figure 1.2. In a curved spacetime, there exist several non-equivalent ways to define a causal diamond Wang [2019]. However, their distinction is irrelevant for the purposes of our discussion here (we return to this issue in the next chapter).

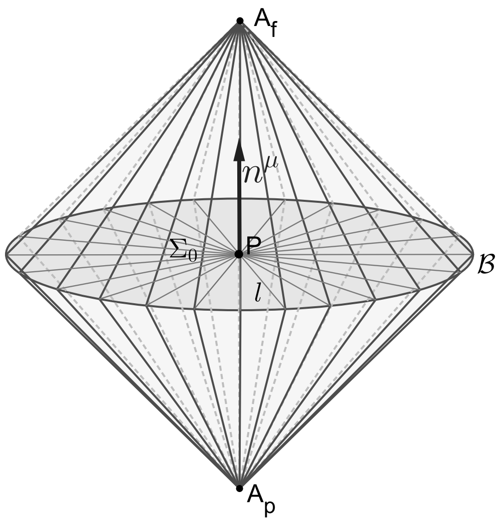

Figure 1.2: A sketch of a causal diamond centred in a spacetime point P. We suppress $n-3$ angular coordinates. The unit, future-directed timelike vector n^μ specifies the local direction of time. Diamond's base is an $(n-1)$-dimensional spatial ball Σ_0 of radius l, whose boundary $\mathcal{B}$ is an approximate $(n-2)$-sphere. The tilted lines starting in the diamond's past apex $A_{\rm p}$ ($t=-l/c$) and ending in the future apex $A_{\rm f}$ ($t=l/c$) demonstrate the null generators of the diamond's boundary. One can see that Σ_0 is the intersection of the future domain of dependence of $A_{\rm p}$ and the past domain of dependence of $A_{\rm f}$.

The metric for the causal diamond can be conveniently written in terms of the Riemann normal coordinates expansion Brewin [2009], choosing the coordinate origin in P, and the local time coordinate so that $n^\mu = (\partial/\partial t)^\mu$. The Riemann normal coordinate expansion around the flat metric reads

$$g_{\mu\nu}(x) = \eta_{\mu\nu} - \frac{1}{3} R_{\mu\alpha\nu\beta}(P)\, x^\alpha x^\beta + O\left(x^3\right), \tag{1.162}$$

where the flat spacetime metric $\eta_{\mu\nu}$ is in arbitrary coordinates. The Christoffel symbols then vanish in point P and, near it, they obey

$$\Gamma^\mu_{\rho\sigma}(x) = -\frac{2}{3} R^\mu{}_{[\rho\sigma]\nu}(P)\, x^\nu + O\left(x^2\right). \tag{1.163}$$

Employing this coordinate expansion, we can identify an approximate (up to $O\left(l^3\right)$) conformal isometry of the causal diamonds generated by conformal Killing vector

$$\zeta^\mu = C\left[\left(l^2 - t^2 - r^2\right)\left(\frac{\partial}{\partial t}\right)^\mu - 2rt\left(\frac{\partial}{\partial r}\right)^\mu\right], \tag{1.164}$$

where r stands for the radial distance from point P, and C is an arbitrary constant. The conformal Killing vector ζ^μ becomes null on the boundary of the causal diamond and vanishes at $\mathcal{B}$. Thus, the diamond boundary corresponds to a conformal Killing horizon and $\mathcal{B}$ to its bifurcation surface. As we discuss in the following, one can study thermodynamics of this horizon using the covariant phase space formalism.

The first law

In Weyl transverse gravity, we, of course, have to define the causal diamond and its conformal Killing vector ζ^μ with respect to the auxiliary, WTDiff-invariant metric $\tilde{g}_{\mu\nu}$. The Riemann normal coordinate expansion also has to be applied to the auxiliary metric. With these subtleties in mind, we can easily derive the first law for the conformal horizon of the causal diamond. We start with a flat spacetime causal diamond and introduce a small perturbation of the metric which solves the vacuum equations of motion of Weyl transverse gravity. To derive the first law, we analyse the perturbation (1.161) of the Hamiltonian generating the evolution along the conformal Killing vector ζ^μ. As $\delta_\zeta \tilde{g}_{\mu\nu} = 0$, it follows that the symplectic current $\Omega^\mu [\delta_\zeta, \delta]$ vanishes and, therefore, $\delta H_\zeta = 0$. Then, equation (1.161) directly leads to the first law of causal diamonds (since the determinant of $\eta_{\mu\nu}$ equals ω^2, we are in the unimodular gauge)

$$\frac{1}{8\pi}\kappa\delta\tilde{\mathcal{A}} - \frac{1}{8\pi}k\kappa\delta\tilde{\mathcal{V}} + \frac{1}{8\pi}\frac{n-1}{n}k\kappa\int_{\Sigma_0}\delta\ln\left(\sqrt{-\mathfrak{g}}/\omega\right)\mathrm{d}^{n-1}x = 0, \qquad (1.165)$$

where $\kappa = 2lC$ denotes the surface gravity corresponding to ζ^μ, and $k = (n-2)/l$ the extrinsic curvature of $\mathcal{B}$ with respect to its embedding in Σ_0. We define $\delta\tilde{\mathcal{A}}$ and $\delta\tilde{\mathcal{V}}$ as the perturbations of the area of $\mathcal{B}$ and the volume of Σ_0, respectively; both measured with respect to the auxiliary metric $\tilde{g}_{\mu\nu}$ and, therefore, WTDiff-invariant. If the determinant does not change under the perturbation we consider, $\delta\mathfrak{g} = 0$, we recover the first law of causal diamonds of general relativity Jacobson and Visser [2019a].

It has been suggested that one can assign temperature $T = \kappa/2\pi$ to the diamond's horizon Jacobson [2015], Jacobson and Visser [2019a,b] (we discuss the question of the normalisation of κ in chapter 2). Then, the first two terms in equation (1.165) become $T\delta S$, with S being the Wald entropy given by equation (1.182). Then, the first law of causal diamonds becomes a genuine first law of thermodynamics, where we now can also identify a work term of the form

$$-\frac{k\kappa}{8\pi}\left[\delta\mathcal{V} + \frac{n-1}{n}\int_{\Sigma_0}\delta\ln\left(\sqrt{-\mathfrak{g}}/\omega\right)\mathrm{d}^{n-1}x\right], \qquad (1.166)$$

where $k\kappa/8\pi$ is an effective pressure and the term in the squared brackets corresponds to the perturbation of the WTDiff-invariant spatial volume of Σ_0.

At the first glance, it might be surprising that we recover the same physical results as in general relativity, since an arbitrary conformal Killing vector generates a local symmetry of general relativity (a diffeomorphism), but it does not correspond to a local symmetry of Weyl transverse gravity (not a transverse diffeomorphism). However, for a flat (or in principle *any* fixed) background, we can fully specify the intersection of Diff and WTDiff groups, which consists of transverse diffeomorphisms and conformal Killing transformations. Then, a conformal Killing vector does generate a WTDiff symmetry of the given background spacetime. It is still somewhat remarkable that this background-dependent statement suffices. It might also indicate that the recent background-dependent analysis of entanglement entropy in arbitrary spacetime regions in terms of von Neumann algebras Jensen et al. [2023] should also directly translate to the WTDiff-invariant setup. We plan to address this question in a future work.

Although the physical content of the first law is the same in Weyl transverse gravity and in general relativity, the volume perturbation enters it in different ways. In Weyl transverse gravity, the volume term comes from the volume integral in the Hamiltonian perturbation (1.161)

$$\frac{n-2}{n}\int_{\mathcal{C}}\left(\frac{1}{16\pi}\tilde{\nabla}_\nu\tilde{\nabla}_\rho\zeta^\rho\delta\tilde{g}^{\mu\nu}+\theta^\mu\left[\delta\right]\tilde{\nabla}_\nu\zeta^\nu\right)\mathrm{d}\mathcal{C}_\mu. \tag{1.167}$$

By contrast, in general relativity, the Hamiltonian is given in as a surface integral. However, unlike in Weyl transverse gravity, the Hamiltonian perturbation corresponding to a conformal Killing vector does not automatically vanish (since $\pounds_\zeta g_{\mu\nu}\neq 0$). Instead, δH_ζ (the "left hand side" of the first law) precisely yields the volume perturbation contribution.

1.4 Covariant phase space formalism for local, WTDiff-invariant theories of gravity

We have seen how the covariant phase space formalism works for a particular (and simplest) WTDiff-invariant theory, Weyl transverse gravity. While this result is of interest by itself, the formalism is much more powerful. Therefore, we now generalise our approach to the case of an arbitrary local, WTDiff-invariant gravitational theory. We have argued that there exists a one to one correspondence between WTDiff- and Diff-invariant Lagrangians (barring some possible loopholes mentioned in subsection 1.1.5). Herein, we thus present a complete WTDiff-invariant alternative to the Diff-invariant covariant phase space formalism developed in the literature Lee and Wald [1990], Wald [1993], Iyer and Wald [1994], Iyer [1997], Wald and Zoupas [2000].

We first derive the symplectic potential for the general WTDiff-invariant action (1.32). A rather lengthy series of manipulations starting from a variation of the Lagrangian yields Alonso-Serrano et al. [2022] (see appendix A.2)

$$\begin{aligned}\theta^\mu\left[\delta\right]=&2\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}}E^{\sigma\nu\rho\mu}\tilde{\nabla}_\sigma\delta\tilde{g}_{\nu\rho}+K^{\mu\nu\rho}\delta\tilde{g}_{\nu\rho}+\sum_{i=2}^{p}M^{\mu\alpha_2\dots\alpha_i}{}_\lambda{}^{\nu\rho\sigma}\delta\tilde{\nabla}_{(\alpha_2}\dots\tilde{\nabla}_{\alpha_i)}\tilde{R}^\lambda{}_{\nu\rho\sigma}\\&+\sum_{i=2}^{q}N^{\mu\alpha_2\dots\alpha_i}\delta\tilde{\nabla}_{(\alpha_2}\dots\tilde{\nabla}_{\alpha_i)}\psi.\end{aligned} \tag{1.168}$$

The WTDiff-invariant tensors $E_\mu{}^{\nu\rho\sigma}$, $K^{\alpha_1\mu\nu}$, $M^{\alpha_1\alpha_2\dots\alpha_i}{}_\mu{}^{\nu\rho\sigma}$, and $N^{\alpha_1\alpha_2\dots\alpha_i}$ are defined so that they have the same symmetries as the terms contracted with them. The precise form of these tensors does not matter for the application of the covariant phase space formalism to black hole spacetimes. All that we need is that they are constructed from $\tilde{g}_{\mu\nu}$, $\tilde{\nabla}_\mu$, $\tilde{R}^\mu{}_{\nu\rho\sigma}$, and ψ. The only exception is $E_\mu{}^{\nu\rho\sigma}$, which possesses the same symmetries as the Riemann tensor and reads

$$E_\mu{}^{\nu\rho\sigma}=\sum_{i=0}^{p}\left(-1\right)^i\tilde{\nabla}_{\alpha_1}\dots\tilde{\nabla}_{\alpha_i}\left(\frac{\partial L}{\partial\tilde{\nabla}_{(\alpha_1}\dots\tilde{\nabla}_{\alpha_i)}\tilde{R}^\mu{}_{\nu\rho\sigma}}\right). \tag{1.169}$$

By construction, the symplectic potential is WTDiff-invariant. In the unimodular gauge and for variations that do not affect the metric determinant, $\delta\mathfrak{g}=0$, equation (1.168) agrees with the symplectic potential of local, Diff-invariant theories of gravity Iyer and Wald [1994].

From the symplectic potential $\theta^{\mu}[\delta]$ we directly obtain the symplectic current using its definition (1.45). However, we are again interested in expressing the symplectic current $\Omega[\pounds_{\xi}, \delta]$[14] in terms of the conserved Noether charges. Thence, we now introduce the Noether currents corresponding to the local symmetries of the action.

We begin by deriving the Noether current for local infinitesimal Weyl transformations, $\delta_{\mathrm{W}} g_{\mu\nu} = 2\sigma g_{\mu\nu}$ (by definition, $\delta_{\mathrm{W}}\psi = 0$). The variation of action (1.32) vanishes by construction, $\delta_{\mathrm{W}} L = 0$. Consequently, vector $\alpha^{\mu}[\delta_{\mathrm{W}}]$ in the general definition of the Noether current (1.53) equals zero, since $\delta_{\mathrm{W}} L = \tilde{\nabla}_{\mu}\alpha^{\mu}[\delta_{\mathrm{W}}] = 0$. Furthermore, symplectic potential (1.168) also vanishes as $\delta_{\mathrm{W}}\tilde{g}_{\mu\nu} = 0$. Therefore, the Noether current corresponding to local, infinitesimal Weyl transformation vanishes in any local, WTDiff-invariant theory, $j^{\mu}[\delta_{\mathrm{W}}] = 0$. This result generalises both our conclusion for Weyl transverse gravity in the previous section and a recent proof valid for an arbitrary WTDiff-invariant theory in four spacetime dimensions Oda [2022]. In the context of Weyl transverse gravity, it has been argued that vanishing $j^{\mu}[\delta_{\mathrm{W}}]$ is closely related to the radiative stability of the cosmological constant and to the absence of an anomaly corresponding to the local Weyl symmetry Oda [2017]. Then, our finding that $j^{\mu}[\delta_{\mathrm{W}}] = 0$ suggests that both properties apply even to arbitrary local, WTDiff-invariant theories of gravity.

Next, we consider infinitesimal transverse diffeomorphisms generated by some vector field ξ^{μ}, such that $\tilde{\nabla}_{\mu}\xi^{\mu} = 0$ and, therefore, $\delta_{\xi} = \pounds_{\xi}$, $\delta_{\xi} g_{\mu\nu} = 2\nabla_{(\mu}\xi_{\nu)}$, $\delta_{\xi}\psi = \pounds_{\xi}\psi$. Lagrangian (1.32) changes under a transverse diffeomorphism as $\pounds_{\xi} L = \xi^{\mu}\tilde{\nabla}_{\mu} L$, and we have $\alpha^{\mu}[\pounds_{\xi}] = L\xi^{\mu}$. Thus, the Noether current reads

$$j^{\mu}_{\xi} = \theta^{\mu}[\pounds_{\xi}] - L\xi^{\mu}. \tag{1.170}$$

Evaluating the WTDiff-invariant divergence of j^{μ}_{ξ}, we obtain, after some work,

$$\begin{aligned}\tilde{\nabla}_{\mu} j^{\mu}_{\xi} = &-\frac{1}{16\pi}\Bigg[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{-2/n}\Bigg(\mathring{A}^{\mu\nu} - (\Phi + \mathcal{J})\,\tilde{g}_{\mu\nu} \\ &+ 8\pi\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k+1}{n}} T^{\mu\nu}\Bigg)\Bigg]\pounds_{\xi} g_{\mu\nu} - \tilde{\nabla}_{\mu}\left(\psi\cdot A_{\psi}\cdot\xi\right)^{\mu},\end{aligned} \tag{1.171}$$

where we recall that $\mathring{A}^{\mu\nu} - (\Phi + \mathcal{J})\,\tilde{g}_{\mu\nu} + 8\pi\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2(k+1)/n} T^{\mu\nu} = \Lambda\tilde{g}_{\mu\nu}$ are the divergence-less gravitational equations of motion. Using the generalised Bianchi identities (1.35) and the transversality condition $\tilde{\nabla}_{\mu}\xi^{\mu} = 0$, we can show that $\tilde{\nabla}_{\mu} j^{\mu}_{\xi}$ vanishes on shell as required of a Noether current. Thence, the Noether

[14] Let us point out a subtle issue. In general, a variation with respect to a vector field ξ^{μ}, δ_{ξ}, applied to a WTDiff-invariant expression does not act like a Lie derivative. On the one hand, the background volume form $\boldsymbol{\omega}$ is non-dynamical. Therefore, its variation by definition vanishes, $\delta_{\xi}\boldsymbol{\omega} = 0$. On the other hand, $\boldsymbol{\omega}$ is an n-form and its Lie derivative equals $\pounds_{\xi}\boldsymbol{\omega} = \boldsymbol{\omega}\tilde{\nabla}_{\mu}\xi^{\mu} \neq 0$. Instead, δ_{ξ} should probably correspond to a suitably defined Lie-gauge derivative Jacobson and Mohd [2015], Aneesh et al. [2020], Elgood et al. [2020], Meessen et al. [2022], Ortín and Pereñíguez [2022], Ballesteros et al. [2023] respecting the WTDiff symmetry. We plan to address this issue in a future work. In any case, for transverse diffeomorphisms, on which we focus here, $\tilde{\nabla}_{\mu}\xi^{\mu} = 0$, and we have that $\delta_{\xi} = \pounds_{\xi}$. We return to the general case $\tilde{\nabla}_{\mu}\xi^{\mu} \neq 0$ in section 1.4.3.

current must obey Wald [1990]

$$j_\xi^\mu = -\left[\frac{1}{8\pi}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{-\frac{2}{n}}\left(\mathring{A}_\nu{}^\mu + 8\pi\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k+1}{n}} T^{\mu\nu}\right)\right.$$
$$\left. - \frac{1}{8\pi}\left(\Phi + \mathcal{J}\right)\delta_\nu^\mu\right]\xi^\nu - \left(A_\psi \cdot \psi \cdot \xi\right)^\mu + \tilde{\nabla}_\nu Q_\xi^{\nu\mu}, \tag{1.172}$$

where the antisymmetric tensor density $Q_\xi^{\nu\mu}$ denotes the Noether charge corresponding to the transverse diffeomorphism generated by the vector field ξ^μ.

Much like in Weyl transverse gravity, the Noether current on shell reduces to

$$j_\xi^\mu = -\frac{1}{8\pi}\Lambda\xi^\mu + \tilde{\nabla}_\nu Q_\xi^{\nu\mu}, \tag{1.173}$$

and again, unlike in the Diff-invariant case, it contains a term proportional to the on-shell value of the cosmological constant.

Let us now take a closer look at the Noether charge. Even without specifying the theory, the structure of $Q_\xi^{\nu\mu}$ can be partially read off from the symplectic potential (1.74) particularised for the case $\theta^\mu\left[\pounds_\xi\right]$. A Lie derivative $\pounds_\xi$ depends at most on the first derivatives of ξ^μ. Therefore, only the term $2E^{\sigma\nu\rho\mu}\tilde{\nabla}_\sigma\left(\pounds_\xi\tilde{g}_{\nu\rho}\right)$ contains second derivatives of ξ^μ, as it represents the only contribution including a derivative of a variation. Thus, $Q_\xi^{\nu\mu}$ has the following structure for any local, WTDiff-invariant theory

$$Q_\xi^{\nu\mu} = 2E^{\nu\mu\rho}{}_\sigma\tilde{\nabla}_\rho\xi^\sigma + W_\rho{}^{\nu\mu}\xi^\rho, \tag{1.174}$$

where $W_\rho{}^{\nu\mu} = W_\rho{}^{[\nu\mu]}$ stands for some WTDiff-invariant tensor, which we cannot further specify without choosing a particular theory.

Lastly, we note that definitions of θ^μ, Ω^μ, j_ξ^μ, and $Q_\xi^{\nu\mu}$ contain the same ambiguities we mentioned for Weyl transverse gravity in the previous subsection. The discussion of the physical impact of these ambiguities again proceeds along the same lines as in Weyl transverse gravity. The upshot is that they do not matter for the physical setups we consider in the following.

1.4.1 Hamiltonian for transverse diffeomorphisms

Following the same strategy as in Weyl transverse gravity, we now find the perturbation of a Hamiltonian generating the evolution along a transverse diffeomorphism generator ξ^μ. We do so for the case when both the original an the perturbed spacetime solve the equations of motion (off-shell results can be found in our paper Alonso-Serrano et al. [2022]). The key point is to equate two on-shell expression for a perturbation of the Noether current, one starting from the general expression (1.170),

$$\delta j_\xi^\mu = \delta\theta^\mu\left[\pounds_\xi\right] - \xi^\mu\tilde{\nabla}_\nu\theta^\nu\left[\delta\right], \tag{1.175}$$

and the other from equation (1.173)

$$\delta j_\xi^\mu = -\frac{1}{8\pi}\xi^\mu\delta\Lambda + \tilde{\nabla}_\nu\delta Q_\xi^{\nu\mu}. \tag{1.176}$$

The equivalence of both expressions yields, after some straightforward manipulations, an equation for the symplectic current corresponding to the transverse diffeomorphism $\pounds_\xi$ and the small perturbation δ

$$\Omega^\mu\left[\pounds_\xi,\delta\right]=\tilde{\nabla}_\nu\left(\delta Q_\xi^{\nu\mu}-2\xi^{[\nu}\theta^{\mu]}\left[\delta\right]\right)-\frac{1}{8\pi}\xi^\mu\delta\Lambda. \tag{1.177}$$

By integrating $\Omega^\mu\left[\pounds_\xi,\delta\right]$ over a suitable Cauchy surface $\mathcal{C}$ in the unperturbed spacetime, we obtain the symplectic form, $\Omega\left[\pounds_\xi,\delta\right]$. Lastly, the Hamilton equations of motion equate the symplectic form with the Hamiltonian perturbation δH_ξ (assuming condition (1.50) is fulfilled and the Hamiltonian exists), and we have

$$\delta H_\xi=\int_{\partial\mathcal{C}}\left(\delta Q_\xi^{\nu\mu}-2\xi^\nu\theta^\mu\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu}-\int_{\mathcal{C}}\frac{1}{8\pi}\delta\Lambda\xi^\mu\mathrm{d}\mathcal{C}_\mu, \tag{1.178}$$

where we used the Gauss theorem to convert an integral over $\mathcal{C}$ to an integral over its boundary $\partial\mathcal{C}$. The perturbation of the Hamiltonian has the same basic structure as in Weyl transverse gravity. It consists of a boundary integral of the Noether charge and the symplectic potential, familiar from Diff-invariant theories, and of a volume term proportional to the perturbation of the cosmological constant. As in the particular case of Weyl transverse gravity, we now apply equation (1.178) to derive the first law of black hole mechanics.

1.4.2 The first law of black hole mechanics

First, we consider the same stationary, asymptotically flat black hole spacetime setup in vacuum as in subsection 1.3.2. In this spacetime, we define the perturbation of the canonical energy

$$\delta E=\int_{\partial\mathcal{C}_\infty}\left(\delta Q_t^{\mu\nu}-2t^\nu\theta^\mu\left[\delta\right]\right)\mathrm{d}\mathcal{C}_{\mu\nu}, \tag{1.179}$$

as the time translational Hamiltonian perturbation contribution at the asymptotic infinity. Similarly, with respect to rotational symmetries, we have $\left(n-1\right)/2$ canonical angular momenta

$$\delta J=-\int_{\partial\mathcal{C}_\infty}\delta Q_\varphi^{\mu\nu}\mathrm{d}\mathcal{C}_{\mu\nu}. \tag{1.180}$$

Evaluating also the horizon contribution (using the same simplifications as in the Weyl transverse gravity case), we obtain the first law of black hole mechanics in any local, WTDiff-invariant gravitational theory

$$\delta E-\sum_{i=1}^{(n-1)/2}\Omega_{\mathcal{H}}^{(i)}\delta J_{(i)}-2\kappa\int_{\partial\mathcal{C}_{\mathcal{H}}}\delta\left(E^{\nu\mu\rho\sigma}\epsilon_{\rho\sigma}\right)\mathrm{d}\mathcal{C}_{\mu\nu}=0. \tag{1.181}$$

Let us recall that $E^{\nu\mu\rho\sigma}$ is defined by equation (1.169). As expected, if we select the unimodular gauge and consider only perturbations that do not change the metric determinant, $\delta\mathfrak{g}=0$, we recover the first law in local, Diff-invariant theories of gravity Wald [1993], Iyer and Wald [1994]. Therefore, the physical content of the first law is equivalent in both cases.

If we identify the (WTDiff-invariant) Hawking temperature $T_{\mathrm{H}} = \kappa/2\pi$, we can heuristically define Wald entropy for an arbitrary local, WTDiff-invariant theory of gravity

$$S = -4\pi \int_{\partial\mathcal{C}_{\mathcal{H}}} E^{\mu\nu\rho\sigma}\epsilon_{\rho\sigma}\mathrm{d}\mathcal{C}_{\mu\nu}. \tag{1.182}$$

In the unimodular gauge, we of course get the Wald entropy prescription valid in Diff-invariant theories Wald [1993], Iyer and Wald [1994]. The first law of black hole mechanics then becomes the genuine first law of thermodynamics

$$\delta E - \sum_{i=1}^{n-3} \Omega_{\mathcal{H}}^{(i)}\delta J_{(i)} - T_{\mathrm{H}}\delta S = 0. \tag{1.183}$$

We note that all the relevant features of the cosmological constant contributions to the first law are captured in the Weyl transverse gravity case, which we discussed in some detail. More on the first law in this setting can be found in our paper Alonso-Serrano et al. [2022].

To conclude our discussion of the first law, we address the case of a stationary, asymptotically flat black hole spacetime in the presence of a perfect fluid. The derivation proceeds exactly in the same way as for Weyl transverse gravity, and we obtain

$$\begin{aligned}
&\delta E - \sum_{i=1}^{n-3} \Omega_{\mathcal{H}}^{(i)}\delta J_{\mathcal{H}}^{(i)} + 2\kappa \int_{\partial\mathcal{C}_{\mathcal{H}}} \delta\left(E^{\nu\mu\rho\sigma}\epsilon_{\rho\sigma}\right)\mathrm{d}\mathcal{C}_{\mu\nu} \\
&- \int_{\mathcal{C}} \left(\sqrt{-\mathfrak{g}}/\omega\right)^{-1/n} |U|\mu\delta I^{\mu}\mathrm{d}\mathcal{C}_{\mu} - \int_{\mathcal{C}} \left(\sqrt{-\mathfrak{g}}/\omega\right)^{-1/n} |U|\mathcal{T}\delta\tilde{S}^{\mu}\mathrm{d}\mathcal{C}_{\mu} \\
&- \int_{\mathcal{C}} \sum_{i=1}^{n-3} \Omega^{(i)}\delta\tilde{J}_{(i)}^{\mu}\mathrm{d}\mathcal{C}_{\mu} = 0,
\end{aligned} \tag{1.184}$$

where the fluid contributions (second and third lines) are the same as in Weyl transverse gravity. This form of the first law represents a WTDiff-invariant equivalent of an earlier Diff-invariant result Iyer [1997] and has the same physical content (naturally, as we fixed $\Lambda = \delta\Lambda = 0$). Thence, we clearly see the equivalence between Diff- and WTDiff-invariant formulations of gravitational theories, exactly in the spirit of our discussion in section 1.1 (see also Carballo-Rubio et al. [2022]).

1.4.3 The first law of causal diamonds

To conclude, we want to repeat our analysis of the first law of causal diamonds we performed in Weyl transverse gravity. However, as the first law must be derived by brute force in this case, it becomes prohibitively difficult to study the general case. Instead, we focus on the class of Lagrangians given as $L\left(\tilde{g}_{\mu\nu}, \tilde{R}^{\mu}{}_{\nu\rho\sigma}\right)$ (without any derivatives of the auxiliary Riemann tensor) together with an arbitrary minimally coupled matter Lagrangian (1.21). As in the Weyl transverse gravity case, we first find an expression for the perturbation of the Hamiltonian corresponding to an arbitrary vector field ζ^{μ}. Since, as explained in subsection 1.3.5, we cannot define background independent conserved Noether currents corresponding to such vector fields, we instead proceed directly from the general definition of the symplectic form 1.46. The derivation is very technical and we leave the details for appendix A.3. The final result reads

$$\delta H_\zeta = \int_{\partial\mathcal{C}} \left[2\delta E^{\nu\mu\rho}{}_\sigma \tilde{\nabla}_\rho \zeta^\sigma - 4\zeta^\sigma \delta\left(\tilde{\nabla}_\rho E^{\nu\mu\rho}{}_\sigma\right) + \tilde{\nabla}_\nu \delta Q^{\nu\mu}_{\psi,\zeta} - 2\zeta^\nu \theta^\mu [\delta]\right] \mathrm{d}\mathcal{C}_{\mu\nu}$$
$$- \int_{\mathcal{C}} \frac{1}{8\pi}\delta\Lambda\zeta^\mu \mathrm{d}\mathcal{C}_\mu + \int_{\mathcal{C}} \Bigg(\Pi^\mu[\zeta,\delta] - \zeta^\mu \tilde{\nabla}_\nu \theta^\nu [\delta] + \frac{4}{n}\delta E^{\mu\nu\rho}{}_\nu \tilde{\nabla}_\rho \tilde{\nabla}_\lambda \zeta^\lambda$$
$$- \frac{4}{n}\delta\left(\tilde{\nabla}_\rho E^{\mu\nu\rho}{}_\nu\right)\tilde{\nabla}_\lambda \zeta^\lambda\Bigg)\mathrm{d}\mathcal{C}_\mu, \tag{1.185}$$

where we have $E_\mu{}^{\nu\rho\sigma} = \partial L / \partial \tilde{R}^\mu{}_{\nu\rho\sigma}$ for the class of the theories we consider, and $\Pi^\mu[\zeta,\delta]$ corresponds to a lengthy expression (A.19) given in appendix A.3. Lastly, $Q^{\nu\mu}_{\psi,\zeta}$ is an antisymmetric tensor depending on the matter variables ψ and their, at most, first derivatives, the auxiliary metric $\tilde{g}^{\mu\nu}$ and the vector field ζ^μ (with no derivatives of ζ^μ). If ζ^μ generates a transverse diffeomorphism, i.e., $\tilde{\nabla}_\mu \zeta^\mu = 0$, the the last integral in equation (1.185) vanishes and we recover the transverse diffeomorphism Hamiltonian perturbation (1.95) (for the class of Lagrangians we work with). Then, the terms $2E^{\nu\mu\rho}{}_\sigma \tilde{\nabla}_\rho \zeta^\sigma - 4\zeta^\sigma \tilde{\nabla}_\rho E^{\nu\mu\rho}{}_\sigma$ and $Q^{\nu\mu}_{\psi,\zeta}$ are the gravitational and matter Noether charges.

Our aim is to obtain the first law of causal diamonds. To this end, we particularise the Hamiltonian perturbation (1.185) to the Hamiltonian generating evolution along the conformal Killing vector of the causal diamond (see equation (1.164)). Our Cauchy surface is the geodesic ball Σ_0 which forms the base of the causal diamond. Let us now simplify the Hamiltonian perturbation. First, we use that ζ^μ vanishes on $\partial\Sigma_0$, and its derivatives obey $\tilde{\nabla}_\mu \zeta^\mu = 0$, and $\tilde{\nabla}_\nu \tilde{\nabla}_\mu \zeta^\mu = -2nC\delta^t_\nu$, with all the higher order derivatives of $\tilde{\nabla}_\mu \zeta^\mu$ vanishing identically. Second, we have $\tilde{\nabla}^\rho \zeta^\sigma = \kappa\epsilon^{\rho\sigma}$, with $\epsilon^{\rho\sigma}$ being the bi-normal to $\partial\Sigma_0$. Third, any terms proportional to the Riemann tensor vanish in flat spacetime we consider. In total, we eventually obtain

$$\delta H_\zeta = \int_{\partial\Sigma_0} 2\kappa\epsilon^{\lambda\sigma}\tilde{g}_{\lambda\rho}\delta E^{\nu\mu\rho}{}_\sigma \mathrm{d}\mathcal{C}_{\mu\nu} - \int_{\Sigma_0} \frac{1}{8\pi}\delta\Lambda\zeta^\mu \mathrm{d}\mathcal{C}_\mu$$
$$+ \int_{\Sigma_0} \frac{\kappa}{l}\delta^t_\sigma \left(\frac{1}{8\pi}\delta\tilde{g}^{\sigma\mu} - 4\delta E^{\mu\nu\sigma}{}_\nu\right)\mathrm{d}\mathcal{C}_\mu. \tag{1.186}$$

The Hamiltonian perturbation δH_ζ consists of two contributions, the gravitational one $\delta H_{\mathrm{g},\zeta}$ and the matter one, $\delta H_{\psi,\zeta}$. The former vanishes since $\delta_\zeta \tilde{g}_{\mu\nu} = 0$. To evaluate the latter (generically non-vanishing Iyer [1997], Jacobson and Visser [2019a]), we follow the same procedure as in subsection 1.3.4 and use that the background energy-momentum tensor vanishes. After some straightforward manipulations, we arrive at

$$\delta H_\zeta = \delta H_{\psi,\zeta} = \int_{\Sigma_0} \left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} \delta T_\nu{}^\mu - \delta\mathcal{J}\delta^\mu_\nu\right]\zeta^\nu \mathrm{d}\mathcal{C}_\mu. \tag{1.187}$$

Altogether, the WTDiff-invariant first law of causal diamonds reads

$$\int_{\Sigma_0} \left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2k/n} \delta T_\nu{}^\mu - \delta\mathcal{J}\delta^\mu_\nu\right]\zeta^\nu \mathrm{d}\mathcal{C}_\mu = \int_{\partial\Sigma_0} 2\kappa\epsilon^{\lambda\sigma}\tilde{g}_{\lambda\rho}\delta E^{\nu\mu\rho}{}_\sigma \mathrm{d}\mathcal{C}_{\mu\nu}$$
$$- \int_{\Sigma_0} \frac{1}{8\pi}\delta\Lambda\zeta^\mu \mathrm{d}\mathcal{C}_\mu + \int_{\Sigma_0} \frac{\kappa}{l}\delta^t_\sigma \left(\frac{1}{8\pi}\delta\tilde{g}^{\sigma\mu} - 4\delta E^{\mu\nu\sigma}{}_\nu\right)\mathrm{d}\mathcal{C}_\mu. \tag{1.188}$$

Let us discuss the various terms in this equation. The left hand side simply corresponds to the perturbation of the matter fields (including the contribution $\delta\mathcal{J}$ corresponding to the possible local energy non-conservation). The first term on the right hand side has the (heuristic) interpretation of the $T\delta S$ term, where $T = \kappa/2\pi$ is the diamond's temperature and S its Wald entropy (1.182). The second term gives the cosmological constant contribution, which, as we discussed, occurs generically in WTDiff-invariant gravity. The last term gives us the perturbation of the so called (WTDiff-invariant) generalised volume $\tilde{W}$ of Σ_0 Bueno et al. [2017], Svesko [2019]

$$\delta\tilde{W} = -\frac{8\pi}{n-2}\int_{\Sigma_0} \delta^t_\sigma \left(\frac{1}{8\pi}\delta\tilde{g}^{\sigma\mu} - 4\delta E^{\mu\nu\rho}{}_\nu\right) \mathrm{d}\mathcal{C}_\mu. \tag{1.189}$$

As in Weyl transverse gravity, the generalised volume perturbation comes from the volume term in the Hamiltonian perturbation (1.185). By contrast, in the Diff-invariant case, the generalised volume contribution comes from the non-vanishing symplectic current $\Omega^\mu\left[\delta_\zeta, \delta\right]$ Bueno et al. [2017].

In the unimodular gauge and for metric perturbations preserving the determinant, $\delta\mathfrak{g} = 0$, we, of course, recover the first law of causal diamond in Diff-invariant gravity Bueno et al. [2017] (barring the cosmological constant perturbation, which must be added *ad hoc* in the Diff-invariant case Jacobson and Visser [2019a]). In summary, the first law of causal diamonds (1.165) for WTDiff-invariant gravity has the same physical content as the Diff-invariant version, except for the generically appearing cosmological constant contribution. These results support the notion that the only physical difference between corresponding WTDiff- and Diff-invariant theories of gravity is the status of the cosmological constant.

2. Semiclassical thermodynamics of spacetime and Weyl transverse gravity

The covariant phase space formalism we have focused on so far clearly exposes how the gravitational dynamics determines the thermodynamic behaviour of causal horizons. This observation applies to a number of different notions of a horizon, including the ones associated with black holes, cosmological ones, and even the observer-dependent ones, such as acceleration (Rindler) horizons or boundaries of causal diamonds[1]. Notably, entropy of the horizon is in each case fully specified by the gravitational Lagrangian. As we showed in chapter 1, this relation between the entropy and the Lagrangian holds not only for Diff-invariant theories of gravity, but also for arbitrary local, WTDiff-invariant theories. In the present chapter, we focus on the so called thermodynamics of spacetime which turns this argument around, showing that the entropy prescription contains enough information to recover the gravitational dynamics. In particular, we review how the local equilibrium conditions applied in every spacetime point encode the (semi)classical equations governing gravitational dynamics. Moreover, we argue that the resulting dynamics is in fact consistent with WTDiff-invariant rather than Diff-invariant gravity. More precisely, we show that if one assumes that the local equilibrium conditions together with the strong equivalence principle encode *all* the information for reconstructing the classical gravitational dynamics, they lead to Weyl transverse gravity.

The seminal works on thermodynamics of spacetime realised the local causal horizons in terms of the local Rindler wedges Jacobson [1995], Padmanabhan [2010], Chirco and Liberati [2010]. However, these are not the best tool for the derivation of the equations governing the gravitational dynamics. The construction of a local Rindler wedge requires to rather arbitrarily "cut" a small enough part of the null congruence forming the horizon. While the cut causes no problems in the case of deriving the Einstein equations, its edges yield unwanted (and not easily handled) contributions to the Clausius relation for more general gravitational dynamics Guedens et al. [2012], Parikh and Svesko [2018]. Furthermore, the Rindler wedge does not have a finite interior region. Then, it becomes difficult to associate to it quantum von Neumann entropy of the matter fields (in fact, it has been argued that the Einstein equations cannot be recovered in this way Carroll and Remmen [2016]). Introducing instead spherical local causal horizons with a well defined finite interior region resolves both issues Parikh and Svesko [2018]. They can be realised as the null boundaries either of local approximate light cones Parikh and Svesko [2018], or of local causal diamonds Jacobson [2015], Svesko [2019]. While both options lead to equivalent results and may be used essentially interchangeably Svesko [2019], we find causal diamonds more

[1] Of course, the notion of radiation and, hence, temperature associated with horizons comes from quantum field theory on curved background. However, the expressions for temperature are very robust, determined just by the kinematic features of the spacetime (the surface gravity of the horizon or the observer's acceleration).

convenient to work with given their finite extension in time. Hence, for the remainder of this and the next chapter, we always realise local causal horizons as the approximate conformal Killing horizons of the causal diamonds.

In the following, we first introduce the construction of causal diamonds and the essential ingredients of their thermodynamics description. Then, we derive the equations for gravitational dynamics from the local equilibrium conditions for causal diamonds and show that the result is consistent with Weyl transverse gravity. To strengthen our argument in favour of Weyl transverse gravity, we analyse two independent derivations; one based on tracking the physical entropy fluxes, the other on considering a small perturbation away from the equilibrium state. Lastly, we show that the correspondence of the local equilibrium conditions with WTDiff-invariant gravitational dynamics holds even beyond the simplest case of Weyl transverse gravity. Namely, we show that the equations of motion for a class of WTDiff-invariant gravitational theories are encoded in the corresponding Wald entropy prescription we obtained in chapter 1.

2.1 Thermodynamics of causal diamonds

Local causal diamonds represent a natural arena for thermodynamic derivations of the equations governing gravitational dynamics, being a finite spacetime region enclosed by a spherical horizon. The causal diamonds also offer remarkable simplicity, being fully described by the position of their centre, a single length scale and a choice of the local direction of time.

From here onward, we always consider local causal diamonds whose size parameter l is much smaller than the local curvature length scale. Otherwise, the construction of a causal diamond detailed in subsection 1.3.5 simply fails as it relies on locally approximating the spacetime by a flat one, up to sufficiently small curvature-dependent corrections. Moreover, we require that the spacetime is well approximated by a smooth Lorentzian manifold. It is generally expected that such an approximation breaks down due to quantum effects at very small length scale. Very general arguments (some based only on the combination of the Newtonian gravity and quantum mechanics) suggests that this breakdown occurs at the length scale of the order of the Planck length, $l_{\mathrm{P}} = \sqrt{G\hbar} \approx 1.6 \cdot 10^{-35}\,\mathrm{m}$ Mead [1964], Garay [1995], Hossenfelder [2013][2]. Therefore, we also demand that l must be much larger than l_{P}.

In subsection 2.1.1, we first introduce two distinct generalisations of causal diamonds to curved spacetimes, light-cone cut local causal diamonds and geodesic local causal diamonds. Both constructions play an important role in thermodynamics of spacetime. Then, we discuss several ways to associate thermodynamic properties to causal diamonds. We discuss their temperature in subsection 2.1.2. Subsection 2.1.3 comments on entropy associated with the conformal Killing horizon of the causal diamond. Finally, subsection 2.1.4 considers entropy of the matter fields present inside the causal diamond. Given the vast literature available on the subject, we cannot explore all the topics we comment on in depth. Nevertheless, we want to convey the idea that, in the last decade, thermody-

[2]These arguments in principle do not exclude the possibility that the breakdown happens already at much larger length scales.

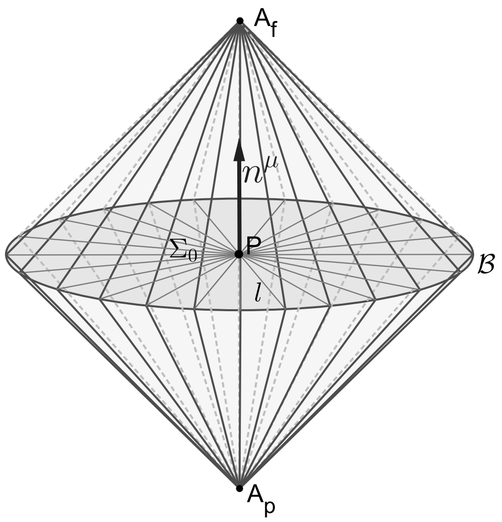

Figure 2.1: A sketch of a causal diamond centred in a spacetime point P. We suppress $n-3$ angular coordinates. The unit, future-directed timelike vector n^μ specifies the local direction of time. Diamond's base is a $(n-1)$-dimensional spatial ball Σ_0 of radius l, whose boundary $\mathcal{B}$ is an approximate $(n-2)$-sphere. The tilted lines starting in the diamond's past apex A_p $(t=-l/c)$ and ending in the future apex A_f $(t=l/c)$ demonstrate the null generators of the diamond's boundary. One can see that Σ_0 is the intersection of the future domain of dependence of A_p and the past domain of dependence of A_f.

namics of causal diamonds has matured into a well established discipline. In fact, essentially all the standard techniques applied to understanding the black hole thermodynamics have also been considered for the case of causal diamonds, yielding very similar results.

2.1.1 Two constructions of a causal diamond in a curved spacetime

Before going to the thermodynamic properties of causal diamonds, we introduce two different ways to define them in curved spacetimes, the light-cone cut local causal diamond and the geodesic local causal diamond. The former turns out to be the natural causal diamond to consider in the physical process version of the derivation, the latter in the equilibrium one. The notation we use and the basic structure of a causal diamond are shown in figure 2.1 (the same figure has originally been displayed in chapter 1, but we also reproduce it here for reader's convenience).

Light-cone cut local causal diamond

The construction of a light-cone cut local causal diamond starts from the past apex A_p. Taking the unit timelike vector n^μ as the local direction of time, we choose a future directed null vector field $k^\mu_\pm$ at A_p normalised so that $n_\mu k^\mu_\pm = -1$. The boundary of the causal diamond is then a congruence of null wordlines tangent to $k^\mu_\pm$. The cross-section of this congruence at the parameter length l

along $k_\pm^\mu$ is an approximate $(n-2)$-sphere $\mathcal{B}$, whose interior, an approximate $(n-1)$-dimensional spacelike ball Σ_0, represents the base of the light-cone cut local causal diamond.

Geodesic local causal diamond

To construct a geodesic local causal diamond, select any regular spacetime point P. In the tangent vector space associated with P, choose an arbitrary unit timelike vector n^μ. In every direction orthogonal to n^μ, send out geodesics of some parameter length l. If we set l to be much smaller than the local curvature length scale (an inverse of the square root of the largest eigenvalue of Riemann tensor), these geodesics are unique and form an $(n-1)$-dimensional spacelike geodesic ball Σ_0 whose boundary is an approximate $(n-2)$-sphere $\mathcal{B}$. The region causally determined by Σ_0 forms the causal diamond.

2.1.2 Temperature

There exist three non-equivalent proposals for assigning a finite temperature to a causal diamond. Since all of them have been considered in various discussions of thermodynamics of spacetime, we review them in the following.

Surface gravity proposal

The first proposal relies on the presence of the conformal Killing horizon and argues that, like the Killing horizons of black holes, it should possess a temperature corresponding to its surface gravity Jacobson and Visser [2019a]

$$T_\kappa = \frac{\hbar\kappa}{2\pi} = \frac{\hbar l C}{\pi}, \tag{2.1}$$

where C is an arbitrary constant which corresponds to the normalisation of the conformal Killing vector ζ^μ. Two choices for C have been put forward. First, setting $C = 1/2l$ has the advantage of having the surface gravity equal to unity Jacobson [2015]. However, the surface gravity (and, hence, temperature) then has incorrect dimensions, making this choice untenable in our view. Second, selecting $C = 1/l^2$ leads to temperature $T_\kappa = \hbar/\pi l$ which is dimensionally correct Svesko [2019].

To fully understand the freedom we have in choosing C, it is instructive to look at the much better explored case of black hole Killing horizons. For an asymptotically flat, stationary black hole spacetime, the same arbitrariness in the Killing vector normal to the horizon exists. It is fixed by requiring that the Killing vector at the asymptotic infinity reduces to a spacetime velocity of an inertial observer. However, we are in principle free to choose any other reference observer with whose velocity we set the Killing vector to agree. Then, the temperature given by the surface gravity of the appropriately normalised Killing horizon yields the temperature measured by this observer[3]. Likewise, the surface gravity itself,

[3]For a generic observer (assuming they even perceive a thermal state with a well defined temperature), this temperature will be sourced by a combination of the Hawking effect (suitably red-shifted) caused by the gravitational pull of the black hole and the Unruh effect occurring due to the observer's acceleration Barbado et al. [2016].

rather than quantifying the force required to hold a unit test mass at rest at the horizon from the asymptotic infinity, becomes the force required to hold such a test mass by the chosen observer. The mass of the black hole is also modified to be the one measured by the observer (while the black hole entropy remains unchanged), ensuring that the first law of black hole thermodynamics continues to hold. A nice way to study the observer-dependent temperature and mass is offered by the Euclidean canonical ensemble construction relying on introducing an artificial York boundary at a finite distance from the black hole horizon. Then, the procedure yields expressions for the temperature and mass measured by an inertial observer at the finite distance boundary and, hence, including the corresponding red-shift factors Braden et al. [1990]. We would get the same result by the covariant phase space approach in the Lorentzian signature by choosing a normalisation of the Killing vector adapted to a finite distance boundary.

The choice of the normalisation of the conformal Killing vector for the causal diamond works in the same way. In particular, taking $C = 1/l^2$ normalises the Killing vector so that, at the origin of the causal diamond it coincides with the velocity of the inertial observer there. For any $C > 1/l^2$, ζ^μ corresponds to the velocity of some accelerating observer moving inside the causal diamond, with $C \to \infty$ representing the limit of the infinite acceleration. Values $C < 1/l^2$ do not have a clear interpretation in terms of observer velocities. It is tempting to interpret the temperatures T_κ corresponding to different values of C as those measured by the various accelerating observers. However, our treatment up to this point disregards the question of whether the local approximate Minkowski vacuum is actually perceived as thermal by the various accelerating observers. To answer this question, it is better to leave the Hawking-like definition of the temperature as being proportional to the surface gravity, and instead study observers moving along accelerated trajectories inside the causal diamond.

Finite time Unruh effect

Consider a uniformly accelerating observer moving inside the causal diamond in the (local approximate) Minkowski vacuum and carrying an Unruh-de Witt particle detector. We show the wordlines of several such observers in figure 2.2. If the Minkowski vacuum were exact and the observer were accelerating for an infinite time, the detector would measure a thermal bath of particles at the Unruh temperature $T_{\mathrm{U}} = \hbar a/2\pi$ Fulling [1973], Bisognano and Wichmann [1976], Unruh [1976]. In our case, the vacuum is only approximate due to curvature effects and the observer only accelerates for a finite time of the order of the diamond's size parameter l. The Unruh effect under such conditions has been analysed in the literature with the result that the detector measures a state well approximated by a thermal bath of particles at the Unruh temperature T_{U} provided that the acceleration a is large enough Barbado and Visser [2012], Rick Perche [2021, 2022]. More precisely, we require $a \gg 1/l$ (the curvature length scale is by assumption very large compared to l and can be neglected altogether). Hence, while the surface gravity dependent temperature T_κ is in principle defined for any constantly accelerating observer inside the causal diamond (for different normalisations of the conformal Killing vector), the corresponding thermal states are apparently only perceived by some of these observers; the ones with sufficiently large accelerations.

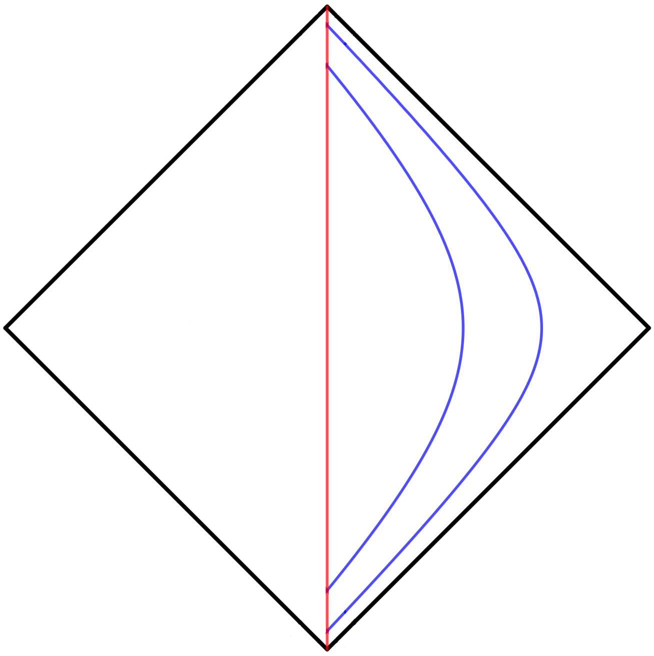

Figure 2.2: Several worldlines of uniformly accelerating observers moving inside a causal diamond are depicted in blue. If their acceleration a is sufficiently large, these observers perform the local Minkowski vacuum as a thermal state at the Unruh temperature $T_{\rm U} = \hbar a/2\pi$. The vertical red line represents the inertial observer with a finite lifetime, $2l$.

We stress that, in our view, this proposal is the only fully physically justified one, as it connects the temperature with a detector response. In subsections 2.2.2 and 2.3.2, we also work with the surface gravity-dependent temperature T_κ. However, we always implicitly assume that κ is sufficiently large so that we can identify T_κ with the temperature $T_{\rm U}$ measured by an Unruh-de Witt detector moving along some uniformly accelerated trajectory inside the causal diamond.

Inertial finite lifetime observers

We have seen that the temperature T_κ remains non-zero even for the normalisation $C = 1/l^2$ corresponding to an inertial observer moving inside the causal diamond. At the first glance, assigning an Unruh temperature to an observer without any acceleration appears rather bizarre. Nevertheless, a finite lifetime observer, being unable to measure the entire spacetime, does perceive a horizon (the future horizon of the causal diamond). Then, the observer cannot access the entire Minkowski vacuum state and it might make sense that their natural notion of vacuum would be distinct from it. That being said, there presently exists no satisfactory connection between the inertial finite lifetime observer's temperature $T_{\rm in} = \hbar/\pi l$ and any detector response. The only evidence available is the existence of a conformal transformation (a combination of a scaling transformation and a special conformal transformation) from an infinite Rindler wedge to a causal diamond, which maps an accelerating trajectory corresponding to an Unruh temperature equal to $\hbar/\pi l$ in the wedge to the inertial one in the diamond Martinetti and Rovelli [2003]. Such a transformation does not affect the

Minkowski vacuum of conformal fields. There are three distinct arguments claiming that the transformation should also not change the temperature. The first proposal relies on the thermal time hypothesis Martinetti and Rovelli [2003]. Second, it has been argued that the conformal Killing vector of the causal diamond can be associated with a Hamiltonian operator in conformal quantum mechanics ($0+1$-dimensional conformal field theory). The eigenstates of this Hamiltonian are then thermal states with temperature T_{in} Arzano [2020, 2021]. Third, the conformal transformation between the Rindler wedge and the diamond has been used to construct a thermofield double state for the diamond's inertial observer with the temperature T_{in} Chakraborty et al..

While the diamond temperature T_{in} is as of yet rather speculative, it does have one attractive feature. Unlike the well established Unruh temperature for the highly accelerated observers or the (also speculative) temperature proportional to the surface gravity, it has a unique value, $T_{\text{in}} = \hbar/\pi l$, fully specified by the size of the causal diamond. As an aside, Bekenstein entropy associated with the causal diamond's horizon equals $S = \pi l^2/l_{\text{P}}^2$. Similarly, for a Schwarzschild black hole, its entropy is $S = \pi r_+^2/l_{\text{P}}^2$ (r_+ being the radius of the event horizon) and its Hawking temperature is $T_{\text{H}} = 1/4\pi r_+$. Then, the entropy and the temperature of the causal diamond would depend on l in the same way as those of the Schwarzschild black hole on r_+ (up to a factor 1/4).

2.1.3 Vacuum entropy

We have seen in the previous chapter that Wald entropy density of the conformal Killing horizon of the causal diamond has the same form as for the event horizon of a black hole. To obtain this result, one simply has to evaluate the Hamiltonian for the region of spacetime bounded by the local causal horizon. A restriction to this region is rather natural from the point of view of the observer perceiving the horizon, as they cannot access the rest of the spacetime. Constructing a (grand)canonical ensemble for the causal diamond (at fixed volume of the spatial ball Σ_0) in the Euclidean signature also leads to the same entropy prescription as one obtains for black holes Jacobson and Visser [2023a,b]. Moreover, the diffeomorphism symmetries of a wide class of null surfaces (including black hole horizons and causal diamond boundaries) form a Virasoro algebra with a central charge Carlip [1999], Chakraborty et al. [2016]. Then, one can compute the entropy from the central charge via the Cardy formula for 2-dimensional conformal field theory. The resulting entropy takes the same form for any type of a horizon. Unfortunately, all of these approaches rely on the knowledge of the gravitational action. Then, invoking them in the process of deriving the equations governing the gravitational dynamics leads to a circular argument.

Nevertheless, a way to define entropy of a local causal horizon independently of the gravitational action exists. The vacuum fluctuations on both sides of the horizon are mutually entangled. This fact follows from the Reeh-Schlieder theorem Reeh and Schlieder [1961] for quantum field theory in a flat spacetime, which asserts that operators defined in an arbitrary $(n-1)$-dimensional spatial subregion acting on the vacuum state can approximate any state defined in the entire $(n-1)$-dimensional spatial slice with arbitrary precision. Hence, there must be quantum entanglement between any subregion and the rest of the spacetime.

Since an observer on one side of the horizon cannot probe the other side, they are unable to access the information present in the entanglement between both sides. Consequently, such an observer perceives a non-zero entanglement entropy. A very simple argument exists for the general form of this entropy Srednicki [1993]. Consider an $(n-1)$-dimensional spatial surface Σ in a flat spacetime that is in vacuum. The vacuum state is by definition pure and its entanglement entropy equals zero. Now consider a separation of Σ into two subregions, Σ_1 and Σ_2. The density operators for these subregions are defined simply as partial traces of the vacuum density operator over the other subregion, i.e., $\rho_{\Sigma_1} = \text{Tr}_{\Sigma_2}\rho_{\text{vac}}$ and vice versa Preskill [2018]. By applying the von Neumann entropy formula

$$S_{\text{vN}} = -\text{Tr}\left(\rho \ln \rho\right), \tag{2.2}$$

we easily find that both subregions have the same entanglement entropy[4], even in the case in which one subregion is infinite and the other finite. Therefore, entanglement entropy certainly cannot scale with volume. Since the only feature shared by both subregions is their boundary, it is natural to expect that entanglement entropy is a function of the boundary's area (although other features of the boundary, such as its Euler characteristic, also apparently play a role in the subleading corrections to entropy Solodukhin [2011]). In fact, calculations of entanglement entropy for various quantum fields show it to be directly proportional to its area to the leading order, $S_{\text{e}} = \eta\mathcal{A}$ Bombelli et al. [1986], Srednicki [1993], Solodukhin [2011], with η being the proportionality constant. Intuitively, this result holds because only the particles very close to the boundary (on different sides of it) become strongly entangled. It follows that the amount of entanglement between both regions simply increases linearly with the size of the boundary.

A naive calculation of the entanglement entropy yields an infinite result. If one introduces some regularisation, the proportionality constant η acquires a finite value which depends both on the quantum fields present and on the regularisation procedure. Most often, a UV cut-off given by some length scale ϵ is introduced. Then, η becomes proportional to the inverse of its square, i.e., $\eta \propto 1/\epsilon^2$. If we take ϵ to be of the order of the Planck length, l_{P}, entanglement entropy is of the same order of magnitude as Bekenstein entropy, $S_{\text{B}} = \mathcal{A}/4l_{\text{P}}^2$. For this reason, quantum entanglement has been put forward as a possible microscopic explanation for black hole entropy Bombelli et al. [1986], Srednicki [1993]. Since entanglement entropy takes the same value for any boundary (to the leading order), it would imply that local causal horizons also possess Bekenstein entropy, as suggested by the gravitational action-dependent approaches we discussed previously. However, some criticism of this interpretation exists as well.

First, entanglement entropy depends on the number of quantum fields of various types present in the spacetime. To address this issue, one may consider approaches that make the cutoff ε (or the Planck length) also sensitive to the matter content of the theory. Then, the dependence completely cancels out in

[4] Von Neumann entropy is a more general concept than entanglement entropy. Whereas entanglement entropy is non-zero only in quantum physics and has no meaningful classical counterpart, von Neumann entropy generalises the classical definitions of both Shannon information entropy and Boltzmann statistical entropy (for deeper insights into the relations of the various entropy concepts, see e.g. Preskill [2018]). However, von Neumann entropy of the vacuum state of some quantum field indeed reduces to entanglement entropy.

entanglement entropy, yielding a universal value Susskind and Uglum [1994], Jacobson [1994], Solodukhin [2011].

Second, while one can motivate the choice of Planck length as the UV cutoff Mead [1964], Garay [1995], Hossenfelder [2013], it lacks a clear justification. Moreover, any explicit UV cutoff breaks the local Lorentz invariance of the theory. However, one can rephrase the calculation with a covariant Pauli-Villars regulator, conforming the cutoff-dependent results Susskind and Uglum [1994].

Third, it has been suggested that, for the entanglement entropy to account for the leading order term in black hole entropy, the vacuum fluctuations must also significantly alter the black hole energy, breaking the self-consistency of the approach 't Hooft [1985], Belgiorno and Liberati [1996], Liberati [1997]. Nevertheless, counter-arguments to this viewpoint were put forward as well Susskind and Uglum [1994], Jacobson [1994], Demers et al. [1995], Banks et al. [2024].

Fourth, a number of approaches to quantum gravity employs some discretisation of the spacetime. In a discrete spacetime, a finite subregion of it can have only finitely many degrees of freedom. At the same time, the Reeh-Schlieder theorem, providing the theoretical justification of quantum entanglement between arbitrary spacelike separated subregions, applies only to systems with infinitely many degrees of freedom. For systems with finitely many degrees of freedom quantum entanglement does not generically occur Agullo et al. [2023]. This observation undermines the entanglement interpretation of Bekenstein entropy if spacetime is discretised. We are aware of no way to refute this objection.

In any case, although the entanglement interpretation of horizon entropy is often assumed in derivations of gravitational dynamics from local equilibrium conditions, the derivation is in fact independent of it. All that we really need is the following. Since the observer perceiving a local causal horizon does not have experimental access to information from the other side, they should measure some Shannon entropy Shannon [1948], Preskill [2018] corresponding to this lack of information. While observer-dependence of entropy seems like an uncomfortable concept, we find it natural. The Unruh effect teaches us that temperature most definitely is an observer-dependent concept. The Unruh temperature also allows us to easily define observer-dependent matter entropy Baccetti and Visser [2014], Carroll and Remmen [2016], Arias et al. [2017]. Moreover, the density operators in quantum field theory are themselves observer-dependent Polo-Gómez et al. [2022], and it follows that the same is true von Neumann entropy (2.2). In view of these insights, we see the observer-dependence of entropy in relativistic physics as a well-established fact. Thus, the local causal horizon should possess entropy. It appears natural that it is determined by the characteristics of the boundary, as it represents the only feature of the "other side" accessible to the observer. Following the logic of the original proposal for the form of black hole entropy Bekenstein [1973], entropy proportional to the area to the leading order is the simplest possibility. Henceforth, we assume that entropy of any local causal horizon obeys $S = \eta\mathcal{A}$, where η is an arbitrary universal constant with dimensions of m^{-2}. For the time being, we need not worry about the microscopic interpretation of this entropy.

2.1.4 Entropy of matter

We miss one last ingredient for the thermodynamic description of causal diamonds; entropy of the matter fields. Herein, we discuss how to compute it. There are two basic ways to define matter entropy. The first one is the semiclassical expression for Clausius entropy, $S_{\mathrm{C}} = \delta Q/T_{\mathrm{U}}$, the second one the quantum von Neumann entropy. In any case, both definitions of entropy lead to equivalent gravitational dynamics Svesko [2019], Alonso-Serrano and Liška [2020a, 2022]. Moreover, for conformally invariant matter fields, their equivalence can be explicitly shown Alonso-Serrano and Liška [2020a].

Let us start with the Clausius entropy flux construction. We follow the procedure introduced in Baccetti and Visser [2014]. Consider a congruence $\mathcal{S}$ of timelike, uniformly accelerating observers with acceleration a who move inside the causal diamond. We choose them so that $\mathcal{S}$ approaches the (null) horizon Σ of the causal diamond. The proper time τ of the uniformly accelerating observers obeys

$$\mathrm{d}t = \cosh\left(a\tau\right)\mathrm{d}\tau, \tag{2.3}$$

where t is the inertial time coordinate inside the diamond (measured along the vector $n^\mu = \left(\partial/\partial t\right)^\mu$) and we only consider $t \in (-l, l)$, as the wordlines of the observers collide at $t = -l$ and $t = l$ (the apices of the diamond). The observers' velocity V^μ and the (spacelike) normal N^μ to $\mathcal{S}$ read

$$V^\mu = \left(\cosh\left(a\tau\right), -\sinh\left(a\tau\right), 0, \dots\right) \tag{2.4}$$

$$N^\mu = \left(-\sinh\left(a\tau\right), \cosh\left(a\tau\right), 0, \dots\right). \tag{2.5}$$

Then, we have for the total heat flux across $\mathcal{S}$

$$\delta Q = -\int_{\mathcal{S}} T_{\mu\nu} V^\nu N^\mu \mathrm{d}^{n-1}\mathcal{S}. \tag{2.6}$$

Since the causal diamond is chosen to be small, we can safely approximate the energy-momentum tensor by its value at the diamond's origin P. As we argued in the previous subsection, the observers measure the Unruh temperature $T_{\mathrm{U}} = \hbar a/2\pi$, provided that $a \gg 1/l$. Thence, the Clausius entropy equals

$$S_C = \frac{\delta Q}{T} = -\frac{2\pi}{\hbar a} T_{\mu\nu}\left(P\right)\int_{\mathcal{S}} V^\nu N^\mu \mathrm{d}^{n-1}\mathcal{S} + O\left(l^{n+2}\right). \tag{2.7}$$

We are interested in the limit of infinite a, where $\mathcal{S}$ coincides with the horizon Σ. In this limit, the heat flux diverges since the integrand grows as $e^{2a\tau}$. However, the coordinate time derivative of the Clausius entropy remains finite and equals Baccetti and Visser [2014]

$$\frac{\mathrm{d}S_C\left(t\right)}{\mathrm{d}t} = \frac{2\pi}{\hbar} t \int_{\mathcal{B}_t} T_{\mu\nu}\left(P\right) k_\pm^\mu k_\pm^\nu \mathrm{d}^{n-2}\mathcal{A} + O\left(l^{n+2}\right), \tag{2.8}$$

where $\mathcal{B}_t$ denotes a spatial cross-section of Σ at time t (an approximate $(n-2)$-sphere) and $k_\pm^\mu = (1, \mp 1, 0, \dots)$ are the future pointing null normals to the future (past) horizon of the causal diamond. This expression applies to any sufficiently small causal diamond.

We can also consider quantum von Neumann entropy of matter fields inside the causal diamond. As a consequence of the Lorentz invariance, the density operator of the $(n-1)$-dimensional spatial ball Σ_0 obeys $\rho = e^{-K/T_\kappa}/\text{Tr}e^{-K/T_\kappa}$, where $T_\kappa = \hbar\kappa/2\pi$ is the temperature of the causal diamond (see the discussion in subsection 2.1.2 for a connection of this temperature with a detector response). The operator K is known as the modular Hamiltonian and equals the boost generator for Σ_0 Bisognano and Wichmann [1976], Jacobson [2015], Arias et al. [2017]. Since Σ_0 is small compared to the local curvature length scale, we can safely use this flat spacetime expression. The modular Hamiltonian K is in general a complicated non-local operator. However, for conformally invariant matter fields, it simply reads

$$K = \int_{\Sigma_0} \langle T^{\mu\nu} \rangle \zeta_\mu \mathrm{d}\Sigma_\nu, \tag{2.9}$$

where $\langle T^{\mu\nu} \rangle$ denotes the quantum expectation value of the energy-momentum tensor operator. Let us now focus on the case of a small perturbation of the vacuum state and linearise in the expectation value of the energy-momentum tensor $\delta\langle T^{\mu\nu} \rangle$. In this case, we can easily apply the von Neumann entropy formula (2.2) to the density operator $\rho = e^{-K/T_\kappa}/\text{Tr}e^{-K/T_\kappa}$, as the exponential can be simply expanded to the linear order in $\delta\langle T^{\mu\nu} \rangle$. We obtain

$$S_{\text{vN}} = \frac{2\pi}{\kappa} \int_{\Sigma_0} \delta\langle T^{\mu\nu} \rangle \zeta_\mu \mathrm{d}\Sigma_\nu + O\left(l^{n+2}\right), \tag{2.10}$$

where the $O\left(l^6\right)$ terms are the curvature-dependent corrections.

One cannot obtain a similarly simple result for general non-conformal quantum field theories. Nevertheless, if the theory has a fixed UV point (around which it is approximately conformal) and Σ_0 is much smaller than all the relevant length scales of the quantum field theory (e.g. the Compton lengths), a local expression for the von Neumann entropy exists Jacobson [2015], Speranza [2016], Casini et al. [2016]. It reads

$$S_{\text{vN}} = \frac{2\pi}{\kappa} \int_{\Sigma_0} \delta\langle T^{\mu\nu} \rangle \zeta_\mu \mathrm{d}\Sigma_\nu + \delta X + O\left(l^{n+2}\right), \tag{2.11}$$

where the new term δX is a rather complicated spacetime scalar that depends on the diamond's size parameter l Speranza [2016], Casini et al. [2016].

2.2 Thermodynamics and Weyl transverse gravity

We have noted in the introduction that the standard derivation of the Einstein equations from thermodynamics of spacetime has a peculiar feature. The local energy conservation needs to be imposed as an additional condition. Consequently, the cosmological constant appears as an arbitrary integration constant that in principle takes different values for different spacetimes. However, the equations of motion of general relativity contain the cosmological constant as a fixed parameter universal for all the solutions of the theory. The equations we obtain from thermodynamics have instead more in common with the divergenceless form of the equations of motion of Weyl transverse gravity (1.26). Herein, we build on our previous works on the subject Alonso-Serrano and Liška [2020a, 2022] as well as on observations of other authors Tiwari [2006], Padmanabhan [2008, 2010]. The results we present follow our most recent paper Alonso-Serrano et al. [2024] and complete the argument for the emergence of Weyl transverse gravity from the local equilibrium conditions. We show that, the local equilibrium conditions and the strong equivalence principle lead to Weyl transverse gravity, provided that they encode *all* the information about the (semi)classical gravitational dynamics.

There is a very simple kinematic reason why one should expect this outcome. The local causal horizons (either of causal diamonds or any other type) constructed in every spacetime point essentially give us the information about the causal structure of the spacetime. It is well known that the causal structure allows one a local kinematic reconstruction of the metric up to an overall conformal factor Hawking and Ellis [1973]. In other words, it is just enough to find the auxiliary metric $\tilde{g}_{\mu\nu}$ (1.10). To completely specify the dynamical metric $g_{\mu\nu}$ including the conformal factor, both the metric reconstruction from the causal structure and the thermodynamics of spacetime require an extra piece of information. Usually, one demands the local energy conservation Hawking and Ellis [1973], Jacobson [1995]. However, there is in principle no need to add any extra condition. Then, it becomes clear that we either need to consider a fixed unimodular gauge $\sqrt{-\mathfrak{g}} = \omega$, or we have to adapt a Weyl invariant description of the spacetime, insensitive to the overall conformal factor of the metric. The idea of the kinematic metric reconstruction from the causal structure can be made sharper with the notion of observers equipped with Unruh-de Witt detectors probing the metric Rick Perche and Martín-Martínez [2022]. We plan to study whether this approach also favours WTDiff invariance in a future work.

Herein, we show that a similar argument can be made on the level of the gravitational dynamics. Specifically, we analyse how Weyl transverse gravity arises in two independent derivations of gravitational dynamics from thermodynamics. The first one is based on a physical process interpretation of the first law and builds on the method presented in Parikh and Svesko [2018], Svesko [2019]. The second one uses the equilibrium interpretation of the first law and follows Jacobson [2015]. In both cases, our starting point is a local causal diamond. However, different constructions of causal diamonds are required in each case (see subsection 2.1.1). A geodesic local causal diamond is considered for the equilibrium approach Jacobson [2015]. It remains unclear whether a different definition of a causal diamond could be used, but it would certainly require significant modifications to the derivation Wang [2019].

In the case of the physical process approach, the type of the causal diamond has not been specified in the seminal work Svesko [2019]. Herein, we rectify this omission. The derivation relies on comparing the entropy of two spatial cross-sections of the diamond's horizon at different times. Therefore, it relies on having the horizon completely determined at any time. However, the geodesic local causal diamond construction fixes the spatial ball at $t = 0$, which forms the base of the causal diamond, leaving the horizon underdetermined. The light-cone cut construction instead defines the diamond by specifying the horizon. Therefore, to our best knowledge, it offers the only definition of a causal diamond usable for the physical process derivation.

Our goal in both approaches is to decide whether the resulting gravitational

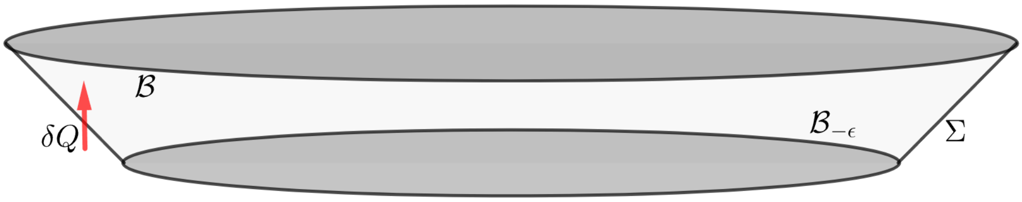

Figure 2.3: The slice of the causal diamond we work with. The spatial cross-sections at times $t = -\epsilon$ and $t = 0$ are denoted by $\mathcal{B}_{-\epsilon}$ and $\mathcal{B}$, respectively. The red arrow shows the physical heat flux δQ across the slice's null boundary Σ.

dynamics are consistent with general relativity or Weyl transverse gravity. Hence, we remain agnostic to whether the causal diamond is defined with respect to the dynamical metric $g_{\mu\nu}$ (as it would be for general relativity), or the auxiliary metric $\tilde{g}_{\mu\nu}$ (for Weyl transverse gravity). To stress that both possibilities are taken into account, we use hatted quantities such as $\hat{g}_{\mu\nu}$, $\hat{\mathcal{A}}$, $\hat{T}_{\mu\nu}$ etc., throughout this section. These can either mean the Diff-invariant expressions, or the corresponding WTDiff-invariant ones. In this way, we avoid having to repeat the analysis twice, once presupposing the Diff invariance and then the WTDiff invariance.

2.2.1 Physical process derivation

In the physical process approach, we study the physical change in entropy of a light-cone cut causal diamond between times $t = -\epsilon$ and $t = 0$. We impose $\epsilon \ll l$, so that we are able to disregard $O\left(\epsilon^3/l^3\right)$ contributions in the following. Hence, we focus on a slice of the diamond's past horizon bounded by the approximate $(n-2)$-sphere $\mathcal{B}_{-\epsilon}$ at $t = -\epsilon$ and by the approximate $(n-2)$-sphere $\mathcal{B}$ at $t = 0$ (see figure 2.3).

On the one side, the total matter entropy inside the causal diamond changes as the matter crosses the horizons. The total entropy flux can be easily computed by integrating equation (2.8) for the time derivative of the Clausius entropy from $t = -\epsilon$ to $t = 0$, i.e.,

$$\Delta S_C = \frac{2\pi}{\hbar}\hat{T}_{\mu\nu}(P)\int_{-\epsilon}^{0}\mathrm{d}t t\,(l+t)^2\int_{\mathcal{B}_t}\mathrm{d}\Omega_{n-2}\hat{k}_-^{\mu}\hat{k}_-^{\nu}. \tag{2.12}$$

Using that $\hat{k}_-^{\mu} = \left(1, \hat{m}^i\right)$ and that

$$\int \hat{m}^i\hat{m}^j\mathrm{d}\Omega_{n-2} = \sigma_{n-2}\frac{\delta^{ij}}{n-1}, \tag{2.13}$$

where σ_{n-2} denotes the area of a unit $(n-2)$-dimensional sphere in flat spacetime, and the latin indices refer to the spatial directions, we obtain

$$\Delta S_C = -\epsilon^2\frac{\pi\sigma_{n-2}l^{n-2}}{\hbar(n-1)}\hat{T}_{\mu\nu}(P)\left(n\hat{n}^{\mu}\hat{n}^{\nu} + \hat{g}^{\mu\nu}(P)\right) + O\left(\epsilon^3\right) + O\left(\epsilon^2 l^n\right). \tag{2.14}$$

We performed an expansion in the small parameter ϵ, discarding all but the leading order terms. The $O\left(\epsilon^2 l^n\right)$ corrections appear due to approximating the energy-momentum tensor by its value in the diamond's centre P and due to neglecting the curvature effects (captured by the Riemann normal coordinate expansion of the metric (1.162)).

On the other side, the horizons expands, which increases its entropy $S = \eta\hat{\mathcal{A}}$. To compute this increase for a light-cone cut causal diamond, the easiest approach is to consider the expansion $\theta = \nabla_\mu \hat{k}_-^\mu$ of the congruence of the null horizon generators $\hat{k}_-^\mu$[5]. By the definition of the expansion, it holds for the change in area between times $t = -\epsilon$ and $t = 0$ Jacobson [1995], Wang [2019]

$$\Delta\hat{\mathcal{A}} = \int_{-\epsilon}^{0} \mathrm{d}\lambda \int \mathrm{d}^{n-2}\hat{\mathcal{A}}\,\theta, \tag{2.15}$$

where λ denotes the null parameter along the horizon generators. The evolution of θ is governed by the Raychaudhuri equation Raychaudhuri [1955]

$$\dot{\theta} = -\frac{1}{n-2}\theta^2 - \sigma^2 - \hat{R}_{\mu\nu}\hat{k}_-^\mu \hat{k}_-^\nu, \tag{2.16}$$

where $\dot{\theta} = \mathrm{d}\theta/\mathrm{d}\lambda$ and $\sigma^2 = \sigma_{\mu\nu}\sigma^{\mu\nu}$, with $\sigma_{\mu\nu} = \hat{\nabla}_{(\mu}\hat{k}_{-,\nu)} - \hat{\nabla}_\rho \hat{k}_-^\rho \hat{h}_{\mu\nu}/\left(n-2\right)$ being the shear tensor and $\hat{h}_{\mu\nu}$ the metric on the null congruence. Since the congruence generates a surface, it follows that its twist must be vanishing and we do not even write the corresponding term in the Raychaudhuri equation. For the shear, we have the following evolution equation

$$\dot{\sigma}_{\mu\nu} = -\frac{2}{n-2}\theta\sigma_{\mu\nu} - \hat{C}_{\lambda\rho\sigma\tau}\hat{k}_-^\lambda \hat{k}_-^\sigma \hat{h}_\mu^\rho \hat{h}_\nu^\tau. \tag{2.17}$$

In a flat spacetime, the shear of the horizon vanishes identically. However, the horizon still expands, at the rate $\theta_{\text{flat}}\left(\lambda\right) = \left(n-2\right)/\left(l+\lambda\right)$ Wang [2019]. Thence, we can solve the equations for the expansion and the shear by expanding them in powers of λ,

$$\theta = \frac{n-2}{l+\lambda} + \theta_{(0)} + \lambda\theta_{(1)} + O\left(\lambda^2\right), \tag{2.18}$$

$$\sigma_{\mu\nu} = \sigma_{(0),\mu\nu} + \lambda\sigma_{(1),\mu\nu} + O\left(\lambda^2\right). \tag{2.19}$$

We set $\theta_{(0)} = \sigma_{(0),\mu\nu} = 0$ as these terms are not sourced by the spacetime curvature and we expect them to take their flat spacetime values[6]. Then, we easily obtain

$$\theta = \frac{n-2}{l+\lambda} - \lambda\hat{R}_{\mu\nu}\hat{k}_-^\mu \hat{k}_-^\nu + O\left(\lambda^2\right), \tag{2.20}$$

$$\sigma_{\mu\nu} = -\lambda\hat{C}_{\lambda\rho\sigma\tau}\hat{k}_-^\lambda \hat{k}_-^\sigma \hat{h}_\mu^\rho \hat{h}_\nu^\tau + O\left(\lambda^2\right). \tag{2.21}$$

[5]At this point, it becomes crucial to consider a light-cone cut causal diamond. Other constructions of a causal diamond do not specify its null boundary, making it difficult to compute the evolution along it.

[6]This assumption is not innocuous. It actually represents a further constraint on the construction of the causal diamond. The freedom in setting up a light-cone cut causal diamond was examined (without using the name) in Parikh and Svesko [2018], Svesko [2019]. The papers show how to fix $n\left(n+1\right)/2$ arbitrary functions that appear in the derivation, which is precisely the number of independent functions represented by $\theta_{(0)}$ and $\sigma_{(0),\mu\nu}$. Therefore, translating the results of this analysis to the language of the Rauchaudhuri equation suggests that we are indeed free to construct it so that $\theta_{(0)} = \sigma_{(0),\mu\nu} = 0$. We expect to address the issue in more detail in a future paper.

Plugging θ into equation (2.15) yields

$$\Delta\hat{\mathcal{A}} = \int_{-\epsilon}^{0} \mathrm{d}\lambda \frac{n-2}{l+\lambda} \int \mathrm{d}^{n-2}\hat{\mathcal{A}} - \int_{-\epsilon}^{0} \mathrm{d}\lambda\lambda \int \mathrm{d}^{n-2}\hat{\mathcal{A}}\hat{R}_{\mu\nu}\hat{k}_{-}^{\mu}\hat{k}_{-}^{\nu} + O\left(\epsilon^{3}\right) + O\left(\epsilon^{2}l^{n}\right). \tag{2.22}$$

The first integral is non-vanishing even for a flat spacetime diamond. In that case, the entropy associated with the horizon increases without any corresponding changes in the matter entropy (we are in vacuum). Therefore, this term cannot correspond to a reversible process for which the total change in entropy must vanish. Instead, we may see the expansion of the causal diamond as an irreversible process, akin to a free expansion of a gas released from a container. In both cases, the system produces entropy without any corresponding transfer of heat[7] (in our case proportional to the energy-momentum tensor). As an aside, it has been argued that this term in fact corresponds to the change in the volume of the spatial cross-section of the causal diamond Svesko [2019]. However, this volume is not unambiguously defined for a light-cone cut causal diamond (unlike for a geodesic local causal diamond), as we specify the diamond via its null boundary Wang [2019]. It might be tempting to actually define the preferred choice of volume in this way, i.e.,

$$\Delta\hat{\mathcal{V}} = \int_{-\epsilon}^{0} \mathrm{d}\lambda \frac{n-2}{n-1} \int \mathrm{d}^{n-2}\hat{\mathcal{A}}. \tag{2.23}$$

However, we do not require such identification and, therefore, we do not pursue this line of reasoning any further.

The reversible contribution $\Delta\hat{\mathcal{A}}_{\mathrm{rev}}$ can be evaluated using essentially the same procedure as for the Clausius entropy flux. We find

$$\begin{aligned}\Delta\hat{\mathcal{A}}_{\mathrm{rev}} =& - \int_{-\epsilon}^{0} \mathrm{d}\lambda\lambda \int \mathrm{d}^{n-2}\hat{\mathcal{A}}\hat{R}_{\mu\nu}\hat{k}_{-}^{\mu}\hat{k}_{-}^{\nu} + O\left(\epsilon^{3}\right) + O\left(\epsilon^{2}l^{n}\right) \\ =& - \epsilon^{2}\frac{\sigma_{n-2}l^{n-2}}{2\left(n-1\right)}\hat{R}_{\mu\nu}\left(P\right)\left(n\hat{n}^{\mu}\hat{n}^{\nu} + \hat{g}^{\mu\nu}\left(P\right)\right) + O\left(\epsilon^{3}\right) + O\left(\epsilon^{2}l^{n}\right).\end{aligned} \tag{2.24}$$

Therefore, the reversible change in the horizon entropy $S = \eta\hat{\mathcal{A}}$ equals

$$\Delta S_{\mathrm{rev}} = -\epsilon^{2}\frac{\eta\sigma_{n-2}l^{n-2}}{2\left(n-1\right)}\hat{R}_{\mu\nu}\left(P\right)\left(n\hat{n}^{\mu}\hat{n}^{\nu} + \hat{g}^{\mu\nu}\left(P\right)\right) + O\left(\epsilon^{3}\right) + O\left(\epsilon^{2}l^{n}\right). \tag{2.25}$$

The total change in entropy for a reversible process must vanish, implying $\Delta S_{\mathrm{rev}} + \Delta S_{\mathrm{C}} = 0$. After some straightforward manipulations, this condition becomes

$$\left(\hat{R}_{\mu\nu}\left(P\right) - \frac{2\pi}{\eta\hbar}\hat{T}_{\mu\nu}\left(P\right)\right)\left(n\hat{n}^{\mu}\hat{n}^{\nu} + \hat{g}^{\mu\nu}\left(P\right)\right) = 0. \tag{2.26}$$

For every unit, timelike vector field $\hat{n}^{\mu}$ defined in P, we can construct a light-cone cut causal diamond and obtain equation (2.26). Therefore, the equation holds for an arbitrary unit, timelike vector. As we prove in appendix A.4, the contractions with $\hat{n}^{\mu}$ can then be removed and it must hold

$$\hat{R}_{\mu\nu}\left(P\right) - \frac{1}{n}\hat{R}\left(P\right)\hat{g}_{\mu\nu}\left(P\right) = \frac{2\pi}{\eta\hbar}\left(\hat{T}_{\mu\nu}\left(P\right) - \frac{1}{n}\hat{T}\left(P\right)\hat{g}_{\mu\nu}\left(P\right)\right). \tag{2.27}$$

[7] If we were to keep non-zero σ_0 and/or θ_0 in equations (2.18) and (2.19), respectively, these would presumably also correspond to entropy production in an irreversible process. This interpretation has been shown in detail for Rindler wedges Chirco and Liberati [2010].

The strong equivalence principle guarantees that η is a universal constant[8] and that the same equation can be derived in any regular spacetime point. Finally, using the Newtonian limit to define $G = 1/4\hbar\eta$, we obtain the traceless equations for gravitational dynamics

$$\hat{R}_{\mu\nu} - \frac{1}{n}\hat{R}\hat{g}_{\mu\nu} = 8\pi G\left(\hat{T}_{\mu\nu} - \frac{1}{n}\hat{T}\hat{g}_{\mu\nu}\right). \tag{2.28}$$

Taking the divergence of the equations and using the Bianchi identities, we find that $\hat{\nabla}^{\nu}\hat{T}_{\mu\nu} = \hat{\nabla}_{\mu}\mathcal{J}$ for some scalar $\mathcal{J}$. Then, following the same steps as for the case of Weyl transverse gravity (see subsection 1.1.1) we finally obtain the divergenceless equations

$$\hat{R}_{\mu\nu} - \frac{1}{2}\hat{R}\hat{g}_{\mu\nu} + \Lambda\hat{g}_{\mu\nu} = 8\pi G\left(\hat{T}_{\mu\nu} - \mathcal{J}\hat{g}_{\mu\nu}\right), \tag{2.29}$$

where Λ is an arbitrary integration constant.

So far, we have been agnostic about the local symmetries of the gravitational dynamics we derived. Equations (2.29) provide a purely metric description of gravity. Hence, they can possess at most $n\left(n+1\right)/2$ local symmetries. Assuming we do not introduce any gauge fixing, we have only two choices (see the discussion at the beginning of section 1.1); the Diff and the WTDiff group. While the thermodynamics of spacetime cannot directly probe the symmetry group, we can nevertheless make a strong case for favouring the WTDiff symmetry. First, only the traceless equations (2.28) can be directly derived from thermodynamics. These do not suffice to recover all $n\left(n+1\right)/2$ components of the metric tensor $g_{\mu\nu}$. However, they suffice to fully specify the auxiliary WTDiff-invariant metric $\tilde{g}_{\mu\nu}$. Thence, the dynamical equations (2.28) are fully consistent just by themselves (together with the matter equations of motion, of course) only if we write them in terms of the WTDiff-invariant auxiliary tensors, recovering the equations of motion of Weyl transverse gravity (A.10). Second, our definition of G makes the horizon entropy $S = \eta\hat{\mathcal{A}}$ coincide with Bekenstein entropy of a black hole event horizon. At the same time, the gravitational equations are encoded in a change of this entropy. Therefore, they are not affected if we shift the entropy by a universal constant (while keeping all the other aspects of the theory unchanged). In subsection 1.3.3 we have seen, on the example of de Sitter spacetime, that such an arbitrary shift of entropy is incorporated into Weyl transverse gravity. However, there is no similar mechanism in Diff-invariant theories of gravity. Lastly, the most striking argument for WTDiff-invariance appears if we demand to also recover a gravitational action from which the traceless equations (2.28) are derived. If we assume that gravity is a fundamental interaction, such an action ought to exist and play an important role in the quantum theory (as it does in loop quantum gravity or path integral quantum gravity). And even if we assume gravity to be emergent, it is still reasonable to expect that some effective classical action can be written. However, there exist no Diff-invariant gravitational action with the metric tensor as the dynamical variable that recovers Λ as an integration constant (although non-metric proposals have been put forward Padmanabhan [2010], Montesinos and Gonzalez [2023]). At the same time, we have seen that this

[8]Otherwise, one could measure entropy to distinguish two test black holes at different spacetime points, breaking the equivalence principle for self-gravitating test particles.

behaviour of Λ is natural for any WTDiff-invariant action principle. Therefore, assuming that we can obtain equations (2.28) as the Euler-Lagrange equations of some action, and that the relevant dynamical variable in the variational principle is the metric, we are uniquely led to Weyl transverse gravity. In conclusion, thermodynamics of spacetime leads to the traceless equations of motion of Weyl transverse gravity (1.22)

$$\tilde{R}_{\mu\nu} - \frac{1}{n}\tilde{R}\tilde{g}_{\mu\nu} = 8\pi \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k-1}{n}} \left(T_{\mu\nu} - \frac{1}{n}Tg_{\mu\nu}\right).$$

Therefore, all the hatted quantities we used throughout the derivation ought to be understood as the WTDiff-invariant ones, defined with respect to the auxiliary metric tensor $\tilde{g}_{\mu\nu} = \hat{g}_{\mu\nu}$. The causal structure of the diamond is likewise defined with respect to $\tilde{g}_{\mu\nu}$.

We, of course, have no a priori reason to expect that thermodynamics of spacetime (together with the equivalence principle) suffices to recover all the information about gravitational dynamics. However, in our view, the apparently very strong connection between gravity and thermodynamics makes the assumption that they are indeed fully equivalent worth exploring. And once we take the encoding of gravity in thermodynamics seriously, we are led to conclude that the appropriate gravitational theory is Weyl transverse gravity. Remarkably, arguments completely independent of thermodynamics also suggest this theory as a serious competitor for general relativity. In particular, we have seen that Weyl transverse gravity naturally arises in the field theoretical approach to gravity Barceló et al. [2014], Carballo-Rubio et al. [2022] (as an alternative to general relativity), and that it offers a robust solution to some of the problems related to the value of the cosmological constant Carballo-Rubio [2015], Barceló et al. [2018].

2.2.2 Entanglement equilibrium derivation

To further strengthen our case for Weyl transverse gravity, we also discuss its consistency with the entanglement equilibrium derivation of gravitational dynamics. This approach phrases the local equilibrium conditions entirely in terms of the quantum von Neumann entropy Jacobson [2015]. In other words, it uses equation (2.11) for von Neumann entropy of matter fields and interprets the horizon entropy $S = \eta\hat{\mathcal{A}}$ in terms of vacuum quantum entanglement. On the one side, we need to rely on a specific interpretation of horizon entropy, and the derivation only applies to quantum field theories with a fixed UV point for which von Neumann entropy is given by equation (2.11). On the other side, it has the advantage of using one definition of entropy for both the matter and the horizon. This approach is particularly natural in the AdS/CFT paradigm Lashkari et al. [2014], Faulkner et al. [2014, 2017], which guarantees that horizon entropy can be completely explained in terms of quantum entanglement (via the Ryu-Takayanagi formula Ryu and Takayanagi [2006]). However, here we apply it in a completely general spacetime setting, following Jacobson [2015].

Our starting point is a geodesic local causal diamond in an equilibrium state. It has been suggested that its equilibrium state corresponds to a vacuum, maximally symmetric spacetime with a cosmological constant λ that in principle depends both on the position of the causal diamond and on its size parameter l Jacobson [2015]. Next, we introduce a small arbitrary perturbation of the metric $\delta\hat{g}_{\mu\nu}$ and of the matter fields. The latter leads to a small non-zero expectation value of the energy-momentum tensor, $\delta\langle\hat{T}^{\mu\nu}\rangle$. Since WTDiff-invariant gravity also allows for variations of the cosmological constant, we further consider a small variation of it, $\delta\lambda$.

Since we perturb the causal diamond away from its equilibrium state, we require that the entropy perturbation vanishes to the leading order, which translates into the condition $\delta S_{\text{vN}} + \eta\delta\hat{\mathcal{A}} = 0$. The matter von Neumann entropy obeys equation (2.11). Regarding the entanglement entropy associated with the horizon, a simple calculation using the Riemann normal coordinate expansion of the metric (1.162) yields Jacobson [2015]

$$\delta\hat{\mathcal{A}}_{\mathcal{B}} = -\frac{\sigma_{n-2}l^n}{3(n-1)}\left(\delta\hat{G}_{\mu\nu}\hat{n}^\mu\hat{n}^\nu - \lambda\right) + (n-2)\,\sigma_{n-2}l^{n-3}\delta l + O\left(l^{n+2}\right), \quad (2.30)$$

where we also allowed for variations of the size parameter l. In subsection 2.1.3, we noted that the Euclidean canonical ensemble for a causal diamond is defined at a fixed volume of the spatial ball Σ_0 Jacobson and Visser [2023a,b]. Only then, can one properly define the partition function of the causal diamond and, thence, its entropy. Therefore, we expect that the equilibrium relation $\delta S_{\text{vN}} + \eta\delta\hat{\mathcal{A}} = 0$ holds only if the volume is held fixed. The volume perturbation equals

$$\delta\hat{\mathcal{V}}_{\Sigma_0} = -\frac{\sigma_{n-2}l^{n+1}}{3(n^2-1)}\left(\delta\hat{G}_{\mu\nu}\hat{n}^\mu\hat{n}^\nu - \lambda\right) + \sigma_{n-2}l^{n-2}\delta l + O\left(l^{n+3}\right). \quad (2.31)$$

Therefore, setting

$$\delta l = \frac{l^3}{3(n^2-1)}\left(\delta\hat{G}_{\mu\nu}\hat{n}^\mu\hat{n}^\nu - \lambda\right), \quad (2.32)$$

ensures $\delta\hat{\mathcal{V}}_{\Sigma_0} = 0$. The area variation at constant volume equals

$$\delta\hat{\mathcal{A}}_{\mathcal{B}}|_{\hat{\mathcal{V}}} = -\frac{\sigma_{n-2}l^n}{n^2-1}\left(\delta\hat{G}_{\mu\nu}\hat{n}^\mu\hat{n}^\nu - \lambda\right) + O\left(l^{n+2}\right). \quad (2.33)$$

Furthermore, we have seen that the first law of causal diamonds (1.165) also includes a term proportional to $\delta\Lambda$. Thence, for the equilibrium condition on entropy to hold, we also need to impose the condition $\delta\Lambda = 0$. This additional requirement is necessary for the WTDiff-invariant setting, whereas the Diff-invariant case automatically implies $\delta\Lambda = 0$.

We are now ready to evaluate the equilibrium condition $\delta S_{\text{vN}} + \eta\delta\hat{\mathcal{A}}|_{\hat{\mathcal{V}}} = 0$. We find

$$\frac{2\pi\sigma_{n-2}l^n}{\hbar(n^2-1)}\left(\delta\langle\hat{T}_{\mu\nu}\hat{n}^\mu\hat{n}^\nu + \delta\hat{X}\rangle\right) - \eta\frac{\sigma_{n-2}l^n}{n^2-1}\left(\delta\hat{G}_{\mu\nu}\hat{n}^\mu\hat{n}^\nu - \lambda\right) + O\left(l^{n+2}\right) = 0. \quad (2.34)$$

Simplifying, we obtain

$$8\pi G\left(\delta\langle\hat{T}_{\mu\nu}\rangle - \delta\hat{X}\hat{g}_{\mu\nu}\right) - \delta\hat{G}_{\mu\nu} - \lambda\hat{g}_{\mu\nu} = 0, \quad (2.35)$$

where we, again, defined the Newton constant $G = 1/4\hbar\eta$ and used the arbitrariness of the unit, timelike vector field $\hat{n}^\mu$ to remove the contractions (see appendix A.4 for the proof of this procedure). We have assumed that the local cosmological constant λ is some function dependent on the spacetime position and the diamond's size. To determine λ, we take a trace of the equations, finding[9]

$$\lambda = 8\pi G \left(\frac{1}{n} \delta \langle \hat{T} \rangle - \delta \hat{X} \right) + \frac{n-2}{2n} \delta \hat{R}. \tag{2.36}$$

Finally, the traceless gravitational equations read

$$\delta \hat{R}_{\mu\nu} - \frac{1}{n} \delta \hat{R} = 8\pi G \left(\delta \langle \hat{T}_{\mu\nu} \rangle - \frac{1}{n} \delta \langle \hat{T} \rangle \hat{g}_{\mu\nu} \right). \tag{2.37}$$

By the virtue of the strong equivalence principle, these equations hold throughout the spacetime. Since we ultimately obtain traceless equations of motion, the entire argument for Weyl transverse gravity we provided in the physical process case holds. Therefore, we should take $\hat{g}_{\mu\nu} = \tilde{g}_{\mu\nu}$ (and likewise for all the other tensors) and we have recovered the semiclassical equations of Weyl transverse gravity, relating the traceless part of the classical Ricci tensor with the traceless part of the quantum expectation value of the energy-momentum tensor.

2.3 Thermodynamics of local causal diamonds and WTDiff-invariant gravity

We have argued that thermodynamics of local causal diamonds encodes gravitational equations equivalent to those of Weyl transverse gravity. However, thermodynamics of spacetime has been found to be even more powerful, allowing one to derive the equations of motion of any purely metric gravitational theory whose Lagrangian is a function of only the metric and the Riemann tensor. The derivation relies on taking the appropriate Wald entropy as the entropy associated with the horizon of the local causal diamond. It is not immediately obvious that we can obtain a similar result for modified WTDiff-invariant theories of gravity, since, as we discussed in subsection 1.3.5, the conformal Killing symmetry of the causal diamond does not correspond to a local symmetry of an arbitrary metric in a WTDiff-invariant theory (in contrast to the Diff-invariant case). Nevertheless, in subsection 1.4.3, we derived the first law of causal diamonds for an arbitrary local, WTDiff-invariant theory of gravity whose Lagrangian is a function of the auxiliary metric $\tilde{g}_{\mu\nu}$ and the auxiliary Riemann tensor $\tilde{R}^{\mu}{}_{\nu\rho\sigma}$. We have also established that entropy of the causal diamond is indeed Wald entropy of the given WTDiff-invariant theory. Here, starting from Wald entropy prescription, we show that thermodynamics of local causal diamonds indeed encodes the appropriate equations of motion. We take WTDiff invariance as our starting assumption, as we already discussed its naturalness in the previous section. We consider both the physical process and the entanglement equilibrium derivations. As we discuss in the following, both approaches are actually no longer completely equivalent in the case of modified theories of gravity Bueno et al. [2017], Svesko [2019].

[9] For conformal matter fields with $\langle \hat{T} \rangle = \delta \hat{X} = 0$ that satisfy the local energy-momentum conservation, we simply have $\lambda = (n-2)\, \delta \hat{R}/n = \Lambda$, where Λ denotes the spacetime cosmological constant.

Before going to the derivations, let us comment on some general features of thermodynamics of spacetime applied to modified theories of gravity. Since the Wald entropy expression is determined by the gravitational Lagrangian, using it to derive the equations for gravitational dynamics represents a circular argument. This problem does not appear with entropy proportional to the horizon area, $S = \eta\tilde{\mathcal{A}}$[10], since we can provide robust, model independent arguments for this form of the leading order entropy contribution (see the discussion in subsection 2.1.3). The derivation of the gravitational equations from Wald entropy then essentially shows that the boundary contribution to the variation of the action (from which one obtains Wald entropy) suffices to recover the gravitational dynamics. This fact is of interest by itself and has been also previously noted in a different context Padmanabhan [2008, 2010]. However, it does not allow us to learn anything truly new about gravitational dynamics (compared to what one learns from the $S = \eta\tilde{\mathcal{A}}$ prescription, as we discussed in the previous section).

An improved version of the argument does not work directly with Wald entropy. Instead, assuming that entropy of the horizon can be interpreted in terms of quantum entanglement, it considers a scheme for renormalising this entropy. The procedure then provides expressions both for the renormalised entanglement entropy and for the effective gravitational action. It turns out that the renormalised entropy agrees with the Wald entropy of the effective action at each order in the effective field theory. While this argument provides a justification for using Wald entropy prescription without the need to *a priori* specify the gravitational Lagrangian, it only works if we interpret horizon entropy completely in terms of quantum entanglement. Herein, we tacitly assume that this is the case. In the next chapter we then provide an alternative viewpoint on deriving modified gravitational dynamics from thermodynamics, one that relies neither on any knowledge of the gravitational Lagrangian, nor on some specific interpretation of entropy.

For Diff-invariant gravity, the Wald entropy approaches lead to another subtle issue. Wald entropy does not depend on the cosmological constant term in the gravitational Lagrangian in any way. Therefore, the equations one obtains from thermodynamics fail to reproduce this term (being, as we have shown, traceless). Instead, the cosmological constant appears as an on-shell integration constant in the process of solving the equations. However, this failure to completely reconstruct the information in the Lagrangian disappears in WTDiff-invariant gravity. In that case, we have noted that any constant term in the Lagrangian has no effect on the equations of motion derived by varying the Lagrangian. It also has no effect on the equations derived from Wald entropy. Therefore, both the variational and the thermodynamic derivation of the equations for gravitational dynamics are perfectly consistent.

In subsection 2.3.1, we discuss the physical process derivation. The method based on the Raychaudhuri equation we developed cannot be straightforwardly generalised beyond the case of $f(R)$ theories of gravity (for an explanation of this limitation in the context of local Rindler wedges, see Guedens et al. [2012]). Instead, we consider a different derivation based on thermodynamics of light-cone cut causal diamonds, which is essentially a WTDiff-invariant version of

[10] $\tilde{\mathcal{A}}$ stands for the WTDiff-invariant area of the horizon measured with respect to the auxiliary metric $\tilde{g}_{\mu\nu}$.

the approach introduced in Parikh and Svesko [2018], Svesko [2019]. Subsection 2.3.2 then builds on the entanglement equilibrium derivation originally developed in Bueno et al. [2017], modifying it to work in the WTDiff-invariant setting.

2.3.1 Physical process derivation

We first discuss the physical process approach in which the gravitational equations are encoded in the change of entropy along the horizon generators of a light-cone cut causal diamond (constructed with respect to the auxiliary, WTDiff-invariant metric $\tilde{g}_{\mu\nu}$). The basic idea of the derivation follows Svesko [2019], although we use a WTDiff-invariant setup and introduce some minor modifications of the argument. Our geometric setup is the same as in subsection 2.2.1, i.e., a slice of the causal diamond sketched in the figure 2.3. We again assume that we are comparing the diamond's entropy between the horizon's spatial cross-sections $\mathcal{B}_{-\epsilon}$ at $t=-\epsilon$ (with $\epsilon \ll l$) and $\mathcal{B}$ at $t=0$. The change in matter Clausius entropy is the same as in the previous case, i.e.,

$$\begin{aligned}\Delta S_C =& -\epsilon^2\frac{n\pi\sigma_{n-2}l^{n-2}}{\hbar\left(n-1\right)}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k-1}{n}}\left(T_{\mu\nu}\left(P\right)-\frac{1}{n}T\tilde{g}_{\mu\nu}\right)\tilde{n}^\mu\tilde{n}^\nu \\ &+O\left(\epsilon^3\right)+O\left(\epsilon^2l^n\right),\end{aligned} \tag{2.38}$$

where the WTDiff-invariant timelike vector $\tilde{n}^\mu$ is normalised to $\tilde{g}_{\mu\nu}\tilde{n}^\mu\tilde{n}^\nu=-1$.

The Wald entropy of a WTDiff-invariant theory whose Lagrangian is a function of $\tilde{g}_{\mu\nu}$ and $\tilde{R}^\mu{}_{\nu\rho\sigma}$ reads

$$S_{\mathrm{W}}\left(t\right)=\frac{2\pi}{\hbar\kappa}\int_{\mathcal{B}_t}Q_\zeta^{\nu\mu}\mathrm{d}\mathcal{B}_{\mu\nu}, \tag{2.39}$$

$$Q_\zeta^{\nu\mu}=2E^{\nu\mu\rho}{}_\sigma\tilde{\nabla}_\rho\zeta^\sigma-4\tilde{\nabla}_\rho E^{\nu\mu\rho}{}_\sigma\zeta^\sigma, \tag{2.40}$$

where $E^{\nu\mu\rho}{}_\sigma=\partial L/\partial\tilde{R}^{\nu\mu\rho}{}_\sigma$. Since entropy of a causal diamond is time-dependent, even the ζ^σ-proportional contribution to the Noether charge, which vanishes for stationary black holes, is non-zero in this case[11]. The change in Wald entropy between the spatial spheres $\mathcal{B}_{-\epsilon}$ and $\mathcal{B}$ then equals

$$\Delta S_{\mathrm{W}}=\int_{-\epsilon}^0\mathrm{d}t\frac{2\pi}{\kappa}\int_{\mathcal{B}_t}\mathrm{d}^{n-2}\tilde{\mathcal{A}}\tilde{k}_-^\mu\tilde{\nabla}_\nu Q_\zeta^{\nu\mu}, \tag{2.41}$$

where $\tilde{k}_-^\mu$ denotes a future-pointing, WTDiff-invariant null normal to the horizon. After some straightforward manipulations, and using the WTDiff-invariant conformal Killing identity

$$\tilde{\nabla}_\nu\tilde{\nabla}_\rho\zeta^\sigma=\tilde{R}_{\lambda\nu\rho}{}^\sigma\zeta^\lambda+\frac{1}{n}\tilde{g}_{\rho\sigma}\tilde{\nabla}_\nu\tilde{\nabla}_\lambda\zeta^\lambda+\frac{1}{n}\tilde{g}_{\nu\sigma}\tilde{\nabla}_\rho\tilde{\nabla}_\lambda\zeta^\lambda-\frac{1}{n}\tilde{g}_{\nu\rho}\tilde{\nabla}_\sigma\tilde{\nabla}_\lambda\zeta^\lambda, \tag{2.42}$$

we obtain

$$\begin{aligned}\Delta S_{\mathrm{W}}=&\int_{-\epsilon}^0\mathrm{d}t\frac{4\pi}{\hbar\kappa}\int_{\mathcal{B}_t}\mathrm{d}^{n-2}\tilde{\mathcal{A}}\tilde{k}_-^\mu\Big[E_\mu{}^{\lambda\rho\sigma}\tilde{R}^\tau{}_{\nu\rho\sigma}\tilde{g}_{\tau\lambda}\zeta^\lambda-2\tilde{\nabla}_\nu\tilde{\nabla}_\sigma E_\mu{}^{\nu\rho\sigma}\tilde{g}_{\rho\lambda}\zeta^\lambda \\ &+\frac{4}{n}\tilde{\nabla}_\rho\tilde{\nabla}_\lambda\zeta^\lambda\tilde{g}_{\nu\sigma}E_\mu{}^{\nu\rho\sigma}\Big]+O\left(\epsilon^3\right).\end{aligned} \tag{2.43}$$

[11] The presence of terms proportional to ζ^σ in entropy is in the same spirit as for the recent proposal for entropy of dynamical black holes Hollands et al. [2024].

Other contributions appear due to the approximate nature of our conformal Killing vector which does not exactly satisfy the conformal Killing identity (2.42). However, these can be cancelled out by adding suitable higher order terms to the definition of the conformal Killing vector Svesko [2019] (the WTDiff-invariant case proceeds exactly in the same way as the Diff-invariant one).

Unlike the rest of the contributions, which are proportional to the spacetime curvature, the term in the second line of equation (2.43) leads to a change of entropy even in a flat spacetime. Thence, just like in the special case of Weyl transverse gravity (see equation 2.22 and the following discussion), we may interpret it as the irreducible entropy production due to the expansion of the causal diamond

$$\Delta S_{\text{irr}} = \int_{-\epsilon}^{0} \mathrm{d}t \frac{4\pi}{\hbar\kappa} \int_{\mathcal{B}_t} \mathrm{d}^{n-2} \tilde{\mathcal{A}} \frac{4}{n} \tilde{\nabla}_\rho \tilde{\nabla}_\lambda \zeta^\lambda \tilde{g}_{\nu\sigma} E_\mu^{\ \nu\rho\sigma}. \tag{2.44}$$

It has been suggested that this term corresponds to the change in the generalised volume of the causal diamond Svesko [2019]. However, we have noted that the volume (or generalised volume) of a light-cone cut causal diamond is not uniquely defined. On the one hand, it would be rather natural to simply use equation (2.44) as its definition, since it does give the correct result in a flat spacetime. On the other hand, we do not need to define the generalised volume at all and simply treat the contribution (2.44) as the non-equilibrium entropy production. The role of the generalised volume becomes crucial in the entanglement equilibrium derivation, as we discuss in the next subsection.

Integrating equation (2.43) and subtracting the entropy production (2.44), we find for the reversible change in Wald entropy

$$\begin{aligned}\Delta S_{\text{W}} = &-\epsilon^2 \frac{8\pi n}{\hbar(n-1)} \sigma_{n-2} l^{n-2} \bigg(E_\mu^{\ \rho\sigma\tau} \tilde{R}^\lambda_{\ \rho\sigma\tau} \tilde{g}_{\lambda\nu} - \frac{1}{n} E_\tau^{\ \lambda\rho\sigma} \tilde{R}^\tau_{\ \lambda\rho\sigma} \tilde{g}_{\mu\nu} \\ &- 2\tilde{\nabla}_\nu \tilde{\nabla}_\sigma E_\mu^{\ \nu\rho\sigma} \tilde{g}_{\rho\lambda} + \frac{2}{n} \tilde{\nabla}_\nu \tilde{\nabla}_\sigma E_\mu^{\ \nu\rho\sigma} \tilde{g}_{\rho\lambda} \bigg) \tilde{n}^\mu \tilde{n}^\nu + O\left(\epsilon^3\right) + O\left(\epsilon^2 l^n\right),\end{aligned} \tag{2.45}$$

where the $O(\epsilon^2 l^n)$ terms correspond to the higher order contributions in the Riemann normal coordinate expansion. Equating the reversible change in Wald entropy with the Clausius entropy flux, we arrive at the equilibrium condition

$$\begin{aligned}\bigg[&E_\mu^{\ \rho\sigma\tau} \tilde{R}^\lambda_{\ \rho\sigma\tau} \tilde{g}_{\lambda\nu} - \frac{1}{n} E_\tau^{\ \lambda\rho\sigma} \tilde{R}^\tau_{\ \lambda\rho\sigma} \tilde{g}_{\mu\nu} - 2\tilde{\nabla}_\nu \tilde{\nabla}_\sigma E_\mu^{\ \nu\rho\sigma} \tilde{g}_{\rho\lambda} + \frac{2}{n} \tilde{\nabla}_\nu \tilde{\nabla}_\sigma E_\mu^{\ \nu\rho\sigma} \tilde{g}_{\rho\lambda} \\ &- \frac{1}{8\pi} \left(\frac{\sqrt{-\mathfrak{g}}}{\omega} \right)^{2\frac{k-1}{n}} \left(T_{\mu\nu}(P) - \frac{1}{n} T \tilde{g}_{\mu\nu} \right) \bigg] \tilde{n}^\mu \tilde{n}^\nu = 0.\end{aligned} \tag{2.46}$$

As in the previous cases, we are free to remove the contractions with the arbitrary, timelike, unit, WTDiff-invariant vector field $\tilde{n}^\mu$ (see the proof in appendix A.4). The Einstein equivalence principle then guarantees the validity of the resulting equations for gravitational dynamics in every regular spacetime point[12]

$$E_{\mu}^{\ \rho\sigma\tau}\tilde{R}^{\lambda}_{\ \rho\sigma\tau}\tilde{g}_{\lambda\nu} - \frac{1}{n}E_{\tau}^{\ \lambda\rho\sigma}\tilde{R}^{\tau}_{\ \lambda\rho\sigma}\tilde{g}_{\mu\nu} - 2\tilde{\nabla}_{\nu}\tilde{\nabla}_{\sigma}E_{\mu}^{\ \nu\rho\sigma}\tilde{g}_{\rho\lambda} + \frac{2}{n}\tilde{\nabla}_{\nu}\tilde{\nabla}_{\sigma}E_{\mu}^{\ \nu\rho\sigma}\tilde{g}_{\rho\lambda}$$
$$- \frac{1}{8\pi}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k-1}{n}}\left(T_{\mu\nu}(P) - \frac{1}{n}Tg_{\mu\nu}\right) = 0. \tag{2.47}$$

We have correctly reproduced the traceless equations of motion of any local, WTDiff-invariant theory of gravity whose Lagrangian is any function of $\tilde{g}_{\mu\nu}$ and $\tilde{R}^{\mu}_{\ \nu\rho\sigma}$, given in equation (A.10) in appendix A.3. We stress that, as we argued in the introduction to this section, both the thermodynamic derivation and the variational principle derivation lead to identical traceless equations for gravitational dynamics and we lose no information in the thermodynamic approach.

2.3.2 Entanglement equilibrium derivation

The entanglement equilibrium approach we analysed in subsection 2.2.2 also allows a generalisation to local, WTDiff-invariant theories of gravity whose Lagrangian is an arbitrary function of $\tilde{g}_{\mu\nu}$ and $\tilde{R}^{\mu}_{\ \nu\rho\sigma}$. Except for assuming WTDiff invariance, the derivation we present here largely follows Bueno et al. [2017].

The renormalised entanglement entropy associated with the horizon of a causal diamond is equal to Wald entropy of certain modified gravity theories. Therefore, we have a well motivated entanglement equilibrium condition

$$\delta S_{\mathrm{W}} + \delta S_{\mathrm{vN}} = 0, \tag{2.48}$$

where δS_{W} is the Wald entropy perturbation and the matter von Neumann entropy perturbation δS_{vN} is given by equation (2.11). We again choose the equilibrium state corresponding to a maximally symmetric spacetime, and allow the value of the local cosmological constant λ to depend on the position in spacetime and on the size parameter of the causal diamond, i.e., $\lambda = \lambda(P, l)$. In the case of Weyl transverse gravity, we have seen that this equilibrium condition only works for perturbations that leave the cosmological constant and the (WTDiff-invariant) volume of the geodesic $(n-1)$-dimensional ball Σ_0 fixed. For modified theories of gravity, the first law of causal diamonds does not contain a variation of the geometric volume but rather of the so-called generalised volume $\delta\tilde{\mathcal{W}}$ (see equation (1.189)). Likewise, the generalised volume is the quantity one must hold fixed to define a Euclidean canonical ensemble for a causal diamond in modified theories of gravity Tavlayan and Tekin [2023]. Therefore, for the equilibrium condition (2.48) to hold, we require $\delta\Lambda = \delta\tilde{\mathcal{W}} = 0$.

Let us explore the expression for the perturbation of Wald entropy of the bifurcate $(n-2)$-surface of the horizon $\mathcal{B}$. Since the conformal Killing vector ζ^{μ} vanishes on the horizon, it holds

$$\delta S_{\mathrm{W}} = \frac{4\pi}{\hbar\kappa}\delta\int_{\mathcal{B}} E^{\nu\mu\rho}_{\ \ \ \ \sigma}\tilde{\nabla}_{\rho}\zeta^{\sigma}\mathrm{d}\mathcal{B}_{\mu\nu}. \tag{2.49}$$

[12] As we previously mentioned, the strong equivalence principle does not hold for modified theories of gravity di Casola et al. [2014, 2015]. In thermodynamics of spacetime, it shows up in the fact that Wald entropy is no longer proportional to the (WTDiff-invariant) horizon area with a universal proportionality constant.

Expanding $E^{\nu\mu\rho}{}_{\sigma}$ in powers of l around the diamond's centre P, we obtain

$$\delta S_{\mathrm{W}} = \frac{4\pi}{\hbar\kappa}\delta\int_{\mathcal{B}}\Bigg(E^{\nu\mu\rho}{}_{\sigma}(P) + l\tilde{m}^{i}\tilde{\nabla}_{i}E^{\nu\mu\rho}{}_{\sigma}(P) + \frac{1}{2}l^{2}\tilde{m}^{i}\tilde{m}^{j}\tilde{\nabla}_{i}\tilde{\nabla}_{j}E^{\nu\mu\rho}{}_{\sigma}(P)\Bigg)\tilde{\nabla}_{\rho}\zeta^{\sigma}\mathrm{d}\mathcal{B}_{\mu\nu}, \tag{2.50}$$

where $\tilde{m}^{i}$ denotes the WTDiff-invariant unit, spatial normal to $\mathcal{B}$. For convenience, we split $E_{\mu}{}^{\nu\rho\sigma}$ into the part corresponding to general relativity and the higher order corrections we denote by $F_{\mu}{}^{\nu\rho\sigma}$, i.e.,

$$E_{\mu}{}^{\nu\rho\sigma} = \frac{1}{32\pi G}\left(\delta^{\rho}_{\mu}\tilde{g}^{\nu\sigma} - \delta^{\sigma}_{\mu}\tilde{g}^{\nu\rho} + F_{\mu}{}^{\nu\rho\sigma}\right). \tag{2.51}$$

Then, integrating equation (2.50), and using that $F_{\mu}{}^{\nu\rho\sigma} = 0$ in a maximally symmetric spacetime, yields

$$\delta S_{\mathrm{W}} = \frac{\delta\tilde{\mathcal{A}}_{\mathcal{B}}}{4l_{\mathrm{P}}^{2}} - \frac{\sigma_{n-2}l^{n-2}}{4l_{\mathrm{P}}^{2}(n-1)}\tilde{g}_{\rho\nu}\tilde{n}^{\mu}\tilde{n}^{\nu}\Bigg[\delta F_{\mu}{}^{\lambda\rho\sigma}\tilde{g}_{\lambda\sigma} + \frac{l^{2}}{2(n+1)}\left(\tilde{g}_{\lambda\sigma}\tilde{g}^{\alpha\beta}\tilde{\nabla}_{\alpha}\tilde{\nabla}_{\beta}\delta F_{\mu}{}^{\lambda\rho\sigma} + \tilde{g}_{\lambda\sigma}\tilde{n}^{\alpha}\tilde{n}^{\beta}\tilde{\nabla}_{\alpha}\tilde{\nabla}_{\beta}\delta F_{\mu}{}^{\lambda\rho\sigma} + 2\tilde{\nabla}_{\lambda}\tilde{\nabla}_{\sigma}\delta F_{\mu}{}^{\lambda\rho\sigma}\right)\Bigg]. \tag{2.52}$$

We have seen that the area variation equals

$$\delta\tilde{\mathcal{A}}_{\mathcal{B}} = -\frac{\sigma_{n-2}l^{n}}{3(n-1)}\left(\delta\tilde{G}_{\mu\nu}\tilde{n}^{\mu}\tilde{n}^{\nu} - \lambda\right) + (n-2)\sigma_{n-2}l^{n-3}\delta l + O\left(l^{n+2}\right). \tag{2.53}$$

For the generalised volume variation, we similarly find

$$\delta\tilde{\mathcal{W}}_{\Sigma_0} = \delta\tilde{\mathcal{V}}_{\Sigma_0} + \frac{\sigma_{n-2}l^{n-1}}{4l_{\mathrm{P}}^{2}(n-1)(n-2)}\tilde{g}_{\lambda\sigma}\tilde{g}_{\rho\nu}\tilde{n}^{\mu}\tilde{n}^{\nu}\Bigg[\delta F_{\mu}{}^{\lambda\rho\sigma} + \frac{l^{2}}{2(n+1)}\left(\tilde{g}^{\alpha\beta}\tilde{\nabla}_{\alpha}\tilde{\nabla}_{\beta}\delta F_{\mu}{}^{\lambda\rho\sigma} + \tilde{n}^{\alpha}\tilde{n}^{\beta}\tilde{\nabla}_{\alpha}\tilde{\nabla}_{\beta}\delta F_{\mu}{}^{\lambda\rho\sigma}\right)\Bigg], \tag{2.54}$$

where

$$\delta\tilde{\mathcal{V}}_{\Sigma_0} = -\frac{\sigma_{n-2}l^{n+1}}{3(n^{2}-1)}\left(\delta\tilde{G}_{\mu\nu}\tilde{n}^{\mu}\tilde{n}^{\nu} - \lambda\right) + \sigma_{n-2}l^{n-2}\delta l + O\left(l^{n+3}\right). \tag{2.55}$$

To satisfy the constraint $\delta\tilde{\mathcal{W}} = 0$, we must set

$$\delta l = \frac{l^{3}}{3(n^{2}-1)}\left(\delta\tilde{G}_{\mu\nu}\tilde{n}^{\mu}\tilde{n}^{\nu} - \lambda\right) - \frac{l}{4l_{\mathrm{P}}^{2}(n-1)(n-2)}\tilde{g}_{\lambda\sigma}\tilde{g}_{\rho\nu}\tilde{n}^{\mu}\tilde{n}^{\nu}\Bigg[\delta F_{\mu}{}^{\lambda\rho\sigma} + \frac{l^{2}}{2(n+1)}\left(\tilde{g}^{\alpha\beta}\tilde{\nabla}_{\alpha}\tilde{\nabla}_{\beta}\delta F_{\mu}{}^{\lambda\rho\sigma} + \tilde{n}^{\alpha}\tilde{n}^{\beta}\tilde{\nabla}_{\alpha}\tilde{\nabla}_{\beta}\delta F_{\mu}{}^{\lambda\rho\sigma}\right)\Bigg]. \tag{2.56}$$

Then, we finally obtain for the perturbation of Wald entropy at fixed generalised volume

$$\delta S_{\mathrm{W}} = \frac{\sigma_{n-2}l^{n}}{4l_{\mathrm{P}}^{2}(n^{2}-1)}\tilde{n}^{\mu}\tilde{n}^{\nu}\left[\delta\tilde{G}_{\mu\nu} - \lambda\tilde{g}_{\mu\nu} - 2\tilde{g}_{\rho\nu}\tilde{\nabla}_{\lambda}\tilde{\nabla}_{\beta}\delta F_{\mu}{}^{\lambda\rho\sigma}\right]. \tag{2.57}$$

Now we are ready to derive the equations governing gravitational dynamics. Expanding the entanglement equilibrium condition (2.48) and performing some straightforward manipulations leads to

$$\left[\delta\tilde{G}_{\mu\nu} - \lambda\tilde{g}_{\mu\nu} - 2\tilde{g}_{\rho\nu}\tilde{\nabla}_{\lambda}\tilde{\nabla}_{\beta}\delta F_{\mu}^{\ \lambda\rho\sigma} - 8\pi G\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k-1}{n}}\left(\delta\langle T_{\mu\nu}\rangle - \delta X\right)\right]\tilde{n}^{\mu}\tilde{n}^{\nu} = 0. \tag{2.58}$$

As always, the contractions with an arbitrary, unit timelike vector field $\tilde{n}^{\mu}$ can be removed and the Einstein equivalence principle ensures the validity of the equilibrium condition throughout the spacetime. Lastly, we take the trace of the equations to determine the local cosmological constant λ

$$\lambda = 8\pi G\left(\frac{1}{n}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k-1}{n}}\delta\langle T\rangle - \delta X\right) + \frac{n-2}{2n}\delta\hat{R} + \frac{2}{n}\tilde{\nabla}_{\lambda}\tilde{\nabla}_{\beta}\delta F_{\rho}^{\ \lambda\rho\sigma}. \tag{2.59}$$

If we limit ourselves to conformal matter fields with $\delta\langle\tilde{T}\rangle = \delta\hat{X} = 0$ and assume that they satisfy the local energy-momentum conservation, we find that $\lambda = (n-2)\,\delta\hat{R}/n + 2\tilde{\nabla}_{\lambda}\tilde{\nabla}_{\beta}\delta F_{\rho}^{\ \lambda\rho\sigma}/n$ is a universal constant. Plugging λ back into the entanglement equilibrium condition (2.58)

$$\delta\tilde{R}_{\mu\nu} - \frac{1}{n}\tilde{R}\tilde{g}_{\mu\nu} - 2\tilde{g}_{\rho\nu}\tilde{\nabla}_{\lambda}\tilde{\nabla}_{\beta}\delta F_{\mu}^{\ \lambda\rho\sigma} - 8\pi G\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k-1}{n}}\left(\delta\langle T_{\mu\nu}\rangle - \frac{1}{n}\delta\langle T\rangle g_{\mu\nu}\right) - = 0. \tag{2.60}$$

In total, we have recovered the linearised traceless equations of motion for a local, WTDiff-invariant theory of gravity whose Lagrangian is an arbitrary function of $\tilde{g}_{\mu\nu}$ and $\tilde{R}^{\mu}_{\ \nu\rho\sigma}$. The WTDiff-invariant thermodynamics of spacetime is thus fully consistent both in the physical process and in the equilibrium approach.

2.3.3 Comparison of the derivations

At the first glance, the physical process approach may seem to be superior as it succeeds in recovering the full, non-linearised equations of motion. However, there is a price to pay. We have seen that the contributions to the Noether charge proportional to the conformal Killing vector field do not vanish in this case. In section 1.3 we discussed that these terms contain ambiguities that the covariant phase space construction of entropy cannot straightforwardly fix. Therefore, the physical process derivation requires that we fix these ambiguities somewhat *ad hoc*. In contrast, the entanglement equilibrium approach, while being limited to linearised equations, suffers from no such ambiguities. The physical process derivation might be improved by employing the recently proposed dynamical prescription for black hole entropy Hollands et al. [2024], which should be unambiguous for the class of Lagrangians we consider. We intend to address this possibility in a future work.

3. Thermodynamics and quantum phenomenological gravitational dynamics

So far, we have been concerned with thermodynamics of spacetime in the semi-classical setting. That is, we treated the spacetime itself as fully classical, while we viewed the matter fields either as quantum (the entanglement equilibrium approach), or even as classical (the physical process approach), invoking their quantum properties only to introduce the Unruh effect. However, thermodynamic methods in principle allow us to go even further, by including quantum corrections to entropy of the local causal horizons. Then, the resulting equations governing the gravitational dynamics contain genuine quantum corrections.

In the present chapter, we introduce a particular realisation of this program we have proposed in Alonso-Serrano and Liška [2020b, 2023a,b]. The idea lies in considering just the leading order quantum correction to entropy, which is a term logarithmic in the horizon area $\mathcal{A}$, i.e.,

$$S_{\text{q}} = \frac{\mathcal{A}}{4l_{\text{P}}^2} + \mathcal{C}\ln\frac{\mathcal{A}}{\mathcal{A}_0} + O\left(\frac{\mathcal{A}_0}{\mathcal{A}}\right), \tag{3.1}$$

where $\mathcal{C}$ is a real number, $\mathcal{A}_0$ a constant with dimensions of area (presumably of the order of l_{P}^2) and $O\left(\mathcal{A}_0/\mathcal{A}\right)$ stands for the subleading corrections that decrease with the area. Remarkably, this expression for entropy in $n = 4$ spacetime dimensions arises independently in a number of distinct approaches. It is implied, e.g. by loop quantum gravity Kaul and Majumdar [2000], Meissner [2004], string theory Banerjee et al. [2011], Sen [2013], Karan and Punia [2023], AdS/CFT correspondence Faulkner et al. [2013], entanglement entropy calculations Solodukhin [2010, 2011], phenomenological approaches based on minimal length and minimal area Adler et al. [2001], Medved and Vagenas [2004], Alonso-Serrano and Liška [2021], quantisation of the horizon area Hod [2004], Davidson [2019], nonlocal effective field theory Xiao and Tian [2022], and even by a model independent analysis of statistical fluctuations Gour and Medved [2003], Medved [2005]. Thence, essentially all the candidate theories of quantum gravity incorporate the logarithmic term in entropy as the leading order quantum correction. The only difference between them lies in the value and sign of the coefficient $\mathcal{C}$ which are theory-dependent.

The logarithmic term in entropy has further remarkable properties. First, the coefficient $\mathcal{C}$ is dimensionless. This feature makes its value a universal number once the theory in which we compute entropy is specified. For instance, in loop quantum gravity, one has explicitly $\mathcal{C} = -3/2$, whereas the leading order term in entropy depends on the, in principle arbitrary, Barbero-Immirzi parameter Meissner [2004] (typically, the value of this parameter is fixed by requiring that one recovers Bekenstein entropy). Similarly, while the leading order contribution to vacuum entanglement entropy depends on the UV cut-off, the coefficient $\mathcal{C}$ is a universal number independent of the cut-off and fully determined by the conformal anomaly Solodukhin [2011] (for conformally invariant quantum fields).

Second, the leading order and the logarithmic terms are the only contributions to entropy that grow with the horizon area. In other words, all the other terms tend to 0 in the limit of $l_{\mathrm{P}} \to 0$, leaving only the first two as being relevant for relatively large horizons. This observation suggests that the logarithmic term can significantly affect the low energy quantum gravitational dynamics, whereas the subleading $O\left(\mathcal{A}_0/\mathcal{A}\right)$ contributions likely only play a role in the highly quantum regime.

Herein, we analyse the local equilibrium conditions for causal diamonds with entropy of the horizon obeying equation (3.1) and show that they encode modified equations governing the gravitational dynamics. The logarithmic term as the leading order quantum correction to entropy is unique to the spacetime dimension $n = 4$. In higher dimensions, various subleading terms larger than the logarithmic one appear, whereas the situation in the lower dimensions is rather specific Solodukhin [2011]. As a result, a straightforward application of our approach in other dimensions fails. Hence, in the rest of the thesis, we always work in $n = 4$ dimensions. In the previous chapter, we have argued that thermodynamics of spacetime leads to WTDiff-invariant description gravity. Since we include the leading order quantum correction to Bekenstein entropy, the resulting equations capture the low-energy quantum gravitational corrections to Weyl transverse gravity. Given the universality of the logarithmic correction to entropy, this procedure can be applied regardless of the final theory of quantum gravity. For the sake of clarity, we focus here on the entanglement equilibrium approach to deriving the equations for gravitational dynamics and we consider only conformally invariant matter fields. We have also applied the physical process approach, obtaining equivalent results Alonso-Serrano and Liška [2020b]. In the near future, we are going to provide a more refined and general physical process derivation, applying the tools developed in Alonso-Serrano et al. [2024].

We discuss the derivation in two stages. First, in section 3.1, we focus on the corrections to linearised gravitational equations. In this case, the result we find corresponds to the equations of motion of WTDiff-invariant quadratic gravity. Then, in section 3.2, we derive the nonlinear equations, while neglecting the higher order Riemann normal coordinate corrections to the metric. We plan to address the general analysis of the local equilibrium conditions without any simplifying assumptions in a future work. We briefly comment on the technique of such an analysis and its possible outcomes in conclusions.

Before going to the specifics of the derivations, let us briefly address the scale at which we apply it. The properties of the logarithmic correction to entropy allow us to easily determine it. We have seen that Bekenstein entropy, being of the order $O\left(1/l_{\mathrm{P}}^2\right)$ yields the traceless equations corresponding to the equations of motion of Weyl transverse gravity. Hence, we expect that the logarithmic correction, being $O\left(1\right)$, will introduce $O\left(l_{\mathrm{P}}^2\right)$ corrections to these equations. Such terms only significantly affect the gravitational dynamics if the local curvature length scale is not too large compared to l_{P}. Furthermore, we have seen that, to construct a causal diamond in a generic curved spacetime, its size parameter l must be much smaller than the local curvature length scale. These requirements together imply that l cannot be that much larger than l_{P}. At the same time, the very description of spacetime as a smooth Lorentzian manifold at length scales comparable with l_{P} appears to break down Mead [1964], Garay [1995], Hossenfelder [2013]. There should still be a sufficient range of acceptable values of the diamond's size parameter l satisfying these conditions and thus allowing us to carry out the derivation. Nevertheless, going forward, we should be mindful of this subtlety.

3.1 Linearised analysis

We begin by discussing the entanglement equilibrium conditions in the case of a gravitational field sufficiently weak to disregard the terms of quadratic or higher order in the curvature tensors. Under this assumption, we can only obtain the linearised equations governing gravitational dynamics. However, it greatly simplifies the derivation, allowing us to introduce its necessary ingredients without being overwhelmed by the technical details, while already yielding interesting results.

Our starting point is a geodesic local causal diamond (see subsection 2.1.1) in entanglement equilibrium. We assume that all the matter fields present in the spacetime are conformally invariant. In that case, we have seen in subsection 2.2.2 that the equilibrium state of the diamond corresponds to a maximally symmetric spacetime described by the local cosmological constant $\lambda = \tilde{R}/4$. However, we have also seen that, for non-conformal matter fields, λ is a more complicated function of the position of the causal diamond and its length parameter, i.e., $\lambda = \lambda\left(P, l\right)$. We do not know *a priori* whether the presence of low energy quantum corrections to horizon entropy similarly makes λ a function of P and l. However, we may assume that, since these corrections are suppressed by l_{P}^2, the local cosmological constant will be equal to $\tilde{R}/4$ up to $O\left(l_{\mathrm{P}}^2\right)$ corrections. Therefore, we adapt the following ansatz for λ

$$\lambda\left(P, l\right) = \frac{1}{4}\tilde{R} + l_{\mathrm{P}}^2 \lambda_{\mathrm{c}}\left(P, l\right). \tag{3.2}$$

Now we introduce a small simultaneous perturbation of the spacetime geometry and of the conformal matter fields. Since we perturb an equilibrium state, the corresponding perturbation of the total von Neumann entropy must vanish, $\delta S_{\mathrm{q}} + \delta S_{\mathrm{vN}} = 0$, provided that there is no irreversible entropy production. In section 2.3.2, we have seen that this condition holds for the perturbations that do not affect the local cosmological constant λ and the (WTDiff-invariant) generalised volume $\tilde{\mathcal{W}}$ of the geodesic ball Σ_0, i.e., $\delta\lambda = \delta\tilde{\mathcal{W}} = 0$. The expression for generalised volume $\tilde{\mathcal{W}}$ follows from the gravitational Lagrangian (see equation (1.189)). As we make no *a priori* assumptions about the Lagrangian, we cannot directly determine $\tilde{\mathcal{W}}$. Nevertheless, we know that $\tilde{\mathcal{W}}$ reduces to the geometric volume $\tilde{\mathcal{V}}$ of Σ_0 (defined with respect to the auxiliary metric $\tilde{g}_{\mu\nu}$) in Weyl transverse gravity. Since the logarithmic term in entropy modifies the Weyl transverse gravity only at the order $O\left(l_{\mathrm{P}}^2\right)$, we may impose the following ansatz for $\tilde{\mathcal{W}}$

$$\tilde{\mathcal{W}} = \tilde{\mathcal{V}} + l_{\mathrm{P}}^2 \tilde{\mathcal{W}}_{\mathrm{q}}\left(P, l, \tilde{n}^{\mu}\right), \tag{3.3}$$

where $\tilde{\mathcal{W}}_{\mathrm{q}}\left(P, l, \tilde{n}^{\mu}\right)$, the leading order quantum correction to the generalised volume, in general depends on the local direction of time $\tilde{n}^{\mu}$.

Upon imposing the constraints $\delta\lambda = \delta\tilde{\mathcal{W}} = 0$, the equilibrium condition $\delta S_{\mathrm{q}} + \delta S_{\mathrm{vN}} = 0$ applies, and we can proceed to study it. The perturbation of the matter von Neumann entropy δS_{vN} for conformal fields obeys the WTDiff-invariant form of equation (2.10), i.e.,

$$S_{\mathrm{vN}} = \frac{8\pi^2 l^4}{15\hbar} \left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4} \delta\langle T_{\mu\nu}\rangle \tilde{n}^\mu \tilde{n}^\nu + O\left(l^6\right). \tag{3.4}$$

To compute the perturbation of the entropy associated with the horizon δS_{q}, we start by perturbing equation (3.1)

$$\delta S_{\mathrm{q}} = \frac{\delta\tilde{\mathcal{A}}}{4l_{\mathrm{P}}^2} + \mathcal{C}\left(\ln\frac{\tilde{\mathcal{A}}_\lambda + \delta\tilde{\mathcal{A}}}{\mathcal{A}_0} - \ln\frac{\tilde{\mathcal{A}}_\lambda}{\mathcal{A}_0}\right), \tag{3.5}$$

where $\tilde{\mathcal{A}}_\lambda$ denotes the WTDiff-invariant area of the boundary $\mathcal{B}$ of Σ_0 in a maximally symmetric spacetime described by the local cosmological constant λ, and $\delta\tilde{\mathcal{A}} = \tilde{\mathcal{A}} - \tilde{\mathcal{A}}_\lambda$ is the area perturbation. Since the Riemann normal coordinate expansion has the leading order corrections to the metric proportional to the auxiliary Riemann tensor, $\delta\tilde{\mathcal{A}}$ is clearly linear in the spacetime curvature. Thence, since we systematically neglect the terms quadratic in curvature, we can disregard the $O\left(\delta\tilde{\mathcal{A}}^2\right)$ contributions to δS_{q}, obtaining

$$\delta S_{\mathrm{q}} = \frac{\delta\tilde{\mathcal{A}}}{4l_{\mathrm{P}}^2} + \mathcal{C}\frac{\delta\tilde{\mathcal{A}}}{\tilde{\mathcal{A}}_\lambda} + O\left(\delta\tilde{\mathcal{A}}^2\right). \tag{3.6}$$

To evaluate the area perturbation, we may again use the Riemann normal coordinate expansion of the auxiliary metric

$$\begin{aligned}\tilde{g}_{\mu\nu}(x) =& \tilde{\eta}_{\mu\nu} - \frac{1}{3}x^\alpha x^\beta \tilde{\eta}_{\mu\lambda}\delta\tilde{R}^\lambda{}_{\alpha\nu\beta} - \frac{1}{6}x^\alpha x^\beta x^\lambda \tilde{\eta}_{\mu\rho}\tilde{\nabla}_\lambda \delta\tilde{R}^\rho{}_{\alpha\nu\beta} \\ & x^\alpha x^\beta x^\lambda x^\rho \left(\frac{2}{45}\tilde{\eta}_{\sigma\tau}\delta\tilde{R}^\sigma{}_{\alpha\mu\beta}\delta\tilde{R}^\tau{}_{\lambda\nu\rho} - \frac{1}{20}\tilde{\eta}_{\mu\sigma}\tilde{\nabla}_\alpha\tilde{\nabla}_\beta\delta\tilde{R}^\sigma{}_{\lambda\nu\rho}\right) + O\left(x^5\right).\end{aligned} \tag{3.7}$$

The term quadratic in the Riemann tensor can be neglected. Then, we expand the area element on $\mathcal{B}$ using the Riemann normal coordinates, integrate, and subtract the equilibrium area $\tilde{\mathcal{A}}_\lambda$. In total, this procedure gives us the following expression for $\delta\tilde{\mathcal{A}}$

$$\delta\mathcal{A} = -\frac{4\pi l^4}{18}\delta\bar{R}^{ij}{}_{ij} - \frac{4\pi l^6}{600}\left(\delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}\right)\nabla_i\nabla_j\delta\bar{R}^m{}_{kml} + 8\pi l\delta l, \tag{3.8}$$

where

$$\delta\bar{R}_{\mu\alpha\nu\beta} = \delta\tilde{R}_{\mu\alpha\nu\beta} + \frac{1}{3}\lambda\left(\tilde{\eta}_{\alpha\nu}\tilde{\eta}_{\beta\mu} - \tilde{\eta}_{\alpha\beta}\tilde{\eta}_{\mu\nu}\right), \tag{3.9}$$

represents the difference between the Riemann tensor of the maximally symmetric background and the Riemann tensor of the perturbed spacetime. Following our experience in the semiclassical case (see subsection 2.2.2), we also included a perturbation δl of the diamond's size parameter in the area perturbation (3.8).

To fix δl, we evaluate the condition $\delta\tilde{\mathcal{W}} = 0$. Applying the same strategy as for the area perturbation, we obtain

$$\begin{aligned}\delta\tilde{\mathcal{W}} =& \delta\tilde{\mathcal{V}} + l_{\mathrm{P}}^2\delta\tilde{\mathcal{W}}_{\mathrm{q}} = -\frac{4\pi l^5}{90}\delta\bar{R}^{ij}{}_{ij} - \frac{4\pi l^7}{4200}\left(\delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}\right) \\ & 4\nabla_i\nabla_j\delta\bar{R}^m{}_{kml} + l_{\mathrm{P}}^2\delta\tilde{\mathcal{W}}_{\mathrm{q}} + 4\pi l^2\delta l.\end{aligned} \tag{3.10}$$

Imposing $\delta\tilde{\mathcal{W}}=0$ implies for δl

$$\delta l=\frac{l^3}{90}\delta\bar{R}^{ij}{}_{ij}+\frac{l^5}{4200}\left(\delta^{ij}\delta^{kl}+\delta^{ik}\delta^{jl}+\delta^{il}\delta^{jk}\right)\nabla_i\nabla_j\delta\bar{R}^m{}_{kml}-l_{\mathrm{P}}^2\delta\tilde{\mathcal{W}}_{\mathrm{q}}. \tag{3.11}$$

Setting this value of δl in the area perturbation (3.8) and performing some straightforward simplifications, we find

$$\delta\tilde{\mathcal{A}}=-\frac{4\pi l^4}{30}\delta\bar{R}^{ij}{}_{ij}-\frac{4\pi l^6}{840}\left(\delta^{ij}\delta^{kl}+\delta^{ik}\delta^{jl}+\delta^{il}\delta^{jk}\right)\nabla_i\nabla_j\delta\bar{R}^m{}_{kml}-\frac{2l_{\mathrm{P}}^2}{l}\delta\tilde{\mathcal{W}}_{\mathrm{q}}. \tag{3.12}$$

The local cosmological constant λ (3.2) is a universal constant in the semiclassical case. Thence, its derivatives are $O\left(l_{\mathrm{P}}^2\right)$ and it holds

$$l_{\mathrm{P}}^2\nabla_i\nabla_j\delta\bar{R}^m{}_{kml}=l_{\mathrm{P}}^2\nabla_i\nabla_j\delta\bar{R}^m{}_{kml}+O\left(l_{\mathrm{P}}^4\right), \tag{3.13}$$

where the $O\left(l_{\mathrm{P}}^4\right)$ corrections can be discarded. We may further simplify the area variation (3.8) by replacing the tensors with spatial indices in a covariant way, using contractions with the unit, WTDiff-invariant, timelike vector field $\tilde{n}^\mu$. We provided the details of the procedure in Alonso-Serrano and Liška [2023a]. The final form of the area perturbation (3.8) then reads

$$\begin{aligned}\delta\tilde{\mathcal{A}}=&\frac{4\pi l^4}{15}\Bigg\{\delta\left(\tilde{R}_{\mu\nu}-\frac{1}{4}\tilde{R}\tilde{g}_{\mu\nu}\right)\tilde{n}^\mu\tilde{n}^\nu-l_{\mathrm{P}}^2\lambda_{\mathrm{c}}+\frac{l^2}{56}\bigg[4\tilde{\nabla}_\mu\tilde{\nabla}_\nu\delta\tilde{R}_{\rho\sigma}\tilde{n}^\mu\tilde{n}^\nu\tilde{n}^\rho\tilde{n}^\sigma\\&+\left(2\tilde{g}^{\lambda\sigma}\tilde{\nabla}_\rho\tilde{\nabla}_\sigma\delta\tilde{R}^\rho{}_{\mu\lambda\nu}+4\tilde{g}^{\lambda\rho}\tilde{\nabla}_\mu\tilde{\nabla}_\rho\delta\tilde{R}_{\lambda\nu}+2\tilde{g}^{\lambda\rho}\tilde{\nabla}_\lambda\tilde{\nabla}_\rho\delta\tilde{R}_{\mu\nu}+\tilde{\nabla}_\mu\tilde{\nabla}_\nu\delta\tilde{R}\right)\tilde{n}^\mu\tilde{n}^\nu\\&+2\tilde{g}^{\lambda\rho}\tilde{g}^{\sigma\tau}\tilde{\nabla}_\rho\tilde{\nabla}_\sigma\delta\tilde{R}_{\lambda\tau}+\tilde{g}^{\lambda\rho}\tilde{\nabla}_\lambda\tilde{\nabla}_\rho\delta\tilde{R}\bigg]\Bigg\}-\frac{2l_{\mathrm{P}}^2}{l}\delta\tilde{\mathcal{W}}_{\mathrm{q}}.\end{aligned} \tag{3.14}$$

At this point, we have all the ingredients necessary to evaluate the equilibrium condition $\delta S_{\mathrm{q}}+\delta S_{\mathrm{vN}}=0$ valid for fixed generalised volume and local cosmological constant. We find that

$$\begin{aligned}&\frac{4\pi l^4}{15}\left(\delta\tilde{R}_{\mu\nu}-\frac{1}{4}\tilde{R}\tilde{g}_{\mu\nu}+l_{\mathrm{P}}^2\lambda_{\mathrm{c}}\tilde{g}_{\mu\nu}\right)\tilde{n}^\mu\tilde{n}^\nu+\frac{\mathcal{C}l^2l_{\mathrm{P}}^2}{15}\Bigg\{\left(\delta\tilde{R}_{\mu\nu}-\frac{1}{4}\delta\tilde{R}\tilde{g}_{\mu\nu}\right)\tilde{n}^\mu\tilde{n}^\nu\\&+\frac{l^2}{56}\bigg[\left(2\tilde{g}^{\lambda\sigma}\tilde{\nabla}_\rho\tilde{\nabla}_\sigma\delta\tilde{R}^\rho{}_{\mu\lambda\nu}+4\tilde{g}^{\lambda\rho}\tilde{\nabla}_\mu\tilde{\nabla}_\rho\delta\tilde{R}_{\lambda\nu}+2\tilde{g}^{\lambda\rho}\tilde{\nabla}_\lambda\tilde{\nabla}_\rho\delta\tilde{R}_{\mu\nu}+\tilde{\nabla}_\mu\tilde{\nabla}_\nu\delta\tilde{R}\right)\tilde{n}^\mu\tilde{n}^\nu\\&+4\tilde{\nabla}_\mu\tilde{\nabla}_\nu\delta\tilde{R}_{\rho\sigma}\tilde{n}^\mu\tilde{n}^\nu\tilde{n}^\rho\tilde{n}^\sigma+2\tilde{g}^{\lambda\rho}\tilde{g}^{\sigma\tau}\tilde{\nabla}_\rho\tilde{\nabla}_\sigma\delta\tilde{R}_{\lambda\tau}+\tilde{g}^{\lambda\rho}\tilde{\nabla}_\lambda\tilde{\nabla}_\rho\delta\tilde{R}\bigg]\Bigg\}-\frac{2l_{\mathrm{P}}^2}{l}\delta\tilde{\mathcal{W}}_{\mathrm{q}}\\&+O\left(l^6\right)=\frac{4\pi l^4}{15}8\pi G\left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4}\delta\langle T_{\mu\nu}\rangle\tilde{n}^\mu\tilde{n}^\nu.\end{aligned} \tag{3.15}$$

This condition depends on an arbitrary length scale l and on a unit, timelike, future-directed vector $\tilde{n}^\mu$. The presence of either of them in the equations governing the gravitational dynamics would violate the equivalence principle and/or the locality of the dynamics. Fortunately, these terms can be consistently removed from the final equations. We now introduce a strategy for their removal, which can also be applied in the nonlinear case we discuss in the next section.

First, we show the independence of the equilibrium condition (3.15) of l by the following simple argument. We construct a sequence of M_0+1 causal diamonds whose size parameters obey $l_m = l + m\epsilon l$, with ϵ being a small dimensionless parameter and $m \in [1, M_0]$ a natural number. We pick M_0 so that $l\,(1 + M_0\epsilon)\,l$ is much smaller than the local curvature length scale. Then, for every size parameter l_m we can construct a geodesic local causal diamond centred in the same point P, with the local direction of time given by the same vector $\tilde{n}^\mu$. For each such diamond, we can derive a version of equation (3.15) with l replaced by l_m, but otherwise identical, as all the tensors are evaluated at P. Then, we obtain a set of conditions

$$\sum_{p=1}^{\infty} (1 + m\epsilon)^{2p}\, l^{2p} E^{(p)} = 0, \tag{3.16}$$

where, for simplicity of notation, we write equation (3.15) schematically as

$$\sum_{p=1}^{\infty} l^{2p} E^{(p)} = 0, \tag{3.17}$$

with all the $E^{(p)}$'s being the same for every diamond we construct (as we approximate the tensors be their value in the origin P). We require that equation (3.16) is satisfied for any $m \in [1, M_0]$. Then, we must have $E^{(p)} = 0$ for any natural number p[1]. The important information is contained in the condition $E^{(2)} = 0$, which reads

$$\begin{aligned}
&\left(\delta\tilde{R}_{\mu\nu} - \frac{1}{4}\tilde{R}\tilde{g}_{\mu\nu} + l_{\mathrm{P}}^2\lambda_{\mathrm{c}}\tilde{g}_{\mu\nu}\right)\tilde{n}^\mu\tilde{n}^\nu + \frac{\mathcal{C}l_{\mathrm{P}}^2}{56}\Bigg[4\tilde{\nabla}_\mu\tilde{\nabla}_\nu\delta\tilde{R}_{\rho\sigma}\tilde{n}^\mu\tilde{n}^\nu\tilde{n}^\rho\tilde{n}^\sigma + \tilde{g}^{\lambda\rho}\tilde{\nabla}_\lambda\tilde{\nabla}_\rho\delta\tilde{R} \\
&+ \left(2\tilde{g}^{\lambda\sigma}\tilde{\nabla}_\rho\tilde{\nabla}_\sigma\delta\tilde{R}^\rho{}_{\mu\lambda\nu} + 4\tilde{g}^{\lambda\rho}\tilde{\nabla}_\mu\tilde{\nabla}_\rho\delta\tilde{R}_{\lambda\nu} + 2\tilde{g}^{\lambda\rho}\tilde{\nabla}_\lambda\tilde{\nabla}_\rho\delta\tilde{R}_{\mu\nu} + \tilde{\nabla}_\mu\tilde{\nabla}_\nu\delta\tilde{R}\right)\tilde{n}^\mu\tilde{n}^\nu \\
&+ 2\tilde{g}^{\lambda\rho}\tilde{g}^{\sigma\tau}\tilde{\nabla}_\rho\tilde{\nabla}_\sigma\delta\tilde{R}_{\lambda\tau}\Bigg]\Bigg\} - \frac{30l_{\mathrm{P}}^2}{4\pi l^5}\delta\tilde{\mathcal{W}}_{\mathrm{q}}^{(5)} = 8\pi G\left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4}\delta\langle T_{\mu\nu}\rangle\tilde{n}^\mu\tilde{n}^\nu,
\end{aligned} \tag{3.18}$$

where $l_{\mathrm{P}}^2\delta\tilde{\mathcal{W}}_{\mathrm{q}}^{(5)}$ denotes the part of the generalised volume perturbation proportional to l^5.

At this point, we have a condition independent of l, but still including an arbitrary unit, timelike, future-directed vector $\tilde{n}^\mu$. If we can rewrite the equation in the form $f_{\mu\nu}\tilde{n}^\mu\tilde{n}^\nu = 0$ where $f_{\mu\nu}$ is any tensor independent of $\tilde{n}^\mu$, it follows that $f_{\mu\nu} = 0$ (see appendix A.4). Most of the terms in equation (3.18) have this form (or can be easily rewritten using that $\tilde{g}_{\mu\nu}\tilde{n}^\mu\tilde{n}^\nu = -1$), except for the one proportional to $\tilde{\nabla}_\mu\tilde{\nabla}_\nu\delta\tilde{R}_{\rho\sigma}\tilde{n}^\mu\tilde{n}^\nu\tilde{n}^\rho\tilde{n}^\sigma$. We also still have the remaining unspecified piece of the generalised volume perturbation $\delta\tilde{\mathcal{W}}_{\mathrm{q}}^{(5)}$ that in principle depends on $\tilde{n}^\mu$. If this piece cancels out the term containing four contractions with $\tilde{n}^\mu$, we get the desired equation of the form $f_{\mu\nu}\tilde{n}^\mu\tilde{n}^\nu = 0$. Generically, the generalised volume takes the form $\tilde{\mathcal{W}} = D_{\mu\nu\rho\sigma}\tilde{n}^\mu\tilde{n}^\sigma\tilde{h}^{\nu\rho}$ Bueno et al. [2017], Alonso-Serrano et al. [2022], Here, $D_{\mu\nu\rho\sigma}$ denotes some curvature-dependent tensor and $\tilde{h}^{\nu\rho} = \tilde{g}^{\nu\rho} + \tilde{n}^\nu\tilde{n}^\rho$ the auxiliary spatial 3-metric on Σ_0. Choosing in particular

$$\delta\tilde{\mathcal{W}}_{\mathrm{q}}^{(5)} = -\frac{\mathcal{C}\pi l^5}{105}\nabla_{(\mu}\nabla_\nu\delta\tilde{R}_{\rho\sigma)}\tilde{n}^\mu\tilde{n}^\sigma\tilde{h}^{\nu\rho}, \tag{3.19}$$

[1]The terms suppressed by sufficiently high power of l can be safely neglected, so we in fact only need to set to zero finitely many $E^{(p)}$'s.

does indeed eliminate the four contraction term. The generalised volume similarly eliminates problematic four contractions terms in the entanglement equilibrium derivation using Wald entropy (see equations (2.52) and (2.54)). Therefore, it is reasonable to expect that the same thing happens in our low-energy quantum gravitational case. Of course, lacking an explicit expression for the generalised volume, we cannot say with certainty that the term (3.19) is the only relevant contribution to $\delta\tilde{\mathcal{W}}_{\mathrm{q}}^{(5)}$. However, its generic form $D_{\mu\nu\rho\sigma}\tilde{n}^{\mu}\tilde{n}^{\sigma}\tilde{h}^{\nu\rho}$ is rather restrictive due to containing four contractions with $\tilde{n}^{\mu}$. The only other allowed contributions are of the form $\tilde{g}_{\nu\rho}\tilde{h}^{\nu\rho}D_{\mu\sigma}\tilde{n}^{\mu}\tilde{n}^{\sigma}$ or $\tilde{g}_{\mu\sigma}\tilde{n}^{\mu}\tilde{n}^{\sigma}D_{\nu\rho}\tilde{h}^{\nu\rho}$, where $D_{\mu\sigma}$ must (for dimensional reasons) contain a second derivative of the auxiliary Riemann tensor. The options read $\tilde{\nabla}_{\mu}\tilde{\nabla}_{\sigma}\delta\tilde{R}$, $\tilde{g}_{\mu\sigma}\tilde{g}^{\alpha\beta}\tilde{\nabla}_{\alpha}\tilde{\nabla}_{\beta}\delta\tilde{R}$, $\tilde{g}_{\mu\sigma}\tilde{g}^{\alpha\gamma}\tilde{g}^{\beta\lambda}\tilde{\nabla}_{\alpha}\tilde{\nabla}_{\beta}\delta\tilde{R}_{\gamma\lambda}$, $\tilde{g}^{\alpha\beta}\tilde{\nabla}_{\alpha}\tilde{\nabla}_{\beta}\delta\tilde{R}_{\mu\sigma}$, and $\tilde{g}^{\alpha\beta}\tilde{\nabla}_{\alpha}\tilde{\nabla}_{\mu}\delta\tilde{R}_{\beta\sigma}$. All these terms are already present in the equilibrium condition (3.18). The generalised volume contributions can therefore affect the numerical coefficient in front of them, but not the qualitative form of the condition, in which we are interested.

To sum up, we now have the equilibrium condition (3.18) which takes the form $f_{\mu\nu}\tilde{n}^{\mu}\tilde{n}^{\nu}=0$. Removing contractions with an arbitrary timelike vector $\tilde{n}^{\mu}$, we find

$$\delta\tilde{R}_{\mu\nu}-\frac{1}{4}\tilde{R}\tilde{g}_{\mu\nu}+l_{\mathrm{P}}^{2}\lambda_{\mathrm{c}}\tilde{g}_{\mu\nu}+\frac{\mathcal{C}l_{\mathrm{P}}^{2}}{56}\bigg[\left(-2\tilde{g}^{\lambda\rho}\tilde{g}^{\sigma\tau}\tilde{\nabla}_{\rho}\tilde{\nabla}_{\sigma}\delta\tilde{R}_{\lambda\tau}-\tilde{g}^{\lambda\rho}\tilde{\nabla}_{\lambda}\tilde{\nabla}_{\rho}\delta\tilde{R}\right)\tilde{g}_{\mu\nu}$$

$$+2\tilde{g}^{\lambda\sigma}\tilde{\nabla}_{\rho}\tilde{\nabla}_{\sigma}\delta\tilde{R}^{\rho}{}_{\mu\lambda\nu}+\frac{4}{3}\tilde{g}^{\lambda\rho}\tilde{\nabla}_{(\mu|}\tilde{\nabla}_{\rho}\delta\tilde{R}_{\lambda|\nu)}+\frac{4}{3}\tilde{g}^{\lambda\rho}\tilde{\nabla}_{\lambda}\tilde{\nabla}_{\rho}\delta\tilde{R}_{\mu\nu}+\frac{1}{3}\tilde{\nabla}_{\mu}\tilde{\nabla}_{\nu}\delta\tilde{R}\bigg]\bigg\}$$

$$=8\pi G\left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4}\delta\langle T_{\mu\nu}\rangle. \tag{3.20}$$

We may further simplify the equations using the definition of the auxiliary Weyl tensor and the contracted Bianchi identities. We aim to rewrite the $O\left(l_{\mathrm{P}}^{2}\right)$ part of the equations in terms of the auxiliary Weyl tensor, the auxiliary scalar curvature and their derivatives, in order to compare it with the standard form of the equations of motion of the quadratic gravity. After some lengthy but straightforward manipulations, we arrive at

$$\delta\tilde{R}_{\mu\nu}-\frac{1}{4}\tilde{R}\tilde{g}_{\mu\nu}+l_{\mathrm{P}}^{2}\lambda_{\mathrm{c}}\tilde{g}_{\mu\nu}+\frac{\mathcal{C}l_{\mathrm{P}}^{2}}{56\pi}\bigg(\frac{20}{3}\tilde{g}^{\lambda\sigma}\tilde{\nabla}_{\rho}\tilde{\nabla}_{\sigma}\delta\tilde{C}^{\rho}{}_{\mu\lambda\nu}+\frac{10}{9}\tilde{\nabla}_{\mu}\tilde{\nabla}_{\nu}\delta\tilde{R}$$

$$-\frac{13}{9}\tilde{g}_{\mu\nu}\tilde{g}^{\rho\sigma}\tilde{\nabla}_{\rho}\tilde{\nabla}_{\sigma}\delta\tilde{R}\bigg)=8\pi G\left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4}\delta\langle T_{\mu\nu}\rangle. \tag{3.21}$$

Finally, we determine the quantum correction to the local cosmological constant λ_{c} by taking the trace of the equations, obtaining (the energy-momentum tensor for conformal fields is traceless)

$$\lambda_{\mathrm{c}}=\frac{\mathcal{C}}{48\pi}\tilde{g}^{\rho\sigma}\tilde{\nabla}_{\rho}\tilde{\nabla}_{\sigma}\delta\tilde{R}. \tag{3.22}$$

Plugging the expression for λ_{c} back into equation (3.21), we finally arrive at

$$\delta\tilde{R}_{\mu\nu}-\frac{1}{4}\tilde{R}\tilde{g}_{\mu\nu}+\frac{5\mathcal{C}l_{\mathrm{P}}^{2}}{42\pi}\bigg(\tilde{g}^{\lambda\sigma}\tilde{\nabla}_{\rho}\tilde{\nabla}_{\sigma}\delta\tilde{C}^{\rho}{}_{\mu\lambda\nu}+\frac{1}{6}\tilde{\nabla}_{\mu}\tilde{\nabla}_{\nu}\delta\tilde{R}$$

$$-\frac{1}{24}\tilde{g}_{\mu\nu}\tilde{g}^{\rho\sigma}\tilde{\nabla}_{\rho}\tilde{\nabla}_{\sigma}\delta\tilde{R}\bigg)=8\pi G\left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4}\delta\langle T_{\mu\nu}\rangle. \tag{3.23}$$

These equations have the form of the linearised, semiclassical equations of motion of WTDiff-invariant quadratic gravity coupled to conformal quantum matter fields Salvio [2018], Donoghue and Menezes [2021]. As expected, the higher derivative corrections are suppressed by a factor $\mathcal{C}l_{\mathrm{P}}^2$, where we recall that $\mathcal{C}$ denotes the proportionality constant of the logarithmic term in entropy (3.1). We can actually write the effective WTDiff-invariant action whose linearised equations of motion are equations (3.23)

$$I_{\mathrm{QG}} = \frac{1}{16\pi G}\int\left[\tilde{R} + \frac{5\mathcal{C}l_{\mathrm{P}}^2}{168\pi}\left(-\frac{\tilde{R}^2}{3} + \tilde{g}_{\alpha\lambda}\tilde{g}^{\beta\rho}\tilde{g}^{\gamma\sigma}\tilde{g}^{\delta\tau}\tilde{C}^{\alpha}{}_{\beta\gamma\delta}\tilde{C}^{\lambda}{}_{\rho\sigma\tau}\right)\right]\omega\mathrm{d}^4x + I_{\psi}, \tag{3.24}$$

where I_{ψ} denotes the WTDiff-invariant matter action (1.21) for conformal fields. In the quadratic gravity literature, it has been argued that only the sign $\mathcal{C} < 0$ manages to avoid the tachyonic instabilities of the theory Salvio [2018], Donoghue and Menezes [2021] (although to what extent such ambiguities matter to an effective theory is a subtle issue). The negative value of $\mathcal{C}$ is consistent with the calculations of entanglement entropy of a flat spacetime sphere, that indeed yields negative logarithmic corrections Mann and Solodukhin [1998], Solodukhin [2011].

The linearised equations we found ought to be primarily understood as a proof of concept. We had no *a priori* guarantee that the local equilibrium conditions with the logarithmic correction encode meaningful gravitational dynamics. Our derivation shows that, at the linearised level, they lead to the linearised equations of motion of quadratic gravity, a reasonable result consistent with an effective field theory approach to gravity. In other words, we have shown that the local equilibrium condition allow us to obtain nontrivial insights into gravitational dynamics even with the logarithmic term present. Consequently, it is of interest to also check the nonlinear case, which we report on, in a simplified setting, in the next section.

At the same time, our findings are only qualitative. Their biggest limitation is that we lack an explicit expression for the generalised volume of the causal diamond. Moreover, the quantum expectation value of the energy-momentum tensor is only defined up to higher order curvature terms that have the same form as the quadratic gravity corrections we derived. Therefore, these corrections in principle also depend on the choice of $\langle T_{\mu\nu}\rangle$. Lastly, our approach does not easily generalise beyond the conformal matter fields due to the interplay between the local value of the cosmological constant and the von Neumann entropy of the non-conformal fields (see equation (2.36)). It should be possible to address all these shortcomings in a physical process approach that would generalise the one we introduced in subsection 2.2.1. We are going to develop this derivation in a future work.

3.2 Nonlinear analysis

Upon introducing the necessary formalism and deriving the linearised equations governing the gravitational dynamics, the natural next step lies in studying the full, nonlinear effect of the logarithmic correction to entropy. However, a number of conceptual issues arises. First and foremost, we allow for a perturbation in the size parameter of the causal diamond, determined by the condition that the perturbation of the generalised volume of the causal diamond vanishes. It then follows that we should also consider deformations of the shape of the geodesic ball Σ_0. These shape deformations do not affect the linearised analysis, but become important in the nonlinear case Jacobson et al. [2017]. However, we have no clear guidance for fixing the shape deformations in general. Moreover, the undetermined generalised volume becomes an increasingly significant problem.

To avoid some of the conceptual issues, one can turn to certain simplified, but physically interesting situations. First, the vacuum case has been studied in the literature. It has been found that the results are indeed sensitive to the shape deformations of Σ_0. The upshot is that, either for a specific choice of these deformations Jacobson et al. [2017] or for generic light-cone cut causal diamonds Wang [2019], the area perturbation is proportional to the Bell-Robinson tensor (a certain combination of terms quadratic in Weyl tensor) contracted with $\tilde{n}^{\mu}$ in all four indices. We are going to further explore this setup using our recently developed physical process approach for light-cone cut causal diamonds in an upcoming work (eventually, the same method should allow us to perform the fully general derivation). Second, we have performed a simplified analysis in which we disregard contributions from higher order terms in the Riemann normal coordinates expansion of the metric Alonso-Serrano and Liška [2020b]. Formally, this simplification can be understood as choosing the shape deformations in such a way that they cancel out these higher order contributions. Physically, these extra contributions all depend on the Weyl tensor and on the derivatives of the scalar curvature. Hence, they play no role for Weyl flat spacetime filled with conformal matter fields, such as for a homogeneous, isotropic, radiation-dominated universe. In such cases, the simplified equations we obtain already capture the relevant features of the dynamics. Moreover, they may serve to build qualitative intuition for the quantum phenomenological gravitational dynamics even in more complicated situations, such as the Oppenheimer-Snyder gravitational collapse model or homogeneous, anisotropic spacetimes. In the following, we introduce this simplified derivation using the entanglement equilibrium approach we have applied in the linearised setting. While we largely follow our original derivation Alonso-Serrano and Liška [2020b], we modify it to reflect the improvements made in our later works on the subject Alonso-Serrano and Liška [2023a]. In particular, we consider a perturbation at fixed generalised volume $\tilde{\mathcal{W}}$ rather than at fixed geometric WTDiff-invariant volume of the geodesic ball Σ_0.

We start with a geodesic local causal diamond in an equilibrium state. For simplicity, we again assume that only conformal matter fields are present, but the derivation can be generalised to non-conformal ones Alonso-Serrano and Liška [2020b]. Then, the local cosmological constant corresponding to the equilibrium state obeys equation (3.2), i.e.,

$$\lambda\left(P,l\right)=\frac{1}{4}\tilde{R}+l_{\mathrm{P}}^2\lambda_{\mathrm{c}}\left(P,l\right).$$

A small perturbation of the geometry and the matter fields with $\delta\lambda=\delta\tilde{\mathcal{W}}=0$ then obeys the equilibrium condition $\delta S_{\mathrm{q}}+\delta S_{\mathrm{vN}}=0$. The perturbation of the matter von Neumann entropy δS_{vN} is given by equation (3.4). To evaluate the perturbation of the entropy of the horizon, δS_{q}, we perturb equation (3.1), obtaining

$$\begin{aligned}\delta S_{\mathrm{q}} =&\frac{\delta\tilde{\mathcal{A}}}{4l_{\mathrm{P}}^2}+\mathcal{C}\left(\ln\frac{\tilde{\mathcal{A}}_\lambda+\delta\tilde{\mathcal{A}}}{\mathcal{A}_0}-\ln\frac{\tilde{\mathcal{A}}_\lambda}{\mathcal{A}_0}\right)\\ =&\frac{\delta\tilde{\mathcal{A}}}{4l_{\mathrm{P}}^2}+\mathcal{C}\frac{\delta\tilde{\mathcal{A}}}{\tilde{\mathcal{A}}_\lambda}-\mathcal{C}\frac{\left(\delta\tilde{\mathcal{A}}\right)^2}{2\left(\tilde{\mathcal{A}}_\lambda\right)^2}+O\left(\frac{\left(\delta\tilde{\mathcal{A}}\right)^3}{\left(\tilde{\mathcal{A}}_\lambda\right)^3}\right),\end{aligned} \tag{3.25}$$

where $\tilde{\mathcal{A}}_\lambda$ denotes the equilibrium area of the horizon spatial cross-section $\mathcal{B}$ at $t=0$ and $\delta\tilde{\mathcal{A}}$ its perturbation. We compute $\delta\tilde{\mathcal{A}}$ via the Riemann normal coordinate expansion of the metric (3.7), where we disregard all the $O\left(x^4\right)$ contributions. Then, we find

$$\delta\tilde{\mathcal{A}}=-\frac{4\pi l^4}{9}\left(\delta\tilde{S}_{\mu\nu}+l_{\mathrm{P}}^2\lambda_{\mathrm{c}}\tilde{g}_{\mu\nu}\right)\tilde{n}^\mu\tilde{n}^\nu+8\pi l\delta l+O\left(l^6\right), \tag{3.26}$$

where we introduced the notation $\delta\tilde{S}_{\mu\nu}=\delta\tilde{R}_{\mu\nu}-\delta\tilde{R}\tilde{g}_{\mu\nu}/4$ for the traceless part of the Ricci tensor to simplify the upcoming expressions. As in the linearised case, we determine δl by demanding that the generalised volume perturbation

$$\begin{aligned}\delta\tilde{\mathcal{W}}=&-\frac{4\pi l^5}{45}\left(\delta\tilde{S}_{\mu\nu}-\frac{1}{4}\delta\tilde{R}\tilde{g}_{\mu\nu}+l_{\mathrm{P}}^2\lambda_{\mathrm{c}}\tilde{g}_{\mu\nu}\right)\tilde{n}^\mu\tilde{n}^\nu+l_{\mathrm{P}}^2\delta\tilde{\mathcal{W}}_{\mathrm{q}}\\&+4\pi l^2\delta l+O\left(l^7\right),\end{aligned} \tag{3.27}$$

vanishes. The area perturbation at fixed generalised volume, $\delta\tilde{\mathcal{W}}=0$, reads

$$\delta\tilde{\mathcal{A}}=-\frac{4\pi l^4}{15}\left(\delta\tilde{R}_{\mu\nu}-\frac{1}{4}\delta\tilde{R}\tilde{g}_{\mu\nu}+l_{\mathrm{P}}^2\lambda_{\mathrm{c}}\tilde{g}_{\mu\nu}\right)\tilde{n}^\mu\tilde{n}^\nu-\frac{2l_{\mathrm{P}}^2}{l}\delta\tilde{\mathcal{W}}_{\mathrm{q}}+O\left(l^6\right).$$

Now, we plug this expression for the area perturbation into equation (3.25), computing the change in entropy of the horizon. At this point, we have everything we need to expand the equilibrium condition $\delta S_{\mathrm{q}}+\delta S_{\mathrm{vN}}=0$. The result contains several terms proportional to various powers of the diamond's size parameter l. Following the same argument as in the linearised case (see equation (3.16) and the accompanying discussion), we find that the contribution proportional to l^4 must vanish separately, i.e.,

$$\begin{aligned}&\left(\delta\tilde{S}_{\mu\nu}+l_{\mathrm{P}}^2\lambda_{\mathrm{c}}\tilde{g}_{\mu\nu}\right)\tilde{n}^\mu\tilde{n}^\nu+\frac{30l_{\mathrm{P}}^2}{4\pi l^5}\delta\tilde{\mathcal{W}}_{\mathrm{q}}^{(5)}+\frac{\mathcal{C}l_{\mathrm{P}}^2}{36\pi}\delta\tilde{R}\delta\tilde{S}_{\mu\nu}\tilde{n}^\mu\tilde{n}^\nu\\&+\frac{\mathcal{C}l_{\mathrm{P}}^2}{30\pi}\delta\tilde{S}_{\mu\nu}\delta\tilde{S}_{\rho\sigma}\tilde{n}^\mu\tilde{n}^\nu\tilde{n}^\rho\tilde{n}^\sigma=8\pi G\left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4}\delta\langle T_{\mu\nu}\rangle\tilde{n}^\mu\tilde{n}^\nu.\end{aligned} \tag{3.28}$$

As in the linearised case, we need to remove the term containing four contractions with $\tilde{n}^\mu$. Since, the generalised volume generically reads $D_{\mu\nu\rho\sigma}\tilde{n}^\mu\tilde{n}^\sigma\tilde{h}^{\nu\rho}$, where $D_{\mu\nu\rho\sigma}$ is some WTDiff-invariant tensor, we can use it to remove the problematic term $\mathcal{C}l_{\mathrm{P}}^2\delta\tilde{S}_{\mu\nu}\delta\tilde{S}_{\rho\sigma}\tilde{n}^\mu\tilde{n}^\nu\tilde{n}^\rho\tilde{n}^\sigma/30\pi$. Moreover, it is easy to realise that, perturbatively, the contribution $\mathcal{C}l_{\mathrm{P}}^2\delta\tilde{R}\delta\tilde{S}_{\mu\nu}\tilde{n}^\mu\tilde{n}^\nu/36\pi$ depends on the vacuum energy. This behaviour breaks one of the most important features of Weyl transverse gravity, which we know to be the correct semiclassical theory implied by thermodynamics of spacetime. Fortunately, this term can also be removed by choosing a suitable prescription for the generalised volume perturbation. In total, we need to set it to

$$\delta\tilde{\mathcal{W}}_{\mathrm{q}}^{(5)} = -\frac{\mathcal{C}l^5}{45}\left(\frac{1}{5}\delta\tilde{S}_{(\mu\nu}\delta\tilde{S}_{\rho\sigma)} + \frac{1}{18}\delta\tilde{R}\tilde{g}_{\mu\nu}\delta\tilde{S}_{\mu\nu}\right)\tilde{n}^\mu\tilde{n}^\nu\tilde{h}^{\rho\sigma}. \tag{3.29}$$

Then, equation (3.28) has the form $f_{\mu\nu}\tilde{n}^\mu\tilde{n}^\nu = 0$, valid for an arbitrary, unit, timelike, future-directed vector $\tilde{n}^\mu$. Therefore, it must hold $f_{\mu\nu} = 0$, i.e.,

$$\delta\tilde{S}_{\mu\nu} + l_{\mathrm{P}}^2\lambda_{\mathrm{c}}\tilde{g}_{\mu\nu} - \frac{\mathcal{C}l_{\mathrm{P}}^2}{30\pi}\tilde{g}^{\rho\sigma}\delta\tilde{S}_{\mu\rho}\delta\tilde{S}_{\nu\sigma} = 8\pi G\left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4}\delta\langle T_{\mu\nu}\rangle. \tag{3.30}$$

Lastly, we determine the quantum correction to the local cosmological constant λ_{c} by taking a trace of the equations

$$\lambda_{\mathrm{c}} = \frac{\mathcal{C}}{120\pi}\left(\tilde{g}^{\lambda\tau}\tilde{g}^{\rho\sigma}\delta\tilde{R}_{\lambda\rho}\delta\tilde{R}_{\tau\sigma} - \frac{1}{4}\delta\tilde{R}^2\right). \tag{3.31}$$

The final traceless equations governing the gravitational dynamics read

$$\begin{aligned}\delta\tilde{S}_{\mu\nu} - \frac{\mathcal{C}l_{\mathrm{P}}^2}{30\pi}\left(\tilde{g}^{\rho\sigma}\delta\tilde{S}_{\mu\rho}\delta\tilde{S}_{\nu\sigma} - \frac{1}{4}\tilde{g}^{\lambda\tau}\tilde{g}^{\rho\sigma}\delta\tilde{R}_{\lambda\rho}\delta\tilde{R}_{\tau\sigma}\tilde{g}_{\mu\nu} + \frac{1}{16}\delta\tilde{R}^2\tilde{g}_{\mu\nu}\right)\\ = 8\pi G\left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4}\delta\langle T_{\mu\nu}\rangle.\end{aligned} \tag{3.32}$$

The derivation works along the same lines even for matter fields that are not conformally invariant Alonso-Serrano and Liška [2020b], although it becomes somewhat more involved. In this way, we obtain a more general form of the equations valid for any matter fields. Moreover, to simplify the notation, we do not write the right hand side as a quantum expectation value (although it should be implicitly understood in this way), we remove the perturbation symbol δ in front of the curvature tensors (as the equations apply to the regime of very high curvatures), and introduce a new numerical coefficient $D = \mathcal{C}/30\pi$. From now on, we thus work with the following form of the equations for quantum phenomenological gravitational dynamics

$$\begin{aligned}\tilde{S}_{\mu\nu} - Dl_{\mathrm{P}}^2\left(\tilde{g}^{\rho\sigma}\tilde{S}_{\mu\rho}\tilde{S}_{\nu\sigma} - \frac{1}{4}\tilde{g}^{\lambda\tau}\tilde{g}^{\rho\sigma}\tilde{R}_{\lambda\rho}\tilde{R}_{\tau\sigma}\tilde{g}_{\mu\nu} + \frac{1}{16}\tilde{R}^2\tilde{g}_{\mu\nu}\right)\\ = 8\pi G\left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4}\left(T_{\mu\nu} - \frac{1}{4}Tg_{\mu\nu}\right).\end{aligned} \tag{3.33}$$

We briefly address the interpretation of these equation. The quantum correction terms are quadratic in the spacetime curvature, they contain at most second derivatives of the metric and are suppressed by a factor Dl_{P}^2. Since $D = \mathcal{C}/30\pi$, where $\mathcal{C}$ denotes the proportionality factor in the logarithmic correction to entropy (3.1) expected to be of the order of unity, the suppression factor Dl_{P}^2 is rather small. Hence, the correction terms likely becomes important only in the regimes of very strong gravity. As expected, in the limit $D \to 0$ (or, equivalently, $l_{\mathrm{P}} \to 0$) the equations reduce to the semiclassical equations of motion of Weyl transverse gravity. Notably, since the corrections are determined by the Ricci tensor, the equations imply no perturbative corrections to vacuum gravitational dynamics. Vacuum solutions distinct from those of Weyl transverse gravity exist, but have no meaningful semiclassical limit.

The equations do not imply a divergenceless energy-momentum tensor, nor even the weaker condition (1.25) valid in WTDiff-invariant gravity. Instead, it holds

$$\tilde{\nabla}_\lambda \left(8\pi G \left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k+1)/4} \tilde{g}^{\lambda\nu} T_{\mu\nu}\right) = -\left(2\pi G \left(\sqrt{-\mathfrak{g}}/\omega\right)^{(k-1)/4} T\right) + \frac{1}{4}\tilde{\nabla}_\mu \tilde{R}$$
$$- D l_{\mathrm{P}}^2 \left[\tilde{\nabla}_\lambda \left(\tilde{g}^{\lambda\nu}\tilde{g}^{\rho\sigma}\tilde{S}_{\mu\rho}\tilde{S}_{\nu\sigma}\right) + \tilde{\nabla}_\mu \left(-\frac{1}{4}\tilde{g}^{\lambda\tau}\tilde{g}^{\rho\sigma}\tilde{R}_{\lambda\rho}\tilde{R}_{\tau\sigma} + \frac{1}{16}\tilde{R}^2\right)\right]. \tag{3.34}$$

Therefore, equations (3.33) cannot correspond to a local, WTDiff-invariant action. The question of how to interpret them remains open. It is possible that the quantum corrections we introduce do break the WTDiff invariance. Alternatively, we may notice that the logarithmic term in entropy is non-local, in the sense that we cannot write a local expression for the entropy density (in contrast to Wald entropy). It has even been proposed that entropy with the logarithmic term corresponds to a certain non-local gravitational theory Xiao and Tian [2022]. The non-local nature of the logarithmic term may affect the equations we derived, implying that they cannot be obtained from a local action. Another possible interpretation is offered by the similarity of equations (3.33) with the so called 4D Einstein-Gauss-Bonnet gravity Glavan and Lin [2020]. While this theory apparently suffers from a number of inconsistencies Gurses et al. [2020], Ai [2020], Shu [2020], Arrechea et al. [2021], a healthy scalar-tensor version of it has been proposed Lu and Pang [2020], Hennigar et al. [2020]. Then, it is possible that a correct interpretation of our equations (3.33) also requires an identification of an additional gravitational degree of freedom, e.g. a scalar field.

The key advantage of our program of deriving the quantum phenomenological gravitational dynamics from thermodynamics of causal diamonds lies in the universality of the logarithmic correction to entropy. As we have pointed out at the beginning of this chapter, the logarithmic term represents a nearly universal prediction of any theory of quantum gravity. Moreover, phenomenological, model-independent approaches such as the generalised uncertainty principle also support it. Therefore, the implications of the logarithmic term for low energy quantum gravitational dynamics are in principle relevant regardless of the final theory of quantum gravity, or of whether gravity is ultimately emergent or can be canonically quantised. While equations (3.34) represent only the first step of this program, they already lead to interesting physical predictions, as we briefly discuss in the next chapter.

4. Physical implications of the quantum phenomenological gravitational dynamics

In this chapter, following Alonso-Serrano et al. [2023b], we study the physics implied by the quantum phenomenological gravitational dynamics (specifically, by equations (3.33)). We focus on the case of homogeneous, isotropic cosmological models. This case has the advantage of being simultaneously rather simple and physically important, as it describes the dynamics of the early universe. A number of methods for studying the quantum gravitational effects in the early universe exist. A notable approach of this kind is loop quantum cosmology Bojowald [2001], Ashtekar et al. [2007, 2006], Martin-Benito et al. [2009, 2010], Wilson-Ewing [2010], Agullo and Singh [2016], Bojowald [2020]. It relies on loop quantum gravity quantisation techniques applied to a symmetry reduced sector of the theory (minisuperspace). The most significant prediction of loop quantum cosmology is the replacement of the initial Big Bang singularity by a regular bounce, at which the universe stops contracting and starts expanding again.

Compared to loop quantum cosmology, our approach lacks a direct connection with a full candidate theory of quantum gravity and describes the spacetime as a smooth, differentiable manifold rather than considering its quantum gravitational nature. Nevertheless, the equations are not derived in a symmetry reduced setting, allowing one more flexibility in applying them to various spacetimes. Furthermore, the generality of the logarithmic correction to entropy allows us to employ our approach regardless of whether the correct description of quantum gravity is provided by loop quantum gravity or some other theory.

As we stressed in the previous chapter, equations (3.33) do not capture all the features of the quantum phenomenological gravitational dynamics implied by the logarithmic correction to entropy. That being said, they contain the term quadratic in the traceless Ricci tensor, which does not correspond to a local, WTDiff-invariant action. It appears very likely that the full equations retain this term. Therefore, it is of interest to study the physical implications of equations (3.33) as they already mark a departure from the standard effective field theory approaches to gravity. Of course, the results we obtain should only be understood as qualitative, capturing the key implications of the quantum phenomenological gravitational dynamics, but not providing precise quantitative predictions.

The general ansatz for the auxiliary metric $\tilde{g}_{\mu\nu}$ of a homogeneous, isotropic cosmology reads

$$ds^2 = -dt^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right), \tag{4.1}$$

where $a\left(t\right)$ denotes the scale factor and k the spatial curvature of the universe. We consider a universe filled with a perfect fluid of density ρ and pressure p.

Plugging this ansatz into equations (3.33) yields one non-trivial condition

$$\dot{H} - \frac{k}{a^2} - Dl_{\mathrm{P}}^2 \left(\dot{H} - \frac{k}{a^2} \right)^2 = -4\pi G \left(\rho + p \right), \tag{4.2}$$

where we introduced the Hubble parameter $H = \dot{a}/a$ and the overdot denotes a time derivative. Equation (4.2) represents a direct generalisation of the classical Raychaudhuri equation for a cosmological spacetime. As it is characteristic for WTDiff-invariant gravity, we only get one independent equation, while a Diff-invariant theory would yield two of them. The second equation comes from the divergence of the energy-momentum tensor. WTDiff-invariant gravity does in principle allow non-vanishing divergence of the energy-momentum tensor and the modified equations (3.33) are even more permissive in this regard, see equation (3.34) and the accompanying discussion. However, we presently have no physical motivation to introduce a non-vanishing divergence of the energy-momentum tensor of the perfect fluid we consider. Therefore, for the purposes of this chapter, we assume that its energy-momentum tensor is divergenceless, implying

$$\dot{\rho} + 3H \left(\rho + p \right) = 0. \tag{4.3}$$

4.1 Vacuum solutions

To get some intuition for the behaviour of the solutions, we first analyse the vacuum, spatially flat case described by the equation

$$\dot{H} - Dl_{\mathrm{P}}^2 \dot{H}^2 = 0. \tag{4.4}$$

This quadratic equation for $\dot{H}$ has two roots

$$\dot{H}_1 = 0, \tag{4.5}$$

$$\dot{H}_2 = \frac{1}{Dl_{\mathrm{P}}^2}. \tag{4.6}$$

The general solutions for the scale factor read

$$a_1 \left(t \right) = \exp \left[\pm \sqrt{L} \left(t + t_0 \right) \right], \tag{4.7}$$

$$a_2 \left(t \right) = \exp \left[\pm \frac{1}{2Dl_{\mathrm{P}}^2} \left(t + t_0 \right)^2 - \frac{Dl_{\mathrm{P}}^2}{2} L \right], \tag{4.8}$$

where t_0 and L are arbitrary integration constants. The first solution $a_1(t)$ simply corresponds to a maximally symmetric spacetime with a cosmological constant $|\Lambda| = 3L$. As expected in WTDiff-invariant gravity, Λ arises as an arbitrary integration constant defined only on shell. This branch of the solutions contains no quantum corrections. The second solution $a_2(t)$ does not have a well defined classical limit $D \to 0$. While such non-perturbative solutions can be of interest, they become relevant only in a regime in which equation (4.2) likely no longer applies, as further $O\left(l_{\mathrm{P}}^4\right)$ corrections to it become relevant.

The lesson we can take from this simple example is that equation (4.2) generically has two solutions, but only one of them has a well defined classical limit. In the following, we focus only on these solutions capturing quantum corrections to the standard cosmology that become relevant only in the early times.

4.2 Perturbative cosmological solutions

In this section, we look for perturbative solutions of the modified Raychaudhuri equation (4.2). The idea is to treat the correction term suppressed by $Dl_{\rm P}^2$ as a small correction to the classical dynamics. Then, we can expand the Hubble parameter H around its classical value $H_{(0)}$

$$H = H_{(0)} + H_{(1)} H_{(0)}^2 l_{\rm P}^2 + O\left(H_{(0)}^4 l_{\rm P}^4\right). \tag{4.9}$$

The $O\left(l_{\rm P}^4\right)$ terms were systematically discarded already in our derivation of the quantum phenomenological gravitational equations (3.33). Therefore, it is consistent to discard them here as well. However, we keep the leading order quantum correction $H_{(1)}$. Within this perturbative approach, the correction term in equation (4.2) can be approximated using the standard Raychaudhuri equation, i.e.,

$$Dl_{\rm P}^2 \dot{H}^2 = \frac{Dl_{\rm P}^2}{c^2} \dot{H}_{(0)}^2 + O\left(l_{\rm P}^4\right). \tag{4.10}$$

The resulting perturbative modified Raychaudhuri equation reads

$$\dot{H} = -4\pi G\left(\rho + p\right) + 16\pi^2 G^2 Dl_{\rm P}^2 \left(\rho + p\right)^2 + \frac{k}{a^2}. \tag{4.11}$$

We apply it to a universe filled with a perfect fluid consisting of two components, dust and radiation. This choice represents the usual description of the standard matter content (without dark matter and dark energy) of the universe. The corresponding equations of state read $p = \omega\rho$, where we have $\omega_{\rm m} = 0$ for dust and $\omega_{\rm r} = 1/3$ for radiation. For a general multi-component perfect fluid with the equations of state $p_i = \omega_i \rho_i$, the modified Raychaudhuri equation (4.11) becomes

$$\begin{aligned}\dot{H} =& -4\pi G \sum_i \rho_i \left(\omega_i + 1\right) + \frac{k}{a^2} \\ &+ 16\pi^2 G^2 Dl_{\rm P}^2 \left(\sum_i \left(\omega_i + 1\right)^2 \rho_i^2 + 2\sum_{i>j} \left(\omega_i\omega_j + \omega_i + \omega_j + 1\right)\rho_i\rho_j\right),\end{aligned} \tag{4.12}$$

where, in our two-component case, the indices i, j take values 1 for dust and 2 for radiation. We further assume that matter and radiation interact only negligibly and each component thus obeys the local energy conservation condition (4.3) separately.

Combining the perturbative modified Raychaudhuri equation (4.12) with the conservation condition (4.3) allows us to obtain the perturbative Friedmann equation. We simply multiply equation (4.12) by $2H$ and integrate, obtaining

$$\begin{aligned}H^2 =& \sum_i \frac{8\pi G}{3}\rho_i - \frac{k}{a^2} - \sum_i \frac{16\pi^2 G^2}{3} Dl_{\rm P}^2 \left(\omega_i + 1\right)\rho_i^2 \\ &- \frac{32\pi^2 G^2}{3} Dl_{\rm P}^2 \sum_{i>j} \frac{\left(\omega_i\omega_j + \omega_i + \omega_j + 1\right)}{\omega_i + \omega_j + 2}\rho_i\rho_j + \tilde{\Lambda},\end{aligned} \tag{4.13}$$

where the arbitrary integration constant $\tilde{\Lambda}$ corresponds to the cosmological constant, $\tilde{\Lambda} = \Lambda/3$. Note that, as we discussed in chapter 1, the appearance of the cosmological constant as an arbitrary integration constant is characteristic for WTDiff-invariant gravitational theories.

For our case of the two component fluid, we have the following perturbative modified Friedmann equation

$$H^2 = \frac{8\pi G}{3}\left(\rho_m + \rho_r\right) - \frac{k}{a^2} - \frac{16\pi^2 G^2 D l_{\mathrm{P}}^2}{3}\left(\frac{4}{3}\rho_r^2 + \rho_m^2 + \frac{7}{5}\rho_m\rho_r\right) + \frac{\Lambda}{3}, \tag{4.14}$$

where $\rho_{\mathrm{m}} = \rho_{\mathrm{m},0}/a^3$ and $\rho_{\mathrm{r}} = \rho_{\mathrm{r},0}/a^4$, with $\rho_{\mathrm{m},0}$, $\rho_{\mathrm{r},0}$ being constants with dimensions of energy density. Taking the limit $D \to 0$ of course recovers the classical Friedmann equation.

To further study equation (4.14), it is convenient to introduce new dimensionless variables. We define $\Omega_i = \left(8\pi G/3H_0^2\right)\rho_{i,0}$ for dust ($i = \mathrm{m}$) and radiation ($i = \mathrm{r}$), $\Omega_{\mathrm{k}} = -k/\left(a_0^2 H_0^2\right)$ for the spatial curvature and $\Omega_\Lambda = \Lambda/\left(3H_0^2\right)$ for the cosmological constant. The constants a_0, H_0, and $\rho_{i,0}$ denote some reference values of a, H, and ρ_i at a given time, often taken to correspond to the present day universe. In terms of these variables, the modified Friedmann equation (4.14) reads

$$\Omega_{\mathrm{r}} + \Omega_{\mathrm{m}} + \Omega_{\mathrm{k}} + \Omega_\Lambda - \frac{D l_{\mathrm{P}}^2}{12} H_0^2 \left(\frac{4}{3}\Omega_{\mathrm{r}}^2 + \Omega_{\mathrm{m}}^2 + \frac{7}{5}\Omega_{\mathrm{r}}\Omega_{\mathrm{m}}\right) = 1. \tag{4.15}$$

This equation conveniently showcases the quantum corrections to the relation between the dimensionless quantities Ω_{r}, Ω_{m}, Ω_{k} and Ω_Λ. Note that as the correction term depends on a length scale l_{P}, it introduces the dimensionful Hubble parameter H_0 into equation (4.15), whose classical version contains only dimensionless quantities. Of course, given the suppression factor $D l_{\mathrm{P}}^2 H_0^2/12$, the correction term only becomes relevant in the early universe. The present day values of the parameters are $\Omega_{\mathrm{r}} = 2.47 \times 10^{-5} h^{-2}$ with $h = 0.704 \pm 0.025$, $\Omega_{\mathrm{m}} = 0.3111 \pm 0.0056$, $\Omega_\Lambda = 0.6889 \pm 0.0056$, $\Omega_{\mathrm{k}} = 0.0007 \pm 0.0019$ and $H_0 = 67.66 \pm 0.42$ km s^{-1}Mpc^{-1} Aghanim et al. [2020], Workman et al. [2022]. They remain unaffected by the modifications we introduced well within the precision with which they have been measured.

For a spatially flat universe $\Omega_{\mathrm{k}} = 0$, vanishing cosmological constant $\Omega_\Lambda = 0$, and for a single component perfect fluid (either dust or radiation), equation (4.15) can be analytically solved. For dust, we obtain

$$a_{\mathrm{m}}(t) = a_0 \left(\frac{9}{4}\Omega_{\mathrm{m}} H_0^2 t^2 + \frac{D l_{\mathrm{P}}^2}{12}\Omega_{\mathrm{m}} H_0^2\right)^{1/3}, \tag{4.16}$$

where the limit $D \to 0$ recovers the classical result $a_{\mathrm{m}} \propto t^{2/3}$. For radiation, we find[1]

$$a_{\mathrm{r}}(t) = a_0 \left(4\Omega_{\mathrm{r}} H_0^2 t^2 + \frac{D l_{\mathrm{P}}^2}{9}\Omega_{\mathrm{r}} H_0^2\right)^{1/4}, \tag{4.17}$$

[1] Notably, the radiation dominated universe represents an example of a Weyl flat spacetime sourced by conformally invariant matter, for which equations (3.33) should, on the perturbative level, apply without any further correction terms. The reason is that the higher derivative terms discarded in the derivation of equations (3.33) vanish up to $O\left(l_{\mathrm{P}}^4\right)$. Therefore, we can consider equation (4.17) to be a prediction of the fully general quantum phenomenological gravitational dynamics, although we are yet to derive the precise form of the corresponding equations.

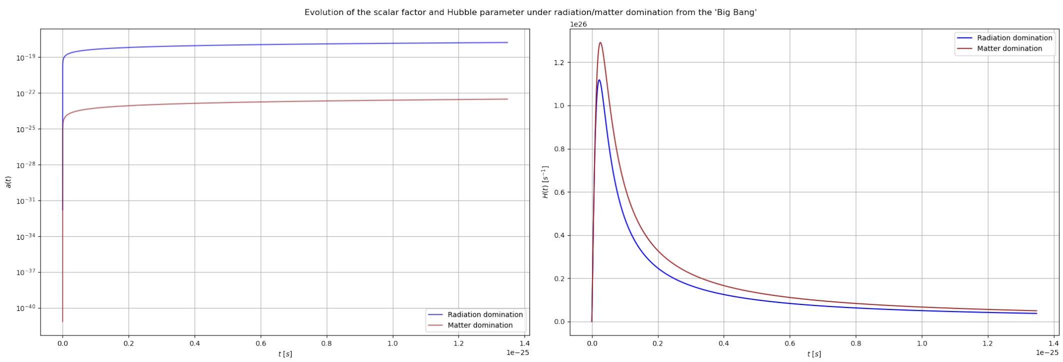

Figure 4.1: On the left: Analytical evolution of the scale factor corresponding to the analytical solutions for the matter (4.16) and radiation (4.17) domination. On the right: The corresponding evolution of the Hubble parameter (4.18). The plot is done for $D = 1$ and $a_0 = 1$.

where the limit $D \to 0$ again yields the classical solution $a_\mathrm{r} \propto t^{1/2}$. Solution (4.17) is relevant in the early universe, as it is characterised both by the radiation dominance (given the faster fall-off of the radiation energy density) and by the importance of the quantum corrections.

It is worth remarking that, for $D > 0$, both solutions (4.16) and (4.17) have no singularity in the limit $t \to 0$. Instead, the scale factor reaches some minimum value proportional to Dl_P^2 and dependent on the values of H_0 and Ω_i. The corresponding Hubble parameters read

$$H_\mathrm{m} = \frac{2}{3}\frac{t}{t^2 + \frac{Dl_\mathrm{P}^2}{27}} \quad \text{and} \quad H_\mathrm{r} = \frac{t}{2t^2 + \frac{Dl_\mathrm{P}^2}{18}}. \tag{4.18}$$

At $t = 0$, the Hubble parameter equals zero in both cases. Moreover, its derivative is positive at $t = 0$, allowing in principle a smooth bounce, i.e., a smooth gluing of the current expanding universe to a past contracting universe for $t < 0$. Nevertheless, the Big Bang curvature singularity is avoided. Interestingly, the early time evolution of the scale factor shown in figure 4.1 is very rapid, giving a short period of an inflation-like behaviour. The fastest change of the scale factor occurs at time $t_\mathrm{max} = \sqrt{Dl_\mathrm{P}^2/36}$ corresponding to the Hubble parameter $H_\mathrm{max} = \sqrt{9/4Dl_\mathrm{P}^2}$.

For $D \leq 0$, the Big Bang singularity does occur. In fact, the approach to it is even faster than the classical gravitational dynamics implies. Since D has the same sign as the logarithmic correction to entropy, it appears that positive quantum corrections to horizon entropy are consistent with the singularity resolution, whereas negative corrections are not. This result is qualitatively consistent with other approaches that consider the impact of the quantum corrections to entropy for cosmology Awad and Ali [2014], Salah et al. [2017], Hernández-Almada et al. [2022]. These works also report a cosmic bounce for positive quantum corrections to entropy.

Aside from the the analytical results, we have also carried out a numerical study of some more realistic cases Alonso-Serrano et al. [2023b]. First, we studied the late time behaviour of the solution of equation (4.14) with different values of the dimensionless parameters $\Omega_\mathrm{r}, \Omega_\mathrm{m}, \Omega_\Lambda, \Omega_\mathrm{k}$. We show the result in figure 4.2.

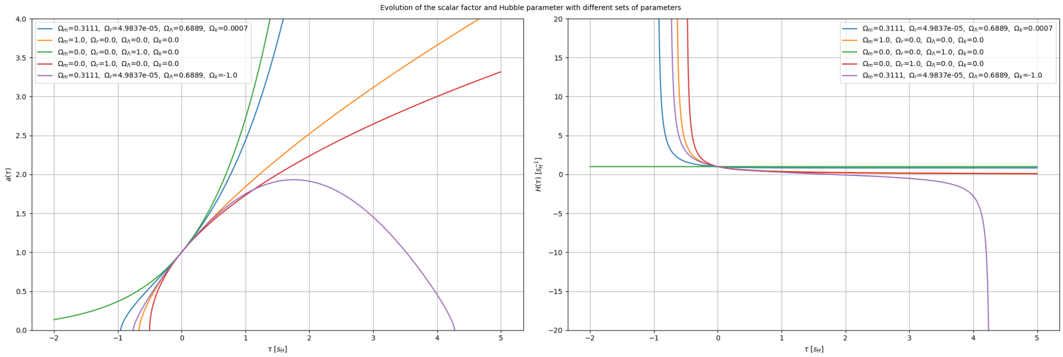

Figure 4.2: On the left: Evolution of the scale factor for different choices of the parameters in equation (4.14). In each case, we set $a_0 = 1$, $D = 1$. We choose the time parameter τ so that $\tau = 0$ corresponds to the present day universe, i.e., $a(\tau = 0) = a_0$ and so on. On the right: Evolution of the Hubble parameter for the same choices of the dimensionless parameters.

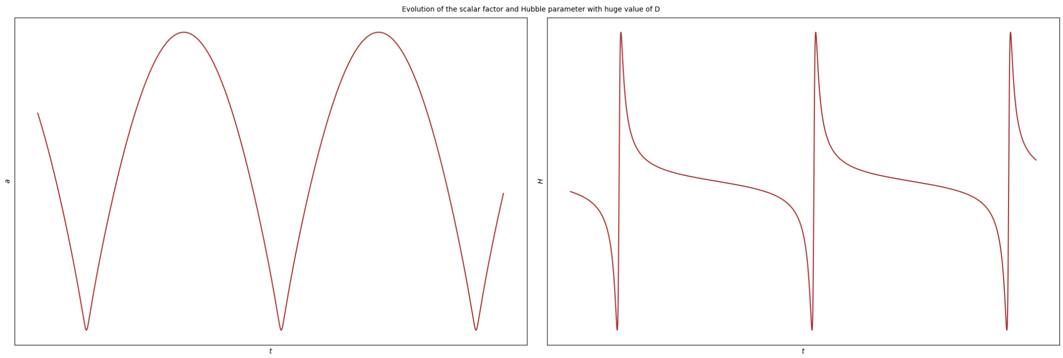

Figure 4.3: On the left: Qualitative evolution of the scale factor in a cyclic universe. On the right: The corresponding qualitative evolution of the Hubble parameter.

As expected, the late time evolution almost precisely agrees with the classical case as the quantum corrections become negligible. Second, we plotted the qualitative cyclic behaviour of the universe suggested by our analytical perturbative solution. The resulting plot 4.3 should be understood only as an illustration, since the bounces lie outside the regime of validity of our perturbative dynamics. Nevertheless, it is noteworthy that the cyclic evolution we found qualitatively agrees with that obtained in loop quantum cosmology.

Of course, our expressions for the scale factors are only consistent up to $O(l_{\mathrm{P}}^2)$, as we systematically neglected all the $O(l_{\mathrm{P}}^4)$ terms in our derivation of the perturbative modified Friedmann equation (4.14). Therefore, the (apparent) bounce lies outside the regime of applicability of equation (4.14) and we cannot consider it a genuine prediction of our model. Nevertheless, one may ask about the implications of the modified Raychaudhuri equation (4.2) if we formally treat it as an exact, non-perturbative description of the gravitational dynamics. In that case, $\dot{H}$ is given by the following equation (we choose the sign of the square root so that we reproduce the classical Raychaudhuri equation in the limit $D \to 0$)

$$\frac{2Dl_{\mathrm{P}}^2}{c^2}\left(\dot{H} - \frac{k}{a^2}\right) = 1 - \sqrt{1 + 16\pi G D l_{\mathrm{P}}^2 (\rho + p)}. \tag{4.19}$$

For $D > 0$ we have $\dot{H} < 0$ at all times and, hence, the Big Bang singularity does occur. For $D < 0$, we again have $\dot{H} < 0$, but the argument of the square root goes to zero as the matter density reaches the Planck scale. At this point, $\dot{H}$ attains a non-zero imaginary part and the further evolution of H becomes ill-defined. Since $\dot{H}$ remains negative at this point, it cannot correspond to a bounce. This breakdown strongly suggests that our equations should only be trusted perturbatively, and all the $O\left(l_{\mathrm{P}}^4\right)$ corrections need to be disregarded.

We can use equation (4.19) to estimate the limit of applicability of our perturbative approach. If we expand it up to $O\left(l_{\mathrm{P}}^4\right)$, we get

$$\dot{H} - \frac{k}{a^2} = -4\pi G\left(\rho + p\right) + 16\pi^2 G^2 D l_{\mathrm{P}}^2 \left(\rho + p\right)^2 - 128\pi^3 G^3 D^2 l_{\mathrm{P}}^4 \left(\rho + p\right)^3 + O\left(l_{\mathrm{P}}^6\right). \tag{4.20}$$

The $O\left(l_{\mathrm{P}}^4\right)$ and $O\left(l_{\mathrm{P}}^2\right)$ contributions have opposite signs. Their absolute values become equal for

$$\rho + p = \frac{1}{8\pi G D l_{\mathrm{P}}^2} = \frac{\rho_{\mathrm{P}}}{8\pi D}, \tag{4.21}$$

where $\rho_{\mathrm{P}} = 1/\left(G l_{\mathrm{P}}^2\right)$ denotes the Planck density. At this density, the perturbative modified Friedmann equation (4.14), valid up to $O\left(l_{\mathrm{P}}^4\right)$ corrections, necessarily fails. At the same time, the bounce corresponds to a density equal to

$$\rho + p = \frac{\rho_{\mathrm{P}}}{4\pi D} > \frac{\rho_{\mathrm{P}}}{8\pi D}, \tag{4.22}$$

and indeed lies out of the realm of validity of equation (4.14). To confirm that the bounce suggested by the perturbative approach really occurs, we would need to include the $O\left(l_{\mathrm{P}}^4\right)$ (and higher) terms in our analysis. Doing so would introduce additional undetermined numerical factors besides D in the quantum phenomenological gravitational equations (3.33) and the singularity resolution would be sensitive to their precise interplay, essentially destroying the predictive power of our model. Therefore, it appears to be more fruitful to instead apply equations (3.33) to phenomena lying inside their regime of validity, but already significantly affected by the quantum corrections. An example of this approach is the recently performed analysis of the primordial power spectrum Alonso-Serrano et al. [2023c]. Nevertheless, the nonperturbative dynamics implied by equation (4.2) in principle represents an interesting toy model for the early universe, and we are currently exploring its physical consequences.

Conclusions

This thesis provides an overview of the main part of the research carried out during my doctoral studies[2] which concerns the relation between gravitational dynamics and thermodynamics. We obtained novel results in four main directions of research, all intertwined by the idea that equilibrium conditions applied to locally constructed causal horizons encode the gravitational dynamics. To each of these main directions of our research is devoted one chapter of the thesis. In the following, we recall our main achievements in each of these research directions

- Covariant phase space formalism for WTDiff-invariant theories of gravity (chapter 1)
  - Inspired by the connection between thermodynamics of spacetime and Weyl transverse gravity we found; we developed an extension of the covariant phase space formalism, including the symplectic structure, to arbitrary local, WTDiff-invariant gravitational theories. To the best of our knowledge, this represents the first ever generalisation of the formalism to a class of theories with non-dynamical structures.
  - We employed this formalism to derive the first law of black hole mechanics for WTDiff-invariant gravity and (heuristically) identified the corresponding Wald entropy. We also obtained the first law for a stationary, asymptotically flat black hole spacetime with a perfect fluid, originally discussed (for general relativity) already in the seminal work on the laws of black hole mechanics Bardeen et al. [1973].
  - We showed that the WTDiff-invariant first law of thermodynamics generically includes a varying cosmological constant (both for de Sitter and anti-de Sitter asymptotics). In the asymptotically anti-de Sitter case we obtained results equivalent to those found in the context of the so-called black hole chemistry Kubiznak and Mann [2015], which also introduces a varying cosmological constant (either *ad hoc*, or by considering some specific dynamical origin of Λ).
  - We further derived the first law of thermodynamics for local causal diamonds and identified the corresponding Wald entropy and the generalised volume.
  - As an aside, we provided a novel analysis of the various formulations of the equivalence principle in WTDiff-invariant gravity, finding that an arbitrary local, WTDiff-invariant theory incorporates the Einstein equivalence principle. We also showed that Weyl transverse gravity and general relativity are the only known metric theories of gravity compatible with the equivalence principle for self-gravitating test particles.

[2] I have also co-authored three more papers not discussed here, one concerning a heuristic approach to horizon thermodynamics and two discussing thermodynamics of scalar-tensor Einstein-Gauss-Bonnet gravity in four spacetime dimensions. I decided against including these topics as it would have make the text far less cohesive.

- All the results we obtained for Weyl transverse gravity were found to be physically equivalent to those valid for general relativity. The only difference is the status of Λ as a global degree of freedom in Weyl transverse gravity. This outcome is fully consistent with previous analysis of the (non)equivalence of both theories Carballo-Rubio et al. [2022].

- WTDiff-invariant gravity from thermodynamics of spacetime (chapter 2)
  - We presented a complete, self-contained argument which shows that the local equilibrium conditions together with the strong equivalence principle imply (semi)classical gravitational dynamics equivalent to Weyl transverse gravity.
  - We also showed that the expressions for Wald entropy and generalised volume of causal diamonds we derived in the previous chapter encode the gravitational equations of motion for any local, purely metric, WTDiff-invariant theory of gravity, whose Lagrangian does not depend on the derivatives of the Riemann tensor. The same result was previously obtained for Diff-invariant theories. However, in that case the procedure fails to reconstruct the value of the cosmological constant fixed in the Lagrangian. By contrast, in the WTDiff-invariant setting, the local equilibrium conditions reproduce all the information present in the gravitational Lagrangian, since Λ appears as an arbitrary integration constant.
  - The arguments presented in this chapter complete the (semi)classical branch of the thermodynamics of spacetime program outlined already in the author's master thesis Liška [2020].

- Quantum phenomenological gravitational dynamics from thermodynamics of spacetime (chapter 3)
  - In the master thesis Liška [2020] we also opened the exploration of the impact of the quantum logarithmic correction to horizon entropy on the gravitational dynamics derived from thermodynamics. In this way, we were able to obtain equations 3.33 which incorporate quantum gravitational corrections suppressed by the factor l_{P}^2.
  - In the present thesis, we reviewed the derivation of equations[3] (3.33) in the entanglement equilibrium approach. The equations are derived under certain simplifying assumptions and do not capture the full quantum phenomenological gravitational dynamics implied by the local equilibrium conditions. Nevertheless, they include a term quadratic in the traceless Ricci tensor, which cannot be obtained from any local, WTDiff-invariant (or Diff-invariant) and purely metric action. In this way, our thermodynamic approach marks a departure from the usual effective field theory results. The full understanding of this departure remains an open issue.

[3] To be precise, the exact form of equations 3.33 only appeared in our subsequent paper Alonso-Serrano and Liška [2020b], the version derived in the master thesis was very similar, but suffered from a subtle error which we only identified afterwards Liška [2020].

    - Compared to the master thesis Liška [2020], we carried out a new linearised analysis of the quantum phenomenological gravitational dynamics. We found a result equivalent to the linearised traceless equations of motion of quadratic gravity. This derivation serves as a consistency check for our approach and also shows that the complete equations definitely include correction terms quartic in the derivatives of the metric.

- Physical implications of the quantum phenomenological gravitational dynamics from thermodynamics (chapter 4)

    - We applied equations (3.33) to a homogeneous, isotropic cosmological spacetime. The analytical solutions we obtained are perturbatively equivalent to the effective dynamics of the loop quantum cosmology. Our perturbative results only significantly deviate from the classical cosmology in the very early universe. In particular, they suggest the replacement of the Big Bang singularity by a regular bounce, although our approach unfortunately cannot be consistently extended all the way to the bounce.

The results we obtained also open a number of new paths for research which we expect to follow in the near future. We thus conclude by summarising these future perspectives and their significance

- In regards to WTDiff-invariant covariant phase space formalism, we expect to address the following open questions:

    - Rephrase the formalism in terms of the gauge-invariant generalisation of Lie derivatives. While this rephrasing represents only a technical issue, it offers a different perspective and might lead to new important insights.
    - In principle the symplectic structures of Weyl transverse gravity and general relativity with a dynamical source for the cosmological constant (realised, e.g. as an n-form) should be equivalent on shell. It would be interesting to see whether that is the indeed the case.
    - The von Neumann algebras of symmetries of arbitrary null surfaces have been recently extensively studied in the literature, yielding important insights into the nature of the corresponding entanglement entropy Jensen et al. [2023]. Apparently, the Diff invariance plays a key role in this approach. We plan to reproduce these results in the WTDiff-invariant setting using our covariant phase space formalism. Our aim is to look for possible physical differences between the Diff-invariant and the WTDiff-invariant cases.

- We plan to further develop the quantum phenomenological gravitational dynamics:

    - We are going to include the logarithmic corrections to entropy in the physical process derivation we recently developed (see Alonso-Serrano et al. [2024] and chapter 2).

- As the first part of this program, we analyse the vacuum case. Previous exploration of this case suggests the presence of the correction terms quadratic in the Weyl tensor in the equations governing the gravitational dynamics Jacobson et al. [2017], Wang [2019]. However, the previous approaches were unable to derive the modified equations in this context. Using our approach, we should be able to address this case fairly straightforwardly. The result is going to tell us whether the quantum phenomenological gravitational dynamics predict any modifications to the classical gravitational dynamics in vacuum, or whether nontrivial corrections only appear in the presence of matter fields.
- We plan to reproduce the linearised analysis leading to quadratic gravity as well as the derivation of simplified nonlinearised equations (3.33) in the physical process approach. While we do not expect any significant new insights, the new derivation should be more streamlined, with fewer *ad hoc* steps.
- The above listed phases of the project are essentially aimed at developing a full control over the new approach. After they are concluded, we are going to derive the quantum phenomenological gravitational dynamics to the $O\left(l_{\mathrm{P}}^2\right)$ without any simplifying assumptions. The outcome should give us new (and at the moment unforeseeable) insights into the interpretation of the quantum phenomenological gravitational dynamics and their relation to other approaches to modifying gravity.

• Lastly, we plan to study the implications of the quantum phenomenological gravitational dynamics for certain more complicated physically relevant spacetimes, in which novel aspects of the modified dynamics might manifest:

- We continue to explore the homogeneous isotropic setting, with the aim to see whether the full, nonperturbative dynamics have any physically interesting features. Of course, since equations (3.33) only work up to the $O\left(l_{\mathrm{P}}^2\right)$ order, such nonperturbative results can only serve as a toy model. Nevertheless, they might reveal a novel scenario for the dynamics of the early universe.
- Another research group is studying the implications of the equations for neutron star physics Prasetyo et al. [2022, 2023]. The homogeneous, isotropic cosmologies are also being analysed by other research groups Alonso-Serrano et al. [2023c], de Cesare and Gubitosi [2024]. We are currently discussing a new joint project together with the latter group.
- We plan to apply equations (3.33) to spacetimes with non-vanishing Weyl tensor. Probably the most tractable such case is the Kantowski-Sachs metric (a special form of Bianchi I metric), which is relevant both as a simple homogeneous, anisotropic model for early universe and as the metric describing the interior (dynamical) region of a Schwarzschild black hole.
- More generally, we are interested in the homogeneous, anisotropic Bianchi metrics, especially Bianchi I and Bianchi II. These are again

relevant as anisotropic cosmological models. Even more importantly, according to the Belinski-Khalatnikov-Lifshitz conjecture, these metrics combined represent a good local approximation of the spacetime in the vicinity of a generic spacelike singularity (either a black hole or a cosmological one).

– Another area to apply the quantum phenomenological gravitational dynamics are black hole solutions. While our equations imply no perturbative corrections for vacuum spacetime, they should in principle lead to non-trivial corrections for electrovacuum black hole solutions (e.g. Reissner-Nordström) and, especially, for the Oppenheimer-Snyder model of a gravitational collapse.

In summary, we carried out a complex research program connecting thermodynamics of spacetime, the classical structure of gravity, and phenomenology of quantum gravity. This program ranges from understanding the (semi)classical physics involved (including a novel mathematical apparatus we developed for this purpose), through a theoretical analysis of low-energy quantum gravitational effects, to the study of quantum corrections to physically relevant spacetimes. Notably, the results we obtained offer an interesting perspective on several open questions, most notably the cosmological constant problems, the symmetries of the equations governing the gravitational dynamics (both on the semiclassical and on the quantum level), and the possible singularity resolution. Of course, it remains unclear whether and how thermodynamics of spacetime fits into our fundamental understanding of gravity, but it certainly offers an original and intellectually stimulating perspective. We hope the thesis does this subject justice.

Bibliography

R. J. Adler, P. Chen, and D. I. Santiago. The generalized uncertainty principle and black hole remnants. *Gen. Rel. Grav.*, 33:2101–2108, 2001. doi: 10.1023/A:1015281430411.

N. Aghanim et al. Planck 2018 results. vi. cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020. doi: 10.1051/0004-6361/201833910.

I. Agullo and P. Singh. Loop quantum cosmology: A brief review. 2016. doi: 10.48550/arXiv.1612.01236.

I. Agullo, B. Bonga, P. Ribes-Metidieri, D. Kranas, and S. Nadal-Gisbert. How ubiquitous is entanglement in quantum field theory? *Phys. Rev. D*, 108:085005, 2023. doi: 10.1103/PhysRevD.108.085005.

W.-Y. Ai. A note on the novel 4D Einstein–Gauss–Bonnet gravity. *Commun. Theor. Phys.*, 72:095402, 2020. doi: 10.1088/1572-9494/aba242.

D. Alonso-López, J. de Cruz Pérez, and A. L. Maroto. Unified transverse diffeomorphism invariant field theory for the dark sector. *Phys. Rev. D*, 2023. doi: 10.48550/arXiv.2311.16836.

A. Alonso-Serrano and M. Liška. New perspective on thermodynamics of spacetime: The emergence of unimodular gravity and the equivalence of entropies. *Phys. Rev. D*, 102:104056, 2020a. doi: 10.1103/PhysRevD.102.104056.

A. Alonso-Serrano and M. Liška. Quantum phenomenological gravitational dynamics: a general view from thermodynamics of spacetime. *J. High Energ. Phys.*, 2020:196, 2020b. doi: 10.1007/JHEP12(2020)196.

A. Alonso-Serrano and M. Liška. Thermodynamics of spacetime from minimal area. *Phys. Rev. D*, 104:084043, 2021. doi: 10.1103/PhysRevD.104.084043.

A. Alonso-Serrano and M. Liška. Thermodynamics of spacetime and unimodular gravity. *IJGMMP*, 19:2230002, 2022. doi: 10.1142/S0219887822300021.

A. Alonso-Serrano, L. J. Garay, and M. Liška. Noether charge formalism for Weyl invariant theories of gravity. *Phys. Rev. D*, 106:064024, 2022. doi: 10.1103/PhysRevD.106.064024.

A. Alonso-Serrano, L. J. Garay, and M. Liška. Noether charge formalism for Weyl transverse gravity. *Class. Quant. Grav.*, 40:025012, 2023a. doi: 10.1088/1361-6382/acace3.

A. Alonso-Serrano, M. Liška, and A. Vicente-Becerril. Friedmann equations and cosmic bounce in a modified cosmological scenario. *Phys. Lett. B*, 835:137827, 2023b. doi: 10.1016/j.physletb.2023.137827.

A. Alonso-Serrano, G. A. Mena Merugan, and A. Vicente-Becerril. Primordial power spectrum in modified cosmology: From thermodynamics of spacetime to loop quantum cosmology. 2023c. doi: 10.48550/arXiv.2307.06813.

A. Alonso-Serrano, L. J. Garay, and M. Liška. Emergence of Weyl invariant gravity from thermodynamics. *in preparation*, 2024.

Ana Alonso-Serrano and Marek Liška. Emergence of quadratic gravity from entanglement equilibrium. *Phys. Rev. D*, 108:084057, 2023a. doi: 10.1103/PhysRevD.108.084057.

Ana Alonso-Serrano and Marek Liška. Thermodynamics as a tool for (quantum) gravitational dynamics. *IJMPD*, 32, 2023b. doi: 10.1142/S021827182342018X.

E. Álvarez and M. Herrero-Valea. Unimodular gravity with external sources. *JCAP*, 2013:014, 2013a. doi: 10.1088/1475-7516/2013/01/014.

E. Álvarez and M. Herrero-Valea. No conformal anomaly in unimodular gravity. *Phys. Rev. D*, 87:084054, 2013b. doi: 10.1103/PhysRevD.87.084054.

E. Álvarez, D. Blas, J. Garriga, and E. Verdaguer. Transverse Fierz–Pauli symmetry. *Nucl. Phys. B.*, 756:148, 2006. doi: 10.1016/j.nuclphysb.2006.08.003.

E. Álvarez, S. González-Martín, and C. P. Martín. Note on the gauge symmetries of unimodular gravity. *Phys. Rev. D*, 93:123018, 2016. doi: 10.1103/PhysRevD.93.123018.

E. Álvarez, J. Anero, and I. Sanchez-Ruiz. Physical charges versus conformal invariance in unimodular gravity. *IJMPA*, 38:2350132, 2023. doi: 10.1142/S0217751X23501324.

P. B. Aneesh, S. J. Hoque, and A. Virmani. Conserved charges in asymptotically de Sitter spacetimes. *Class. Quant. Grav.*, 36:205008, 2019. doi: 10.1088/1361-6382/ab3be7.

P. B. Aneesh, S. Chakraborty, S. J. Hogue, and A. Virmani. First law of black hole mechanics with fermions. *Class. Quant. Grav.*, 37:205014, 2020. doi: 10.1088/1361-6382/aba5ab.

R. E. Arias, D. D. Blanco, H. Casini, and M. Huerta. Local temperatures and local terms in modular Hamiltonians. *Phys. Rev. D*, 95:065005, 2017. doi: 10.1103/PhysRevD.95.065005.

J. Arrechea, A. Delhom, and A. Jiménez-Cano. Inconsistencies in four-dimensional Einstein-Gauss-Bonnet gravity. *Chinese Phys. C*, 45:013107, 2021. doi: 10.1088/1674-1137/abc1d4.

M. Arzano. Conformal quantum mechanics of causal diamonds. *J. High Energ. Phys.*, 72, 2020. doi: 10.1007/JHEP05(2020)072.

M. Arzano. Vacuum thermal effects in flat space-time from conformal quantum mechanics. *J. High Energ. Phys.*, 2021:3, 2021. doi: 10.1007/JHEP07(2021)003.

A. Ashtekar, T. Pawlowski, and P. Singh. Quantum nature of the Big Bang: Improved dynamics. *Phys. Rev. D*, 74:084003, 2006. doi: 10.1103/PhysRevD.74.084003.

A. Ashtekar, T. Pawlowski, P. Singh, and K. Vandersloot. Loop quantum cosmology of k=1 FRW models. *Phys. Rev. D*, 75:024035, 2007. doi: 10.1103/PhysRevD.75.024035.

A. Awad and A. F. Ali. Minimal length, Friedmann equations and maximum density. *J. High Energ. Phys.*, 2014(6), 2014. doi: 10.1007/jhep06(2014)093.

V. Baccetti and M. Visser. Clausius entropy for arbitrary bifurcate null surfaces. *Class. Quant. Grav*, 31, 2014. doi: 10.1088/0264-9381/31/3/035009.

R. Ballesteros, Gómez-Fayrén, C., T. Ortín, and M. Zatti. On scalar charges and black hole thermodynamics. *J. High Energ. Phys.*, 2023:158, 2023. doi: 10.1007/JHEP05(2023)158.

J. Balsells and M. Bojowald. Adherence and violation of the equivalence principle from classical to quantum mechanics. *Phys. Rev. D*, 108:084030, 2023. doi: 10.1103/PhysRevD.108.084030.

S. Banerjee, R. K. Gupta, I. Mandal, and A. Sen. Logarithmic Corrections to N=4 and N=8 Black Hole Entropy: A One Loop Test of Quantum Gravity. *JHEP11*, 143, 2011. doi: 10.1007/JHEP11(2011)143.

T. Banks, P. Draper, and M. Karydas. Breakdown of field theory in near-horizon regions. 2024. doi: 10.48550/arXiv.2401.03572.

L. C. Barbado and M. Visser. Unruh-DeWitt detector event rate for trajectories with time-dependent acceleration. *Phys. Rev. D*, 86, 2012. doi: 10.1103/PhysRevD.86.084011.

L. C. Barbado, C. Barceló, L. J. Garay, and G. Jannes. Hawking versus Unruh effects, or the difficulty of slowly crossing a black hole horizon. *JHEP10*, 2016, 2016. doi: 10.1007/JHEP10(2016)161.

J. F. Barbero G., J. Margalef-Bentabol, V. Varo, and E. J. S. Villaseñor. On-shell equivalence of general relativity and holst theories with nonmetricity, torsion, and boundaries. *Phys. Rev. D*, 105:064066, 2022. doi: 10.1103/PhysRevD.105.064066.

C. Barceló, L. J. Garay, and R. Carballo-Rubio. Unimodular gravity and general relativity from graviton self-interactions. *Phys. Rev. D*, 89, 2014. doi: 10.1103/PhysRevD.89.124019.

C. Barceló, L. J. Garay, and R. Carballo-Rubio. Absence of cosmological constant problem in special relativistic field theory of gravity. *Annals of Physics*, 398:9, 2018. doi: 10.1016/j.aop.2018.08.016.

C. Barceló, J. Eguia Sánchez, and G. García-Moreno. Chronology protection implementation in analogue gravity. *Eur. Phys. J. C*, 82:299, 2022. doi: 10.1140/epjc/s10052-022-10275-3.

J. D. Bardeen, B. Carter, and S. W. Hawking. Black holes and entropy. *Phys. Rev. D*, 7:2333, 1973. doi: 10.1007/BF01645742.

J.D. Barrow and D.J. Shaw. The value of the cosmological constant. *Gen. Relativ. Gravit.*, 43:2555, 2011. doi: 10.1007/s10714-011-1199-1.

J. D. Bekenstein. Black Holes and Entropy. *Phys. Rev. D*, 7:2333, 1973. doi: 10.1103/PhysRevD.7.2333.

F. Belgiorno and S. Liberati. Divergence problem in the black hole brick-wall model. *Phys. Rev. D*, 53:3172, 1996. doi: 10.1103/PhysRevD.53.3172.

P. Berglund, J. Bhattacharyya, and D. Mattingly. Mechanics of universal horizons. *Phys. Rev. D*, 85:124019, 2012. doi: 10.1103/PhysRevD.85.124019.

J. J. Bisognano and E. H. Wichmann. On the duality condition for quantum fields. *J. Math. Phys.*, 17:303, 1976. doi: 10.1063/1.522898.

M. Bojowald. Absence of singularity in loop quantum cosmology. *Phys. Rev. Lett.*, 86:5227, 2001. doi: 10.1103/PhysRevLett.86.5227.

M. Bojowald. Critical evaluation of common claims in loop quantum cosmology. *Universe*, 6:36, 2020. doi: 10.3390/universe6030036.

L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin. Quantum source of entropy for black holes. *Phys. Rev. D*, 34:373, 1986. doi: 10.1103/PhysRevD.34.373.

Harry W. Braden, J. David Brown, Bernard F. Whiting, and James W. York Jr. Charged black hole in a grand canonical ensemble. *Phys. Rev. D*, 42:3376–3385, 1990. doi: 10.1103/PhysRevD.42.3376.

L. Brewin. Riemann normal coordinate expansions using Cadabra. *Class. Quant. Grav.*, 26, 2009. doi: 10.1088/0264-9381/26/17/175017.

J. D. Brown. Action functionals for relativistic perfect fluids. *Class. Quant. Grav.*, 10:1579, 1993. doi: 10.1088/0264-9381/10/8/017.

P. Bueno, V. S. Min, A. J. Speranza, and M. R. Visser. Entanglement equilibrium for higher order gravity. *Phys. Rev. D*, 95:046003, 2017. doi: 10.1103/PhysRevD.95.046003.

R. Bufalo, M. Oksanen, and A. Tureanu. How unimodular gravity theories differ from general relativity at quantum level. *Eur. Phys. J. C*, 75:477, 2015. doi: 10.1140/epjc/s10052-015-3683-3.

C. P. Burgess. The cosmological constant problem: Why it's hard to get dark energy from micro-physics. 2013. doi: 10.48550/arXiv.1309.4133.

R. Carballo-Rubio. Longitudinal diffeomorphisms obstruct the protection of vacuum energy. *Phys. Rev. D*, 91:124071, 2015. doi: 10.1103/PhysRevD.91.124071.

R. Carballo-Rubio, L. J. Garay, and G. García-Moreno. Unimodular gravity vs general relativity: A status report. *Class. Quant. Grav.*, 39:243001, 2022. doi: 10.1088/1361-6382/aca386.

S. Carlip. Black hole entropy from conformal field theory in any dimension. *Phys. Rev. Lett.*, 82:2828–2831, 1999. doi: 10.1103/PhysRevLett.82.2828.

S. M. Carroll and G. N. Remmen. What is the entropy in entropic gravity? *Phys. Rev. D*, 93:124052, 2016. doi: 10.1103/PhysRevD.93.124052.

H. Casini, D. A. Galante, and R. C. Myers. Comments on Jacobson's "Entanglement equilibrium and the Einstein equation". *J. High Energ. Phys.*, page 194, 2016. doi: 10.1007/JHEP03(2016)194.

S. Chakraborty, C. R. Ordóñez, and G. Valdivia-Mera. Path integral derivation of the thermofield double state in causal diamonds. doi: 10.48550/arXiv.2312.03541.

S. Chakraborty, S. Bhattacharya, and T. Padmanabhan. Entropy of a generic null surface from its associated Virasoro algebra. *Phys. Lett. B*, 763:347, 2016. doi: 10.1016/j.physletb.2016.10.059.

R. Chan, M. F. A. da Silva, and V. H. Satheeshkumar. Thermodynamics of Einstein-aether black holes. *Eur. Phys. J. C*, 82:943, 2022. doi: 10.1140/epjc/s10052-022-10912-x.

G. Chirco and S. Liberati. Non-equilibrium Thermodynamics of Spacetime: the Role of Gravitational Dissipation. *Phys. Rev. D*, 81, 2010. doi: 10.1103/PhysRevD.81.024016.

G. Chirco, H. M. Haggard, A. Riello, and C. Rovelli. Spacetime thermodynamics without hidden degrees of freedom. *Phys. Rev. D*, 90, 2014. doi: 10.1103/PhysRevD.90.044044.

G. Compère and A Fiorucci. Advanced lectures on general relativity. 2018. doi: 10.1007/978-3-030-04260-8.

A. Davidson. From Planck area to graph theory: Topologically distinct black hole microstates. *Phys. Rev. D*, 100:081502, 2019. doi: 10.1103/PhysRevD.100.081502.

G.P. de Brito, O. Melichev, R. Percacci, and A. D. Pereira. Can quantum fluctuations differentiate between standard and unimodular gravity? *J. High Energ. Phys.*, 2021:90, 2021. doi: 10.1007/JHEP12(2021)090.

M. de Cesare and G. Gubitosi. Cosmological evolution from modified bekenstein entropy law. *J. Cosmol. Astropart. Phys.*, 2024:046, 2024. doi: 10.1088/1475-7516/2024/01/046.

J.-G. Demers, R. Lafrance, and R. C. Myers. Black hole entropy without brick walls. *Phys. Rev. D*, 52:2245, 1995. doi: 10.1103/PhysRevD.52.2245.

S. Deser. Self-interaction and gauge invariance. *Gen. Relat. Gravit.*, 1:9, 1970. doi: 10.1007/BF00759198.

E. di Casola, S. Liberati, and S. Sonego. Weak equivalence principle for self-gravitating bodies: A sieve for purely metric theories of gravity. *Phys. Rev. D*, 89, 2014. doi: 10.1103/PhysRevD.89.084053.

E. di Casola, S. Liberati, and S. Sonego. Nonequivalence of equivalence principles. *Am. J. Phys.*, 20:39, 2015. doi: 10.1119/1.4895342.

H.-F. Ding and X.-H. Zhai. Entropies and the first laws of black hole thermodynamics in Einstein-aether-Maxwell theory. *Class. Quant. Grav.*, 37:185015, 2020. doi: 10.1088/1361-6382/aba31d.

X. Dong. Holographic entanglement entropy for general higher derivative gravity. *J. High Energ. Phys.*, 2014:44, 2014. doi: 10.1007/JHEP01(2014)044.

J. F. Donoghue. Cosmological constant and the use of cutoffs. *Phys. Rev. D*, 104: 045005, 2021. doi: 10.1103/PhysRevD.104.045005.

J. F. Donoghue and G. Menezes. On quadratic gravity. 2021. doi: 10.48550/arXiv.2112.01974.

A. Einstein. On the influence of gravitation on the propagation of light. *Ann. Phys.*, 35:898, 1911.

Z. Elgood, P. Meessen, and T. Ortín. The first law of black hole mechanics in the Einstein-Maxwell theory revisited. *J. High Energ. Phys.*, 2020:026, 2020. doi: 10.1007/JHEP09(2020)026.

C. Eling, R. Guedens, and T. Jacobson. Non-equilibrium thermodynamics of spacetime. *Phys. Rev. Lett.*, 96, 2006. doi: PhysRevLett.96.121301.

T. Faulkner, A. Lewkowycz, and J. Maldacena. Quantum corrections to holographic entanglement entropy. *J. High Energ. Phys.*, 2013:74, 2013. doi: 10.1007/JHEP11(2013)074.

T. Faulkner, M. Guica, T. Hartman, R. C. Myers, and M. Van Raamsdonk. Gravitation from entanglement in holographic CFTs. *J. High Energ. Phys.*, 2014:051, 2014. doi: 10.1007/JHEP03(2014)051.

T. Faulkner, F. M. Haehl, E. Hijano, O. Parrikar, C. Rabideau, and M. Van Raamsdonk. Nonlinear gravity from entanglement in conformal field theories. *J. High Energ. Phys.*, 2017:057, 2017. doi: 10.1007/JHEP08(2017)057.

J. Fernando Barbero G., J. Margalef-Bentabol, V. Varo, and E. J. S. Villaseñor. Covariant phase space for gravity with boundaries: Metric versus tetrad formulations. *Phys. Rev. D*, 104:044048, 2021. doi: 10.1103/PhysRevD.104.044048.

C. J. Fewster, B. A. Juárez-Aubrey, and L. Jorma. Waiting for Unruh. *Class. Quant. Grav.*, 33:165003, 2016. doi: 10.1088/0264-9381/33/16/165003.

D. R. Finkelstein, A. A. Galiautdinov, and J. E. Baugh. Unimodular relativity and cosmological constant. *J. Math. Phys.*, 42:340, 2001. doi: 10.1063/1.1328077.

B. Fiol and J. Garriga. Semiclassical Unimodular Gravity. *JCAP*, 015, 2010. doi: 10.1088/1475-7516/2010/08/015.

B. Z. Foster. Noether charges and black hole mechanics in Einstein-aether theory. *Phys. Rev. D*, 73:024005, 2006. doi: 10.1103/PhysRevD.73.024005.

S. A. Fulling. Nonuniqueness of canonical field quantization in Riemannian space-time. *Phys. Rev. D*, 7:2850, 1973. doi: 10.1103/PhysRevD.7.2850.

L. J. Garay. Quantum gravity and minimum length. *Int. J. Mod. Phys. A*, 10: 145, 1995. doi: 10.1142/S0217751X95000085.

L. J. Garay and G. García-Moreno. Embedding unimodular gravity in string theory. *J. High Energ. Phys.*, 2023:27, 2023. doi: 10.1007/JHEP03(2023)027.

G. García-Moreno and A. Jiménez Cano. On the existence of a parent theory for general relativity and unimodular gravity. doi: 10.48550/arXiv.2309.06903.

R. Geroch. Colloquium at Princeton University. 1971.

R. Geroch and P. S. Jang. Motion of a body in general relativity. *J. Math. Phys.*, 16:65, 1975. doi: 10.1063/1.522416.

F. Giacomini and Č. Brukner. Einstein's equivalence principle for superpositions of gravitational fields and quantum reference frames. doi: 10.48550/arXiv.2012.13754.

G. W. Gibbons and S. W. Hawking. Action integrals and partition functions in quantum gravity. *Phys. Rev. D*, 15:2752–2756, 1977. doi: 10.1103/PhysRevD.15.2752.

D. Glavan and C. Lin. Einstein-Gauss-Bonnet gravity in four-dimensional space-time. *Phys. Rev. Lett.*, 124:081301, 2020. doi: 10.1103/PhysRevLett.124.081301.

G. Gour and A. J. M. Medved. Thermal fluctuations and black-hole entropy. *Class. Quant. Grav.*, 20105:3307, 2003. doi: 10.1088/0264-9381/20/15/303.

R. Guedens, T. Jacobson, and S. Sarkar. Horizon entropy and higher curvature equations of state. *Phys. Rev. D*, 85, 2012. doi: 10.1103/PhysRevD.85.064017.

M. Gurses, T. C. Sisman, and B. Tekin. Is there a novel Einstein-Gauss-Bonnet theory in four dimensions? *Eur. Phys. J. C*, 80:647, 2020. doi: 10.1140/epjc/s10052-020-8200-7.

S. W. Hawking. Particle creation by black holes. *Commun. Math. Phys.*, 43:199, 1975. doi: 10.1007/BF02345020.

S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge University Press, Cambridge, United Kingdom, 1973. doi: 10.1017/CBO9780511524646.

S. W. Hawking and G. T. Horowitz. The gravitational Hamiltonian, action, entropy and surface terms. *Class. Quant. Grav.*, 13:1487, 1996. doi: 10.1088/0264-9381/13/6/017.

S. W. Hawking and D. N. Page. Thermodynamics of black holes in anti-de Sitter space. *Commun. Math. Phys.*, 87:577, 1996. doi: 10.1007/BF01208266.

M. Henneaux and C. Teitelboim. The Cosmological Constant and General Covariance. *Phys. Lett. B*, 222:195, 1989. doi: 10.1016/0370-2693(89)91251-3.

R.A. Hennigar, D. Kubizňák, R.B. Mann, and C. Pollack. On taking the D $\to$ 4 limit of Gauss-Bonnet gravity: theory and solutions. *J. High Energ. Phys.*, 2020:27, 2020. doi: doi.org/10.1007/JHEP07(2020)027.

A. Hernández-Almada, G. Leon, J. Magaña, M. A. García-Aspeitia, V. Motta, E. N. Saridakis, K. Yesmakhanova, and A. D. Millano. Observational constraints and dynamical analysis of Kaniadakis horizon-entropy cosmology. *MNRAS*, 512:5122, 2022. doi: 10.1093/mnras/stac795.

F.-H. Ho, S.-J. Zhang, H. Liu, and A. Wang. Smarr integral formula of D-dimensional stationary spacetimes in Einstein-æther–Maxwell theory. *Phys. Lett. B*, 782:723, 2018. doi: 10.1103/PhysRevD.85.124019.

S. Hod. High-order corrections to the entropy and area of quantum black holes. *Class. Quant. Grav.*, 21, 2004. doi: 10.1088/0264-9381/21/14/L01.

S. Hollands, R. M. Wald, and V. G. Zhang. The entropy of dynamical black holes. 2024. doi: 10.48550/arXiv.2402.00818.

S. Hossenfelder. Minimal length scale scenarios for quantum gravity. *Living Rev. Relativ.*, 16:2, 2013. doi: 10.12942/lrr-2013-2.

S. Hossenfelder. Screams for explanation: finetuning and naturalness in the foundations of physics. *Synthese*, 198:3727, 2021. doi: 10.1007/s11229-019-02377-5.

V. Iyer. Lagrangian perfect fluids and black hole mechanics. *Phys. Rev. D*, 55: 3411–3426, 1997. doi: 10.1103/PhysRevD.55.3411.

V. Iyer and R. M. Wald. Some properties of Noether charge and a proposal for dynamical black hole entropy. *Phys. Rev. D*, 50:846–864, 1994. doi: 10.1103/PhysRevD.50.846.

R. Jackiw and S.-Y. Pi. Fake conformal symmetry in conformal cosmological models. *Phys. Rev. D*, 91:067501, 2015. doi: 10.1103/PhysRevD.91.067501.

T. Jacobson. Black hole entropy and induced gravity. 1994. doi: 10.48550/arXiv.gr-qc/9404039.

T. Jacobson. Thermodynamics of space-time: The Einstein equation of state. *Phys. Rev. D*, 75:1260, 1995. doi: 10.1103/PhysRevLett.75.1260.

T. Jacobson. Entanglement Equilibrium and the Einstein equation. *Phys. Rev. Lett.*, 116:1260, 2015. doi: 10.1103/PhysRevLett.116.201101.

T. Jacobson and A. Mohd. Black hole entropy and Lorentz-diffeomorphism Noether charge. *Phys. Rev. D*, 92:124010, 2015. doi: 10.1103/PhysRevD.92.124010.

T. Jacobson and R. Parentani. Horizon entropy. *Found. Phys.*, 33:323, 2003. doi: 10.1023/A:1023785123428.

T. Jacobson and M. R. Visser. Gravitational thermodynamics of causal diamonds in (A)dS. *SciPost Phys.*, 7, 2019a. doi: 10.21468/SciPostPhys.7.6.079.

T. Jacobson and M. R. Visser. Spacetime equilibrium at negative temperature and the attraction of gravity. *IJMPD*, 28:194016, 2019b. doi: 10.1142/S0218271819440164.

T. Jacobson and M. R. Visser. Partition function for a volume of space. *Phys. Rev. Lett.*, 130:221501, 2023a. doi: 10.1103/PhysRevLett.130.221501.

T. Jacobson and M. R. Visser. Entropy of causal diamond ensembles. *SciPost Phys.*, 15:023, 2023b. doi: 10.21468/SciPostPhys.15.1.023.

T. Jacobson, G. Kang, and R. C. Myers. On black hole entropy. *Phys. Rev. D*, 49:6587–6598, 1994. doi: 10.1103/PhysRevD.49.6587.

T. Jacobson, J. M. M. Senovilla, and A. J. Speranza. Area deficits and the Bel-Robinson tensor. *Class. Quant. Grav.*, 35, 2017. doi: 10.1088/1361-6382/aab06e.

D. Jaramillo-Garrido, A. L. Maroto, and P. Martín-Moruno. TDiff in the dark: Gravity with a scalar field invariant under transverse diffeomorphisms. 2024. doi: 10.48550/arXiv.2307.14861.

K. Jensen, J. Sorce, and A.J. Speranza. Generalized entropy for general subregions in quantum gravity. *J. High Energ. Phys.*, 2023:20, 2023. doi: 10.1007/JHEP12(2023)020.

T. Josset, A. Perez, and D. Sudarsky. Dark Energy from Violation of Energy Conservation. *Phys. Rev. Lett.*, 118, 2017. doi: 10.1103/PhysRevLett.118.021102.

V. Kagramanova, J. Kunz, and C. Lämmerzahl. Solar system effects in Schwarzschild–de Sitter space–time. *Nucl. Phys. B*, 634:465, 2006. doi: 10.1016/j.physletb.2006.01.069.

S. Karan and G.S. Punia. Logarithmic correction to black hole entropy in universal low-energy string theory models. *J. High Energ. Phys.*, 2023:28, 2023. doi: 10.1007/JHEP03(2023)028.

D. Kastor, S. Ray, and J. Traschen. Enthalpy and the mechanics of AdS black holes. *Class. Quant. Grav.*, 26:195011, 2009. doi: 10.1088/0264-9381/26/19/195011.

D. Kastor, S. Ray, and J. Traschen. Smarr formula and an extended first law for Lovelock gravity. *Class. Quant. Grav.*, 27:235014, 2010. doi: 10.1088/0264-9381/27/23/235014.

R. K. Kaul and P. Majumdar. Logarithmic Correction to the Bekenstein-Hawking Entropy. *Phys. Rev. D*, 97, 2000. doi: 10.1103/PhysRevLett.84.5255.

D. Kothawala. Minimal length and small scale structure of spacetime. *Phys. Rev. D*, 88:104029, 2013. doi: 10.1103/PhysRevD.88.104029.

D. Kothawala and T. Padmanabhan. Entropy density of spacetime as a relic from quantum gravity. *Phys. Rev. D*, 90:124060, 2014. doi: 10.1103/PhysRevD.90.124060.

D. Kothawala and T. Padmanabhan. Entropy density of spacetime from the zero point length. *Phys. Lett. B*, 748:67, 2015. doi: 10.1016/j.physletb.2015.06.066.

D. Kubiznak and R. B. Mann. Black hole chemistry. *Can. J Phys.*, 93:999, 2015. doi: 10.1139/cjp-2014-0465.

D. Kubizňák and R. B. Mann. Black hole chemistry. *Can. J. Phys.*, 93, 2014. doi: 10.1139/cjp-2014-0465.

D. Kubizňák, R. Mann, and M. Teo. Black hole chemistry: thermodynamics with Lambda. *Class. Quant. Grav.*, 34:063001, 2017. doi: 10.1088/1361-6382/aa5c69.

Y. Kucukakca and A.R. Akbarieh. Noether symmetries of einstein-aether scalar field cosmology. *Eur. Phys. J. C*, 80:1019, 2020. doi: 10.1140/epjc/s10052-020-08583-7.

N. Lashkari, M. B. McDermott, and M. Van Raamsdonk. Gravitational dynamics from entanglement "thermodynamics". *J. High Energ. Phys.*, 2014:195, 2014. doi: 10.1007/JHEP04(2014)195.

J. Lee and R. M. Wald. Local symmetries and constraints. *Journal of Mathematical Physics*, 31:725–743, 1990. doi: 10.1063/1.528801.

S. Liberati. Problems in black-hole entropy interpretation. *Nuov. Cim. B*, 112: 405, 1997. doi: 10.48550/arXiv.gr-qc/9601032.

M. Liška. Thermodynamics of spacetime: A new perspective from the quantum realm. 2020.

H. Lu and Y. Pang. Horndeski gravity as D $\to$ 4 limit of Gauss-Bonnet. *Phys. Lett. B*, 809:135717, 2020. doi: 10.1016/j.physletb.2020.135717.

R. B. Mann and S. Solodukhin. Universality of Quantum Entropy for Extreme Black Holes. *Nucl. Phys. B*, 523:293, 1998. doi: 10.1016/S0550-3213(98)00094-7.

J. Margalef-Bentabol and E. J. S. Villaseñor. Geometric formulation of the covariant phase space methods with boundaries. *Phys. Rev. D*, 103:025011, 2021. doi: 10.1103/PhysRevD.103.025011.

J. Margalef-Bentabol and E. J. S. Villaseñor. Proof of the equivalence of the symplectic forms derived from the canonical and the covariant phase space formalisms. *Phys. Rev. D*, 105:L101701, 2022. doi: 10.1103/PhysRevD.105.L101701.

M. Martin-Benito, G. A. Mena-Marugán, and J. Olmedo. Further improvements in the understanding of isotropic loop quantum cosmology. *Phys. Rev. D*, 80: 104015, 2009. doi: 10.1103/PhysRevD.80.104015.

M. Martin-Benito, G. A. Mena-Marugán, E. Wilson-Ewing, and J. Olmedo. Hybrid quantization: From Bianchi I to the Gowdy model. *Phys. Rev. D*, 82: 084012, 2010. doi: 10.1103/PhysRevD.82.084012.

P. Martinetti and C. Rovelli. Diamonds's temperature: Unruh effect for bounded trajectories and thermal time hypothesis. *Class. Quant. Grav.*, 20:4919, 2003. doi: 10.1088/0264-9381/20/22/015.

C. A. Mead. Possible connection between gravitation and fundamental length. *Phys. Rev.*, 135, 1964. doi: 10.1103/PhysRev.135.B849.

A. J. M. Medved. A follow-up to 'Does Nature abhor a logarithm?' (and apparently she doesn't). *Class. Quant. Grav.*, 22:595, 2005.

A. J. M. Medved and E. C. Vagenas. When conceptual worlds collide: The GUP and the BH entropy. *Phys. Rev. D*, 70, 2004. doi: 10.1103/PhysRevD.70.124021.

P. Meessen, D. Mitsios, and T. Ortín. Black hole chemistry, the cosmological constant and the embedding tensor. *J. High Energ. Phys.*, 2022:155, 2022. doi: 10.1007/JHEP12(2022)155.

K. A. Meissner. Black hole entropy in loop quantum gravity. *Class. Quant. Grav.*, 21:5245, 2004. doi: 10.1088/0264-9381/21/22/015.

C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. Princeton University Press, Princeton, 2017. ISBN 9780691177793.

M. Montesinos and D. Gonzalez. Diffeomorphism-invariant action principles for trace-free Einstein gravity. *Phys. Rev. D*, 108:124013, 2023. doi: 10.1103/PhysRevD.108.124013.

S. Mooij and M. Shaposhnikov. QFT without infinities and hierarchy problem. *Nucl. Phys. B*, 990:116172, 2021a. doi: 10.1016/j.nuclphysb.2023.116172.

S. Mooij and M. Shaposhnikov. Finite Callan-Symanzik renormalisation for multiple scalar fields. *Nucl. Phys. B*, 990:116176, 2021b. doi: 10.1016/j.nuclphysb.2023.116176.

E. Mottola. The effective theory of gravity and dynamical vacuum energy. *J. High Energ. Phys.*, 2022:37, 2022. doi: 10.1007/JHEP11(2022)037.

I. Oda. Classical Weyl transverse gravity. *Eur. Phys. J. C*, 77:284, 2017. doi: 10.1140/epjc/s10052-017-4843-4.

I. Oda. Vanishing Noether current in Weyl invariant gravities. *Int. J. Mod. Phys. A*, 37:2250213, 2022. doi: 10.1142/S0217751X2250213X.

T. Ortín and D. Pereñíguez. Magnetic charges and wald entropy. *J. High Energ. Phys.*, 2022:81, 2022. doi: 10.1007/JHEP11(2022)081.

J. Overduin, F. Everitt, J. Mester, and P. Worden. The science case for STEP. *Advances in Space Research*, 43:1532, 2009. doi: https://doi.org/10.1016/j.asr.2009.02.012.

C. Pacilio and S. Liberati. Improved derivation of the Smarr formula for Lorentz-breaking gravity. *Phys. Rev. D*, 95:124010, 2017. doi: 10.1103/PhysRevD.95.124010.

A. Padilla and I. D. Saltas. A note on classical and quantum unimodular gravity. *Eur. Phys. J. C*, 75:561, 2014. doi: 10.1140/epjc/s10052-015-3767-0.

T. Padmanabhan. From Gravitons to Gravity: Myths and Reality. *IJMPD*, 17:367, 2008. doi: 10.1142/S0218271808012085.

T. Padmanabhan. Thermodynamical Aspects of Gravity: New insights. *Rep. Prog. Phys.*, 73, 2010. doi: 10.1088/0034-4885/73/4/046901.

T. Padmanabhan. Geodesic distance: A descriptor of geometry and correlator of pregeometric density of spacetime events. *Mod. Phys. Lett. A*, 35:2030008, 2020. doi: 10.1142/S0217732320300086.

T. Padmanabhan and S. Chakraborty. Microscopic origin of Einstein's field equations and the raison d'être for a positive cosmological constant. *Mod. Phys. Lett. A*, 824:136828, 2022. doi: 10.1016/j.physletb.2021.136828.

M. Parikh and A. Svesko. Einstein's equations from the stretched future light cone. *Phys. Rev. D*, 98, 2018. doi: 10.1103/PhysRevD.98.026018.

A. Perez, D. Sudarsky, and J. D. Bjorken. A microscopic model for an emergent cosmological constant. *IJMPD*, 27:1846002, 2018. doi: 10.1142/S0218271818460021.

J. Polo-Gómez, L. J. Garay, and E. Martín-Martínez. A detector-based measurement theory for quantum field theory. *Phys. Rev. D*, 105:065003, 2022. doi: 10.1103/PhysRevD.105.065003.

K. Prabhu. The first law of black hole mechanics for fields with internal gauge freedom. *Class. Quant. Grav.*, 34:035011, 2017. doi: 10.1088/1361-6382/aa536b.

I. Prasetyo, I. H. Belfaqih, A. B. Wahidin, A. Suroso, and A. Sulaksono. Minimal length, nuclear matter, and neutron stars. *Eur. Phys. J. C*, 82:884, 2022. doi: 10.1140/epjc/s10052-022-10849-1.

I. Prasetyo, I. H. Belfaqih, A. Suroso, and A. Sulaksono. Anisotropic ultracompact object in Serrano–Liska gravity model. *Eur. Phys. J. C*, 83:780, 2023. doi: 10.1140/epjc/s10052-023-11954-5.

J. Preskill. Quantum Shannon theory, (course notes, quantum information). 2018. doi: 10.48550/arXiv.1604.07450.

A. Raychaudhuri. Relativistic cosmology. I. *Phys. Rev.*, 98:1123–1126, 1955. doi: 10.1103/PhysRev.98.1123.

H. Reeh and S. Schlieder. Bemerkungen zur unitäräquivalenz von lorentzinvarianten feldern. *Nuov. Cim.*, 22:1051, 1961. doi: 10.1007/BF02787889.

T. Rick Perche. General features of the thermalization of particle detectors and the unruh effect. *Phys. Rev. D*, 104:065001, 2021. doi: 10.1103/PhysRevD.104.065001.

T. Rick Perche. Localized nonrelativistic quantum systems in curved spacetimes: A general characterization of particle detector models. *Phys. Rev. D*, 106:025018, 2022. doi: 10.1103/PhysRevD.106.025018.

T. Rick Perche and E. Martín-Martínez. Geometry of spacetime from quantum measurements. *Phys. Rev. D*, 105:066011, 2022. doi: 10.1103/PhysRevD.105.066011.

N. Rosen. General relativity and flat space. I. *Phys. Rev.*, 57:147, 1940. doi: 10.1103/PhysRev.57.147.

N. Rosen. A bi-metric theory of gravitation. *Gen. Relat. Gravit.*, 4:435, 1973. doi: 10.1007/BF01215403.

S. Ryu and T. Takayanagi. Holographic derivation of entanglement entropy from the anti–de Sitter space/conformal field theory correspondence. *Phys. Rev. Lett.*, 96:181602, 2006. doi: 10.1103/PhysRevLett.96.181602.

M. Salah, F. Hammad, M. Faizal, and A. F. Ali. Non-singular and cyclic universe from the modified GUP. *J. Cosmol. Astropart. Phys.*, 2017:035, 2017. doi: 10.1088/1475-7516/2017/02/035.

A. Salvio. Quadratic gravity. *Front. in Phys.*, 6:77, 2018. doi: 10.3389/fphy.2018.00077.

S. Sarkar and A. C. Wall. Second law violations in Lovelock gravity for black hole mergers. *Phys. Rev. D*, 83:26, 2011. doi: 10.1103/PhysRevD.83.124048.

A. Sen. Logarithmic corrections to Schwarzschild and other non-extremal black hole entropy in different dimensions. *JHEP04*, 156, 2013. doi: 10.1007/JHEP04(2013)156.

C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.

V. Shevchenko. Finite time measurements by Unruh–de Witt detector and Landauer's principle. *Ann. Phys.*, 381:17, 2017. doi: 10.1016/j.aop.2017.03.014.

F.-W. Shu. Vacua in novel 4D Einstein-Gauss-Bonnet gravity: pathology and instability? *Phys. Lett. B*, 811:135907, 2020. doi: 10.1016/j.physletb.2020.135907.

S. Solodukhin. The conical singularity and quantum corrections to entropy of black hole. *Phys.Rev. D*, 51:609, 1995. doi: 10.1103/PhysRevD.51.609.

S. Solodukhin. Entanglement entropy of black holes. *Living Rev. Rel.*, 214, 2011. doi: 10.12942/lrr-2011-8.

S. N. Solodukhin. Entanglement entropy of round spheres. *Phys. Lett. B*, 693:605, 2010. doi: 10.1016/j.physletb.2010.09.018.

A. J. Speranza. Entanglement entropy of excited states in conformal perturbation theory and the Einstein equation. *J. High Energ. Phys.*, page 105, 2016. doi: 10.1007/JHEP04(2016)105.

M. Srednicki. Entropy and area. *Phys. Rev. Lett.*, 71:666, 1993.

L. Susskind and J. Uglum. Black hole entropy in canonical quantum gravity and superstring theory. *Phys. Rev. D*, 50:2700, 1994. doi: 10.1103/PhysRevD.50.2700.

A. Svesko. Equilibrium to Einstein: Entanglement, thermodynamics, and gravity. *Phys. Rev. D*, 99, 2019. doi: 10.1103/PhysRevD.99.086006.

G. 't Hooft. On the quantum structure of a black hole. *Nucl. Phys. B*, 256:727, 1985. doi: 10.1016/0550-3213(85)90418-3.

A. Tavlayan and B. Tekin. Partition function of a volume of space in a higher curvature theory. *Phys. Rev. D*, 108:L041902, 2023. doi: 10.1103/PhysRevD.108.L041902.

S. C. Tiwari. Thermodynamics of spacetime and unimodular relativity. 2006. doi: https://doi.org/10.48550/arXiv.gr-qc/0612099.

W. G. Unruh. Notes on black hole evaporation. *Phys.Rev. D*, 14:870, 1976. doi: 10.1103/PhysRevD.14.870.

W. G. Unruh. Unimodular theory of canonical quantum gravity. *Phys. Rev. D*, 40:1048, 1989. doi: 10.1103/PhysRevD.40.1048.

E. P. Verlinde. On the origin of gravity and the laws of Newton. *JHEP04*, 2011, 2011. doi: 10.1007/JHEP04(2011)029.

M. Visser. Essential and inessential features of Hawking radiation. *Int. J. Mod. Phys. D*, 12:649, 2003. doi: 10.1142/S0218271803003190.

F. Wagner, G. Varão, I. P. Lobo, and V. B. Bezerra. Quantum-spacetime effects on nonrelativistic Schrödinger evolution. *Phys. Rev. D*, 108:066008, 2023. doi: 10.1103/PhysRevD.108.066008.

R. M. Wald. On identically closed forms locally constructed from a field. *Journal of Mathematical Physics*, 31:2378–2384, 1990. doi: 10.1063/1.528839.

R. M. Wald. Black hole entropy is Noether charge. *Phys. Rev. D*, 48:3427, 1993. doi: 10.1103/PhysRevD.48.R3427.

R. M. Wald. The thermodynamics of black holes. *Living Rev. Rel.*, 6, 2001. doi: 10.12942/lrr-2001-6.

R. M. Wald and A. Zoupas. A general definition of ”conserved quantities” in general relativity and other theories of gravity. *Phys. Rev. D*, 61:084027, 2000. doi: 10.1103/PhysRevD.61.084027.

A. C. Wall. Proof of the generalized second law for rapidly changing fields and arbitrary horizon slices. *Phys. Rev. D*, 85:104049, 2012. doi: 10.1103/PhysRevD.85.104049.

A. C. Wall. A second law for higher curvature gravity. *IJMPD*, 24:1544014, 2015. doi: 10.1142/S0218271815440149.

J. Wang. Geometry of small causal diamonds. *Phys. Rev. D*, 100, 2019. doi: 10.1103/PhysRevD.100.064020.

S. Weinberg. The cosmological constant problem. *Rev. Mod. Phys.*, 61:1, 1989. doi: 10.1103/RevModPhys.61.1.

J. A. Wheeler and K. Ford. *Geons, Black Holes, and Quantum Foam*. W. W. Norton and Company, NY, U.S.A., 1998. doi: 10.1119/1.19497.

E. Wilson-Ewing. Loop quantum cosmology of Bianchi type IX models. *Phys. Rev. D*, 82:043508, 2010. doi: 10.1103/PhysRevD.82.043508.

R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022. doi: 10.1093/ptep/ptac097.

Y. Xiao and Y. Tian. Logarithmic correction to black hole entropy from the nonlocality of quantum gravity. *Phys. Rev. D*, 105:044013, 2022. doi: 10.1103/PhysRevD.105.044013.

B. Zhang. Is equivalence principle valid for quantum gravitational field? *Results Phys.*, 57:107345, 2024. doi: 10.1016/j.rinp.2024.107345.

List of publications

The thesis is based on the following publications in peer-reviewed journals

(1) Ana Alonso-Serrano and Marek Liška, “New perspective on thermodynamics of spacetime: The emergence of unimodular gravity and the equivalence of entropies”, Phys. Rev. D 102 (2020), arXiv:gr-qc/2008.04805.

(2) Ana Alonso-Serrano and Marek Liška, “Quantum phenomenological gravitational dynamics: A general view from thermodynamics of spacetime”, JHEP 2020 (2020), arXiv:2009.03826.

(3) Ana Alonso-Serrano and Marek Liška, “Thermodynamics of spacetime and unimodular gravity”, IJGMMP (2022), arXiv:2112.06301.

(4) Ana Alonso-Serrano, Luis J. Garay and Marek Liška, “Noether charge formalism for Weyl invariant theories of gravity”, Phys. Rev. D 106 (2022), arXiv:2206.08746.

(5) Ana Alonso-Serrano, Luis J. Garay and Marek Liška, “Noether charge formalism for Weyl transverse gravity”, Class. Quant. Grav. 40 (2023), arXiv:2204.08245.

(6) Ana Alonso-Serrano, Marek Liška and Antonio Vicente-Becerril, “Friedmann equations and cosmic bounce in a modified cosmological scenario, Phys. Lett. B 839 (2023), arXiv:2212.10928.

(7) Ana Alonso-Serrano, and Marek Liška, “Thermodynamics as a tool for (quantum) gravitational dynamics”, IJMPD (2023), arXiv:2305.11756, essay received Honorable Mention at the Gravity Research Foundation 2023 Awards for Essays on Gravitation.

(8) Ana Alonso-Serrano, and Marek Liška, “Emergence of quadratic gravity from entanglement equilibrium”, Phys. Rev. D 108 (2023) 084057, arXiv:2212.03168.

Other publications in peer-reviewed journals

(1) Ana Alonso-Serrano and Marek Liška, “Thermodynamics of spacetime from minimal area”, Phys. Rev. D 104 (2021), arXiv:2107.08749.

(2) Marek Liška and David Kubizňák, Shall Bekenstein’s Area Law Prevail?”, Phys. Rev. D 108 (2023), arXiv:2307.16201.

(3) Marek Liška, Robie A. Hennigar and David Kubizňák, “No logarithmic corrections to entropy in shift-symmetric Gauss-Bonnet gravity”, JHEP 2023 (2023) 195, arXiv:2309.05629.

Conference proceedings

(1) Ana Alonso-Serrano and Marek Liška, “Quantum gravity phenomenology from thermodynamics of spacetime”, Proceedings of the MG16 Meeting on General Relativity (2023).

(2) Ana Alonso-Serrano and Marek Liška, “Quantum Gravity Phenomenology from the Thermodynamics of Spacetime”, Universe 8 (2022).

A. Appendices

A.1 Gravitational weak equivalence principle

Herein, we prove the key requirement for the validity of the gravitational weak equivalence principle in Weyl transverse gravity. Expanded, the condition for the principle to hold, $\bar{\nabla}_\nu \mathcal{E}^\nu_\mu = 0$, reads

$$\begin{aligned}&\frac{1}{2}\tilde{g}^{\nu\rho}\tilde{\nabla}_\nu\Big[2\tilde{g}^{\lambda\sigma}\tilde{\nabla}_\lambda\tilde{\nabla}_{(\mu}\tilde{\gamma}_{\rho)\sigma} - \tilde{g}^{\lambda\sigma}\tilde{\nabla}_\lambda\tilde{\nabla}_\sigma\tilde{\gamma}_{\mu\rho} + \tilde{g}_{\mu\rho}\tilde{g}^{\alpha\lambda}\tilde{g}^{\beta\sigma}\left(-\tilde{\nabla}_\lambda\tilde{\nabla}_\sigma\tilde{\gamma}_{\alpha\beta} + \tilde{R}_{\lambda\sigma}\tilde{\gamma}_{\alpha\beta}\right)\\ &- \tilde{R}\tilde{\gamma}_{\mu\rho} + 2\delta\Lambda\tilde{g}_{\mu\rho} + 2\Lambda\tilde{\gamma}_{\mu\rho}\Big] = 0, \qquad \text{(A.1)}\end{aligned}$$

where $\tilde{\gamma}_{\mu\rho}$ denotes a perturbation of the auxiliary metric[1] and $\tilde{g}_{\mu\nu}$, $\tilde{\nabla}_\mu$, $\tilde{R}_{\mu\nu}$ are to be understood as background quantities. Since the value of Λ generically changes between solutions of Weyl transverse gravity, we must also consider its perturbation $\delta\Lambda$. Expressing the derivative commutators in terms of the auxiliary Riemann tensor and simplifying yields

$$\tilde{g}^{\nu\rho}\left[\tilde{g}^{\lambda\sigma}\left(2\tilde{R}_{\mu\lambda}\tilde{\nabla}_\nu\tilde{\gamma}_{\rho\sigma} + 2\tilde{\gamma}_{\rho\sigma}\tilde{\nabla}_\nu\tilde{R}_{\mu\lambda} + \tilde{R}_{\lambda\rho}\tilde{\nabla}_\mu\tilde{\gamma}_{\nu\sigma}\right) - \tilde{\nabla}_\nu\left(\tilde{R}\tilde{\gamma}_{\mu\rho}\right) + 2\Lambda\tilde{\nabla}_\nu\tilde{\gamma}_{\mu\rho}\right] = 0. \qquad \text{(A.2)}$$

Finally, we use the vacuum equations (1.20) for the background to express the Ricci tensor in terms of Λ, i.e. $\tilde{R}_{\mu\nu} = 2\Lambda\tilde{g}_{\mu\nu}/\left(n-2\right)$. Then, the left hand side of equation (A.2) indeed vanishes. Therefore, Weyl transverse gravity incorporates the gravitational weak equivalence principle.

A.2 Derivation of the symplectic potential for WTDiff-invariant gravity

Herein, we derive in detail the symplectic potential presented for an arbitrary local, WTDiff-invariant theory. We begin by varying Lagrangian (1.9) independently with respect to $g_{\mu\nu}$, $\tilde{R}^{\mu}{}_{\nu\rho\sigma}$ and ψ, obtaining,

$$\begin{aligned}\delta L =&\frac{\partial L}{\partial g_{\mu\nu}}\delta g_{\mu\nu} + \sum_{i=0}^{p}\frac{\partial L}{\partial\tilde{\nabla}_{(\alpha_1}\ldots\tilde{\nabla}_{\alpha_i)}\tilde{R}^{\mu}{}_{\nu\rho\sigma}}\delta\tilde{\nabla}_{(\alpha_1}\ldots\tilde{\nabla}_{\alpha_i)}\tilde{R}^{\mu}{}_{\nu\rho\sigma}\\ &+\sum_{i=0}^{q}\frac{\partial L}{\partial\tilde{\nabla}_{(\alpha_1}\ldots\tilde{\nabla}_{\alpha_i)}\psi}\delta\tilde{\nabla}_{(\alpha_1}\ldots\tilde{\nabla}_{\alpha_i)}\psi. \qquad \text{(A.3)}\end{aligned}$$

[1]Of course, we actually perturb the dynamical metric, $\tilde{\gamma}_{\mu\rho}$ is simply a convenient bookkeeping device.

We apply the following approach to modify the terms with variations of derivatives of the auxiliary Riemann tensor

$$\frac{\partial L}{\partial \tilde{\nabla}_{(\alpha_1} \ldots \tilde{\nabla}_{\alpha_i)} \tilde{R}^{\mu}{}_{\nu\rho\sigma}} \delta \tilde{\nabla}_{(\alpha_1} \ldots \tilde{\nabla}_{\alpha_i)} \tilde{R}^{\mu}{}_{\nu\rho\sigma}$$
$$= \frac{\partial L}{\partial \tilde{\nabla}_{(\alpha_1} \ldots \tilde{\nabla}_{\alpha_i)} \tilde{R}_{\mu\nu\rho\sigma}} \tilde{\nabla}_{\alpha_1} \delta \tilde{\nabla}_{\alpha_2} \ldots \tilde{\nabla}_{\alpha_i} \tilde{R}_{\mu\nu\rho\sigma}$$
$$+ \text{“terms proportional to } \tilde{\nabla}\delta g_{\mu\nu} \text{ and } \delta g_{\mu\nu}\text{”}$$
$$= \tilde{\nabla}_{\alpha_1} \left(\frac{\partial L}{\partial \tilde{\nabla}_{(\alpha_1} \ldots \tilde{\nabla}_{\alpha_i)} \tilde{R}^{\mu}{}_{\nu\rho\sigma}} \delta \tilde{\nabla}_{\alpha_2} \ldots \tilde{\nabla}_{\alpha_i} \tilde{R}^{\mu}{}_{\nu\rho\sigma} \right) - \tilde{\nabla}_{\alpha_1} \left(\frac{\partial L}{\partial \tilde{\nabla}_{(\alpha_1} \ldots \tilde{\nabla}_{\alpha_i)} \tilde{R}^{\mu}{}_{\nu\rho\sigma}} \right)$$
$$\delta \tilde{\nabla}_{\alpha_2} \ldots \tilde{\nabla}_{\alpha_i} \tilde{R}^{\mu}{}_{\nu\rho\sigma} + \tilde{\nabla}_{\alpha_1} \text{“terms proportional to } \delta g_{\mu\nu}\text{”}$$
$$+ \text{“terms proportional to } \delta g_{\mu\nu}\text{”} = \tilde{\nabla}_{\alpha_1} \left(\frac{\partial L}{\partial \tilde{\nabla}_{(\alpha_1} \ldots \tilde{\nabla}_{\alpha_i)} \tilde{R}^{\mu}{}_{\nu\rho\sigma}} \delta \tilde{\nabla}_{\alpha_2} \ldots \tilde{\nabla}_{\alpha_i} \tilde{R}^{\mu}{}_{\nu\rho\sigma} \right)$$
$$- \tilde{\nabla}_{\alpha_1} \left(\frac{\partial L}{\partial \tilde{\nabla}_{(\alpha_1} \ldots \tilde{\nabla}_{\alpha_i)} \tilde{R}^{\mu}{}_{\nu\rho\sigma}} \right) \delta \tilde{\nabla}_{\alpha_2} \ldots \tilde{\nabla}_{\alpha_i} \tilde{R}^{\mu}{}_{\nu\rho\sigma}$$
$$+ \tilde{\nabla}_{\alpha_1} \text{“terms proportional to } \delta g_{\mu\nu}\text{”} + \text{“terms proportional to } \delta g_{\mu\nu}\text{”}. \tag{A.4}$$

In total, we have a lower derivative term, some Weyl covariant divergences and terms proportional to $\delta g_{\mu\nu}$ and contributing to the equations of motion. An analogous procedure can be applied to terms containing variations of derivatives of the matter fields. Using this algorithm repeatedly, we find

$$\begin{aligned} \delta L = & \hat{A}^{\mu\nu} \delta g_{\mu\nu} + E_{\mu}{}^{\nu\rho\sigma} \delta \tilde{R}^{\mu}{}_{\nu\rho\sigma} + A_{\psi} \delta \psi \\ & + \tilde{\nabla}_{\alpha_1} \Bigg[\left(K^{\alpha_1\mu\nu} + 2\tilde{g}^{\alpha_1\sigma} \tilde{\nabla}_{\rho} E_{\sigma}{}^{\mu\nu\rho} \right) \delta \tilde{g}_{\mu\nu} + \sum_{i=2}^{p} M^{\alpha_1\alpha_2\ldots\alpha_i}{}_{\mu}{}^{\nu\rho\sigma} \delta \tilde{\nabla}_{(\alpha_2} \ldots \tilde{\nabla}_{\alpha_i)} \tilde{R}^{\mu}{}_{\nu\rho\sigma} \\ & + \sum_{i=2}^{q} N^{\alpha_1\alpha_2\ldots\alpha_k} \delta \tilde{\nabla}_{(\alpha_2} \ldots \tilde{\nabla}_{\alpha_i)} \psi \Bigg]. \end{aligned} \tag{A.5}$$

Tensor densities $K^{\alpha_1\mu\nu}$, $M^{\alpha_1\alpha_2\ldots\alpha_i}{}_{\mu}{}^{\nu\rho\sigma}$, and $N^{\alpha_1\alpha_2\ldots\alpha_k}$ arise in the above introduced process of rewriting the variations of derivatives of $\tilde{R}^{\mu}{}_{\nu\rho\sigma}$ and ψ. The explicit form of these tensors would be exceedingly complicated in the fully general case and we do not require it for our purposes.

If we had $\tilde{R}^{\mu}{}_{\nu\rho\sigma}$ as an independent field, $\hat{A}^{\mu\nu} = 0$ would have been the equations of motion for the metric. In fact, $\delta \tilde{R}^{\mu}{}_{\nu\rho\sigma}$ of course depends on the variation of the auxiliary metric,

$$\delta \tilde{g}_{\mu\nu} = \left(\frac{\sqrt{-g}}{\omega} \right)^{-\frac{2}{n}} \left(\delta g_{\mu\nu} - \frac{1}{n} g_{\mu\nu} \frac{\delta g}{g} \right). \tag{A.6}$$

Therefore, we have to re-express it in terms of $\delta \tilde{g}_{\mu\nu}$ and its derivatives, obtaining

$$E_{\mu}{}^{\nu\rho\sigma} \tilde{R}^{\mu}{}_{\nu\rho\sigma} = 2E_{\mu}{}^{\nu\rho\sigma} \tilde{g}^{\mu\lambda} \tilde{\nabla}_{\lambda} \tilde{\nabla}_{\sigma} \delta \tilde{g}_{\nu\rho} + E_{\mu}{}^{\nu\rho\sigma} \tilde{R}^{\mu}{}_{\nu\rho\lambda} \tilde{g}^{\lambda\tau} \delta \tilde{g}_{\sigma\tau}, \tag{A.7}$$

where we used that $E_{\mu}{}^{\nu\rho\sigma}$ is defined to have the same symmetries as the Riemann tensor. Then, the variation of the Lagrangian finally reads

$$\begin{aligned}\delta L =&\frac{1}{16\pi}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{-\frac{2}{n}}\left[\mathring{A}^{\mu\nu}+8\pi\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k+1}{n}}\left(T_{\mu\nu}-\frac{1}{n}Tg_{\mu\nu}\right)\right]\delta g_{\mu\nu}+A_\psi\delta\psi\\&+\tilde{\nabla}_{\alpha_1}\left[2\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}}E^{\mu\nu\rho\alpha_1}\tilde{\nabla}_\mu\delta\tilde{g}_{\nu\rho}+K^{\alpha_1\mu\nu}\delta\tilde{g}_{\mu\nu}\right.\\&\left.+\sum_{i=2}^{p}M^{\alpha_1\alpha_2\ldots\alpha_i}{}_{\mu}{}^{\nu\rho\sigma}\delta\hat{\nabla}_{(\alpha_2}\ldots\tilde{\nabla}_{\alpha_i)}\tilde{R}^{\mu}{}_{\nu\rho\sigma}+\sum_{i=2}^{q}N^{\alpha_1\alpha_2\ldots\alpha_i}\delta\tilde{\nabla}_{(\alpha_2}\ldots\hat{\nabla}_{\alpha_i)}\psi\right],\end{aligned}\tag{A.8}$$

and the traceless equations of motion are

$$\begin{aligned}\mathring{A}^{\mu\nu}+8\pi\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k+1}{n}}\left(T_{\mu\nu}-\frac{1}{n}Tg_{\mu\nu}\right)=&8\pi\left(\frac{\sqrt{-g}}{\omega}\right)^{-\frac{2}{n}}\left[\hat{A}^{(\mu\nu)}+2E_\iota{}^{\rho\sigma(\mu|}\tilde{R}^{\iota}{}_{\rho\sigma}{}^{|\nu)}\right.\\&-\frac{1}{n}E_\iota{}^{\rho\sigma\lambda}\tilde{R}^{\iota}{}_{\rho\sigma\lambda}g^{\mu\nu}+4\tilde{\nabla}_\sigma\tilde{\nabla}^\rho E_\rho{}^{(\mu\nu)\sigma}\\&\left.-\frac{4}{n}g_{\iota\lambda}\left(\tilde{\nabla}_\sigma\tilde{\nabla}^\rho E_\rho{}^{\iota\lambda\sigma}\right)g^{\mu\nu}\right]=0.\end{aligned}\tag{A.9}$$

A.3 Hamiltonians corresponding to general vector fields in WTDiff-gravity

In this appendix, we derive the perturbation (1.185) of the Hamiltonian corresponding to an arbitrary vector field ζ^μ for a WTDiff-Lagrangian of the form $L\left(\tilde{g}_{\mu\nu},\tilde{R}^{\mu}{}_{\nu\rho\sigma}\right)$ together with minimally coupled matter fields given by the action (1.21). The corresponding traceless gravitational equations of motion are

$$\begin{aligned}0=&\mathring{A}^{\alpha\beta}+8\pi\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k+1}{n}}\left(T_{\mu\nu}-\frac{1}{n}Tg_{\mu\nu}\right)\\=&-\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}}E^{\mu\lambda\rho\sigma}\tilde{R}^{\nu}{}_{\lambda\rho\sigma}+2\tilde{\nabla}_\rho\tilde{\nabla}_\sigma\left[\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}}E^{\mu\rho\sigma\nu}\right]+\frac{1}{n}\Bigg(E^{\lambda\rho\sigma\tau}\tilde{R}^{\lambda\rho\sigma\tau}\\&-2\tilde{\nabla}_\rho\tilde{\nabla}_\sigma E_\lambda{}^{\rho\sigma\lambda}\Bigg)\tilde{g}^{\mu\nu}+8\pi\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k+1}{n}}\left(T_{\mu\nu}-\frac{1}{n}Tg_{\mu\nu}\right),\end{aligned}\tag{A.10}$$

where $E_\mu{}^{\nu\rho\sigma}$ reads

$$E_\mu{}^{\nu\rho\sigma}=\frac{\partial L}{\partial\tilde{R}^{\mu}{}_{\nu\rho\sigma}}.\tag{A.11}$$

The divergence-free equations (1.37) then become

$$\begin{aligned}0=A^{\mu\nu}=&-\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}}E^{\mu\lambda\rho\sigma}\tilde{R}^{\nu}{}_{\lambda\rho\sigma}+2\tilde{\nabla}_\rho\tilde{\nabla}_\sigma\left[\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}}E^{\mu\rho\sigma\nu}\right]+\frac{1}{2}L\tilde{g}^{\mu\nu}\\&-\Lambda\tilde{g}^{\mu\nu}+8\pi\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2(k+1)}{n}}T^{\mu\nu}-\mathcal{J}\tilde{g}^{\mu\nu},\end{aligned}\tag{A.12}$$

where $\mathcal{J}$ is defined by equation (1.1.5). As always in WTDiff-invariant theories, Λ is an arbitrary integration constant.

We can straightforwardly work out the symplectic potential from the general expression (1.74). As the Lagrangian does not contain derivatives of the auxiliary Riemann tensor, we have $M^{\mu\alpha_2\ldots\alpha_i\ \nu\rho\sigma} = 0$, $K^{\mu\nu\rho} = -2\tilde{\nabla}_\sigma\left[\left(\sqrt{-\mathfrak{g}}/\omega\right)^{2/n} E^{\mu\nu\rho\sigma}\right]$. The matter symplectic potential reads $\theta^\mu_\psi = \left[\partial L_\psi / \partial\left(\partial_\mu\psi\right)\right]\delta\psi$. Altogether, we have

$$\theta^\mu\left[\delta\right] = 2\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} E^{\sigma\nu\rho\mu}\tilde{\nabla}_\sigma\delta\tilde{g}_{\nu\rho} - 2\tilde{\nabla}_\sigma\left[\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} E^{\mu\nu\rho\sigma}\right]\delta\tilde{g}_{\nu\rho} + \frac{\partial L_\psi}{\partial\left(\partial_\mu\psi\right)}\delta\psi. \tag{A.13}$$

For the symplectic current corresponding to a transformation generated by any vector field ζ^μ and an arbitrary perturbation of the metric we have

$$\Omega^\mu\left[\delta_\zeta,\delta\right] = \delta_\zeta\theta^\mu\left[\delta\right] - \delta\theta^\mu\left[\delta_\zeta\right]. \tag{A.14}$$

As we explained in subsection 1.3.5, δ_ζ applied to a WTDiff-invariant expression is not in general a Lie derivative. Differences between δ_ζ and $\pounds_\zeta$ acting on the building blocks of $\theta^\mu\left[\delta\right]$ yield

$$\left(\delta_\zeta - \pounds_\zeta\right)\frac{\sqrt{-\mathfrak{g}}}{\omega} = \frac{\sqrt{-\mathfrak{g}}}{\omega}\tilde{\nabla}_\mu\zeta^\mu, \tag{A.15}$$

$$\left(\delta_\zeta - \pounds_\zeta\right)\tilde{\Gamma}^\mu{}_{\nu\rho} = -\frac{1}{n}\left(\delta^\mu_\nu\delta^\alpha_\rho + \delta^\mu_\rho\delta^\alpha_\nu - g_{\nu\rho}g^{\mu\alpha}\right)\tilde{\nabla}_\alpha\tilde{\nabla}_\lambda\zeta^\lambda, \tag{A.16}$$

$$\left(\delta_\zeta - \pounds_\zeta\right)\tilde{R}^\mu{}_{\nu\rho\sigma} = \frac{2}{n}\left(\delta^\alpha_\nu\delta^\mu_{[\rho}\delta^\beta_{\sigma]} - g^{\mu\alpha}g_{\nu[\rho}\delta^\beta_{\sigma]}\right)\tilde{\nabla}_\alpha\tilde{\nabla}_\beta\tilde{\nabla}_\lambda\zeta^\lambda, \tag{A.17}$$

$$\left(\delta_\zeta - \pounds_\zeta\right)\tilde{\nabla}_\tau\tilde{R}^\mu{}_{\nu\rho\sigma} = \frac{2}{n}\left(\delta^\alpha_\nu\delta^\mu_{[\rho}\delta^\beta_{\sigma]} - g^{\mu\alpha}g_{\nu[\rho}\delta^\beta_{\sigma]}\right)\tilde{\nabla}_\tau\tilde{\nabla}_\alpha\tilde{\nabla}_\beta\tilde{\nabla}_\lambda\zeta^\lambda. \tag{A.18}$$

Thence, after some straightforward manipulations, we get a somewhat lengthy expression $\Pi^\mu\left[\zeta,\delta\right] = \left(\delta_\zeta - \pounds_\zeta\right)\theta^\mu\left[\delta\right]$,

$$\begin{aligned}\Pi^\mu\left[\zeta,\delta\right] = &-\frac{16}{n}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} F^{\sigma\nu\rho\mu}{}_\alpha{}^{\beta\alpha\gamma}\tilde{\nabla}_\sigma\tilde{\nabla}_\beta\tilde{\nabla}_\gamma\tilde{\nabla}_\lambda\zeta^\lambda\delta\tilde{g}_{\nu\rho} \\ &+\frac{16}{n}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} F^{\sigma\nu\rho\mu}{}_\alpha{}^{\beta\alpha\gamma}\tilde{\nabla}_\beta\tilde{\nabla}_\gamma\tilde{\nabla}_\lambda\zeta^\lambda\tilde{\nabla}_\sigma\delta\tilde{g}_{\nu\rho} \\ &-\frac{16}{n}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} G^{\sigma\nu\rho\mu}{}_\alpha{}^{\beta\alpha\gamma}{}_\xi{}^{\tau\vartheta\omega}\tilde{\nabla}_\sigma\tilde{R}^\xi{}_{\tau\vartheta\omega}\tilde{\nabla}_\beta\tilde{\nabla}_\gamma\tilde{\nabla}_\lambda\zeta^\lambda\delta\tilde{g}_{\nu\rho} \\ &-\frac{4}{n}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} E^{\sigma\nu\rho\mu}\tilde{\nabla}_\sigma\tilde{\nabla}_\lambda\zeta^\lambda\delta\tilde{g}_{\nu\rho} \\ &+\frac{8}{n}\left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} F^{\sigma\nu\rho\mu}{}_\alpha{}^{\beta\gamma\eta}\tilde{R}^\alpha{}_{\tau\vartheta\omega}\tilde{\nabla}_\kappa\tilde{\nabla}_\lambda\zeta^\lambda\delta\tilde{g}_{\nu\rho} \\ &\left(\delta^\kappa_\beta\delta^\tau_\sigma\delta^\vartheta_\gamma\delta^\omega_\eta + \delta^\kappa_\eta\delta^\tau_\beta\delta^\vartheta_\gamma\delta^\omega_\sigma + \delta^\kappa_\sigma\delta^\tau_\beta\delta^\vartheta_\gamma\delta^\omega_\eta - g_{\sigma\beta}g^{\kappa\tau}\delta^\vartheta_\gamma\delta^\omega_\eta - g_{\sigma\eta}g^{\kappa\omega}\delta^\tau_\beta\delta^\vartheta_\gamma\right) \\ &+\frac{2k}{n}\frac{\partial L_\psi}{\partial\left(\partial_\mu\psi\right)}\delta\psi\tilde{\nabla}_\lambda\zeta^\lambda,\end{aligned} \tag{A.19}$$

where

$$F_\mu{}^{\nu\rho\sigma}{}_\alpha{}^{\beta\gamma\eta} = \frac{\partial E_\mu{}^{\nu\rho\sigma}}{\partial \tilde{R}^\alpha{}_{\beta\gamma\eta}} = \frac{\partial^2 L}{\partial \tilde{R}^\alpha{}_{\beta\gamma\eta} \partial \tilde{R}^\mu{}_{\nu\rho\sigma}}, \tag{A.20}$$

$$G_\mu{}^{\nu\rho\sigma}{}_\alpha{}^{\beta\gamma\eta}{}_\xi{}^{\tau\vartheta\omega} = \frac{\partial^2 E_\mu{}^{\nu\rho\sigma}}{\partial \tilde{R}^\xi{}_{\tau\vartheta\omega} \partial \tilde{R}^\alpha{}_{\beta\gamma\eta}} = \frac{\partial^3 L}{\partial \tilde{R}^\xi{}_{\tau\vartheta\omega} \partial \tilde{R}^\alpha{}_{\beta\gamma\eta} \partial \tilde{R}^\mu{}_{\nu\rho\sigma}}. \tag{A.21}$$

Next, we use that $\delta_\zeta \theta^\mu [\delta] = \pounds_\zeta \theta^\mu [\delta] + \Pi^\mu [\zeta, \delta]$ and expand $\pounds_\zeta \theta^\mu [\delta]$

$$\begin{aligned} \delta_\zeta \theta^\mu [\delta] =& \pounds_\zeta \theta^\mu [\delta] + \Pi^\mu [\zeta, \delta] = \zeta^\nu \tilde{\nabla}_\nu \theta^\mu [\delta] - \theta^\nu [\delta] \tilde{\nabla}_\nu \zeta^\nu + \Pi^\mu [\zeta, \delta] \\ =& -2\tilde{\nabla}_\nu \left(\theta^{[\nu} [\delta] \zeta^{\mu]}\right) + \zeta^\mu \frac{1}{16\pi} \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} \left[\mathring{A}^{\alpha\beta} + 8\pi \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k+1}{n}} \right. \\ & \left. \left(T_{\mu\nu} - \frac{1}{n} T g_{\mu\nu}\right)\right] + \zeta^\mu A_\psi \delta\psi - \zeta^\mu \delta L - \theta^\mu [\delta] \tilde{\nabla}_\nu \zeta^\nu + \Pi^\mu [\zeta, \delta]. \end{aligned} \tag{A.22}$$

For the second term in $\Omega^\mu [\delta_\zeta, \delta]$ we directly obtain

$$\begin{aligned} -\delta\theta^\mu [\delta_\zeta] =& -\frac{1}{8\pi} \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} A_\nu{}^\mu \zeta^\nu + \zeta^\mu \delta L - \delta \left(\zeta \cdot A_\psi \cdot \psi\right)^\mu \\ & + \frac{4}{n} \delta E^{\mu\nu\rho}{}_\nu \tilde{\nabla}_\rho \tilde{\nabla}_\lambda \zeta^\lambda - \frac{4}{n} \delta \left(\tilde{\nabla}_\rho E^{\mu\nu\rho}{}_\nu\right) \tilde{\nabla}_\lambda \zeta^\lambda \\ & + \tilde{\nabla}_\nu \left[2\tilde{\nabla}_\rho \zeta^\sigma \delta E^{\nu\mu\rho}{}_\sigma - 4\zeta^\sigma \delta \left(\tilde{\nabla}_\rho E^{\nu\mu\rho}{}_\sigma\right)\right] + \tilde{\nabla}_\nu \delta Q^{\nu\mu}_{\psi,\zeta}, \end{aligned} \tag{A.23}$$

where $Q^{\nu\mu}_{\psi,\zeta}$ is an antisymmetric tensor density appearing from the matter part of $\theta^\mu [\delta_\zeta]$. It can be shown that $Q^{\nu\mu}_{\psi,\zeta}$ depends on ζ^μ (see subsection 1.3.4). If ζ^μ generates a transverse diffeomorphism, the last line becomes the Weyl covariant divergence of the corresponding Noether charge.

In total, the symplectic current $\Omega^\mu [\delta_\zeta, \delta]$ obeys

$$\begin{aligned} \Omega^\mu [\delta_\zeta, \delta] =& \tilde{\nabla}_\nu \left[2\tilde{\nabla}_\rho \zeta^\sigma \delta E^{\nu\mu\rho}{}_\sigma - 4\zeta^\sigma \delta \left(\tilde{\nabla}_\rho E^{\nu\mu\rho}{}_\sigma\right)\right] - 2\tilde{\nabla}_\nu \left(\theta^{[\nu} [\delta] \zeta^{\mu]}\right) + \tilde{\nabla}_\nu \delta Q^{\nu\mu}_{\psi,\zeta} \\ & + \Pi^\mu [\zeta, \delta] - \theta^\mu [\delta] \tilde{\nabla}_\nu \zeta^\nu + \frac{4}{n} \delta E^{\mu\nu\rho}{}_\nu \tilde{\nabla}_\rho \tilde{\nabla}_\lambda \zeta^\lambda - \frac{4}{n} \delta \left(\tilde{\nabla}_\rho E^{\mu\nu\rho}{}_\nu\right) \tilde{\nabla}_\lambda \zeta^\lambda \zeta^\mu \\ & + \frac{1}{16\pi} \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} \left[\mathring{A}^{\alpha\beta} + 8\pi \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{2\frac{k+1}{n}} \left(T_{\mu\nu} - \frac{1}{n} T g_{\mu\nu}\right)\right] \delta g_{\alpha\beta} \\ & - \frac{1}{8\pi} \left(\frac{\sqrt{-\mathfrak{g}}}{\omega}\right)^{\frac{2}{n}} A_\nu{}^\mu \zeta^\nu + \zeta^\mu A_\psi \delta\psi - \delta \left(\zeta \cdot A_\psi \cdot \psi\right)^\mu. \end{aligned} \tag{A.24}$$

So far, all the expressions are off shell. We are mainly interested in the setup in which both the original and the perturbed spacetime are solutions of the equations of motion. Then, it holds

$$\begin{aligned} \Omega^\mu [\delta_\zeta, \delta] =& \tilde{\nabla}_\nu \left[2\tilde{\nabla}_\rho \zeta^\sigma \delta E^{\nu\mu\rho}{}_\sigma - 4\zeta^\sigma \delta \left(\tilde{\nabla}_\rho E^{\nu\mu\rho}{}_\sigma\right)\right] + \tilde{\nabla}_\nu \delta Q^{\nu\mu}_{\psi,\zeta} - 2\tilde{\nabla}_\nu \left(\theta^{[\nu} [\delta] \zeta^{\mu]}\right) \\ & - \frac{1}{8\pi} \zeta^\mu \delta\Lambda + \Pi^\mu [\zeta, \delta] - \theta^\mu [\delta] \tilde{\nabla}_\nu \zeta^\nu + \frac{4}{n} \delta E^{\mu\nu\rho}{}_\nu \tilde{\nabla}_\rho \tilde{\nabla}_\lambda \zeta^\lambda - \frac{4}{n} \delta \left(\tilde{\nabla}_\rho E^{\mu\nu\rho}{}_\nu\right) \tilde{\nabla}_\lambda \zeta^\lambda. \end{aligned} \tag{A.25}$$

Finally, integration over a suitable Cauchy surface $\mathcal{C}$ yields the symplectic form $\Omega [\delta_\zeta, \delta]$ and, therefore, also the perturbation of the Hamiltonian corresponding to the evolution along ζ^μ (1.185).

A.4 Removing contractions with an arbitrary timelike vector

In this appendix, we prove that the if the equation $f_{\mu\nu}n^\mu n^\nu = 0$ holds in a spacetime with dimension $n \geq 2$ for every timelike, unit, future-pointing vector, it implies $f_{\mu\nu} = 0$ (with no loss of generality, we assume $f_{\mu\nu} = f_{\nu\mu}$). To that end, we construct a local orthonormal coordinate system defined so that the metric locally reduces to the Minkowski one, i.e., $g_{\mu\nu} = \eta_{\mu\nu}$. We set the local direction of time so that $n^\mu = (\partial/\partial t)^\mu$ and we denote the spatial coordinate vectors by $e_i = \partial/\partial x^i$. We stress that, since $f_{\mu\nu}$ is a tensor, we are free to choose any coordinate system without any loss of generality. We introduce the following set of unit timelike vectors

$$t^\mu_{ij} = \sqrt{(1+p^2+q^2)}n^\mu + pe^\mu_i + qe^\mu_j, \tag{A.26}$$

where i, j, are natural numbers such that $0 < i < j \leq n-1$, and p, q are arbitrary real numbers. Since $f_{\mu\nu}t^\mu_{ij}t^\nu_{ij} = 0$ for every t^μ_{ij}, we have

$$\begin{aligned}&\left(1+p^2+q^2\right) f_{00} + p^2 f_{ii} + q^2 f_{jj} + 2p\sqrt{(1+p^2+q^2)}f_{0i}\\&+ 2q\sqrt{(1+p^2+q^2)}f_{0j} + 2pqf_{ij} = 0,\end{aligned} \tag{A.27}$$

which must hold for any real p, q. Thence, every coefficient in the expansion of the left hand side in the powers of p, q must be zero. We only need the first few conditions implied by this procedure

$$f_{00} = 0, \tag{A.28}$$
$$2pf_{0i} = 0, \tag{A.29}$$
$$2qf_{0j} = 0, \tag{A.30}$$
$$p^2\left(f_{00} + f_{ii}\right) = 0, \tag{A.31}$$
$$q^2\left(f_{00} + f_{jj}\right) = 0, \tag{A.32}$$
$$2pqf_{ij} = 0. \tag{A.33}$$

To satisfy these equations for every i, j, it must hold $f_{\mu\nu} = 0$.