Emergent Gravitation

Krhač, Kaja

Master's thesis / Diplomski rad

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:166781

UNIVERSITY OF ZAGREB
FACULTY OF SCIENCE
DEPARTMENT OF PHYSICS

Kaja Krhač

EMERGENT GRAVITATION

Master Thesis

Zagreb, 2021

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MAEMTATIČKI FAKULTET
FIZIČKI ODSJEK

Kaja Krhač

IZRANJAJUĆA GRAVITACIJA

Diplomski rad

Zagreb, 2021

UNIVERSITY OF ZAGREB
FACULTY OF SCIENCE
DEPARTMENT OF PHYSICS

INTEGRATED UNDERGRADUATE AND GRADUATE UNIVERSITY PROGRAMME IN PHYSICS

Kaja Krhač

Master Thesis

Emergent Gravitation

Advisor: Professor Ivica Smolić, Associate Professor

Master Thesis grade: ______________________

Committee: 1. ________________________

2. ________________________

3. ________________________

Master Thesis defence date: ________________

Zagreb, 2021

I would like to thank my supervisor Prof. Ivica Smolić for his guidance, patience and support.
I thank my family for all their encouragement throughout my university years.
Last but not least, I thank my friends (in alphabetical order) Ivana, Marija, Marta and Željka.

Izranjajuća gravitacija

Sažetak

Opća teorija relativnosti opisuje gravitaciju kao zakrivljenost prostovremena, glatke mnogostrukosti s Lorentzovom metrikom. Otkriće da crne rupe zadovoljavaju zakone analogne zakonima termodinamike te da zrače termalnim spektrom bacilo je drugačije svjetlo na opću teoriju relativnosti i utvrdilo vezu između opće relativnosti, termodinamike i kvantne teorije. Ozbiljno shvaćanje te veze dovelo je do pokušaja da se gravitacija opiše kao izranjajući fenomen, rezultat kolektivnog gibanja nekih trenutno nepoznatih mikroskopskih stupnjeva slobode, a paradigma je dobila ime “izranjajuća gravitacija". U ovom radu dan je pregled različitih modela izranjajuće gravitacije, kao i ograničenja koje takva teorija mora zadovoljavati kako bi uspješno reproducirala opću teoriju relativnosti na makroskopskoj skali. U radu razmatramo različite pristupe paradigmi izranjajuće gravitacije na primjeru kauzalne teorije skupova gdje se prostorvrijeme uzima kao izveden koncept koje je rezultat uprosječivanja diskretnog skupa točaka između kojih postoje kauzalne relacije, Weinberg-Wittenovog teorema koji postavlja ograničenja na ideju da je graviton, medijator gravitacijske interakcije, "izranjajuć" i ulogu termodinamike u općoj teoriji relativnosti u sklopu crnih rupa, ali i horizonata općenito.

Ključne riječi: opća teorija relativnosti, kauzalna teorija skupova, Weinberg-Wittenov teorem, termodinamika crnih rupa, termodinamika prostorvremena

Emergent Gravitation

Abstract

The general theory of relativity describes gravitation as a curvature of spacetime, smooth manifold with Lorentzian metric. A different light was shed on the theory with the discovery that black holes behave as thermodynamic objects, satisfying laws analogous to the one of thermodynamics and radiating with a thermal spectrum, confirming that gravitation, thermodynamics, and quantum theory are intimately related. Taking this connection seriously led to attempts, united under the name “Emergent Gravitation", to describe gravitation as an emergent phenomenon, a result of collective motion of some yet unknown microscopic degrees of freedom. In this work, we give an overview of various models of Emergent Gravitation, as well as the constraints such models should satisfy to reproduce General Relativity at a larger scale. We consider different lines of thinking in the Emergent Gravitation approach by studying Causal Set Theory as a representative of Emergent Gravitation models where spacetime is considered as a derived notion, resulting from coarse-graining of a discrete set of points with causal relations between them, Weinberg-Witten theorem as a constraint on an idea that graviton, a mediator of gravitational interaction is emergent, and the role of thermodynamics in General Relativity in the scope of black holes and their horizons, but also horizons in general.

Keywords: general relativity, emergence, causal set theory, Weinberg-Witten theorem, black hole thermodynamics, spacetime thermodynamics

Contents

1 Introduction

Emergent gravitation (EG) is an idea that General Relativity (GR) is an effective theory, similar to hydrodynamic or thermodynamic description, underlined by a novel, more fundamental, yet unknown microstructure. It is novelty in particular that distinguishes EG viewpoint from various approaches trying to quantize gravity.

The reason why one would change the perspective on GR in the first place is a number of obstacles faced from both observational and theoretical standpoint. On the one side, although GR is a very successful theory capable of explaining a great number of gravitational phenomena, there is an increasing number of observations GR fails to capably address, such as accelerated expansion of the universe unless large quantity of some non-Standard Model type of matter and energy is considered. On the other side, GR seems to fail to be quantized in any standard way. Motivated by these issues, there are many possible alterations and extensions of GR, and EG approach may be considered as one of them.

The paradigm of EG is motivated by the fact that some properties of gravitation resemble the behaviour of thermodynamic systems. One of the first such indications, appearing in 1970s, is the similarity between the laws describing black holes, called the four laws of black hole mechanics and the four laws of thermodynamics. It was first shown by Hawking that area of the event horizon of a black hole can never decrease. This is analogous to the second law of thermodynamics stating that the entropy can never decrease. Soon after, Bardeen, Carter and Hawking introduced the zeroth, the first and the third law. They showed a certain quantity called surface gravity is constant for stationary black holes, similarly to how temperature is constant in equilibrium in thermodynamics. Moreover, the first law describing how parameters of the black hole change bears resemblance to the first law of thermodynamics, a version of conservation law. The third law states that the surface gravity cannot be reduced to zero by any finite number sequence of operations. However, the four laws were seen to be only formally analogous to thermodynamics, since black hole, as perfect absorber, can only have a temperature of absolute zero. This changed after Hawking showed that black holes radiate with a black body temperature, solidifying the analogy to equivalence.

Further property suggestive of emergent nature is the universality of gravity. It attracts all bodies carrying energy with the same strength, i.e., irrespectively of other properties the bodies might have. This reminds of the situation in molecular physics where all molecules attract one another due to dipole moments induced by fluctuations of charge distribution in the molecules. Hence, the idea is that attractive nature of gravity is caused by similar fluctuations in something.

Another reason is that GR is perturbatively non-renormalizable. The situation is similar to the case of Fermi interactions explaining beta decay, which was later replaced by a theory of weak interactions. In addition, a common result obtained by linearizing classical scalar field theory around some non-trivial background is emergence of "effective Lorentzian geometry" of a curved spacetime, which is in the premise of “analog models" of general relativity. In a similar manner, as showed by Sakharov, one-loop effective action of quantized matter fields propagating on Lorentzian manifold considered as background contains a term proportional to Einstein-Hilbert action. Simply put, dynamics of general relativity emerges from quantum field theory. Thus, there is wide variety of indications hinting towards gravity as emergent phenomenon.

That being said, theories of EG are concerned with the way the fundamental degrees of freedom interact to produce GR. The models can be sorted into two large groups based on whether or not this micro-constituents live in some ambient space or not. Models also differ depending on which aspects of GR are emergent — spacetime, metric, or dynamics of spacetime, although the latter kind is commonly not classified as emergent theory. The classification is not very sharp since the constituents might have some properties of the macroscopic structure one is trying to reproduce.

On that account, it should be pointed out that emergence is not a new concept, nor it is tight solely to physics. As highlighted by Anderson’s 1971 paper “More is different", there are new emergent laws governing dynamics of the basic constituents and new modes of interaction at every level of structure. The idea originated in biology and is prevalent in condensed matter physics. Consequently, ideas in EG are largely influenced by the concepts from this area of research. This is the rough idea behind EG viewpoint.

This paper provides an overview of different EG models, but it does not focus on their methods and ideas in much detail. Instead, certain specific results and constraints will be considered, as representations of different lines of thinking in EG.

Particularly, we will discuss the Causal Set Theory as a good example of EG model in which spacetime is considered emergent. Next, we will look at Weinberg-Witten theorem as a constraint of the kind of particles can live on flat spacetime. Finally, we will rederive some of the important results relating GR and thermodynamics.

The paper is organized as follows. In Section 2 we review foundations of General Relativity by following Ehlers-Pirani-Schild formalism, physically motivating the mathematical structures of General Relativity by considering the behaviour of light and matter. The emphasis is on the implications the choices that have been made to construct the theory may have on some of the shortcomings of General Relativity, signaling the need for a new approach. This is where the Emergent Gravity paradigm comes into play. We clarify what emergence is (not) using various examples from condensed state physics and put forward the hints that point towards emergent nature of gravity. At the end of the section we give an overview of EG models and explain the constraints they should satisfy to have a chance in reproducing GR.

Section 3 describes Causal Set Theory (CST) approach without much detail. CST is based on hypothesis that the structure of spacetime is discrete. As the claim cannot be verified at this point we first provide theoretical arguments why this should be so. Then the main premise of CST is put forward. Namely, that the causal order is fundamental and spacetime consists of points among which exist causal relations. The smooth structure called spacetime is just coarse-grained version of causal set. Lastly, we also very briefly discuss the dynamics of causal set

Section 4 is dedicated to Weinberg-Witten theorem. It constraints spin of particles based on symmetry considerations. We first review the important concepts needed to prove the theorem and discuss its implications on graviton, hypothetical mediator of gravitational interaction. The consequence of the theorem is that graviton on flat spacetime cannot be an emergent particle in a sense that it cannot be a composite particle of other matter fields.

The goal of section 5 is to explain the relationship between thermodynamics and General Relativity. We derive the four laws of black hole thermodynamics and give an overview of the many versions of the derivation showing that the theory is consistent and the four laws are not just accidental. We also derive Hawking radiation using WKB approximation and Hawking temperature of black holes using path integral formulation. Next, we discuss how well the analogy between black holes and thermodynamics is established. More precisely, we discuss if one can really view the laws of black hole mechanics as extension of thermodynamics to black holes. Finally, following Jacobson, we show that thermodynamics is not constrained only to black holes. The Einstein's field equations can be obtained from geometrical considerations of the geodesics, the second law of thermodynamics and its relationship to the area of the horizon, and Clausius' relation.

The final section contains summary of the obtained result and outlook.

2 Emergent nature of gravitation

In this section the goal is to explain what is understood under the term “emergent gravitation", providing context for topics that will be discussed in detail below. The discussion follows [1] and [2]. First, we briefly review the foundations of general relativity, current description of gravitation, mentioning along the way some problems from both theoretical and experimental standpoints, which signal possible need for modifications of current theory. One of such alternative approaches is given by theories which can be grouped under the name “emergent gravitation". We will clarify what it means for gravitation to be emergent and the properties which suggest that this may be the case. Finally, we give an overview of models which classify as emergent.

General Relativity is a description of spacetime, a four dimensional manifold with a Lorentzian metric on top describing gravitational field [3]. We start by explaining observations which lead to portrayal of gravitation as metric theory. This will be done by presenting the results of [4], referred to as EPS (based on the names of the authors of the article) in the text, a typical constructive approach to the subject.

EPS take light rays and freely falling objects described classically as fundamental entities. Light rays are treated as small packets of electromagnetic waves, and objects refer to any body whose dimension and structure can be neglected in normal circumstances. When referring to “particle" in the text, we mean a worldline of a freely falling object as explained above. Light rays and particles propagate in otherwise empty region of spacetime. One starts from the following axioms,

- Set of events $\mathcal{M}$ with set of worldlines of light rays L and particles P is considered as given.
- The wordlines define differential topology on $\mathcal{M}$. Topology is defined because the observed light and particle worldlines are continuous, i.e., light rays and particles don’t disappear at one point and appear at the other, but move along continuous trajectories. Next, light ray messages between particles P and Q are assumed as smooth and can be used as “radar soundings", assigning four coordinates to each event. Once the radar coordinates are defined one may use any compatible chart, providing differential structure to the topological space.

Thus, light rays and particles naturally provide topology and smooth atlas, making the set of events $\mathcal{M}$ into smooth manifold by definition. Furthermore, the light rays determine the causal structure, light cones i.e., which events can influence each other as well as relations measured by clocks, rods and lengths of four-vectors.

- Propagation of light determines quadric cone, which we call a light cone, in tangent space of each point of $\mathcal{M}$. Such manifold, where at each point there is a light cone singled out is said to have a conformal structure, an equivalence class of metrics proportional to one another, so they determine the same light cones. The conformal structure is assumed to be of Lorentzian signature. The local causality established by light rays is observed as local Lorentz invariance.

- As a result of conformal structure one may distingush between curves and vectors that are timelike, lightlike and spacelike at a point. This classification comes from the way light rays at an event separate vectors into different classes, i.e., from topology. It can then be proved that each class of vectors have the norm one would expect. Furthermore, it follows that tangent vector of a light ray is a null vector, and accordingly, the light rays are null curves. What's more, it is proved that null curves are null geodesics of any representative of the conformal structure.

The results so far uniquely determine differential and causal structure of set of events. To describe matter one needs additional structure.

- Out of all possible conformal classes, particles belong to singled out family of worldlines, to family of timelike curves. In other words, particles will always move along timelike trajectories, and not any other. What one should recognize is that no restrictions are imposed on the particle in the previous statement, i.e., particle carrying electric charge for example, moves along a timelike curve in the presence or absence of electromagnetic field.

- There is a class of particles, uncharged under any "non-geometrical" field, whose family of curves behaves in the second-order infinitesimal neighbourhood of each point of $\mathcal{M}$ like the straight line[1] of an ordinary projective

[1] To be more precise, what is understood here is that for each event e in $\mathcal{M}$ there exists a coordinate system x^μ defined in the neighbourhood of e such that any particle P through e, $\frac{d^2x^\mu}{d\lambda^2}|_e = 0$, where λ is parameter of the curve. Such coordinate system is said to be projective at e

four-space. In other words, trajectories of freely falling objects are autoparallels[2], defining a projective structure of connections — an equivalence class of connections which refer to the same worldline. Each representative of the class parametrizes the same autoparallel differently (and there is no preferred parametrization). The postulate is called *(weak) equivalence principle*.

We now have a projective structure imposed by autoparallel curves fixing the free-fall of objects, and conformal structure required by light rays, defining a family of null geodesics determining a causal structure and measurement.

Next, motivated by the fact that a massive particle, although slower then light, can chase a photon arbitrarily close, or in other words, trajectories of particles fill up the light cone, EPS demands that conformal and projective structure are compatible, meaning that Lorentzian metric underlies both conformal and projective structure.

- The compatibility requirement implies that autoparallels of projective equivalence class of connections and family of null geodesics of the conformal equivalence class of metrics are related, i.e., null geodesics must be autoparallels of the connection.

Space endowed with compatible conformal and projective structure is a Weyl space. From these data — the fact that null geodesics determined by conformal structure belong to the class of geodesics determined by projective structure — it is possible to derive existence and uniqueness of affine connection, such that its geodesics coincide with geodesics of projective structure and the properties of vectors which are null with respect to conformal structure are conserved under parallel transport defined by affine structure. Consequently, one may look at conformal structure and affine connections instead of projective structure, but it is important to point out that in order to get to affine connections, one uses both conformal and projective structure. Affine structure differs from projective structure in that the geodesics carry affine parameters, defined up to a linear transformation.

Finally, to obtain Lorentzian spacetime from Weyl spacetime, EPS impose one final condition by requiring that time runs equally fast along all paths. This is exclusion of the so-called second-clock effect.

[2]In [4], the curves are called geodesics, but here, we will refer to geodesics as both the straightest and the shortest lines, while autoparallels are the straightest lines, so one does not need a metric to define them.

- To rule out second clock effect one imposes a condition that magnitude of a vector under parallel transport is path independent.

There are also equivalent conditions of ruling out the second clock effect using Einstein simultaneity [5].

It is observed that “speed" of time is independent of the path. That is, the time interval depends only on the taken path. The absence of second clock effect implies that there exist a single Lorentzian metric[3], unique up to a constant positive factor, compatible with conformal and affine structure — Levi-Civita connection (also called Christoffel's symbols). In other words, Weyl's spacetime reduced to Lorentzian spacetime [4][6][7].

In summary, we have shown that structure of spacetime follows from properties of massive particles (affine connections) and light rays (metric). Due to the compatibility condition affine connections are Levi-Civita connections of metric. As a result, metric alone determines the causal structure and free fall. With this setup, the metric field is used to encode gravitational effects. This is only kinematical description. To obtain dynamics of the metric field from kinematics given above one needs to provide Hamiltonian or Lagrangian. One way to obtain it is by postulating diffeomorphism invariance, eliminating the possibility of nondynamical objects.

Nevertheless, the requirement of diffeomorphism invariance does not exclude the existence of additional dynamical fields, besides the metric field. We have shown that the response of matter to gravitational field is through metric, but this does not imply that the reverse is also true, i.e., that the response of gravitational field to matter occurs only through metric. Thus, one should add another postulate requiring that gravitational dynamics depends only on metric. As a result, Lagrangian density of gravitation is given by Ricci scalar (to the lowest order), and varying the action with respect to the metric yields Einstein's field equations. Such theory reproduces Newton's law in weak-energy limit, Kepler's laws related to the planetary motion, and is consistent with the observations in the Solar System [8]. On the other hand, there are number of problems that the theory fails to explain in a meaningful way. Before addressing them, let's reflect back and emphasize choices that were made. This will make it easier to understand from where the discrepancy may be coming

[3]The proof consists of using the fact that affine structure determines curvature tensor, which is Riemannian only if the second clock effect is ruled out.

from.

One of the things that should not be overlooked is the choice of light rays as probes, treated classically. Electromagnetic field is the only type of field found in nature in classical field theory so there is not much to choose from in the first place. What we probed was the background on which light propagates. From the formal point of view, it is a structure of smooth manifold that underlines electromagnetism. In other words, smooth manifold provides enough structure to put tensor fields, such as electromagnetic four-potential on it, although one usually does not speak of it when discussing electromagnetism. The role of the background is most prominent in the covariant description of matter[4] Lagrangian. Maxwell's equations describing dynamics of electromagnetism besides the four-potential and its derivatives contain also metric coefficients, describing the structure of the background on which the electromagnetism is formulated. Thus, another important point that was only implied is that the probe (light) must couple to whatever we are trying to describe (background). In essence, the behaviour of light depends on the choice of background. Furthermore, since the metric coefficients appear in the matter Lagrangian one gives the background its own dynamics chosen so that it captures gravitational effects in the curvature. Hence, the background itself carries a geometric structure, only determined by matter chosen to probe it.

As a consequence, Lorentzian metric is a geometric structure in line with Maxwell's theory. Turning the argument the other way around, if one starts with Lorentzian spacetime, the matter coupled to it will not show phenomena like birefringence for example, where different polarization states of photon propagate differently, since such behaviour would require a tensor field of rank four[5][10][11]. These kind of effects are possible in some anisotropic materials, implying non-linear electrodynamics, but also in quantum vacuum which acts as a dispersive medium [9][12][13].

What's more, as a consequence of the Lorentzian signature of the metric the observed matter field dynamics coupling to it is predictive[6], i.e., initial conditions are

[4]Matter refers to non-gravitational fields. In the classical picture that is only electromagnetism.

[5]It should be pointed out that birefrigence does not necessarily lead to change of the tensorial structure. Moreover, in some cases it leads to bimetricity, where different photon polarization see distinct metrics [9]

[6]Requirement of predictivity can be formulated as algebraic condition. Namely, that principal polynomial obtained from matter action is hyperbolic. In case of metric this condition mandates Lorentzian signature [14].

translated into later values of the field equations, as should be the case for any sensible theory. For example, electromagnetism formulated on a manifold with Riemannian metric would not be a predictable theory [15][14]. The point is, geometric structure of spacetime is constrained by the matter that interacts with it and vice versa.

Moreover, it was propagation of light that made us perceptive of metric, but it has even more prominent role when it comes to massive particles. All known matter can be divided into two large groups: bosons, described by tensorial fields and fermions, described by spinor fields. The latter can only be defined with the help of metric[7]. In summary, all type of matter we observe stems from the same geometric structure as required by electromagnetic theory.

Further remark concerns the compatibility condition. It should be emphasized that connection is a structure governing free fall, independent of causal structure determined by metric. It is only after imposing compatibility condition, motivated from the standpoint of physics, such that null geodesics have to be autoparallels of the connection, that the metric connection specializes to Levi-Civita connection. *A priori*, these structures are independent. Although we have physical motivations for imposing compatibility, restrictions on structures should follow from the dynamics, not being fixed in advance.

The next note is about the choice of dynamical variable. We have chosen only metric as a dynamical variable, while connection which governs free fall was given no dynamics. However, it was shown that in the case where metric is minimally coupled to matter, and both connection and the metric are treated as dynamical variables, the Levi-Civita connection is a result of the field equations. Thus, when coupling is minimal, it does not matter whether we fix the metric and connection *a priori*, or treat both as dynamical variables, although the latter is a more honest approach. What's more, although the metric is a dynamical variable it does not mean it is a fundamental degree of freedom, only that it naturally emerges under the circumstances explained above.

This concludes discussion about theoretical framework of General Relativity. We now put forward some of the shortcomings of GR. It fails to capably address galactic,

[7]The fastest way to see this, and avoid too much technicality, is to remember that spin representation is given by representation of Clifford algebra, as for example Dirac (gamma) matrices γ^μ. To generate the Clifford algebra they must satisfy $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$.

extra-galactic and cosmic dynamics, like the observed accelerated expansion of the universe unless a huge quantity of some non-Standard Model type of matter-energy and energy is considered, called “dark energy" [7]. To be precise, approximately 83% of matter and 95% of total mass-energy in the universe is of the unknown type [16]. Further problems are related to the standard Big Bang cosmological model based on GR and SM. Namely, the flatness problem, stating that the universe is flat, as confirmed by for example WMAP. The problem is that if universe started as not precisely flat, the curvature would increase with time, making universe today strongly curved. In other words, to obtain the observed flat universe one needs to extremely fine tune the initial conditions in the standard big bang model, without physically motivated explanation. Next is the Horizon problem, describing the fact that although the universe consists of causally disconnected regions, measurements of COBE and WMAP show that temperature of black body radiation of these regions is finely tuned. To restate, the universe is assumed to be homogeneous, yet it consists of regions that are not in causal contact so it is not clear how homogeneity could be accomplished [17]. The third is the Monopole problem. The standard Big Bang model predicts a large number of massive magnetic monopoles, which should be observed today. However, this is not the case [18]. This leads to a conclusion that GR and Standard Model of particles are unsatisfactory in describing the universe in extreme energy-curvature regimes. The outcome may not be surprising, since GR was formulated by observing matter inspired by classical electromagnetism, which brings us to our next point. GR is a classical theory and currently, there is no successful theory describing gravitation as a quantum theory. For one, the methods using which all other theories have been successfully quantized failed when applied to gravity. One of the problems is that the quantization of field theories requires a fixed background, while in GR spacetime itself is a dynamical variable, there is no fixed background. Although there are ways to deal with such obstacles, they ultimately lead to a theory that is non-renormalizabile at higher orders, implying that the theory is unkwnown at the Planck scale.

Motivated by the above issues, there are many possible alternations and extensions of GR. On the one hand, one may be prompted to examine the geometric structure of spacetime, since a tensor with more degrees of freedom then Lorentzian metric may allow one to describe matter with different causal behaviour then electromagnetism. One may also add other dynamical variables, like the connection to Lagrangian, known as Palatini formalism. Then there are Extended Theories of Gravity, modifying the field equations by coupling fields non-minimally, or adding higher order derivatives of the metric. One could also add higher order curvature terms, since there is no reason to restrict the gravitational Lagrangian only to linear function of Ricci scalar, which is in the domain of $f(R)$ theories, etc.

The focus of this study is to look at models of emergent gravitation, whose approach is to consider gravity as a macroscopic, coarse-grained theory such as thermodynamics or hydrodynamics. The basic premise is that the spacetime — which includes all the structures, topology, differentiable structure and metric — is derivable from some more fundamental theory, ultimately leading to new structures and providing answers to current discrepancies as the one mentioned.

Emergence Emergence can be summed up as “more is different". As pointed out by [19], the behaviour of large and complex systems of elementary particles cannot be understood in terms of a simple extrapolation of the properties of a few particles, but at each level of complexity, entirely new structures appear, unexpected by the underlying theory. As an example, consider QED which describes how electrons interact at the most fundamental level. However, in $2D$ crystal where one has 10^{20} of pairs of electrons, and interaction of each of these pairs is dictated by QED, one finds a new phenomena, not predicted by QED in an obvious manner, Fractional Quantum Hall Effect (FQHE), where excitations of collective state of electrons behave as quasiparticles of fractional charge, and the reason for this is not because QED is lacking in some way. Discovering FQHE did not introduce any amends into QED. What's more, if one starts from FQHE, a better understanding of quasiparticles is gained only after electron is discovered. In other words, quasiparticles are not made of electrons, but are rather collective motions of electrons.

A similar example is Euler's equation, $\partial_t^2\rho - v^2\partial_x^2\rho = 0$, describing small deformations, such as waves, in a (zero temperature) liquid, where ρ is density of the liquid. The underlying entity of liquid waves are atoms, described by Schrödinger equation. Moreover, the dynamics of atoms depends on how they are organized in the ground state. When they are organized into crystal, a solid instead of liquid, their deformations are described by a different equation, namely, Navier's equation $\partial_t^2 u^i - T_j^{ikl}\partial_k\partial_l u^j = 0$, where $u^i(\boldsymbol{x}, t)$ describes the local displacement of the solid

[20]. Thus, it is difficult to conclude what is the underlying structure and how it behaves based on emergent phenomena.

The examples of emergence are not tied to condensate matter physics but also appear in nuclear physics, chemistry, or biology [19]. In short, they can be characterized as a collective behaviour not obvious from the fundamental laws.

In a similar manner to the examples above, emergent gravitation is a viewpoint considering Einstein's equations as non-fundamental, analogous to the Euler's equation for example, with different underlying degrees of freedom. We start by giving some examples of the definition.

"...the basic picture is that gravity, and perhaps space, or spacetime themselves are collective manifestations of very different underlying degrees of freedom." [2]

"... space and time are emergent concepts, i.e., not present in the fundamental formulation of the theory but appear as approximate semiclassical notions of macroscopic world..." [21]

"...we will intend as emergence of a given theory as a reorganization of the degrees of freedom of a certain underlying model in a way that leads to a regime in which the relevant degrees of freedom are qualitatively different from the microscopic ones." [22]

"The alternative viewpoint is that General Relativity is a low energy effective theory, and the metric and connection forms are the collective or hydrodynamic variables of some unknown microscopic theory. These variables will lose their meaning at shorter wavelengths and higher energies." [23]

"...these approaches take gravity to be an intrinsically classical, large-scale phenomenon arising from the collective action of the dynamics of more fundamental, non-gravitational degrees of freedom." [24]

The main takeaway of the examples from condensed state physics and the characterizations of emergent gravitation provided above is that

- Emergence is related to underlying “microstructure", i.e., the emergent system is composed of more fundamental degrees of freedom, obeying different laws.
- It is unlikely for emergent phenomena to be predicted before it is observed. Even if the fundamental degrees of freedom are known, the emergent phenomena is not likely to be “derived" from the properties of microtheory.

Consequently, [1] proposes the following definition of emergence,

We call a theory M_1 underlying a theory T, type I microtheory to T if and only if it is inspired from T (for instance through discretization, quantization or renormalization).

An underlying theory M_2 to theory T is called a type II microtheory if and only if it is not directly inspired or motivated from T.

The structure linked to a microtheory is called microstructure.

A theory is called emergent if and only if there exists an underlying type II microtheory to it.

“Motivated from T" in the definition above, for the case of gravitation does not ask whether or not the produced spacetime coincides with the general relativistic spacetime. That is, just because the theory T may be drastically different from General Relativity, does not automatically make it emergent in the sense of type II theories.

It should also be emphasized that at in certain cases discretization, quantization and renormalization may produce theories that classify as type II. We will explore this further below.

Moreover, we note that for example [2] differentiates between models based on the existence or lack of a medium in which the underlying degrees of freedom live, focusing more on the question which structures emerge. The emphasise does not seem to be on how and whether these degrees of freedom are fundamentally different or not. As a result, [2] takes into account some emergent models that are ruled out by [1]. Although the approach of [1] seems to be more in the spirit of what emergent theories of gravity try to convey, and as the authors point out, the definition of emergent gravitation must not be too wide concept, because otherwise it is useless, our intention here is ultimately not to rule out some of the theories, but provide overview and context important especially concerning the subjects that will be explored in the following chapters. We will be mindful of the definition above, but it will not be taken as constraint.

In the next section are given examples which show how the above definition is applied.

2.1 Types of coarse-graining

In this section we will look at discretization, quantization and renormalization, as a common examples of procedures which affect the degrees of freedom. We call "coarse-graining" anything that takes us from what is considered as fundamental description to some sort of "bigger" picture.

Discretization is most easily explained on an example. For this reason, consider macroscopic object of some mass density which varies in space. One knows that this is due to the number density of atoms varying in space (assuming that the object is made of identical atoms). Thus, mass density is a continuous approximation of a discrete quantity obtained by dividing object into regions and integrating over them. Even more, mass density is not emergent since the notion of mass exists for the most fundamental unit, i.e., every atom has mass.

In analogue way as mass density, one may try to discretise Lorentzian manifold. One such approach is causal set theory [25], where the role of mass is played by volume of a region. The atoms of spacetime, simply called “elements", are structureless objects organized in some way on the basis of causal relations among them, so that continuous limit produces a Lorentzian manifold. We will explore this further in a chapter below. For current discussion it is important to note that in causal set theory, spacetime, although discrete, is “not a different substance" [25].

Another example of discretization is Loop Quantum Gravity (LQG) where physically existing structure should be a quantum superposition of spin networks, three dimensional structures of intertwined loops with spin representation on the network’s nodes and edges. Each spin representation quantifies so-called “quantum volume" if it corresponds to a node, while discretely valued “quantum area" of the edge corresponds to the surface of the adjacency of the connected volumes. One should point out that is not yet understood how a topological structure like spacetime could emerge from it [24].

In [1], both causal set theory and LQG are ruled out as emergent theories. It seems to me that causal set theory fits into type I theory, but it is in the best scenario not clear what is the case for LQG since the coarse-graining method that would yield spacetime is not yet known. Spin-networks, to me, come across as uninspired from continuum they should produce, and as such, I would put them into type II category.

Quantization It is considered that a fundamental description of any theory is of quantum nature at atomic scale, so the same is expected to be true for spacetime. There are few difficulties when speaking about quantization and quantum phenomenon. First of all, how to quantize, and whether that is a well defined procedure in the first place is a complicated question. So, to explain what we mean under "quantization" we will say that in order to mathematically capture the observed wave-particle duality leading to probabilistic nature of Quantum Mechanics, the behaviour of particles, or more precisely wave-particles is given by wave functions, while observables are promoted to operators. As a result, quantities which are classically continuous take only certain values, as for example, angular momentum. On the other hand, for energy this is true only in the case of a bound system, but we still consider it as quantized. This is in fact important distinction between discretization, which always leads to discrete quantities, and quantisation, which although related, is a different concept.

Quantized theories are not emergent, because they are expected. In other words, as stated by [19], emergence does not oppose reductionism, i.e., there will be new fundamental laws and phenomena at an atomic scale, compared to macroscopic scale. For theory to be considered emergent, the change must be more than quantum behaviour explained above. What we mean is best understood on an example. Let's once again consider condensed matter physics. Quantizing vibrational modes of an atomic crystal yields phonons, but if we "zoom in", we do not see smaller parts of a phonon, but atoms that fill the entire space. The phonons are not formed by those atoms, the phonons are simply collective motions the atoms [20][23]. Hence, phonons can be considered as emergent, since the underlying degrees of freedom are different.

Consequently, we rule out theory of linearized gravity leading to GR as emergent. It can be shown that metric field and its dynamics arise through self-coupling of graviton, spin-2 particle that is a mediator of gravitational force, assuming the graviton field possesses gauge symmetry [26]. We do not view spin-2 field as non-gravitational, although extra steps are needed to relate it to GR.

Renormalization In quantum field theory, or any perturbative calculation, the observables quantities, like cross-sections, are obtained by summing over all possible intermediate states. It turns out however that beyond the lowest order, the perturbation expansion is ill-defined due to appearance of divergent quantities, originating as a result of intermediate states carrying arbitrarily large momenta. The renormalization takes care of these infinities. In the Wilsonian approach, one starts from "bare" Lagrangian, such that the parameters do not represent physical quantites, and cuts-off the high-energy states at momentum Λ, which makes divergent terms finite. The value of cutoff Λ is chosen so that it is above the energy scale of interest, i.e., the energy scale of experiment probing the system Λ_R, but below the "ultimate" high-energy scale. The goal is to obtain the same physics at Λ_R as at Λ. That is, we wish to obtain theory that is independent of chosen Λ, and without the divergences. This is done by decreasing the cutoff on momenta to $\Lambda - \delta\Lambda$, called "integrating out a momentum shell" and then compensating the effect of those momenta by adjusting the parameters in Lagrangian, so that they now depend on Λ, and adding new interactions. The dependance of the parameters on the cut-off is characterized by a beta function, and called renormalization group (RG) flow, as each iteration, i.e., each procedure of integrating out a momentum shell, is a point in parameter space, generating a trajectory. Ultimately, it is expected that a theory reaches a fixed point in UV region (at large energies), related to zeroes of the beta-function, where the theory behaves either as a free theory if the fixed point is trivial, or, if the fixed point is non-trivial, obtains a conformal group as a symmetry group, so it becomes independent of the scale. Such theories are renormalizable. [27]. Gravity is (perturbatively) non-renormalizable, but, there are theories, such as asymptotic safety, not based on perturbative approach, with idea that non-trivial UV fixed point exists even for non-renormalizable theories. Some techniques that search for asymptotic safety in gravity are Causal Dynamical Triangulation [28], and Regge calculus [29][30][31].

In short, the Wilsonian renormalization gives low-energy approximation to high-energy dynamics by integrating out high-energy modes. Although this results in the change of the parameters and Lagrangian with scale, it is expected that the theory eventually reaches a fixed point, such that the Lagrangian and values of parameters remain unchanged.

Consequently, concerning the question whether such procedure produces emergent theory — since the high-energy modes are integrated out, the degrees of freedom of effective Lagrangian are just low-energy degrees of freedom of the original Lagrangian. Thus, it makes sense not to consider them as fundamentally different. On the other hand, according to [32], elimination of degrees of freedom produces novelty if the effective Lagrangian is dynamically distinct from the original Lagrangian. An example is the Lagrangian of superfluid Helium 3-A. Non-relativistic Lagrangian of Helium-3 atoms is after a phase transition to superfluid phase described by effective Lagrangian describing hydrodynamical sound waves. Furthermore, as the temperature is lowered, the hydrodynamical Lagrangian becomes formally identical to the Lagrangian density for massless $(3 + 1)$ dimensional electrodynamics. Consequently, [32] concludes that "Dynamical distinctness entails a failure of law-like deducibility from Lagrangian of the properties described by effective Lagrangian, and a difference in field variables suggests the properties and entities described by effective Lagrangian and Lagrangian are ontologically distinct."

Hence, there is no unique answer, whether a procedure can or cannot result in fundamental change in degrees of freedom. The criteria for emergent theories is not very sharp and depends on what kind of differences one focuses on. Hopefully, the given examples still provided a notion of what can be regarded as emergence.

2.2 Indications of emergence

In this section we put forward some arguments implying that gravitation is emergent phenomenon. We will look at universality of gravity, it's perturbative non-renormalizability and black hole thermodynamics, which will be explored in more details in subsequent chapters.

2.2.1 Universality

There are two aspects of universality concerning gravitation. The first is universality of free fall, also called weak equivalence principle (WEP). It was first formulated in the scope of Newton's law of gravity as equivalence of inertial and gravitational mass, which suggests that it is impossible to locally distinguish between the effects of gravitational field from those of uniformly accelerated frames using the observation of the free fall of massive objects. Then, to be in accordance with Special Relativity, WEP is generalized so that mass is replaced with energy, since according to special relativity principle, mass is a manifestation of energy and momentum. As a result, it is impossible to distinguish between a gravitational field and uniform acceleration based on any experiment. Thus, the usual statement of WEP is that the trajectory of a freely falling object is independent of both its mass and internal structure, accounting for the fact that energy also gravitates. It depends only on starting position and velocity. The second aspect is universal strength of gravitation.

These universal aspects perhaps have the best explanation if gravitation is viewed as emergent, since similar universal notions are present in some coarse-grained processes and hydrodynamics [33].

It is noted in [34] that universal law of attraction reminds of London forces, a type of Van der Waals forces, in molecular physics. In system, consisting of (neutral) molecules any pair interacts with potential r^{-6}. This is due to the fact that electron distribution within a molecule fluctuates, so it may happen at an instant of time that electrons in a molecule are asymmetrically distributed, producing temporary dipole, which, in turn, induces a dipole moment in adjacent molecules. The force between such induced dipoles is attractive[8], and is equally strong, i.e., it is independent of the substance. Condensate of any type of matter will manifest the same behaviour. In the same manner, one could expect gravitation to be a manifestation of similar fluctuation of unknown substance similar to charge. In fact, it has been shown in [35] and further in [36], that gravity could be understood as force induced by fluctuations of the virtual particles of quantum vacuum. That is, starting from one-loop contribution to the effective action of quantum field theory on a Lorentzian manifold, where

[8]As the molecules come further together the force changes sign and becomes repulsive, since Pauli exclusion principle comes into play. As a result, the interaction between molecules is given by so-called Lennard Jones potential.

geometry is treated as classical background, by extremizing the action one obtains terms proportional to cosmological constant, Einstein-Hilbert action and curvature squared. In other words, dynamics of gravity will automatically emerge at one-loop level, which unlike the zero-loop level, takes quantum fluctuations into account [37].

Furthermore, in hydrodynamics of uncharged fluids the trajectory of every test body is the same, i.e., the trajectory does not depend of internal composition of the body, analogous to WEP. Since hydrodynamical laws are the result of coarse-graining, giving the same description of different underlying microstructure, one could expect the same is true for gravitation. On the other hand, since gravitation relates to the general framework of spacetime, universality of may be just a natural consequence [1].

2.2.2 Pertubative non-renormalizability

QFT approach to gravitation is to treat metric as a fundamental dynamical field, and quantize it like any other field. Before explaining the implications of (non)renormalizability, it is worth to emphasize that first of all, this kind of treatment of metric is based on formal analogy. Einstein's equations are second order differential equations obtained by varying metric, just as Maxwell's equations are for four-potential A_μ. Thus, it seems natural to promote metric to quantum field in QFT. Then, since quantization procedure requires a background on which the fields propagates, the metric is expanded around the flat metric (or any other metric is the background is not flat), $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, so that only $h_{\mu\nu}$ field is treated as dynamical. In other words, one chooses a background metric and linearizes fluctuations around it. The particle associated with $h_{\mu\nu}$ is spin- 2 graviton. However, one should bear in mind that assuming that metric is a dynamical field equivalent to external fields propagating on spacetime is a nontrivial step [38][39]. The question of fundamentality and role of metric is a complicated question, which we will graze a little bit in the further discussion. Nevertheless, this is the context in which we discuss perturbative non-renormalizability.

As already mentioned, in all field theories divergences appear at high-energy levels. However, we don't consider these divergences to be "true" divergences. Rather, they are the manifestation of new degrees of freedom belonging to some new theory. If this were not the case, no procedure could eliminate them.

In the traditional approach to renormalization — counterterm renormalization, the divergencies are taken care of by introducing a cutoff by some method of regularization, which respects the symmetry of the initial Lagrangian. The final result is independent of the cutoff, since these terms are cancelled by counterterms of renormalized Lagrangian, order by order. Such theories are called renormalizable. All the divergences can be removed by absorbing them into finite number of parameters, whose value is determined by experiment. For non-renormalizable theories on the other hand, infinite number of parameters would have to be measured to remove the divergencies, so such theories have no predictive power.

It turns out, gravitation is in the category of non-renormalizable theories. This can be seen from the criteria of superficial divergences, which gives correct predictions in the most cases [27]. In reality, non-renormalizability has been established up to second loop order [40].

The reason non-renormalizability has to do with possible emergence of gravity comes from the view that if the theory is fundamental, it must be renormalizable. The argument comes from requirement that the aim of physics is to formulate theories able to make prediction, and such theories are only possible if the theory contains only finite number of parameters to be determined by measurement. An example which further solidifies this position is the fact that Fermi theory of four-fermion direct interaction, non-renormalizable theory, proposed to describe weak interactions, was eventually replaced by a theory of weak interactions with a vector boson as mediator, which is renormalizable [41][42].

However, it should be pointed out that even non-renormalizable theory is able to make predictions. In today's approach to renormalization, i.e., Wilsonian approach, any theory is viewed as effective theory. By previously explained procedure, effective field theory makes prediction at energy scale that is of interest by implementing the effect of high-energy modes that have been integrated out in the form of additional terms in Lagrangian, which may be non-renormalizable and through the values of the coupling constants. Such interactions imitating high energy modes are local, they are polynomial in the fields and their derivatives, with the same symmetry as underlying theory. Furthermore, as Lagrangian has (energy) dimension four, so is the case for the interaction terms. As a consequence, interaction term of energy dimension $n+4$ must have coefficient of $\mathcal{O}(\Lambda^{-n})$. Thus, term of dimension $n+4$ only affects results in order $(p/\Lambda)^n$ or less. Since in reality we are interested in physics at a certain scale, one can determine the order which is required to achieve the desired accuracy and use the explained powercounting to identify which terms in the effective Lagrangian give the main contribution. At any fixed order only finite number of terms in effective Lagrangian can contribute, since the number of possible terms is limited, as it must respect the symmetry on underlying Lagrangian [43][44]. The point is, non-renormalizability does not obstruct predictivness, as long as the calculations are done at sufficiently low energy scale. There, the dominant contribution comes only from finite number of terms. The predictivity only fails when $p \propto \Lambda$, which is for gravity expected to be around the Planck energy scale, $M_P \propto 10^{19} m_p$, where m_p is mass of the proton.

In summary, in perturbative approach gravitation is non-renormalizable theory, and as such is not predictive at high-energy scales since to determine it's dynamics, i.e., Lagrangian, requires measuring infinite number of parameters. This implies, analogous to Fermi theory, that gravitation should be replaced by a more fundamental theory which would be renormalizable, and as such posses the predictive power at every energy scale.

2.2.3 Black hole thermodynamics

In the 1970s, a group of laws similar to the laws of thermodynamics were derived, describing behaviour of black holes, a region of spacetime nothing can escape from. The laws, and their respective thermodynamic analogue are stated below [45].

- The zeroth law:
  - Surface gravity κ is constant on the horizon of a stationary black hole.
  - Body in equilibrium has constant temperature.
- The first law:
  - $\delta M = \frac{\kappa}{8\pi}\delta A + \Omega_H \delta J + \Phi_H \delta Q$,
  - $\delta U = T\delta S - p\delta V + \mu dN$.
- The second law:
  - $\delta A \geq 0$,

- $\delta S \geq 0$.

- The third law:
  - No physical process can reduce the surface gravity of a black hole to zero by a finite number of operations.
  - No physical process can reduce the temperature of a system to zero by a finite number of operations.

The analogy is clear. Equilibrium corresponds to stationarity, temperature to surface gravity, and area to entropy. Moreover, black holes are generically described only by a small number of parameters, similarly to how systems in thermodynamics, containing large number of degrees of freedom can be described with a few macroscopic parameters. The analogy was further solidified with the discovery of Hawking radiation — black holes radiate with a black body spectrum of temperature $T_H = \kappa/2\pi$. As the Hawking radiation is of the order of mikrokelvin for generic black holes, the result is out of scope for experimental verification. For comparison, Cosmic Microwave Background radiation is 2.73 K. However, as it will be shown below, the laws are not just a coincidence, in a sense that they can be derived on many different ways, showing that the theory is consistent.

If the laws of black hole mechanics truly are extension of thermodynamics to GR, one is lead to conclude that there must exist more fundamental description of black holes, some new law underlying Einstein's equations, since thermal properties of the system reflect the statistical mechanics of underlying microstates.

Hence, the question is, how seriously should one take this analogy. For one, there is more to thermodynamics then the four laws. Moreover, entropy is extensive, scaling as volume, while black hole entropy scales an area. Another objection is that Hawking radiation is not black body radiation in the full sense. A black body, such as a lump of hot coal, radiates due to dynamics of microscopic degrees of freedom, while Hawking radiation is produced by quanta living nearby black hole horizon, not by microscopic constituents of black hole itself [46][47].

These issues are more thoroughly investigated in the chapters below. In summary, one can say that there certainly is a correspondence between thermodynamics and GR, but how far the resemblance goes is an open question.

2.3 Overview of emergent gravitation approaches

Emergent gravitation models are usually classified in one of two ways. The first one is by differentiating between top-down and bottom-up approaches as for example [23] and [1], where top-down refers to energies of the theories, i.e., microtheory corresponds to higher energies or vice versa. The second way is based on nature of the underlying microstructure. For example in [2] and [32], one distinguishes between emergent theories where degrees of freedom live in some setting (type I[9]), and the one where space and spacetime themselves are emergent (type II). The classification is not very sharp, as there are theories for which it is not evident into which group they belong, for example causal set theory, where continuum — spacetime, is emergent, but the microstructure contains some of its elements. Here, we will follow [2] and categorize models based on whether or not they live in some setting. We will add a category of type III models where dynamics is emergent, although they do not satisfy either the emergence criteria of [1] or [2].

- Type I — fundamental degrees of freedom live in an "environment"
  - *Analog models* such as [48][49], where, as the names suggests, one probes aspects of gravity by developing analogies, usually based on condensed matter physics.
    - The most well-known of these analogies is the use of sound waves in a moving fluid as analogy for light waves in curved spacetime. As pointed out by [1], such models do not establish gravity as emergent, but rather hint to its emergent nature.
    - Similarly, in [23], spacetime physics is considered as low temperature hydrodynamics, and the metric or connections are hydrodynamics variables. Macroscopic gravitational phenomena can be explained as collective modes, while spacetime itself is considered a condensate, a collective quantum state of many atoms with macroscopic quantum coherence.
  - *Models inspired by condensed state physics*

[9] These types do not indicate the type of microtheory from the definition of emergent theories.

* *Excitations near Fermi point*, for example [50]. The idea is that degenerate state of some condensed matter system may imitate the properties of quantum vacuum, and excitations of such state may possess some gravity-like properties. Since particle physics deals with interacting Fermi (spinors) and Bose (bosons) quantum fields, the condensed matter system ought to be fermionic — metals, superconductors, normal or superfluid He-3, since only they can capture both fermionic and bosonic nature. Fermionic degrees of freedom come from fermions, electrons or helium atoms respectively, while Bose fields appear as low energy collective modes. Furthermore, one differentiates between fermionic systems based on topology of the energy spectrum of fermionic quasiparticles. The most familiar is gapless system with a Fermi surface, in which category belongs non-interacting fermionic gas, whose Fermi surface is a sphere. In the case of gravity, however, we are interested in gapless systems with topologically stable point nodes instead of surfaces, called Fermi points. Gauge and gravitational fields are then collective bosonic modes of such systems [51].
* *Boundary excitations of quantum Hall effect* [52], where elementary particles, for example photons and gravitons, are modeled as collective boundary excitations of four dimensional quantum Hall effect. The system consists of fermions moving on four dimensional sphere, with three dimensional boundary introduced by applying a confining potential. Then, one can construct local bosonic operators for creating collective excitations at the boundary. The time evolution of these operators, given by Heisenberg equation, i.e., by the commutator of Hamiltonian and bosonic operator satisfies relativistic wave equations, which for helicities $j = 1$ and $j = 2$ coincide with free Maxwell and linearized Einstein equations respectively.
* *Defects in crystal*, such as [53] or [54], where universe is viewed as a crystal with defect tensor playing the role of Einstein curvature tensor, since curvature and torsion induced by elastic deformations and topological defects can be formulated in the scope of differential geometry.

– *Spin system on a lattice*, as given by [55], where fundamental degrees of freedom are qubits, used in the current implementation of quantum computers. In this model, qubit is a fundamental degree of freedom, where vacuum corresponds to the ground state of qubits, while collective motions of qubits organized in a certain way can be described by Maxwell of Einstein equation. Different organizations lead to different collective excitations.

- Type II — spacetime is emergent
  - *Graph-based models*, where graph refers to mathematical structures used to model pairwise relations between objects. In this context, a graph is made of vertices, which are connected by so-called edges. Examples include Loop quantum gravity, causal set theory, quantum graphity [56] [57], Discrete Unitary Causal Theory (DUCT) [58] etc.,
  - *Group field theory approach to quantum gravity* [59], whose approach is similar to the familiar field theories, with one of the main differences being that it is background independent. The fundamental building blocks are atoms of space, combined so to give rise to geometry and topology of space, while spacetime is discrete history of creation and annihilation of these atoms. All spacetime information is reconstructed from information carried by the atoms.
  - *AdS/CFT correspondence* [60] in versions where CFT is considered primary and bulk spacetime emergent. The statement of AdS/CFT correspondence is that string theory, which is a gravity theory, in asymptotically Anti de-Sitter spacetime is equivalent to quantum field theory in flat spacetime on its boundary.
- Type III — only dynamics of gravity is emergent
  - *Induced gravity models* [36][36], where Einstein-Hilbert action appears from by extremizing one-loop effective matter action minimally coupled to metric,
  - *Spontaneous symmetry breaking models*, given in [61][62][63][64], based on analogy of how electromagnetic and weak interactions were unified

after realizing that the intermediate boson is massive, which is captured with spontaneous symmetry breaking,

  - *Thermodynamical approach*, such as [65], [33] or [66], where metric is determined by thermodynamics.

Constraints of emergent models Although emergent models are currently beyond experimental verification, there are certain experimental constraints that hold even on the scales where spacetime is expected to be discrete and potentially described by some emergent gravitation model.

- Local Lorentz Invariance — terrestrial experiments, as well as astrophysical observations severely limit possibility of LLI violation. Current bounds suggest that it would be very difficult to violate Lorentz invariance at the Planck scale up to the sixth order of field theory [67]. The point is, there is currently no observational evidence of LLI violation so we expect that emergent model theories should not break LLI.

  From theoretical aspect, to obtain Lorentz transformations, one starts from two principles — relativity principle, according to which the laws of physics are irrespective of the reference frame, and the principle that states the speed of light is independent of the velocity of the source. However, it was shown that one does not need the second principle, if one assumes isotropy and homogeneity of space. This implies that two observers must agree on their relative velocities, or equivalently the transformations and their inverse are linear. In such approach Lorentz transformations contain limiting speed, which needs further assumptions to identify it with the speed of light, like for example, invariance of electrodynamics under Lorentz transformations [68]. In summary, results depends on what assumptions are taken as fundamental.

  In the context of emergent gravitation models, we will take the latter, as it is more general and makes it easier to see what kind of model will satisfy Lorentz invariance.

  - Type I: the background environment usually violates either relativity principle or isotropy. This will be most clear in an example of causal set theory, which succeeds in making the background Lorentz invariance. Moreover, a generic result of linearizing field theory around non-trivial background results in emergence of curved spacetime with effective Lorentzian geometry. In this category are for example analog models [49]. However, unless the field theory contains only scalar field, one ends up with bi-metricity, multi-metricity, or even Finsler instead of Lorentz geometry. This can be "mended" by adding certain constraints, but they don't have poor, if any physical motivation, other then producing a single Lorentzian metric.

  - Type II: LLI is a lesser challenge, but the problem usually arises if LLI is attainable by averaging over Lorentz non-invariant configurations because the Lorentz group is non-compact, so integral over boosts diverges.

- Universal coupling — meaning that all forms of matter couple to a single gravitational field with equal strength, i.e., with strength independent of position or velocity of matter. According to Weinberg's soft-graviton theorem, a single metric together with LLI ensures that the coupling of matter to gravitational field is always with equal strength.

  - type I models usually deal with one of the two problems. In models of gravity as fluctuation around a background many gravitons, i.e., spin-2 fields exist. Similarly, in models where gravitons are composite particles there are multiple potential metrics. The first problem may be solved by imposing invariance of effective action under diffeomorphisms and Local Lorentz transformations, resulting in only one massless spin-2 field. The issue is that massive spin-2 fields then in most cases contain negative energy Boulware-Deser ghosts.

  - Type II models are in most cases not yet able to describe coupling of matter to gravity. Nevertheless, since such models should reproduce a dynamical spacetime from the fundamental degrees of freedom, a metric would be a natural way to describe its dynamics. Universal coupling could once again be obtained from LLI and Weinberg's soft graviton theorem.

- Diffeomorphism invariance — which we will define as absence of non-dynamical background. If we are looking for a complete theory, then no *a priori* fixed background fields should exist. In other words, fixed background leads to a question why i.e., what fixes this background structure. One should have a reason why a background structure is not dynamical. Furthermore, diffeomorphisms can be understood as gauge transformations of general relativity, important for projecting out additional degrees of freedom of graviton field in a Lorentz covariant manner.

- Type I models are mostly formulated on the fixed background so the only way to obtain diffeomorphism invariance is if the background can decouple from observable quantities
- Type II models must first be able to reproduce smooth manifold before one can talk about diffeomorphism invariance. A possible complication regarding diffeomorphism invariance is that these models contain a time parameter describing evolution of underlying degrees of freedom, which poses a question about relationship of time in which fundamental degrees of freedom evolve and emergent notion of time. Once again, fundamental and emergent time must decouple, not to define a preferred time, breaking diffeomorphism invariance in the process.

In one of the following chapters we will look at result which further restricts emergent models of type I, Weinberg–Witten's theorem.

3 Causal set theory

The causal set theory (CST) is based on hypothesis that the structure of spacetime is discrete. There are no experimental evidence yet to support this claim as the scale at which the spacetime is expected to breakdown is the Planck length $l_p = \sqrt{8\pi G\hbar}$, where $c = 1$, which is of the order $10^{-35}\,\mathrm{m}$, or about 10^{-20} times the diameter of a proton. Such scale is beyond any experiment. However, there exist theoretical reasons why it is expected that spacetime is discrete, which we will review in this section. Next, it will be explained how the causal structure of continuous Lorentzian spacetime, i.e. relationships which specify which events can influence which, points to a discrete structure made of causal set — elements with causal order between them, which up to conformal factor of the metric determines geometry of spacetime. Finally, we give a brief outline of dynamics of causal set.

Motivation Theoretical arguments which point toward discrete spacetime, are what [69] refers to as three of four infinities. They are

- Divergences in QFT which, as it was already explained, can be taken care of through the process of renormalization. However, since such theories are expected to become trivial at some energy scale, i.e., they reach a fixed point in order to make sense, it is best to treat them as effective, valid only up to cutoff, until it is proved that fixed point indeed exist
- Singularities in GR, such as BH or Big Bang, since the known physical laws all break at such points,
- Non-renormalizability of naively quantized gravity, i.e., perturbative canonical quantization leads to non-renormalizable theory, which, unless understood as effective theory with a cut-off, is not predictive,
- Infinite value of BH entropy when no cutoff is present[10].

All these problems indicate that spacetime may be discrete [70]. The main challenge is then the following. If the spacetime is indeed reduced to some discrete structure, what would that structure be in order to reproduce topology, differentiable structure and metric which one observes at a larger scale. To answer this question, one must first presume that there is some sort of connection between the underlying microstructure and spacetime, otherwise it is hard to draw conclusions. As CST argues, that is the causal order. In EPS approach causality was determined by metric. CST, on the other hand, takes causal order as fundamental and derives all other structures, except the conformal factor, from it.

Formally, in the standard approach to GR one chooses the most minimal topology, just so to provide enough structure to capture continuity of curves. In order to also speak about velocities, one adds differentiable structure compatible with the topology in a sense that continuous and differential structure are in compliance with each other. Finally, one can define a metric field from which it is possible to derive causal order. CST “weakens" compatibility between topology and differentiable structure by imposing a different topology, called Alexandrov topology, determined by a causal structure. It is defined as the smallest topology in which chronological

[10]This last point is an open question, this is why this is a question of three or four infinities.

future and past[11] of some set are open. As a result, it is coarser (smaller) then manifold topology. Thus, the differentiable structure is kept, but the topology is changed [72]. Moreover, since the light cones can be defined in causal terms in (continuous) Lorentizan spacetime — event $p \in J^+(q)(p \in J^-(q))$ is said to be in causal future (past) of q if there is a future (past) directed causal curve (timelike or null) from p to q [71], then it is also possible to recover metric.

To sum up, the “standard" approach is to impose enough structure on a set to be able to speak about tensor fields, one of which is a metric field which provides causal structure. The shift from standard to CST point of view is that causal relations are considered as primary. Causal structure exists without the metric field [73].

Now that we understand the motivation, let's turn to formalizing this idea, starting from kinematics.

Kinematics Causal set (or causet for short) $(C, \prec)$ is locally finite partially ordered set (poset). In other words, it is defined as set C with a binary relation $\prec$ called “precedes", i.e., “is in the past of", with the following properties

1. Transitivity — if $x \prec y$ and $y \prec z$, $\forall x, y, z \in C$
2. Non-circularity — if $x \prec y$ and $y \prec x$ then $x = y \in C$
3. Local finitness — $(\forall x, z \in C)$ $(\text{card}\{y \in C | x \prec y \prec z\} < \infty)$

where “card" stands for cardinality of the set, number of it's elements. That is, every order-interval has finite cardinality. Local finitness is a formal way of saying that a causet is discrete, while partial order implies that not every pair of elements are related. Sometimes, instead of non-circularity one requires irreflexivity, $x \not\prec x$, meaning that no element precedes itself, which together with transitivity implies non-circularity, i.e., there are no cycles $x_0 \prec x_1 \prec ... \prec x_n = x_0$. The relationship $x \prec y$ is described by saying “x precedes y", or x is an ancestor of y, and y is descendant of x. Similarly, it is said that x lies in the past of y, or y in the future of x. The structure of a causet as represented on so-called Hasse diagram, Fig. 3.1, where elements of C are represented as vertices, while relationships are represented as edges. If the elements are related, they are connected with a line. If there are no elements between p and

[11]Chronological future (past) of event $p \in M$, $I^+(p)(I^-(p))$ is defined as the set of point which can be connected to p by future (past) directed timelike curve [71].

q, and $p \prec q$, we say there is a link between p and q. It should be pointed out that link is just an order of relation and has no length. Usually, not all relationships are drawn, only those not implied by transitivity. Another way of thinking about causets is as a family tree, which is clear from the jargon used.

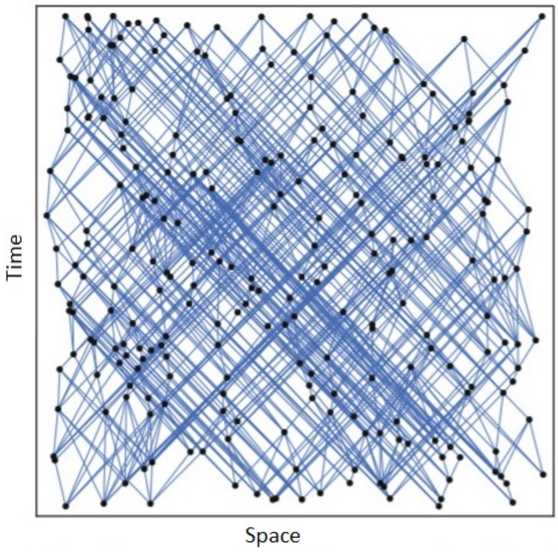


Figure 3.1: The Hasse diagram of a causet generated by sprinkling into $1+1$ dimensional Minkowski spacetime. The elements are black dots, while the blue edges are links, the nearest neighbour relations [69]

As already explained, the causal structure is almost enough to provide all the geometrical structure, since light cones, which determine the metric up to conformal transformation, can be defined in the causal terms, i.e., J^+. What is missing is the scale which fixes the conformal factor, i.e., volume. One cannot obtain the volume from the causal order, but one can postulate that a finite volume of spacetime contains a large but finite number of elements that measure the volume $\sqrt{-g}d^4x$ of a region in spacetime. That is, number of points in some region is equal to volume of that region $N = V$. This is true only is the proportionality factor is set to unity, so in general $N = \nu V$, and we must determine the proportionality factor, i.e., the value of the unit volume. One way is to consider BH entropy given by $S = \frac{A}{4G\hbar} = 2\pi\frac{A}{\kappa\hbar}$, $\kappa = 8\pi G$. According to this relationship, one bit of entropy belonging to fragments of horizon of roughly the size $l_p = \kappa\hbar$, reflecting underlying discreteness, so $\nu = l_p^{-4}$. In short, order and volume provide geometry [74].

To sum up, underlying structure of spacetime consists of causet, with volume as a number of elements in a region and macroscopic causality as reflection of order of the causet. An important remark is that since causal order plays such an important role, for a smooth manifold to be a realization of causal set, the macroscopic metric field ought to allow one to distinguish between past and future. This rules out Riemannian metric for example.

The next goal is to relate the picture of causal set to a manifold with metric, if such exists.

We begin from causet C with a large number of elements. A faithful embedding $f : C \to M$ of causal set into manifold M with a time oriented Lorentzian metric g without null or timelike closed curves is defined if the following conditions are satisfied:

1. $f(x) \in J^{-}(f(y))$ iff $x \prec y$, where $J^{-}(p)$ is the set of points of M of the causal past of p,

2. Embedded points are distributed uniformly with unit density, where uniform means that the points are realized by the Poisson process — the probability of finding a point in the region of finite volume depends only on the volume of the region, ensuring that the manifold contains a number of points of C equal to its volume,

3. Characteristic length over which the continuous geometry varies is everywhere much greater than the mean spacing between embedded points.

The first condition makes sure that causal relations induced by embedding agree with those of C. In the second condition, points are embedded with unit density if $\int \sqrt{-g} d^4x$ counts the elements of C. What's more, the reason the points are realized by Poisson process is to preserve Lorentz invariance. Were the causal set in Fig. 3.1 displayed in a different frame, the elements would have different positions but the distribution would still be uniform with respect to the volume, which is Lorentz invariant, along with the causal order of points. To explain why this is so, consider the case where the spacetime is flat. Here, the volume element is equal to Euclidean volume element d^4x. Lorentz invariance of is a consequence of the fact on the one hand, Lorentz transformations preserve the volume, and on the other hand, the Poisson process is invariant under volume preserving map in $\mathbb{R}^n$. This would not be the case were the points "sprinkled" in a shape of cubic lattice for example. The boosted frames would view different distribution of elements. The third condition is needed so that setting of the conformal factor to unity does not introduce unreasonably large curvature or similar small characteristic lengths, where small means the size of unity or smaller.

It should be pointed out that (M, g) in which one can embed C faithfully needs not exit, but if it does, one expects it is unique. This is referred to as *Hauptvermutung*, meaning conjecure. It is the idea that if two continua are good approximations of the same discretum, they should be close to each other, $M_1 \approx C$ and $M_2 \approx C$, then $M_1 \approx M_2$, where the last relation means that M_1 and M_2 are “approximately isometric". The quotation marks indicate that it is hard to give a rigorous meaning to what this means [74][25].

The explained process of embedding is called “sprinkling", since from the point of view of M, causal set C looks like it is obtained by sprinkling in points until Planckian density is reached [75].

As kinematics is not complete without dynamics, we now give some account of dynamics of causal set.

Dynamics Usually, dynamical law is specified by a Hamiltonian, since it is a generator of the time evolution. Because this presumes the existence of continuous time, thinking in terms of Hamiltonian is not a good start. However, in the broader sense, what we are looking for when speaking of dynamics is to prescribe the evolution. In the case of causal set, evolution is considered as stohastic growth of the causal set, described by probabilities in classical sense [70].

Dynamics of causal sets is called *sequential growth* — an element of causal set comes into being as “offspring" of a definite set of existing elements — defined by giving transition probability for each element of a causet, from it to each of its possible children. The growth of the causet is the passage of time. This also implies that there is no meaningful order of birth of the elements other than that implied by $\prec$. However, the elements are treated as if they happened in a definite order with respect to some fictitious “external time", represented by labeling of the elements by integers. The labels are considered not to carry any physical significance so calls this “discrete general covariance", as it is analogue to independence of Einstein$-$Hilbert action of the choice of coordinates. Moreover, as a consequence of discrete general covariance, the growth process is Markovian — the future of the growth process is independent of the past, and transition probabilities obey the Markov sum rule — the sum of transition probabilities emerging from a given causet is equal to unity. One further restriction, without which the possibilities of laws of growth would be infinite is Bell causality, which captures the intuition that a birth taking place in one region cannot be influenced by the other births occurring in regions that are spacelike to the first. With these requirements on the dynamics one can obtain a probability for transition $C \to C'$ [70][74].

To relate to the previous discussion about how causality determines matter, it is worth to mention that a certain form of Ising-like state emerges in indirect way from the dynamical laws, showing once again that matter and gravitation are intimately related [76].

In summary, considering causal order of elements of a set as fundamental, with number of elements in any region of spacetime corresponding to volume, one can obtain obtain geometric structure of Lorentzian spacetime, providing a discrete view of spacetime, which is a step toward quantum nature. Although there are some attempts, the theory needs to be further develop to see whether it can solve the problems from the beginning of the chapter.

4 Weinberg-Witten theorem

Weinberg−Witten (WW) theorem [77] constrains the spin[12] j of (composite or elementary) massless particles, charged under conserved Lorentz covariant four-vector currents[13]. As formulated by Weinberg and Witten, it consists of two parts and states that

(a) A theory that allows construction of conserved Lorentz covariant four-vector current J^μ *cannot contain massless particles of spin* $j > \frac{1}{2}$ *with nonvanishing values of conserved charge* $Q \equiv \int J^0 d^3x$*,*

(b) A theory that allows construction of conserved Lorentz covariant energy-momentum

[12]To be precise, when speaking of massless particles spin refers to helicity, projection of spin onto direction of momentum.

[13]It should be pointed out that objects carrying indices will be referred to as vectors, or tensor, irrespective of their transformation law under general coordinate transformations, contrary to the usual definition. Hence, one should not assume anything about transformation properties of an object solely based on the name.

tensor $T^{\mu\nu}$ for which $P^\nu \equiv \int T^{0\nu} d^3x$ is the conserved energy-momentum four-vector cannot contain massless particles of spin $j > 1$.

The goal of the next few sections is to unpack the key elements of the theorem —

- Definition of spin and massless particles,
- Conserved currents and respective charges,
- Lorentz covariance of the conserved currents,

and ultimately prove it. It will also be shown how the constraints of the theorem apply to Standard Model, currently most accurate theory describing known particles and interactions between them, and to gravitation, i.e., graviton, mediator of gravitational force, which is at present hypothetical particle.

4.1 Mass and spin

In this section we show that elementary relativistic particles described as physical states characterized by mass and spin arise from irreducible representation of isometry group of flat spacetime — Poincaré group. First, it is explained how isometries relate to quantities describing physical system. Next, it is shown that mass and spin are properties associated with Poincaré group by considering its algebra.

What one has in mind when speaking of describing a system is finding solutions of the field equations which determine its dynamics. However, instead of solving the equations directly, the solutions can be obtained by finding the group of spacetime transformations that transform the equations so that their functional form stays the same, i.e. covariantly. Since the transformation didn't change the equations the solution of the old and the new equation belong to the same equation. In other words, the elements of the group generate new solutions from the old ones. Finally, in certain cases, the group narrows down to symmetry group of spacetime on which the theory is formulated. This will become clear in the next chapter which is dedicated to symmetries. To give some explanation on how this happens, with details in the next chapter, the special case mentioned in the text refers to theories in which metric field is held fixed, i.e., it's dynamics is considered as given. Subsequently, any spacetime transformation changing the field equations covariantly must also be an element of isometry, group that leaves the metric field invariant. For the case when flat metric is fixed the isometries are elements of Poincaré group.

Furthermore, it should be pointed out that by studying symmetry group we are dealing with all field equations with a certain symmetry, since any theory on the corresponding fixed background will have the same symmetry group. Distinguishing between solutions belonging to a certain equation is possible with the help of invariants of the group, as all solutions generated by the symmetry group share the value of property corresponding to the invariants. In fact, it makes more sense to turn the argument around. Existence of invariants implies one can differentiate between different types of solutions, where each type describes a different entity, with its own field equations whose solutions are generated by the symmetry group. In short, one can use the invariants to classify different types of solutions. What's more, it is the elements of the symmetry group, or more precisely, its generators that correspond to measurable quantities using which we can characterize the solutions. We now show how the discussion applies to quantum systems.

For quantum systems, such as relativistic particles, solutions are fields, operator valued-spacetime functions. They don't have a direct physical interpretation but they can be mapped to space of physical states, so a solution is represented by a state in the state space. Thus, finding the solutions consists of determining all states in which the particle can be found, i.e., its space of physical states. As already mentioned, on flat spacetime the space is generated by Poincaré group. As it will be shown below, invariants of the group are mass and spin. As a result, there are different types of relativistic particles living on flat spacetime, distinguished by mass and spin.

Before we show this below, as a final remark note that even though the fields don't have a physical interpretation, they are very useful. Relationship between fields and particles, i.e., their states, will be given at the end of the chapter as it is important in the following discussion.

To start, states of a quantum system span a Hilbert space, with each state represented as state vector. According to the previous discussion it is generated by the symmetry group G, or more precisely its representation U,

$$g \in G \xrightarrow{U} U(g)\,, \tag{4.1}$$

since group acting on vector space are defined to be given by representation, a mapping U, where every element of the group is represented by $U(g)$, an operator on the state space[14]. For finite n dimensional representations it can realized as $n \times n$ matrix $D(g)$

$$U(g)|\psi_i\rangle = D(g)_i^j|\psi_j\rangle\,, \tag{4.2}$$

where $\{|\psi_i\rangle\}$ form a basis of the state space. One can choose any basis, but as it will be shown below, there is a basis naturally related to measurable quantities which uniquely determine the state.

Next, physical states transform under unitary operators, as shown by Wigner, making sure the normalization of states is preserved under transformations. To understand what unitarity condition entails, let $|\psi\rangle$ and $|\psi'\rangle$ be two possible states such that $|\psi'\rangle = U(g)|\psi\rangle$. The unitarity makes sure that

$$|\langle\psi|\psi\rangle|^2 = 1 \implies |\langle\psi'|\psi'\rangle|^2 = 1\,. \tag{4.3}$$

As it is known, absolute square of scalar product of two states is called transition probability. Scalar product of the state with itself will gives the probability that the state exists at all. So, starting in state $|\psi\rangle$, we require that the state obtained by a symmetry transformation $U(g)$, also exists[15] which translates to $|\psi'\rangle$ being obtained from $|\psi\rangle$ by a unitary operator $U(g)$. Moreover, Poincaré group is a Lie group so every representation of the group is also a representation of the algebra and vice versa. Consequently, for the most part we deal with generators of the algebra which are hermitian if the representation is unitary. This has important consequences. First, eigenvectors of hermitian operators constitute a basis of state space. Second, eigenvalues of hermitian operators are real and represent measurable quantities, which makes such basis naturally suited for describing space of states.

What's more, there may exist operators, constructed out of generators, that commute with all the generators of the group. We refer to them as Casimir operators.

[14]The vector space on which representation acts on is called representation space. We will use the terms state space and representation space interchangeably, as the representation space in the current context is the state space.

[15]In relativistic mechanics the particle can decay. Thus, we are assuming that one is dealing with a free particle, there is no interaction.

They are the invariants of the group, because the fact they commute with all the generators entails the properties they describe don't change under the action of the group. Each set of states sharing the same Casimir eigenvalue represents one type of solutions. As a result, the state space is partitioned into blocks, i.e., subspaces invariant under Casimir operators, where invariant means that all state vectors belonging to a block transform under action of the group only among themselves, transforming the space into itself, i.e., the space is invariant. What's more, as a consequence of Schur's lemma, each such subspace is minimal invariant subspace, i.e., it contains no further invariant subspaces besides the trivial ones. Representation on such space is referred to as irreducible and serves as a building block, with eigenvalues of Casimir operators used to label it. Even more, states belonging to minimal invariant subspace are associated with elementary particles, stemming from the notion that a fundamental entity should be "indivisible"[16]. Accordingly, its states belong to the smallest, indivisible subspace of state space of relativistic particles. Composite particles are then associated with representations built from irreducible ones.

In summary, state space of free, elementary relativistic particles is generated by action of group elements represented as unitary operators. The invariants of the group, Casimir operators, organize the state space into minimal invariant subspaces, such that in each subspace operators $U(g)$ form an irreducible representation of the symmetry group[17] [78][79][80].

Reminders of important properties of Poincaré group and Poincaré algebra are given in App. B and App. C respectively. We now turn our attention to showing how the basis of state space is generated using the method of induced representation.

4.1.1 Irreducible unitary representation of Poincaré group

In this chapter we find a basis spanning minimal invariant subspace of representation space of Poincaré group and the matrix elements of the unitary representation. As the group is non-compact its irreducible representations are infinite dimensional. This is because unitarity of representations for compact Lie groups follows from finite dimensionality. However, one can, at least partly, avoid working with infinitesimal

[16]What is considered fundamental, i.e., indivisible can depends on energy scale. For example, proton or neutron are often treated as point particles, although they consist of quarks. Elementary particles are considered to be as fundamental as it gets.

[17]It should be pointed out that the term irreducible representation often refers to its representation space rather then the map.

representation. As it will be shown, since Poincaré group has an Abelian invariant subgroup – the group of translations, entire representation of the Poincaré group can be constructed from little group of appropriately, i.e., easy to work with, momentum eigenvector. Little group consists of elements which leave the momentum eigenvalue invariant. Furthermore, it turns out to be compact, so it is finite dimensional. Entire irreducible representation, labeled by Casimir operators C.15, can be obtained by application of the rest of the generators. The procedure is known as the method of induced representation [81].

Since we know translation subgroup is Abelian from C.11, let's show it is also invariant subgroup. According to definition, we must check that $(\Lambda, b)(\mathbb{1}, a)(\Lambda, b)$ produces another element of translation subgroup. This can in fact be seen from C.5, but we deal with it now in more general terms. To begin with, note that as a consequence of multiplication rule B.2, general element of Poincaré group can be written as Lorentz transformation followed by translation,

$$(\Lambda, a) = (\mathbb{1}, a)(\Lambda, 0)\,. \tag{4.4}$$

As a result,

$$\begin{aligned}(\Lambda, b)(\mathbb{1}, a)(\Lambda, b)^{-1} &= (\mathbb{1}, b)(\Lambda, 0)(\mathbb{1}, a)(\Lambda, b)^{-1}(\mathbb{1}, b)^{-1}\\ &= (\mathbb{1}, b)(\mathbb{1}, \Lambda a)(\mathbb{1}, b)^{-1}\\ &= (\mathbb{1}, \Lambda a)\,,\end{aligned} \tag{4.5}$$

This completes the proof. As representation is a homomorphism by definition, the group structure remains preserved.

Next, we choose as basis vectors of state space the eigenvectors of momentum. This is possible because all components of momentum mutually commute. Further, due to existence of Casimir operators, the action of generators is partitioned into minimal invariant subspaces, which amounts to studying irreducible representations. As a result, domain of the eigenvector is also restricted. To explain this, let's study one such subspace invariant under the Casimir operators, starting with C_1. It must be true that

$$C_1(p) = C_1(\Lambda p)\,, \tag{4.6}$$

since for an operator to commute with translation generator it must be a function of momentum. Then, for it to also commute with generator of Lorentz transformation is must be of the form of the Lorentz invariant product, which leads to C.15, and in turn to 4.6. Then, since Casimir operators are multiples of the unity for irreducible representations we have

$$C_1|p,\zeta\rangle = m^2|p,\zeta\rangle\,, \tag{4.7}$$

where m^2 is some constant and ζ refers to possible additional labels. Conditions 4.6 and 4.7 can be met only if $|p,\zeta\rangle$ is different from zero only for momenta p which can be obtained from each other by Lorentz transformations $p' = \Lambda p$. The goal is to show that each such irreducible representation can be obtained from a single momentum eigenvector. However, one should note that there are different classes of standard momentum, which categorizes irreducible representations into four classes,

- $P^2 > 0$,
- $P^2 = 0$, $P = 0$
- $P^2 = 0$, $P \neq 0$
- $P^2 < 0$.

Thus, to be precise, all irreducible representations of a certain class can be obtained from one standard vector. Null vector, $P^2 = 0$ with $P = 0$ is a separate case because such states are invariant under all Poincaré transformations, which is not true for states of any other class. Consequently, representation of the second class cannot be obtained in any other class.

Moreover, in unitary representation generators are hermitian operators. Since $C_1 = P^2$ is constructed out of hermitian operators, C_1 itself is hermitian. As a consequence its eigenvalue m^2 must be real. Due to relativistic energy-momentum relation m^2 is interpreted as mass. Finally, as there is an infinite number of p^μ satisfying $p^2 = m^2$ for each m^2, irreducible representations are infinite dimensional. Casimir operator $C_2 = W^2$ is more complicated, so it will be discussed later and we will focus only on this classification for now. Lastly, states with $m^2 > 0$, corresponding to massive particles, belong to different representation from massless $m^2 = 0$ particles.

We will consider only these classes as only they are physical. We now return to the question of the basis vectors.

Mutually commuting operators share the eigenspace. As a result, according to commutator relation C.16-C.18, one of the components of Pauli-Lubanski pseudovector can be chosen as label which determines state vector uniquely. It remains to determine the meaning of Pauli-Lubanski pseudovector and show how the irreducible representation is generated from standard vector. Taking into consideration the discussion above, let's start with massive representations $m^2 > 0$.

We choose the standard momentum as the one with eigenvalue $p = (m, 0, 0, 0)$, describing particle in its rest frame. Dimension of space belonging to the rest frame momentum is determined by existence of generators which leave the rest momentum invariant, i.e., its little group. Here, the Pauli-Lubanski components reduce to

$$W^0 = \boldsymbol{P} \cdot \boldsymbol{J} = 0\,, \tag{4.8}$$

$$\boldsymbol{W} = -P^0 \boldsymbol{J} + \boldsymbol{P} \times \boldsymbol{K} = -P^0 \boldsymbol{J} = m \boldsymbol{J}\,, \tag{4.9}$$

Thus, in the rest frame Pauli-Lubanski pseudovector has only spatial components, proportional to angular momentum. As a consequence,

$$[W_\mu, W_\nu] = i\epsilon_{\mu\nu\rho} W^\rho\,, \tag{4.10}$$

from which one may recognize that the little group is $SO(3)$. This is true for all massive representations. It should be noted that angular momentum in rest frame can only come from spin. What's more, Casimir operator C_2 is in the rest frame equal to

$$C_2 = W^2 = -m^2 \boldsymbol{J}^2\,, \tag{4.11}$$

Thus, eigenvalues of C_2 represent spin, and W_i is its projection. Moreover, since C_2 is Poincaré invariant it can be evaluated in any frame. The value of spin of massive particles is the same in any state within the representation. In conclusion, this shows that mass and spin are concepts related to symmetry of flat spacetime given by Poincaré group.

Finally, the basis vectors of the subspace corresponding to $p = (m, 0, 0, 0)$ are determined as

$$
\begin{aligned}
P^\mu|m,j;\mathbf{0},j_3\rangle &= p^\mu|m,j;\mathbf{0},j_3\rangle \\
W_3|m,j;\mathbf{0},j_3\rangle &= -mj_3|m,j;\mathbf{0},j_3\rangle
\end{aligned}
\tag{4.12}
$$

where component of Pauli-Lubanski pseudovector we have chosen is W_3, projection of spin onto z axis, as it is common in quantum mechanics. From there we also know that the rest of basis vectors of the little group is obtained by acting with raising and lowering operators $J_\pm$. In addition, dimension of subspace belonging to momentum eigenvector is $2j+1$. The first two labels refer to the representation and are usually left out.

Elements of little group mix only the states with the same value of momentum. To generate entire irreducible invariant space one must act with transformations which produce new eigenvalue of momentum. Let's start by considering pure Lorentz boost defined as

$$
U(L(p)) \equiv [R^{-1}(p)]'' L_{z'}(p) R(p)\,, \tag{4.13}
$$

where $R(p)$ rotates z axis, projection axis of spin, of the rest frame into direction of p. The new frame obtained by such rotation is denoted by primes. The z' axis is then boosted in direction of p. Finally, inverse rotation $[R^{-1}(p)]''$, i.e., rotation in the opposite direction of $R(p)$ but by the same amount, maps z' to z'', which is parallel to z. Hence, the action on standard state vector results in

$$
U(L(p))|m,j;\mathbf{0},j_3\rangle = |m,j;\boldsymbol{p},j_3\rangle\,. \tag{4.14}
$$

Pure Lorentz boosts transform the particle into state with arbitrary momentum but same projection of spin onto z axis, i.e., the boosted frame has the same orientation as the initial one.

In summary, one can generate basis vectors of representation space by acting with $J_\pm$ and $U(L(p))$ on $|m,j;\mathbf{0},j_3\rangle$. It is left to prove that vectors obtained in such way span irreducible invariant subspace of the Poincaré group. To start, note that the vector space spanned by $\{|m,j;\boldsymbol{p},j_3\rangle\}$ must be irreducible as it's generated from one single vector by above explained procedure. To show that the space is invariant one must show that vectors $\{|m, j; \boldsymbol{p}, j_3\rangle\}$ for fixed value of m and j transform only among themselves under Poincaré transformations. Due to 4.4 we can consider translations and Lorentz transformations separately.

Starting with translations, the first step is to show that $|m, j; \boldsymbol{p}, j_3\rangle$ is an eigenstate of P^μ,

$$
\begin{aligned}
P^\mu |m, j; \boldsymbol{p}, j_3\rangle &= P^\mu U(L(p)) |m, j; \boldsymbol{0}, j_3\rangle \\
&= U(L(p)) U(L(p))^{-1} P^\mu U(L(p)) |m, j; \boldsymbol{0}, j_3\rangle \\
&= U(L(p)) U(L(p))^\mu_\nu P^\nu |m, j; \boldsymbol{0}, j_3\rangle \\
&= U(L(p)) |m, j; \boldsymbol{0}, j_3\rangle L(\boldsymbol{p})^\mu_\nu p^\nu \\
&= p^\mu |m, j; \boldsymbol{p}, j_3\rangle \,.
\end{aligned}
\tag{4.15}
$$

The first equality follows from 4.14, the second by inserting $\mathbb{1} = U(L(p))U(L(p))^{-1}$, and the third from C.5. Note that C.5 follows from 4.5, so it is true because the translation subgroup is invariant subgroup. Thus, matrix element of irreducible representation is

$$
e^{-iaP} |m, j; \boldsymbol{p}, j_3\rangle = e^{-iap} |m, j; \boldsymbol{p}, j_3\rangle \,. \tag{4.16}
$$

This follows directly from 4.15. Since we are dealing with abstract Hilbert space we cannot write down the explicit expression for the momentum generator, but it is hermitian so that the representation is unitary. As expected, because the translation subgroup is Abelian, its irreducible representation is one-dimensional. What's more, under translations, states are invariant up to a phase, i.e., under translations the basis vectors are mapped into themselves.

Next, we turn to Lorentz group. Consider

$$
\begin{aligned}
U(\Lambda) |m, j; \boldsymbol{p}, j_3\rangle &= \sum_{j_3'} \mathcal{D}^{(j)}_{j_3' j_3} \left(R_W\right) L(\Lambda p) |m, j; \boldsymbol{0}, j_3'\rangle \\
&= \sum_{j_3'} \mathcal{D}^{(j)}_{j_3' j_3} \left(R_W\right) |m, j; \boldsymbol{p'}, j_3'\rangle \,,
\end{aligned}
\tag{4.17}
$$

with $p'^{\mu} = \Lambda^{\mu}_{\nu} p^{\nu}$. This stems from

$$
\begin{aligned}
U(\Lambda)|m, j; \boldsymbol{p}, j_3\rangle &= U(\Lambda)\mathrm{L}(p)|m, j; \mathbf{0}, j_3\rangle \\
&= U(\Lambda L(p))|m, j; \mathbf{0}, j_3\rangle \\
&= U(\mathrm{L}(\Lambda p) \underbrace{\mathrm{L}^{-1}(\Lambda p)\Lambda \mathrm{L}(p)}_{\equiv R_W})|m, j; \mathbf{0}, j_3\rangle \,.
\end{aligned}
\tag{4.18}
$$

The first line is obtained using 4.14. The second line follows from the composition rules and the third by inserting identity $\mathrm{L}(\Lambda p)\mathrm{L}^{-1}(\Lambda p) = \mathbb{1}$. Wigner rotation R_W is an element of little group since it transforms the momentum so that $\mathring{p} \to p \to \Lambda p \to \mathring{p}$, where $\mathring{p}$ is the rest momentum. As a result $\mathcal{D}^j(R_W)$ is unitary representation matrix of $SO(3)$ group corresponding to eigenvalue j, called Wigner D-matrix. Thus, we have shown that under Lorentz transformations state vectors $\{|m, j; \boldsymbol{p}, j_3\rangle\}$ transform only among themselves.

In summary, space of states belonging to particle of mass m and spin j is obtained from unitary irreducible representation of Poincaré group. The space is infinite dimensional as there are infinitely many p^{μ} obeying $p^2 = m^2$, with elements of Poincaré group permuting them along the mass shell. To reflect on the discussion from the beginning, equation that is transformed under elements of the Poincaré group yields a state that is either rotated, boosted or translated with respect to the old state. All these states share the same value of mass and spin, i.e., describe the same particle in different state.

We now turn to massless particles. The consideration follows the same steps, with the only difference being the little group. As standard vector we choose the one with eigenvalue

$$
p_0 = (E, 0, 0, E) \,. \tag{4.19}
$$

One can immediately expect that the little group contains $SO(2)$ group. To determine it, notice that for 4.19

$$
W_{\mu} P^{\mu} = \frac{1}{2} \varepsilon_{\mu\rho\sigma\nu} M^{\rho\sigma} P^{\nu} P^{\mu} \,. \tag{4.20}
$$

Right-hand side is zero because Levi-Civita is antisymmetric, while the product of momentum is symmetric. Expanding the left-hand side,

$$EW^0 - EW^3 = 0 \implies W^0 = W^3 = EJ_3 \tag{4.21}$$

The rest of components are

$$\begin{aligned} W_1 &= E(K_2 - J_1)\,, \\ W_2 &= -E(K_1 + J_2)\,. \end{aligned} \tag{4.22}$$

The algebra of the little group is

$$\begin{aligned} [W_1, W_2] &= 0\,, \\ [W_2, J_3] &= iW_1\,, \\ [W_1, J_3] &= -iW_2\,. \end{aligned} \tag{4.23}$$

These are commutation relations of algebra of Euclidean group in two dimensions $E2$ consisting of rotation specified by angle θ and translations specified by parameters δ and η. General element $g \in E2$ can be written as

$$g(\delta, \eta, \theta) = T(\delta, \eta)R(\theta)\,, \tag{4.24}$$

where $T(\delta, \eta) \equiv g(\delta, \eta, 0)$ refers to translations and $R(\theta) \equiv g(0, 0, \theta)$ to rotations. For four-vector one can express the transformations as

$$R(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.25}$$

and

$$T(\delta, \eta) = \begin{pmatrix} 1+\zeta & \delta & \eta & -\zeta \\ \delta & 1 & 0 & -\delta \\ \eta & 0 & 1 & -\eta \\ \zeta & \delta & \eta & 1-\zeta \end{pmatrix} \tag{4.26}$$

with $\zeta = (\delta^2 + \eta^2)/2$. In fact, one can conclude from 4.23 that W_1 and W_2 are generators of translations as their commutation relations are the same as those of momentum C.11, while J_3 is generator of rotations. Furthermore, the multiplication rule is

$$g(\delta, \eta, \theta)g(\delta', \eta', \theta') = g(R(\theta)\delta + \delta', R(\theta)\eta + \eta', \theta + \theta')\,, \tag{4.27}$$

so the inverse is $g(-R(-\theta)\delta, -R(-\theta)\eta, -\theta)$. One more property we will need in the following derivation is the fact that group of translations is invariant Abelian subgroup of $E2$,

$$\begin{aligned} g(\delta, \eta, \theta)T(\delta', \eta')g(\delta, \eta, \theta)^{-1} &= T(\delta, \eta)R(\theta)T(\delta', \eta')R(-\theta)T(-\delta, -\eta) \\ &= T(\delta, \eta)T(R(\theta)\delta', R(\theta)\eta')T(-\delta, -\eta) \\ &= T\left[R(\theta)\delta', R(\theta)\eta'\right]\,. \end{aligned} \tag{4.28}$$

Returning now to the little group, note that besides momentum, one can also label the states with eigenvalues of W_1 and W_2 since they mutually commute according to 4.23. Hence, the states are defined as

$$\begin{aligned} P^\mu|p_0, w_1, w_2\rangle &= p_0^\mu|p_0, w_1, w_2\rangle\,, \\ W_1|p_0, w_1, w_2\rangle &= w_1|p_0, w_1, w_2\rangle\,, \\ W_2|p_0, w_1, w_2\rangle &= w_2|p_0, w_1, w_2\rangle\,. \end{aligned} \tag{4.29}$$

It remains to determine the meaning of W_1 and W_2. Let's begin with the eigenvalues. Because translation subgroup is an invariant subgroup according to 4.28, $U(R)|p, w_1, w_2\rangle$ is also an eigenstate of W_i. So, we start by determining what $U(R(\theta))^{-1}W_iU(R(\theta))$ is equal to. One can follow the same procedure as when deriving Poincaré algebra in C.5, but there is a shortcut. As it can be seen from the multiplication rule, rotation group is additive. Furthermore, as the group only has one generator, the group is Abelian. So, just as we derived for momentum, we have for rotations

$$U(R(\theta)) = e^{-i\theta J_3}\,. \tag{4.30}$$

The representation is unitary because J_3 is hermitian. Thus, $U(R(\theta)^{-1}W_iU(R(\theta))$ is equal to

$$f(\theta) \equiv e^{-i\theta J_3} W_1 e^{i\theta j_3}\,, \tag{4.31}$$

$$g(\theta) \equiv e^{-i\theta J_3} W_2 e^{i\theta j_3}\,, \tag{4.32}$$

where we defined $f(\theta)$ and $g(\theta)$, operator valued functions. Taking derivative with respect to θ and using the commutator relations 4.23, one obtains

$$\frac{df}{d\theta} = -g(\theta)\,, \tag{4.33}$$

$$\frac{dg}{d\theta} = f(\theta)\,, \tag{4.34}$$

which can be solved by differentiating one of the equations and inserting it into the other, with initial conditions $f(\theta = 0) = W_1$ and $g(\theta = 0) = W_2$, to be in accordance with 4.29. The result is

$$f(\theta) = W_1 \mathrm{cos}\theta - W_2 \mathrm{sin}\theta\,, \tag{4.35}$$

$$g(\theta) = W_1 \mathrm{sin}\theta + W_2 \mathrm{cos}\theta\,. \tag{4.36}$$

Thus, eigenvalues of $U(R(\theta))|p_0, w_1, w_2\rangle$ are

$$\begin{aligned} W_1 U(R(\theta))|p, w_1, w_2\rangle &= U(R(\theta))U(R(\theta))^{-1} W_1 U(R(\theta))|p_0, w_1, w_2\rangle \\ &= (w_1 \mathrm{cos}\theta - w_2 \mathrm{sin}\theta) U(R(\theta))|p_0, w_1, w_2\rangle \\ W_2 U(R(\theta))|p_0, w_1, w_2\rangle &= U(R(\theta))U(R(\theta))^{-1} W_2 U(R(\theta))|p_0, w_1, w_2\rangle \\ &= (w_1 \mathrm{sin}\theta + w_2 \mathrm{cos}\theta) U(R(\theta))|p_0, w_1, w_2\rangle \end{aligned} \tag{4.37}$$

Spectrum of W_1 and W_2 is continuous, as there is an eigenvalue for each value of θ. This was expected as they generate translations. It turns out however, that particles do not have properties corresponding to such continuous degrees of freedom. For example, photon is massless particle of spin-1, with two possible polarizations, i.e., only discrete eigenvalues. This is why we limit physical states by requiring $W_1 = W_2 = 0$. Thus, the little group is reduced to only one generator, $J_3 = \frac{1}{E} W_3$ and the group is $SO(2)$. To determine eigenvalues of W_3 note that,

$$C_2 = (W^0)^2 - (W^3)^2 = 0\,. \tag{4.38}$$

From the fact that Pauli-Lubanski pseudovector has only zeroth and third component, which are equal, and because $W^2 = C_2 = 0$, the Pauli-Lubanski pseudovector must be proportional to 4.19, i.e.,

$$W^\mu = \lambda P^\mu \implies \lambda = \frac{\boldsymbol{JP}}{P^0}\,, \tag{4.39}$$

where the last line follows from 4.9. One may recognize λ as helicity, projection of spin onto momentum of massless particle. Unlike spin, one can show that helicity commutes with generators of Poincaré group, i.e., it is Poincaré invariant. This is because massless particles always travels at the speed of light, so one cannot boost to the frame where direction of momentum would be reversed. Furthermore, helicity can have only two values, depending on whether spin is parallel or antiparallel to direction of motion. As a result, eigenvalues of W_3 are $\lambda = \pm j$. Finally, basis vectors of little group are labeled as

$$P^\mu |m=0, j=0; \boldsymbol{p}_0, \pm j\rangle = p^\mu |m=0, j=0; \boldsymbol{p}_0, \pm j\rangle \tag{4.40}$$

$$W_3 |m=0, j=0; \boldsymbol{p}_0, \pm j\rangle = \pm j |m=0, j=0; \boldsymbol{p}_0, \pm j\rangle \tag{4.41}$$

As label, one may choose the projection with positive or negative sign. Starting from any of them will generate the basis of irreducible invariant subspace of massless particles, related by parity transformation. We now generate the rest of the basis vectors. As rotation group in two dimensions is Abelian, the subspace belonging to 4.19 is one-dimensional, while general momentum can be obtained by applying

$$|m=0, j=0; \boldsymbol{p}, \pm j\rangle = R(\phi, \theta, 0) U(L_0(p)) |m=0, j=0; \boldsymbol{p}_0, \pm j\rangle\,, \tag{4.42}$$

where $L_0(p)$ boost the momentum to value $|\boldsymbol{p}| = e^{\tanh^{-1}(\beta)} E$ without changing its direction. Although massless particle always travels with the speed of light, it's energy depends on frequency. As a result, energy is an observer dependent quantity, so $U(L_0(p))$ changes the value of reference momentum, and the rotation $R(\phi)$ brings it to arbitrary direction specified by angles ϕ and θ.

Finally, following the same procedure as for massive case, one can show that $\{|\boldsymbol{p}, \pm j\rangle\}$ spans irreducible, invariant subspace of representation labeled by $m = 0$ and $j = 0$, so the derivation will not be repeated. The matrix elements of the representation are

$$U(\mathbb{1}, a)|\boldsymbol{p}, \pm j\rangle = e^{-iap}|\boldsymbol{p}, \pm j\rangle\,, \tag{4.43}$$

$$U(\Lambda, 0)|\boldsymbol{p}, \pm j\rangle = e^{\pm j i\theta(\Lambda,p)}|\Lambda\boldsymbol{p}, \pm j\rangle\,. \tag{4.44}$$

We have left out the labels of representation, C_1 and C_2 since they vanish. Angle $\theta(\Lambda, p)$ can be determined from

$$e^{\pm i\theta(\Lambda,p)} = \langle \boldsymbol{p}_0, \pm j|L_0(\Lambda p)^{-1}\Lambda L_0(p)|\boldsymbol{p}_0, \pm j\rangle\,. \tag{4.45}$$

As mentioned, helicity is Poincaré invariant quantity, so Lorentz transformation can only change the state up to phase [81][80][82].

To sum up, physical states of mass and spin labelled by momentum and spin or helicity arise from irreducible representations of Poincaré group, the symmetry group of flat spacetime.

4.2 Fields and particles

As it was explained, objects living on flat spacetime arise from some representation of Poincaré group. The particle states are physical states and as such must be associated with unitary representation. This however, is not the case for fields. They are operator valued spacetime functions that are not physical observables. As result, there is no reason to demand unitarity, so the field transform under finite dimensional representation. Thus, we briefly discuss how the finite dimensional representations of Lorentz group are characterized.

One can show that Lie algebra in basis consisting of M_i and N_i defined as

$$\begin{aligned} M_i &= \frac{J_i + iK_i}{2}\,, \\ N_i &= \frac{J_i - iK_i}{2}\,, \end{aligned} \tag{4.46}$$

reduces to direct product of two subalgebras,

$$
\begin{aligned}
[M_i, M_j] &= i\varepsilon_{ijk} M^k \,, \\
[N_i, N_j] &= i\varepsilon_{ijk} N^k \,, \\
[N_i, M_j] &= 0 \,.
\end{aligned}
\tag{4.47}
$$

One may recognize that this algebras belong to $SU(2) \cong SO(3)$. Hence, the is algebra of Lorentz group is equivalent to $SU(2) \times SU(2)$. Also, one can see from 4.46 that such generators will not result in unitary representation, since both M_i and N_i cannot simultaneously be hermitian. As a consequence, such representation cannot correspond to physical states. Moving on, the Casimir operators of the group in such basis are M^2 and N^2, whose eigenvalues are $u(u+1)$ and $v(v+1)$ respectively. So, one can label representation by (u, v), and the basis vectors are direct product of $|k, l\rangle, k = -u, ..., u$, and $|v, l\rangle, l = -v, ..., v$. With these vectors one can form a new basis $|j, m\rangle$ where $j(j+1)$ is eigenvalue of J^2, and m is eigenvalue of J_3, with $|u - v| \geq j \geq u + v$, so that we no longer have to distinguish between the two subgroups $SU(2)$.

As fields and physical states are in different representations, there is no automatic relation between them. Instead, the correspondence between field and particle is established by connecting the field with a particle state it creates or destroys. The goal is to show that a field transforming according to 4.48 describes a state of definite mass and spin.

$$
U(\Lambda)\Psi(x)^a U(\Lambda^{-1}) = D(\Lambda^{-1})^a_b \Psi^b(\Lambda x) \,,
\tag{4.48}
$$

where $U(\Lambda)$ is non-unitary representation of Lorentz group on Hilbert space where Ψ lives, and $D(\Lambda)$ is $n \times n$ matrix representation of Lorentz group. What's more, field is by definition a function of spacetime. As a result, translation generators act as[18]

$$
P_\mu \Psi^a = -i\partial_\mu \Psi^a(x) \,.
\tag{4.49}
$$

Thus, the translations are already taken into account, so we focus our attention to

[18] In the previous chapter we have worked in an “abstract" Hilbert space, i.e., we haven't stated what the vectors of Hilbert space are, which is also why we couldn't write down the explicit form of the generators. Here, we chose Hilbert space of position base of square integrable functions.

the Lorentz group.

The starting point in relating unitary irreducible representations of Poincaré group with solutions of the field equations Ψ is the field equation which Ψ satisfies. Since, as we have seen, fields transform under finite dimensional Lorentz representation (u, v), fields for both u and v both different from zero have multiple spin content. Then, to ensure fields correspond to particle states of some irreducible representation (which has only single spin) the differential equation should act as a projection matrix selecting out desired components of spin from Ψ. In short, field equation must impose enough Poincaré invariant conditions onto fields, without over constraining the solution space. By solving the field equations as usual, by Fourier transform, one can show that the solution space is equivalent to particle of spin-j propagating on flat spacetime.

Hence, to ensure the representation is irreducible, one should impose

$$(\partial^2 - m^2)\Psi = 0\,. \tag{4.50}$$

The differential equation, when converted into algebraic form using Fourier transform, converts to relativistic energy-momentum relationship,

$$\Psi^a(x) = \int \frac{d^4p}{(2\pi)^3}\Phi^a(p)e^{ipx}\,, \qquad \Longrightarrow \; (p^2 - m^2)\Phi(p) = 0\,. \tag{4.51}$$

With this condition we fix P^2. Next, to account for spin degrees of freedom, the equation should impose transversality $\partial_a\Psi^a = 0$, tracelessness $\eta_{ab}\Psi^{ab} = 0$, symmetricity $\Psi^{ab} = \Psi^{ba}$ etc. These conditions make sure we are dealing with simplest finite, irreducible representations of Lorentz group, i.e., that the fields have the same degrees of freedom as the particles. We will come back to this later.

The solutions of the equation are referred to as positive and negative energy solutions,

$$E = \pm(\boldsymbol{p}^2 + m^2)^{\frac{1}{2}}\,. \tag{4.52}$$

Finally, if $\varepsilon^a(\boldsymbol{p}, \lambda)$ are elementary solutions of the field equation, defined as

$$\varepsilon^a(\boldsymbol{p}, \lambda) = D(L(p))^a_b\varepsilon^b(\boldsymbol{0}, \lambda)\,, \tag{4.53}$$

the general solution is obtained by separating positive and negative types of solutions as a linear combination of the elementary solutions,

$$\Psi(x) = \sum_{\lambda} \int \frac{d^3p}{(2\pi)^3\sqrt{2E}} \left(a(\boldsymbol{p},\lambda)\varepsilon^a(\boldsymbol{p},\lambda)e^{ipx} + a^\dagger(\boldsymbol{p},\lambda)\varepsilon^{*a}(\boldsymbol{p},\lambda)e^{-ipx} \right) , \tag{4.54}$$

where $a(\boldsymbol{p},\lambda)$ are operator valued functions. We use λ to label helicity in the massless case and projection of spin in massive case. The Poincaré states we defined in the previous section can be written as

$$|\boldsymbol{p},\lambda\rangle = a^\dagger(\boldsymbol{p},\lambda)|0\rangle\,, \tag{4.55}$$

where $|0\rangle$ is a vacuum state. To shorten the notation we leave out the labels of the representation. From previous discussion, we know how the states should transform under infinite dimensional unitary representation of the Lorentz group. As a result, the creation operators should transform as

$$U(\Lambda)a^\dagger(\boldsymbol{p},\lambda)U(\Lambda^{-1}) = D^{(j)}(R(\Lambda,p))^{\lambda'}_{\lambda}a^\dagger(\Lambda\boldsymbol{p},\lambda')\,, \tag{4.56}$$

$$U(\Lambda)a(\boldsymbol{p},\lambda)U(\Lambda^{-1}) = D^{(j)}(R(\Lambda,p)^{-1})^{\lambda'}_{\lambda}a(\Lambda\boldsymbol{p},\lambda')\,. \tag{4.57}$$

From 4.54 one concludes that complex wave function $\varepsilon^a(\boldsymbol{p},\lambda)e^{ipx}$ are the coefficient functions connecting the set of operators $a(\boldsymbol{p},\lambda)$, transforming as irreducible unitary representation (m,j) of the Poincaré group, to set of field operators $\Psi^a(x)$ transforming as finite dimensional non-unitary representation of Lorentz group. Their transformation law can be obtained by comparing the left-hand side of 4.48,

$$U(\Lambda)\Psi(x)^aU(\Lambda^{-1}) = D(\Lambda^{-1})^a_b\Psi^b(\Lambda x) \tag{4.58}$$

$$= D(\Lambda^{-1})^a_{a'}\sum_{\lambda}\int \frac{d^3q}{(2\pi)^3\sqrt{2E}} \left(a(\boldsymbol{q},\lambda)\varepsilon^{a'}(\boldsymbol{q},\lambda)e^{iq\Lambda x} + \text{c.c.} \right) \tag{4.59}$$

$$= D(\Lambda^{-1})^a_{a'}\sum_{\lambda}\int \frac{d^3p}{(2\pi)^3\sqrt{2E}} \left(a(\boldsymbol{\Lambda p},\lambda)\varepsilon^{a'}(\Lambda\boldsymbol{p},\lambda)e^{ipx} + \text{c.c.} \right) , \tag{4.60}$$

where the last equality is obtained by substituting $p = \Lambda^{-1}q$, and the right-hand side which is equal to

$$\sum_{\lambda,\lambda'}\int\frac{d^3p}{(2\pi)^3 2E}D^{(j)}(R(\Lambda,p)^{-1})^{\lambda'}_{\lambda}a(\Lambda\boldsymbol{p},\lambda)\varepsilon^a(\boldsymbol{p},\lambda')e^{ipx}+\text{c.c.}\,. \tag{4.61}$$

As a result,

$$D(\Lambda^{-1})^a_{a'}\varepsilon^{a'}(\Lambda\boldsymbol{p},\lambda)=D^{(j)}(R(\Lambda,p)^{-1})^{\lambda'}_{\lambda}\varepsilon^a(\boldsymbol{p},\lambda')\,, \tag{4.62}$$

or equivalently,

$$D(\Lambda)^a_{a'}\varepsilon^{a'}(\boldsymbol{p},\lambda)=D^{(j)}(R(\Lambda,p))^{\lambda}_{\lambda'}\varepsilon^a(\Lambda\boldsymbol{p},\lambda')\,. \tag{4.63}$$

Expression 4.63 is necessary and sufficient condition for field to transform according to 4.48. Let's now explicitly calculate how fields corresponding to massless particles transform under Lorentz transformation. First, condition 4.63 turns to

$$D(\Lambda)^a_{a'}\varepsilon^{a'}(\boldsymbol{p},\lambda)=e^{\pm j,\theta(\Lambda p)}\varepsilon^a(\Lambda\boldsymbol{p},\lambda')\,, \tag{4.64}$$

with the help of 4.44. As explained, to satisfy 4.63 it is enough to consider little group of the standard vector and boosts using which one reaches all the other basis vectors. Hence, consider $\Lambda=L(p_0)$

$$D(L(p_0))^a_{a'}\varepsilon^{a'}(\boldsymbol{p_0},\lambda)=\varepsilon^a(\boldsymbol{p},\lambda')\,, \tag{4.65}$$

since $e^{\pm j\theta(L(p_0),p)}=1$. Next, under the action of elements of little group we have

$$D(R_W)^a_{a'}\varepsilon^{a'}(\boldsymbol{p_0},\lambda)=e^{\pm j\theta(W,p_0)}\varepsilon^a(\boldsymbol{p_0},\lambda')\,. \tag{4.66}$$

One can treat the rotations and translations separately, so

$$\begin{aligned}D(R(\theta))^a_{a'}\varepsilon^{a'}(\boldsymbol{p_0},\lambda)&=e^{\pm j\theta}\varepsilon^a(\boldsymbol{p_0},\lambda')\,,\\ D(T(\eta,\delta))^a_{a'}\varepsilon^{a'}(\boldsymbol{p_0},\lambda)&=\varepsilon^a(\boldsymbol{p_0},\lambda')\,,\end{aligned} \tag{4.67}$$

Let's now narrow down the consideration even further to massless particles of helicity ±1. Such states correspond to fields in finite representation called four-vector representation, labeled by ($u = \frac{1}{2}, v = \frac{1}{2}$), as they are the simplest objects that can account for spin content, since it follows that $0 \leq j \leq 1$. In four-vector representation

$$D(\Lambda) = \Lambda\,. \tag{4.68}$$

As a result, from 4.65 and 4.67 we have

$$\begin{aligned} L(p_0)^\mu_\nu \varepsilon^\nu(\boldsymbol{p_0}, \pm 1) &= \varepsilon^\mu(\boldsymbol{p}, \pm 1)\,, \\ R(\theta)^\mu_\nu \varepsilon^\nu(\boldsymbol{p_0}, \pm 1) &= e^{\pm i\theta} \varepsilon^\mu(\boldsymbol{p}, \pm 1)\,, \\ T(\eta, \delta)^\mu_\nu \varepsilon^\nu(\boldsymbol{p_0}, \pm 1) &= \varepsilon^\mu(\boldsymbol{p}, \pm 1)\,, \end{aligned} \tag{4.69}$$

where the rotation is given by 4.25 and translation by 4.26.

One can satisfy the rotation requirement with

$$\varepsilon^\mu(\boldsymbol{p_0}, \pm 1) = \frac{1}{\sqrt{2}}(1, 0, 0, \pm i)\,, \tag{4.70}$$

but then, as a result of action of the little group we get

$$D^\mu_\nu(R_W(\theta, \eta, \delta))\varepsilon^\nu(\boldsymbol{p_0}, \pm 1) = T^\mu_\rho(\eta, \delta) R^\rho_\nu(\theta) \varepsilon^\nu(\boldsymbol{p_0}, \pm 1) \tag{4.71}$$

$$= e^{\pm\theta} \left(\varepsilon^\mu(\boldsymbol{p_0}, \pm 1) + p_0^\mu \frac{\eta \pm i\delta}{\sqrt{2}|p_0|} \right)\,. \tag{4.72}$$

For a general momentum this reduces to

$$\varepsilon^\mu(\Lambda \boldsymbol{p}, \pm 1) \varepsilon^{\pm i\theta(p,\Lambda)} = \Lambda^\mu_\nu \varepsilon^\nu(\boldsymbol{p}, \pm 1) + p^\mu \Omega_\pm(\boldsymbol{p}, \Lambda)\,. \tag{4.73}$$

We are not interested in the explicit form of $\Omega_\pm$. Finally, the field of massless spin-1 particle is $A^\mu(x)$

$$A^\mu(x) = \sum_\lambda \int \frac{d^3 p}{(2\pi)^3 \sqrt{2E}} \left(a(\boldsymbol{p}, \lambda) \varepsilon^a(\boldsymbol{p}, \lambda) e^{ipx} + a^\dagger(\boldsymbol{p}, \lambda) \varepsilon^{*a}(\boldsymbol{p}, \lambda) e^{-ipx} \right)\,, \tag{4.74}$$

which, according to 4.73, transforms as

$$U(\Lambda) A_\mu(x) U^{-1}(\Lambda) = \Lambda^\nu_\mu A_\nu(\Lambda x) + \partial_\mu \Omega(x, \Lambda)\,. \tag{4.75}$$

Hence, the field related to massless particle does not transform covariantly under

Lorentz transformations. The second term, which ruins the covariance, is a consequence of the fact that the little group of massless states contains translations. As we have seen, these transformations do not correspond to observable quantities. Instead, they are called gauge transformations and we interpret them as relating different descriptions of the same physical state.

This concludes the discussion of Poincaré group. We now turn to studying conserved quantities, which is just a continuation of discussion about symmetries.

4.3 Noether's theorem and conservation laws

The goal of this section is to find conserved currents and charges of a theory described by action $S[\psi_i]$, given by 4.76.

$$S[\psi_i] = \int d^4x \mathcal{L}(\psi_i, \partial_\mu \psi_i)\,, \quad i = 1, ..., N\,. \tag{4.76}$$

We will refer to Lagrangian density $\mathcal{L}$ as Lagrangian. It is a function of N fields $\psi_i(x)$[19] and their derivatives[20] $\partial_\mu \psi_i(x)$. Fields are functions of points labeled by coordinates x^μ.

As it is known, conservation of a quantity is a consequence of some symmetry transformation of the action. But, before we define what kind of transformations are considered to be symmetry transformations, let's discuss transformations of action in general. First, transformation under which the action transforms acts on fields, since action is a functional. Secondly, transformations may be continuous or discrete, both resulting in conserved quantity if they represent symmetry. However, Weinberg–Witten theorem applies only to charges derived from conserved currents and they can result only from continuous symmetry, as only such transformations allow one to define currents and densities, quantities varying continuously in space. As an example of discrete symmetry transformation consider parity transformation in non-weak interactions. Conserved charge, a consequence of this symmetry, is internal parity, which exists only as a global notion — one cannot define density of internal parity. Consequently, since the current would describe the flow of quantity

[19]Field ψ_i is not necessary a scalar field. Indices are suppressed to simplify the notation.

[20]The Lagrangian may in general be a function of higher order derivatives and coordinates, but as we are interested in Standard Model theories it is enough to limit the Lagrangian to form 4.76. Lagrangian describing gravity does contain derivatives of the field of the second order, but as it will be shown below, it is enough to consider 4.76.

as density changes, there can be no current. From this point on, transformations will refer only to continuous one.

Formally, continuous transformations form a Lie group, introduced in the previous section. We will distinguish between global Lie groups containing n finite (or countably infinite) number of independent, constant parameters $\omega_a = const.$, and local Lie groups with continuously infinite number of parameters, i.e. parameters that are arbitrary functions of coordinates $\omega_a = \omega_a(x)$. Accordingly, one refers to elements of infinite dimensional Lie group, which act independently at different points, as local transformations, and to elements of finite dimensional Lie group as global transformations, since they transform the field by the same value at each point. What's more, since Lie group can be obtained from its generators, it is enough to study how the fields transform under infinitesimal transformations. The change of the field under infinitesimal transformation is called variation $\delta\psi_i$,

$$\begin{aligned} \psi_i(x) \to \psi_i'(x) = \psi_i(x) + \delta\psi_i(x) \\ \implies \delta\psi_i(x) = \psi_i'(x) - \psi_i(x)\,. \end{aligned} \tag{4.77}$$

It contains terms of the lowest power of the parameters and (in the case of local transformation) their derivatives,

$$\delta\psi_i(x) = \alpha_i(\psi_i, \partial\psi_i)\omega + \beta_i^\mu(\psi_i, \partial\psi_i)\partial_\mu\omega\,, \tag{4.78}$$

where α_i and β_i are some functions, whose explicit form depends on the form of transformation. Furthermore, one can divide the transformation between those that act on spacetime points, called spacetime transformations, and all others referred to as internal transformations. Hence, let's determine the variations of fields, starting with spacetime transformations.

One begins by defining how points transform under spacetime transformation. Consider vector field ξ. At any given point it generates a unique local integral flow Φ^ξ. Using the flow, one can move a point a parameter distance $\epsilon \in I \subseteq \mathbb{R}$ as defined by map Φ^ξ_ϵ,

$$\begin{aligned} \Phi^\xi_\epsilon : U \to M \\ P \mapsto \Phi^\xi_\epsilon(P) = Q\,. \end{aligned} \tag{4.79}$$

By "flowing" for the amount ϵ starting from each point $P \in U$, along integral curve through P, the point is moved to a new position denoted as Q. If the vector field is complete, $I = \mathbb{R}$, one has $U = M$, so the flow is defined on an entire manifold, and it is called global flow. Concerning integral curves, this means that they don't have an end point. In such case, the map Φ^{ξ}_{ϵ} is called diffeomorphism, and the family $\{\Phi^{\xi}_{\epsilon}\}_{\epsilon \in \mathbb{R}}$ is called one-parameter group[21] of ξ. It is a Lie group, with vector field ξ as the generator. In a chart induced basis $\xi = \xi^{\mu}\partial_{\mu}$, where ξ^{μ} is parameter of the group. As explained, if $\xi^{\mu} = \xi^{\mu}(x)$ the group has continuously infinite number of parameters, i.e., it is local Lie group. Sometimes we will refer to it as group of general diffeomorphisms. For infinitesimal transformations we will write $\xi^{\mu}(x)\epsilon = \epsilon^{\mu}(x)$, where ϵ is infinitesimal. Hence, infinitesimal spacetime transformation is in a chart x^{μ} described as

$$x(P) = x^{\mu} \to x(Q) = x^{\mu} + \epsilon(x)\xi^{\mu} = x^{\mu} + \epsilon^{\mu}(x)\,, \tag{4.80}$$

Since we're dealing with infinitesimal transformations, one does not move very far away under the flow and only one chart can be used. Thus, under infinitesimal transformation, point P with coordinates x^{μ} is moved along the flow to a new position Q with coordinates $x^{\mu} + \epsilon^{\mu}$.

Let's now see what happens to scalar field $\phi(x)$, belonging to $(u = 0, v = 0)$ representation, when dragged along the flow Ψ^{ξ}_{ϵ}. We say that the transformation generates a new field $\phi'(x)$, called pushed-forward of $\phi(x)$, such that

$$\begin{aligned} \phi'(x(Q)) &= \phi(x(P)) \\ \phi'(x+\epsilon) &= \phi(x)\,. \end{aligned} \tag{4.81}$$

The spacetime transformation 4.80 mapped ϕ into new field ϕ', which has the same value at Q as the original field at P, i.e., ϕ' describes a new configuration of points. This is why the new field is called the pushed-forward field. The relationship 4.81

[21]The group can be defined only for complete vector fields because otherwise the composition of $\Phi^{\xi}_{s} \circ \Phi^{\xi}_{t}$ may not be defined, as one may exist the domain of integral curve.

from the point of view of the new field is obtained by

$$\begin{aligned} \phi'(x') &= \phi(x' - \epsilon(x')) \\ &= \phi(x') - \epsilon^\mu(x')\partial_\mu\phi(x') \,, \end{aligned} \tag{4.82}$$

since, according to 4.80, $x(x') = x' - \epsilon$. Usually, in the last step we relabel coordinates $x' \leftrightarrow x$,

$$\begin{aligned} \phi'(x) &= \phi(x) - \epsilon\xi^\mu(x)\partial_\mu\phi(x) \\ &= \phi(x) - \mathfrak{L}_\xi\phi \,. \end{aligned} \tag{4.83}$$

Infinitesimal change of field under diffeomorphism is described by Lie derivative $\mathfrak{L}$. Spacetime transformation generates a new field ϕ' by pushing field ϕ in the direction of vector field ξ by amount $(-\epsilon)$. We get the minus sign because 4.83 is from point of view of the new field[22].

Let's now see how to deal with fields of higher tensor rank. Unlike for the scalar field, we must take into account that the components change. Consider vector field $V^\mu_{(x)}(x)$ and pushed-forward vector field $V'^\mu_{(x)}(x)$. The subscript denotes the chart in which coordinates are evaluated. The relationship between the vector fields is

$$V'^\mu_{(x)}(x(Q)) = V^\mu_{(y)}(y(P)) \,, \tag{4.84}$$

where y is a chart such that[23]

$$y^\mu = x^\mu + \epsilon^\mu \,. \tag{4.85}$$

[22]Let's say that scalar field ϕ describes temperature field, which is generated by an oven, and that the temperature field has as extremum at point P, the position of the oven. If the oven is moved to new point Q, then this is the point of extremum and the new distribution of temperature is described by field ϕ'. One can see that the old field had the same value at P as the new field has at Q, which is expression 4.81. Furthermore, $\phi'(Q) = \phi((\Phi^\xi_\lambda)^{-1}Q)$, as $P = (\Phi^\xi_\lambda)^{-1}Q$. The last equality describes ϕ from the point of view of the new field.

[23]It is easiest to explain 4.84 on an example. Consider vector field in three dimensions that one rotates by a right angle about z axis. After the transformation the vectors point into different direction and their base point changed if it was not on rotation axis. Consider a vector V_P whose base point is $x(P) = (1,0,0)$, so that it's not on rotation axis. After the transformation, the base point is at $x(Q) = (0,1,0)$. Furthermore, let's say the components of V_P prior to the rotation are $V_{(x)}(x(P)) = (a,0,0)$. After the transformation the new configuration is described by a vector field V'. As a result $V_{(x)P} = (a,0,0)$, and $V'_{(x)Q} = (0,a,0)$. To find relationship between the new and the old components in the sense of 4.81 we use the freedom of using any (compatible) chart. Hence, in y chart given by 4.85 one can see that $V_{(y)}(y(P)) = (0,a,0) = V'_{(x)Q}$.

However, one can only see how the field actually changed when both are written in the same chart. To transform $V_{(y)}$ back to x chart we use the rule for vector components transformation,

$$V_{(y)}(y(P))^{\mu} = \frac{\partial y^{\mu}}{\partial x^{\nu}} V^{\nu}_{(x)}(x(P))\,. \tag{4.86}$$

Inserting 4.86 into right-hand side of 4.84 leads to

$$\begin{aligned} V'^{\mu}_{(x)}(x(Q)) &= \frac{\partial y^{\mu}}{\partial x^{\nu}} V^{\nu}_{(x)}(x(P)) \\ V'^{\mu}_{(x)}(x+\epsilon) &= \frac{\partial y^{\mu}}{\partial x^{\nu}} V^{\nu}_{(x)}(x) \\ &= \Big(\delta^{\mu}_{\nu} + \partial_{\nu}\epsilon^{\mu}(x)\Big) V^{\nu}_{(x)}(x)\,. \end{aligned} \tag{4.87}$$

As both sides of 4.87 are in the same chart we will drop the subscript. As before, we wish to express relationship 4.87 from the point of view of the new field, so

$$\begin{aligned} V'^{\mu}(x') &= \Big(\delta^{\mu}_{\nu} + \partial_{\nu}\epsilon^{\mu}(x')\Big) V^{\nu}(x' - \epsilon(x')) \\ &= \Big(\delta^{\mu}_{\nu} + \partial_{\nu}\epsilon^{\mu}(x')\Big)\Big(V^{\nu}(x') - \epsilon^{\rho}(x')\partial_{\rho}V^{\nu}(x')\Big) \\ V'^{\mu}(x) &= V^{\mu}(x) - \Big(\epsilon^{\nu}(x)\partial_{\nu}V^{\mu}(x) - V^{\nu}(x)\partial_{\nu}\epsilon^{\mu}(x)\Big) \\ &= V^{\mu}(x) - \mathfrak{L}_{\xi}V^{\mu}\,. \end{aligned} \tag{4.88}$$

Note that $\epsilon(x) = \epsilon(x') + O(\epsilon^2)$. In the last step we just relabeled the coordinates. The first term in 4.88 in the curly brackets is the same as in 4.83. It corresponds to shifting of the field, or more precisely, the base points. The second term corresponds to transformation of components. The above procedure is analogous for tensors of different rank.

In summary, variation of the field caused by infinitesimal spacetime transformation δ_{ϵ} is given by Lie derivative.

$$\psi_i(x) \to \psi_i(x) - \mathfrak{L}_{\xi}\psi_i\,, \implies \delta_{\epsilon}\psi_i = -\mathfrak{L}_{\xi}\psi_i\,. \tag{4.89}$$

One can check, by transforming the coordinates that Lie derivative is a tensorial quantity, independent of chosen chart.

Aside from spacetime transformations, there are symmetries related to internal degrees of freedom important for the first part of the theorem. In case of Standard

Model they are elements of $SU(N)$ group, whose generators X_a of the underlying algebra satisfy

$$[X_a, X_b] = i f_{abc} X_c \, , \tag{4.90}$$

where f_{abc} are structure constants. In such case, one usually works with fields in either fundamental or adjoint representation. In the fundamental representation the generators are represented as traceless, hermitian $N \times N$ matrices that act on column vectors with N elements,

$$\psi_i^{\text{fund}} \to U \psi_i^{\text{fund}} \implies \delta_0 \psi_i^{\text{fund}} = i\omega_a X^a \psi_i^{\text{fund}} = i\omega \psi_i^{\text{fund}} \, , \tag{4.91}$$

where $U = e^{i\omega_a X^a}$ is en element of global group. Infinitesimal change of the field under $SU(N)$ group is denoted by δ_0. On the other hand, in the adjoint representation generators are represented on vector space of the algebra. As a consequence, elements of the algebra A transforms by conjugation

$$A \to U A U^{-1} \, , \tag{4.92}$$

and the variation is obtained by looking at infinitesimal transformation $U = 1 + i\omega_a X^a$ leading to

$$\delta_0 A = \omega^a f_{abc} A^b X^c \, . \tag{4.93}$$

Expressions 4.91 and 4.92 continue to hold for local case where $\omega = \omega(x)$, except for gauge field A which in such case transform as

$$A \to U_x A U_x^{-1} + i U_x \partial_\mu U_x^{-1} \implies \delta_0 A = X^a \partial_\mu \omega_a + \omega^a f_{abc} A^b \, , \tag{4.94}$$

The subscript x on U denotes the fact that we are dealing with local Lie groups. Comparing the above expressions to 4.78, one can see that gauge field will for example have both α and β coefficients non-zero, while for fields in fundamental representation have only α as nonvanishing coefficient. Further, variation of action $S[\psi_i]$

defined in 4.76 is

$$\delta S = \int d^4x \frac{\delta S}{\delta \psi_i} \delta \psi_i \,, \tag{4.95}$$

where $\frac{\delta S}{\delta \psi_i}$ is the functional derivative of S, whereas the effect of any infinitesimal transformation, symmetry or not, on Lagrangian is

$$\begin{aligned} &\mathcal{L} \to \mathcal{L} + \delta\mathcal{L} \,, \\ \delta\mathcal{L} &\equiv \frac{\partial \mathcal{L}}{\partial \psi_i} \delta \psi_i + \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)} \delta \left(\partial_\mu \psi_i\right) \\ &= \left[\frac{\partial \mathcal{L}}{\partial \psi_i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)}\right)\right] \delta \psi_i + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)} \delta \psi_i\right) \,. \end{aligned} \tag{4.96}$$

To obtain the second line one uses the fact that partial and variational derivative commute,

$$\delta(\partial_\mu \psi_i) = \partial_\mu \psi_i' - \partial_\mu \psi_i = \partial_\mu (\psi_i' - \psi_i) = \partial_\mu (\delta \psi_i) \,, \tag{4.97}$$

and then partially integrates the second term in the first line of 4.96. The term in square brackets are the left-hand side of Euler-Lagrange equations, also referred to as equations of motion, or field equations. The second term is the boundary term.

Finally, we define symmetry group of the theory as the one that changes the action at most by a total derivative.

$$\delta_s \mathcal{L} = \partial_\mu K^\mu \implies \quad \delta_s S = S[\psi_i + \delta_s \psi_i] - S[\psi_i] = \int d^4x \, \partial_\mu K^\mu \,, \tag{4.98}$$

where K^μ is some function. When we allow that Lagrangian changes up to boundary term it is assumed that this new term vanishes under the same boundary conditions which result in field equations in the first place. Symmetry variation is denoted by index s. Thus, variation of the field $\delta \psi_i$ satisfying condition 4.98 is a symmetry transformation. In other words, 4.98 is an equation for $\delta_s \psi_i$, and as such is true for any field, on-shell or off-shell[24]. As a result of 4.98 and 4.96, one has

$$\begin{aligned}\int d^4x\partial_\mu K^\mu &= \int d^4x \Bigg\{ \left[\frac{\partial \mathcal{L}}{\partial \psi_i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \psi_i\right)}\right)\right]\delta_s\psi_i + \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \psi_i\right)}\delta_s\psi_i\right)\Bigg\} \\ 0 &= \int d^4x \Bigg\{ \left[\frac{\partial \mathcal{L}}{\partial \psi_i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \psi_i\right)}\right)\right]\delta_s\psi_i + \partial_\mu j^\mu \Bigg\}\,,\end{aligned} \tag{4.99}$$

with j^μ, defined as

$$j^\mu \equiv \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \psi_i\right)}\delta_s\psi_i - K^\mu\,. \tag{4.100}$$

To simplify notation, let

$$\left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL} \equiv \left[\frac{\partial \mathcal{L}}{\partial \psi_i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \psi_i\right)}\right)\right]\,. \tag{4.101}$$

Thus, the symmetry transformations leads to the following integral-free relationship

$$\left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL}\delta_s\psi_i = -\partial_\mu j^\mu\,. \tag{4.102}$$

In general, the symmetry variation of the fields is of the form given by 4.78. Inserting the expression into the left-hand side of 4.102 yields

$$\begin{aligned}\left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL}\delta_s\psi_i &= \left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL}\left[\alpha_i\omega + \beta_i\partial_\mu\omega\right] \\ &= \left\{\left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL}\alpha_i - \partial_\mu\left(\beta_i\left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL}\right)\right\}\omega + \partial_\mu b^\mu\,,\end{aligned} \tag{4.103}$$

where $\partial_\mu b^\mu = \partial_\mu\left(\beta_i\left[\frac{\delta\mathcal{L}}{\delta\psi_i}\right]_{EL}\omega\right)$. Hence, 4.102 turns into

$$\left\{\left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL}\alpha_i - \partial_\mu\left(\beta_i\left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL}\right)\right\}\omega = -\partial_\mu(j^\mu + b^\mu)\,. \tag{4.104}$$

To draw some further conclusions we will get rid of the right-hand side[25]. Note that ω_a and it's derivatives are arbitrary functions. Thus, we integrate over 4.104, and choose the functions that vanish at the boundary. As a consequence, the right-hand

[24]Field is on shell if it satisfies equations of motion.

[25]Although there are some additional relationships one can obtain from 4.104 [83], they are not important for our purposes.

side vanishes since it is an integral over divergence. It then follows, for every ω_a in the interior,

$$\left[\frac{\delta\mathcal{L}}{\delta\psi_i}\right]_{EL}\alpha_i - \partial_\mu\left(\left[\frac{\delta\mathcal{L}}{\delta\psi_i}\right]_{EL}\beta_i\right) = 0\,. \tag{4.105}$$

What one can conclude from 4.105 is that when symmetry group of the theory is local,

- Field equations and their derivatives are not independent. Since 4.105 is true for any field configuration it is an identity,
- Relationship 4.105 can be viewed as condition on the form of Lagrangian for the variation of action to vanish under local transformations,
- If all fields satisfy field equations the result is tautology,
- If there are two set of fields, one such that the variations depends only on ω, but not on it's derivatives $\partial\omega$ and vice versa, and either of the sets (but not both at the same time) satisfy field equations, one obtains conservation law.

Some application is shown further below, but one may recognize that 4.105 is the origin of what is known as covariant conservation laws. Expression 4.104, or sometimes 4.105, is known as the second Noether's theorem.

We now turn to the case where the action is invariant under global Lie groups, i.e., the situation where parameters are constant. We right away consider variation of fields under diffeomorphism and $SU(N)$ group (which doesn't impose any restriction on generality of the discussion),

$$\begin{aligned}\delta\psi_i &= \delta_0\psi_i - \mathfrak{L}_\xi\psi_i \\ &= \alpha_i\omega_a X^a - \xi^\mu\partial_\mu\psi_i\,.\end{aligned} \tag{4.106}$$

Because parameter is constant Lie derivative in 4.106 reduces to partial derivative. Next, only diffeomorphisms change Lagrangian[26]. It transforms as a scalar density, so

$$\mathfrak{L}_\xi\mathcal{L} = \partial_\mu(\mathcal{L}\xi^\mu) \implies K^\mu = \mathcal{L}\xi^\mu\,. \tag{4.107}$$

[26]If there are no background fields. This will be discussed below.

Returning to 4.102, and after inserting 4.106 into 4.100 the right-hand side looks like

$$\begin{aligned} j^\mu &= \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)} \delta_0 \psi_i - \left[\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)} (\xi^\nu \partial_\nu \psi_i) - \delta^\mu_\nu \mathcal{L} \xi^\nu\right] \\ &= \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)} \alpha_i \omega^a X_a - \left[\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)} (\partial_\nu \psi_i) - \delta^\mu_\nu \mathcal{L}\right] \xi^\nu \\ &\equiv J^\mu_\alpha \omega^\alpha - T^\mu_\nu \xi^\nu \,. \end{aligned} \tag{4.108}$$

We defined J^μ_α as

$$J^\mu_\alpha \equiv \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)} \alpha_i X_a \,, \tag{4.109}$$

and the term in square brackets as T^μ_ν,

$$T^\mu_\nu \equiv \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)} (\partial_\nu \psi_i) - \delta^\mu_\nu \mathcal{L} \,, \tag{4.110}$$

As a result, the expression 4.102 reduces to

$$\begin{aligned} \left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL} (\alpha_i \omega^a X_a - \xi^\mu \partial_\mu \psi_i) = -\partial_\mu j^\mu \implies \\ \left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL} \alpha_{i,a} = -\partial_\mu J^\mu_\alpha \,, \\ \left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL} \partial_\mu \psi_i = -\partial_\mu T^\mu_\nu \,. \end{aligned} \tag{4.111}$$

If the symmetry group of the theory is global, divergence of j^μ is proportional to the left-hand side of equations of motion. What's more, if all fields are on-shell, i.e., fields satisfy Euler-Lagrange equations, 4.111 leads to conservation laws,

$$\partial_\mu j^\mu = 0 \,. \tag{4.112}$$

Furthermore, as ω and ϵ are independent,

$$\partial_\mu J^\mu_\alpha = \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)} \alpha_{a,i}\right] = 0 \,, \tag{4.113}$$

$$\partial_\mu T^\mu_\nu = \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i\right)} (\partial_\nu \psi_i) - \delta^\mu_\nu \mathcal{L}\right] = 0 \,. \tag{4.114}$$

These are local conservation laws, existing for each parameter of the group. Note that on-shell equations of motion follow from Hamilton's principle which fixes variation at the boundary, otherwise functional derivative is not defined, i.e., variation of Lagrangian is not in the form 4.95.

This is the Noether's first theorem, or just Noether's theorem — symmetries of action associated with global Lie groups result in conserved currents when the equations of motion are on-shell for all fields on which Lagrangian depends on. We will refer to conserved quantities J^μ_α and T^μ_ν as Noether's currents[27], or canonical currents. Furthermore, $T^{\mu\nu}$ is called energy-momentum tensor[28] (EMT), because under spacetime translations the additional index ν transforms under Lorentz transformations, so the resulting Noether current is a tensor of rank two.

It is left to derive the conserved charges, i.e., globally conserved quantities. Consider 4.113 for example. One can write

$$\partial_\mu J^\mu_\alpha = \partial_0 J^0_a + \partial_i J^i_\alpha \,. \tag{4.115}$$

We then define charge Q_α as

$$Q_\alpha(t) \equiv \int d^3x J^0_a \,, \tag{4.116}$$

The zeroth component of current is local density of charge Q_a. Assuming all fields vanish at spatial infinity, one may integrate 4.115 to obtain,

$$\begin{aligned} 0 &= \int d^3x \left[\partial_0 J^0_\alpha(\mathbf{x},t) + \partial_i J^i_\alpha(\mathbf{x},t)\right] \\ &= \int d^3x \partial_0 J^0_\alpha(\mathbf{x},t) \\ &= \partial_0 Q_\alpha(t) \,, \end{aligned} \tag{4.117}$$

where the second line follows from the mentioned boundary conditions. Hence, charge is conserved in time as a consequence of local conservation, and the converse is also true. Similarly, for EMT one defines the charge P^μ

$$P^\mu \equiv \int d^3x T^{0\mu} \,, \tag{4.118}$$

[27]More precisely, $J^\mu = J^\mu_\alpha \omega^\alpha$ and $T^\mu_\nu \xi^\nu$ are currents, but we usually leave out the parameters since they' re constant.

[28]The name is misleading as EMT obtained by Noether's procedure does not always transform as second rank tensor under general coordinate transformations. We will come back to this point later.

where P^μ is total energy-momentum of the fields. Specifically, P^0 agrees with canonical Hamiltonian $\mathcal{H}$, so the zeroth component of EMT $T^{0\mu}$ is energy density. Moreover, P^μ is a Lorentz vector. This result is not obvious and it is known as von Laue's theorem. Proof of it is based on assumption that $T^{\mu\nu}$ vanishes sufficiently fast at the boundary [84].

Conserved charges resulting from Noether's laws are called Noether's charges. We refer to them as global conservation laws as they state that the amount of charge is constant in time. This on the other hand is not true for density. Since $J_a^0 = J_a^0(x)$, even though charge is conserved, density is not. It can increase at some point, which means that it decreased at another and as a result, there is a current, J_a^i, flowing from one point to the other.

These two theorems determine conservation laws of a theory. It can be proved that the converse of the theorems is also true. If there is a dependency between field equations and its derivatives the symmetry group of the action is local. Furthermore, if field equations can be written as divergences the symmetry group of the theory is global. Moreover, if a symmetry group of a theory consists of local transformations that contain a non-trivial subgroup of global transformations then the second Noether's theorem applies to local transformations, resulting in identities 4.105, and the first one to global, resulting in conserved quantities 4.112. However, since the second theorem considers all possible parameters, it includes also the case when parameters are constant. In other words, conservation laws arising from global subgroup are just a consequence of 4.105. This means that divergent term in 4.111 is a linear combination of field equations and it's derivatives [29]. The point is that in such case, there will be no conservation law. To reformulate, to obtain the conservation law one should be sure that the global symmetry group is not a subgroup of a bigger group of local transformations [83][85][86][87][88][71][89].

Some further subtleties concerning the Noether's theorems are in App A. We'll now apply the theorems to current theories describing charged massless particles — QCD, describing spin-1 particles and gravity which describes spin-2 particles.

[29] The combination is *linear* because field equations occur linearly in both 4.111 and 4.104. Thus, in general, the right-hand side of 4.102 is linear combination of field equations, their derivatives and functions whose divergence vanishes.

4.3.1 QCD

Gluon action describes massless spin-1 particles charged under $SU(3)$ group. Thus, the theory is in the scope of the first part of Weinberg–Witten theorem, with Lagrangian $\mathcal{L}_g$ given by 4.119.

$$\begin{aligned} \mathcal{L}_g(G^a_\mu) &= -\frac{1}{4}F^a_{\mu\nu}(x)F^{\mu\nu}_a(x)\,, \\ F^a_{\mu\nu}(x) &= \partial_\mu G^a_\nu(x) - \partial_\nu G^a_\mu(x) + f_{abc}G^b_\mu(x)G^c_\nu(x) \\ &= D_\mu G^a_\nu(x) - D_\nu G^a_\mu(x) + f_{abc}G^b_\mu(x)G^c_\nu(x)\,. \end{aligned} \tag{4.119}$$

Field G^a_μ is gluon gauge field and D_μ is gauge covariant derivative defined as

$$D_\mu\psi = \partial_\mu\psi + iG_\mu\psi\,. \tag{4.120}$$

To see if there is a conservation law we must determine the symmetry group. First, the theory does not contain the background fields. Next, note that $F^{\mu\nu}_a$ is in the adjoint representation, so it transforms under elements of $SU(3)$ group $U = \exp^{-i\omega_a(x)T_a}$ as

$$F_{\mu\nu} \to UF_{\mu\nu}U^{-1}\,, \tag{4.121}$$

by definition. We have used $F_{\mu\nu} = F^a_{\mu\nu}T_a$. Since 4.119 can be written as

$$\mathcal{L}_g(G^a_\mu) = -\frac{1}{2}\mathrm{Tr}\left(F_{\mu\nu}F^{\mu\nu}\right)\,, \tag{4.122}$$

using the cyclic properties of trace one can easily see that Lagrangian is invariant under local transformations. Even though the local group contains a global subgroup which is also a symmetry of the theory, according to the second Noether's theorem we do not expect to obtain a conservation law, but an identity. Nevertheless, let's see what kind of identity one obtains. Starting from variation of the gluon field,

$$\delta_0 G_{\mu,a} = f_{abc}G^c_\mu\omega^b + \partial_\mu\omega_a\,, \tag{4.123}$$

and comparing 4.123 to 4.78, we see that

$$\begin{aligned} \alpha_{\mu,a} &= f_{abc} G^c_\mu \,, \\ \beta^{\nu,b}_{\mu,a} &= \delta^\nu_\mu \delta^b_a \,. \end{aligned} \tag{4.124}$$

Then, by varying 4.119,

$$\begin{aligned} \frac{\partial \mathcal{L}_g}{\partial(\partial_\mu G^a_\nu)} = \frac{\partial \mathcal{L}_g}{\partial(D_\mu G^a_\nu)} &= -F^{\mu\nu}_a \,, \\ \frac{\partial \mathcal{L}_g}{\partial G^a_\nu} &= -f_{bac} F^{\nu\rho,b} G^c_\rho \,. \end{aligned} \tag{4.125}$$

Finally, using 4.101 one obtains the left-hand side of Euler-Lagrange equations,

$$D_\mu F^{\mu\nu}_a = \partial_\mu F^{\mu\nu}_a - f_{bac} F^{\nu\rho,b} G^c_\rho \,. \tag{4.126}$$

Inserting obtained expressions into the second Noether's theorem 4.105, results in

$$(D_\mu F^{\mu\nu}_a)\, f_{abc} G^c_\nu = \partial_\nu (D_\mu F^{\mu\nu}_a) \,. \tag{4.127}$$

Now, for the case of global transformation, coefficient β in 4.123 vanishes. Hence,

$$(D_\mu F^{\mu\nu}_a)\, f_{abc} G^c_\nu = 0 \,. \tag{4.128}$$

The expression coincides with 4.111. Since structure constants and gluon field are nonvanishing (otherwise, one would have $F^{\mu\nu} = 0$, which leads to vanishing of the Lagrangian 4.119), expression 4.128 implies that gluons satisfy

$$D_\mu F^{\mu\nu}_a = 0 \,. \tag{4.129}$$

The second Noether's theorem for global transformations reduces to condition equivalent to field equations. What's more, this is so-called covariant conservation law. One can expand the covariant derivative using 4.126, which leads to

$$\partial_\mu F^{\mu\nu}_a = f_{bac} F^{\nu\rho,b} G^c_\rho \,. \tag{4.130}$$

This is a continuity equation, with the term on the right-hand side acting as a source. In other words, covariant conservation law describes exchange of color charge among the gluons. However, one may obtain a conservation law because tensor $F_a^{\mu\nu}$ is antisymmetric, which one may see from 4.119. Due to antisymmetricity,

$$\partial_\nu \partial_\mu F_a^{\mu\nu} = \partial_\nu \left(f_{bac} F^{\nu\rho,b} G_\rho^c \right) = 0 \,. \tag{4.131}$$

Thus, current $\mathcal{J}_a^\nu$, defined as

$$\mathcal{J}_a^\nu = \partial_\mu F_a^{\mu\nu} = f_{bac} F^{\nu\rho,b} G_\rho^c \implies \partial_\mu \mathcal{J}_a^\mu = 0 \,. \tag{4.132}$$

is conserved, and the charges are given by

$$Q_a = \int d^3x \partial_\mu F_a^{\mu 0} = \int d^3x F^{\mu 0,b} G_\mu^c f_{abc} \,. \tag{4.133}$$

We obtained conserved current $\mathcal{J}_a$ due to antisymmetric properties of $F^{\mu\nu}$. What remains to explain is the meaning of covariant conservation law and it's reduction to “ordinary" conservation law. Local symmetry is a way to implement interaction. In 4.119 there is only one field, but the Lagrangian is non-linear, which means there is self-interaction — gluons exchange the conserved charge among themselves, described by covariant conservation law 4.129. What the conservation law 4.132 states is that the system of gluons is closed, i.e., there are no other fields, besides the gluons themselves, with whom the gluons can exchange the charge. Simply put, there are no other fields that could act as a source. By adding interaction with quark field to 4.119 for example, we would obtain

$$\partial_\nu \left(f_{bac} F^{\nu\rho,b} G_\rho^c \right) = -\partial_\mu (\bar{\psi} \gamma^\mu T^a \psi) \,. \tag{4.134}$$

As a result, the current would no longer be conserved. One could define conserved current by moving everything to the left-hand side, but that would be a sort of an empty statement. In such case, we could view the left-hand side as describing one big field, whose current is conserved because there is nothing else in the universe the field can exchange the charge with. Such current would also be charged under both gluons and quarks, so it is not of interest, as quarks are massive particles [27][88].

4.3.2 Gravitation

In quantum field theory, particle describing gravitation, considered to be excitation of the metric field, is called graviton. The properties we are interested in with regard to Weinberg–Witten theorem, are mass and spin. As it was shown, they can be defined only on flat background, but let's first deal with what kind of properties one expects of graviton at all. First, the graviton should be massless, to account for the fact that gravitation is a long-range force decreasing with distance[30]. Next, the spin decides whether the force is attractive or repulsive, or more precisely, the polarization sum does [41]. It turns out particles of even spin meditate attractive force between particles of like charge, while particles with odd spin meditate repulsive force. Since gravitation is always attractive, only particles of even spin can be its mediators. What's more, graviton couples to anything carrying energy[31], charge of gravitational action, which includes itself and matter. As it was discussed above, local energy (and momentum) distribution of a system is represented by its energy-momentum tensor T^{μ}_{ν}. The only way for scalar field to couple to EMT is in the form ϕT^{μ}_{μ}, but for the electromagnetic field for example, the trace of EMT vanishes. This would mean gravity doesn't couple to electromagnetic field, which we know is false, as gravity bends light. As a result, the gravitational force must be mediated by spin-2 field[32], and is as such related to field $h_{\mu\nu}$ transforming in symmetric, traceless $(u = 1, v = 1)$ representation.

In summary, graviton, particle meditating gravitational force should be massless, spin-2 particle, corresponding to symmetric, traceless field $h_{\mu\nu}$. Since the graviton as massless particles has energy, it is charged. Thus, it is in the scope of the second part of Weinberg–Witten theorem. What we are looking for is energy-momentum tensor $T^{\mu\nu}$ whose $T^{0\nu}$ components are related to energy-momentum four-vector, which, as it is known results from translation invariance.

As for the gluon case, to find conserved currents, we should start from action

[30]Although gluons are massless, strong force is not a long range force due to asymptotic freedom — interaction between colour charged particles becomes weaker at higher energies. i.e., shorter distances, and because of colour confinement — only colour singlet particles can exist as free. As a result, separating quarks results in production of quark-antiquark pair, and gluons end up confined within hadrons. Hence, it is mesons, massive particles, which meditate force between hadrons. Furheremore, it is hypothesized that by separating gluons, so called glueballs form, which are yet to be detected. Glueballs should be massive, even though they are made of massless gluons. This is part of so called mass-gap problem.

[31]Mass and energy are equivalent, according to the special equivalence principle.

[32]Theories with massless higher spin particles are over-constrained [90].

describing gravitation. However, it turns out there are quite a few approaches, so instead, we start from left-hand side of field equations, resulting from extremizing the action.

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \, . \tag{4.135}$$

Tensor $G_{\mu\nu}$ is Einstein's tensor, and $R_{\mu\nu}$ is Ricci tensor constructed by contraction of Riemann curvature tensor, while R is Ricci or curvature scalar,

$$R = g^{\mu\nu}(\Gamma^{\sigma}_{\mu\sigma,\nu} - \Gamma^{\sigma}_{\mu\nu,\sigma}) - g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \, , \tag{4.136}$$

with Christoffel symbols Γ defined as

$$\Gamma^{\rho}_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} \left(\partial_\nu g_{\sigma\mu} + \partial_\mu g_{\sigma\nu} - \partial_\sigma g_{\mu\nu} \right) \, , \tag{4.137}$$

dependent on metric and it's first derivative. Although it isn't obvious from the compact notation, field equations are of the second order. If the right-hand side is zero the equations describe free gravitational field. If there is a source term on the right-hand side then there is interaction between gravity and matter. Because we wish to find energy and momentum of graviton, we are not interested in source terms.

Let's start the most standard way, from Einstein–Hilbert action

$$S_{EH}[g^{\mu\nu}] = \int d^4x \sqrt{-g} R \, , \tag{4.138}$$

by varying the metric,

$$\begin{aligned} \delta S_{EH} &= \int d^4x \delta \left(g^{\mu\nu} R_{\mu\nu} \sqrt{-g} \right) \\ &= \int d^4x \sqrt{-g} \left((\delta g^{\mu\nu}) R_{\mu\nu} \sqrt{-g} + (\delta R_{\mu\nu}) g^{\mu\nu} \sqrt{-g} + (\delta \sqrt{-g}) R \right) \\ &= \int d^4x \sqrt{-g} \left(G_{\mu\nu} \delta g^{\mu\nu} + g^{\mu\nu} \delta R_{\mu\nu} \right) \, . \end{aligned} \tag{4.139}$$

As for any scalar, symmetry group of the theory described by 4.138 is general diffeomorphism group, so any global sub-group is just a special case of the second Noether's theorem, leading to identity. However, as with the gluon case, we proceed with the calculation. The left-hand side of 4.139 is under general diffeomorphisms equal to

$$\begin{aligned}\mathfrak{L}_\xi S_{EH} &= \int d^4x \mathfrak{L}_\xi \left(\sqrt{-g}R\right) \\ &= \int d^4x \partial_\mu (R\sqrt{-g}\xi^\mu) \\ &= \int d^3\Sigma_\mu \sqrt{-g} R \xi^\mu \,.\end{aligned} \tag{4.140}$$

The second line follows from the fact that Lagrangian is a scalar density. The first term on the right hand side of 4.139 is

$$\delta g^{\mu\nu} = \mathfrak{L}_\xi g^{\mu\nu} = 2\nabla_\mu \xi_\nu \,, \tag{4.141}$$

while variation of Ricci tensor gets us the Palatini identity,

$$\delta R_{\mu\nu} = \nabla_\rho \left(\delta \Gamma^\rho_{\nu\mu}\right) - \nabla_\nu \left(\delta \Gamma^\rho_{\rho\mu}\right) \,, \tag{4.142}$$

so the second term in 4.139 is

$$\begin{aligned} g^{\mu\nu}\delta R_{\mu\nu} &= \nabla_\rho \left(g^{\mu\nu}\delta\Gamma^\rho_{\nu\mu} - g^{\mu\rho}\delta\Gamma^\alpha_{\alpha\mu}\right) \\ &= \nabla_\rho A^\rho \\ &= \frac{1}{\sqrt{-g}} \partial_\rho \left(\sqrt{-g} A^\rho\right) \,, \end{aligned} \tag{4.143}$$

where A^ρ is defined as

$$A^\rho \equiv g^{\mu\nu}\delta\Gamma^\rho_{\nu\mu} - g^{\mu\rho}\delta\Gamma^\alpha_{\alpha\mu} \,. \tag{4.144}$$

As a result, integral free identity is

$$\sqrt{-g} G_{\mu\nu}(2\nabla_\mu \xi_\nu) = -\partial_\mu \left(\sqrt{-g}A^\mu - \sqrt{-g} R \xi^\mu\right) \,. \tag{4.145}$$

Expression 4.145 is analogous to 4.102, which is a step prior to obtaining the second Noether's theorem. To continue, we must find suitable boundary conditions which leave us only with the interior contribution. What's more, appropriate boundary conditions are needed so that field equations can be obtained by extremizing the action. It follows from 4.140 that for variations that vanish at the boundary the contribution of $\partial_\mu(\sqrt{-g}\xi^\mu)$ vanishes, unlike the term $\partial_\mu(\sqrt{-g}A^\mu)$. We solve this problem by adding Gibbons-Hawking-York (GHY) boundary term to Einstein−Hilbert action,

$$S_G[g_{\mu\nu}] = S_{EH} + S_{GHY} = \int_M d^4x\sqrt{-g}R + 2\int_{\partial M} d^3y\varepsilon\sqrt{h}K\,, \tag{4.146}$$

where y are coordinates on the boundary of the manifold, h_{ab} is metric induced on the boundary, ε is equal to $+1$ if the normal to ∂M is timelike and -1 if it's spacelike. Scalar K is extrinsic curvature defined as

$$K = h^{\alpha\beta}\nabla_\beta n_\alpha\,, \quad \alpha,\beta = 1,2,3\,, \tag{4.147}$$

As a consequence, variation of 4.146 results in

$$\delta S_G = \int d^4x\sqrt{-g}G^{\mu\nu}\delta g_{\mu\nu}\,. \tag{4.148}$$

Finally, under general diffeomorphisms such that variations vanish at the boundary,

$$\begin{aligned} 0 &= \int d^4x\sqrt{-g}G^{\mu\nu}2\nabla_\mu\xi_\nu \\ &= \int d^4x\sqrt{-g}G^{\mu\nu}\left(2\partial_\mu\xi_\mu - 2\xi_\lambda\Gamma^\lambda_{\nu\mu}\right)\,. \end{aligned} \tag{4.149}$$

We can now apply the second Noether's theorem. Comparing 4.149 to 4.103, it follows that

$$\begin{aligned} \alpha^\lambda_{\nu\mu} &= -2\Gamma^\lambda_{\nu\mu}\,, \\ \beta^\mu_\nu &= 2\delta^\mu_\nu\,. \end{aligned} \tag{4.150}$$

The second Noether's theorem for gravitation results in

$$\begin{aligned} -2\Gamma^\lambda_{\nu\mu}(\sqrt{-g}G^{\mu\nu}) - \partial_\mu\left(2\sqrt{-g}G^{\mu\lambda}\right) &= 0 \\ \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}G^{\mu\lambda}\right) + \Gamma^\lambda_{\nu\mu}G^{\mu\nu} &= 0 \\ \nabla_\mu G^{\mu\nu} &= 0\,. \end{aligned} \tag{4.151}$$

As expected, we obtained a covariant conservation law, known as (twice) contracted Bianchi identity. The expression is analogous to 4.129 in QCD, but valid in general, not just for global subgroup. For global subgroup $\xi^\mu(x) = const.$, 4.151 reduces to

$$\sqrt{-g}\Gamma^\lambda_{\nu\mu}G^{\mu\nu} = 0 \implies G_{\mu\nu} = 0\,. \tag{4.152}$$

Determinant of the metric and Christoffel symbols are non-zero, otherwise the action identically vanishes. The result is the same as for QCD. The identity 4.151 in the case of global transformations reduces to condition requiring that field equations are satisfied. However, it was explained that $\xi^\mu(x) = const.$ as condition for global subgroup is not general, but applies only to translations. Instead one should start from

$$\sqrt{-g}G_{\mu\nu}\mathfrak{L}_\xi g^{\mu\nu} = \partial_\mu(\sqrt{-g}T^\mu_\nu\xi^\nu) \equiv \partial_\mu(\sqrt{-g}J^\mu)\,. \tag{4.153}$$

and integrate it with boundary conditions such that variations vanish at the boundary, resulting in

$$\sqrt{-g}G_{\mu\nu}\mathfrak{L}_\xi g^{\mu\nu} = 0\,. \tag{4.154}$$

General manifold does not have Killing vectors, which means that for global subgroup we obtain 4.152. The result is the same since $\xi^\mu(x) = const.$ is just a special case belonging to global subgroup.

Now, if Einstein's tensor was antisymmetric instead of symmetric one could proceed as before, defining a conserved current as partial derivative of antisymmetric part of field equations. This is not possible due to symmetric property, but we can learn one more thing from 4.130 — the conserved current is equal to non-linear part of field equations and they result from non-linear terms in action. Thus, we start by separating Einstein-Hilbert action into linear and non-linear terms,

$$
\begin{aligned}
R\sqrt{-g} &= \sqrt{-g}g^{\mu\nu}(\Gamma^{\sigma}_{\mu\sigma,\nu} - \Gamma^{\sigma}_{\mu\nu,\sigma}) - \sqrt{-g}g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \\
&= \partial_{\nu}(\sqrt{-g}g^{\mu\nu}\Gamma^{\sigma}_{\mu\sigma}) - \partial_{\sigma}(\sqrt{-g}g^{\mu\nu}\Gamma^{\sigma}_{\mu\nu}) - \partial_{\nu}(\sqrt{-g}g^{\mu\nu})\Gamma^{\sigma}_{\mu\sigma} + \partial_{\sigma}(\sqrt{-g}g^{\mu\nu})\Gamma^{\sigma}_{\mu\nu} - \\
&\quad \sqrt{-g}g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \\
&= \partial_{\nu}(\sqrt{-g}g^{\mu\nu}\Gamma^{\sigma}_{\mu\sigma}) - \partial_{\sigma}(\sqrt{-g}g^{\mu\nu}\Gamma^{\sigma}_{\mu\nu}) + g^{v\beta}\Gamma^{\mu}_{\beta v}\Gamma^{\sigma}_{\mu\sigma}\sqrt{-g} + \\
&\quad \left(-2g^{\nu\beta}\Gamma^{\mu}_{\beta\sigma} + g^{\mu\nu}\Gamma^{\beta}_{\sigma\beta}\right)\Gamma^{\sigma}_{\mu\nu} - \sqrt{-g}g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \\
&= \partial_{\nu}(\sqrt{-g}g^{\mu\nu}\Gamma^{\sigma}_{\mu\sigma}) - \partial_{\sigma}(\sqrt{-g}g^{\mu\nu}\Gamma^{\sigma}_{\mu\nu}) + \sqrt{-g}g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \,.
\end{aligned}
\tag{4.155}
$$

The second line is obtained by partial integration of the first term. The third line follows from identities

$$\partial_{\sigma}(\sqrt{-g}g^{\mu\nu}) = \sqrt{-g}(-g^{\nu\beta}\Gamma^{\mu}_{\beta\sigma} - g^{\mu\alpha}\Gamma^{\nu}_{\alpha\sigma} + g^{\mu\nu}\Gamma^{\beta}_{\sigma\beta}) \,, \tag{4.156}$$

$$\partial_{\nu}(\sqrt{-g}g^{\mu\nu}) = -\sqrt{-g}g^{\nu\beta}\Gamma^{\mu}_{\beta\nu} \,. \tag{4.157}$$

The first two terms of 4.155 are boundary terms, and they are linear in derivatives of metric, while the last term is non-linear. As field equations are derived by extremizing the action with boundary terms vanishing under appropriate boundary conditions, the first two terms will not contribute to field equations. In other words, one can derive the same field equations using only non-linear terms of the action. What's more, non-linear part contains only metric and it's first derivatives, so all the expression obtained by deriving the Noether's theorems can be directly applied. Hence, action consisting of only non-linear parts is

$$S'_G = \int d^4x\sqrt{-g}g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \,. \tag{4.158}$$

It is invariant under global transformations,

$$\delta g_{\alpha\beta} = \xi^{\mu}\partial_{\mu}g_{\alpha\beta} \,. \tag{4.159}$$

What's more, the action 4.158 does not have a larger local symmetry group, i.e., it is not invariant under general diffeomorphisms. This happened because we left out the boundary terms. Although they are not important for field equations, their role is to make sure the action transforms as a scalar under general diffeomorphisms. Anyhow, the conserved current obtained by Noether's first theorem is

$$\tau_\mu^\nu\sqrt{-g} = \left(\frac{\partial\mathcal{L}'_G}{\partial(\partial_\nu g_{\alpha\beta})}\right)\partial_\mu g_{\alpha\beta} - \delta_\mu^\nu\mathcal{L}'_G\,, \tag{4.160}$$

where $\mathcal{L}'_G = \sqrt{-g}g^{\lambda\tau}(\Gamma^\sigma_{\lambda\tau}\Gamma^\rho_{\sigma\rho} - \Gamma^\rho_{\lambda\sigma}\Gamma^\sigma_{\tau\rho})$. Calculating the derivatives leads to

$$\tau_\mu^\nu\sqrt{-g} = \left[\left(\Gamma^\nu_{\alpha\beta} - \delta^\nu_\beta\Gamma^\sigma_{\alpha\sigma}\right)\partial_\mu\left(g^{\alpha\beta}\sqrt{-g}\right) - \delta_\mu^\nu\sqrt{-g}g^{\lambda\tau}(\Gamma^\sigma_{\lambda\tau}\Gamma^\rho_{\sigma\rho} - \Gamma^\rho_{\lambda\sigma}\Gamma^\sigma_{\tau\rho})\right]. \tag{4.161}$$

Note that energy-momentum pseudotensor consists only of metric and its first derivatives. This is in accordance with equivalence principle which states that locally, in a point, it is possible to choose a frame such that the laws of physics are the same as in special relativity, i.e., gravitational field cannot be detected at a point. Moreover, since the Christoffel symbols and partial derivatives of tensors don't transform as tensors,

$$\Gamma^\alpha_{\beta\gamma} \xrightarrow{y(x)} \frac{\partial x^\mu}{\partial y^\beta}\frac{\partial x^\nu}{\partial y^\gamma}\Gamma^\sigma_{\mu\nu}\frac{\partial y^\alpha}{\partial x^\sigma} + \frac{\partial y^\alpha}{\partial x^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\beta\partial x^\gamma}\,, \tag{4.162}$$

and

$$\partial_\mu g^{\alpha\beta} \xrightarrow{y(x)} \frac{\partial x^\rho}{\partial y^\mu}\frac{\partial y^\alpha}{\partial x^\sigma}\frac{\partial y^\beta}{\partial x^\tau}\partial_\rho g^{\sigma\tau} + \frac{\partial x^\rho}{\partial y^\mu}\frac{\partial^2 y^\beta}{\partial x^\rho x^\tau}\frac{\partial y^\alpha}{\partial x^\sigma}g^{\sigma\tau} + \frac{\partial x^\rho}{\partial y^\mu}\frac{\partial^2 y^\alpha}{\partial x^\rho x^\sigma}\frac{\partial y^\beta}{\partial x^\tau}g^{\sigma\tau}\,, \tag{4.163}$$

the expression does not transform covariantly under general coordinate transformations, unless the transformation is linear. Thus, we have found a gravitational energy-momentum tensor that transforms covariantly under Lorentz transformations as Weinberg–Witten theorem requires. One can check that it is indeed conserved,

$$\begin{aligned}\partial_\nu(\sqrt{-g}\tau_\mu^\nu) &= \left[\frac{\partial\mathcal{L}'_G}{\partial g_{\alpha\beta}} - \partial_\rho\left(\frac{\partial\mathcal{L}'_G}{\partial(\partial_\rho g_{\alpha\beta})}\right)\right]\partial_\mu g_{\alpha\beta} \\ &= G^{\alpha\beta}\partial_\mu g_{\alpha\beta}\,.\end{aligned} \tag{4.164}$$

If field equations are satisfied, one obtains conserved energy-momentum tensor. The explicit calculation is skipped as the variation is straightforward. The four-momentum is then defined as integral of the zeroth components of the pseudotensor.

$$P^{\mu} = \int d^3 x \tau^{0\mu} \,. \tag{4.165}$$

One should note that it is not possible to obtain an expression for energy and momentum of gravitational field satisfying at the same time claims that

- When gravitational energy is added to other forms of energy the total energy is conserved
- Energy and momentum within a three-dimensional region at some point in time are independent of coordinate system.

What's more, because we are dealing with pseudo-tensor instead of tensor, it has no meaning to speak about localization of gravitational energy in space, since by performing non-linear transformations one may obtain the value of τ different from zero in flat spacetime.

As a final remark, there are various attempts to find local energy-momentum density for gravitation. They all give non-covariant pseudo-tensors. Some examples include Landau and Lifshitz [91] and Weinberg [89]. We opted for this procedure because it is consistent with the following step.

The next goal is to find an expression for τ in terms of gravitational field propagating on flat background. This is done by considering

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \,, \quad |h_{\mu\nu}| << 1 \,, \tag{4.166}$$

where is dynamical field $h_{\mu\nu}$, representing graviton, and flat background is considered consider as, i.e., $h_{\mu\nu}$ is small perturbation on fixed background. The actual metric field is then given by $g_{\mu\nu}$ in 4.166. By plugging 4.166 into 4.161 one obtains gravitational energy-momentum tensor in terms of field $h_{\mu\nu}$. What's more, it can be shown that starting from 4.166 one may obtain the action 4.158 [26][34].

Next, notice that, as it was shown for four-vector field related to massless particles 4.75, field $h_{\mu\nu}$ under Lorentz transformations transforms as

$$h_{\mu\nu}(x) \rightarrow \Lambda^{\alpha}_{\mu} \Lambda^{\beta}_{\nu} h_{\alpha\beta} + \partial_{\mu}\Omega_{\nu} + \partial_{\nu}\Omega_{\mu} \,. \tag{4.167}$$

Lorentz transformations change the gauge of $h_{\mu\nu}$ [41][92][34][93].

We have obtained conserved currents for both QCD and gravitation. This finishes the discussion about conserved quantities. Before we explain how these theories avoid the Weinberg–Witten theorem, consider the proof.

4.4 *Proof of Weinberg–Witten theorem*

The proof consists of looking at the following matrix elements

$$\langle p', \pm j | J^{\mu} | p, \pm j \rangle \,, \tag{4.168}$$

$$\langle p', \pm j | T^{\mu\nu} | p, \pm j \rangle \ , \tag{4.169}$$

where $|p, \pm j\rangle$ and $|p', \pm j\rangle$ are massless, one-particle states, carrying a non-zero conserved charge, labeled by four-momentum and helicity, which is the same for both states. The motivation for these matrix elements is that charges, energies and momenta are experimentally determined by looking at nearly forward scattering caused by exchange of spacelike, but nearly lightlike massless gauge boson, corresponding to $(p' - p) \to 0$, i.e. $p' \to p$.

$$\lim_{p' \to p} \langle p', \pm j | J^{0} | p, \pm j \rangle \propto q \,, \tag{4.170}$$

$$\lim_{p' \to p} \langle p', \pm j | P^{\mu} | p, \pm j \rangle \propto p^{\mu} \,, \tag{4.171}$$

where charges q and p are defined by measurement process. By defining it as the limit, we do not require continuity of the matrix elements at $p' = p$. This is important because there are examples of currents with discontinuity at $p' = p$ caused by the change in the sign of the polarization when momentum transfer changes from spacelike to null-like.

To start with the proof, note that due to assumption that particles carry charges $Q = \int d^3x J^0$ and $P^{\mu} = \int d^3x T^{0\mu}$, it is true that

$$Q | p, \pm j \rangle = q | p, \pm j \rangle \,, \tag{4.172}$$

$$P^{\mu} | p, \pm j \rangle = p^{\mu} | p, \pm j \rangle \,, \tag{4.173}$$

which leads to

$$\langle p', \pm j|Q|p, \pm j\rangle = q\langle p' \pm j|p, \pm j\rangle\,, \tag{4.174}$$

$$\langle p', \pm j|P^\mu|p, \pm j\rangle = p^\mu\langle p' \pm j|p, \pm j\rangle\,. \tag{4.175}$$

One should be careful when evaluating the right-hand side. The states are usually normalised by $\langle p' \pm j|p, \pm j\rangle = \delta^{(3)}(\boldsymbol{p'} - \boldsymbol{p})$, i.e., momentum is sharply defined. But we know that in reality, physical one-particle states are smeared. As a result physically correct normalization is

$$\langle p', \pm j|p, \pm j\rangle = \delta_a^{(3)}(\boldsymbol{p'} - \boldsymbol{p})\,, \tag{4.176}$$

where δ_a is an approximate delta function, such that

$$\lim_{a\to 0} \delta_a^{(3)}(\boldsymbol{p'} - \boldsymbol{p}) = \delta^{(3)}(\boldsymbol{p'} - \boldsymbol{p})\,. \tag{4.177}$$

The parameter a is determined by level of sharpness in the experiment. What's more, although physical states consist of superposition of pure momentum states, we will identify them with eigenvalue p^μ, where they have a sharp peak. From this moment on, all one-particle states will refer to physical states normalized by 4.176. Hence, 4.175 reduces to

$$\langle p', \pm j|Q|p, \pm j\rangle = q\delta_a^{(3)}(\boldsymbol{p'} - \boldsymbol{p})\,, \tag{4.178}$$

$$\langle p', \pm j|P^\mu|p, \pm j\rangle = p^\mu\delta_a^{(3)}(\boldsymbol{p'} - \boldsymbol{p})\,. \tag{4.179}$$

By integrating 4.179 using the definition of charge as integral of time-components of conserved current, one obtains

$$\begin{aligned}\langle p', \pm j|Q|p, \pm j\rangle &= \int_{V_a} d^3x \left\langle p', \pm j \left| J^0(t, \boldsymbol{x}) \right| p, \pm j\right\rangle \\ &= \int_{V_a} d^3x \left\langle p', \pm j \left| e^{i\boldsymbol{P}\cdot\boldsymbol{x}} J^0(t, 0) e^{-i\boldsymbol{P}\cdot\boldsymbol{x}} \right| p, \pm j\right\rangle \\ &= \int_{V_a} d^3x e^{i(\boldsymbol{p'}-\boldsymbol{p})\boldsymbol{x}} \left\langle p', \pm j \left| J^0(t, 0) \right| p, \pm j\right\rangle \\ &= (2\pi)^3 \delta_a^3 (\boldsymbol{p'} - \boldsymbol{p}) \left\langle p', \pm j \left| J^0(t, 0) \right| p, \pm j\right\rangle\,.\end{aligned} \tag{4.180}$$

And completely analogous,

$$\begin{aligned}
\langle p', \pm j | P^\mu | p, \pm j\rangle &= \int_{V_a} d^3x \left\langle p', \pm j \left| T^{0\mu}(t, \boldsymbol{x}) \right| p, \pm j \right\rangle \\
&= \int_{V_a} d^3x \left\langle p', \pm j \left| e^{i\boldsymbol{P}\cdot\boldsymbol{x}} T^{0\mu}(t, 0) e^{-i\boldsymbol{P}\cdot\boldsymbol{x}} \right| p, \pm j \right\rangle \\
&= \int_{V_a} d^3x e^{i(\boldsymbol{p}'-\boldsymbol{p})\boldsymbol{x}} \left\langle p', \pm j \left| T^{0\mu}(t, 0) \right| p, \pm j \right\rangle \\
&= (2\pi)^3 \delta^3_a (\boldsymbol{p}' - \boldsymbol{p}) \left\langle p', \pm j \left| T^{0\mu}(t, 0) \right| p, \pm j \right\rangle \,.
\end{aligned} \tag{4.181}$$

Integral is evaluated over a large but finite volume. Comparing 4.178 with 4.180, and 4.179 with 4.181, one is lead to conclusion that

$$\lim_{p' \to p} \left\langle p', \pm j \left| J^0(t, 0) \right| p, \pm j \right\rangle = \frac{q}{(2\pi)^3} \,, \tag{4.182}$$

$$\lim_{p' \to p} \left\langle p', \pm j \left| T^{0\mu}(t, 0) \right| p, \pm j \right\rangle = \frac{p^\mu}{(2\pi)^3} \,. \tag{4.183}$$

Hence, for general components of the currents,

$$\lim_{p' \to p} \left\langle p', \pm j \left| J^\mu(t, 0) \right| p, \pm j \right\rangle = \frac{q p^\mu}{(2\pi)^3 E} \,, \tag{4.184}$$

$$\lim_{p' \to p} \left\langle p', \pm j \left| T^{\mu\nu}(t, 0) \right| p, \pm j \right\rangle = \frac{p^\mu p^\nu}{(2\pi)^3 E} \,. \tag{4.185}$$

Let's assume that the currents transform covariantly under Lorentz transformations. As a consequence, then so should the right-hand side of 4.185. Finally, if we assume that the currents are properly conserved the Lorentz covariant quantity that should appear on the right-hand side is momentum p. Too see this, we multiply 4.184 and 4.185 by p_μ, which results in

$$\langle J^\mu(t, 0)\rangle p_\mu = 0 \,, \tag{4.186}$$

$$\langle T^{\mu\nu}(t, 0)\rangle p_\mu = 0 \,, \tag{4.187}$$

where the right-hand side vanishes because particles are massless, $p_\mu p^\mu = 0$, while the left-hand side can be understood as Fourier-transform of $\partial_\mu J^\mu$ and $\partial_\mu T^{\mu\nu}$. In other words, writing the currents in terms of creation and annihilation operators and acting with the differential operator lowers p_μ from phase factor multiplying the annihilation and creation operators. Hence, the form 4.184 and 4.185 is valid if the currents are Lorentz covariant and conserved. One more thing to notice is that, if we would look to only sharply defined states, 4.184 and 4.185 would hold only where the delta function does not vanish, for $\boldsymbol{p'} = \boldsymbol{p}$. This completes the first part which shows that if particle is charged under conserved, Lorentz covariant Noether currents, the matrix elements do not vanish. Next, we show that this statement leads to contradiction if helicity of the particle is $j > \frac{1}{2}$ according to first part, and $j > 1$ according to second part of the theorem. Thus, consider light-like p and p',

$$\begin{aligned}(p'+p)^2 &= 2(p'p) \\ &= 2\left(|\boldsymbol{p}'|\,|\boldsymbol{p}| - \boldsymbol{p}'\cdot\boldsymbol{p}\right) \\ &= 2\,|\boldsymbol{p}'|\,|\boldsymbol{p}|(1-\cos\phi) \leq 0\,,\end{aligned} \tag{4.188}$$

where ϕ is an angle between $\boldsymbol{p}$ and $\boldsymbol{p'}$. If $\phi \neq 0$ the total momentum is timelike and we can choose a frame such that the space component of total momentum vanishes,

$$p = (|\boldsymbol{p}|, \boldsymbol{p}), \quad p' = (|\boldsymbol{p}|, -\boldsymbol{p})\,. \tag{4.189}$$

Because we are only interested in the limit $p' \to p$, it is always true that $\phi \neq 0$. Further, in such frame consider rotation of the particles by an angle θ around axis in direction of $\boldsymbol{p}$,

$$\begin{aligned}|p, \pm j\rangle &\to U(R_W(\theta))|p, \pm j\rangle = e^{\pm i\theta j}|p, \pm j\rangle\,, \\ |p', \pm j\rangle &\to U(R_W(\theta))\,|p', \pm j\rangle = e^{\mp i\theta j}|p', \pm j\rangle\,,\end{aligned} \tag{4.190}$$

where the p' state has opposite phase compared to p because rotation by θ in the direction of $\boldsymbol{p}$ is the same as rotation around $\boldsymbol{p}'$ in $-\theta$ direction, since their momenta are antiparallel. On the other hand, instead of transforming the states, one can transform the operators. As J and T are Lorentz covariant quantities by assumption the right-hand side is

$$e^{\pm 2i\theta j}\left\langle p', \pm j\left|J^{\mu}(t,0)\right|p, \pm j\right\rangle = R_W(\theta)^{\mu}_{\nu}\left\langle p', \pm j\left|J^{\nu}(t,0)\right|p, \pm j\right\rangle\,, \tag{4.191}$$

$$e^{\pm 2i\theta j}\left\langle p', \pm j\left|T^{\mu\nu}(t,0)\right|p, \pm j\right\rangle = R_W(\theta)^{\mu}_{\rho}R_W(\theta)^{\nu}_{\sigma}\left\langle p', \pm j\left|T^{\rho\sigma}(t,0)\right|p, \pm j\right\rangle\,, \tag{4.192}$$

where $R_W(\theta)$ is the proper rotation matrix in two dimensions in three dimensional space. By explicit calculation, one can check that eigenvalues of rotation matrix are

$e^{\pm i\theta}$ and 1, so

$$e^{\pm 2i\theta j} \langle p', \pm j \left| J^{\mu}(t,0) \right| p, \pm j \rangle = e^{\pm i\theta} \langle p', \pm j \left| J^{\mu}(t,0) \right| p, \pm j \rangle \ , \tag{4.193}$$

$$e^{\pm 2i\theta j} \langle p', \pm j \left| T^{\mu\nu}(t,0) \right| p, \pm j \rangle = e^{\pm 2i\theta} \langle p', \pm j \left| T^{\mu\nu}(t,0) \right| p, \pm j \rangle \ , \tag{4.194}$$

or,

$$e^{\pm 2i\theta j} \langle p', \pm j \left| J^{\mu}(t,0) \right| p, \pm j \rangle = \langle p', \pm j \left| J^{\mu}(t,0) \right| p, \pm j \rangle \ , \tag{4.195}$$

$$e^{\pm 2i\theta j} \langle p', \pm j \left| T^{\mu\nu}(t,0) \right| p, \pm j \rangle = \langle p', \pm j \left| T^{\mu\nu}(t,0) \right| p, \pm j \rangle \ , \tag{4.196}$$

From 5.119 and 4.195 one can see that $j = \{0, \frac{1}{2}\}$, while from 4.194 and 4.196 $j = \{0, \frac{1}{2}, 1\}$. This completes the proof. Under the assumptions of Lorentz covariance and conservation of currents,

$$\lim_{p' \to p} \langle p', \pm j \left| J^{\mu}(t,0) \right| p, \pm j \rangle = 0 \,, \quad j > \frac{1}{2} \,, \tag{4.197}$$

$$\lim_{p' \to p} \langle p', \pm j \left| T^{\mu\nu}(t,0) \right| p, \pm j \rangle = 0 \,, \quad j > 1 \,. \tag{4.198}$$

According to the first part of the theorem, no elementary or composite massles, charged particle with $j > \frac{1}{2}$ exists. The second part states that elementary or composite massless particle with spin $j > 1$ cannot carry energy or momentum. To reformulate, the first part of the theorem allows the existence of uncharged massless particles of spin $j > \frac{1}{2}$, while the second part, as it is stated, does not permit the existence of massless particles of spin $j > 1$ even if they are uncharged with respect to translation symmetry, i.e., even if they do not carry energy and momentum. First, it seams that such particles can only correspond to vacuum, as they belong to a class of irreducible representations invariant under entire Poincaré group. However, in theories with non-minimal coupling, there appear so-called stealth fields, non-trivial fields whose energy-momentum vanishes [94].

In the Standard Model (where the Lagrangian contains only minimal coupling), out of particles proved to exist, only photons and gluons are massless and are both spin-1 particles. Photons are not charged under electromagnetic current thus they are in accordance with the Weinberg–Witten theorem, while gluons carry color charge. Particle hypothesized to exist is graviton, massless spin-2 particle. Let's explain how gluons and gravitons avoid the theorem.

4.5 Lorentz covariance and gauge transformations

It was shown that massless fields do not transform covariantly under Lorentz transformations. As a result, nor the gluon current, nor the gravitational energy-momentum pseudotensor transform covariantly.

The source of this behaviour is in the little group of massless particles. The field transform under $E2$, since there is no reason to restrain the group to $SO(2)$, as fields aren't physical states. Lorentz covariance is saved by proclaiming that the fields related by transformations of the form

$$A_\mu \to A_\mu + \partial_\mu \xi \,, \tag{4.199}$$

or, for symmetric second rank field

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu \,, \tag{4.200}$$

describe the same physical configuration, and are in this sense not relevant. Furthermore, such fields appear in Lagrangian either as coupled to conserved currents or used to construct Lorentz covariant tensors. In the first case, the theory remains Lorentz invariant because

$$A_\mu j^\mu \to A_\mu j^\mu + \partial_\mu \xi j^\mu = A_\mu j^\mu + \partial_\mu (\xi j^\mu) - \xi \partial_\mu j^\mu \,. \tag{4.201}$$

By partial integration one obtains a boundary term, which is a symmetry transformation by definition, while the second term vanishes due to current conservation. An example of the second case is the gluon Lagrangian itself, as $F^{\mu\nu}$ is generally covariant tensor constructed out of derivatives of A_μ.

In short, gluons and gravitons evade the Weinberg–Witten theorem because the conserved currents are not Lorentz covariant. The most important consequence is the fact that graviton cannot be composed of gluons for example. Gluons are matter, so they would contribute to conserved Lorentz covariant energy-momentum tensor A.16. As a result, momentum of graviton composed of gluons would be described by $P^\mu = \int T^{0\mu} d^3x$, i.e., charged under the conserved current. Consequently, theories living on flat background cannot mimic the properties of gravity, so the graviton cannot be a composite particle [95][96][97].

5 Thermodynamics and General Relativity

In this section the goal is to explain the connection between gravitation and thermodynamics. First, we derive the four laws of black hole mechanics. We also discuss their analogy with the four laws of thermodynamics. Next, Hawking derivation is derived. Finally, we consider a possible thermodynamic behaviour of the spacetime itself.

5.1 Classical black hole mechanics

In this section we briefly review the most important theorems and properties related to black holes, starting with some mathematical preliminaries needed to understand the definitions, after which we will discuss the four laws of black hole mechanics. Throughout the text we will make comparison with thermodynamics to motivate the choices and restriction we make, assuming the reader is familiar with the mentioned concepts from thermodynamics.

5.1.1 Mathematical preliminaries

The goal of this section is to explain what is a hypersurface and focus on null hypersurfaces that are generated by Killing vector fields. We also derive Raychaudhuri's equation, describing how the neighbouring geodesics evolve as one moves along the congruence.

Hypersurface Σ is defined as $(n-1)$-dimensional manifold embedded into n-dimensional manifold. To define a hypersurface one can restrict the spacetime coordinates in some way, usually by imposing that some function of coordinates $\Phi(x^\mu)$ is constant or zero[33],

$$\Phi(x^\mu) = 0\,. \tag{5.1}$$

With such definition of hypersurface, one-form n_μ is normal to Σ if

$$n_\mu \propto \partial_\mu \Phi\ . \tag{5.2}$$

If the normal is null vector, i.e., $n^2 = 0$, the hypersurface is said to be *null* hypersurface. Furthermore, a vector l is tangent to Σ if $l \cdot n = 0$. Since the null vectors satisfy this by definition, in the null case, k is both normal and tangent to the hypersurface.

We will now consider a special case. Let χ^μ be a normal to null hypersurface defined as $\Phi = \chi^\mu \chi_\mu = 0$, where χ^μ satisfies the Killing equation

$$\nabla_\mu \chi_\nu + \nabla_\nu \chi_\mu = 0\,. \tag{5.3}$$

Because everywhere on the null hypersurface, $\Phi = \chi^\mu \chi_\mu = 0$, for any vector l tangent to it, $l^\mu \partial_\mu \Phi = 0$. Consequently, since χ is both normal and parallel to the hypersurface,

$$\partial_\mu \Phi = -2\kappa \chi_\mu\,, \tag{5.4}$$

with the proportionality constant of 5.2 chosen as -2κ, where κ is a scalar that can depend on coordinates. Moreover, as mentioned, one can associate integral curves $x^\mu(\lambda)$ to a vector field, such that tangent to the curve is equal to the vector field at the same point,

$$\chi^\mu = \frac{dx^\mu}{d\lambda}\,, \tag{5.5}$$

The null curves cover the hypersurface so they are referred to as generators of the hypersurface.

Relation 5.4 is a statement of Frobenius theorem, according to which a congru-

[33] For example a two-sphere in three dimensional spacetime can be defined as $\Phi(x, y, z) = x^2 + y^2 + z^2 = R^2$, i.e., $\Phi(x, y, z) = x^2 + y^2 + z^2 - R^2 = 0$.

ence of curves whose tagnent vector is u is hypersurface orthogonal if

$$u_{[\alpha}\nabla_{\beta}u_{\gamma]} = 0\,, \tag{5.6}$$

This follows by direct calculation from $u_\mu = -\alpha(x^\mu)\partial_\mu\Phi$, for some scalar field α. One can see that this is true for χ from 5.4. In other words, when tangent of the curves is everywhere proportional to normal of the hypersurface, the congruence is everywhere orthogonal to hypersurface. This will have important consequence of behaviour of neighbouring curves.

It also follows immediately from 5.3 and 5.4 that χ satisfies non-affinely parametrized geodesic equation,

$$\partial_\nu(\chi^\mu\chi_\mu) = 2\chi_\mu\nabla_\nu\chi^\mu = -2\chi^\mu\nabla_\mu\chi_\nu \implies \chi^\mu\nabla_\mu\chi_\nu = \kappa\chi_\nu\,, \tag{5.7}$$

where all the expressions are evaluated at the horizon. The first equality follows from the fact that the action of partial derivative is the same as covariant derivative for a scalar. The second equality is the result of 5.3. The implication follows from 5.4. The parameter κ is called *surface gravity* for reasons that will be explained below.

To sum up, the congruence of null curves associated with a Killing vector field satisfy non-affinely parametrized geodesic equation. Moreover, the congruence is hypersurface orthogonal.

We will now describe with the behaviour of the neighbouring geodesics in the congruence. To begin, we will set up a chart adapted to congruence generating the null hypersurface. In other words, one can set a “natural" chart y^μ covering a patch of spacetime, with the help of integral curves, starting from the hypersurface. The labels are then carried along the geodesics as we move along them.

Assume that each curve is parametrised by λ, running along the curve, and θ^A, $A = 2, 3$, constant on each curve. Hence, $y = (\lambda, \theta^A)$, and the basis $\{e^\mu_a\}$, $a = 1, 2, 3$ induced by the chart is such that

$$e^\mu_1 = \left(\frac{\partial x^\mu}{\partial\lambda}\right)_{\theta^A} = \chi^\mu\,. \tag{5.8}$$

Parameter λ is chosen so that the partial derivative with respect to it coincides with

χ^μ. The remaining coordinates are chosen so that the respective vectors e^μ_A are normal to χ, i.e., $\chi_\mu e^\mu_A = 0$,

$$e^\mu_A = \left(\frac{\partial x^\mu}{\partial \theta^A}\right)_\lambda . \tag{5.9}$$

To have a complete basis we need one more vector (tangent space is four dimensional), which the previous conditions of χ being null and orthogonality specify uniquely,

$$N^\mu N_\mu = 0\,, \quad \chi^\mu N_\mu = -1\,, \quad N^\mu e^A_\mu = 0\,, \tag{5.10}$$

Vector N^μ point in the other null direction and is called auxiliary null vector. Thus, the basis consists of vectors $\{e^\mu_a, N^\mu\}$.

Next, we introduce ξ, called a deviation vector. Since it is to describe how the geodesics change relative to one another, it should point from some arbitrarily picked geodesic in the congruence to the neighbouring one. This is achieved by introducing a set of auxiliary curves, “crossing" the congruence, parametrized by s, so that $\frac{\partial x^\mu}{\partial s} = \xi^\mu$ and

$$\xi^\mu \chi_\mu = 0\,. \tag{5.11}$$

This condition rules out component of ξ in the direction of N, confining the vector onto the null hypersurface. In other words, $\xi^\mu = a\chi^\mu + b^A e^\mu_A$ in the basis the congruence induced on the null hypersurface. We are not interested in the deviation vector itself, but the way it evolves as we move along the congruence, i.e.

$$\chi^\mu \nabla_\mu \xi^\nu = B^\nu_\mu \xi^\mu\,, \tag{5.12}$$

where the change in ξ is given by tensor $B_{\mu\nu}$[34]. Thus, the goal is to determine $B_{\mu\nu}$, or more precisely, its transverse component. First of all, note that since $\frac{\partial x^\mu}{\partial s} = \xi^\mu$ and

[34]To motivate 5.12 picture a rubber sheet with displacement between two neighboring points $\boldsymbol{x}_P$ and $\boldsymbol{x}_Q$ on the sheet as given by $\boldsymbol{\xi} = \boldsymbol{x}_Q - \boldsymbol{x}_P$. Suppose the sheet evolves in some way, for example, it is being stretched. The change in displacement is then given by $\frac{d\xi^i}{dt} = v^i(\boldsymbol{\xi} + \boldsymbol{x}_P) - v^i(\boldsymbol{x}_P) \approx \xi^j \partial_j v^i(\boldsymbol{x}_P) \equiv B^i_j(\boldsymbol{x}_P)\xi^j$.

$\frac{\partial x^\mu}{\partial \lambda} = \chi^\mu$, it follows that

$$\left.\begin{aligned} \frac{\partial \xi^\mu}{\partial \lambda} &= \frac{\partial^2 x^\mu}{\partial \lambda \partial s} = \chi^\nu \nabla_\nu \xi^\mu \,, \\ \frac{\partial \chi^\mu}{\partial s} &= \frac{\partial^2 x^\mu}{\partial s \partial \lambda} = \xi^\nu \nabla_\nu \chi^\mu \end{aligned}\right\} \quad \chi^\nu \nabla_\nu \xi^\mu = \xi^\nu \nabla_\nu \chi^\mu \,. \tag{5.13}$$

Eq. 5.13 follows from Schwarz's theorem — one can exchange the order of the second partial derivatives. Hence, inserting 5.13 into 5.12 we can conclude that

$$B_{\mu\nu} = \nabla_\nu \chi_\mu \,. \tag{5.14}$$

The transverse component of 5.14 can be obtained using only transverse part of the metric $h_{\mu\nu}$ given by 5.15,

$$h_{\mu\nu} = g_{\mu\nu} + \chi_\mu N_\nu + N_\mu \chi_\nu \,. \tag{5.15}$$

As one can check by explicit calculation, $h_{\mu\nu}\chi^\nu = 0$ and $h_{\mu\nu}N^\nu = 0$. Consequently, $h^\mu_\nu \xi^\nu$ will give only the transverse component of ξ as we wanted. Likewise,

$$\begin{aligned} \widetilde{B}_{\mu\nu} &= h^\rho_\mu h^\tau_\nu B_{\rho\tau} \\ &= B_{\mu\nu} + \chi_\mu N^\rho B_{\rho\nu} + \chi_\nu B_{\mu\rho} N^\rho + \chi_\mu k_\nu B_{\rho\tau} N^\rho N^\tau \,, \end{aligned} \tag{5.16}$$

is only transverse part of $B_{\mu\nu}$. The second line follows from expanding h^μ_ν using 5.15. It is also useful to decompose $\widetilde{B}_{\mu\nu}$ into antisymmetric and symmetric part,

$$\widetilde{B}_{\mu\nu} = \frac{1}{2}\theta h_{\mu\nu} + \sigma_{\mu\nu} + \omega_{\mu\nu} \,, \tag{5.17}$$

where the symmetric part is further decomposed into trace $\theta = g^{\mu\nu}\widetilde{B}_{\mu\nu}$ and symmetric traceless part $\sigma_{\mu\nu} = \widetilde{B}_{\mu\nu} - \frac{1}{2}\theta h_{\mu\nu}$. The antisymmetric part is $\omega_{\mu\nu} = \widetilde{B}_{[\mu\nu]}$. The trace θ is referred to as the (fractional) *expansion*[35] of the congruence, describing how the volume or the area changes as we move along the congruence, while $\sigma_{\mu\nu}$ and $\omega_{\mu\nu}$ describe the change in shape. The part $\sigma_{\mu\nu}$ is called the *shear* describing "stretch-

[35]The expansion $\theta = \nabla_\mu \chi^\mu - \kappa$ when explicitly evaluated. The second term follows from the congruence being parametrized by non-affine parameter. One can recognize that expansion corresponds to divergence of some vector field $\boldsymbol{E}$, i.e., $\boldsymbol{\nabla E} = \partial_i E^i$. Hence, expansion captures focusing or defocusing of the geodesics.

ing" or "compression" of the shape (without the change of area)[36], and $\omega_{\mu\nu}$ is *twist*[37]. The names are suggestive of their role which is familiar from continuum mechanics where one describes deformation of some medium. In general, the congruence will evolve as a combination of expansion, shear and rotation. However, according to the Frobenius theorem, when the congruence is hypersurface orthogonal the twist, anti-symmetric part of $\widetilde{B}_{\mu\nu}$ must vanish, $\omega_{\mu\nu} = 0$. This can be shown by direct calculation if one applies 5.6 to 5.16.

Moreover, one can derive the equation describing how each of the terms in the expansion change as we move along the congruence, that is $\chi^\rho \nabla_\rho \widetilde{B}_{\mu\nu}$. For what follows, we are interested only in the evolution of the expansion θ, which is obtained by taking the trace — $g^{\mu\nu}\chi^\rho\nabla_\rho\widetilde{B}_{\mu\nu}$. Let's start from

$$\begin{aligned} g^{\mu\nu}\widetilde{B}_{\mu\nu} &= g^{\mu\nu}B_{\mu\nu} + N^\rho B_{\rho\nu}\chi^\nu \\ &= g^{\mu\nu}B_{\mu\nu} - \kappa\,. \end{aligned} \tag{5.18}$$

The last line is the result of χ satisfying non-affinely parametrized geodesic equation 5.7 and the equation 5.10. Hence,

$$\begin{aligned} \frac{d\theta}{d\lambda} &= \chi^\rho\nabla_\rho(g^{\mu\nu}\widetilde{B}_{\mu\nu}) \\ &= \chi^\rho\nabla_\rho g^{\mu\nu}B_{\mu\nu} - \chi^\rho\nabla_\rho\kappa\,. \end{aligned} \tag{5.19}$$

The first term can further be written as

$$\begin{aligned} \chi^\rho\nabla_\rho g^{\mu\nu}B_{\mu\nu} &= \chi^\rho\nabla_\rho\nabla_\mu\chi^\mu \\ &= \chi^\rho\nabla_\mu\nabla_\rho\chi^\mu - R_{\rho\mu}\chi^\rho\chi^\mu \\ &= \nabla_\mu(\chi^\rho\nabla_\rho\chi^\mu) - (\nabla_\mu\chi^\rho)(\nabla_\rho\chi^\mu) - R_{\rho\mu}\chi^\rho\chi^\mu \\ &= \nabla_\mu(\kappa\chi^\mu) - B^{\rho\mu}B_{\mu\rho} - R_{\rho\mu}\chi^\rho\chi^\mu\,. \end{aligned} \tag{5.20}$$

The second line follows from the definition of the Ricci tensor $-R^\alpha_{\nu\alpha\beta}\chi^\nu = \nabla_\alpha\nabla_\beta\chi^\alpha - \nabla_\beta\nabla_\alpha\chi^\alpha$, the third from the derivative rules and the last one from recognizing the geodesic equation and definition of $B_{\mu\nu}$, 5.14.

[36] In continuum mechanics there is a (linear) strain tensor ε_{ij} which combines the effects of expansion and shear, $\varepsilon_{ij} = \frac{1}{2}\left(\partial_i u_j + \partial_j u_i\right)$, where u_i is a component of displacement vector.

[37] As the name suggests, twist, or also referred to as *rotation* corresponds to curl i.e., $\omega_{\mu\nu} = \varepsilon_{\mu\nu\rho\tau}\widetilde{B}^{\rho\tau}$. The curl of some vector field $\boldsymbol{B}$ is $(\nabla\times\boldsymbol{B})_i = \varepsilon_{ijk}\partial^j B^k$.

Inserting 5.20 into 5.19 results in

$$\begin{aligned}\frac{d\theta}{d\lambda} &= \nabla_\rho(\kappa\chi^\rho) - B^{\mu\nu}B_{\mu\nu} - R_{\mu\nu}\chi^\mu\chi^\nu - \chi^\rho\nabla_\rho\kappa \\ &= \kappa\theta - B^{\mu\nu}B_{\nu\mu} - R_{\mu\nu}\chi^\mu\chi^\nu \,.\end{aligned} \tag{5.21}$$

Next, by explicit calculation from 5.16, one can check that $B^{\mu\nu}B_{\nu\mu} = \widetilde{B}^{\mu\nu}\widetilde{B}_{\nu\mu}$. Then, using 5.17,

$$\begin{aligned}B^{\mu\nu}B_{\nu\mu} &= \left(\frac{1}{2}\theta h^{\mu\nu} + \sigma^{\mu\nu} + \omega^{\mu\nu}\right)\left(\frac{1}{2}\theta h_{\nu\mu} + \sigma_{\nu\mu} + \omega_{\nu\mu}\right) \\ &= \frac{1}{2}\theta^2 + \sigma^{\mu\nu}\sigma_{\nu\mu} - \omega^{\nu\mu}\omega_{\nu\mu} \,.\end{aligned} \tag{5.22}$$

The last line follows from $h^{\mu\nu}h_{\nu\mu} = 2$, antisymmetry of $\omega_{\mu\nu}$ and the fact that all terms are normal to each other. Finally, the equation that describes evolution of expansion θ along the congruence is

$$\frac{d\theta}{d\lambda} = \kappa\theta - \frac{1}{2}\theta^2 - \sigma^{\mu\nu}\sigma_{\nu\mu} + \omega^{\nu\mu}\omega_{\nu\mu} - R_{\mu\nu}k^\mu k^\nu \,, \tag{5.23}$$

which is known as Raychaudhuri's equation.

Let's now consider the case in which the null hypersurface is generated by Killing vector field. As mentioned, null hypersurface generated by Killing vector field is hypersurface orthogonal, so the twist vanishes. Furthermore, as χ satisfies the Killing equation 5.3, the symmetric part of $B_{\mu\nu}$ has to vanish. Consequently, 5.23 reduces to

$$R_{\mu\nu}\chi^\mu\chi^\nu = 0 \,. \tag{5.24}$$

In summary, we have shown that the expansion, shear and twist of the null congruence associated with Killing vector field vanish. We now turn to discussing black holes.

5.1.2 Black holes

Black hole (BH) is a *region* of spacetime where gravitation is so strong nothing can escape. Anything that enters the black hole region will remain there. In other words, to escape the black hole one would need to exceed the the speed of light, which, according to special theory of relativity, no physical entity can. The boundary of BH is called an *event horizon*. These notions are made precise by specifying where the escape could happen to — usually, to infinity — capturing the idea that light cannot propagate an arbitrary distance away from the BH. The exterior region can then be thought of as being infinitely far away from black hole. That is, in the exterior anything can at least in principle escape to infinity [98][99].

We will consider only asymptotically flat[38] spacetime. Loosely speaking, for asymptotically flat spacetime, the region of spacetime "near infinity" has the causal structure like flat spacetime. Asymptotically flat spacetime represent gravitationally isolated system. That is, a system we are considering is the only thing in spacetime, so that far away from it the spacetime is flat [98]. As a consequence of asymptotic flatness, we have a well defined notion of past and future null infinity $\mathcal{J}^{\pm}$, so we can now give a formal definition of black hole [71][98][100][101].

In asymptotically flat spacetime, black hole $\mathcal{B}$ is defined as the set of all events of spacetime (M, g) that are not part of the causal past J^{-} of the future null infinity $\mathcal{J}^{+}$, an end point of future directed null curves.

$$\mathcal{B} = M - J^{-}(\mathcal{J}^{+})\,. \tag{5.25}$$

The event horizon $\mathcal{H}$ is a null hypersurface defined as the boundary of $\mathcal{B}$,

$$\mathcal{H} = \partial\mathcal{B} = \partial(J^{-}(\mathcal{J}^{+}))\,. \tag{5.26}$$

As we are describing the region of spacetime, the definition is geometrical. Moreover, it is the causal structure, as determined by light rays, which dictates what can escape to infinity. Since the event horizon is a causal boundary, it is a *null* hypersurface. Light rays generating the horizon are neither captured by the black hole, nor they can escape to future null infinity. This can be nicely represented using Penrose diagram, with example in Fig. 5.1. It is also worth to notice that the event horizon possesses no local properties that would distinguish it from the rest of the spacetime. In other words, the provided definition of event horizon is not “operational", since an observer would not be able to conclude, based on local measurement, that he is passing the horizon [98][46].

[38] More precisely, *weakly asymptotically simple* spacetime [100].

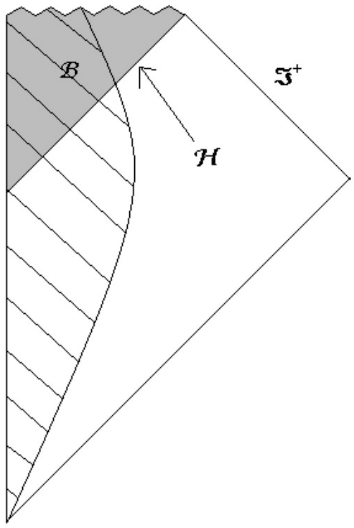


Figure 5.1: Penrose diagram of spherically symmetric stellar collapse in asymptotically flat spacetime. Each point is 2-dimensional sphere. Light rays propagate along $45°$ diagonals. The star region is hatched. The black hole region $\mathcal{B}$ is in grey. The future null infinity is denoted by $\mathcal{J}^+$. The event horizon $\mathcal{H}$ is the boundary of $\mathcal{B}$. The singularity is represented by a wavy line [102].

Black holes are unavoidable in some "normal" circumstances. They occur as a result of stellar collapse, such as the one depicted in Fig. 5.1. A typical star during its life burns nuclear fuel, supporting itself against gravity by thermal and radiation pressure. When the fuel is spent the star starts imploding. It can be shown that if the mass of the star M is larger then some limit mass $M_L \approx 1.5 M_S$, where M_S is mass of the Sun, nor electron nor neutron degeneracy pressure can stop the collapse. At this stage the star must either eject sufficient matter so that its mass is reduced to less than M_L or it reaches a critical point where the event horizon is formed. At this stage, nothing can prevent it from collapsing further and ultimately a singularity is formed [103][71].

Singularity is tight to some kind of pathological behaviour and a black hole will generically contain one. As proven by Penrose and Hawking in singularity theorems, under suitable causality assumption, non-negative energy density condition and existence of trapped surfaces[39] implying strong gravitational field, spacetime will admit a singularity in a sense of geodesic incompleteness. Loosely speaking, this means that particle following the worldline "runs out of world" in finite amount of time. It signalizes the existence of infinite curvature. Moreover, in the mentioned circumstances

[39] Outgoing light rays originating after the formation of event horizon converge and the area of the wavefront decreases, reaching zero when the rays reach singularity. Such wavefront is called a trapped surface, because it is captured by gravitational field [100].

the singularities are either behind an event horizon or naked, causally "visible". The latter case is problematic because such singularities might imply a breakdown of predictability [103][104]. Consequently, it is expected that in reality no naked singularities form. This is yet to be proven and it is known as *cosmic censorship conjecture*. On the other hand, singularity hidden beyond an event horizon poses no such problems, since black hole region is causally disconnected from the exterior. On that account, singularities are usually viewed as a breakdown of classical theory. In other words, it is thought that the singular behaviour does not represent true behaviour in physical world, but the end of classical description. Note that at some point of the collapse the spacetime is bend on the scale of Planck's length requiring quantum theory of gravitation for proper treatment. So, taking into consideration the quantum effects is supposed to lead to regular behaviour of spacetime[40], since for one, energy-positivity condition, assumed by the singularity theorems can be violated locally in quantum field theory [99][106][107].

Further description of black holes will consider a *stationary black hole*. That is, the collapse of the star during which the black hole forms is a dynamical process. However, one expects that the geometry and gravitational field eventually settle down to a so-called stationary state, which we consider as an analogue of equilibrium in thermodynamics. In rough terms, a system is in equilibrium when state of the system does not change on the relevant time scale. In geometrical setting in GR, the nature of equilibrium is that a system is independent of time with respect to an observer for which the system is at rest. This is captured by the following definition [71].

If an asymptotically flat spacetime (M, g) *contains a black hole* $\mathcal{B}$*, then* $\mathcal{B}$ *is said to be stationary if there exists a one-parameter isometry group* $\phi_t : M \to M$ *on spacetime* (M, g) *generated by a Killing vector* t^μ *which is timelike at infinity.*

In practice, this means that we can find coordinates such that metric components describing the exterior region of stationary black hole are independent of the coordinate labeled as time. A stronger property is *static black hole* defined as

A black hole is said to be static if it is stationary and if, in addition, t^μ *is hypersurface orthogonal*[41].

[40] Not everyone agree that this is the case [105].

[41] Hypersurface orthogonal vector field satisfies Frobenius' theorem, which states that a vector field u is hypersurface orthogonal if it satisfies the condition $u_{[\mu}\nabla_\nu u_{\rho]} = 0$.

Equivalently, one may demand that metric is invariant under reflections $t \to -t$. In other words, there exists an isometry which changes time orientation. Static black hole is "motionless", while stationary black hole may exhibit a behaviour that does not change in time. An example is rotation, but such that the angular velocity is constant in time. The latter is a more realistic situation, since, as mentioned, black holes are formed by a stellar collapse. As the star usually rotates, due to conservation of angular momentum, the resulting black hole will continue to do so. In addition, there are important theorems concerned with stationary black holes for which we must also introduce *axisymmetric black holes* and Killing horizons.

A black hole is said to be axisymmetric if there exists a one-parameter group of isometries generated by a Killing vector field ϕ^μ which correspond to rotations at infinity, whose orbits are 2π periodic.

A stationary axisymmetric black hole is said to possess the "$t-\phi$ orthogonality property" if the 2-planes spanned by t^μ and the rotational Killing field ϕ^μ are orthogonal to a family of 2-dimensional hypersurfaces. The $t-\phi$ orthogonality property holds for all stationary-axisymmetric black hole solutions to the vacuum Einstein or Einstein-Maxwell equations.

As before, the $t-\phi$ orthogonality manifests as $t-\phi$ reflection isometry of the metric. Note that stationary-axisymmetric spacetime does not necessarily possesses $t-\phi$ orthogonality property.

Besides the event horizons, there is a completely independent concept of Killing horizons. As the name suggests, Killing horizon is also a null hypersurface. A definition is the following.

A null hypersurface $\mathcal{K}$, whose null generators coincide with the orbits of a one-parameter group of isometries (so there is a Killing field ξ^μ normal to $\mathcal{K}$) is called a Killing horizon.

Moreover, we will also mention a special case of Killing horizon, called *bifurcate Killing horizon*, depicted in Fig. 5.2.

A bifurcate Killing horizon is a pair of null hypersurfaces, $\mathcal{K}_A$ and $\mathcal{K}_B$, which intersect on a spacelike 2-surface B called the bifurcation surface, such that $\mathcal{K}_A$ and $\mathcal{K}_B$ are each Killing horizons with respect to the same Killing field ξ^μ. The converse is also true.

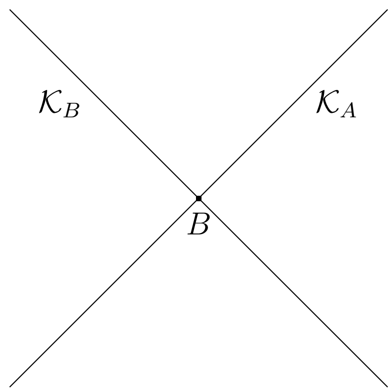


Figure 5.2: The bifurcate Killing horizon consisting of null hypersurfaces $\mathcal{K}_A$ and $\mathcal{K}_B$, intersecting on a spacelike hypersurface B called the bifurcation surface.

It follows from the definition of bifurcate Killing horizon that ξ^μ vanishes at the bifurcation sphere. That is, Killing vector field has unique value at every point. This can be satisfied only if Killing vector field vanishes at the bifurcation sphere.

Although Killing and event horizons are, in general, not related, there are circumstances under which the horizons coincide. They are given by two results, called rigidity theorems. One was proved by Carter and states

For a static black hole, the Killing field t^μ must be normal to the horizon.

For a stationary-axisymmetric black hole with the $t-\phi$ orthogonality property there exists a Killing field χ^μ of the form

$$\chi^\mu = t^\mu + \Omega_H \phi^\mu \ , \tag{5.27}$$

which is normal to the event horizon. The constant Ω_H is called angular velocity of the horizon.

This result does not rely on Einstein field equation. According to the theorem, Killing horizon coincides with the event horizon if the black hole is either static, or stationary axisymmetric with $t-\phi$ orthogonality property. As stated by Carter, Killing horizon will not be an event horizon for any stationary black hole. The second result concerning relationship between event and Killing horizons is called Hawking's strong rigidity theorem, according to which [108]

If the matter fields obey well behaved hyperbolic field equations and the energy-

momentum tensor satisfies weak energy condition[42]*, the event horizon of any stationary black hole must be a Killing horizon.*

According to Hawking's theorem, for stationary black holes there is a Killing vector field that is null at the event horizon. In other words, there exists some Killing vector field generating a Killing horizon which coincides with the event horizon for stationary black holes, under assumptions the matter is well behaved. This further implies that if stationary Killing field t^μ fails to be null at the horizon, there must exist an additional Killing field which is. It was then shown that stationary spacetime (under the assumptions of the theorem), is also axisymmetric (without necessarily the $t - \phi$ orthogonality property). That is, a stationary spacetime possesses two Killing vector fields t^μ and ϕ^μ. Since linear combination of Killing vector fields is also a Killing vector field, it is their combination

$$\chi^\mu = t^\mu + \Omega_H \phi^\mu \,, \tag{5.28}$$

that is null at the event horizon. As before, Ω_H is angular velocity of the horizon. Simply put, if a black hole is stationary it is either static or axially symmetric [101][98].

Moreover, the strong rigidity theorem is important when discussing uniqueness of black hole solutions. In the case of (electro)vacuum, the most general solution, given by Kerr-Newman metric, describes gravitational field exterior of electrically charged, rotating black hole. The presence of electric charge is solely due to charged matter collapsing to the black hole. The solution is unique and depends only on three parameters — mass M, angular momentum J and charge Q. The case where $Q = J = 0$ is described by Schwarzschild solution. This result is surprising because the initial stars are very complex objects, differing in internal structure, shape, pressure, density, etc. After the collapse, none of these details matter, and the resulting black holes may differ only in the above parameters. It is also worth mentioning that charged black holes are only of theoretical interest. Any charged object that is not in vacuum will attract opposite charge and neutralize. The mentioned uniqueness theorems are the premise of so-called *'no-hair' theorems*, according to which all asymptotically flat, stationary

[42] Weak energy condition states that the matter density ρ observed by an observer traveling with velocity u^μ, where u^μ is timelike vector field (by definition of an observer) is always non-negative, $\rho = T_{\mu\nu} u^\mu u^\nu \geq 0$.

black hole spacetimes are characterized only by those parameters[43]. Thereof, the properties of a black hole being independent of the details of the collapsing matter is the first indication that black holes could be a thermodynamics limit of underlying degrees of freedom. As it is known, in thermodynamics, a system has a large number of degrees of freedom, but its state is described by a small number of macroscopic parameters, like energy, entropy, volume, etc., not all necessarily independent [110].

We will now derive the laws describing the behavior of black holes in stationary, asymptotically flat spacetime. As mentioned, asymptotically flat spacetime is analogous to isolated systems, while stationary black holes are analogue of a system in equilibrium. To discuss the zeroth law, we must first explain what is surface gravity.

5.1.3 Surface gravity

As it was shown, generators of Killing horizon — null hypersurface whose normal satisfies the Killing equation — satisfy non-affinely parametrized geodesic equation 5.7. Hence, one interpretation of κ is that it is a measure of extent to which the geodesics fail to be affinely parametrized. Consequently, the value of surface gravity is in principle arbitrary, as one can always scale a Killing vector by a real constant, or equivalently, reparametrize the integral curve $\lambda \to \lambda'$, so that $d\lambda/d\lambda' = const.$ and obtain a different value of surface gravity. The choice is fixed by setting the renormalization, usually by requiring that $\chi^2 \to -1$ at infinity.

Equivalent relation for κ, which will help us to interpret it as surface gravity, is obtained from Frobenius' theorem 5.6 with the help of Killing equation, implying that

$$\chi_\mu \nabla_\nu \chi_\rho + \chi_\nu \nabla_\rho \chi_\mu + \chi_\rho \nabla_\mu \chi_\nu = 0\,. \tag{5.29}$$

Contracting with $\nabla^\nu \chi^\rho$ gives

$$\chi_\mu \nabla^\nu \chi^\rho \nabla_\nu \chi_\rho = -2\kappa^2 \chi_\mu\,, \tag{5.30}$$

[43]There are exceptions, for example when Einstein's equations are coupled to Yang-Mills fields [109][108].

where we used 5.7 and Killing equation. Thus one obtains

$$\kappa^2 = -\frac{1}{2}\nabla^\mu \chi^\nu \nabla_\mu \chi_\nu \,, \tag{5.31}$$

evaluated at the horizon. The surface gravity is the norm of the divergence of Killing vector field at the horizon. Moreover, using the Killing equation and starting from

$$3(\chi^{[\mu}\nabla^\nu \chi^{\rho]})(\chi_{[\mu}\nabla_\nu \chi_{\rho]}) = \chi^\mu \chi_\mu (\nabla^\nu \chi^\rho)(\nabla_\nu \chi_\rho) - 2\chi^\mu (\nabla^\nu \chi^\rho)\chi_\nu (\nabla_\mu \chi_\rho) \,, \tag{5.32}$$

it follows from l'Hospital's rule that

$$\lim \frac{3(\chi^{[\mu}\nabla^\nu \chi^{\rho]})(\chi_{[\mu}\nabla_\nu \chi_{\rho]})}{\chi^\mu \chi_\mu} \to 0 \,, \tag{5.33}$$

in the limit where one approaches horizon. The gradient of the numerator is zero because χ satisfies Frobenius' theorem at the horizon, while the gradient of denominator is different from zero if $\kappa \neq 0$. Hence, using 5.31 we obtain

$$\kappa^2 = \lim \frac{-(\chi_\nu \nabla^\nu \chi^\rho)(\chi_\mu \nabla^\mu \chi^\rho)}{\chi^\mu \chi_\mu} \,. \tag{5.34}$$

Next, one may recognize the acceleration $a^\rho = \chi_\nu \nabla^\nu \chi^\rho / (-\chi^\mu \chi_\mu)$, and define the norm of Killing vector field $V = \sqrt{-\chi^\mu \chi_\mu}$, leading to another expression for surface gravity,

$$\kappa = \lim (Va) \,, \tag{5.35}$$

where $a = \sqrt{a^\rho a_\rho}$. The obtained expression can be interpreted as follows. Let's first consider static, asymptotically flat spacetime. As in such case $\Omega_H = 0$, so $\chi^\mu = t^\mu$. Suppose a particle of unit mass is at rest near horizon, which means that it follows the orbit of t^μ. The orbit of Killing vector that generates the horizon in stationary spacetime are geodesics only on the horizon. Thus, the particle must be accelerating to remain at rest, which requires a force. The proper acceleration of the particle is a, and it diverges if the particle is near the horizon, since the norm of the Killing vector field goes to zero. Then, by introducing the factor V, which goes to zero as one approaches the horizon the product of the quantities in 5.35 is finite. The factor V is referred to “redshift" factor and it converts the change in velocity per unit proper time of the particle to change in velocity per unit coordinate time of observer at infinity. In other words, κ is the force per unit mass as measured at infinity, required to hold a particle at place near horizon. What one has in mind is that the particle is attached to one end of a massless inelastic string, with the other end held by static observer at infinity[44]. Then, the tension the observer measures approaches κ as the particle approaches horizon. In stationary case it is not possible to hold particle at rest near the black hole. More precisely, t^μ is spacelike near event horizon, so a particle at rest would have to move along spacelike trajectories, which is not possible for physical particle. Thus, the same line of thinking fails. One could, however, consider a particle which corotates with the black hole with the angular velocity Ω_H, so its four-velocity is proportional to χ^μ. One can once again calculate the proper acceleration and convert it to acceleration as measured by an observer at by multiplying the acceleration by V. The result tends to κ when the particle is infinitesimally close to the event horizon [45] [71][111].

5.1.4 The zeroth law

We will now state and prove the so-called zeroth law. There are many variations of the proof, with slightly different assumptions. That is, the assumptions needed to prove the claim depend on the rigidity theorems mentioned above. As we will see, if one starts from Hawking's strong rigidity theorem, which does not require the $t - \phi$ orthogonality, one needs to impose a dominant energy condition to prove the statement. On the other hand, starting from Carter's rigidity theorem, the zeroth laws follows as purely geometric claim. The zeroth law states that

The surface gravity of stationary black hole is constant.

It was first proved in [45] for stationary, axisymmetric black holes. The derivation following [45] as adapted by [101] is given in App. D. Here we will follow [71].

Proof of the zeroth law The goal is to show that κ doesn't change in any direction tangent to the horizon. However, note that κ is defined only on the horizon, so we cannot just calculate $\nabla_\mu \kappa$, as divergence of κ may have some components orthogonal to the horizon. The part of $\nabla_\mu \kappa$ that is confined to the horizon is $\chi_{[\mu} \nabla_{\nu]} \kappa$. This is

[44]An observer is called static if his four-velocity is aligned with the Killing vector t^μ.

because $\chi_{[\mu}\nabla_{\nu]}\kappa = \varepsilon^{\mu\nu\rho\tau}\chi_\rho\nabla_\tau\kappa$, and $\varepsilon^{\mu\nu\rho\tau}\chi_\rho\chi_\tau = 0$. In other words, $\varepsilon^{\mu\nu\rho\tau}\chi_\tau$ is tangent to the horizon, working as a kind of projector to the horizon. Thus, κ is constant if we can show that

$$\chi_{[\mu}\nabla_{\nu]}\kappa = 0 \; . \tag{5.36}$$

One starts by multiplying the definition of surface gravity $\chi^\mu\nabla_\mu\chi_\nu = \kappa\chi_\nu$ by $\chi_{[\rho}\nabla_{\tau]}$,

$$\begin{aligned} \chi_\nu\chi_{[\rho}\nabla_{\tau]}\kappa + \kappa\chi_{[\rho}\nabla_{\tau]}\chi_\nu = \chi_{[\rho}\nabla_{\tau]}(\chi^\mu\nabla_\mu\chi_\nu) \\ \chi_\nu\chi_{[\rho}\nabla_{\tau]}\kappa &= -\kappa\chi_{[\rho}\nabla_{\tau]}\chi_\nu + (\chi_{[\rho}\nabla_{\tau]}\chi^\mu)(\nabla_\mu\chi_\nu) + \chi^\mu(\chi_{[\rho}\nabla_{\tau]}\nabla_\mu\chi_\nu) \\ &= -\chi^\mu R_{\mu\nu[\tau}{}^{\lambda}\chi_{\rho]}\chi_\lambda \; . \end{aligned} \tag{5.37}$$

The last line follows from the first two terms cancelling each other our and identity $\nabla_\mu\nabla_\nu\chi_\rho = -R_{\nu\rho\mu}{}^{\tau}\chi_\tau$ which is the result of definition of Riemann tensor and Killing equation. To show the mentioned cancellation, note that from Frobenius' theorem 5.6 and Killing's equation we have

$$\chi_\mu\nabla_\rho\chi_\tau = -2\chi_{[\rho}\nabla_{\tau]}\chi_\mu \; . \tag{5.38}$$

On the right-hand side one may recognize the second term in 5.37. Multiplying the obtained relation with $\nabla^\mu\chi_\nu$ from both sides leads to

$$\begin{aligned} (\chi_{[\rho}\nabla_{\tau]}\chi^\mu)(\nabla_\mu\chi_\nu) &= -\frac{1}{2}(\nabla^\mu\chi_\nu)\chi_\mu\nabla_\rho\chi_\tau \\ &= -\frac{1}{2}\kappa\chi_\nu\nabla_\rho\chi_\tau \\ &= \kappa\chi_{[\rho}\nabla_{\tau]}\chi_\nu \; , \end{aligned} \tag{5.39}$$

in the second line we recognize the definition of κ and the last line follows from 5.38. This is the same as the first term in 5.37. We will now rewrite the last term in 5.37 into a more convenient form. First, we multiply 5.38 by $\chi_{[\nu}\nabla_{\lambda]}$, obtaining

$$(\chi_{[\nu}\nabla_{\lambda]}\chi_\mu)(\nabla_\rho\chi_\tau) + \chi_\mu\chi_{[\nu}(\nabla_{\lambda]}\nabla_\rho\chi_\tau) = -2(\chi_{[\nu}\nabla_{\lambda]}\chi_{[\rho})\nabla_{\tau]}\chi_\mu - 2(\chi_{[\nu}\nabla_{\lambda]}\nabla_{[\tau}\chi_{|\mu|})\chi_{\rho]} \; . \tag{5.40}$$

By repeated use of 5.38 one can show that the first term on the left-hand side and the first term on the right-hand side cancel each other out. Hence, we are left with

$$\begin{aligned}\chi_\mu \chi_{[\nu}\nabla_{\lambda]}\nabla_\rho \chi_\tau &= -2(\chi_{[\nu}\nabla_{\lambda]}\nabla_{[\tau}\chi_{|\mu|})\chi_{\rho]} \\ -\chi_\mu R_{\rho\tau[\lambda}{}^{\sigma}\chi_{\nu]}\chi_\sigma &= 2\chi_{[\rho}R_{\tau]\mu[\lambda}{}^{\sigma}\chi_{\nu]}\chi_\sigma \ ,\end{aligned} \tag{5.41}$$

where we once again made use of the definition of Riemann tensor. Next, multiplying the expression with $g^{\mu\lambda}$, contracting the indices μ and λ leads to

$$\begin{aligned}-\chi^\sigma \chi^\mu R_{\rho\tau[\mu\sigma}\chi_{\nu]} &= 2\chi_{[\rho}R_{\tau]}{}^{\lambda}{}_{[\lambda|\sigma|}\chi_{\nu]}\chi^\sigma \\ 0 &= \chi_{[\rho}R_{\tau]}^{\sigma}\chi_\sigma\chi_\nu - \chi_{[\rho}R_{\tau]\lambda\nu\sigma}\chi^\lambda\chi^\sigma \\ \chi^\sigma R_{\sigma\nu[\tau|\lambda|}\chi_{\rho]}\chi^\lambda &= -\chi_{[\rho}R_{\tau]}^{\sigma}\chi_\sigma\chi_\nu\end{aligned} \tag{5.42}$$

The left-hand side vanishes due to symmetries of Riemann tensor. The second term on the right-hand side has the same form as the last term in 5.37. Hence,

$$\chi_{[\rho}\nabla_{\tau]}\kappa = -\chi_{[\rho}R_{\tau]\sigma}\chi^\sigma \ . \tag{5.43}$$

To show that the right-hand side vanishes we will use Einstein's equation and dominant energy condition, according to which the current $j^\mu = T^\mu_\nu \chi^\nu$ can only be future null or timelike directed. On the other hand, for stationary spacetime, Raychaudhuri's equation gives $R_{\mu\nu}\chi^\mu\chi^\nu = 0$. Using Einstein's equation, this implies that $T^\mu_\nu \chi^\nu \chi_\mu = 0$. In other words, the current can only point in the direction of χ^μ, which we write as $\chi_{[\rho}T_{\mu]\nu}\chi^\nu = 0$. Using Einstein's equation again we get the result that the right hand side has to vanish.

$$\chi_{[\mu}\nabla_{\nu]}\kappa = 0 \ . \tag{5.44}$$

We have shown that the surface gravity of stationary black hole is constant on the horizon if the dominant energy condition is satisfied. It is expected that κ does not change along the geodesic, since it is an integral curve of Killing vector field, along which metric doesn't change. The surprising part is that κ also doesn't change as we move from one geodesic to another, since in general, the horizon is deformed due to rotation for example. Also note that the surface gravity exists only for stationary black holes, as only then the Killing horizon coincides with the event horizon. In other words, κ is well defined only for black holes in equilibrium. The zeroth law of black hole mechanics is analogous temperature being constant in thermodynamics for a system in equilibrium.

Other versions of the proof of the zeroth law Second version of the proof, which we will only sketch, does not require energy condition starts from stationary black hole with $t-\phi$ orthogonality property. As proved in [112], necessary and sufficient condition for the constancy of surface gravity on Killing horizon is that the exterior derivative of the twist of the stationary Killing field t^μ vanishes at the horizon.

$$\nabla_{[\mu}\omega_{\nu]} = 0 \ , \quad \omega_\mu = \varepsilon_{\mu\nu\rho\tau} t^\nu \nabla^\rho t^\tau \ . \tag{5.45}$$

One can show that, for $t^\mu \nabla_\mu t_\nu = \kappa t_\nu$,

$$t_{[\mu}\nabla_{\nu]}\kappa = -\frac{1}{4}\varepsilon_{\mu\nu\rho\tau}\nabla^{[\rho}\omega^{\tau]} \ . \tag{5.46}$$

If t^μ is hypersurface orthogonal, by definition $\omega_\mu = 0$, so κ is constant. Consequently, κ is constant for static black hole, as Killing horizon generated by t^μ coincides with the event horizon. Moreover, in the stationary case there exists a Killing field ϕ^μ, linearly independent of t^μ that commutes with t^μ. Using these properties it is shown that

$$\phi^\mu \nabla_\mu \kappa = 0 \ , \quad \varepsilon^{\mu\nu\rho\sigma}\phi_\nu t_\rho \nabla_\sigma \kappa = \phi_\nu \nabla^{[\mu}\omega^{\nu]} = \frac{1}{2}\nabla^\mu(\phi_\nu \omega^\nu) \ . \tag{5.47}$$

The $t-\phi$ orthogonality property requires that $\phi^\mu\omega_\mu = 0$, which means that both $\phi^\mu\nabla_\mu\kappa = 0$ and $\varepsilon^{\mu\nu\rho\sigma}\phi_\nu t_\rho \nabla_\sigma\kappa = 0$, implying that $t_{[\mu}\nabla_{\nu]}\kappa = 0$. In other words, surface gravity is constant for stationary black holes satisfying $t-\phi$ orthogonality property.

The third version considers bifurcate Killing horizon. One proves, starting from 5.31, that divergence of surface gravity is zero as we move along the generators, and constant as one moves from generator to generator. The derivation is similar to the one in App.D. Furthermore, because we are dealing with bifurcate Killing horizon, we know that the Killing vector field vanishes at the bifurcation 2-sphere. As a result, the divergence of surface gravity must be zero as one moves from generator to generator, which means that κ is constant on the entire horizon. Conversely, one can also show [112] if κ is constant and non-zero over a Killing horizon, then the horizon can be extended so that it is one of the null hypersurfaces of bifurcate Killing horizon.

In summary, we have the following theorems:

- Surface gravity κ is constant on Killing horizon of stationary black hole if Einstein's equation holds with matter satisfying the dominant energy condition.

- The surface gravity κ is constant if the spacetime is either static or stationary with $t - \phi$ orthogonality property, as a consequence of the vanishing of the divergence of the twist.

- The surface gravity κ of stationary black hole is constant on the bifurcate Killing horizon. Conversely, if it can be shown that κ is constant and non-zero on Killing horizon, then the horizon can be extended to bifurcate Killing horizon.

When $\kappa = 0$ the horizon is said to be degenerate, and such black holes are called extreme. They lie in the boundary between black holes and naked singularities and are important in study of cosmic censorship conjecture. Another important question concerning extreme black holes is their stability [113][114]. However, discussion regarding extreme black holes is out of scope of this paper so we will not discuss the issues further.

Apart from the degenerate case, it follows from the zeroth law that the bifurcate horizons are the only types of Killing horizons in GR.

5.1.5 The first law

The first law of thermodynamics is an identity relating the change in mass M, angular momentum J and horizon area A of stationary black hole when it is perturbed. It is often said that the first law is just a statement of energy conservation, and we will see why this is so, but its real importance lies in the fact that it is the Clausius relation.

It has been proven for nearby stationary solutions, where one considers variations in phase space, and for the case when matter fields fall across the horizon. The former version of the proof is called *equilibrium state* version, while the latter is referred to as *physical process* version. Moreover, the first law can most easily be derived using the exact solution in vacuum. However, the significance of the first law is that it holds in very general setting.

The first order variation of stationary state of black hole of mass M, angular momentum J and charge Q is

$$\delta M = \frac{\kappa}{8\pi}\delta A + \Omega_H \delta J + \Phi_H \delta Q \ , \tag{5.48}$$

where Ω_H is angular velocity of the horizon, Φ_H is the electric potential on the horizon, κ is the surface gravity and A is area of 2-dimensional cross section of the horizon.

To simplify the derivation of the first law we will not consider charged black holes.

Equilibrium process version The first law was first derived in [45], and the proof is given in App. E, since the derivation is lengthy. One considers two infinitesimally different stationary black hole solutions, with matter, regarded as ideal fluid for simplicity, in circular orbit outside of the black hole. Besides the differential formula, one also obtains the so-called integral mass formula,

$$M = \int_S \left(2T^\mu_\nu - T\delta^\mu_\nu\right) t^\nu d\Sigma_\mu + 2\Omega_H J_H + \frac{\kappa}{4\pi}A \, . \tag{5.49}$$

The first term in 5.49 is the contribution of matter outside the black hole. In vacuum 5.49 reduces to so-called Smarr formula,

$$M = \frac{\kappa}{4\pi}A + 2\Omega_H J_H \, . \tag{5.50}$$

The expression is analogous to Euler equation in thermodynamics (also called Gibbs-Duhem equation),

$$U(S, V, N) = TS - pV + \mu N \, , \tag{5.51}$$

where U is internal energy of the system, S is entropy, T is temperature, p is pressure, V is volume, μ is chemical potential and N is number of particles. The Euler equation follows from extensivity of the thermodynamic variables, and is different for system described by a different set variables. The first term in 5.50 represents total rotational energy of the black hole, while the last term is analogous to TS term of the Euler equation. Moreover, since κ is constant on the horizon for stationary black hole, just as temperature in thermodynamics, one can deduce that entropy S corresponds to area of the horizon A, $S \propto A$. The differential formula obtained by varying the integral mass formula (in vacuum) is

$$\delta M = \frac{\kappa}{8\pi}\delta A + \Omega_H \delta J \ . \tag{5.52}$$

The analogous statement in thermodynamics is

$$\delta U = T\delta S - pdV \ . \tag{5.53}$$

The second term is the work term and it depends on the kind of system we are considering. Usually, in classical thermodynamics one considers work done by gas which is expanding. One could for example consider work required to change angular momentum of the gas, so the work term would be more similar to the one in 5.52.

Another version of equilibrium process derivation of the first law uses Noether's procedure. It can be shown that 5.52 follows directly from Lagrangian or Hamiltonian formulation of general relativity [115][116]. In the Hamiltonian formulation, 5.52 follows from considering variations of the Hamiltonian associated with the Killing vector. It turns out that variation is a boundary integral whose value at infinity can be related to the mass and angular momentum of black hole so that 5.52 holds.

Even more so, it is shown in [116] that 5.52 is valid if one starts from Lagrangian of any diffeomorphism invariant theory of gravity with arbitrary matter fields. The details of the derivation are out of the scope of this paper, so we show the important results, following [115] and [116]. One starts with Lagrangian in its general form

$$L = L(g_{ab}, R_{abcd}, \nabla_e R_{abcd}, ...; \psi, \nabla_a \psi) \ , \tag{5.54}$$

in n-dimensional spacetime, where ψ represents all matter fields. AN arbitrary but finite number of derivatives of R_{abcd} and ψ are allowed to appear. Moreover, as the theory is diffeomorphism invariant, vector fields on spacetime $\mathcal{M}$ constitute a collection of infinitesimal local symmetries. Let ξ^a be any vector field on $\mathcal{M}$. Then, to each ξ^a one can associate Noether current $(n-1)$-form $\boldsymbol{j}$. Furthermore, when equations of motion hold we have a relationship $\boldsymbol{j} = d\boldsymbol{Q}$, where an $(n-2)$-form $\boldsymbol{Q}$ is referred to as Noether charge. The first law follows from variation of the Noether current. For Killing vector χ it follows,

$$\delta \int_\Sigma \boldsymbol{Q}[\chi] = \delta M - \Omega_H \delta J \ , \tag{5.55}$$

where Σ is bifurcation $(n-2)$-surface at which χ vanishes. Further analysis also shows that

$$\delta \int_\Sigma \boldsymbol{Q}[\chi] = \frac{\kappa}{2\pi} \delta S \ , \tag{5.56}$$

with S, some geometrical quantity, given by

$$S \equiv -2\pi \int_\Sigma E^{abcd} n_{ab} n_{cd} \ , \tag{5.57}$$

where n_{ab} denotes binormal to Σ and E^{abcd} is functional derivative of Lagrangian L with respect to R_{abcd}

$$E^{abcd} \equiv \frac{\delta L}{\delta R_{abcd}} \ , \tag{5.58}$$

such that g_{ab} and ∇_a are held fixed. For the case where the Lagrangian describes GR in vacuum $L = \frac{1}{16\pi} R$, one obtains that

$$E^{abcd} = \frac{1}{16\pi} g^{ac} g^{bd} \ , \tag{5.59}$$

and S is

$$S = -2\pi \int_\Sigma \frac{1}{16\pi} g^{ac} g^{bd} n_{ab} n_{cd} = -\frac{1}{8} \int_\Sigma n^{ab} n_{ab} = \frac{1}{4} A \ . \tag{5.60}$$

Hence, the first law 5.52 holds if we can replace $A/4$ by S. Note that this derivation also fixes the proportionality constant, so $S \leftrightarrow A/4$ and $T \leftrightarrow \kappa/2\pi$. This concludes the discussion about equilibrium processes and we will now turn to the physical process version of the derivation.

Physical process version This version of the first law describes the change in black hole parameters when infinitesimal amount of matter crosses the event horizon so it is closer in spirit to the first law of thermodynamics. We assume that the process if quasi-static and that stationary black hole settles down to a stationary state after the process is over. The original derivation was done in [117], where they considered slowing down of a black hole caused by non-axisymmetric matter and non-axisymmetric gravitational perturbation produced by a distant mass.

When the perturbation is produced by a matter field, the change of parameters can be calculated using formulas for flux of energy and angular momentum across the event horizon, given by 5.61 and 5.62.

$$\delta M = -\int_{\mathcal{H}} \Delta T^{\mu}_{\nu} t^{\nu} d\Sigma_{\mu} \,, \tag{5.61}$$

$$\delta J = \int_{\mathcal{H}} \Delta T^{\mu}_{\nu} \phi^{\nu} d\Sigma_{\mu} \,. \tag{5.62}$$

Small amount of matter is represented by ΔT^{μ}_{ν}. The first expression is energy flux through the horizon as seen by an observer at infinity, and we have a similar expression for the flux of angular momentum[45]. Combining the relations, one obtains

$$\begin{aligned} \delta M - \Omega_H \delta J &= \int_{\mathcal{H}} T_{\mu\nu}(t^{\nu} + \Omega_H \phi^{\nu}) \chi^{\mu} dS d\lambda \\ &= \int d\lambda \int_{\mathcal{H}(\lambda)} T_{\mu\nu} \chi^{\mu} \chi^{\nu} dA \,. \end{aligned} \tag{5.63}$$

The infinitesimal surface element of the horizon is $d\Sigma_{\mu} = \chi^{\mu} dS d\lambda$, where dS is an element of the area of cross section, and $d\lambda$ moves along the generators. The integration is done so that one integrates over the cross section first, which is what is understood by $\mathcal{H}(\lambda)$, and then along the generators. Note that we are dealing with perturbative approach to the first order, so we treat the initial black hole spacetime as a background. In other words, the Killing vectors in 5.61 and 5.62 belong to initial state of black hole (to the first order). Moreover, we assume that sources of the matter field are far away from the black hole and have low enough mass that their gravitational effect on the black hole is negligible compared to the effect of matter fields.

The integral 5.63 is calculated using Raychaudhuri's equation 5.23. When deriving the zeroth law, we were looking at the equilibrium case, so expansion, shear and twist all vanished. As we now have a flow of matter across the horizon, the geodesics

[45]The expression 5.61 and 5.62 are “generalizations" of formula for flux as a current, i.e., a vector field, that is crossing a surface.

deviate from one another as described by 5.64 to the first order. That is, the black hole is not stationary while the process takes place.

$$\frac{d\theta}{d\lambda} = \kappa\theta - 8\pi T_{\mu\nu}\chi^\mu\chi^\nu \, , \tag{5.64}$$

where we neglected the quadratic terms in θ and σ and used Einstein's equations. With 5.64, the expression 5.63 is

$$\begin{aligned} \delta M - \Omega_H \delta J &= -\frac{1}{8\pi}\int d\lambda \int_{\mathcal{H}(\lambda)} \left(\frac{d\theta}{d\lambda} - \kappa\theta\right) dA \\ &= -\frac{1}{8\pi}\int d\lambda \int_{\mathcal{H}(\lambda)} \frac{d\theta}{d\lambda} dA + \frac{\kappa}{8\pi}\int d\lambda \int_{\mathcal{H}(\lambda)} \theta dA \, . \end{aligned} \tag{5.65}$$

In the second line, κ can be taken in front of the integral because to the first order κ is surface gravity of the initial state of black hole, and is constant across the horizon as shown by the zeroth law. Let's look at the first integral,

$$\int d\lambda \int_{\mathcal{H}(\lambda)} \frac{d\theta}{d\lambda} dA = \left(\int_{\mathcal{H}(\lambda)} \theta dA\right)\Big|_{\lambda_1}^{\lambda_2} = 0 \, . \tag{5.66}$$

The boundaries refer to initial and final states of black hole, which are taken to be stationary. Thus, $\theta = 0$ at the lower and upper boundary, so the integral vanishes. To evaluate the first term in 5.65 remember that expansion describes fractional rate of expansion of cross sectional area as we are moving along the generators, i.e.,

$$\theta = \frac{1}{dA}\frac{dA}{d\lambda} \implies \theta dA = \frac{dA}{d\lambda} \, , \tag{5.67}$$

where δA is cross section of a bundle of generators. As a result,

$$\int d\lambda \int_{\mathcal{H}(\lambda)} \frac{d}{d\lambda} A = A(\lambda_2) - A(\lambda_1) = \delta A \, . \tag{5.68}$$

The change in area is infinitesimal. Finally, 5.65 is

$$\delta M - \Omega_H \delta J = \frac{\kappa}{8\pi}\delta A \implies \delta M = \Omega_H \delta J + \frac{\kappa}{8\pi}\delta A \, . \tag{5.69}$$

Hence, we reproduced the same formula as with equilibrium process version [101].

5.1.6 The second law

The second law, unlike the zeroth and the first law, is not related only to stationary black holes. In other words, one does not need any property of Killing vector field to prove the theorem. It states that

Area of event horizon of a predictable[46] black hole never decrease with time, assuming the dominant energy condition $R_{\mu\nu}k^\mu k^\nu \geq 0$ is satisfied for all null vectors k^μ.

$$\delta A \geq 0 \, , \tag{5.70}$$

where $\delta A = 0$ corresponds to the stationary case.

The assumption of predictability comes down to that cosmic censorship conjecture holds. The second law is often called the area increase theorem since, as the name suggest, it states that the area of the black hole cannot decrease. It is proved in [118][103], by showing that the expansion of generators on the horizon is non-negative, leading to condition that the surface area of black hole can never decrease. The derivation in this section follows [115] and [102]. Before proving the second law we need to mention some properties of the event horizons. The proofs can be found in [103].

1. No two point on the horizon can be connected by a timelike curve.
2. The null geodesic generators of the event horizon may have past end-points in the sense that the continuation of the geodesic further into the past is no longer part of the event horizon.
3. The generators of event horizons have no future end-points, i.e., no generator may leave the horizon.

These properties are depicted in Fig. 5.3. The first property is also called *achronicity property*. As a consequence of second and third property the generators may enter the horizon but not leave it. To prove the area increase theorem, let's start from

[46] An asymptotically flat spacetime is said to contain a predictable black hole if there exists a globally hyperbolic region containing both the exterior to the black hole and event horizon of the black hole. There are several notions of globally hyperbolic spacetime, which are all equivalent. One is the existence of Cauchy surface — an achronal (cannot be linked by a causal curve) spacelike hypersurface, which causal curve intersects only once [71][115].

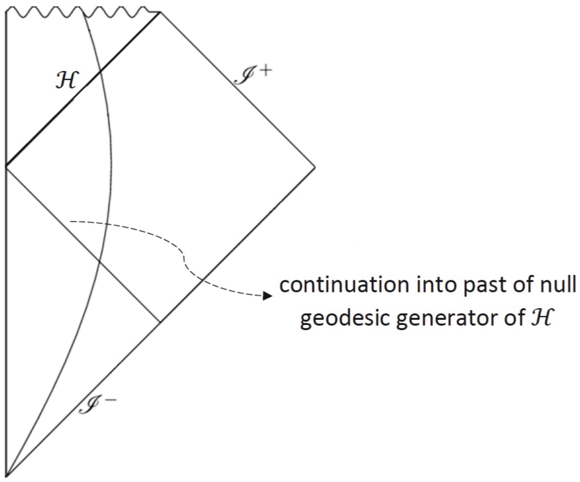


Figure 5.3: Penrose diagram of spherically symmetric collapsing star. The horizon $\mathcal{H}$ is a null hypersurface generated by null geodesics with no future end-point. The generators can be continued to the past null infinity [119].

affinely parametrized Raychaudhuri equation,

$$\frac{d\theta}{d\tilde{\lambda}} = -\frac{1}{2}\theta^2 - \sigma_{\mu\nu}\sigma^{\mu\nu} - R_{\mu\nu}k^\mu k^\nu \ . \tag{5.71}$$

Affine parameter of null geodesic on the horizon is denoted by $\tilde{\lambda}$, and k^μ is its tangent vector, which is proportional to Killing vector χ^μ. However, k^μ doesn't satisfy the Killing equation. Now, assuming the null energy condition holds,

$$R_{\mu\nu}k^\mu k^\nu \geq 0 \ , \tag{5.72}$$

and because shear is spatial, so $\sigma_{\mu\nu}\sigma^{\mu\nu}$ must be positive, it must hold that

$$\begin{gathered} \frac{d\theta}{d\tilde{\lambda}} \leq -\frac{1}{2}\theta^2 \\ \int_{\theta_0}^{\theta} \frac{d\theta}{\theta^2} \leq -\int_0^{\tilde{\lambda}} \frac{d\tilde{\lambda}}{2} \implies \frac{1}{\theta(\tilde{\lambda})} \geq \frac{1}{\theta_0} + \frac{\lambda}{2} \, . \end{gathered} \tag{5.73}$$

The obtained result tells us that if geodesics are converging at some point, i.e., $\theta_0 < 0$, then, as the parameter $\tilde{\lambda}$ increases, the geodesics reach a point where the right-hand side is zero. This corresponds to $\theta \to -\infty$. In other words, the geodesics reach a point of infinite convergence for finite value of parameter, they intersect. One says that the geodesics converge to a focus, or caustic. The physical explanation is that geodesics get focused due to attractive nature of gravitation. The result if known as the focusing theorem. This is the first step towards the area increase theorem [101].

Next, to prove that the area of the horizon can never decrease, we must show that each area element of the horizon a has this property. As expansion is by definition the fractional rate of change of the cross section, we have that

$$\frac{da}{d\tilde{\lambda}} = a\theta \ . \tag{5.74}$$

Thus, the area element is non-decreasing along $\tilde{\lambda}$ if $\theta \geq 0$ everywhere on the horizon. This would follow immediately if the horizon would satisfy geodesic completeness, i.e., if all the generators of the horizon could be extended for all values of the affine parameter. However, the predictability condition imposes no such constraints. Instead, the proof is obtained by contradiction — consider geodesic γ and assume that $\theta < 0$. Then, as it was just proven, geodesics near γ will form a caustic at finite distance away. That is (see Fig. 5.4), nearby geodesic passing through point p must intersect γ a finite distance along it. The first point for which this happens is called

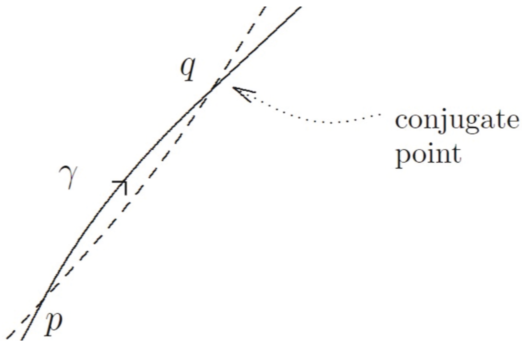


Figure 5.4: Solid line is geodesic γ. Geodesics near γ (dashed line), pass through point p and form a caustic a finite distance away. The point q is conjugate to p. Points on γ beyond q are timelike separated from p [119].

the point conjugate to p on γ. Let's call it point q. It can then be proven[47] that if there exists such point, a variation of γ will give a timelike curve from q to p. That is, points on γ beyond q are timelike separated from p. This is in contradiction with the first property of event horizon. Hence, it must be true that $\theta \geq 0$ everywhere on the horizon, otherwise the horizon would cease to exist [119]. Hence, we have showed that

$$\delta A \geq 0 \ . \tag{5.75}$$

[47]The proof of the claim is Proposition 4.5.12 in [103].

Moreover, as stated in [45], if two black holes coalesce, the area of the final event horizon is greater than the sum of the areas of the initial horizons, i.e.,

$$A_3 > A_1 + A_2 \ . \tag{5.76}$$

The statement of the second law is similar to the entropy only increasing in thermodynamics.

In summary, we have shown that area of the black hole horizon can only grow larger or stay the same if the black hole is stationary, i.e., where no matter crosses the horizon. Simply put, matter crossing the event horizon will cause an increase of the area of event horizon. This result is surprising because it is possible to extract energy from spinning black hole via Penrose process. As already mentioned, in Kerr spacetime outside the event horizon, t^μ is spacelike in region called ergosphere. Consequently, matter in ergosphere corotates with the black hole. Then, if a piece of matter splits into two parts in the ergosphere under the right circumstances, one piece of matter can escape to infinity with greater energy than the original piece of matter, while the other piece falls into black hole with negative energy. The net effect is that energy is extracted from the black hole, with the difference provided by the black hole. The angular momentum of black hole decreases by the amount corresponding to energy transferred to the escaping part of matter. It then seems at the first sight that it may be possible to decrease the area of the black hole. However, this is not the case [46].

5.1.7 The third law

The third law is concerned with the limiting behaviour of systems as the temperature, or equivalently, surface gravity approaches absolute zero. Both in GR and thermodynamics, it does not have the same importance as the other three laws, since (for some formulations of the third law) there are counterexamples. The third law as stated by [45] says,

No physical process can reduce the surface gravity of a black hole to zero by a finite sequence of operations.

The formal proof is given in [120] by Israel, with assumption that matter satisfies the weak energy condition. The third law for black hole mechanics is similar to Nernst's formulation of the third law of thermodynamics, which state that it is not possible to reduce the temperature of system to zero. We should also mention that Nernst's formulation is the most accepted version of the third law.

Another version is given by Simon [121], and states that for system in an equilibrium the entropy goes to zero, as the temperature tends to absolute zero. For black holes, temperature is analogous to surface gravity and entropy to area of the horizon. Thus, we know that in GR this version of the third law does not hold because there exist extreme Kerr black holes for which $\kappa = 0$ and $A \neq 0$ [98]. The importance of this version is its relationship with statistical physics, where entropy is $S = \ln\Omega$, and Ω is the number of microstates corresponding to the same macrostate. A system at temperature zero is expected to be in its ground state, so it then follows that if the ground state is non-degenerate, $\Omega = 1$ and $S = 0$. Because this is not always the case, the law was reformulated. It states that entropy approaches a constant as the absolute temperature tends to zero [122], which is similar to the black hole case.

This concludes the discussion about the four laws ob black hole mechanics. When they were first derived they were considered nothing more then a formal analogy because black hole could only have a temperature of absolute zero. This is because energy can only flow into black hole, but never out. So the "equivalence" would work only for extreme black holes. However, there is an indication that the relationship of the second law of black hole mechanics with the entropy increase should be taken seriously. As noticed first by Wheeler, if one drops matter into black hole, entropy of the visible universe decreases in the process. Then, since the outside of the black hole is causally isolated system, the observer outside the black hole would not be able to verify that the total entropy of the universe increased, surpassing the second law of thermodynamics. On the other hand, as it was just showed, the area of the black hole increased. Then, if one considers the area increase of the black hole horizon as the actual increase in entropy, it is natural to generalize the second law of thermodynamics as — *the sum of entropy in the black hole exterior and the entropy of black hole can never decrease*. In other words, to salvage the second law of thermodynamics one must add the black hole entropy to the entropy content of the universe. This is as far the classical picture takes us. The status of the four laws of black hole mechanics changed when Hawking showed that, taking quantum effect into account, black holes do radiate with a black body spectrum corresponding to some finite temperature. The result is called Hawking radiation.

First we will derive Hawking radiation and then we will discuss the implications it has on the four laws.

5.2 *Hawking radiation*

In this section we consider semi-classical picture, where the matter fields surrounding the black hole are quantized. Consequently, as shown by Hawking in [123], when one takes into account the quantum behaviour of matter outside the black hole horizon, black holes radiate with a spectrum of a black body of a temperature equal to T_H, called Hawking temperature.

$$T_H = \frac{\kappa}{2\pi} \ . \tag{5.77}$$

The Hawking temperature of the black hole at the center of the Milky Way, Sagittarius A*, having a mass of approximately $4 \times 10^6 M_{\odot}$, is approximately 10^{-14} K [46]. Since the temperature of cosmic microwave background is ≈ 2.73 K, the Hawking radiation is substantially overpowered, making the possible experimental verification quite difficult. Nevertheless, Hawking radiation has been derived in many different ways, which although is not the same as experimental verification, shows that theory is consistent and one has good reasons to expect that in reality, black holes radiate.

The original derivation of the result is highly technical and requires one to consider quantum field theory in curved spacetime, since spacetime at the event horizon is curved. We have discussed quantum field theory on flat spacetime when talking about Weinberg-Witten theorem, but generalization to curved spacetime is non-trivial as some of the important concepts are deeply rooted in Poincare symmetries of flat spacetime, such as the existence of unique vacuum state. In fact, the key feature on which Hawking's calculation is based on is that the choice of vacuum in curved spacetime depends on time. Thus, we only briefly explain how Hawking obtained the result. The details are worked out in [124]. Hawking considered a collapse to a Schwarzschild black hole and a free quantum field that is propagating on the background spacetime. Comparing the waves at infinitely late times, when the black hole settled to a stationary state, with the ones at infinitely early times before the collapse has begun, it was shown that the expected number of particles at late infinity corresponds to emission from a perfect black body of temperature T_H.

The heuristic argument explaining Hawking radiation is the following. Vacuum is not empty as virtual particle-antiparticle pairs are continuously created and annihilated, violating the energy conservation for short period of time as allowed by Heisenberg's uncertainty principle. In the case when the pair is created just outside the event horizon, during its short existence one of the particles may cross the horizon, ending up with negative energy since its Killing vector becomes spacelike, reducing the mass of black hole. The other particle, since it remains outside the event horizon, escapes to infinity with positive energy. The net effect is that to an observer outside it appears as the black hole emits particles. This is the Hawking radiation. Moreover, if the pair is created just inside the event horizon, it is classically confined to remain there, once again, because Killing vector field is spacelike beyond the horizon. However, quantum mechanics allows the particles to tunnel out of region that is classically forbidden, so once again we can have a situation where a particle escapes to infinity carrying with it a fraction of mass of the black hole, so that energy is conserved. Derivation based on this picture is given by [125][126], which we now reproduce.

Derivation of Hawking radiation using WKB approximation One stars by considering a virtual pair just inside the horizon. The probability of tunneling, i.e., an emission (tunneling) rate Γ is calculated using WKB approximation, where it is given by

$$\Gamma \propto e^{-2\mathrm{Im}S} , \tag{5.78}$$

and S is action for the trajectory,

$$\mathrm{Im}S = \mathrm{Im}\int_{r_i}^{r_f} p(r)dr . \tag{5.79}$$

Quantity p is the momentum of particle that can be obtained from Hamiltonian. Before proceeding, one should justify the WKB approximation. This is the case when the phase of wave function oscillates fast, i.e., when[48] $\hbar \to 0$. In other words, WKB is valid when de Broglie wavelength λ goes to zero. The wavelength of the emitted radiation as measured by someone at infinity is of the order of the size of black hole, $\lambda \propto M$. However, tracing the geodesic back to the horizon, the wavelenghth is blue-shifted (decreased in wavelength), approaching zero at the horizon. In other words, local observer at the horizon measures infinite blue-shift, so $\lambda \to 0$.

Furthermore, since we are considering WKB approximation, one needs to specify the potential barrier, classically forbidden region of finite size the particles tunnels through. Although not obvious at first where the potential barrier is, the answer comes from conservation of energy. The emission process lowers the mass of black hole by the amount of energy carried by the particle being emitted. Consequently, the horizon radius decreases. Hence, the barrier is taken as the region between the initial and new position of the horizon. It is created by the outgoing particle itself.

We now return to the derivation. To evaluate the integral 5.79 we need a specific trajectory. Let's consider the case of Schwarzschild black hole, and take the particle to be a photon. Now, because the particle escapes from the inside of event horizon r_i to the outside r_f, one needs to use coordinates which are regular on the event horizon. The spherical coordinates are not suitable for this purpose. Instead, we will use Gullstrand-Painlevé coordinates,

$$ds^2 = -\left(1 - \frac{2(M-\omega)}{r}\right) dt^2 + 2\sqrt{\frac{2(M-\omega)}{r}} dt dr + dr^2 + r^2 d\Omega^2 \; . \tag{5.80}$$

Note that M is exchanged by $M - \omega$ in the metric. The result follows from [127], where it is shown that if self-gravitating shell of energy ω moves on the geodesic, while total mass of spacetime is considered fixed, the shell travels on the geodesic given by line element 5.80. That is, with M exchanged by $M - \omega$. Moreover, assuming that the photon moves along radial null geodesic, its trajectory is

$$\frac{dr}{dt} = \pm 1 - \sqrt{\frac{2(M-\omega)}{r}} \; , \tag{5.81}$$

with upper (lower) sign corresponding to outgoing (ingoing) geodesics, assuming t increases towards future. Also, $r_i = 2M$ and $r_f = 2(M - \omega)$. One can see that the

[48]We are working in units where $\hbar = 1$. That is, the expression in the exponent is $S/\hbar$.

width of the barrier depends on energy of the particle. Evaluation of 5.79 is done using the methods of complex integration [125]. We provide the final result,

$$\Gamma \propto e^{-8\pi M\omega(1-\frac{\omega}{2M})} \ . \tag{5.82}$$

This is the emission rate of the Schwarzschild black hole. Neglecting the term $\propto \omega^2$, the expression takes the form $e^{-\beta E}$, with $E = \omega$ which confirms the intuitive picture that Hawking radiation can be viewed as tunneling. Moreover, by explicit calculation one can check that for Schwarzschild black hole the surface gravity is

$$\kappa = \frac{1}{4M} \implies T = \frac{1}{8\pi M} = \frac{\kappa}{2\pi} \ , \tag{5.83}$$

which leads to the result stated at the beginning of the section. The temperature of the radiations is the Hawking temperature T_H.

Lastly, we give a few remarks. The additional term $\propto \omega^2$ is the result of energy conservation in a sense that the negative sign ensures the temperature of black hole increases the more energy it radiates. Secondly, the derivation in this section is given for Schwarzschild black hole in specific coordinates for a special case of trajectory. This is a very particular setting. Thus, it is worth to mention that the approach has been generalized to other types of black holes, with the null geodesic method replaced by Hamilton-Jacobi method [128][129].

As mentioned there are other ways to derive Hawking's radiation. One such method is by trace anomalies, proved in [130]. It consists of calculating the vacuum expectation value of EMT of massless scalar field $\langle T_{\mu\nu}\rangle$ in 2-dimensional spacetime evaluated in a 2-dimensional model of gravitational collapse. The method is called trace anomaly because classically, the trace of EMT has to vanish on-shell, which is no longer true when quantum corrections are taken into account. Then with the help of covariant conservation law it follows that there is a flux of positive energy density at infinity, describing thermal radiation, balanced by fluxed of negative energy through the horizon. The idea was generalized by [131][132].

Another method relies on the fact that in quantum field theory (on flat spacetime) one can obtain a thermal partition function from path integral approach. The basic idea that the partition function $Z_C = \mathrm{Tr} e^{-\beta H}$, where β is inverse of temperature, is formally equivalent to probability amplitude $Z = \langle F|e^{-iHt}|I\rangle$ for initial state $|I\rangle$ to end up in the final state $|F\rangle$ if $|I\rangle = |F\rangle$ and one analytically continues (real) time t to imaginary time with period β, i.e., $t \to -i\beta$. The cyclic property makes sure that every state goes back to itself, i.e., that $|I\rangle = |F\rangle$ holds [41].

In GR, as explained before, there is in general no preferred time coordinate, unless the spacetime is stationary. Then the Killing vector field defines a preferred time coordinate. We will now derive Hawking temperature for Schwarzschild black hole using methods explained above, called Euclidean path integral approach following [41][119][133].

Derivation of Hawking temperature using Euclidean path integral approach To begin with, we briefly explain the idea of Euclidean path integrals within quantum mechanics. As we know, the amplitude to propagate from initial state q_I to the final state q_F is given by $\langle q_F|e^{-iHt}|q_F\rangle$, where t is time and H is the Hamiltonian. The path integral representation of the amplitude is given by

$$\langle q_F|e^{-iHt}|q_I\rangle = \int \mathcal{D}q\, e^{i\int d^4\mathcal{L}} , \tag{5.84}$$

where $\mathcal{D}q$ is path integral measure. This is obtained by dividing time t into N segments, each lasting $\delta t = \frac{t}{N}$, and making use of

$$\int_{-\infty}^{\infty} |q\rangle\langle q| = 1 , \tag{5.85}$$

so that the amplitude can be written as

$$\begin{aligned}
&\langle q_F|e^{-iHt}|q_I\rangle = \\
&\quad \left\langle q_F \middle| e^{-\mathrm{i}H\delta t} \left(\int_{-\infty}^{\infty} |q_{N-1}\rangle \langle q_{N-1}| \,\mathrm{d}q_{N-1} \right) e^{-\mathrm{i}H\delta t} \left(\int_{-\infty}^{\infty} |q_{N-2}\rangle \langle q_{N-2}| \,\mathrm{d}q_{N-2} \right) e^{-\mathrm{i}H\delta t} \times \right. \\
&\qquad \left. \cdots e^{-\mathrm{i}H\delta t} \left(\int_{-\infty}^{\infty} |q_1\rangle \langle q_1| \,\mathrm{dq}_1 \right) e^{-\mathrm{i}H\delta t} \middle| q_I \right\rangle \\
&\quad = \left(\prod_{j=1}^{N-1} \int dq_j \right) \left\langle \psi_F \left| e^{-\mathrm{i}H\delta t} \right| q_{N-1} \right\rangle \left\langle q_{N-1} \left| e^{-\mathrm{i}H\delta t} \right| q_{N-2} \right\rangle \times \\
&\qquad \left\langle q_{N-2} \left| e^{-\mathrm{i}H\delta t} \cdots e^{-\mathrm{i}H\delta t} \right| q_1 \right\rangle \left\langle q_1 \left| e^{-\mathrm{iH}\delta t} \right| q_I \right\rangle .
\end{aligned} \tag{5.86}$$

If one assumes that the particle is free, $H = p^2/2m$. Evaluating the individual integral $\langle q_{j+1}|e^{-i\delta t\cdot p^2/2m}|q_j\rangle$ gives

$$\left\langle q_{j+1}\left|e^{-\mathrm{i}H\delta t}\left(\int_{-\infty}^{\infty}\frac{\mathrm{dp}}{2\pi}|p\rangle\langle p|\right)\right|q_j\right\rangle = \int_{-\infty}^{\infty}\frac{\mathrm{dp}}{2\pi}e^{-\mathrm{i}p^2\delta t/2m}\left\langle q_{j+1}\mid p\right\rangle\left\langle p\mid q_j\right\rangle \tag{5.87}$$

$$= \int_{-\infty}^{\infty}\frac{\mathrm{dp}}{2\pi}e^{-\mathrm{i}p^2\delta t/2m}e^{\mathrm{i}p(q_{j+1}-q_j)} \tag{5.88}$$

$$= \left(\frac{-i2\pi m}{\delta t}\right)^{\frac{1}{2}} e^{i\delta t(m/2)[(q_{j+1}-q_j)/\delta t]^2}\ . \tag{5.89}$$

Consequently, the amplitude is

$$\langle q_F|e^{-iHt}|q_I\rangle = \left(\frac{-im}{2\pi\delta t}\right)^{\frac{N}{2}}\left(\sum_{k=1}^{N-1}\int dq_k\right)e^{i(m/2)\delta t\sum_{j=0}^{N-1}[(q_{j+1}-q_j)/\delta t]^2} \tag{5.90}$$

In the limit $N \to \infty$ and $\delta t \to 0$ one can replace $(q_{j+1}-q_j)/\delta t$ with $\dot{q}$ and $\delta t\sum_{j=0}^{N-1}$ with $\int_0^t dt$. Defining integral over paths as

$$\int \mathcal{D}q = \lim_{N\to\infty}\left(\frac{-im}{2\pi\delta t}\right)^{\frac{N}{2}}\left(\sum_{k=1}^{N-1}\int dq_k\right)\ , \tag{5.91}$$

one obtains 5.84. Finally, to evaluate the integral one performs Wick rotation by substituting $t \to -it$ and rotating the integration contour in the complex plane. Moreover, in most cases one is interested in the ground state which is denoted by $|0\rangle$, and the amplitude is called $Z \equiv \langle 0|e^{-iHt}|0\rangle$.

$$Z = \int \mathcal{D}q\, e^{-\int d^4x\mathcal{L}}\ . \tag{5.92}$$

In QFT, the dynamical variable is field ϕ, so

$$Z = \int \mathcal{D}\phi\, e^{-S_E[\phi]}\ , \tag{5.93}$$

where S_E is Euclidean action. The Wick rotation changed Lorentzian signature to Riemannian signature, which in flat spacetime gives the metric of Euclidean plane.

On the other hand, partition function of canonical ensemble, representing possible states of a system in thermal equilibrium with the surrounding at fixed temperature of a quantum and discrete system is given by

$$Z_C = \sum_n \langle n|e^{-\beta H}|n\rangle = \sum_n e^{-\beta E_n} = \mathrm{Tr} e^{-\beta H} \ , \tag{5.94}$$

where E_n is energy of state $|n\rangle$. We will now show how path integral formulation and partition function are related. Remember that the amplitude is given by Green function,

$$G\left(q', t; q, 0\right) = \left\langle q' \left| e^{-iHt} \right| q\right\rangle \ . \tag{5.95}$$

If we take t to be pure imaginary, we can write $t = -i\beta$, where β is real. Then

$$\begin{aligned} G\left(q', -i\beta; q, 0\right) &= \left\langle q' \left| e^{-iH(-i\beta)} \right| q\right\rangle \\ &= \left\langle q' \left| e^{-\beta H} \sum_j \right| j \right\rangle \langle j||q\rangle \\ &= \sum_j e^{-\beta E_j} \left\langle q' \mid j\right\rangle \langle j \mid q\rangle \\ &= \sum_j e^{-\beta E_j} \langle j \mid q\rangle \left\langle q' \mid j\right\rangle \ . \end{aligned} \tag{5.96}$$

Considering $q' = q$ and integrating over q we get

$$\int dq G(q, -i\beta; q, 0) = \sum_j e^{-\beta E_j} \left\langle j| \underbrace{\int dq|q\rangle\langle q}_{=1} |j \right\rangle = Z_C \ . \tag{5.97}$$

This shows that to relate path integral formulation to partition function we should make time imaginary, $t \to -i\beta$ and consider the state that goes back to itself. Moreover, in QFT for finite temperature the Green function is periodic. To show this, consider the Green function of a scalar field ϕ in thermal ensemble at temperature

$T = 1/\beta$.

$$\begin{aligned} G_\beta\left(q', t; q, 0\right) &\equiv \mathrm{Tr}\left(e^{-\beta H}\varphi(q', t)\varphi\left(q, 0\right)\right) \\ &= \mathrm{Tr}\left(e^{-\beta H}\varphi(q', t)e^{-\beta H}e^{\beta H}\varphi\left(q, 0\right)\right) \\ &= \mathrm{Tr}\left(\varphi(q', t)e^{-\beta H}e^{\beta H}\varphi\left(q, 0\right)e^{-\beta H}\right) \\ &= \mathrm{Tr}\left(\varphi(q', t)e^{-\beta H}\varphi\left(q, t + i\beta\right)\right) \\ &= G_\beta\left(q, 0; q', t + i\beta\right) \ . \end{aligned} \tag{5.98}$$

We used the cyclic property of trace and $e^{\beta H}\varphi\left(q, 0\right)e^{-\beta H} = \varphi\left(q, t + i\beta\right)$ because Hamiltonian generates time translations. It then follows that a Green function symmetric in its arguments is periodic in imaginary time. In other words, field living in spacetime with imaginary and cyclic time is living in a temperature bath proportional to inverse of the imaginary time period[49].

Using this observation we will determine the temperature of Hawking radiation. In spherical coordinates the metric os Schwarzschild black hole is

$$ds^2 = -f\,dt^2 + \frac{1}{f}dr^2 + r^2 d\Omega^2 \ , \quad f = 1 - \frac{2M}{r} \ . \tag{5.99}$$

We are interested in what is happening near the horizon. Hence, let's make the substitution $r - 2M = x^2/8M$ and look at the case when $x \to 0$. Also taking into consideration that surface gravity $\kappa = 1/4M$ for Schwarzschild black hole, we have that

$$f = \frac{(\kappa x)^2}{1 + (\kappa x)^2} \approx (\kappa x)^2 \ , \quad dr^2 = (\kappa x)^2 dx^2 \ . \tag{5.100}$$

The approximation is valid near the horizon where $x \approx 0$. The metric near the horizon is

$$ds^2 \approx -(\kappa x)^2 dt^2 + dx^2 + \frac{1}{4\kappa^2}d\Omega^2 \ . \tag{5.101}$$

We will be interested only in the first two terms. If we now set $t = -it_E$, we get

$$ds^2_{2E} \approx x^2 d(\kappa dt_E)^2 + dx^2 \ , \tag{5.102}$$

[49]The explanation is based on the mathematical procedure. Whether or not this connection is only formal or it truly has a physical interpretation is an open question [41].

which we can recognize as the metric of a plane in polar coordinates,

$$ds^2 = dr^2 + r^2 d\theta \ . \tag{5.103}$$

We can conclude that the angular coordinate is κt_E. Consequently, κt_E has the period equal to 2π,

$$t_E \propto t_E + \frac{2\pi}{\kappa} \ . \tag{5.104}$$

The plane $x - t_E$ is flat only if κt_E is periodic, as otherwise the point $x = 0$ is a conical singularity, assuming that the vertex of the cone is at the origin [134].

On the other hand, as previously explained, the period of complex time is equal to β. We can then make the following identification,

$$\beta = \frac{2\pi}{\kappa} \ , \quad T_H = \frac{\kappa \hbar}{2\pi k_B} \ , \tag{5.105}$$

where we reintroduced all the Boltzmann and Planck's constants, k_B and $\hbar$. We see that classically, the Hawking temperature vanishes. Moreover, from the first law we can now determine the proportionality constant between area of the black hole and entropy,

$$\delta M = \frac{\kappa}{8\pi} \delta A \implies A = \frac{1}{4} S_{BH} \ , \tag{5.106}$$

from analogy with thermodynamics, $\delta U = T \delta S$. The black hole entropy is called Bekenstein-Hawking entropy,

$$S_{BH} = \frac{A}{4\hbar G} \ , \tag{5.107}$$

and its nature is quantum gravitational in a sense that it contains Newton's constant G and Planck's constant $\hbar$.

In summary, black hole radiate as black bodies with temperature T_H. As a consequence of this result and Hawking radiation, the entropy of black hole is S_{BH}. We will now discuss the impact this result has on the second law.

5.3 Generalized second law

As already explained, from completely classical consideration it seems that the second law should be generalized, in a sense that one should assign entropy to the black hole.

Moreover, a problem with the second law of black hole mechanics arises taking Hawking radiation into account. Quantum matter does not satisfy null energy condition, assumed in the derivation of the second law. Consequently, as the black hole radiated its area decreases, it “evaporates". Hence, the generalized statement of the second law was proposed, stating that

$$\delta(S + \frac{1}{4}A) \geq 0 \ , \tag{5.108}$$

where S is entropy of matter outside the black hole. The entropy outside the black hole compensates for decrease in area of the black hole as it radiates [98].

Thus, the second law of thermodynamics fails in the presence of black hole, while the second law of black hole mechanics fails if quantum effects are taken into account. It is the generalised second law that holds[50].

5.4 Commentary on the four laws of black hole mechanics

In this section we consider some objections concerned with analogy between black holes and thermodynamics. Let’s start by stating the four laws of black hole mechanics obtained in the previous sections, next to their thermodynamic analogues [46].

Zeroth Law

[Thermodynamics] The temperature T is constant throughout a body in thermal equilibrium.

[Black Holes] The surface gravity κ is constant over the event horizon of a stationary black hole.

[50] Although we have strong indications that the second law of black hole is valid, some ideas have been proposed violating the second law, which led to proposition of an entropy bound [135]. We will not discuss these issues here, as complete treatment relies on statistical interpretation of entropy, which is out of scope of this paper.

First Law

[Thermodynamics]

$$\mathrm{d}E = T\,\mathrm{d}S + p\,\mathrm{d}V + \Omega \mathrm{d}J$$

where E is the total energy of the system, T the temperature, S the entropy, p the pressure, V the volume, Ω the rotational velocity and J the angular momentum.

[Black Holes]

$$\delta M = \frac{1}{8\pi}\kappa\delta A + \Omega_{\mathrm{II}}\delta J_{\mathrm{II}}$$

where M is the total black hole mass, A the surface area of its horizon, Ω_{II} the rotational velocity of its horizon, J_{II} its total angular momentum, and δ denotes the result of a first-order, linear perturbation of the spacetime.

Second Law

[Thermodynamics] $\delta S \geq 0$ for any process in an isolated system.

[Black Holes] $\delta A \geq 0$ in any process if null energy condition holds.

Third law

[Thermodynamics] No physical process can reduce the surface gravity of a black hole to zero by a finite sequence of operations.

[Black Holes]No physical process can reduce the temperature of the system to zero by a finite sequence of operations.

As already explained, in thermodynamics, a system is some type of matter matter confined in space and separated from the its surrounding by a wall, allowing for transfer of various quantities. In a similar manner, black hole is a region of spacetime that is causally disconnected from the rest, with the event horizon playing the role of the boundary. Equilibrium is extended to geometry as existence of stationary Killing vector field that is timelike near infinity. Moreover, the state of thermodynamical system is specified by a few macroscopic parameters, depending on the nature of the system, just like black holes can be characterized with a small number of parameters. The behaviour of the system in (quasi)equilibrium is determined by thermodynamical laws given above. However, unlike equations of motion which are limited to specific systems and deterministic, prescribing exactly how the system evolves (at least classically), thermodynamics works as a constraint imposed by probability theory. Thus, if the analogy of gravitation and thermodynamics can be considered as equivalence, it would imply that Einstein's field equations can be considered as a constraint, underlined by some more fundamental laws.

Lastly, as we are discussing thermodynamic properties, note that heat capacity of black hole is negative. Consider Hawking temperature of Schwarzschild black hole,

$$T_H = \frac{1}{8\pi M} \,. \tag{5.109}$$

As the black hole radiates, its mass decreases, meaning that the temperature increases. The result is not unique to black holes, since some stars can also have negative heat capacity.

Let's now see what are the differences between black holes and thermodynamics.

The first difference is that the laws of thermodynamics, the zeroth law especially, are not usually stated at this form. The zeroth law of thermodynamics states that if two thermodynamic systems A and B are separately in thermal equilibrium with a third system C, then they are in thermal equilibrium with each other. It defines thermal equilibrium as an equivalence relation between thermodynamic systems [136].

The point is that that the Zeroth law in thermodynamics is not just a statement that temperature is constant in equilibrium, but of transitivity of the equilibrium. The constancy of temperature is in fact a consequence of transitivity.

The 'in equilibrium with' relation for the black holes exists thanks to Hawking radiation [137]. Black holes are in thermal contact just as any other self-gravitating system. Imagine putting two systems, one or both of which is a black hole, in a large box far enough from each other so that their gravitational attraction can be neglected. Each box will fill up with radiation corresponding to some temperature so the heat will flow from the hotter to the colder body.

Another discrepancy is that in thermodynamics, entropy is an extensive quantity that scales with volume. On the other hand for black holes entropy scales with area. There is no arguing with this claim. Although this represents departure from thermodynamics, as stated in [137], there is nothing in thermodynamics that requires extensivity of entropy.

Finally, it should be pointed out that Hawking radiation of black hole does not originate from microscopical degrees of freedom of the black hole itself, as it is generally the case for black body radiation.

This concludes the discussion about relationship between black holes and thermodynamics. Although there are many similarities, it is not clear that the analogy can be considered equivalence. Note that we haven't discussed the origin of the entroy. The discussion of the subject is out of scope of this paper.

We now turn to more general setting motivated by the fact that the event horizons are not constrained only to black holes. In other words, although the four laws we discussed concern black holes, the crucial role was played by the event horizon, but they exist even in spacetimes that do not contain a black hole. Following [65] we will refer to horizons that are not the boundary of the black hole as causal horizons. More precisely, we will consider Rindler horizons. In the next section we will introduce them in the scope of so-called Unruh effect.

5.5 *Unruh effect*

An effect closely related to Hawking radiation, but independent is the Unruh effect. A uniformly accelerated observer in flat spacetime, called Rindler observer perceives a thermal state of temperature T_U, what an inertial observer perceives as vacuum.

$$T_U = \frac{a}{2\pi} \, , \tag{5.110}$$

where a is the magnitude of acceleration of Rindler observer. In other words, Rindler observer perceives thermal bath of particles. For example, if there is a scalar field propagating in flat spacetime, Rindler observer would measure density of particles as given by Bose-Einstein distribution with $T = T_U$, while at the same time, the expected number of particles according to an inertial observer is zero. Moreover, the Unruh effect can be generalized to curved spacetimes with a bifurcate Killing horizon [115].

Before dealing with the subject further, we need to discuss spacetime from the point of view of Rindler observer. For simplicity, we will consider 2-dimensional spacetime. Let the Rindler observer describe spacetime using coordinates (x, t), while inertial observer uses Cartesian coordinates (X, T). Then, the trajectory of uniformly accelerating observer is hyperbola (see Fig. 5.5). The relationship between the coor-

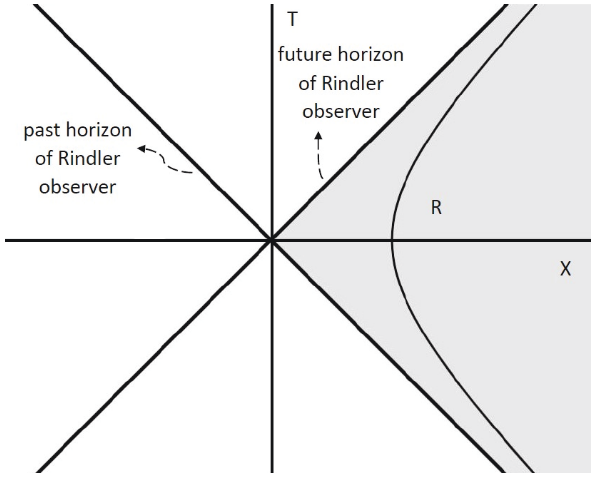


Figure 5.5: Spacetime diagram of flat spacetime. Hyperbola is the trajectory of Rindler observer. The region in grey is part of spacetime accesible to Rindler observer, called right Rindler wedge (marked by R). The horizons are Killing horizons for vector field ∂_t and represent boundaries of past and future as experienced by Rindler observer.

dinates is

$$X = x\cosh(at) \ , \quad T = x\sinh(at) \ , \quad x > |t| \ . \tag{5.111}$$

The range of Rindler's coordinates is $-\infty < t, x < +\infty$ and covers only the right wedge. Moreover, the line element is

$$ds^2 = -dT^2 + dX^2 = -a^2x^2dt^2 + dx^2 \ . \tag{5.112}$$

One may recognize that the metric is the same as the one approximating geometry in the neighbourhood of Schwarzschild spacetime, 5.101. As 5.112 is independent of coordinate t we know that ∂_t is a Killing vector field. We are in flat spacetime so the Killing vector field can be either spacetime translation or rotation (including boosts and spatial rotations). Expressing ∂_t in (T, X) coordinates gives

$$\partial_t = a(X\partial_T + T\partial_X) \ . \tag{5.113}$$

Thus, ∂_t is boost Killing vector field. its orbits are hyperbolas. That is, they correspond to the worldlines of uniformly accelerated observer with proper acceleration $a = 1/x$. It is the boost Killing field that generates the bifurcate Killing horizon on

Fig. 5.5. The horizon is called the Rindler horizon, and represents causal boundary for Rindler observer. We can now give the formal definition of Unruh effect.

Consider a classical spacetime that contains a bifurcate Killing horizon $\mathcal{K}$, so that there is one parameter group of isometries whose associated Killing vector field is normal to $\mathcal{K}$. Consider a free quantum field on this spacetime. Then there exists at most one globally nonsingular state of the field which is invariant under the isometries. Furthermore, in the "wedges" of the spacetime where the isometries have timelike orbits, this state (if it exists) is a thermal equilibrium state at temperature T_U.

Rindler spacetime —- flat spacetime in Rindler coordinates — is the simplest examples satisfying the definition. In flat spacetime, any one-parameter group of Lorentz boosts has an associated bifurcate Killing horizon comprised of two intersecting planes. The unique, globally nonsingular state which is invariant under these isometries is the "usual" vacuum state, the one observed by inertial observer. In the "right" and "left" wedge of flat spacetime defined by Killing horizon, the orbits of Lorentz boost isometries are timelike and correspond to worldlines of uniformly accelerating observers.

Just as Hawking radiation, Unruh effect can be derived in many ways — from Bogoliubov coefficient method, path integral approach, structure of the propagator [138], to name a few. Moreover, the derivation of Unruh effect is mathematically very similar to the derivation of Hawking's effect. This is the reason the Unruh temperature T_U has the same form as T_H.

If one is not interested in the thermal spectrum (we are not), the fastest way to derive the Unruh temperature is to start from explicit form of the Green function in flat spacetime and apply the coordinate transformations to Rindler coordinates. Incorporating the condition that the momenta of states must lie either inside or on the future light cone one obtains a Green function that is periodic with $\beta = 1/T_U$. There is even shorter way, along the same lines. Note that Euclideanized flat spacetime metric is periodic in imaginary Rindler time,

$$\begin{aligned} & ds^2 = -dT^2 + dX^2 \ , X = a\cosh(at) \ , T = x\sinh(at) \xrightarrow{T=iT_E \ , t=it_E} \\ & ds^2 = dT_E^2 + dX^2 \ , X = a\cos(at_E) \ , T = x\sin(at_E) \ . \end{aligned} \tag{5.114}$$

Euclidenized coordinates are periodic in t_E with period $2\pi/a$. Consequently, the

Green functions are related by (remember 5.98)

$$G_E(T_E, X) \equiv G_E(t_E, x) = G_E(t_E + \frac{2\pi}{a}, x) \ . \tag{5.115}$$

After the Wick rotation the Green function for inertial observer can be interpreted as thermal Green function in Rindler time with $\beta = 2\pi/a$, i.e., $T = a/2\pi$. In other words, vacuum of inertial observer looks like thermal state for Rindler observer [139].

5.6 Thermodynamics of spacetime

What we have shown so far is that thermodynamics is tied to a black hole spacetime. It was first shown in [65] that one can obtain Einstein's equations from geometrical and thermodynamic considerations for causal horizons. In what follows, we will show with the help of thermodynamics as perceived by Rindler observer, that the geometry is constrained in a way that it satisfies Einstein's equations.

We start by defining a local Rindler horizon at an arbitrary point p (see Fig. 5.6). This will define one part of the spacetime as our system. First, we choose a small

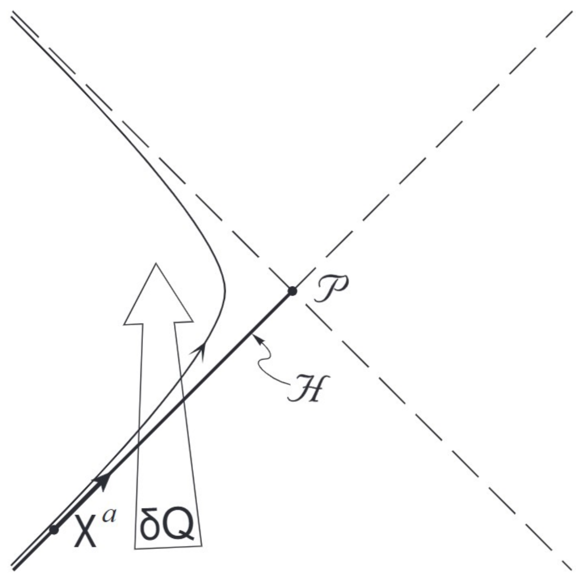


Figure 5.6: $\mathcal{P}$ is 2-dimensional spacelike hypersurface containing p. The dashed line in the lower part is the boundary of the past of $\mathcal{P}$, whose one side (bold line) is chosen as local Rindler horizon $\mathcal{H}$. We think of the left wedge as a system, with Unruh temperature as measured by an observer whose worldline is the hyperbola in the left wedge, asymptotically approaching $\mathcal{H}$. The heat flow across LCH is given by δQ. The vector field χ is the approximate Killing vector field [65].

patch of 2-dimensional spacelike hypersurface that contains p, called $\mathcal{P}$. Because of the equivalence principle, spacetime at p is flat, so one can introduce the normal coordinates x^μ such that $g_{\mu\nu}(x) = \eta_{\mu\nu}(x) + \mathcal{O}(x^2)$, which we transform to Rindler coordinates,

$$ds^2 = -dT^2 + dX^2 + ds_2^2 = -\kappa^2 x^2 dt^2 + dx^2 + ds_2^2 \,, \tag{5.116}$$

by uniformly accelerating along X axis with an acceleration κ. Because we are considering left Rindler wedge (Fig. 5.6), one should flip the sign in 5.111.

Next, note that the *boundary of the past* of $\mathcal{P}$ consists of two components, each generated by a congruence of null geodesics perpendicular to $\mathcal{P}$. We choose one of the components which we will call local Rindler horizon (LRH) and denote it by $\mathcal{H}$.

The next step is to set up a coordinate system using the congruence that generates the LRH. We will consider future directed null vectors k^μ tangent to the congruence that generates the horizon, parametrized by an affine parameter λ, chosen so that it vanishes as p.

Finally, we will consider the region behind LRH, i.e., the left wedge, as system whose temperature is the Unruh temperature as measured by uniformly accelerated observer asymptotically approaching the the horizon.

We are interested in how the area of the cross section of the horizon behaves when infinitesimal amount of energy (i.e., heat, to be in line with thermodynamic terminology) flows across it.

The notion of heat flux is related to energy-density current. As we know, in order to define one we need a Killing vector field. In Rindler spacetime, as shown in the previous section, the Killing vector field χ generates boosts. However, remember that we are not in flat spacetime. Nevertheless, since our consideration is local, spacetime in a small neighbourhood of $\mathcal{P}$ is approximately flat and consequently, there is an approximate Killing vector field generating boosts. Moreover, Killing vector field is uniquely determined by specifying its value and derivative at an arbitrary point. Let χ vanish at $\mathcal{P}$ and generate boosts orthogonal to $\mathcal{P}$ in the small neighbourhood of $\mathcal{P}$. Furthermore, χ is normalized to unity so that the acceleration of the Killing orbit is equal to κ. Ultimately, the heat flow related to the approximate Killing vector field is

$$\delta Q = \int T_{\mu\nu}\chi^{\mu}d\Sigma^{\nu}\,. \tag{5.117}$$

The integral is over a small bundle of generators of $\mathcal{H}$. The heat flow is defined as boost-energy current of the matter $T_{\mu\nu}\chi^{\nu}$, where $T_{\mu\nu}$ is energy-momentum tensor. Because the Killing vector field is spacelike in the region outside the horizon [51], the flow of heat is from the outside towards the system behind the LRH.

Next, since k is negative to the past of $\mathcal{P}$, the approximate boost Killing vector field χ is related to the vector field of null generators as

$$\chi^{\mu} = -\kappa\lambda k^{\mu}\ , \tag{5.118}$$

where $\kappa = const.$ As one can check, multiplying 5.118 by $\chi^{\nu}\nabla_{\nu}$,

$$\begin{aligned}\chi^{\mu}\nabla_{\mu}\chi^{\nu} &= -\kappa\left(k^{\nu}\chi^{\mu}\partial_{\nu}\lambda + \lambda(-\kappa\lambda)k^{\mu}\nabla_{\mu}k^{\nu}\right)\\ &= -\kappa(\chi^{\mu}\partial_{\mu}\lambda)k^{\nu}\ .\end{aligned} \tag{5.119}$$

The second term in the curly brackets vanishes because k satisfies affinely parametrized geodesic equation. Moreover, for the same reason

$$k^{\mu}\partial_{\mu}\lambda = 1 \implies \chi^{\mu}\partial_{\mu}\lambda = -\kappa\lambda\ , \tag{5.120}$$

where we used 5.118. Consequently,

$$\chi^{\mu}\nabla_{\mu}\chi^{\nu} = \kappa\chi^{\nu}\ . \tag{5.121}$$

The approximate Killing vector field χ satisfies the Killing's equation at p. Moreover, as the surface element is $d\Sigma^{\mu} = k^{\mu}d\lambda dA$ we have that

$$\delta Q = \kappa\int T_{\mu\nu}(-\lambda)k^{\mu}k^{\nu}d\lambda dA\,. \tag{5.122}$$

We will now relate the heat flux to entropy using Clausius' relation $\delta Q = T\delta S$. This relationship is thermodynamical in nature. Next, from the consideration of black

[51]The region of spacetime that is accessible to an observer which is in the past of p, the lower wedge in Fig. 5.6.

hole thermodynamics, we will assume that entropy is related to the area of the causal horizon as

$$\delta S = \alpha \delta A \,. \tag{5.123}$$

If we were looking at the event horizon of the black hole, $\alpha = 1/4$. This follows from consideration of Hawking radiation. Here, we do not know what α is. In other words, α is determined by microscopic theory of spacetime. Moreover, the only assumption in the proof of the second law which cannot be generalized to causal horizons is the cosmic censorship conjecture, replacing the assumption that the generators must be future complete. Thus, under assumption of null energy condition and future completeness of the horizon generators, using the same steps as in the original derivation, one can show that the second law holds for causal horizons [140]. This makes area of the causal horizon analogous to entropy.

On the other hand, from purely geometrical considerations, we known that the change in cross sectional area of the bundle of generator of the horizon is

$$\delta A = \int_{\mathcal{H}} \theta d\lambda dA \,, \tag{5.124}$$

which follows from the definition of expansion,

$$\theta = \frac{1}{\delta A}\frac{d\delta A}{d\lambda} \,. \tag{5.125}$$

We will now impose an instantaneous equilibrium condition — shear σ and expansion θ vanish on $\mathcal{P}$. This is necessary to justify the use of Clausius relation, as the relation is valid only in equilibrium. In other words, energy flux δQ focuses the horizon generators near $\mathcal{P}$ in just the right rate that the expansion vanishes at $\mathcal{P}$. Hence, using affinely parametrized Raychaudhuri's equation,

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \sigma_{\mu\nu}\sigma^{\mu\nu} - R_{\mu\nu}k^\mu k^\nu \,, \tag{5.126}$$

and taking into consideration the instantaneous equilibrium condition, one obtains

$$\theta = -\lambda R_{\mu\nu}k^\mu k^\nu \,, \tag{5.127}$$

integrating over the bundle of generators near $\mathcal{P}$. The terms θ^2 and σ^2 are considered as higher order contributions so we neglected them. Inserting 5.127 into 5.124 leads to

$$\delta A = \int R_{\mu\nu} k^\mu k^\nu (-\lambda) d\lambda dA \, . \tag{5.128}$$

On the other hand, Clausius' relation requires that $\delta Q = T\delta S$,

$$\delta Q = T\delta S = \frac{\kappa}{2\pi} \alpha \delta A \, , \tag{5.129}$$

resulting with the use of 5.123. Inserting 5.122 into the left-hand side and 5.127 into the right-hand side we have that

$$T_{\mu\nu} k^\mu k^\nu = \frac{1}{2\pi} \alpha R_{\mu\nu} k^\mu k^\nu \implies \frac{2\pi}{\alpha} T_{\mu\nu} = R_{\mu\nu} + f g_{\mu\nu} \, , \tag{5.130}$$

where the implication follows because $g_{\mu\nu} k^\mu k^\nu = 0$, and f is some function which we now determine. Taking divergence of 5.130 gives us

$$\begin{aligned} \frac{2\pi}{\alpha} \nabla^\mu T_{\mu\nu} &= \nabla^\mu R_{\mu\nu} + \nabla_\nu f \\ 0 &= \nabla_\nu (\frac{1}{2} R + f) \implies \frac{1}{2} R + f = \Lambda \, . \end{aligned} \tag{5.131}$$

In the second line we used the fact that EMT is covariantly conserved and contracted Bianchi identity. The term Λ is a constant. Consequently, we obtain that

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{2\pi}{\alpha} T_{\mu\nu} \, . \tag{5.132}$$

This is Einstein's equation. We also know that the constant of proportionality $\alpha = 1/4\text{G}$, and we interpret Λ as cosmological constant whose value remains undetermined.

It is worth to remark one more time that the derivation of Einstein's equations from the Clausius's relation assumes the existence of local equilibrium condition, as Clausius' relation applies only to variations between nearby stationary states. For example, in adiabatic gas expansion, entropy is not associated with the heat flow and one cannot use Clausius' relation. Another reason to use the equilibrium assumption is that temperature and entropy are not well defined concepts away from the equilibrium. In other words, LRH that are instantaneously stationary are in local equilibrium.

One more comment concerns the fact that we have considered the left wedge instead of the right one. That is, it is more natural to consider an observer asymptotically approaching one component of future horizon from the right side (see Fig. 5.7). The problem is that such observer would perceive energy flow across the horizon out of the right region as positive. Consequently, the heat flow as defined in 5.122 would have an opposite sign, which would lead to Einstein's equations with the wrong sign. To fix this, the following adjustment is suggested in [141]. One should move the

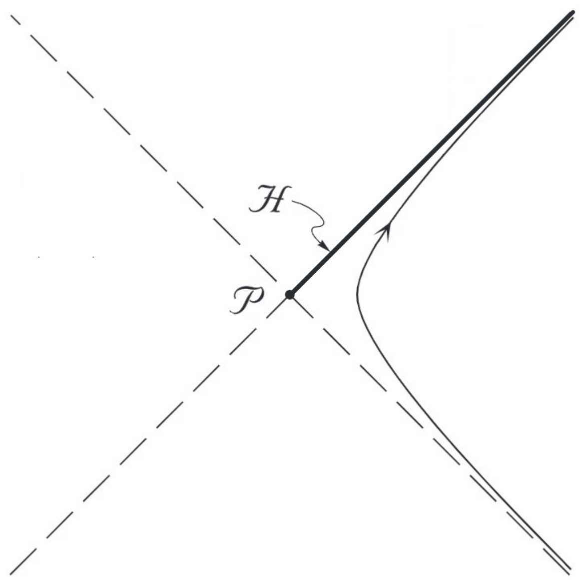


Figure 5.7: $\mathcal{P}$ is 2-dimensional spacelike hypersurface containing p. The local Rindler horizon is the bold line [65].

bifurcation point, the point at which χ vanishes, to the past (see Fig. 5.8). Let's now

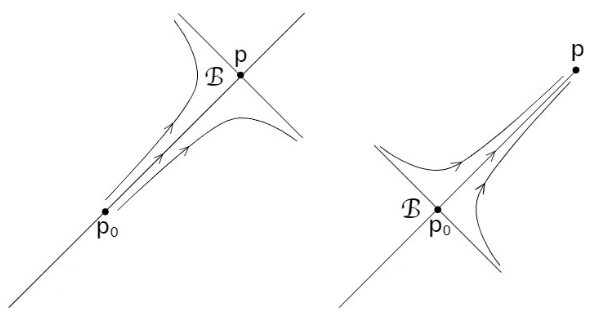


Figure 5.8: LRH and Killing vector field in the old (left) and new (right) setup. The arrows indicate the flow lines of χ, which vanishes at p on the left, and at p_0 on the right [141].

see how modifying the location of the bifurcation point changes the heat flux. The new Killing vector field vanishes at p_0, so $\chi^\mu_N = \kappa(\lambda - \lambda_0)k^\mu$, where λ_0 is the value of λ at p_0. This time there is no minus sign because k and χ_N both point to the future. The boost energy current is

$$T^{\mu\nu}\chi_\nu = (\lambda - \lambda_0)T^{\mu\nu}k_\nu \; . \tag{5.133}$$

Integral from the bifurcation point to p is

$$\int_{\lambda_0}^{0} (\lambda - \lambda_0) = -\frac{\lambda_0^2}{2} \; . \tag{5.134}$$

which is the same result as would have been obtained in the previous case,

$$\int_{\lambda_0}^{0} (-)\lambda d\lambda = -\frac{\lambda_0^2}{2} \; . \tag{5.135}$$

Then, by applying Clausius relation to interval from $\lambda = \lambda_0$ where χ vanishes, to $\lambda = 0$ where expansion and shear vanish yields Einstein's equations (in the limit where $\lambda_0 \to 0$).

To summarize, it was shown using the Clausius relation and the second law on one side, and Raychaudhuri's equation on the other, that the geometry satisfies Einstein's equations at arbitrary point of spacetime. Clausius relation is purely thermodynamical in nature, relating the change in heat to the change in entropy for systems in local equilibrium. The bridge between thermodynamics and geometry is analogy between area of the cross section of the causal horizon and entropy. The final piece comes from Raychaudhuri's equation which allows one to determine the change in area based only on the behaviour of geodesics. This result suggests that Einstein's equation may be viewed as an equation of state. In thermodynamics, if the entropy $S(E,V)$ is known, as a function of energy E and volume V, one can deduce the equation of state with the help of the first law $dQ = dE + pdV$. Differentiation of $S(E,V)$ gives

$$dS = \frac{\partial S}{\partial E}dE + \frac{\partial S}{\partial V}dV \; . \tag{5.136}$$

As a consequence of the first law it can be inferred that

$$\frac{1}{T} = \frac{\partial S}{\partial E} \ , \quad p = T\frac{\partial S}{\partial V} \ . \tag{5.137}$$

The last equation is the equation of state.

Finally, it should be pointed out that the question we haven't managed to discuss is what kind of implications the non-equilibrium situation would have on Einstein's equation. This case was considered in [142].

6 Summary and outlook

6.1 Summary

In the paper we reviewed some of the indications that gravitation may be an emergent phenomenon. Its universal nature, perturbative non-renormalizability and the analogy of four laws of black hole mechanics with the four laws of thermodynamics.

Motivated by these properties, which are usually attributed to effective, low-energy theories, a wide variety of models was developed, mostly based on the ideas from solid state physics, trying to reproduce some part of General Relativity — spacetime, and/or its dynamics, from more fundamental constituents. An overview of different approaches is given, along with the constraints the models should satisfy to reproduce General Relativity.

One such idea, representing the central theme of aforementioned emergent theories, is Causal Set Theory. Its premise is that spacetime as a smooth Lorentzian manifold is just a coarse-grained consequence of the causal relations between 'atoms' of spacetime — points without structure. In other words, causal relations are considered fundamental, while all the other mathematical structures on which General Relativity is based on can be derived from there. The main concepts of the approach are explored, without much technical details.

Another interesting proposition thoroughly discussed in the paper is Weinberg-Witten theorem. It constraints the spin of particles living on flat spacetime. We proved the theorem and discussed the implications it has on graviton in the light of emergent theories. Namely, on flat spacetime graviton cannot be a constructed as a composite particle from elementary particles of the Standard Model.

A different category of emergent models is based on relationship between thermodynamics and general relativity. We derived the four laws of black hole mechanics which initiated the emergent viewpoint. Some weaknesses in the analogy are pointed out and discussed. Lastly, Einstein's field equations are derived from Clausius' relation, the second law of thermodynamics and Raychaudhuri's equation, extending the relationship from black holes to causal horizons.

6.2 *Outlook*

Most Emergent Gravitation approaches are far from reproducing all aspects of General Relativity. In most cases they manage to replicate some elements under certain circumstances. With experimental verification of the ideas currently out of reach, there is no decisive answer on whether they are moving in the right direction or not. Furthermore, although there certainly are indications that gravitation is emergent phenomenon, the arguments rely mostly on analogies and it is not clear to what extent these similarities should be taken seriously.

Appendices

Appendix A Remarks concerning Noether's theorems

Another subtlety concerning the (first) Noether's theorem is that it the fact that *all* fields must satisfy equations of motion in order for a conservation law to exist. The case when field may not satisfy the field equations is when it is a background field. To be more precise, its dynamics is not given by the particular action. This entails that the action we are considering isn't "whole", it does not contain all the information. We use such actions when we are interested only in certain aspects of the theory. It is easiest to show what this means on an example. Consider

$$S[\psi; A_\mu, \eta_{\mu\nu}] = \int d^4x \left[\overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi + \overline{\psi}\gamma^\mu\psi A_\mu\right] . \tag{A.1}$$

This is an action of a theory describing dynamics of electron on *fixed* electromagnetic field when there is no presence of gravity. The background fields are A_μ and $\eta_{\mu\nu}$, which is stressed by the semi-colon. If, for the sake or argument, one tries to obtain field equations for A_μ from A.1 for example, the result is

$$\overline{\psi}\gamma^\mu\psi = 0 . \tag{A.2}$$

As a consequence, the interaction term in Lagrangian should vanish. This insensible result tells us that something is missing, a term that would contain derivatives of A_μ. Thus, it makes sense to vary only dynamical fields of the action when looking for field equations. Only their dynamics is correctly predicted in the presence of the background fields, while they are considered as fixed.

On the other hand, performing a transformation of the action to check if it is a symmetry affects *all* fields which have the right degrees of freedom, even if the fields are background fields. It turns out, one may still obtain conservation laws only with dynamical fields satisfying the field equations if the global symmetry group is not a subgroup and the background field is invariant under the transformations. For In other words, a group of symmetry transformations of the action is restricted to the group leaving the background fields invariant.

It should be pointed out that it is not just the question of conserved quantities, but whether or not some transformation can be considered a symmetry transformation according to definition 4.98. As it was mentioned in the previous section, symmetry transformations permute solutions of the field equations. There we considered the spacetime transformations of theories with metric as the background fields, so we now look at such case. First, notice that there is a problem with the way the global symmetry is defined. Lie derivative, which is tensorial quantity reduced to partial derivative *in Cartesian coordinates*. As a result, global spacetime transformations include only translation. For example, in Cartesian coordinates vector field generating rotation in a plane is $\xi = \xi^\mu \partial_\mu = -y\partial_x + x\partial_y$. Since components of the vector field are coordinate dependent, rotation doesn't fit into definition of global transformations, although all points are rotated by the same amount, which suits the notion of global transformations. Hence, in general, in case of *global* spacetime transformations, relationship between symmetry and current should be

$$\left[\frac{\delta \mathcal{L}}{\delta \psi_i}\right]_{EL} \mathfrak{L}_\xi \psi_i = -\partial_\mu (T^\mu_\nu \xi^\nu(x)) \,, \tag{A.3}$$

which looks the same as when dealing with local group, so it seems that it does not lead to conservation law. However, for global transformations one can always find coordinates in which the parameter of transformations is not a function of coordinates. In the example of rotation in the plane, in polar coordinates one has $\xi = \phi\partial_\phi$, where ϕ is constant. Thus, when field equations are satisfied the conserved current resulting from spacetime symmetry is (in Cartesian coordinates)

$$\partial_\mu (T^\mu_\nu \xi^\nu(x)) = 0 \,. \tag{A.4}$$

In the context of Weinberg–Witten theorem, only translations are important as they lead to energy-momentum four-vector.

Now, for any theory where metric is the background field $S[\psi_i; g_{\mu\nu}]$, one has according to 4.96,

$$\begin{aligned}
\mathfrak{L}_\xi \mathcal{L} &= \frac{\partial \mathcal{L}}{\partial \psi_i} \mathfrak{L}_\xi \psi_i + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_i)} \mathfrak{L}_\xi (\partial_\mu \psi_i) + \frac{\partial \mathcal{L}}{\partial g_{\mu\nu}} \mathfrak{L}_\xi g_{\mu\nu} \\
&= \left[\frac{\partial \mathcal{L}}{\partial \psi_i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_i)}\right)\right] \mathfrak{L}_\xi \psi_i + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_i)} \mathfrak{L}_\xi \psi_i\right) + \frac{\partial \mathcal{L}}{\partial g_{\mu\nu}} \mathfrak{L}_\xi g_{\mu\nu} \,,
\end{aligned} \tag{A.5}$$

where $\delta = \mathfrak{L}_\xi$ for spacetime transformations. For dynamical fields satisfying field equations, the above expression reduces to

$$\partial_\mu \left[\left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i \right)} \mathfrak{L}_\xi \psi_i \right) - \mathfrak{L}_\xi \mathcal{L} \right] + \frac{\partial \mathcal{L}}{\partial g_{\mu\nu}} \mathfrak{L}_\xi g_{\mu\nu} = 0 \,. \tag{A.6}$$

Thus, only if $\mathfrak{L}_\xi g^{\mu\nu} = 0$ the transformation is a symmetry. What's more,

$$\mathfrak{L}_\xi g^{\mu\nu} = 0 \implies \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0 \,. \tag{A.7}$$

Vector fields that satisfies A.7 are called Killing vector fields. They are are infinitesimal generators of isometries, symmetries of metric tensor. For the flat metric, they are translations and spacetime rotations. Hence, as it was mentioned, if metric field is background field, it is its isometry group that determines the symmetry group of the theory.

Final remark is the question of uniqueness of the conserved currents. Let's look at EMT, although the same is true for internal current. As one may have noticed, one can always add a so-called super-potential term to Noether's currents, i.e., divergence of an antisymmetric tensor,

$$T'^{\mu\nu} = T^{\mu\nu} + \partial_\rho \chi^{\mu\rho\nu} \,, \quad \chi^{\mu\rho\nu} = -\chi^{\rho\mu\nu} \,. \tag{A.8}$$

The antisymmetry of super-potential in the first two indices ensures that if $T^{\mu\nu}$ is conserved, then so is $T'^{\mu\nu}$. On the other hand, the conserved charged formed from either of the conserved currents is the same, provided that $\chi^{0i\nu}$ decreases with distance fast enough. This freedom can be used to form EMT with certain useful properties. The problem is that different improvements of EMT represent different localizations of energy and momentum, which leads to a question if there is a way to determine the correct current and energy-momentum tensor. Before we answer that, let's discuss what problems the super-potential can fix.

The first problem the super-potential can fix is that Noether's theorem does not produce gauge invariant EMT or current. This is problematic since EMT should describe measurable quantities, which are gauge invariant.

Furthermore, if EMT is not symmetric, the angular momentum tensor $M^{\mu\nu\rho}$ obtained from rotational symmetry of the theory is not related to momentum the same way as in classical mechanics, i.e., by the form

$$M^{\mu\nu\rho} = x^{\nu} T^{\mu\rho} - x^{\rho} T^{\mu\nu} \,. \tag{A.9}$$

From here it follows that

$$0 = \partial_{\mu} M^{\mu\nu\rho} = T^{\nu\rho} - T^{\rho\nu} \,. \tag{A.10}$$

Otherwise, for non-symmetric EMT, there is an extra term,

$$M^{\mu\nu\rho} = x^{\nu} T^{\mu\rho} - x^{\rho} T^{\mu\nu} + s^{\mu\nu\rho} \,, \quad s^{\mu\nu\rho} = -s^{\mu\rho\nu} \equiv \frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\psi)} M^{\nu\rho}\psi \,, \tag{A.11}$$

where $M_{\nu\rho}$ is a finite dimensional representation of Lie algebra of Lorentz group under which the fields transform. Thus, $M_{\nu\rho} = 0$ for scalar field, $M_{\nu\rho} = M_{\nu\rho}$ for a vector field, etc. Non-symmetry of EMT is an indicator that the fields contributing to it don't transform covariantly under Lorentz group. Moreover, the symmetric property of EMT is important when coupling it to gravity. Action should be a Lorentz scalar and the only way to implement Lorentz invariant interaction terms is by coupling the source to conserved current. In electromagnetism for example, photon field A_{μ}, which is a four-vector, couples to four-vector current $j^{\mu} = \bar{\psi}\gamma^{\mu}\psi$. In the same way, the source of gravity should be a symmetric second rank tensor to be able to couple to gravitational field $g_{\mu\nu}$, symmetric second rank tensor. On flat spacetime, the most common method of improving non-symmetric EMT to symmetric is by Belifante [143], although the same result can be obtained by exploiting the gauge invariance of Lagrangian [144][87], which in fact exploits the second Noether's theorem. In short, the problem with EMT missing some important properties is not in the Noether's procedure, but in variations we choose. In other words, starting with variation that lacks gauge invariance results in currents that are not gauge invariant. Either way, even though EMT can be made symmetric with the help of super-potential, it is still not unique. There are infinite number of super-potentials that result in symmetric EMT.

Finally, the ambiguity is removed with the fact that gravity couples to real EMT. In other words, gravity identifies EMT. Using this fact, the EMT one obtains is so-called

Einstein–Hilbert EMT,

$$
\begin{aligned}
\delta S_M[g_{\mu\nu};\psi_i] &= \int d^4x\delta(\sqrt{-g}\mathcal{L}_M) \\
&= \int d^4x\left[\delta(\sqrt{-g})\mathcal{L}_M + \sqrt{-g}\frac{\delta\mathcal{L}_M}{\delta g_{\mu\nu}}\delta g_{\mu\nu} + \sqrt{-g}\frac{\delta\mathcal{L}_M}{\delta\psi_i}\delta\psi_i\right] \\
&= -\int d^4x\frac{\sqrt{-g}}{2}\left[-2\frac{\delta\mathcal{L}_M}{\delta g_{\mu\nu}} + g^{\mu\nu}\mathcal{L}_M\right]\delta g_{\mu\nu}\,,
\end{aligned}
\tag{A.12}
$$

where S_M is the action that takes into account all fields, i.e., matter (which includes the gauge fields) as dynamical, except gravitational field. The third line follows from assumption that all dynamical fields satisfy field equations. Then, the action is diffeomorphism invariant if

$$
0 = -\int d^4x\frac{\sqrt{-g}}{2}T^{\mu\nu}\mathfrak{L}_\xi g_{\mu\nu}\,,\quad T^{\mu\nu} \equiv \frac{-2}{\sqrt{-g}}\frac{\delta(\sqrt{-g}\mathcal{L}_M)}{\delta g_{\mu\nu}} = -2\frac{\delta\mathcal{L}_M}{\delta g_{\mu\nu}} + g^{\mu\nu}\mathcal{L}_M\,. \tag{A.13}
$$

One obtains covariant conservation law

$$
0 = \int d^4x\sqrt{-g}T^{\mu\nu}\nabla_\mu\xi_\nu = -\int d^4x\sqrt{-g}\nabla_\mu T^{\mu\nu}\xi_\nu\,, \tag{A.14}
$$

since for diffeororphisms,

$$
\delta g_{\mu\nu} = -\mathfrak{L}_\xi g^{\mu\nu} = 2\nabla_\mu\xi_\nu\,. \tag{A.15}
$$

EMT obtained this way is manifestly symmetric, due to the variation being symmetric, and gauge invariant, i.e., invariant under general diffeomorphisms, since the variations are given by Lie derivatives, tensorial quantities. In flat spacetime in appropriate coordinates the covariant conservation law reduces to “ordinary" conservation law,

$$
g_{\mu\nu} = \eta_{\mu\nu} \implies \partial_\mu T^{\mu\nu} = 0\,. \tag{A.16}
$$

What’s more, EMT obtained in such way coincides with EMT obtained by Belifante procedure [145][146]. Notice however that this EMT contains only matter fields, i.e., all fields except gravity, which is expected since covariant conservation laws describe exchange of conserved quantities, i.e., energy and momentum between matter field and gravitation fields in this case. This doesn’t mean that gravity doesn’t have its energy-momentum tensor. In fact, since sources exchange energy and momentum locally, one expects the gravitational field has its own EMT. This is in fact the goal of the next section. What's more, one can obtain conservation law by adding the gravitational energy-momentum tensor to covariant conservation.

In flat spacetime, no gravitational field is present so covariant conservation law reduces to ordinary conservation law [83][85][86][87][88][71][89].

Appendix B Poincaré group

Poincaré group in four-dimensional spacetime is a ten-dimensional, noncompact Lie group that is a semiproduct of translation group $\mathbb{R}^4$ consisting of four spacetime translations, and homogeneous Lorentz group, pseudo-orthogonal group $O(3,1)$ consisting of three rotations and three Lorentz boosts,

$$\mathbb{R}^4 \rtimes O(3,1)\,. \tag{B.1}$$

The multiplication rule of the group is given by

$$(\tilde{\Lambda}, \tilde{a})(\Lambda, a) = (\tilde{\Lambda}\Lambda,\ \tilde{\Lambda}a + \tilde{a})\,, \tag{B.2}$$

where (Λ, a) denotes general element of the group. From B.2 one concludes that the neutral element must be $(\mathbb{1}, 0)$, and the inverse element is $(\Lambda^{-1}, -\Lambda^{-1}a)$. Translation subgroup is obtained by setting $\Lambda = \mathbb{1}$, while with $a = 0$ one obtains the homogeneous Lorentz group. Furthermore, homogeneous Lorentz group consists of all invertible, linear matrices $\Lambda \in GL(4, \mathbb{R})$ satisfying

$$\Lambda^T \eta \Lambda = \eta\,, \tag{B.3}$$

where $\eta = \mathrm{diag}(1, -1, -1, -1)$. To shorten the notation $(\Lambda, 0) \equiv \Lambda$. The group consists of four components. This is shown by taking the determinant of B.3,

$$\begin{aligned} \det(\Lambda^T \eta \Lambda) &= \det(\eta) \\ \det(\Lambda^T)\det(\eta)\det(\Lambda) &= \det(\eta) \\ \det(\Lambda)^2 = 1 \implies \det(\Lambda) &= \pm 1\,, \end{aligned} \tag{B.4}$$

where the second line follows from the rule for determinant of matrix product, and the third from the fact that determinant of a matrix is the same as the determinant of its transpose. Finally, determinant of any Λ belonging to homogeneous Lorentz group is either $+1$ or -1. Action of elements with negative determinant flips the orientation of space. Since determinant is continuous function, the subgroups with different signs of determinant are disconnected.

Next, consider $(0,0)$ component of B.3,

$$\begin{gathered}\Lambda^{\mu}_{0}\eta_{\mu\rho}\Lambda^{\rho}_{0} = 1 \\ \left(\Lambda^{0}_{0}\right)^{2} - \sum_{i=0}^{i=3}\left(\Lambda^{i}_{0}\right)^{2} = 1 \implies |\Lambda^{0}_{0}| \geq 1\,.\end{gathered} \tag{B.5}$$

Hence Λ^{0}_{0} is either greater than of equal to $+1$, or less than or equal to -1. For connected component Λ^{0}_{0} is continuous function, so one ends up with two disconnected groups.

Hence, depending on signs of $\det\Lambda$ and Λ^{0}_{0} homogeneous Lorentz group has four disconnected subgroups which are all continuous. It also contains a discrete subgroup $\{\mathbb{1}, P, T, PT\}$, where $T =$ diag$(-1,1,1,1)$, $P =$ diag$(1,-1,-1,-1)$. The continuous subgroup with $\det\Lambda = 1$ and $\Lambda^{0}_{0} \geq 1$ is called proper orthochronous Lorentz group $SO(1,3)^{\uparrow}_{+}$. Proper refers to space orientation, and orthochronous to time direction. Elements of $SO(1,3)^{\uparrow}_{+}$ preserve orientation of time and space. Other subgroups can be obtained by multiplication with elements of the discrete subgroup [27][81],

- Proper orthochronus — $SO(1,3)^{\uparrow}_{+}$: $\det\Lambda = 1$, $\Lambda^{0}_{0} \geq 1$
- Improper orthochronus — $P \times SO(1,3)^{\uparrow}_{-}$: $\det\Lambda = -1$, $\Lambda^{0}_{0} \geq 1$
- Proper nonorthochronus — $T \times SO(1,3)^{\downarrow}_{+}$: $\det\Lambda = 1$, $\Lambda^{0}_{0} \leq -1$
- Improper nonorthochronus — $PT \times SO(1,3)^{\downarrow}_{-}$: $\det\Lambda = 1$, $\Lambda^{0}_{0} \leq -1$

Laws of physics are not invariant under parity transformations (weak force for example violates parity symmetry of the interaction) and indirectly under time reversal transformations [71], hence the focus will be on studying only proper orthochronous subgroup of homogeneous Lorentz group to which it will be referred to as Lorentz group.

It is yet left to determine Λ by solving B.3. This is hard to do in the given form, but can be done by looking at element of the group near identity $(\mathbb{1}, 0)$. They are referred to as infinitesimal elements or transformations. Thus, for Poincaré group[52]

$$(\mathbb{1} + \omega, \varepsilon) = \mathbb{1} + \frac{i}{2}\omega_{\mu\nu}M^{\mu\nu} - i\varepsilon_\mu P^\mu \,, \tag{B.6}$$

where $\omega_{\mu\nu}$ and ε_μ are infinitesimal and real parameters, P^μ is a vector and $M_{\mu\nu}$ is a matrix. They belong to tangent space at the identity of Lorentz group. This will be explained in the next section. With B.6, equation B.3 reduces to

$$\begin{aligned} \eta_{\alpha\beta} &= (\delta^\mu_\alpha + \omega^\mu_\alpha)(\delta^\mu_\beta + \omega^\mu_\beta)\eta_{\mu\nu} \\ &= \eta_{\alpha\beta} + \delta^\mu_\alpha \omega_{\mu\beta} + \delta^\nu_\beta \omega_{\nu\alpha} \implies \omega_{\alpha\beta} = -\omega_{\beta\alpha} \,, \end{aligned} \tag{B.7}$$

where we defined

$$\omega_{\alpha\beta} \equiv \frac{1}{2}\omega_{\mu\nu}\left(M^{\mu\nu}\right)_{\alpha\beta} \,. \tag{B.8}$$

Result B.7 implies that $M^{\mu\nu} = -M^{\nu\mu}$. Hence, near identity B.3 reduces to linear equation. It then follows that elements of Lorentz group near identity should be antisymmetric. Although this is only local description there exists a map, called exponential map, using which one can obtain most of global properties of the group from local ones.

$$\Lambda = e^{\frac{i}{2}\omega_{\mu\nu}M^{\mu\nu}} \,. \tag{B.9}$$

Next, consider translation subgroup. One can show that it is additive,

$$(\mathbb{1}, a)(\mathbb{1}, \bar{a}) = (\mathbb{1}, a + \bar{a}) \,. \tag{B.10}$$

Then, for any integer n,

$$(\mathbb{1}, a) = \left(\mathbb{1}, \frac{a}{n}\right)^n \,. \tag{B.11}$$

[52] Parametrization of infinitesimal element is done this way for later connection with unitary representation. It ensures that generators of unitary representation are hermitian.

For $n \to \infty$ it follows $\frac{a}{n} \to \varepsilon$. As $(\mathbb{1}, \varepsilon) = 1 - i\varepsilon_\mu P^\mu$ according to B.6, one can write

$$(\mathbb{1}, a) = \lim_{n\to\infty} \left(\mathbb{1} - i\frac{a_\mu}{n} P^\mu\right)^n = e^{-ia_\mu P^\mu} . \tag{B.12}$$

Thus, we have obtained general expression for element of translation. Because the group is Abelian and additive the exponential map coincides with series expansion. This is in general true only for matrix Lie groups.

Formally, local description of group properties is in the scope of Lie algebra, which is introduced in the next section. All in all, it was demonstrated that algebra is easier to work with. What's more, except for some global properties, such as topological structure, properties of the group, and naturally its representations, can be extracted by considering its algebra. Thus, the next objective is determining algebra of Poincaré group.

Appendix C Poincaré algebra

In this section we review Lie algebras and determine commutator relations of the Poincaré group.

As it was mentioned, Poincaré group is a Lie group, which means that besides the group structure, it is also a smooth differentiable manifold. This entails that there is a tangent space at each point, i.e., at each element of the group. Furthermore, the group has the same local properties at each point as all tangent spaces are isomorphic [147]. For that reason it is the most common to describe local properties at the identity elements. What's more, every group is guaranteed to have it.

Tangent space is first of all a vector space. Using the basis on vector space one is able to write down element of the group near identity, as B.6. Not only that, certain vectors generate one parameter subgroup of Lie group G via exponential map $\exp$,

$$\exp : \mathfrak{g} \to G \, . \tag{C.1}$$

Thus, they are appropriately referred to as generators. Accordingly, P^μ and $M^{\mu\nu}$ in B.6 are generators of Poincaré group. Beyond that, when it comes to the structure of the group, multiplication rule of the group elements, or more precisely, their failure to commute, locally narrows down to so-called Lie bracket. Additionally, associativity of the group elements leads to Jacobi identity. Tangent space with respect to this structures constitutes Lie algebra.

Also, Lie brackets have important physical interpretation. Generators in unitary representation are hermitian operators and as such represent physical observables, quantities one can measure to determine the state of the system. Non-vanishing Lie bracket of the generators indicates that the order in which measurements are performed is important. In other words, performing consecutive measurement represented by non-commuting generators changes initial state of the system. Consequently, state of the system is determined by carrying out measurements represented by generators that mutually commute, so state vectors will be labeled by eigenvalues of mutually commuting generators.

This being said, before deriving the Lie brackets of Poincaré group let's show how the generators transform under Poincaré transformations, which in fact gives us the interpretation of the generators. Start by looking at series of transformations

$$(\Lambda, a)(\mathbb{1}+\omega, \varepsilon)(\Lambda, a)^{-1} = (\Lambda, a)\left(\mathbb{1} + \frac{i}{2}\omega_{\mu\nu}M^{\mu\nu} - i\varepsilon_\mu P^\mu\right)(\Lambda, a)^{-1}\,. \tag{C.2}$$

The total effect of the transformations is obtained by using composition rule B.2,

$$\begin{aligned}(\Lambda, a)(\mathbb{1}+\omega, \varepsilon)(\Lambda, a)^{-1} &= (\Lambda\omega\Lambda^{-1}, \Lambda\varepsilon - \Lambda\omega\Lambda^{-1}a)\\ &= \mathbb{1} + \frac{i}{2}(\Lambda\omega\Lambda^{-1})_{\mu\nu}M^{\mu\nu} - i(\Lambda\varepsilon - \Lambda\omega\Lambda^{-1}a)_\mu P^\mu\,,\end{aligned} \tag{C.3}$$

where the last equality follows from B.6 since resulting transformation is infinitesimal because ε and ω are infinitesimal parameters. The right-hand side of C.2 is

$$\begin{aligned}(\Lambda, a)\left(\mathbb{1} + \frac{i}{2}\omega_{\mu\nu}M^{\mu\nu} - i\varepsilon_\mu P^\mu\right)(\Lambda, a)^{-1} = \mathbb{1} &+ \frac{i}{2}\omega_{\mu\nu}(\Lambda, a)M^{\mu\nu}(\Lambda, a)^{-1}\\ &- i\varepsilon_\mu(\Lambda, a)P^\mu(\Lambda, a)^{-1}\,.\end{aligned} \tag{C.4}$$

Finally, comparing coefficients next to ε and ω of C.2 and C.4 leads to

$$\varepsilon: \quad (\Lambda, a)P^\mu(\Lambda, a)^{-1} = \Lambda^\mu_\rho P^\rho \tag{C.5}$$

$$\begin{aligned}\omega: \quad \frac{i}{2}(\Lambda, a)M^{\mu\nu}(\Lambda, a)^{-1} &= \frac{i}{2}\Lambda^\mu_\alpha\Lambda^\nu_\beta(M^{\alpha\beta} + 2a^\beta P^\alpha) \implies\\ (\Lambda, a)M^{\mu\nu}(\Lambda, a)^{-1} &= \Lambda^\mu_\alpha\Lambda^\nu_\beta(M^{\alpha\beta} + a^\beta P^\alpha - a^\alpha P^\beta)\,.\end{aligned} \tag{C.6}$$

One can write $a^\beta P^\alpha = \frac{1}{2}(a^\beta P^\alpha - a^\alpha P^\beta)$, since the right-hand side is contracted with antisymmetric tensor ω, obtaining the last line of C.6. Hence, Poincaré transformation of generator of translations leads to translation by ΛP. On the other hand, Poincaré transformation of generator of Lorentz transformation turns out to be combination of Lorentz transformation and translation. Moreover, under pure Lorentz transformations, i.e., setting $a = 0$, P^μ and $M^{\mu\nu}$ transform as tensors, while for pure translations, $\Lambda = \mathbb{1}$, one may recognize $M^{\mu\nu}$ transforms as angular momentum of a rigid body a distance a from the origin. Accordingly, the first term in the parentheses of C.6 describes spacetime rotation of an object around it's center of mass, and the second part is angular momentum with respect to the origin. With that being said, the objects considered are elementary particle, point-like object without structure, so the first term is interpreted as spin. On the other hand, P^μ is invariant under pure translations. One can conclude that P^μ is four-momentum, as it is a Lorentz four-vector invariant under translations, which one expects from generator of translations, whilst $M^{\mu\nu}$ generates spacetime rotations.

Finally, Lie brackets stem from local composition law, so it follows that if all transformations in C.5 and C.6 are infinitesimal,

$$\left(\mathbb{1} + \frac{i}{2}\omega_{\mu\nu}M^{\mu\nu} - i\varepsilon_\mu P^\mu\right) P^\rho \left(\mathbb{1} - \frac{i}{2}\omega_{\alpha\beta}M^{\alpha\beta} + i\varepsilon_\alpha P^\alpha\right) = (\delta^\rho_\sigma + \omega^\rho_\sigma)P^\sigma\,, \tag{C.7}$$

$$\left(\mathbb{1} + \frac{i}{2}\omega_{\alpha\beta}M^{\alpha\beta} - i\varepsilon_\alpha P^\alpha\right) M^{\mu\nu} \left(\mathbb{1} - \frac{i}{2}\omega_{\alpha\beta}M^{\alpha\beta} + i\varepsilon_\alpha P^\alpha\right) = (\delta^\mu_\rho + \omega^\mu_\rho)(\delta^\nu_\sigma + \omega^\nu_\sigma)M^{\rho\sigma}\,. \tag{C.8}$$

Keeping only the terms linear in ε and ω,

$$\begin{aligned} i\varepsilon_\mu[P^\rho, P^\mu] + \frac{i}{2}\omega_{\mu\nu}[M^{\mu\nu}P^\rho] &= \omega^\rho_\nu P^\nu \\ &= \eta^{\mu\rho}\omega_{\mu\nu}P^\nu \\ &= \frac{1}{2}\omega_{\mu\nu}(\eta^{\mu\rho}P^\nu - \eta^{\nu\rho}P^\mu)\,, \end{aligned} \tag{C.9}$$

$$\begin{aligned} \frac{i}{2}\omega_{\alpha\beta}[M^{\alpha\beta}, M^{\mu\nu}] &= \omega_{\alpha\beta}(\eta^{\mu\alpha}M^{\beta\nu} - \eta^{\nu\alpha}M^{\mu\beta}) \\ &= \frac{1}{2}\omega_{\alpha\beta}(\eta^{\mu\alpha}M^{\beta\nu} - \eta^{\beta\mu}M^{\alpha\nu} - \eta^{\nu\alpha}M^{\mu\beta} + \eta^{\nu\beta}M^{\mu\alpha})\,, \end{aligned} \tag{C.10}$$

where the last line of C.9 and C.10 follows because of antisymmetric property. Comparing coefficients next to ε and ω leads to

$$[P^\rho, P^\mu] = 0\,, \tag{C.11}$$

$$i[M^{\mu\nu}, P^\rho] = \eta^{\mu\rho}P^\nu - \eta^{\nu\rho}P^\mu\,, \tag{C.12}$$

$$i[M^{\alpha\beta}, M^{\mu\nu}] = \eta^{\mu\alpha}M^{\beta\nu} - \eta^{\beta\mu}M^{\alpha\nu} - \eta^{\nu\alpha}M^{\mu\beta} + \eta^{\nu\beta}M^{\mu\alpha}\,. \tag{C.13}$$

These are Lie brackets of the Poincaré group. Furthermore, one can check by explicit calculation that the Casimir operators one can construct from P^μ and $M^{\mu\nu}$ are

$$C_1 = P^\mu P_\mu\,, \tag{C.14}$$

$$C_2 = W^\mu W_\mu\,, \tag{C.15}$$

where $W_\mu = \frac{1}{2}\varepsilon_{\mu\rho\sigma\nu}M^{\rho\sigma}P^\nu$ is Pauli-Lubanski pseudovector. It's Lie brackets are

$$[P^\mu, W^\nu] = 0\,, \tag{C.16}$$

$$i[M^{\mu\nu}, W^\rho] = (\eta^{\rho\mu}W^\nu - \eta^{\rho\nu}W^\mu)\,, \tag{C.17}$$

$$i[W_\mu, W_\nu] = \varepsilon_{\mu\nu\rho\sigma}W^\rho P^\sigma\,. \tag{C.18}$$

It is also useful to write down the commutator relations for components,

$$M_{ij} = -\varepsilon_{ijk}J^k \quad \Longrightarrow \quad J_i = -\frac{1}{2}\varepsilon_{ijk}M^{jk}\,, \tag{C.19}$$

$$M_{0i} = K_i\,, \tag{C.20}$$

Generator **J** is interpreted as angular momentum (which includes spin and orbital angular momentum), generating space rotations, while **K** is interpreted as generator of boosts, rotations mixing space and time. Relations C.11-C.13 narrow down to

$$\begin{aligned}
[J_i, J_j] = i\varepsilon_{ijk}J_k\,, \quad [J_i, P_j] = i\varepsilon_{ijk}P_k\,, \quad [P_i, P_j] = 0\,. \\
[J_i, K_j] = i\varepsilon_{ijk}K_k\,, \quad [K_i, P_j] = i\delta_{ij}P_0\,, \quad [J_i, P_0] = 0\,. \\
[K_i, K_j] = -i\varepsilon_{ijk}J_k\,, \quad [K_i, P_0] = iP_i\,, \quad [P_i, P_0] = 0\,.
\end{aligned} \tag{C.21}$$

It should be mentioned that time component of four-momentum corresponds Hamiltonian, so generators that commute with it are constants of motion.

This concludes discussion about Lie algebra. Using these relations one may determine the set of operators that mutually commute, but the set is not unique. Any such set is equally valid, but the method used to generate the basis of minimal invariant subspace will favor one.

Appendix D Proof of the zeroth law

To prove that $\kappa = const.$, we must show that is true along the geodesic, but also from one geodesic to another. The conditions are mathematically formulated as

$$\left(\frac{\partial\kappa}{\partial\lambda}\right)_{\theta_A} = \chi^\mu \nabla_\mu \kappa = 0\,, \tag{D.1}$$
$$\left(\frac{\partial\kappa}{\partial\theta_A}\right)_{\lambda} = e^\mu_A \nabla_\mu \kappa = 0\,, \tag{D.2}$$

where both of the equations are evaluated at the horizon. From these relations it follows that κ doesn't change in any direction along the null hypersurface. Since $\kappa = -N^\nu \chi^\mu \nabla_\mu \chi_\nu$ from $\chi^\mu \nabla_\mu \chi_\nu = \kappa \chi_\nu$ we start by first calculating $\nabla_\mu \chi_\nu$. Because it is a tensor, we can decompose it the basis as

$$\begin{aligned}\nabla_\mu \chi_\nu = c_1 \chi_\mu \chi_\nu + c_2 \chi_\mu N_\nu + c_3^A \chi_\mu e_{A\nu} + \\ c_4 N_\mu \chi_\nu + c_5 N_\mu N_\nu + c_6^A N_\mu e_{A\nu} + \\ c_7^A e_{A\mu} \chi_\nu + c_8^A e_{A\mu} N_\nu + c_9^{AB} e_{A\mu} e_{B\nu}\,.\end{aligned} \tag{D.3}$$

The components c_i, $i = 1, ..., 9$ will be determined using the properties $\nabla_\mu \chi_\nu$ must satisfy. The first property we will use is the fact that χ satisfies the Killing equation. As a result

$$\nabla_\mu \chi_\nu = -\nabla_\nu \chi_\mu \implies c_1 = c_5 = c_9 = 0\,, \quad c_2 = -c_4\,,\, c_3^A = -c_7^A\,,\, c_6^A = -c_8^A\,. \tag{D.4}$$

All the terms on the "diagonal" of D.3 are zero, while "upper" and "lower triangle" differ up to a sign. Next, we use the geodesic equation $\chi^\mu \nabla_\mu \chi_\nu = \kappa \chi_\nu$. When D.3 is contracted with χ^μ only terms which combine to $\chi_\mu N^\nu$ survive, so

$$\kappa \chi_\nu = -c_4 \chi_\nu + c_6^A e_{A\nu} \implies -c_4 = c_2 = \kappa\,, c_6^A = -c_8^A = 0\,. \tag{D.5}$$

The only terms we are left with are

$$\nabla_\mu \chi_\nu = \kappa(\chi_\mu N_\nu - \chi_\nu N_\mu) + c_3^A(\chi_\mu e_{A\nu} - \chi_\nu e_{A\mu})\,. \tag{D.6}$$

From here we also obtain, by contracting the above relation with e_B^ν and N^μ,

$$N^\nu e_B^\mu \nabla_\mu \chi_\nu = -c_3^A e_{A\mu} e_B^\mu (N^\nu \chi_\nu) = c_{3B}\,. \tag{D.7}$$

The last equality follows from $\sigma_{AB} = g_{\mu\nu} e_A^\mu e_B^\nu$, where σ_{AB} is spatial part of the metric, and $h^{\mu\nu} = \sigma^{AB} e_A^\mu e_B^\nu$. The next goal is to calculate $\nabla_\rho \kappa$,

$$\begin{aligned}
\nabla_\rho \kappa &= \nabla_\rho(-N^\nu \chi^\mu \nabla_\mu \chi_\nu) \\
&= -(\nabla_\rho N^\nu)\chi^\mu \nabla_\mu \chi_\nu - N^\nu \nabla_\rho(\chi^\mu \nabla_\mu \chi_\nu) \\
&= -\kappa \chi_\nu \nabla_\rho N^\nu - N^\nu (\nabla_\rho \chi^\mu)(\nabla_\mu \chi_\nu) - N^\nu \chi^\mu \nabla_\rho \nabla_\mu \chi_\nu \\
&= \kappa N^\nu \nabla_\rho \chi_\nu - N^\nu (\nabla_\rho \chi^\mu)(\nabla_\mu \chi_\nu) - R_{\nu\mu\rho\tau} \chi^\tau \chi^\mu N^\nu\,.
\end{aligned} \tag{D.8}$$

The first term in the last line is the result of $\nabla_\rho(\chi_\nu N^\nu) = 0 = N^\nu \nabla_\rho \chi^\nu + \chi_\nu \nabla_\rho N^\nu$, where we also used $\nabla_\rho \nabla_\mu \chi_\nu = R_{\nu\mu\rho\tau}\chi^\tau$, which follows from the definition of Riemann tensor and Killing equation. We will first show that D.1 holds.

$$\begin{aligned}
\chi^\rho \nabla_\rho \kappa &= \kappa N^\nu \chi^\rho \nabla_\rho \chi_\nu - N^\nu \chi^\rho (\nabla_\rho \chi^\mu)(\nabla_\mu \chi_\nu) - R_{\nu\mu\rho\tau} \chi^\rho \chi^\tau \chi^\mu N^\nu \\
&= \kappa^2 \chi_\nu N^\nu - N^\nu \kappa \chi^\mu \nabla_\mu \chi_\nu \\
&= -\kappa^2 + \kappa^2 \\
&= 0\,.
\end{aligned} \tag{D.9}$$

To get the second line we used the geodesic equation and antisymmetry of Riemann tensor in the second pair of indices. The third line follows from 5.10 and geodesic equation. We have proven that scalar field κ does not change along the congruence of geodesics. This result is not surprising because metric doesn't change in the direction of Killing vector, which is probably the reason why this part of the proof is not shown in [45]. It is left to prove D.2. The procedure is analogous to the previous case.

$$\begin{aligned}
e^\rho_A \nabla_\rho \kappa &= \kappa N^\nu e^\rho_A \nabla_\rho \chi_\nu - N^\nu e^\rho_A (\nabla_\rho \chi^\mu)(\nabla_\mu \chi_\nu) - R_{\nu\mu\rho\tau} e^\rho_A \chi^\tau \chi^\mu N^\nu \\
&= -\kappa c_{3A} - N^\nu c_{3A} \chi^\mu \nabla_\mu \chi_\nu - R_{\mu\nu\rho\tau} N^\nu \chi^\mu e^\rho_A \chi^\tau \\
&= -\kappa c_{3A} + \kappa c_{3A} - R_{\mu\nu\rho\tau} e^\rho_A \chi^\mu N^\nu \chi^\tau \\
&= R_{\nu\mu\tau\rho}(-g^{\nu\tau} - \chi^\nu N^\tau - h^{\nu\tau}) e^\rho_A \chi^\mu \\
&= -R_{\mu\rho} e^\rho_A \chi^\mu - R_{\mu\nu\tau\rho} \chi^\mu e^\nu_B e^\tau_C e^\rho_A \sigma^{BC}\,.
\end{aligned} \tag{D.10}$$

The second line follows from D.7, the third from geodesic equation, and the fourth from 5.15. The first term in the last line is by definition the Ricci tensor. Moreover, contracting Riemann tensor with the second term in the brackets vanishes because Riemann tensor is antisymmetric in the first two indices while $\chi^\nu \chi^\mu$ is symmetric. The last term follows from expanding $h^{\nu\tau}$ in the basis.

To show that the second term of the last line in D.10 vanishes consider $\nabla_\mu \chi_\nu e^\mu_A e^\nu_B = 0$. This is the statement that transverse components of the velocity gradient is zero, which we know is true because $\widetilde{B}_{\mu\nu} = 0$ as expansion, shear and rotation all vanish. It also holds that everywhere on the horizon $e^\rho_C \nabla_\rho (\nabla_\mu \chi_\nu e^\mu_A e^\nu_B) = 0$. Expanding this equation leads to

$$\begin{aligned}
0 &= (\nabla_\rho \nabla_\mu \chi_\nu) e^\mu_A e^\nu_B e^\rho_C + e^\rho_C e^\nu_B (\nabla_\mu \chi_\nu) \nabla_\rho e^\mu_A + e^\rho_C e^\mu_A (\nabla_\mu \chi_\nu) \nabla_\rho e^\nu_B \\
&= R_{\nu\mu\rho\tau} \chi^\tau e^\mu_A e^\nu_B e^\rho_C - e^\rho_C c_{3B} \chi_\mu \nabla_\rho e^\mu_A + e^\rho_C c_{3A} \chi_\nu \nabla_\rho e^\nu_B \\
&= R_{\nu\mu\rho\tau} \chi^\tau e^\mu_A e^\nu_B e^\rho_C + e^\rho_C c_{3B} e^\mu_A \nabla_\rho \chi^\mu - e^\rho_C c_{3A} e^\nu_B \nabla_\rho \chi^\nu \\
&= R_{\nu\mu\rho\tau} \chi^\tau e^\mu_A e^\nu_B e^\rho_C - e^\rho_C c_{3B} e^\mu_A \nabla_\mu \chi^\rho + e^\rho_C c_{3A} e^\nu_B \nabla_\nu \chi^\rho \\
&= R_{\nu\mu\rho\tau} \chi^\tau e^\mu_A e^\nu_B e^\rho_C\,.
\end{aligned} \tag{D.11}$$

The second line follows from D.7, that is $c_{3A}\chi^\nu = e^\mu_A \nabla_\mu \chi_\nu$. Killing equation was also used. In the third line we used that on the horizon $\nabla_\rho(\chi_\mu e^\mu_A) = 0$. Thus, the second term in D.10 vanishes. Regarding the first term needs some additional conditions.

We rewrite it as

$$
\begin{aligned}
e^{\rho}_{A}\nabla_{\rho}\kappa &= -R_{\mu\rho}\chi^{\mu}e^{\rho}_{A} \\
&= -8\pi(T_{\mu\rho} + \frac{1}{2}Rg_{\mu\rho})\chi^{\mu}e^{\rho}_{A} \\
&= -8\pi T_{\mu\rho}\chi^{\mu}e^{\rho}_{A} \\
&= 8\pi j_{\mu}e^{\mu}_{A} \,.
\end{aligned} \tag{D.12}
$$

In the second line we used Einstein's equations, while in the third lines one can recognize the definition of current. Using Einstein's field equations, it follows from 5.24 that $T_{\mu\nu}\chi^{\mu}\chi^{\nu} = -j_{\nu}\chi^{\nu} = 0$. We further impose dominant energy condition $g_{\mu\nu}j^{\mu}j^{\nu} \leq 0$. This means that the current density of matter is either timelike or null-like, i.e., the current density cannot flow faster then light. Let's write the current vector j in the basis

$$
j^{\mu} = a\chi^{\mu} + bN^{\mu} + c^{A}e^{\mu}_{A} \,. \tag{D.13}
$$

From the null energy condition we have that

$$
j^{\mu}\chi_{\mu} = 0 \implies b = 0 \,, \tag{D.14}
$$

which, using the dominant energy condition yields

$$
\begin{aligned}
g_{\mu\nu}j^{\mu}j^{\nu} &= g^{\mu\nu}(a\chi^{\mu} + c^{A}e^{\mu}_{A})(a\chi^{\nu} + c^{B}e^{\nu}_{B}) \\
&= \sigma_{AB}c^{A}c^{B} \leq 0 \,.
\end{aligned} \tag{D.15}
$$

However, since we obtain only the spatial component, $\sigma_{AB}c^{A}c^{B} \geq 0$. So, the only possible solution is that $c^{A} = 0$. As a result, the current D.13 has only null component. Consequently D.10 is

$$
e^{\rho}_{A}\nabla_{\rho}\kappa = 8\pi j_{\mu}e^{\mu}_{A} = 0 \,. \tag{D.16}
$$

As one moves across the geodesics the value of κ doesn't change. This concludes the proof.

In summary, we have shown that if spacetime admits a Killing vector which is the generator of the event horizon and Einstein's equations hold with matter satisfying the dominant energy condition, the surface gravity κ is constant over the event horizon.

Appendix E Proof of the first law

As mentioned, the first laws is a statement of energy conservation. Before proceeding with the derivation, it needs to be clarified what is meant by energy conservation, since it was shown that it is not possible to talk about gravitational energy in a coordinate independent way. Total energy of asymptotically flat, stationary spacetime is given by Komar integral,

$$\mathcal{Q} = \int_{\partial S_\infty} \nabla^\mu K^\nu d\Sigma_{\nu\mu} \,, \tag{E.1}$$

where ∂S_∞ is a spacelike 2-dimensional hypersurface at infinity, whose surface element is $d\Sigma_{\nu\mu}$. Komar integral associated with Killing vector t^μ of stationary spacetime can be interpreted as total energy (also called mass). In a similar manner, taking $K^\mu = \phi^\mu$, one obtains the expression proportional to total angular momentum J of a stationary axisymmetric spacetime. The reasoning behind Komar's integral is as follows. First of all, energy is associated with time translations. Note that Komar's integral works for both static and stationary spacetime, although Killing vector t^μ in stationary spacetime is timelike only at infinity. It turns out this is all that matters because E.1 is evaluated at spatial infinity. Thus, we are concerned only with what is happening at the boundary. The integrand looks somewhat arbitrary, but further insight is gained by transforming the surface integral E.1 into volume integral E.2.

$$\mathcal{Q} = -\int_{S_\infty} R^\mu_\nu K^\nu d\Sigma_\mu \,, \tag{E.2}$$

where S_∞ is a hypersurface bounded by ∂S and $d\Sigma_\mu$ is a surface element of S_∞. We used Stokes' theorem[53] and property $\nabla_\mu \nabla^\mu K^\nu = -R^\nu_\mu K^\mu$ satisfied by Killing vectors. The integrand $j_K \equiv R^\mu_\nu K^\nu$ reminds of Noether's current $j^\mu = T^\mu_\nu K^\nu$, and one can show that $\nabla_\mu j^\mu_K = 0$. Hence, j_K is some covariantly conserved quantity. Moreover, if one takes hypersurface S_∞ consisting of two disconnected parts, one at infinity, and the other at the horizon, the boundary term at infinity turns out to be proportional

[53] Stokes' theorem states that $\int_{\partial S} \omega^{\mu\nu} d\Sigma_{\nu\mu} = \int_S \nabla_\mu \omega^{\mu\nu} d\Sigma_\nu$.

to mass M, which can be checked by explicit calculations [111]. In summary, it is possible to define total energy (mass) in asymptotically flat spacetime as one can define boundary integrals, concerned only with the behaviour at infinity, where one can meaningfully talk about time and spatial translations which we relate to conservation of energy, momentum and angular momentum. We now turn to proving the first law.

One starts with Komar's integral over hypersurface S, bounded by ∂S

$$\int_{\partial S} \nabla^\mu t^\nu d\Sigma_{\nu\mu} = -\int_S R^\mu_\nu t^\nu d\Sigma_\mu \,, \tag{E.3}$$

where $d\Sigma_{\nu\mu}$ is a surface element of ∂S, and $d\Sigma_\mu$ is a surface element of S. We are considering a black hole spacetime that is stationary axisymmetric asymptotically flat. Thus, we choose S that is spacelike and tangent to axisymmetric Killing vector field ϕ^μ. The boundary ∂S consists of 2-surface ∂B that intersects the event horizon and 2-surface S_∞ at infinity. Hence,

$$\left(\int_{S_\infty} + \int_{\partial B}\right) \nabla^\mu t^\nu d\Sigma_{\nu\mu} = -\int_S R^\mu_\nu t^\nu d\Sigma_\mu \,. \tag{E.4}$$

One can recognize the mass of spacetime M as measured from infinity,

$$\int_{S_\infty} \nabla^\mu t^\nu d\Sigma_{\nu\mu} = -4\pi M \,. \tag{E.5}$$

Rearranging the right-hand side using of E.4 using $R^\mu_\nu = \frac{1}{2} 8\pi \left(2T^\mu_\nu - T\delta^\mu_\nu\right)$, results in

$$M = \int_S \left(2T^\mu_\nu - T\delta^\mu_\nu\right) t^\nu d\Sigma_\mu + \frac{1}{4\pi}\int_{\partial B} \nabla^\mu t^\nu d\Sigma_{\nu\mu} \,. \tag{E.6}$$

The first term can be regarded as contribution of the matter outside the horizon to the total mass, while the second term is viewed as the mass of the black hole. By completely analogous procedure for Killing vector ϕ^μ one obtains

$$\begin{aligned}\left(\int_{S_\infty} + \int_{\partial B}\right) \nabla^\mu \phi^\nu d\Sigma_{\nu\mu} &= -4\pi \int_S \left(2T^\mu_\nu - T\delta^\mu_\nu\right) \phi^\nu d\Sigma_\mu \\ &= -8\pi \int_S T^\mu_\nu \phi^\nu d\Sigma_\mu \,.\end{aligned} \tag{E.7}$$

The second term vanishes due to the choice of hypersurface — S is spacelike, tangent to ϕ^μ, so the normal of its surface element $d\Sigma_\mu$ is timelike. Consequently, $\phi^\mu d\Sigma_\mu = 0$.

The total angular momentum J is

$$8\pi J = \int_{S_\infty} \nabla^\mu \phi^\nu d\Sigma_{\nu\mu}\,, \tag{E.8}$$

which leads to

$$\begin{aligned} J &= -\int_S T^\mu_\nu \phi^\nu d\Sigma_\mu - \frac{1}{8\pi}\int_{\partial B} \nabla^\mu \phi^\nu d\Sigma_{\nu\mu} \\ &= -\int_S T^\mu_\nu \phi^\nu d\Sigma_\mu + J_H\,, \end{aligned} \tag{E.9}$$

where we defined

$$J_H \equiv -\frac{1}{8\pi}\int_{\partial B} \nabla^\mu \phi^\nu d\Sigma_{\nu\mu}\,. \tag{E.10}$$

The first term in E.9 is angular momentum of matter, and the second term is regarded as angular momentum of the black hole. To put everything together, remember that Killing vector field that is null on the horizon is

$$\chi^\mu = t^\mu + \Omega_H \phi^\mu\,. \tag{E.11}$$

Using $\nabla^\mu t^\nu = \nabla^\mu \chi^\nu - \Omega_H \nabla^\mu \phi^\nu$ in E.6, together with E.9, yields

$$M = \int_S \left(2T^\mu_\nu - T\delta^\mu_\nu\right) t^\nu d\Sigma_\mu + 2\Omega_H J_H + \frac{1}{4\pi}\int_{\partial B} \nabla^\mu \chi^\nu d\Sigma_{\nu\mu}\,. \tag{E.12}$$

Let's expand the last integral in the expression. The surface element at the horizon is $d\Sigma_{\nu\mu} = \chi_{[\nu}N_{\mu]}dA$. The normal of the surface element is $\chi_{[\nu}N_{\mu]}d$, and and dA is surface element of the horizon. As before, N_μ is the auxillary null vector, normalized so that $\chi^\mu N_\mu = -1$. It follows that

$$\begin{aligned} \frac{1}{4\pi}\int_{\partial B} \nabla^\mu \chi^\nu d\Sigma_{\nu\mu} &= \frac{1}{4\pi}\int_{\partial B} \nabla^\mu \chi^\nu \chi_{[\nu}N_{\mu]} dA \\ &= \frac{1}{4\pi}\int_{\partial B} N_\mu \chi_\nu \nabla^\mu \chi^\nu dA \\ &= -\frac{1}{4\pi}\int_{\partial B} N_\mu \chi_\nu \nabla^\nu \chi^\mu dA \\ &= \frac{1}{4\pi}\int_{\partial B} \kappa dA \\ &= \frac{\kappa}{4\pi} A\,. \end{aligned} \tag{E.13}$$

In the second line we omitted the antisymmetric braces because $\nabla^{\mu}\chi^{\nu}$ is already antisymmetric, since χ^{μ} satisfies the Killing equation. In the third line we used the Killing's equation and finally recognized the definition of surface gravity, $\kappa = -N^{\nu}\chi^{\mu}\nabla_{\mu}\chi_{\nu}$. The zeroth law states that κ is constant on the horizon, which is why κ can be taken in front of the integral. The total mass E.12 is then

$$M = \int_S \left(2T^{\mu}_{\nu} - T\delta^{\mu}_{\nu}\right) t^{\nu} d\Sigma_{\mu} + 2\Omega_H J_H + \frac{\kappa}{4\pi}A\,. \tag{E.14}$$

This is the integral mass formula. The first term in E.14 is the contribution of matter outside the black hole. Since we are interested in making the correspondence between black holes and thermodynamics, we will ignore it for the time being. In such case, E.14 reduces to so-called Smarr formula,

$$M = 2\Omega_H J_H + \frac{\kappa}{4\pi}A\,. \tag{E.15}$$

The expression is analogous to Euler equation in thermodynamics (also called Gibbs-Duhem equation),

$$U(S,V,N) = TS - pV + \mu N\,, \tag{E.16}$$

where U is internal energy of the system, S is entropy, T is temperature, p is pressure, V is volume, μ is chemical potential and N is number of particles. The Euler equation follows from extensivity of the thermodynamic variables, and is different for system described by a different set variables. The first term in E.15 represents total rotational energy of the black hole, while the last term is analogous to TS term of the Euler equation.

We will now derive the expression of how formula for M changes when infinitesimal, stationary, axisymmetric change is made to the solution. As stated by [45], when comparing two slightly different solutions there is a freedom when choosing which points are meant to correspond. In other words, one can make use of diffeomorphisms to ensure that the surface S, event horizons and the Killing vectors t^{μ} and

ϕ^μ stay the same in two solutions [71]. Hence,

$$\delta t^\mu = \delta\phi^\mu = 0\,, \tag{E.17}$$

$$\delta g_{\mu\nu} \equiv \gamma_{\mu\nu} = -g_{\mu\rho}g_{\nu\tau}\delta g^{\rho\tau}\,, \tag{E.18}$$

$$\delta t_\mu = \gamma_{\mu\nu}t^\nu\,, \quad \delta\phi_\mu = \gamma_{\mu\nu}\phi^\mu\,, \tag{E.19}$$

$$\delta\chi^\mu = \delta\Omega_H\phi^\mu\,, \tag{E.20}$$

$$\delta\chi_\mu = \gamma_{\mu\nu}\chi^\nu + \delta\Omega_H\phi_\mu\,. \tag{E.21}$$

Moreover, as event horizons stay the same, Killing vector of the first solution must be parallel with the Killing vector of the new solution. Consequently, the Lie derivative of $\delta\chi_\mu$ with respect to χ vanishes,

$$\mathfrak{L}_\chi\delta\chi_\mu = \chi^\nu\nabla_\nu\delta\chi_\mu + \delta\chi_\nu\nabla_\mu\chi^\nu = 0\,, \tag{E.22}$$

as $\delta\chi_\mu \propto \chi_\mu$, and we used Killing equation. To evaluate δM we first express E.14 using Einstein's equation as

$$M = \int_S \left(\frac{1}{8\pi}R\delta^\mu_\nu + 2T^\mu_\nu\right)t^\nu d\Sigma_\mu + 2\Omega_H J_H + \frac{\kappa}{4\pi}A\,. \tag{E.23}$$

The procedure is to vary each term. Starting with the first,

$$\begin{aligned}\frac{1}{8\pi}\delta\left(\int_S Rt^\mu d\Sigma_\mu\right) &= \frac{1}{8\pi}\delta\left(\int_S Rt^\mu n_\mu\sqrt{h}d^3y\right)\\ &= \frac{1}{8\pi}\int_S \delta\left(R_{\rho\tau}g^{\rho\tau}\sqrt{h}\right)t^\mu n_\mu d^3y\\ &= \frac{1}{8\pi}\int_S \left(g^{\rho\tau}\sqrt{h}\delta R_{\rho\tau} + R_{\rho\tau}\sqrt{h}\delta g^{\rho\tau} + R_{\rho\tau}g^{\rho\tau}\delta\sqrt{h}\right)t^\mu n_\mu d^3y\\ &= \frac{1}{8\pi}\int_S \left(g^{\rho\tau}\sqrt{h}\delta R_{\rho\tau} - \gamma^{\rho\tau}R_{\rho\tau}\sqrt{h} + R\frac{1}{2}g_{\alpha\beta}\gamma^{\alpha\beta}\sqrt{h}\right)t^\mu n_\mu d^3y\\ &= -\frac{1}{8\pi}\int_S \left(R_{\rho\tau} - \frac{1}{2}Rg_{\rho\tau}\right)h^{\rho\tau}t^\mu d\Sigma_\mu + \frac{1}{8\pi}\int_S g^{\rho\tau}\delta R_{\rho\tau}t^\mu d\Sigma_\mu\,.\end{aligned} \tag{E.24}$$

As S is a spacelike hypersurface, in the first line the surface element was expended as $d\Sigma_\mu = n_\mu\sqrt{h}d^3y$, where n_μ is normal to the hypersurface, and $h_{\mu\nu} = g_{\mu\nu} - n_\mu n_\nu$ is projection of metric on hypersurface S, so that $\sqrt{h}d^3y$ is invariant three-dimensional surface element (or volume element, depending how you look at it). When varying the surface element we made use of the fact that both S and t^μ do not change under variation. It is now left to calculate the variation of Ricci tensor. We start from Palatini identity,

$$\delta R_{\mu\nu} = \nabla_\rho \delta\Gamma^\rho_{\mu\nu} - \nabla_\nu \delta\Gamma^\rho_{\rho\mu} \,. \tag{E.25}$$

The variation of Christoffel symbol defined in 4.137 is

$$\begin{aligned}\delta\Gamma^\rho_{\mu\nu} &= g^{\rho\tau}\frac{1}{2}\left(\partial_\mu \delta g_{\nu\tau} + \partial_\nu \delta g_{\mu\tau} - \partial_\tau \delta g_{\mu\nu}\right) + \delta g^{\rho\tau}\frac{1}{2}\left(\partial_\mu g_{\nu\tau} + \partial_\nu g_{\mu\tau} - \partial_\tau g_{\mu\nu}\right) \\ &= g^{\rho\tau}\frac{1}{2}\left(\nabla_\mu \delta g_{\nu\tau} + \nabla_\nu \delta g_{\mu\tau} - \nabla_\tau \delta g_{\mu\nu}\right) \\ &= \frac{1}{2}\left(\nabla_\mu \gamma^\rho_\nu + \nabla_\nu \gamma^\rho_\mu - \nabla^\rho \gamma_{\mu\nu}\right) \,.\end{aligned} \tag{E.26}$$

The second line follows because variation of Christoffel symbol must be a tensor. In other words, since the variation is difference between two infinitesimally different connections with the same indices, the non-tensorial parts will rule each other out. As a result, we obtain

$$\begin{aligned}g^{\mu\nu}\delta R_{\mu\nu} &= \frac{1}{2}g^{\mu\nu}\nabla_\rho\left(\nabla_\mu \gamma^\rho_\nu + \nabla_\nu \gamma^\rho_\mu - \nabla^\rho \gamma_{\mu\nu}\right) - \frac{1}{2}g^{\mu\nu}\nabla_\nu\left(\nabla_\rho \gamma^\rho_\mu + \nabla_\mu \gamma^\rho_\rho - \nabla^\rho \gamma_{\rho\mu}\right) \\ &= \frac{1}{2}\Big(2\nabla_\rho \nabla_\mu \gamma^{\rho\mu} - 2\nabla_\rho \nabla^\rho \gamma^\mu_\mu\Big) \\ &= -2\nabla^\rho \nabla_{[\rho}\gamma^\mu_{\mu]} \,.\end{aligned} \tag{E.27}$$

Finally, putting everything together, the variation of the first term of E.23 is

$$\frac{1}{8\pi}\delta\left(\int_S R t^\mu d\Sigma_\mu\right) = -\frac{1}{8\pi}\int_S \left(R_{\rho\tau} - \frac{1}{2}R g_{\rho\tau}\right) h^{\rho\tau} t^\mu d\Sigma_\mu - \frac{1}{4\pi}\int_S \nabla^\rho \nabla_{[\rho}\gamma^\nu_{\nu]} t^\mu d\Sigma_\mu \,. \tag{E.28}$$

We can further rearrange the last term to be in the form of boundary integral. Starting from

$$
\begin{aligned}
\nabla^\rho \nabla_{[\rho}\gamma^\nu_{\nu]}t^\mu &= \nabla^\rho(\nabla_{[\rho}\gamma^\nu_{\nu]}t^\mu) - \nabla_{[\rho}\gamma^\nu_{\nu]}\nabla^\rho t^\mu \\
&= \nabla^\rho(\nabla_{[\rho}\gamma^\nu_{\nu]}t^\mu) + \nabla_{[\rho}\gamma^\nu_{\nu]}\nabla^\mu t^\rho \\
&= \nabla^\rho(\nabla_{[\rho}\gamma^\nu_{\nu]}t^\mu) + \underbrace{\nabla^\mu(\nabla_{[\rho}\gamma^\nu_{\nu]}t^\rho)}_{(1)=0} - (\nabla^\mu\nabla_{[\rho}\gamma^\nu_{\nu]})t^\rho \\
&= \nabla^\rho(\nabla_{[\rho}\gamma^\nu_{\nu]}t^\mu) - \frac{1}{2}\Big(\nabla_\rho(\nabla^\mu\gamma^\nu_\nu t^\rho) - \underbrace{\nabla^\mu\gamma^\nu_\nu\nabla_\rho t^\rho}_{(2)=0} - \underbrace{\nabla^\mu\nabla_\nu\gamma^\nu_\rho t^\rho}_{(3)}\Big)\,.
\end{aligned} \tag{E.29}
$$

The first term has the “boundary" form. The last line is obtained with the help of product rule and Killing equation. The term (2) vanishes because of Killing's equation. To show the same is true for (1) we will use the identities

$$
\nabla_\rho\gamma_{\mu\nu}t^\rho + \gamma_{\rho\mu}\nabla_\nu t^\rho + \gamma_{\rho\nu}\nabla_\mu t^\rho = 0\,, \tag{E.30}
$$

$$
\nabla_\rho\gamma^\nu_\nu t^\rho = 0\,. \tag{E.31}
$$

The first relation is the variation of Lie derivative of the metric in direction of t^μ, $\mathfrak{L}_t g = \nabla_\rho g_{\mu\nu}t^\rho + g_{\rho\mu}\nabla_\nu t^\rho + g_{\rho\nu}\nabla_\mu t^\rho = 0$. The second is obtained by taking the trace of E.30 and using the fact that variation of the metric is symmetric, while tensor $\nabla_\mu t_\nu$ is antisymmetric. Thus, it follows that

$$
\begin{aligned}
(1) = \nabla^\mu(\nabla_{[\rho}\gamma^\nu_{\nu]}t^\rho) &= \frac{1}{2}\nabla^\mu\left(\nabla_\rho\gamma^\nu_\nu t^\rho - \nabla_\nu\gamma^\nu_\rho t^\rho\right) \\
&= -\frac{1}{2}\nabla^\mu\left(\nabla_\nu(\gamma^\nu_\rho t^\rho) - \gamma^\nu_\rho\nabla_\nu t^\rho\right) \\
&= 0\,.
\end{aligned} \tag{E.32}
$$

The first term vanishes as a consequence of E.31. The second line follows from the product rule. Finally, the first term in the second line is zero because $\delta t^\nu = \gamma^\nu_\rho t^\rho = 0$, and the second because $\gamma_{\nu\rho}$ is symmetric and multiplied by an antisymmetric tensor.

Let's now further expand the term (3) in E.29

$$
\begin{aligned}
(3) = \nabla^\mu \nabla_\nu \gamma^\nu_\rho t^\rho &= \nabla^\mu(\nabla_\nu \gamma^\nu_\rho t^\rho) - \nabla_\nu \gamma^\nu_\rho \nabla^\mu t^\rho \\
&= \nabla^\mu \nabla_\nu(\gamma^\nu_\rho t^\rho) - \nabla^\mu(\gamma^\nu_\rho \nabla_\nu t^\rho) - \nabla_\nu(\gamma^\nu_\rho \nabla^\mu t^\rho) + \gamma^\nu_\rho \nabla_\nu \nabla^\mu t^\rho \\
&= -\nabla^\nu(\gamma_{\rho\nu} \nabla^\mu t^\rho) + \gamma_{\nu\rho} R^{\nu\mu\rho}{}_\tau t^\tau \\
&= \nabla^\nu(\nabla_\rho \gamma^\mu_\nu t^\rho + \gamma^\mu_\rho \nabla_\nu t^\rho) + \gamma_{\nu\rho} R^{\nu\mu\rho}{}_\tau t^\tau \\
&= \nabla^\nu \left(\nabla_\rho \gamma^\mu_\nu t^\rho - \nabla^\rho(\gamma^\mu_\rho t_\nu) + t_\nu \nabla^\rho \gamma^\mu_\rho\right) + \gamma_{\nu\rho} R^{\nu\mu\rho}{}_\tau t^\tau \\
&= \nabla^\nu(\nabla^\rho \gamma^\mu_\rho t_\nu) + \underbrace{\nabla^\nu \left(\nabla_\rho \gamma^\mu_\nu t^\rho - \nabla^\rho(\gamma^\mu_\rho t_\nu)\right) - R^{\mu\nu\rho\tau} \gamma_{\nu\rho} t_\tau}_{(4)=0} \\
&= \nabla_\rho(\nabla^\nu \gamma^\mu_\nu t^\rho)\,.
\end{aligned}
\tag{E.33}
$$

Once again, the final result is obtained by repetitive use of product rule and Killing equation. To get the third line one recognizes the definition of Riemann tensor. We also used E.31 and $\delta t^\nu = \gamma^\nu_\rho t^\rho = 0$. The fourth line follows from E.30, and the last line is obtained by rearranging the terms and using the product rule. To show that (4) is zero we need the following identities,

$$
\nabla^\rho \nabla^\nu A^\mu_\rho - \nabla^\nu \nabla^\rho A^\mu_\rho = R^\mu{}_\tau{}^{\rho\nu} A^\tau_\rho - R^\tau{}_\rho{}^{\rho\nu} A^\mu_\tau\,, \tag{E.34}
$$

$$
\nabla^\mu \nabla_\nu K_\rho = R^\tau{}_{\mu\nu\rho} K_\tau\,. \tag{E.35}
$$

One starts by expanding the derivatives,

$$
\begin{aligned}
(4) &= \nabla^\nu \nabla_\rho \gamma^\mu_\nu t^\rho + \nabla_\rho \gamma^\mu_\nu \nabla^\nu t^\rho - \nabla^\nu \nabla^\rho \gamma^\mu_\rho t_\nu - \nabla^\rho \gamma^\mu_\rho \nabla^\nu t_\nu - \nabla^\nu \gamma^\mu_\rho \nabla^\rho t_\nu - \gamma^\mu_\rho \nabla^\nu \nabla^\rho t_\nu - R^{\mu\nu\rho\tau} \gamma_{\nu\rho} t_\tau \\
&= \nabla^\nu \nabla^\rho \gamma^\mu_\nu t_\rho - \nabla^\nu \nabla^\rho \gamma^\mu_\rho t_\nu - \gamma^\mu_\rho \nabla^\nu \nabla^\rho t_\nu - R^{\mu\nu\rho\tau} \gamma_{\nu\rho} t_\tau \\
&= R^\mu{}_\tau{}^{\rho\nu} \gamma^\tau_\rho t_\nu - R^\tau{}_\rho{}^{\rho\nu} \gamma^\mu_\tau t_\nu - \gamma^\mu_\rho R^{\tau\nu\rho}{}_\nu t_\tau - R^{\mu\nu\rho\tau} \gamma_{\nu\rho} t_\tau \\
&= R^{\mu\tau\rho\nu} \gamma_{\tau\rho} t_\nu + R^{\tau\rho\nu}{}_\rho \gamma^\mu_\tau t_\nu - R^{\tau\nu\rho}{}_\nu \gamma^\mu_\rho t_\tau - R^{\mu\nu\rho\tau} \gamma_{\nu\rho} t_\tau \\
&= 0\,.
\end{aligned}
\tag{E.36}
$$

In the first line the second and the fifth term cancel each other out, and the fourth is zero because of Killing equation. The third line is obtained by renaming indices of the first term and using E.34, while the third term follows from E.35. Finally, using the symmetries of the Riemann tensor one obtains the result. Hence, E.29 can be written as

$$\begin{aligned}\nabla^\rho\nabla_{[\rho}\gamma^\nu_{\nu]}t^\mu &= \nabla^\rho(\nabla_{[\rho}\gamma^\nu_{\nu]}t^\mu) - \frac{1}{2}\Big(\nabla_\rho(\nabla^\mu\gamma^\nu_\nu t^\rho) - \nabla_\rho(\nabla^\nu\gamma^\mu_\nu t^\rho)\Big) \\ &= \nabla_\rho(\nabla^{[\rho}\gamma^{\nu]}_\nu t^\mu) - \nabla_\rho(\nabla^{[\mu}\gamma^{\nu]}_\nu t^\rho) \\ &= \nabla_\rho(\nabla^{[\rho}\gamma^{\nu]}_\nu t^\mu - \nabla^{[\mu}\gamma^{\nu]}_\nu t^\rho)\,.\end{aligned} \tag{E.37}$$

The relation is valid for any Killing vector field of stationary spacetime, as we have not used any properties of t^μ specifically. Consequently, using E.37 the last term of E.28 can be written as

$$\begin{aligned}\frac{1}{4\pi}\int_S \nabla^\rho\nabla_{[\rho}\gamma^\nu_{\nu]}t^\mu d\Sigma_\mu &= \frac{1}{4\pi}\int_S \nabla_\rho(\nabla^{[\rho}\gamma^{\nu]}_\nu t^\mu - \nabla^{[\mu}\gamma^{\nu]}_\nu t^\rho)d\Sigma_\mu \\ &= \frac{1}{4\pi}\int_{\partial S}(\nabla^{[\rho}\gamma^{\nu]}_\nu t^\mu - \nabla^{[\mu}\gamma^{\nu]}_\nu t^\rho)d\Sigma_{\mu\rho} \\ &= \delta M + \frac{1}{4\pi}\int_{\partial B} 2\nabla^{[\rho}\gamma^{\nu]}_\nu t^\mu \chi_{[\mu}N_{\rho]}dA\end{aligned} \tag{E.38}$$

,

where we transferred the volume on the left to integral over 2-hypersurface ∂S and integrated over the components of ∂S, the horizon and boundary at infinity. According to [45], the integral over S_∞ gives δM, with no other arguments given. It seems to me that this is the case simply because the integral is evaluated at infinity, but I do not know how to show this precisely. A possible justification, looking at explicit variation of Kerr solution in [111] or [148], may be that as we are infinitely far away from matter, so one is only left with variation of total energy. In the term that is integrated over horizon we expended the surface element and took advantage of the antisymmetry of indices. Continuing the derivation,

$$\begin{aligned}\frac{1}{4\pi}\int_S \nabla^\rho\nabla_{[\rho}\gamma^\nu_{\nu]}t^\mu d\Sigma_\mu &= \delta M + \frac{1}{4\pi}\int_{\partial B}(\nabla^{[\rho}\gamma^{\nu]}_\nu\chi^\mu - \Omega_H\nabla^{[\rho}\gamma^{\nu]}_\nu\phi^\mu)(\chi_\mu N_\rho - \chi_\rho N_\mu)dA \\ &= \delta M + \frac{1}{4\pi}\int_{\partial B}\nabla^{[\rho}\gamma^{\nu]}_\nu\chi_\rho dA - \underbrace{\frac{1}{4\pi}\int_{\partial B}\Omega_H\nabla^{[\rho}\gamma^{\nu]}_\nu\phi^\mu d\Sigma_{\mu\rho}}_{(5)=0} \\ &= \delta M + \frac{1}{4\pi}\int_{\partial B}(-)\frac{1}{2}\nabla^\nu\gamma^\rho_\nu\chi_\rho dA\,.\end{aligned} \tag{E.39}$$

The first line is obtained using E.11 and the second with 5.10. Finally, in the last line we used the identity E.31. To show (5) is zero we will need

$$\begin{aligned}\nabla^{\rho}\nabla_{[\rho}\gamma^{\nu}_{\nu]}\phi^{\mu}d\Sigma_{\mu} &= \nabla_{\rho}(\nabla^{[\rho}\gamma^{\nu]}_{\nu}\phi^{\mu} - \nabla^{[\mu}\gamma^{\nu]}_{\nu}\phi^{\rho})d\Sigma_{\mu} \\ &= (\nabla^{[\rho}\gamma^{\nu]}_{\nu}\phi^{\mu} - \nabla^{[\mu}\gamma^{\nu]}_{\nu}\phi^{\rho})d\Sigma_{\mu\rho} \\ &= 2\nabla^{[\rho}\gamma^{\nu]}_{\nu}\phi^{\mu}d\Sigma_{\mu\rho}\,.\end{aligned} \tag{E.40}$$

Using the obtained expression it then follows that

$$\begin{aligned}(5) = \int_{\partial B}\Omega_H\nabla^{[\rho}\gamma^{\nu]}_{\nu}\phi^{\mu}d\Sigma_{\mu\rho} &= \frac{1}{2}\int_{\partial B}\Omega_H\nabla^{\rho}\nabla_{[\rho}\gamma^{\nu}_{\nu]}\phi^{\mu}d\Sigma_{\mu} \\ &= 0\,.\end{aligned} \tag{E.41}$$

The result is zero because 2-surface ∂S is chosen so that it is tangent to ϕ, as previously explained. We now have, from E.28 and expression E.39, that E.23 is

$$\begin{aligned}\delta M = &-\frac{1}{8\pi}\int_S\left(R_{\rho\tau} - \frac{1}{2}Rg_{\rho\tau}\right)h^{\rho\tau}t^{\mu}d\Sigma_{\mu} - \delta M + \frac{1}{4\pi}\int_{\partial B}\frac{1}{2}\nabla^{\nu}\gamma^{\rho}_{\nu}\chi_{\rho}dA + \\ &2\int_S\delta(T^{\mu}_{\nu}t^{\nu}d\Sigma_{\mu}) + 2\delta(\Omega_H J_H) + \frac{1}{4\pi}\delta(\kappa A)\,.\end{aligned} \tag{E.42}$$

Next, consider variation of surface gravity $\kappa = -N^{\nu}\chi^{\mu}\nabla_{\mu}\chi_{\nu} = \frac{1}{2}N^{\nu}\nabla_{\nu}(\chi_{\mu}\chi^{\mu})$,

$$\begin{aligned}\delta\kappa &= \frac{1}{2}\delta N^{\nu}\nabla_{\nu}(\chi_{\mu}\chi^{\mu}) + \frac{1}{2}N^{\nu}\nabla_{\nu}(\delta\chi_{\mu}\chi^{\mu} + \delta\chi^{\mu}\chi_{\mu}) \\ &= \delta N^{\nu}\chi^{\mu}\nabla_{\nu}\chi_{\mu} + \frac{1}{2}N^{\nu}\Big(\nabla_{\nu}(\delta\chi_{\mu}\chi^{\mu}) + \delta\Omega_H\nabla_{\nu}(\phi_{\mu}\chi^{\mu}) - \delta\chi_{\mu}\nabla_{\nu}\chi^{\mu} + \delta\chi_{\mu}\nabla_{\nu}\chi^{\mu}\Big) \\ &= \delta N^{\nu}\chi^{\mu}\nabla_{\nu}\chi_{\mu} + \frac{1}{2}N^{\nu}\Big(\chi^{\mu}\nabla_{\nu}\delta\chi_{\mu} + \delta\Omega_H\nabla_{\nu}(\phi_{\mu}\chi^{\mu}) + \chi^{\mu}\nabla_{\mu}\delta\chi_{\nu} + 2\delta\chi_{\mu}\nabla_{\nu}\chi^{\mu}\Big) \\ &= \nabla_{\nu}\chi_{\mu}(\delta N^{\nu}\chi^{\mu} + N^{\nu}\delta\chi^{\mu}) + \frac{1}{2}(\nabla_{\nu}\delta\chi_{\mu})(N^{\nu}\chi^{\mu} + N^{\mu}\chi^{\nu}) + \frac{1}{2}N^{\nu}\delta\Omega_H\nabla_{\nu}(\phi_{\mu}\chi^{\mu})\end{aligned} \tag{E.43}$$

The second line follows from E.20 and the derivative product rule. To obtain the third line we used E.22 and expended the derivative of the first term in curly brackets. The fourth line is a result of re-combining the terms. The first term vanishes because the bracket is symmetric, but χ satisfies Killing equation which is antisymmetric. In the brackets of the second term we recognize the induced metric $g^{\mu\nu} = -\chi^{\mu}N^{\nu} - \chi^{\nu}N^{\mu} + h^{\mu\nu}$. However, as $\delta\chi_{\mu} \propto \chi_{\mu}$ on the horizon $(\nabla_{\nu}\delta\chi_{\mu})h^{\mu\nu} = 0$. We can use E.44 to rewrite the last term,

$$
\begin{aligned}
\delta\Omega_H \nabla_\nu(\phi_\mu \chi^\mu) &= \nabla_\nu(\delta\chi_\mu\chi^\mu) \\
&= \chi^\mu\nabla_\nu\delta\chi_\mu + \delta\chi_\mu\nabla_\nu\chi^\mu \\
&= \chi^\mu\nabla_\nu\delta\chi_\mu - \chi^\mu\nabla_\mu\delta\chi_\nu \\
&= \chi^\mu\nabla_\nu\delta\chi_\mu + \chi^\mu\nabla_\nu\delta\chi_\mu \\
&= 2\delta\Omega_H\chi^\mu\nabla_\nu\phi_\mu \, ,
\end{aligned}
\tag{E.44}
$$

which is a consequence of the vanishing of E.22. Consequently,

$$
\begin{aligned}
\delta\kappa &= \frac{1}{2}(\nabla_\nu\delta\chi_\mu)(-g^{\mu\nu}) + N^\nu\delta\Omega_H\chi^\mu\nabla_\nu\phi_\mu \\
&= -\frac{1}{2}\nabla^\mu\delta(g_{\mu\rho}\chi^\rho) + N_\nu\delta\Omega_H\chi_\mu\nabla^\nu\phi^\mu \\
&= -\frac{1}{2}\nabla^\mu\gamma_{\mu\rho}\chi^\rho - \frac{1}{2}\nabla_\rho\delta\chi^\rho + \delta\Omega_H N_{[\nu}\chi_{\mu]}\nabla^\nu\phi^\mu \\
&= -\frac{1}{2}\nabla^\nu\gamma^\rho_\nu\chi_\rho + \delta\Omega_H N_{[\nu}\chi_{\mu]}\nabla^\nu\phi^\mu \, .
\end{aligned}
\tag{E.45}
$$

Metric can be put under variation because the additional term from the product rule vanishes as a result of antisymmetry of Killing equation. The second term in the third line is zero for the same reason. Moreover, since $\nabla^\nu\phi^\mu$ is already antisymmetric in indices, adding the antisymmetry brackets doesn't change anything. Finally, integrating E.45 over the horizon and recognizing E.10, leads to

$$
\frac{1}{4\pi}\int_{\partial B}(-)\frac{1}{2}\nabla^\nu\gamma^\rho_\nu\chi_\rho = \frac{\delta\kappa}{4\pi}A + 2\delta\Omega_H J_H \, . \tag{E.46}
$$

Thus, inserting the obtained expression into E.42 results in

$$
\begin{aligned}
\delta M = -\frac{1}{8\pi}\int_S\left(R_{\rho\tau} - \frac{1}{2}Rg_{\rho\tau}\right)h^{\rho\tau}t^\mu d\Sigma_\mu - \delta M - \frac{\delta\kappa}{4\pi}A - 2\delta\Omega_H J_H + \\
2\int_S\delta(T^\mu_\nu t^\nu d\Sigma_\mu) + 2\delta(\Omega_H J_H) + \frac{1}{4\pi}\delta(\kappa A) \\
2\delta M = -\int_S T_{\rho\tau}h^{\rho\tau}t^\mu d\Sigma_\mu + \frac{\kappa}{4\pi}\delta A + 2\Omega_H\delta J_H + 2\int_S\delta(T^\mu_\nu t^\nu d\Sigma_\mu) \, .
\end{aligned}
\tag{E.47}
$$

To vary the EMT we need to specify with what kind of matter we are dealing with. Consider the case of perfect fluid described by energy density $\rho = \rho(n, s)$ which is a function of number density n and energy density s. The temperature θ, chemical potential μ and pressure p are defined by

$$\theta = \frac{\partial \rho}{\partial s} \,, \tag{E.48}$$

$$\mu = \frac{\partial \rho}{\partial n} \,, \tag{E.49}$$

$$p(n, s) = \mu n + \theta s - \rho \,. \tag{E.50}$$

Hence, EMT is given by

$$T_{\mu\nu} = (\rho + p) v_\mu v_\nu + p g_{\mu\nu} \,, \tag{E.51}$$

where v is defined as

$$v^\mu = \frac{u^\mu}{\sqrt{-g_{\rho\tau} u^\rho u^\tau}} \,, \quad u^\mu = t^\mu + \Omega \phi^\mu \,. \tag{E.52}$$

Note that EMT is given with respect to an observer moving with the fluid. Moreover, the matter outside the black hole is in circular orbit outside the black hole, so that the spacetime is axisymmetric. We will also use the expressions

$$dJ = -T^\mu_\nu \phi^\mu d\Sigma_\nu \,, \tag{E.53}$$

$$dS = \frac{s}{\sqrt{-u_\nu u^\nu}} t^\mu d\Sigma_\mu \,, \tag{E.54}$$

$$dN = \frac{n}{\sqrt{-u_\nu u^\nu}} t^\mu d\Sigma_\mu \,, \tag{E.55}$$

describing the change in angular momentum, entropy, and number of particles of the fluid corssing the surface element $d\Sigma_\mu$ respectively. Moreover, from the previous definitions we see that

$$\delta p = n \delta\mu + s \delta\theta \tag{E.56}$$

$$\delta u^\mu = 0 \,, \tag{E.57}$$

$$u^\mu \delta\Big(\frac{u_\mu}{-g_{\rho\tau} u^\rho u^\tau}\Big) = \frac{1}{2} v^\alpha v^\beta h_{\alpha\beta} \,. \tag{E.58}$$

The variation of the last term in E.47 is then

$$
\begin{aligned}
\int_S \delta(T^\mu_\nu t^\nu d\Sigma_\mu) &= \int_S \delta(T^\mu_\nu u^\nu d\Sigma_\mu) - \int_S \delta(T^\mu_\nu \phi^\nu d\Sigma_\mu) \\
&= \int_S \delta\left[\left((\rho+p)\frac{u^\mu u_\nu u^\nu}{-g_{\rho\tau}u^\rho u^\tau} + pu^\mu\right) d\Sigma_\mu\right] + \int_S \Omega\delta dJ \\
&= \int_S \delta(pt^\mu d\Sigma_\mu) - \int_S \frac{u_\nu u^\nu}{u^\rho u_\rho}\delta\Big[(\mu n + \theta s)t^\mu d\Sigma_\mu\Big] + \\
&\quad \int_S (\mu n + \theta s)u^\nu \delta\Big[\frac{u_\nu}{-g_{\rho\tau}u^\rho u^\tau}\Big] t^\mu d\Sigma_\mu + \int_S \Omega\delta dJ \\
&= \frac{1}{2}\int_S p g^{\alpha\beta} h_{\alpha\beta} t^\mu d\Sigma_\mu + \int_S \delta p t^\mu d\Sigma_\mu - \int_S \delta\Big[(\mu n + \theta s)t^\mu d\Sigma_\mu\Big] + \\
&\quad \frac{1}{2}\int_S (\mu n + \theta s) v^\alpha v^\beta h_{\alpha\beta} t^\mu d\Sigma_\mu + \int_S \Omega\delta dJ \ .
\end{aligned}
\tag{E.59}
$$

The first line follows from definition of u, E.52. The second line uses the definition of EMT given in E.51 and also E.53. We further use $\phi^\mu d\Sigma_\mu = 0$. The variations of surface elements were derived before, and we make use of E.58, hence the last line. Further, we defined the “red-shifted" chemical potential and temperature respectively,

$$
\bar{\mu} = \sqrt{-u_\mu u^\mu}\mu \ , \quad \bar{\theta} = \sqrt{-u_\mu u^\mu}\theta \ . \tag{E.60}
$$

This is because, as mentioned, quantities in EMT are defined with respect to ab observer which is moving together with the fluid, while everything else is formulated with respect to an observer at infinity. To restate, we have introduced the red-shift factor to “convert" the quantities of EMT from the one an observer which moves with the fluid measures, to the one measured by an observer at infinity, so that everything is consistent. Thus, the variation E.59 is

$$
\begin{aligned}
\int_S \delta(T^\mu_\nu t^\nu d\Sigma_\mu) &= \frac{1}{2}\int_S p g^{\alpha\beta} h_{\alpha\beta} t^\mu d\Sigma_\mu + \int_S \bar{\mu}\delta dN + \int_S \bar{\theta}\delta dS + & \text{(E.61)} \\
&\quad \frac{1}{2}\int_S T^{\alpha\beta} h_{\alpha\beta} t^\mu d\Sigma_\mu - \frac{1}{2}\int_S p g^{\alpha\beta} h_{\alpha\beta} t^\mu d\Sigma_\mu + \int_S \Omega\delta dJ & \text{(E.62)} \\
&= \int_S \bar{\mu}\delta dN + \int_S \bar{\theta}\delta dS + \frac{1}{2}\int_S T^{\alpha\beta} h_{\alpha\beta} t^\mu d\Sigma_\mu + \int_S \Omega\delta dJ & \text{(E.63)}
\end{aligned}
$$

The second line follows from definition of EMT and previously defined definitions. Inserting the obtained expression into E.47 results in

$$
\delta M = \int_S \bar{\mu}\delta dN + \int_S \bar{\theta}\delta dS + \int_S \Omega\delta dJ + \Omega_H \delta J_H + \frac{\kappa}{8\pi}\delta A \ . \tag{E.64}
$$

This is the differential mass formula. The last two terms describe the change in black hole energy, while the rest of the terms describe the perfect fluid outside the black hole as measured from infinity. In vacuum, the first law is

$$\delta M = \frac{\kappa}{8\pi}\delta A + \Omega_H \delta J_H \ . \tag{E.65}$$

The mass formula expresses how parameters of two nearby stationary solutions are related.

7 Prošireni sažetak

7.1 Uvod

Izranjajuća gravitacija polazi od ideje da je opća teorija relativnosti efektivna teorija, a gravitacija fenomen koji proizlazi iz kolektivnog gibanja mikroskopskih stupnjeva slobode neke trenutno nepoznate teorije. Razlog zašto uopće postoji potreba za novim pristupom jest to što, iako je opća teorija relativnosti vrlo uspješna u reproduciranju mnogih opažanja, postoje problemi s teorijske i eksperimentalne strane koji indiciraju da opća teorija relativnosti nije cijela priča.

Neka svojstva gravitacije koje govore u prilog izranjajućoj paradigmi su zakoni mehanike crnih rupa koji su analogni zakonima termodinamike, isključivo privlačna priroda gravitacije i perturbativna ne-renormalizabilnost. S druge strane, osim tih indicija, postoje modeli, uglavnom temeljeni na fizici kondenzirane tvari, koji uspješno reproduciraju neke od svojstava gravitacije.

U ovom radu dan je pregled različitih pristupa paradigmi izranjajuće gravitacije bez zalaženja u detalje različitih modela. Fokus je na određenim rezultatima i teorijama koje ilustriraju glavne ideje i smjerove ovog pristupa. Konkretno, kratko razmatramo kauzalnu teoriju skupova kao primjer teorije koja polazi od toga da je prostorvrijeme diskretno. Nadalje, objašnjena je uloga Weinberg-Wittenovog teorema koji ograničava vrste čestica koje mogu postojati na ravnom prostorvremenu. Konačno, izvedeni su zakoni termodinamike crnih rupa i diskutirana je njihova uloga u uspostavljanju veze između termodinamike i gravitacije. Osim toga, pokazano je da se Einsteinova jednadžba može dobiti iz veze geometrije i termodinamike za kauzalne horizonte.

7.2 Izranjajuća priroda gravitacije

U ovom poglavlju cilj je objasniti što podrazumijevamo pod pojmom “izranjajuća gravitacija", stavljajući u kontekst teme koje ćemo obraditi u narednim poglavljima. Diskusija prati [1] i [2]. Najprije ćemo kratko ponoviti temelje opće teorije relativnosti. Spomenut ćemo neke nedostatke teorije koji postoje s teorijske i eksperimentalne strane i upućuju na to da se teorija treba "nadograditi". Jedan od pristupa dan je upravo teorijama koje možemo grupirati pod nazivom "izranjajuća gravitacija".

Spomenut ćemo svojstva gravitacije koja indiciraju da je to slučaj i dati pregled različitih modela koji pripadaju toj paradigmi.

Opća teorija relativnosti je opis prostorvremena, 4-dimenzionalne mnogostrukosti s Lorentzovom metrikom koja opisuje gravitacijsko polje [3]. Objasnit ćemo koja opažanja vode na opis gravitacije kao metričke teorije. Taj pristup prezentiran je u [4] i naziva se EPS formulacija.

Fundamentalni objekti EPS formulacije su svjetlost i tijelo u slobodnom padu, koji se tretiraju klasično. Svjetlost kao valni paket elektromagnetskog zračenja, dok se tijelo odnosi na bilo koje tijelo zanemarivih dimenzija i strukture. U tekstu ćemo se na takvo tijelo u slobodnom padu referirati kao na česticu. Formulacija kreće od sljedećih aksioma,

- Postoji skup događaja $\mathcal{M}$ zajedno sa skupom svjetskih krivulja zraka svjetlosti L i čestica P.

- Svjetske linije definiraju topologiju na $\mathcal{M}$. Trajektorije svjetlosti i čestica su kontinuirane. Nadalje, čestice mogu međusobno slati svjetlosne signale što možemo iskoristiti da svakom događaju pridijelimo koordinate.

- Svjetlosne zrake određuju kauzalnu strukturu, tj. definiraju svjetlosne stošce u svakoj točki skupa $\mathcal{M}$. Kaže se da takva mnogostrukost ima konformalnu strukturu. Posljedično, možemo govoriti o vektorima i krivuljama vremenskog, svjetlosnog ili prostornog tipa. Štoviše, može se pokazati da su krivulje svjetlosnog tipa geodezici.

- Čestice se gibaju po vremenolikim krivuljama. Postoji poseban tip čestica koje ne nose naboj nijedne druge sile osim gravitacije. Takve čestice gibaju se po autoparaleli i definiraju projektivnu strukturu koneksija na mnogostrukosti.

- Konfromalna i projektivna struktura su kompatibilne u smislu da svjetlosni geodezici moraju biti autoparalele koneksije. Takav prostor zove se Weylov prostor.

- Kako bi “brzina" vremena bila neovisna o putanji, u smislu da vremenski interval ovisi samo o putanji kojom se gibamo, nameće se uvjet da je iznos vektora koji je paralelno transportiran neovisan o putanji. Iz tog uvjeta slijedi da koneksija mora biti tzv. Levi-Civita koneksija i Weylov prostor svodi se na Lorentzovo prostorvrijeme.

Ovime smo odredili kinematiku teorije relativnosti. Dinamika je određena Lagrangianom ili Hamiltonijanom. U tu svrhu, postuliramo da je teorija invarijantna na difeomorfizme i zahtijevamo da dinamika ovisi samo o metriki. Tada je Lagrangian (u najnižem redu) dan Riccijevim skalarom, a jednadžbe polja dane su Einsteinovom jednadžbom.

Time smo došli do kraja diskusije o formulaciji opće teorije relativnosti. Iako teorija uspješno reproducira Keplerove zakone i mnoga druga opažanja, postoje neke nedosljednosti. Kako bi teorija bila u skladu s opaženim ubrzanim širenjem svemira mora postojati velika količina “tamne" materije. Nadalje, postoje problemi sa standardnim kozmološkim modelom baziranim na općoj teoriji relativnosti i standardnom modelu. Konkretno, problem ravne geometrije svemira, tzv. problem horizonta, te problem magnetskog monopola.

Kako bi teorija bolje adresirala navedene probleme predložena su mnoga proširenja opće teorije relativnosti, npr. $f(R)$ teorije, ne-minimalna vezanja u Lagrangianu, zamjena metrike nekakvom drugačijom tenzorskom strukturom, itd. S druge strane, izranjajuća gravitacija gleda na opću teoriju relativnosti kao na makroskopski limes neke fundamentalnije teorije.

Izranjajući fenomeni Izranjajuće fenomene može se sumirati kao "više je drugačije". Opis sustava sastavljenoga od velikog broja čestica nije dan samo ekstrapolacijom fundamentalnih zakona dobivenih promatranjem malog broja čestica, već se na drugačijoj skali javljaju novi fenomeni [19]. Primjer je pojava frakcionog kvantnog Hallovog efekta, gdje se elektroni u $2D$ kristalu ponašaju kao čestice frakcionog naboja. Ovakav fenomen je u potpunosti neočekivan. Još jedan primjer su deformacije u kristalu opisane Navierovom jednadžbom, dok je ponašanje pojedinog atoma dano Schrödingerovom jednadžbom. Dakle, na različitim skalama imamo neočekivane fenomene, opisane novim zakonima. Navedeni primjeri obuhvaćaju srž ideje izranjajućih fenomena. Formalna definicija je stoga sljedeća.

Mikroteorija M_1 teorije T je tipa I ako i samo ako je M_1 inspirirana teorijom T, npr.

diskretizacijom, kvantizacijom ili renormalizacijom.

Mikroteorija M_2 teorije T je tipa II ako i samo ako M_2 nije motivirana ili inspirirana teorijom T.

Teorija je izranjajuća ako i samo ako njezina mikroteorija pripada tipu II.

Daljni cilj bit će objasniti kad je mikroteorija inspirirana makroskopskom teorijom. Nerijetko je slučaj da mikroteorija ima svojstva tipa I i tipa II, stoga podjela nije oštra.

7.2.1 Tipovi usrednjavanja (coarse-graining)

U ovom poglavlju razmatramo vezu između mikroskopskog i makroskopskog opisa, koja je u principu dana nekakvom metodom usrednjavanja. Obrnut smjer dan je diskretizacijom u širem smislu riječi.

- Postupak diskretizacije u užem smislu sastoji se od toga da neku kontinuiranu veličinu "podijelimo" na manje dijelove. Primjer je veza ukupne mase i funkcije koja opisuje raspodjelu mase u tijelu, dobivene tako da podijelimo tijelo čiju masu opisujemo na male komadiće konstantne gustoće. U ovom primjeru komadić ima ista svojstva kao masa ukupnog tijela. Diskretizacija nije dovela do nekog novog fenomena pa u većini slučajeva diskretizacija vodi do mikroteorije tipa I. Metoda diskretizacije bitna je kod kauzalne teorije skupova koju ćemo kasnije razmatrati.

- Kvantizacija u pravilu odgovara teorijama tipa I. Iako je ponašanje kvantiziranog sustava drugačije od klasičnog, takva promjena je očekivana. Na primjer, kvantizacija vibracijskih modova kristalne rešetke daje fonone, no njih ne smatramo izranjajućim fenomenom. Drugim riječima, fononi nisu fundamentalni stupnjevi slobode, već atomi. Posljedično, teorije linealizirane gravitacije ne smatramo izranjajućima.

- Renormalizacija je postupak kojim dobivamo efektivnu teoriju. Ona daje dobar opis samo na određenim energetskim skalama jer su visoko-energetski modovi izintegrirani. Prema [1] renormalizacija spada u tip I, no [32] daje protuprimjer.

Možemo zaključiti će diskretizacija i kvantizacija u pravilu voditi to mikroteorije tipa I, dok kod renormalizacije nije sasvim jasno.

7.2.2 Indicije emergentnosti

U ovom poglavlju razmatramo svojstva gravitacije koja upućuju na to da se radi o izranjajućoj pojavi.

- Univerzalnost — gravitacija je uvijek privlačna što podsjeća na situaciju iz molekularne fizike. Neutralne molekule privlače se Londonovim silama koje uzrokuje fluktuacija naboja u molekuli što inducira dipole. Analogno, privlačna priroda gravitacije mogla bi potjecati od fluktuacija neke nepoznate veličine.
- Perturbativna ne-renormalizabilnost gravitacije upućuje na to da interakcija nije fundamentalna zbog slične situacije s Fermijevom teorijom koja opisuje beta raspad i nije renormalizabilna. Fermijeva teorija zamijenjena je slabom interakcijom koja jest.
- Termodinamika crnih rupa opisuje mehaniku crnih rupa zakonima koji imaju isti oblik kao zakoni termodinamike. Osim toga, karakteristike crnih rupa slične su termodinamičkim sustavima — opisane su malim brojem makroskopskih parametara i zrače kao crna tijela temperature Hawkingovog zračenja. Implikacije ove analogije razmatrane su u kasnijim poglavljima.

7.2.3 Pregled izranjajućih modela

Teorije izranjajuće gravitacije mogu se ugrubo podijeliti u dvije skupine. U jednoj su teorije gdje fundamentalni stupnjevi slobode žive u nekakvom mediju (tip I). U drugoj skupini modela i samo prostorvrijeme je izranjajuće (tip II). Postoji i treća skupina, koja se uglavnom ne ubraja u emergentne teorije, gdje je dinamika prostorvremena izranjauća, a metrika i mnogostrukost su zadani (tip III).

- tip I
  - Analogni modeli [48][49], kao što ime kaže, pokušavaju reproducirati svojstva gravitacije na primjerima iz čvrstog stanja, npr. valovi zvuka u fluidu ponašaju se kao svjetlost u zakrivljenom prostorvremenu.

- Modeli inspiriranim fizikom kondenzirane tvari, npr. ekscitacije blizu Fermijeve točke [50], rubne escitacije kvantnog Hallovog efekta [52], defekti u kristalu [53].

- tip II
  - Modeli koji se temelje na teoriji grafova [56] — prostorvrijeme je diskretno u smislu da se sastoji od vrhova grafa koji su spojeni bridovima. Ovdje spadaju npr. kauzalna teorija skupova i Loop Quantum Gravity (kvantna teorija petlji).
  - Kvantna gravitacija u formulaciji grupne teorije polja (group field theory) [59] koristi sve koncepte teorije polja. Razlika je u tome što ne ovisi o pozadini.
  - AdS/CFT korespondencija [60] u verziji gdje se CFT smatra primarnim, a prostorvrijeme emergentnim.

- tip III
  - Inducirana gravitacija [36] — dobije se član proporcionalan Einstein-Hilbertovoj akciji ekstremiziranjem efektivne akcije.
  - Modeli koji se baziraju na vezi termodinamike i gravitacije [66][65].

Kako bi modeli izranjajuće gravitacije reproducirali opću teoriju relativnosti trebaju zadovoljavati sljedeće uvjete.

- Lokalna Lorentz invarijantnost (LLI) eksperimentalno je potvrđena da vrijedi barem do Planckove skale. Ona je u modelima tipa I narušena jer postojanje medija u kojem fundamentalni stupnjevi slobode žive narušava princip relativnosti, tj. postoji preferirani referentni sustav. Modeli tipa II stoga nemaju taj problem.

- Univerzalno vezanje zahtijeva da je konstanta vezanja jednako jaka za sve tipove materije. Prema Weinbergovom teoremu za gravitone niskih energija (soft graviton theorem), to je slučaj kada postoji jedinstvena metrika i vrijedi LLI. Modeli tipa I često sadrže više od jedne metrike. Modeli tipa II u većini slučajeva nisu još u mogućnosti opisati vezanje materije i gravitacije.

- Invarijantnost na difeomorfizme, tj. zahtjev da ne postoji pozadina koja nije dinamička. Kod modela tipa I to se može postići ako veličine koje možemo opažati nisu vezane ("coupled") s pozadinom. Većina modela tipa II još nije u mogućnosti proizvesti glatku mnogostrukost u kontekstu koje bi imalo smisla razmatrati difeomorfizme.

Osim uvjeta povezanih s pitanjem reproduciranja opće teorije relativnosti, postoje druge vrste ograničenja na modele. Jedan takav primjer je Weinberg-Wittenov teorem koji ograničava vrste čestica koje mogu postojati u ravnom prostorvremenu.

7.3 Kauzalna teorija skupova

Kauzalna teorija skupova (CST)[25][73] razmatra ideju da se mikrostruktura prostorvremena sastoji od točaka u kauzalnom odnosu, tzv. kauzeta. Trenutno ne postoji eksperimentalni dokaz kojim bismo utvrdili da je prostorvrijeme diskretno, jer se takvo ponašanje očekuje tek na Planckovoj skali. Međutim, postoje naznake u teoriji koje govore tome u prilog. Konkretno, to što nailazimo na divergencije i singularnosti u općoj teoriji relativnosti i kvantnoj teoriji polja, što obično implicira da smo izašli iz režima u kojem teorija vrijedi.

Motivacija iza kauzalne teorije skupova je sljedeća. Ako prihvatimo da je prostorvrijeme diskretno, pitanje je kakva struktura može reproducirati topologiju, diferencijalnu strukturu i metriku koje opažamo na velikim skalama. Prema CST, to su kauzalni odnosi. Za razliku od EPS formulacije gdje je metrika bila primarna i odredila kauzalnu strukturu, CST uzima kauzalne odnose kao fundamentalne i izvodi ostale matematičke strukture iz toga. Formalnim jezikom, CST kreće od drugačije topološke strukture, tzv. Alexandrove topologije, definirana kao najmanja topologija u kojoj su kronološka budućnost i prošlost nekog skupa otvorene. Nadalje, svjetlosne stošce možemo definirati samo pomoću kauzalnih odnosa između događaja što implicira da možemo dobiti metriku (do na konformalnu transformaciju).

Kauzalni skup $(C, \prec)$ definiran je kao konačni parcijalno uređen skup. Drugim riječima, skup C s binarnom relacijom $\prec$ "prethodi" zadovoljava sljedeća svojstva,

1. Tranzitivnost — ako $x \prec y$ i $y \prec z$, $\forall x, y, z \in C$

2. Ne-cirkularnost — ako $x \prec y$ i $y \prec x$ onda $x = y \in C$

3. Konačnost — $(\forall x, z \in C)$ $(\text{card}\{y \in C | x \prec y \prec z\} < \infty)$

gdje “card" označuje kardinalnost skupa. Kao što je objašnjeno, kauzalna struktura određuje gotovo svu geometrijsku strukturu, osim konformalnog faktora. Kako bismo ga fiksirali postuliramo da konačan volumen prostorvremena sadrži velik, ali konačan broj elemenata tako da je $\sqrt{-g}d^4x$ mjera volumena. Drugim riječima, broj elemenata nekog dijela prostovremena odgovara volumenu tog prostorvremena, $N = V$.

Nadalje, kako bi kauzet reproducirao neko prostorvrijeme koje je Lorentz invarijantno točke tog kauzeta moraju biti raspodijeljene prema Poissonovoj raspodjeli. Tako neće postojati preferirani smjer u prostoru. Taj proces zove se “prskanje" (“sprinkling"). Treba naglasiti da ne postoji nužno prostorvrijeme koje odgovara nekom kauzetu. Pitanje evolucije kauzeta vodi nas do dinamike. U CST-u dinamika se naziva sekvencijalni rast, tj. rast skupa jest ono što se smatra "tokom vremena".

Ovo su glavne ideje kauzalne teorije skupova. Kauzalni odnosi i volumen zajedno nose dovoljno informacija da odrede geometriju.

7.4 Weinberg-Wittenov theorem

Weinberg-Wittenov teorem ograničava spin, tj. helicitet j bezmasenih čestica koje nose naboj očuvane Lorentz kovarijantne struje [95][77].

(a) Teorija u kojoj možemo konstruirati očuvanu Lorentz kovarijantnu struju J^μ ne može sadržavati bezmasene čestice spina $j > \frac{1}{2}$ ako ona nosi naboj $Q \equiv \int J^0 d^3x \neq 0$.

(b) Teorija u kojoj možemo konstruirati Lorentz kovarijantni tenzor energije i impulsa za koji je $P^\nu \equiv \int T^{0\nu} d^3x$ očuvani impuls ne može sadržavati čestice spina $j > 1$.

Kako bismo razumjeli dokaz ovog teorema moramo najprije definirati spin i bezmasene čestice.

7.4.1 Masa i spin

Jednočestična stanja opisana masom i spinom su ireducibilne reprezentacije Poincaréove grupe, grupe izometrija ravnog prostorvremena. Stanja kvantnog sustava razapinju Hilbertov prostor gdje je svako stanje sustava reprezentirano vektorom stanja.

Njih možemo odrediti ako znamo grupu simetrija teorije koja opisuje čestice. Ona ostavlja jednadžbe gibanja invarijantne, a dotični Lagrangian invarijantan do na rubni član. Na ravnom prostorvremenu to je Poincaréova grupa, što će biti objašnjeno kasnije. Nadalje, ako želimo da grupa djeluje na vektore stanja, moramo naći njezinu reprezentaciju. Kad radimo s vektorima stanja želimo da reprezentacija bude unitarna kako bi vjerojatnost bila očuvana. Štoviše, ako postoje Casimirovi operatori, Hilbertov prostor stanja možemo organizirati u minimalne invarijantne podprostore, tako da je svaki podprostor ireducibilna reprezentacija. Poincaréova grupa ima dva Casimirova operatora koja označuju ireducibilnu reprezentaciju. Cilj je stoga odrediti Casimirove operatore i ireducibilne reprezentacije Poincaréove grupe.

Ireducibilne reprezentacije Poincaréove grupe Svojstva Poincaréove grupe i Poincaréove algebre dana su u Dodatcima. Jer Poincaréova grupa nije kompaktna, ireducibilna reprezentacija je beskonačno dimenzionalna. Poincaréova grupa sadržava invarijantnu podgrupu koja je ujedno Abelova, grupu translacija, pa ireducibilne reprezentacije možemo naći metodom induciranih reprezentacija [80].

Stanje kvantnog sustava jedinstveno određeno svojstvenim vrijednostima potpunog skupa operatora koji međusobno komutiraju. Stoga, krećemo od proizvoljno odabranog svojstvenog stanja operatora impulsa kojeg nazivamo standardni impuls. Zatim odredimo malu grupu standardnog vektora. Ona se sastoji od generatora koji ostavljaju svojstvenu vrijednost standardnog impulsa invarijantnom. Broj generatora u maloj grupi određuje dimenziju podprostora koji pripada standardom impulsu. Potpunu ireducibilnu reprezentaciju onda dobivamo tako da generatori koji nisu elementi male grupe djeluju na stanja podprostora.

Casimirovi operatori P^2 i W^2 označuju Poincaréovu reprezentaciju. Impuls pripada jednoj od četiri moguće klase,

- $P^2 > 0$,
- $P^2 = 0$, $P = 0$
- $P^2 = 0$, $P \neq 0$
- $P^2 < 0$.

Nas zanimaju bezmasene čestice pa ćemo promatrati samo slučaj kad je kvadrat impulsa $P^2 = 0$, tj. $m^2 = 0$ i impuls $P \neq 0$. Drugi Casimirov operator W^2 je Pauli-Lubanski pseudovektor. Kao standardni impuls bezmasenih čestica odabrat ćemo

$$p_0 = (E, 0, 0, E) \,. \tag{7.66}$$

Odredimo malu grupu tog vektora. Iz 7.66 vidimo da mora uključivati $SO(2)$. Algebra male grupe je

$$\begin{aligned} [W_1, W_2] &= 0 \,, \\ [W_2, J_3] &= iW_1 \,, \\ [W_1, J_3] &= -iW_2 \,. \end{aligned} \tag{7.67}$$

Prepoznajemo da su ovo komutacijske relacije Euklidske grupe u dvije dimenzije. Stoga slijedi da je stanje definirao kao

$$\begin{aligned} P^\mu |p_0, w_1, w_2\rangle &= p_0^\mu |p_0, w_1, w_2\rangle \,, \\ W_1 |p_0, w_1, w_2\rangle &= w_1 |p_0, w_1, w_2\rangle \,, \\ W_2 |p_0, w_1, w_2\rangle &= w_2 |p_0, w_1, w_2\rangle \,. \end{aligned} \tag{7.68}$$

Može se pokazati da su svojstvena stanja W_1 i W_2 kontinuirana. Jer nisu opažene čestice koje bi bile opisane takvim kvantnim brojevima, zahtijevamo da $W_1 = W_2 = 0$. Iz toga slijedi da je generator male grupa bezmasene čestice J_3 koji odgovara helicitetu. Jer je grupa Abelova,

$$U(R(\theta)) = e^{-i\theta J_3} \,. \tag{7.69}$$

Dakle, stanje je opisano kao

$$P^\mu |m = 0, j = 0; \boldsymbol{p}_0, \pm j\rangle = p^\mu |m = 0, j = 0; \boldsymbol{p}_0, \pm j\rangle \tag{7.70}$$

$$W_3 |m = 0, j = 0; \boldsymbol{p}_0, \pm j\rangle = \pm j |m = 0, j = 0; \boldsymbol{p}_0, \pm j\rangle \tag{7.71}$$

Da bismo dobili cijelu ireducibilnu reprezentaciju moramo djelovati operatorima koji ne pripadaju maloj grupi. To je Lorentzov potisak $L_0(p)$.

$$|m=0,j=0;\boldsymbol{p},\pm j\rangle = R(\phi,\theta,0)U(L_0(p))|m=0,j=0;\boldsymbol{p}_0,\pm j\rangle\,, \tag{7.72}$$

gdje je R rotacija i $|\boldsymbol{p}| = e^{\tanh^{-1}(\beta)}E$. Konačno, matrični elementi ireducibilne reprezentacije su

$$U(\mathbb{1},a)|\boldsymbol{p},\pm j\rangle = e^{-iap}|\boldsymbol{p},\pm j\rangle\,, \tag{7.73}$$

$$U(\Lambda,0)|\boldsymbol{p},\pm j\rangle = e^{\pm ji\theta(\Lambda,p)}|\Lambda\boldsymbol{p},\pm j\rangle\,, \tag{7.74}$$

gdje $\theta(\Lambda,p)$ možemo odrediti iz

$$e^{\pm i\theta(\Lambda,p)} = \langle\boldsymbol{p}_0,\pm j|L_0(\Lambda p)^{-1}\Lambda L_0(p)|\boldsymbol{p}_0,\pm j\rangle\,. \tag{7.75}$$

Time je zaključena rasprava o ireducibilnim reprezentacijama Poincaréove grupe [81][80][82].

7.4.2 Polja i čestice

U ovom poglavlju objašnjeno je koja je veza između polja i čestica, te kako se polja transformiraju na Lorentzove transformacije.

Čestice opisuju fizikalna stanja, stoga proizlaze iz unitarne reprezentacije. Isto ne vrijedi za polja, stoga su ona povezana s ne-unitarnom reprezentacijom. Iz tog razloga ne postoji direktna veza između čestica i polja, već je ona dana jednadžbom gibanja. Ona je zapravo relativistički izraz za energiju, čijim rješavanjem dobivamo

$$\Psi(x) = \sum_\lambda \int \frac{d^3p}{(2\pi)^3\sqrt{2E}}\left(a(\boldsymbol{p},\lambda)\varepsilon^a(\boldsymbol{p},\lambda)e^{ipx} + a^\dagger(\boldsymbol{p},\lambda)\varepsilon^{*a}(\boldsymbol{p},\lambda)e^{-ipx}\right)\,, \tag{7.76}$$

gdje je $a(\boldsymbol{p},\lambda)$ operator poništenja i

$$|\boldsymbol{p},\lambda\rangle = a^\dagger(\boldsymbol{p},\lambda)|0\rangle\,, \tag{7.77}$$

gdje je $|0\rangle$ vakuum. Jer znamo kako se jednočestična stanja transformiraju na Lorentzove transformacije, slijedi da za polja mora vrijediti

$$U(\Lambda)\Psi(x)^a U(\Lambda^{-1}) = D(\Lambda^{-1})^a_b \Psi^b(\Lambda x)\,, \tag{7.78}$$

gdje je $U(\Lambda)$ ne-unitarna reprezentacija i $D(\Lambda)$ je reprezentacija Lorentzove grupe. No, dovoljan i nužan uvjet da gornja relacija vrijedi jest

$$D(\Lambda)^a_{a'}\varepsilon^{a'}(\boldsymbol{p},\lambda) = D^{(j)}(R(\Lambda,p))^\lambda_{\lambda'}\varepsilon^a(\Lambda\boldsymbol{p},\lambda')\,. \tag{7.79}$$

Za konkretan slučaj bezmasene čestice heliciteta 1, odgovarajuće polje A^μ

$$A^\mu(x) = \sum_\lambda \int \frac{d^3p}{(2\pi)^3\sqrt{2E}}\left(a(\boldsymbol{p},\lambda)\varepsilon^a(\boldsymbol{p},\lambda)e^{ipx} + a^\dagger(\boldsymbol{p},\lambda)\varepsilon^{*a}(\boldsymbol{p},\lambda)e^{-ipx}\right)\,, \tag{7.80}$$

transformira se na Lorentzove transformacije kao

$$U(\Lambda)A_\mu(x)U^{-1}(\Lambda) = \Lambda^\nu_\mu A_\nu(\Lambda x) + \partial_\mu\Omega(x,\Lambda)\,. \tag{7.81}$$

Bezmasena polja ne transformiraju se kovarijantno na Lorentzove transformacije. Rezultat zapravo slijedi zbog toga što mala grupa sadrži translacije.

7.4.3 Noetherin teorem i zakoni očuvanja

Očuvane struje i naboje nalazimo pomoću Noetherinog teorema koji kaže da očuvane struje slijede iz akcije koja je invarijantna na kontinuirane globalne simetrije. Dokaz slijedi iz varijacionog postupka. Polazimo od akije $S[\psi_i]$,

$$S[\psi_i] = \int d^4x\mathcal{L}(\psi_i, \partial_\mu\psi_i)\,, \quad i = 1, ..., N\,. \tag{7.82}$$

Veličina $\mathcal{L}$ je gustoća Lagrangiana koju zovemo Lagrangian, ψ je polje (koje nije nužno skalarno).

Cilj je primijeniti Weinberg-Wittenov teorem na standardni model i gravitaciju. Jedine čestice standardnog modela koje su u domeni Weinberg-Wittenovog teorema su gluoni koji nose naboj boju i imaju spin $j = 1$. Medijator gravitacijske sile je hipotetska čestica graviton spina $j = 2$. Njegov naboj jest energija, tj. impuls. Grupa simetrija QCD-a je $SU(3)$, dok je Einstein-Hilbertova akcija invarijantna na generalne difeomorfizme, koje ćemo sada ukratko objasniti.

Difeomorfizmi su prostornovremenske transformacije. Vektorsko polje ξ u svakoj točki generira tok Φ^{ξ}. Pomoću tog toka možemo pomaknuti točku P tako da slijedimo tok za neku vrijednost parametra ϵ,

$$\begin{aligned} \Phi^{\xi}_{\epsilon} &: U \to M \\ &P \mapsto \Phi^{\xi}_{\epsilon}(P) = Q \,. \end{aligned} \tag{7.83}$$

U koordinatama je 7.83 dan s

$$x(P) = x^{\mu} \to x(Q) = x^{\mu} + \epsilon(x)\xi^{\mu} = x^{\mu} + \epsilon^{\mu}(x) \,. \tag{7.84}$$

Nadalje, zanima nas kako izgleda polje koje opisuje ovu novu konfiguraciju točaka. Uzmimo primjer skalarnog polja $\phi(x)$ koje opisuje situaciju prije prostornovremenske transformacije i povežimo ga s $\phi'(x)$ koje opisuje novu konfiguraciju. Mora vrijediti da je

$$\begin{aligned} \phi'(x(Q)) &= \phi(x(P)) \\ \phi'(x+\epsilon) &= \phi(x) \,. \end{aligned} \tag{7.85}$$

S gledišta novog polja imamo vezu

$$\begin{aligned} \phi'(x) &= \phi(x) - \epsilon\xi^{\mu}(x)\partial_{\mu}\phi(x) \\ &= \phi(x) - \mathfrak{L}_{\xi}\phi \,, \end{aligned} \tag{7.86}$$

gdje je $\mathfrak{L}_{\xi}\phi$ Liejeva derivacija. Drugim riječima, Liejeve derivacije generiraju infinitezimalne difeomorfizme. Sličnu analizu možemo provesti za vektorsko polje, tj. tenzor bilo kojeg višeg ranga.

Vratimo se na pitanje očuvanih struja. Kao što je spomenuto, one slijede iz varijacije akcije,

$$\delta S = \int d^4x \frac{\delta S}{\delta \psi_i} \delta\psi_i \,, \tag{7.87}$$

Varijacija Lagrangiana je dana s

$$
\begin{aligned}
&\mathcal{L} \to \mathcal{L} + \delta\mathcal{L}\,, \\
\delta\mathcal{L} &\equiv \frac{\partial\mathcal{L}}{\partial\psi_i}\delta\psi_i + \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\delta\left(\partial_\mu\psi_i\right) \\
&= \left[\frac{\partial\mathcal{L}}{\partial\psi_i} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\right)\right]\delta\psi_i + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\delta\psi_i\right)\,,
\end{aligned}
\tag{7.88}
$$

gdje smo iskoristili

$$
\delta(\partial_\mu\psi_i) = \partial_\mu\psi_i' - \partial_\mu\psi_i = \partial_\mu(\psi_i' - \psi_i) = \partial_\mu(\delta\psi_i)\,. \tag{7.89}
$$

Kad se radi o simetrijskoj transformaciji akcija se može promijeniti do na rubni član,

$$
\delta_s\mathcal{L} = \partial_\mu K^\mu \implies \quad \delta_s S = S[\psi_i + \delta_s\psi_i] - S[\psi_i] = \int d^4x\,\partial_\mu K^\mu\,, \tag{7.90}
$$

gdje je K^μ neka funkcija. Varijacija koja je simetrija označena je s indeksom s. Valja primijetiti da kad dozvolimo da se akcija promijeni do na rubni član podrazumijevamo da taj član iščezava za iste rubne uvjete koji su vrijedili prije transformacije. Ako iskoristimo prethodne izraze, dobivamo da

$$
\begin{aligned}
\int d^4x\partial_\mu K^\mu &= \int d^4x\left\{\left[\frac{\partial\mathcal{L}}{\partial\psi_i} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\right)\right]\delta_s\psi_i + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\delta_s\psi_i\right)\right\} \\
0 &= \int d^4x\left\{\left[\frac{\partial\mathcal{L}}{\partial\psi_i} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\right)\right]\delta_s\psi_i + \partial_\mu j^\mu\right\},
\end{aligned}
\tag{7.91}
$$

gdje je j^μ definiran kao

$$
j^\mu \equiv \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\delta_s\psi_i - K^\mu\,. \tag{7.92}
$$

Da bismo pojednostavili notaciju uvodimo pokratu

$$
\left[\frac{\delta\mathcal{L}}{\delta\psi_i}\right]_{EL} \equiv \left[\frac{\partial\mathcal{L}}{\partial\psi_i} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\right)\right]\,. \tag{7.93}
$$

Slijedi da kad je transformacija simetrija akcije,

$$
\left[\frac{\delta\mathcal{L}}{\delta\psi_i}\right]_{EL}\delta_s\psi_i = -\partial_\mu j^\mu\,. \tag{7.94}
$$

Nadalje, promotrimo slučaj kad je varijacija globalna za slučaj difeomorfizama i $SU(N)$ grupe,

$$\begin{aligned}\delta\psi_i &= \delta_0\psi_i - \mathfrak{L}_\xi\psi_i \\ &= \alpha_i\omega_a X^a - \xi^\mu\partial_\mu\psi_i\,.\end{aligned} \tag{7.95}$$

Lagrangian se transformira kao

$$\mathfrak{L}_\xi\mathcal{L} = \partial_\mu(\mathcal{L}\xi^\mu) \implies K^\mu = \mathcal{L}\xi^\mu\,, \tag{7.96}$$

jer je skalarna gustoća. Konačno, dobivamo rezultat

$$\begin{aligned}j^\mu &= \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\delta_0\psi_i - \left[\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}(\xi^\nu\partial_\nu\psi_i) - \delta^\mu_\nu\mathcal{L}\xi^\nu\right] \\ &= \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\alpha_i\omega^a X_a - \left[\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}(\partial_\nu\psi_i) - \delta^\mu_\nu\mathcal{L}\right]\xi^\nu \\ &\equiv J^\mu_\alpha\omega^\alpha - T^\mu_\nu\xi^\nu\,,\end{aligned} \tag{7.97}$$

gdje smo definirali J^μ_α

$$J^\mu_\alpha \equiv \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}\alpha_i X_a\,, \tag{7.98}$$

i veličinu T^μ_ν,

$$T^\mu_\nu \equiv \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi_i\right)}(\partial_\nu\psi_i) - \delta^\mu_\nu\mathcal{L}\,. \tag{7.99}$$

Relacija 7.94 se svodi na oblik,

$$\begin{aligned}\left[\frac{\delta\mathcal{L}}{\delta\psi_i}\right]_{EL}(\alpha_i\omega^a X_a - \xi^\mu\partial_\mu\psi_i) = -\partial_\mu j^\mu \implies \\ \left[\frac{\delta\mathcal{L}}{\delta\psi_i}\right]_{EL}\alpha_{i,a} = -\partial_\mu J^\mu_\alpha\,, \\ \left[\frac{\delta\mathcal{L}}{\delta\psi_i}\right]_{EL}\partial_\mu\psi_i = -\partial_\mu T^\mu_\nu\,.\end{aligned} \tag{7.100}$$

Ako sva polja zadovoljavaju jednadžbe gibanja vrijedi da je

$$\partial_\mu j^\mu = 0\,. \tag{7.101}$$

Jer su varijacije nezavisne,

$$\partial_\mu J^\mu_\alpha = \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i \right)} \alpha_{a,i} \right] = 0\,, \tag{7.102}$$

$$\partial_\mu T^\mu_\nu = \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \psi_i \right)} \left(\partial_\nu \psi_i \right) - \delta^\mu_\nu \mathcal{L} \right] = 0\,. \tag{7.103}$$

Ovo su lokalni zakoni očuvanja koji vrijede za svaki parametar grupe. Rezultat je poznat kao Noetherin teorem. Veličine J^μ_α i T^μ_ν su Noetherine struje. Veličina T^μ_ν je tenzor energije i impulsa. Iz ovih veličina možemo dobiti i globalne zakone očuvanja, tzv. Noetherine naboje.

$$\partial_\mu J^\mu_\alpha = \partial_0 J^0_a + \partial_i J^i_\alpha\,. \tag{7.104}$$

Naboj Q_α definiramo kao

$$Q_\alpha(t) \equiv \int d^3x J^0_a\,, \tag{7.105}$$

Nulta komponenta struje odgovara gustoći. Ako polja trnu dovoljno brzo u beskonačnosti,

$$\begin{aligned} 0 &= \int d^3x \left[\partial_0 J^0_\alpha(\mathbf{x},t) + \partial_i J^i_\alpha(\mathbf{x},t) \right] \\ &= \int d^3x \partial_0 J^0_\alpha(\mathbf{x},t) \\ &= \partial_0 Q_\alpha(t)\,. \end{aligned} \tag{7.106}$$

Na potpuno analogan način možemo pokazati da je

$$P^\mu \equiv \int d^3x T^{0\mu}\,, \tag{7.107}$$

gdje je P^μ impuls.

Bitno je reći da postoji i drugi Noetherin teorem koji bismo dobili da nismo pretpostavili da su transformacije globalne.

$$\left\{ \left[\frac{\delta \mathcal{L}}{\delta \psi_i} \right]_{EL} \alpha_i - \partial_\mu \left(\beta_i \left[\frac{\delta \mathcal{L}}{\delta \psi_i} \right]_{EL} \right) \right\} \omega = -\partial_\mu (j^\mu + b^\mu)\,. \tag{7.108}$$

Drugi Noetherin teorem vrijedi uvijek, i kad jednadžbe gibanja nisu zadovoljene. U slučaju da jesu, dobivamo tautologiju. Posljedica drugog Noetherinog teorema jest da ako je globalna grupa simetrija podgrupa veće grupe, grupe lokalnih simetrija, tada ne postoje zakoni očuvanja, već identiteti. Dobivene rezultate primijenit ćemo na QCD-u i gravitaciji.

7.4.4 QCD

Lagrangian QCD-a dan je izrazom

$$\begin{aligned}\mathcal{L}_g(G^a_\mu) &= -\frac{1}{4}F^a_{\mu\nu}(x)F^{\mu\nu}_a(x)\,,\\ F^a_{\mu\nu}(x) &= \partial_\mu G^a_\nu(x) - \partial_\nu G^a_\mu(x) + f_{abc}G^b_\mu(x)G^c_\nu(x)\\ &= D_\mu G^a_\nu(x) - D_\nu G^a_\mu(x) + f_{abc}G^b_\mu(x)G^c_\nu(x)\,.\end{aligned} \tag{7.109}$$

Polje G^a_μ je gluonsko polje, a D_μ je kovarijantna derivacija

$$D_\mu\psi = \partial_\mu\psi + iG_\mu\psi\,. \tag{7.110}$$

Grupa simetrija je $SU(3)$ i Lagrangian je invarijantan na lokalne baždarne transformacije. Iako iz drugog Noetherinog teorema slijedi da ne bi trebao postojati zakon očuvanja, to ipak nije slučaj zbog toga što je $F^{\mu\nu}_a$ antisimetričan. Odnosno, dobije se rezultat da je drugi Noetherin teorem uvjet da jednadžbe gibanja budu zadovoljene. Postupkom varijacije dobivamo

$$\partial_\mu F^{\mu\nu}_a = f_{bac}F^{\nu\rho,b}G^c_\rho\,. \tag{7.111}$$

Možemo definirati očuvanu struju

$$\mathcal{J}^\nu_a = \partial_\mu F^{\mu\nu}_a = f_{bac}F^{\nu\rho,b}G^c_\rho \implies \partial_\mu\mathcal{J}^\mu_a = 0\,, \tag{7.112}$$

i naboj koji je dan izrazom

$$Q_a = \int d^3x\partial_\mu F^{\mu 0}_a = \int d^3x F^{\mu 0,b}G^c_\mu f_{abc}\,. \tag{7.113}$$

Pokazali smo da postoji očuvana struja i naboj za kvantnu kromodinamiku [27].

7.4.5 Gravitacija

U slučaju gravitacije postoji više akcija koje daju iste jednadžbe gibanja

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \,. \tag{7.114}$$

Tenzor $G_{\mu\nu}$ je Einsteinov tenzor, $R_{\mu\nu}$ je Riccijev tenzor i R je Riccijev skalar,

$$R = g^{\mu\nu}(\Gamma^{\sigma}_{\mu\sigma,\nu} - \Gamma^{\sigma}_{\mu\nu,\sigma}) - g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \,, \tag{7.115}$$

gdje su Christoffelovi simboli Γ dani s

$$\Gamma^{\rho}_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} \left(\partial_\nu g_{\sigma\mu} + \partial_\mu g_{\sigma\nu} - \partial_\sigma g_{\mu\nu} \right) \,. \tag{7.116}$$

Akcija koja se najčešće pojavljuje u literaturi je Einstein-Hilbertova (EH) akcija

$$S_{EH}[g^{\mu\nu}] = \int d^4x \sqrt{-g} R \,. \tag{7.117}$$

Grupa simetrija EH akcije je grupa generalnih difeomorfizama pa ne postoji očuvana struja. Točnije, zakoni očuvanja nisu kompatibilni s generalnim difeomorfizmima EH akcije. Kako bismo dobili akciju koja daje iste jednadžbe gibanja, ali ima globalnu grupu simetrija zapišemo integrand EH akcije u obliku

$$\begin{aligned} R\sqrt{-g} &= \sqrt{-g} g^{\mu\nu}(\Gamma^{\sigma}_{\mu\sigma,\nu} - \Gamma^{\sigma}_{\mu\nu,\sigma}) - \sqrt{-g} g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \\ &= \partial_\nu(\sqrt{-g} g^{\mu\nu}\Gamma^{\sigma}_{\mu\sigma}) - \partial_\sigma(\sqrt{-g} g^{\mu\nu}\Gamma^{\sigma}_{\mu\nu}) - \partial_\nu(\sqrt{-g} g^{\mu\nu})\Gamma^{\sigma}_{\mu\sigma} + \partial_\sigma(\sqrt{-g} g^{\mu\nu})\Gamma^{\sigma}_{\mu\nu} - \\ &\quad \sqrt{-g} g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \\ &= \partial_\nu(\sqrt{-g} g^{\mu\nu}\Gamma^{\sigma}_{\mu\sigma}) - \partial_\sigma(\sqrt{-g} g^{\mu\nu}\Gamma^{\sigma}_{\mu\nu}) + g^{v\beta}\Gamma^{\mu}_{\beta v}\Gamma^{\sigma}_{\mu\sigma}\sqrt{-g} + \\ &\quad \left(-2g^{\nu\beta}\Gamma^{\mu}_{\beta\sigma} + g^{\mu\nu}\Gamma^{\beta}_{\sigma\beta} \right) \Gamma^{\sigma}_{\mu\nu} - \sqrt{-g} g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \\ &= \partial_\nu(\sqrt{-g} g^{\mu\nu}\Gamma^{\sigma}_{\mu\sigma}) - \partial_\sigma(\sqrt{-g} g^{\mu\nu}\Gamma^{\sigma}_{\mu\nu}) + \sqrt{-g} g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \,. \end{aligned} \tag{7.118}$$

Einsteinove jednadžbe gibanja mogu se dobiti samo iz zadnjeg člana.

$$S'_G = \int d^4x \sqrt{-g} g^{\mu\nu}(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\nu\rho}) \,. \tag{7.119}$$

Akcija 7.119 invarijantna je na globalne transformacije,

$$\delta g_{\alpha\beta} = \xi^{\mu}\partial_{\mu}g_{\alpha\beta}\,. \tag{7.120}$$

Primijetimo da nova akcija sadrži samo prve derivacije metrike. No, iako izgleda kao skalar, ne transformira se dobro na generalne difeomorfizme. Iz Noetherinog teorema slijedi da je

$$\tau^{\nu}_{\mu}\sqrt{-g} = \left(\frac{\partial \mathcal{L}'_G}{\partial(\partial_\nu g_{\alpha\beta})}\right)\partial_\mu g_{\alpha\beta} - \delta^{\nu}_{\mu}\mathcal{L}'_G\,, \tag{7.121}$$

gdje $\mathcal{L}'_G = \sqrt{-g}g^{\lambda\tau}(\Gamma^{\sigma}_{\lambda\tau}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\lambda\sigma}\Gamma^{\sigma}_{\tau\rho})$. Kad se sve uvrsti dobivamo

$$\tau^{\nu}_{\mu}\sqrt{-g} = \left[\left(\Gamma^{\nu}_{\alpha\beta} - \delta^{\nu}_{\beta}\Gamma^{\sigma}_{\alpha\sigma}\right)\partial_\mu\left(g^{\alpha\beta}\sqrt{-g}\right) - \delta^{\nu}_{\mu}\sqrt{-g}g^{\lambda\tau}(\Gamma^{\sigma}_{\lambda\tau}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\rho}_{\lambda\sigma}\Gamma^{\sigma}_{\tau\rho})\right]. \tag{7.122}$$

Veličina τ je pseudovektor energije i impulsa. Možemo provjeriti da doista je očuvan kad su jednadžbe gibanja zadovoljene. Očuvani naboj je impuls P,

$$P^{\mu} = \int d^3x\tau^{0\mu}\,. \tag{7.123}$$

Iako je τ pseudovektor transformira se kao vektor na Lorentzove transformacije. Primijetimo da je ova diskusija klasična. Da bismo primijenili Weinberg-Wittenov teorem želimo promatrati graviton koji se propagira na ravnom prostrovremenu. To se postiže tako da se graviton promatra kao mala perturbacija ravne metrike,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}\,, \quad |h_{\mu\nu}| << 1\,, \tag{7.124}$$

gdje je dinamičko polje $h_{\mu\nu}$ graviton [92].

7.4.6 Dokaz Weinberg-Wittenovog teorema

Dokaz Weinberg-Wittenovog teorema kreće od matričnih elemenata

$$\langle p', \pm j|J^{\mu}|p, \pm j\rangle\,, \tag{7.125}$$

$$\langle p', \pm j|T^{\mu\nu}|p, \pm j\rangle\,. \tag{7.126}$$

Izrazi su motivirani dalje u tekstu. Pretpostavimo da bezmasene čestice koje promatramo nose naboj,

$$Q|p, \pm j\rangle = q|p, \pm j\rangle \,, \tag{7.127}$$

$$P^{\mu}|p, \pm j\rangle = p^{\mu}|p, \pm j\rangle \,, \tag{7.128}$$

što vodi do izraza

$$\langle p', \pm j|Q|p, \pm j\rangle = q\langle p' \pm j|p, \pm j\rangle \,, \tag{7.129}$$

$$\langle p', \pm j|P^{\mu}|p, \pm j\rangle = p^{\mu}\langle p' \pm j|p, \pm j\rangle \,. \tag{7.130}$$

Treba biti pažljiv kad izvrjednjujemo desnu stranu. Obično se uzima da je $\langle p' \pm j|p, \pm j\rangle = \delta^{(3)}(\boldsymbol{p'} - \boldsymbol{p})$, no u stvarnosti je impuls “razmazan". Dakle, fizikalno je točnije uzeti

$$\langle p', \pm j|p, \pm j\rangle = \delta_a^{(3)}(\boldsymbol{p'} - \boldsymbol{p}) \,, \tag{7.131}$$

gdje smo definirali

$$\lim_{a \to 0} \delta_a^{(3)}(\boldsymbol{p'} - \boldsymbol{p}) = \delta^{(3)}(\boldsymbol{p'} - \boldsymbol{p}) \,. \tag{7.132}$$

Parametar a je određen preciznošću eksperimenta. Stoga, slijedi da je

$$\langle p', \pm j|Q|p, \pm j\rangle = q\delta_a^{(3)}(\boldsymbol{p'} - \boldsymbol{p}) \,, \tag{7.133}$$

$$\langle p', \pm j|P^{\mu}|p, \pm j\rangle = p^{\mu}\delta_a^{(3)}(\boldsymbol{p'} - \boldsymbol{p}) \,. \tag{7.134}$$

Ako raspišemo lijevu stranu ovih jednadžbi dobivamo

$$\begin{aligned}
\langle p', \pm j|Q|p, \pm j\rangle &= \int_{V_a} d^3x \left\langle p', \pm j \left| J^0(t, \boldsymbol{x}) \right| p, \pm j \right\rangle \\
&= \int_{V_a} d^3x \left\langle p', \pm j \left| e^{i\boldsymbol{P}\cdot\boldsymbol{x}} J^0(t, 0) e^{-i\boldsymbol{P}\cdot\boldsymbol{x}} \right| p, \pm j \right\rangle \\
&= \int_{V_a} d^3x e^{i(\boldsymbol{p'}-\boldsymbol{p})\boldsymbol{x}} \left\langle p', \pm j \left| J^0(t, 0) \right| p, \pm j \right\rangle \\
&= (2\pi)^3 \delta_a^3 \left(\boldsymbol{p'} - \boldsymbol{p}\right) \left\langle p', \pm j \left| J^0(t, 0) \right| p, \pm j \right\rangle \,.
\end{aligned} \tag{7.135}$$

Analogno, izraz za impuls je

$$
\begin{aligned}
\langle p', \pm j | P^\mu | p, \pm j\rangle &= \int_{V_a} d^3x \left\langle p', \pm j \left| T^{0\mu}(t, \boldsymbol{x}) \right| p, \pm j \right\rangle \\
&= \int_{V_a} d^3x \left\langle p', \pm j \left| e^{i\boldsymbol{P}\cdot\boldsymbol{x}} T^{0\mu}(t, 0) e^{-i\boldsymbol{P}\cdot\boldsymbol{x}} \right| p, \pm j \right\rangle \\
&= \int_{V_a} d^3x e^{i(\boldsymbol{p}'-\boldsymbol{p})\boldsymbol{x}} \left\langle p', \pm j \left| T^{0\mu}(t, 0) \right| p, \pm j \right\rangle \\
&= (2\pi)^3 \delta_a^3 \left(\boldsymbol{p}' - \boldsymbol{p}\right) \left\langle p', \pm j \left| T^{0\mu}(t, 0) \right| p, \pm j \right\rangle \ .
\end{aligned}
\tag{7.136}
$$

Integral je po nekom konačnom volumenu V_a. Ako sad usporedimo lijevu i desnu stranu dobivenih izraza, dobivamo da mora vrijediti da je za proizvoljnu komponentu

$$
\lim_{p' \to p} \langle p', \pm j \left| J^\mu(t, 0) \right| p, \pm j\rangle = \frac{q p^\mu}{(2\pi)^3 E}\,, \tag{7.137}
$$

$$
\lim_{p' \to p} \langle p', \pm j \left| T^{\mu\nu}(t, 0) \right| p, \pm j\rangle = \frac{p^\mu p^\nu}{(2\pi)^3 E}\,. \tag{7.138}
$$

Ovi izrazi odgovaraju eksperimentalnoj definiciji naboja i impulsa. Nadalje, promotrimo slučajeve kad je $j > 1/2$ i $j > 1$ za impulse p i p',

$$
\begin{aligned}
(p' + p)^2 &= 2(p'p) \\
&= 2\left(|\boldsymbol{p}'|\,|\boldsymbol{p}| - \boldsymbol{p}' \cdot \boldsymbol{p}\right) \\
&= 2\,|\boldsymbol{p}'|\,|\boldsymbol{p}|(1 - \cos\phi) \leq 0\,,
\end{aligned}
\tag{7.139}
$$

gdje je ϕ kut između prostornih komponenti impulsa. Ako je $\phi \neq 0$ ukupni impuls je vremenoliki i možemo odabrati referentni sustav u kojem je ukupna prostorna komponenta impulsa jednaka nuli,

$$
p = (|\boldsymbol{p}|, \boldsymbol{p}), \quad p' = (|\boldsymbol{p}|, -\boldsymbol{p})\,. \tag{7.140}
$$

U tom sustavu promotrimo rotaciju za kut θ oko osi u smjeru $\boldsymbol{p}$,

$$
\begin{aligned}
|p, \pm j\rangle &\to U(R_W(\theta))|p, \pm j\rangle = e^{\pm i\theta j}|p, \pm j\rangle\,, \\
|p', \pm j\rangle &\to U(R_W(\theta))\,|p', \pm j\rangle = e^{\mp i\theta j}|p', \pm j\rangle\,,
\end{aligned}
\tag{7.141}
$$

S druge strane, umjesto da rotiramo česticu možemo rotirati operatore. Pretpostavimo da su oni Lorentz kovarijantni. Tada slijedi da

$$e^{\pm 2i\theta j} \langle p', \pm j | J^{\mu}(t,0) | p, \pm j \rangle = R_W(\theta)^{\mu}_{\nu} \langle p', \pm j | J^{\nu}(t,0) | p, \pm j \rangle \ , \tag{7.142}$$

$$e^{\pm 2i\theta j} \langle p', \pm j | T^{\mu\nu}(t,0) | p, \pm j \rangle = R_W(\theta)^{\mu}_{\rho} R_W(\theta)^{\nu}_{\sigma} \langle p', \pm j | T^{\rho\sigma}(t,0) | p, \pm j \rangle \ , \tag{7.143}$$

gdje je $R_W(\theta)$ matrica rotacije sa svojstvenim vrijednostima $e^{\pm i\theta}$ i 1. Zbog toga imamo

$$e^{\pm 2i\theta j} \langle p', \pm j | J^{\mu}(t,0) | p, \pm j \rangle = e^{\pm i\theta} \langle p', \pm j | J^{\mu}(t,0) | p, \pm j \rangle \ , \tag{7.144}$$

$$e^{\pm 2i\theta j} \langle p', \pm j | T^{\mu\nu}(t,0) | p, \pm j \rangle = e^{\pm 2i\theta} \langle p', \pm j | T^{\mu\nu}(t,0) | p, \pm j \rangle \ , \tag{7.145}$$

ili,

$$e^{\pm 2i\theta j} \langle p', \pm j | J^{\mu}(t,0) | p, \pm j \rangle = \langle p', \pm j | J^{\mu}(t,0) | p, \pm j \rangle \ , \tag{7.146}$$

$$e^{\pm 2i\theta j} \langle p', \pm j | T^{\mu\nu}(t,0) | p, \pm j \rangle = \langle p', \pm j | T^{\mu\nu}(t,0) | p, \pm j \rangle \ , \tag{7.147}$$

Konačno, slijedi rezultat da

$$\lim_{p' \to p} \langle p', \pm j | J^{\mu}(t,0) | p, \pm j \rangle = 0 \, , \quad j > \frac{1}{2} \, , \tag{7.148}$$

$$\lim_{p' \to p} \langle p', \pm j | T^{\mu\nu}(t,0) | p, \pm j \rangle = 0 \, , \quad j > 1 \, . \tag{7.149}$$

U dokazu smo pretpostavili da su struje Lorentz kovarijantne i da je struja očuvana.

7.4.7 Lorentz kovarijantnost i baždarne transformacije

Pokazali smo da gluon i graviton imaju očuvane struje. Nadalje, gluon nosi boju, a graviton ima impuls. Iako je graviton hipotetska čestica, znamo da gluon mora postojati, no Weinberg-Witten teorem to zabranjuje. Odgovor je u tome što se struje bezmasenih čestica ne transformiraju kovarijantno na Lorentzove transformacije jer ovise o baždarnim poljima. Podsjetimo se, za česticu spina 1,

$$A_{\mu} \to A_{\mu} + \partial_{\mu} \xi \, , \tag{7.150}$$

dok za česticu spina 2,

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu \ . \tag{7.151}$$

Weinberg-Wittenov teorem važan je jer eliminira mogućnost da je graviton kompozitna čestica. Na primjer, graviton se ne može sastaviti od dva gluona jer bi doprinosili Lorentz kovarijantnom, očuvanom tenzoru energije i impulsa (materije).

7.5 Termodinamika i opća teorija relativnosti

U ovom poglavlju izvedeni su zakoni mehanike crnih rupa sa naglaskom na njihovu analogiju s četiri zakona termodinamike. Zatim je izvedeno Hawkingovo zračenje. Naposljetku, razmatramo termodinamiku samog prostorvremena.

7.5.1 Klasični zakoni mehanike crnih rupa

U ovom poglavlju najprije uvodimo definicije koje su potrebne za razumijevanje zakona mehanike crnih rupa. Zatim dajemo pregled različitih načina na koji se oni mogu izvesti. Istovremeno uspoređujemo uvedene koncepte s termodinamikom.

Matematički temelji Hiperploha je $(n-1)$-dimenzionalna mnogostrukost uronjena u n dimenzionalnu mnogostrukost. Definirana je tako da zahtijevamo da je funkcija koordinata $\Phi(x^\mu)$,

$$\Phi(x^\mu) = 0 \, . \tag{7.152}$$

Hiperploha je svjetlosnog tipa ako je njezina normala n vektor svjetlosnog tipa.

$$n_\mu \propto \partial_\mu \Phi \ . \tag{7.153}$$

U tom slučaju, normala je ujedno i tangenta. Promotrimo slučaj kada normala koju ćemo zvati χ zadovoljava Killingovu jednadžbu,

$$\nabla_\mu \chi_\nu + \nabla_\nu \chi_\mu = 0 \, . \tag{7.154}$$

Pripadne integralne krivulje zadovoljavaju jednadžbu geodezika koji nije afino parametriziran. Štoviše, integralne krivulje su ortogonalne na hiperplohu, tj. zadovoljavaju Frobeniusov teorem.

Dalje nas zanima evolucija susjednih geodezika. Opisana je tenzorom $B_{\mu\nu}$,

$$B_{\mu\nu} = \nabla_\nu \chi_\mu \,. \tag{7.155}$$

Transverzalni dio dobije se tako da djelujemo na tenzor transverzalnom metrikom $h_{\mu\nu}$,

$$h_{\mu\nu} = g_{\mu\nu} + \chi_\mu N_\nu + N_\mu \chi_\nu \,. \tag{7.156}$$

Vektor N je pomoćni vektor svjetlosnog tipa. Tenzor možemo rastaviti na njegov simetrični i antisimetrični dio,

$$\widetilde{B}_{\mu\nu} = \frac{1}{2}\theta h_{\mu\nu} + \sigma_{\mu\nu} + \omega_{\mu\nu} \,, \tag{7.157}$$

gdje tilda označava da se radi o transverzalnoj komponenti. Veličina θ je ekspanzija geodezika, σ je smik, a ω je rotacija. Kao posljedica Frobeniusovog teorema, rotacija iščezava. Od velike je važnosti tzv. Raychaudhurijeva jednadžba koja opisuje evoluciju ekspanzije,

$$\begin{aligned} \frac{d\theta}{d\lambda} &= \chi^\rho \nabla_\rho (g^{\mu\nu} \widetilde{B}_{\mu\nu}) \\ &= \chi^\rho \nabla_\rho g^{\mu\nu} B_{\mu\nu} - \chi^\rho \nabla_\rho \kappa \,. \end{aligned} \tag{7.158}$$

Može se pokazati da odavde slijedi

$$\frac{d\theta}{d\lambda} = \kappa\theta - \frac{1}{2}\theta^2 - \sigma^{\mu\nu}\sigma_{\nu\mu} + \omega^{\nu\mu}\omega_{\nu\mu} - R_{\mu\nu}k^\mu k^\nu \,. \tag{7.159}$$

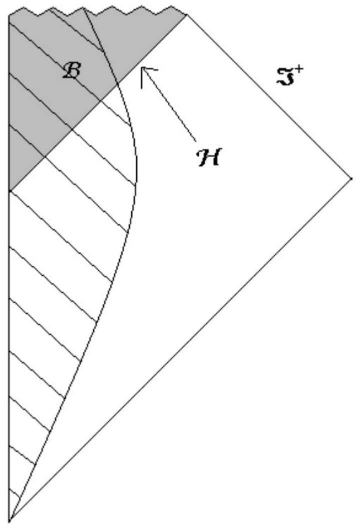
To je Raychaudhurijeva jednadžba. Uzevši u obzir da je χ Killingov vektor, slijedi da ekspanzija i smik iščezavaju [101].

7.5.2 Crne rupe

Crna rupa je dio prostorvremena gdje je gravitacija toliko jaka da ništa ne može pobjeći u beskonačnost. U asimptotski ravnom prostoru formalna definicija crne rupe je

$$\mathcal{B} = M - J^-(\mathcal{J}^+)\,, \tag{7.160}$$

gdje je $\mathcal{B}$ crna rupa, M prostorvrijeme, a $J^-(\mathcal{J}^+)$ kauzalna prošlost buduće svjetlosne beskonačnosti. Granica crne rupa je horizont događaja $\mathcal{H}$. Navedeni koncepti mogu se lijepo prikazati pomoću Penroseovog dijagrama prikazanog na slici 7.1. U



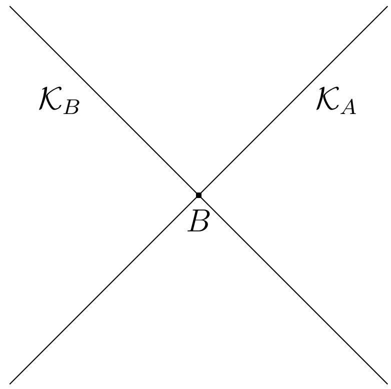
Slika 7.1: Penroseov dijagram sferično simetričnog urušavanja zvijezde (iscrtkani dio) [102]. Sivom bojom označena je crna rupa $\mathcal{B}$, a horizont je $\mathcal{H}$.

stvarnosti, crne rupe nastaju urušavanjem zvijezde čija je masa dovoljno velika da ništa ne može zaustaviti urušavanje. Na kraju se formira singularnost, točka beskonačne zakrivljenosti. Kad se sve smiri, kažemo da je u okolini crne rupe postignuto stacionarno stanje. Ono je analogno ravnotežnom stanju u termodinamici. Formalno, to znači da postoji Killingov vektor t^μ koji je vremenoliki u beskonačnosti. Štoviše, crna rupa je statična ako Killingov vektor t^μ zadovoljava Frobeniusov teorem. Za dokaze koji slijede moramo još uvesti pojam aksijalno simetričnog prostorvremena i Killingov hoizont.

Analogno definiciji za stacionarno prostorvrijeme, crna rupa je aksijalno simetrična ako postoji Killingov vektor ϕ^μ koji u beskonačnosti odgovara rotacijama. Ako je 2-dimenzionalna hiperploha razapeta vektorima t^μ i ϕ^μ ortogonalna na 2-

dimenzionalnu familiju hiperploha, kažemo da stacionarna aksisimetrična crna rupa ima $t-\phi$ svojstvo ortogonalnosti.

Killingov horizon je koncept koji je neovisan o horizontu događaja, a definiran je kao svjetlosna hiperploha čija je normala Killingov vektor. Specijalni slučaj Killingovog horizonta je Killingov horizont s bifurkacijom, prikazan na slici 7.2. Veza



Slika 7.2: Killingov horizont s bifurkacijom sastoji se od svjetlosnih hiperploha $\mathcal{K}_A$ i $\mathcal{K}_B$, koje se sjeku u bifurkacijskoj sferi B.

između Killingovog horizonta i horizonta događaja dana je tzv. teoremima o rigidnosti (krutosti). Killingov horizont poklapa se s horizontom događaja za

$$\chi^\mu = t^\mu + \Omega_H \phi^\mu \, . \tag{7.161}$$

Carterova verzija teorema o rigidnosti odnosi se na stacionarno aksijalno simetrične crne rupe sa svojstvo $t-\phi$ ortogonalnosti, dok se Hawkingova verzija odnosi na stacionarne crne rupe [98]. Vrijedi još spomenuti 'no-hair' teorem — crne rupe nemaju kosu. Odnosno, za opis crne rupe potreban je mali broj parametara.

Iz ovog razmatranja vidimo da su stacionarne crne rupe sustavi u ravnoteži. Granica tog sustava je horizont događaja, a opis crne rupe dan je malim brojem parametara. Ovo je početak veze s termodinamikom.

7.5.3 Površinska gravitacija

Na Killingovom horizontu Killingov vektor zadovoljava jednadžbu geodezika koja nije afino parametrizirana.

$$\chi_\mu \nabla^\mu \chi_\nu = \kappa \chi_\nu \, . \tag{7.162}$$

Veličina κ zove se površinska gravitacija. Ekvivalentna definicija površinske gravitacije dana je izrazom

$$\kappa^2 = -\frac{1}{2}\nabla^\mu \chi^\nu \nabla_\mu \chi_\nu \,. \tag{7.163}$$

Nadalje, može se pokazati da odavde slijedi

$$\kappa = \lim(Va) \,, \tag{7.164}$$

gdje je $a^\rho = \chi_\nu \nabla^\nu \chi^\rho / (-\chi^\mu \chi_\mu)$ akceleracija, a $V = \sqrt{-\chi^\mu \chi_\mu}$ takozvani faktor pomaka prema crvenom. Iz ovog izraza jasno je zašto se κ zove površinska gravitacija. U statičnom slučaju odgovara sili po jedinici mase koju mora primijeniti opažač u beskonačnosti da drži česticu koja se približava horizontu na mjestu.

7.5.4 Nulti zakon

Izjava prvog zakona mehanike crnih rupa jest

Površinska gravitacija stacionarne crne rupe je konstantna.

Prvi put je dokazan u [45]. Postoji nekoliko verzija dokaza, tj. ovisno o tome pretpostavi li se Hawkingova ili Carterova verzija teorema o rigidnosti potreban je dominantni uvjet na energiju [45][112].

Kako bismo dokazali nulti zakon moramo pokazati da kovarijantna derivacija površinske gravitacije iščezava. Potrebna je mjera opreza jer je površinska gravitacija κ definirana samo na horizontu. Stoga, promatramo $\chi_{[\mu}\nabla_{\nu]}\kappa$.

$$\chi_{[\mu}\nabla_{\nu]}\kappa = 0 \,. \tag{7.165}$$

Prvo pomnožimo $\chi^\mu \nabla_\mu \chi_\nu = \kappa \chi_\nu$ sa $\chi_{[\rho}\nabla_{\tau]}$,

$$\begin{aligned} \chi_\nu \chi_{[\rho}\nabla_{\tau]}\kappa + \kappa \chi_{[\rho}\nabla_{\tau]}\chi_\nu &= \chi_{[\rho}\nabla_{\tau]}(\chi^\mu \nabla_\mu \chi_\nu) \\ \chi_\nu \chi_{[\rho}\nabla_{\tau]}\kappa &= -\kappa \chi_{[\rho}\nabla_{\tau]}\chi_\nu + (\chi_{[\rho}\nabla_{\tau]}\chi^\mu)(\nabla_\mu \chi_\nu) + \chi^\mu(\chi_{[\rho}\nabla_{\tau]}\nabla_\mu \chi_\nu) \\ &= -\chi^\mu R_{\mu\nu[\tau}{}^{\lambda}\chi_{\rho]}\chi_\lambda \,. \end{aligned} \tag{7.166}$$

Korišten je identitet $\nabla_\mu \nabla_\nu \chi_\rho = -R_{\nu\rho\mu}{}^{\tau} \chi_\tau$. Sličnim postupkom pokaže se da

$$\begin{aligned} \chi_\mu \chi_{[\nu} \nabla_{\lambda]} \nabla_\rho \chi_\tau &= -2(\chi_{[\nu} \nabla_{\lambda]} \nabla_{[\tau} \chi_{|\mu|}) \chi_{\rho]} \\ -\chi_\mu R_{\rho\tau[\lambda}{}^{\sigma} \chi_{\nu]} \chi_\sigma &= 2\chi_{[\rho} R_{\tau]\mu[\lambda}{}^{\sigma} \chi_{\nu]} \chi_\sigma \; . \end{aligned} \tag{7.167}$$

Zatim kontrahiramo indekse μ i λ što nam daje

$$\begin{aligned} -\chi^\sigma \chi^\mu R_{\rho\tau[\mu\sigma} \chi_{\nu]} &= 2\chi_{[\rho} R_{\tau]}{}^{\lambda}{}_{[\lambda|\sigma|} \chi_{\nu]} \chi^\sigma \\ 0 &= \chi_{[\rho} R_{\tau]}{}^{\sigma} \chi_\sigma \chi_\nu - \chi_{[\rho} R_{\tau]\lambda\nu\sigma} \chi^\lambda \chi^\sigma \\ \chi^\sigma R_{\sigma\nu[\tau|\lambda|} \chi_{\rho]} \chi^\lambda &= -\chi_{[\rho} R_{\tau]}{}^{\sigma} \chi_\sigma \chi_\nu \end{aligned} \tag{7.168}$$

Sve zajedno vodi do

$$\chi_{[\rho} \nabla_{\tau]} \kappa = -\chi_{[\rho} R_{\tau]\sigma} \chi^\sigma \; . \tag{7.169}$$

Kako bi se pokazalo da desna strana izraza iščezava treba nam dominantni uvjet na energiju i Einsteinova jednadžba [71] iz čega onda slijedi nulti zakon. Navest ćemo još ostale verzije nultog zakona.

- Pokazali smo da je κ konstantna na horizontu stacionarne crne rupe ako vrijedi Einsteinova jednadžba i materija zadovoljava dominantni uvjet na energiju.
- Površinska gravitacija je konstantna ako je prostorvrijeme statično ili stacionarno i zadovoljava $t - \phi$ svojstvo ortogonalnosti [112].
- Površinska gravitacija κ stacionarne crne rupe konstantna je na bifurkacijskom Killingovom horizontu. Vrijedi i obrnuta tvrdnja [112].

Kad je $\kappa = 0$ kaže se da je horizont degeneriran, a crna rupa ekstremna. Ovaj slučaj nećemo razmatrati. Štoviše, iz nultog zakona slijedi da su jedini tipovi Killingovog horizonta u općoj teoriji relativnosti ili degenerirani ili bifurkacijski.

Nulti zakon je u analogiji s tvrdnjom da je temperatura tijela u ravnoteži stalna. Stoga, površinska gravitacija analogna je temperaturi. Bitno je napomenuti [136] kako nulti zakon termodinamike sadrži i tranzitivnost.

7.5.5 Prvi zakon

Prvi zakon povezuje promjenu parametara crne rupe ako ju perturbiramo.

$$\delta M = \frac{\kappa}{8\pi}\delta A + \Omega_H \delta J + \Phi_H \delta Q \ , \tag{7.170}$$

gdje je M masa crne rupe κ je površinska gravitacija, Ω_H je angularna brzina na horizontu, J je angularni moment crne rupe, Φ_H je električni potencijal na horizontu, a Q je naboj. Postoje dvije verzije dokaza. Jedna se u literaturi naziva ravnotežna, a druga fizikalna. U prvom slučaju uspoređujemo dvije infinitezimalno različite stacionarne crne rupe. Dokaz slijedi iz varijacije po parametrima. Osim toga, u ravnotežnu verziju spada metoda koja je slična Noetherinom pristupu.

U fizikalnoj verziji dokaza, kao što ime sugerira, promatramo fizikalni proces, tj. infinitezimalna količina materija prelazi horizont događaja. Ova verzija je “u duhu" najsličnija termodinamici. Kreće se od izraza za tok,

$$\delta M = -\int_{\mathcal{H}} \Delta T^{\mu}_{\nu} t^{\nu} d\Sigma_{\mu} \, , \tag{7.171}$$

$$\delta J = \int_{\mathcal{H}} \Delta T^{\mu}_{\nu} \phi^{\nu} d\Sigma_{\mu} \, . \tag{7.172}$$

Izraze iskombiniramo na način,

$$\begin{aligned} \delta M - \Omega_H \delta J &= \int_{\mathcal{H}} T_{\mu\nu}(t^{\nu} + \Omega_H \phi^{\nu})\chi^{\mu} dS d\lambda \\ &= \int d\lambda \int_{\mathcal{H}(\lambda)} T_{\mu\nu}\chi^{\mu}\chi^{\nu} dA \, . \end{aligned} \tag{7.173}$$

Element površine horizonta dan je s $d\Sigma_{\mu} = \chi^{\mu} dS d\lambda$. Kako bismo izvrijednili integral koristimo Raychaudhurijevu jednadžbu,

$$\frac{d\theta}{d\lambda} = \kappa\theta - 8\pi T_{\mu\nu}\chi^{\mu}\chi^{\nu} \, , \tag{7.174}$$

gdje smo iskoristili Einsteinovu jednadžbu i zanemarili više doprinose. Slijedi da je

$$\begin{aligned} \delta M - \Omega_H \delta J &= -\frac{1}{8\pi}\int d\lambda \int_{\mathcal{H}(\lambda)} \left(\frac{d\theta}{d\lambda} - \kappa\theta\right) dA \\ &= -\frac{1}{8\pi}\int d\lambda \int_{\mathcal{H}(\lambda)} \frac{d\theta}{d\lambda} dA + \frac{\kappa}{8\pi}\int d\lambda \int_{\mathcal{H}(\lambda)} \theta dA \, . \end{aligned} \tag{7.175}$$

Prvi integral je nula jer su početno i konačno stanje stacionarni. Da izvrijednimo drugi integral upotrijebit ćemo definiciju ekspanzije,

$$\theta = \frac{1}{dA}\frac{dA}{d\lambda} \implies \theta dA = \frac{dA}{d\lambda} \,, \tag{7.176}$$

Konačno, dobivamo da je

$$\delta M - \Omega_H \delta J = \frac{\kappa}{8\pi}\delta A \implies \delta M = \Omega_H \delta J + \frac{\kappa}{8\pi}\delta A \,. \tag{7.177}$$

Ovaj zakon analogan je prvom zakonu termodinamike,

$$\delta U = T\delta S - pdV \,. \tag{7.178}$$

Jer iz nultog zakona očekujemo da je κ analogna temperaturi, slijedi da je površina horizonta analogna entropiji. Kako bismo odredili konstantu proporcionalnosti moramo uključiti kvantne efekte u sliku.

7.5.6 Drugi zakon

Drugi zakon, za razliku od nultog i prvog zakona, nije vezan samo uz stacionarne crne rupe. Izjava drugog zakona jest da se površina horizonta ne može smanjiti

$$\delta A \geq 0 \,, \tag{7.179}$$

ako vrijedi hipoteza o kozmičkoj cenzuri ("cosmic censorship conjecture") i materija zadovoljava dominantni uvjet na energiju. Dokaz kreće od afino parametrizirane Raychaudhuriejeve jednadžbe koja zajedno s dominantnim uvjetom daje

$$\begin{gathered} \frac{d\theta}{d\tilde{\lambda}} \leq -\frac{1}{2}\theta^2 \\ \int_{\theta_0}^{\theta} \frac{d\theta}{\theta^2} \leq -\int_0^{\tilde{\lambda}} \frac{d\tilde{\lambda}}{2} \implies \frac{1}{\theta(\tilde{\lambda})} \geq \frac{1}{\theta_0} + \frac{\lambda}{2} \,, \end{gathered} \tag{7.180}$$

gdje je $la\tilde{m}bda$ afini parametar. Sljedeći korak je pokazati da svaki element površine horizonta a ima pozitivnu ekspanziju. Kao što se može vidjeti iz definicije ekspanzije

[119],

$$\frac{da}{d\tilde{\lambda}} = a\theta \ . \tag{7.181}$$

Dokaz ove tvrdnje slijed iz po kontradikciji sa svojstvima horizonta ako gornja tvrdnja ne vrijedi [103]. Štoviše, ako se crne rupe sjedine, mora vrijediti da je

$$A_3 > A_1 + A_2 \ . \tag{7.182}$$

Drugi zakon sličan je zakonu porasta entropije. Ovaj zakon ne vrijedi više kad se uzmu u obzir kvantni efekti jer oni ne zadovoljavaju dominantni uvjet.

7.5.7 Treći zakon

Treći zakon dokazan je u [120]. Verzija trećeg zakona koja je najprihvaćenija u literaturi kaže sljedeće.

Nijedan fizikalni proces ne može sniziti površinsku gravitaciju crne rupe na nulu u konačnom broju koraka.

Treći zakon potpuno je analogan termodinamičkom izričaju ako se temperatura zamijeni površinskom gravitacijom.

Ovime završavamo pregled klasičnih zakona mehanike crnih rupa. U sljedećem poglavlju promatrat ćemo poluklasičnu sliku i izvesti Hawkingovo zračenje.

7.6 Hawkingovo zračenje

Kad se uzmu u obzir kvantni efekti crne rupe zrače kao crna tijela na temperaturi T_H,

$$T_H = \frac{\kappa}{2\pi} \ . \tag{7.183}$$

Hawkingovo zračenje za generičke crne rupe je reda veličine mikrokelvin pa ne postoji eksperimentalna potvrda. Hawkingov originalni izvod [123] bazira se na činjenici da u zakrivljenom prostorvremenu ne postoji jedinstveni vakuum. Ovdje nećemo izlagati taj izvod. Umjesto toga, držat ćemo se “rukomahajućeg" obrazloženja ovog rezultata, prema kojem su za Hawkingovo zračenje odgovorne kvantne fluktuacije blizu horizonta crne rupe. Antičestica koja za vrijeme svojeg postojanja prijeđe horizont imat će negativnu energiju, dok će čestica koja je ostala izvan horizonta moći pobjeći u beskonačnost s pozitivnom energijom što opažamo kao Hawkingovo zračenje. Nadalje, kako bi energija bila očuvana, masa, tj. energija crne rupe se smanji što nakon dovoljno vremena dovodi do toga da crna rupa potpuno "ispari". Ovdje se nećemo baviti tim problem, već ćemo samo kratko izložiti kako iz ove slike možemo izvesti Hawkingovo zračenje.

Zamislimo da je par čestice i antičestice nastao unutar Schwarzschildove crne rupe. Klasično ništa ne može pobjeći, no kvantna fizika dozvoljava da se čestica tunelira u klasično zabranjeno područje. Jer se masa crne rupe smanji radijus horizonta se pomakne što predstavlja kvantnu barijeru. Iz WKB aproksimacije možemo izračunati emisiju, Γ,

$$\Gamma \propto e^{-2\mathrm{Im}S} , \tag{7.184}$$

gdje je veličina u eksponent definirana kao

$$\mathrm{Im}S = \mathrm{Im} \int_{r_i}^{r_f} p(r) dr . \tag{7.185}$$

Valna duljina čestice na horizontu pomaknuta je prema plavom. Drugim riječima, valna duljina čestice je mala što opravdava korištenje WKB aproksimacije. Ako pretpostavimo da je čestica foton koji se po putanji radijalnog geodezika (u Gullstrand-Painlevé koordinatama)

$$\frac{dr}{dt} = \pm 1 - \sqrt{\frac{2(M-\omega)}{r}} , \tag{7.186}$$

tunelira iz crne rupe, slijedi da je

$$\Gamma \propto e^{-8\pi M\omega(1-\frac{\omega}{2M})} , \tag{7.187}$$

odnosno, ako zanemarimo ω^2 član,

$$\kappa = \frac{1}{4M} \implies T = \frac{1}{8\pi M} = \frac{\kappa}{2\pi} . \tag{7.188}$$

Temperatura zračenja crne rupe je T_H. Valja primijetiti [46] da temperatura ne dolazi od mikroskopskih stupnjeva slobode same crne rupe, već od kvanata materije koji uspiju iztunelirati. U svakom slučaju, jer možemo govoriti o temperaturi crne rupe, nulti zakon mehanike crnih rupa dobiva tranzitivnu prirodu. Nadalje, usporedbom sa prvim zakonom možemo zaključiti da je $S = 1/4A$. Dakle, konstanta proporcionalnosti između entropije i površine horizonta je $1/4$. Ovaj izraz zove se Bekenstein-Hawking entropija. Osim toga, kao što smo već spomenuli, površina crne rupe se smanjuje, što znači da drugi zakon ne vrijedi. Iz tog razloga, a i nekih klasičnih razmatranja, predložen je generalizirani drugi zakon koji glasi

$$\delta(S + \frac{1}{4}A) \geq 0 \; . \tag{7.189}$$

Suma doprinosa entropije izvan crne rupe S i entropije crne rupe $1/4A$ se nikad ne smanjuje.

Od ostalih načina dobivanja Hawkingovog zračenja, zanimljiv je pristup koji se temelji na vezi kvantne amplitude i kanonske funkcije izvodnice. Za kvantni sustav na inverznoj temperaturi β, kanonska funkcija izvodnica je

$$Z_C = \sum_n \langle n|e^{-\beta H}|n\rangle = \sum_n e^{-\beta E_n} = \mathrm{Tr} e^{-\beta H} \; , \tag{7.190}$$

gdje je E_n energija stanja $|n\rangle$. S druge strane, Greenova funkcija je

$$G\left(q', t; q, 0\right) = \left\langle q' \left| e^{-iHt} \right| q \right\rangle \; . \tag{7.191}$$

Ako umjesto realnog promatramo kompleksno vrijeme $t = -i\beta$, dobivamo

$$\begin{aligned} G\left(q', -i\beta; q, 0\right) &= \left\langle q' \left| e^{-iH(-i\beta)} \right| q \right\rangle \\ &= \left\langle q' \left| e^{-\beta H} \sum_j \right| j \right\rangle \langle j||q\rangle \\ &= \sum_j e^{-\beta E_j} \left\langle q' \mid j \right\rangle \left\langle j \mid q \right\rangle \\ &= \sum_j e^{-\beta E_j} \left\langle j \mid q \right\rangle \left\langle q' \mid j \right\rangle \; . \end{aligned} \tag{7.192}$$

Za slučaj kad je $q = q'$ slijedi

$$\int dqG(q, -i\beta; q, 0) = \sum_j e^{-\beta E_j} \left\langle j| \underbrace{\int dq|q\rangle\langle q}_{=1} |j \right\rangle = Z_C \ . \tag{7.193}$$

Osim toga, može se pokazati da je Greenova funkcija periodička s periodom β,

$$\begin{aligned} G_\beta (q', t; q, 0) &\equiv \mathrm{Tr}\left(e^{-\beta H}\varphi(q', t)\varphi(q, 0)\right) \\ &= \mathrm{Tr}\left(e^{-\beta H}\varphi(q', t)e^{-\beta H}e^{\beta H}\varphi(q, 0)\right) \\ &= \mathrm{Tr}\left(\varphi(q', t)e^{-\beta H}e^{\beta H}\varphi(q, 0)\, e^{-\beta H}\right) \\ &= \mathrm{Tr}\left(\varphi(q', t)e^{-\beta H}\varphi(q, t + i\beta)\right) \\ &= G_\beta (q, 0; q', t + i\beta) \ . \end{aligned} \tag{7.194}$$

Dakle, polje koje živi u prostorvremenu s kompleksnim vremenom koje je periodičko, smatra da živi u termalnom spremniku inverzne temperature β.

Time smo zaključili razmatranje zakona termodinamike crnih rupa. Bitno je primjetiti da je ključnu ulogu u svemu igrao horizont događaja. Kako horizonti nisu striktno vezani uz crne rupe, postavlja se pitanje mogu li se ovi rezultati generalizirati na proizvoljne horizonte. Prije nego se vratimo na to pitanje, promotrit ćemo prostorvrijeme u kojem postoji Rindlerov horizont.

7.7 Unruhov efekt

Opažač koji se giba uniformnom akceleracijom a zove se Rindlerov opažač i opaža termalni spektar koji odgovara temperaturi T_U.

$$T_U = \frac{a}{2\pi} \ . \tag{7.195}$$

Prije nego što pokažemo kako se dođe do Unruhove temperature (spektar zračenja nećemo izvoditi), potrebno je objasniti kakvo prostorvrijeme opaža Rindlerov opažač.

Neka su koordinate Rindlerovog opažača (x, t), dok inercijalni opažač koristi koordinate (X, T). Veza između njih dana je s

$$X = x\cosh(at) \ , \quad T = x\sinh(at) \ , \quad x > |t| \ . \tag{7.196}$$

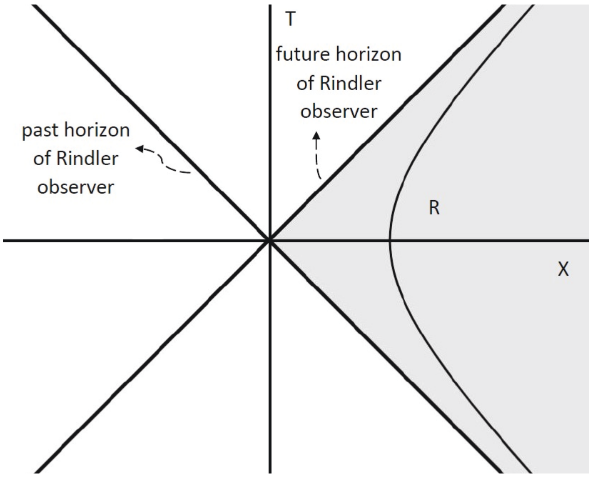
Prostornovremenski interval je

$$ds^2 = -dT^2 + dX^2 = -a^2x^2dt^2 + dx^2 \ . \tag{7.197}$$

U Rindlerovim koordinatama imamo Killingovo vektorsko polje

$$\partial_t = a(X\partial_T + T\partial_X) \ . \tag{7.198}$$

koje generira bifurkacijski horizont, prikazan na slici 7.3. Možemo prepoznati da se radi o Lorentzovom potisku. Kako bismo izveli Unruhovu temperaturu koristit ćemo



Slika 7.3: Dijagram ravnog prostorvremena. Hiperbola je putnja Rindlerovog opažača. Dio prostorvremena koji mu je dostupan označen je sivom bojom. Rindlerov horizont je bifurkacijski horizont generiran Lorentzovim potiskom.

vezu Greenove funkcije i kanonske funkcije izvodnice. Ako analitički produljimo vrijeme u kompleksnu ravninu na način da

$$\begin{aligned} &ds^2 = -dT^2 + dX^2 \ , X = a\cosh(at) \ , T = x\sinh(at) \xrightarrow{T=iT_E \ ,t=it_E} \\ &ds^2 = dT_E^2 + dX^2 \ , X = a\cos(at_E) \ , T = x\sin(at_E) \ , \end{aligned} \tag{7.199}$$

vidimo da t_E ima period $2\pi/a$. Stoga,

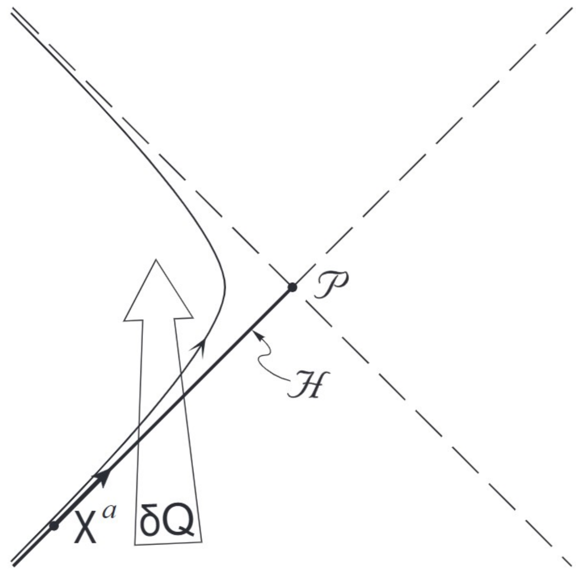
$$G_E(T_E, X) \equiv G_E(t_E, x) = G_E(t_E + \frac{2\pi}{a}, x) \ . \tag{7.200}$$

Nakon Wickove rotacije slijedi da Rindlerov opažač mjeri temperaturu koja odgovara Unruhovoj temperaturi [132].

7.8 Termodinamika prostorvremena

U ovom poglavlju pokazat ćemo da se Einsteinova jednadžba može izvesti iz zakona termodinamike i Raychaudhurijeve jednadžbe.

Najprije definiramo sustav tako da promotrimo proizvoljnu točku p iz perspektive Rindlerovog opažača. Odaberemo jednu stranu horizonta kao što je prikazano na slici 7.4. Sustav je onda lijevi “stožač". Jer je prostor lokalno ravan u okolini točke p



Slika 7.4: Lijeva strana prostorvremena je sustav u koji ulazi toplina δQ. Podebljana crta je horizont [65].

imamo aproksimativno Killingovo vektorsko polje χ koje odaberemo tako da iščezava u p. Geodezici koji generiraju horizon parametrizirani su afinim parametrom λ i k je vektor svjetlosnog tipa tangentan na geodezike. Pomoću ove konstrukcije možemo definirati tok topline, dan sljedećom relacijom

$$\delta Q = \kappa \int T_{\mu\nu}(-\lambda) k^\mu k^\nu d\lambda dA\,. \tag{7.201}$$

Iz analogije zakona termodinamike znamo da vrijedi

$$\delta S = \alpha \delta A\,, \tag{7.202}$$

gdje je α konstanta proporcionalnosti. S druge strane, pomoću Raychaudhurijeve jednadžbe možemo dobiti

$$\theta = -\lambda R_{\mu\nu} k^\mu k^\nu \,, \tag{7.203}$$

Nadalje, iz definicije ekspanzije imamo

$$\delta A = \int_{\mathcal{H}} \theta d\lambda dA \,, \tag{7.204}$$

gdje smo pretpostavili da horizont u trenutnoj ravnoteži. Ove relacije možemo povezati pomoću Clausiusove relacije,

$$\delta Q = T\delta S = \frac{\kappa}{2\pi}\alpha\delta A \,, \tag{7.205}$$

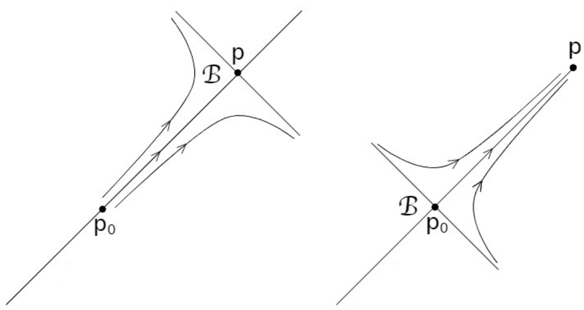
Iz čega slijedi Einsteinova jednadžba.

$$T_{\mu\nu} k^\mu k^\nu = \frac{1}{2\pi}\alpha R_{\mu\nu} k^\mu k^\nu \implies \frac{2\pi}{\alpha} T_{\mu\nu} = R_{\mu\nu} + f g_{\mu\nu} \,. \tag{7.206}$$

Funkcija f odredi se pomoću zakona očuvanja tenzora energije i impulsa i Bianchievog identiteta.

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{2\pi}{\alpha} T_{\mu\nu} \,. \tag{7.207}$$

Isto rezultat dobije se za desni stožac ako točku u kojoj Killingovo polje iščezava pomaknemo u prošlost (slika 7.5).



Slika 7.5: Lijevo je “stari", a desno novi postav. Killingovo polje iščezava u točki p_0 [141].

Pokazali smo da se Einsteinova jednadžba može izvesti iz drugog zakona termodinamike, Raychaudhurijeve jednadžbe i Clausiusove relacije. Stoga se može reći da je Einsteinova jednadžba u neku ruku jednadžba stanja.

7.9 *Sažetak i osvrt*

7.9.1 Sažetak

U radu smo razmatrali indicije koje upućuju na to da je gravitacija izranjajući fenomen — univerzalna priroda gravitacije, perturbativna ne-renormalizabilnost i analogija zakona mehanike crnih rupa sa zakonima termodinamike.

S druge strane, razvijeni su mnogi modeli, temeljeni na analogiji s fizikom kondenzirane tvari, koji uspješno reproduciraju neka svojstva gravitacije. U radu je dan pregled različitih modela i neka ograničenja koja modeli moraju zadovoljavati kako bi mogli rekonstruirati opću teoriju relativnosti na makroskopskim skalama.

Jedan od takvih modela je teorija kauzalnih skupova gdje je prostorvrijeme diskretno. Na takvoj skali preživljavaju kauzalni odnosi između točaka kauzalnog skupa što je dovoljno da odredimo metriku do na konformalni faktor. Nakon što fiksiramo i volumen dobije se jedinstvena metrika.

Ograničenje na modele izranjajuće gravitacije jest Weinberg-Wittenov teorem koji postavlja ograničenje na vrste čestica koje mogu postojati na ravnom prostorvremenu. Štoviše, kao posljedica teorema graviton ne može biti kompozitna čestica.

Još jedan pravac proizlazi iz analogije termodinamike i opće teorije relativnosti. Izveli smo zakone mehanike crnih rupa i pokazali da su analogni zakonima termodinamike. Veza je učvršćena otkrićem Hawkingovog zračenja. Istaknuli smo neke problematične aspekte analogije. Naposljetku izveli smo Einsteinovu jednadžbu stanja iz Clausiusove relacije, drugog zakona termodinamike i Raychaudhurijeve jednadžbe, čime je analogija proširena izvan domene crnih rupa.

7.9.2 Osvrt

Većina emergentnih modela daleko je od toga da reproducira opću teoriju relativnosti. Iako postoje indicije da gravitacija jest emergentna pojava, nije jasno koliko se ozbiljno postojeće analogije trebaju shvatiti.

Bibliography

[1] Niels S. Linnemann and Manus R. Visser. "Hints towards the emergent nature of gravity". In: *Stud. Hist. Phil. Sci. B* 64 (2018), pp. 1–13. DOI: 10.1016/j.shpsb.2018.04.001. arXiv: 1711.10503 [physics.hist-ph].

[2] S. Carlip. "Challenges for Emergent Gravity". In: *Stud. Hist. Phil. Sci. B* 46 (2014), pp. 200–208. DOI: 10.1016/j.shpsb.2012.11.002. arXiv: 1207.2504 [gr-qc].

[3] Kip S. Thorne, David L. Lee, and Alan P. Lightman. "Foundations for a Theory of Gravitation Theories". In: *Phys. Rev. D* 7 (12 1973), pp. 3563–3578. DOI: 10.1103/PhysRevD.7.3563. URL: https://link.aps.org/doi/10.1103/PhysRevD.7.3563.

[4] Jürgen Ehlers, Felix Pirani, and Alfred Schild. "Republication of: The geometry of free fall and light propagation". In: *General Relativity and Gravitation* 44 (June 2012). DOI: 10.1007/s10714-012-1353-4.

[5] Éric Gourgoulhon. *Special Relativity in General Frames*. Graduate Texts in Physics. Berlin, Heidelberg: Springer, 2013. DOI: 10.1007/978-3-642-37276-6.

[6] J. Ehlers and A. Schild. "Geometry in a manifold with projective structure". In: *Comm. Math. Phys.* 32.2 (1973), pp. 119–146. URL: https://projecteuclid.org:443/euclid.cmp/1103859104.

[7] S. Capozziello et al. "The physical foundations for the geometric structure of relativistic theories of gravitation. From General Relativity to Extended Theories of Gravity through Ehlers-Pirani-Schild approach". In: *Int. J. Geom. Meth. Mod. Phys.* 9 (2012), p. 1250072. DOI: 10.1142/S0219887812500727. arXiv: 1202.5699 [gr-qc].

[8] Clifford M. Will. "The Confrontation between general relativity and experiment". In: *Living Rev. Rel.* 9 (2006), p. 3. DOI: 10.12942/lrr-2006-3. arXiv: gr-qc/0510072.

[9] Matt Visser, Carlos Barcelo, and Stefano Liberati. "Birefringence versus bimetricity". In: 2003. arXiv: gr-qc/0204017.

[10] Jonas Schneider et al. “Gravitational closure of weakly birefringent electrodynamics”. In: (Aug. 2017). arXiv: 1708.03870 [hep-th].

[11] Maximilian Düll et al. “Gravitational closure of matter field equations”. In: *Phys. Rev. D* 97.8 (2018), p. 084036. DOI: 10.1103/PhysRevD.97.084036. arXiv: 1611.08878 [gr-qc].

[12] Alexander B. Balakin and Winfried Zimdahl. “Optical metrics and birefringence of anisotropic media”. In: *Gen. Rel. Grav.* 37 (2005), pp. 1731–1751. DOI: 10.1007/s10714-005-0155-3. arXiv: gr-qc/0504051.

[13] G.M. Shore. “Causality and superluminal light”. In: *Time and Matter: An International Colloquium on the Science of Time*. Feb. 2003, pp. 45–66. arXiv: gr-qc/0302116.

[14] Frederic P. Schuller. “All spacetimes beyond Einstein (Obergurgl Lectures)”. In: (Nov. 2011). arXiv: 1111.4824 [gr-qc].

[15] Kristina Giesel et al. “Gravitational dynamics for all tensorial spacetimes carrying predictive, interpretable and quantizable matter”. In: *Phys. Rev. D* 85 (2012), p. 104042. DOI: 10.1103/PhysRevD.85.104042. arXiv: 1202.2991 [gr-qc].

[16] D.N. Spergel et al. “Wilkinson Microwave Anisotropy Probe (WMAP) three year results: implications for cosmology”. In: *Astrophys. J. Suppl.* 170 (2007), p. 377. DOI: 10.1086/513700. arXiv: astro-ph/0603449.

[17] Alan H. Guth. “The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems”. In: *Adv. Ser. Astrophys. Cosmol.* 3 (1987). Ed. by Li-Zhi Fang and R. Ruffini, pp. 139–148. DOI: 10.1103/PhysRevD.23.347.

[18] George Lazarides. “Basics of inflationary cosmology”. In: *J. Phys. Conf. Ser.* 53 (2006). Ed. by K. Anagnostopoulos et al., pp. 528–550. DOI: 10.1088/1742-6596/53/1/033. arXiv: hep-ph/0607032.

[19] P.W. Anderson and Jeffrey Goldstein. “More is different: Broken symmetry and the nature of the hierarchical structure of science”. In: 16 (Sept. 2014), pp. 117–134.

[20] Xiao-Gang Wen. “Quantum Field Theory of Many-body Systems – from the Origin of Sound to an Origin of Light and Fermions”. In: 2004.

[21] Nathan Seiberg. "Emergent spacetime". In: *23rd Solvay Conference in Physics: The Quantum Structure of Space and Time*. Jan. 2006, pp. 163–178. DOI: 10.1142/9789812706768_0005. arXiv: hep-th/0601234.

[22] Lorenzo Sindoni. "Emergent Models for Gravity: an Overview of Microscopic Models". In: *SIGMA* 8 (2012), p. 027. DOI: 10.3842/SIGMA.2012.027. arXiv: 1110.0686 [gr-qc].

[23] B.L. Hu. "Emergent/Quantum Gravity: Macro/Micro Structures of Spacetime". In: *J. Phys. Conf. Ser.* 174 (2009). Ed. by Lajos Diosi et al., p. 012015. DOI: 10.1088/1742-6596/174/1/012015. arXiv: 0903.0878 [gr-qc].

[24] Nick Huggett and Christian Wüthrich. "Emergent spacetime and empirical (in)coherence". In: *Stud. Hist. Phil. Sci. B* 44 (2013), pp. 276–285. DOI: 10.1016/j.shpsb.2012.11.003. arXiv: 1206.6290 [physics.hist-ph].

[25] Fay Dowker. "Causal sets and the deep structure of spacetime". In: *100 Years Of Relativity: space-time structure: Einstein and beyond*. Ed. by Abhay Ashtekar. 2005, pp. 445–464. DOI: 10.1142/9789812700988_0016. arXiv: gr-qc/0508109.

[26] T. Padmanabhan. "From gravitons to gravity: Myths and reality". In: *Int. J. Mod. Phys. D* 17 (2008), pp. 367–398. DOI: 10.1142/S0218271808012085. arXiv: gr-qc/0409089.

[27] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5.

[28] J. Ambjorn, J. Jurkiewicz, and R. Loll. "Reconstructing the universe". In: *Phys. Rev. D* 72 (2005), p. 064014. DOI: 10.1103/PhysRevD.72.064014. arXiv: hep-th/0505154.

[29] Herbert W. Hamber. "Quantum Gravity on the Lattice". In: *Gen. Rel. Grav.* 41 (2009), pp. 817–876. DOI: 10.1007/s10714-009-0769-y. arXiv: 0901.0964 [gr-qc].

[30] Roberto Percacci and Gian Paolo Vacca. "Asymptotic Safety, Emergence and Minimal Length". In: *Class. Quant. Grav.* 27 (2010), p. 245026. DOI: 10.1088/0264-9381/27/24/245026. arXiv: 1008.3621 [hep-th].

[31] Astrid Eichhorn. "Asymptotically safe gravity". In: *57th International School of Subnuclear Physics: In Search for the Unexpected*. Feb. 2020. arXiv: 2003.00044 [gr-qc].

[32] Jonathan Bain. "The emergence of spacetime in condensed matter approaches to quantum gravity". In: *Stud. Hist. Phil. Sci. B* 44 (2013), pp. 338–345. DOI: 10.1016/j.shpsb.2012.05.001.

[33] Erik P. Verlinde. "On the Origin of Gravity and the Laws of Newton". In: *JHEP* 04 (2011), p. 029. DOI: 10.1007/JHEP04(2011)029. arXiv: 1001.0785 [hep-th].

[34] R.P. Feynman. *Feynman lectures on gravitation*. Ed. by F.B. Morinigo, W.G. Wagner, and B. Hatfield. Dec. 1996.

[35] A.D. Sakharov. "Vacuum quantum fluctuations in curved space and the theory of gravitation". In: *Usp. Fiz. Nauk* 161.5 (1991). Ed. by Yu.A. Trutnev, pp. 64–66. DOI: 10.1070/PU1991v034n05ABEH002498.

[36] Andrei D Sakharov. "Spectral density of eigenvalues of the wave equation and vacuum polarization". In: *Soviet Physics Uspekhi* 34.5 (1991), pp. 395–400. DOI: 10.1070/pu1991v034n05abeh002499. URL: https://doi.org/10.1070/pu1991v034n05abeh002499.

[37] Matt Visser. "Sakharov's induced gravity: A Modern perspective". In: *Mod. Phys. Lett. A* 17 (2002). Ed. by Dharam Vir Ahluwalia and N.K. Dadhich, pp. 977–992. DOI: 10.1142/S0217732302006886. arXiv: gr-qc/0204062.

[38] Thanu Padmanabhan. "Exploring the Nature of Gravity". In: *35th Max Born Symposium: The Planck Scale II*. Feb. 2016. arXiv: 1602.01474 [gr-qc].

[39] C.J. Isham. "Structural issues in quantum gravity". In: *14TH INTERNATIONAL CONFERENCE ON GENERAL RELATIVITY AND GRAVITATION (GR14)*. Aug. 1995, pp. 167–209. arXiv: gr-qc/9510063.

[40] Marc H. Goroff, Augusto Sagnotti, and Augusto Sagnotti. "Quantum gravity at two loops". In: *Physics Letters B* 160.1 (1985), pp. 81 –86. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(85)91470-4. URL: http://www.sciencedirect.com/science/article/pii/0370269385914704.

[41] A. Zee. *Quantum field theory in a nutshell*. Nov. 2003. ISBN: 978-0-691-14034-6.

[42] Tian Cao and Silvan Schweber. "The conceptual foundations and the philosophical aspects of renormalization theory". In: *Synthese* 97 (Oct. 1993), pp. 33–108. DOI: 10.1007/BF01255832.

[43] G.Peter Lepage. "What is Renormalization?" In: *Theoretical Advanced Study Institute in Elementary Particle Physics*. June 1989, pp. 483–508. arXiv: hep-ph/0506330.

[44] C.P. Burgess. "Quantum gravity in everyday life: General relativity as an effective field theory". In: *Living Rev. Rel.* 7 (2004), pp. 5–56. DOI: 10.12942/lrr-2004-5. arXiv: gr-qc/0311082.

[45] James M. Bardeen, B. Carter, and S. W. Hawking. "The Four laws of black hole mechanics". In: *Commun. Math. Phys.* 31 (1973), pp. 161–170. DOI: 10.1007/BF01645742.

[46] Erik Curiel. "The many definitions of a black hole". In: *Nature Astron.* 3.1 (2019), pp. 27–34. DOI: 10.1038/s41550-018-0602-1. arXiv: 1808.01507 [physics.hist-ph].

[47] Steven B. Giddings. "Hawking radiation, the Stefan–Boltzmann law, and unitarization". In: *Phys. Lett. B* 754 (2016), pp. 39–42. DOI: 10.1016/j.physletb.2015.12.076. arXiv: 1511.08221 [hep-th].

[48] W. G. Unruh. "Sonic analogue of black holes and the effects of high frequencies on black hole evaporation". In: *Phys. Rev. D* 51 (6 1995), pp. 2827–2838. DOI: 10.1103/PhysRevD.51.2827. URL: https://link.aps.org/doi/10.1103/PhysRevD.51.2827.

[49] Carlos Barcelo, Stefano Liberati, and Matt Visser. "Analogue gravity". In: *Living Rev. Rel.* 8 (2005), p. 12. DOI: 10.12942/lrr-2005-12. arXiv: gr-qc/0505065.

[50] Grigory E. Volovik. "Emergent physics: Fermi point scenario". In: *Phil. Trans. Roy. Soc. Lond. A* 366 (2008), pp. 2935–2951. DOI: 10.1098/rsta.2008.0070. arXiv: 0801.0724 [gr-qc].

[51] G.E. Volovik. "Field theory in superfluid He-3: What are the lessons for particle physics, gravity, and high temperature superconductivity?" In: *Proc. Nat. Acad. Sci.* 96 (1999), pp. 6042–6047. DOI: 10.1073/pnas.96.11.6042. arXiv: cond-mat/9812381.

[52] Shou-Cheng Zhang and Jiang-ping Hu. "A Four-dimensional generalization of the quantum Hall effect". In: *Science* 294 (2001), p. 823. DOI: 10.1126/science.294.5543.823. arXiv: cond-mat/0110572.

[53] H. Kleinert. "Gravity as a Theory of Defects in a Crystal with Only Second Gradient Elasticity". In: *Annalen der Physik* 499 (1987), pp. 117–119.

[54] H. Kleinert and J. Zaanen. "World Nematic Crystal Model of Gravity Explaining the Absence of Torsion". In: 2003.

[55] Zheng-Cheng Gu and Xiao-Gang Wen. "Emergence of helicity +- 2 modes (gravitons) from qbit models". In: *Nucl. Phys. B* 863 (2012), pp. 90–129. DOI: 10.1016/j.nuclphysb.2012.05.010. arXiv: 0907.1203 [gr-qc].

[56] Tomasz Konopka, Fotini Markopoulou, and Lee Smolin. "Quantum Graphity". In: (Nov. 2006). arXiv: hep-th/0611197.

[57] Tomasz Konopka, Fotini Markopoulou, and Simone Severini. "Quantum Graphity: A Model of emergent locality". In: *Phys. Rev. D* 77 (2008), p. 104029. DOI: 10.1103/PhysRevD.77.104029. arXiv: 0801.0861 [hep-th].

[58] Aron C. Wall. "A discrete, unitary, causal theory of quantum gravity". In: *Class. Quant. Grav.* 30 (2013), p. 115002. DOI: 10.1088/0264-9381/30/11/115002. arXiv: 1201.2489 [gr-qc].

[59] Daniele Oriti. "The Group field theory approach to quantum gravity". In: (July 2006), pp. 310–331. arXiv: gr-qc/0607032.

[60] Robert de Mello Koch and Jeff Murugan. "Emergent Spacetime". In: *Foundations of Space and Time: Reflections on Quantum Gravity*. Nov. 2009, pp. 164–184. arXiv: 0911.4817 [hep-th].

[61] Stephen L. Adler. "Induced Gravitation". In: *The High-Energy Limit*. Ed. by Antonino Zichichi. Boston, MA: Springer US, 1983, pp. 3–27. ISBN: 978-1-4613-3506-1. DOI: 10.1007/978-1-4613-3506-1_2. URL: https://doi.org/10.1007/978-1-4613-3506-1_2.

[62] Stephen L. Adler. “Einstein Gravity as a Symmetry-Breaking Effect in Quantum Field Theory”. In: *Rev. Mod. Phys.* 54 (1982). [Erratum: Rev.Mod.Phys. 55, 837 (1983)], p. 729. DOI: 10.1103/RevModPhys.54.729.

[63] A. Zee. “Spontaneously generated gravity”. In: *Phys. Rev. D* 23 (4 1981), pp. 858–866. DOI: 10.1103/PhysRevD.23.858. URL: https://link.aps.org/doi/10.1103/PhysRevD.23.858.

[64] A. Zee. “Broken-Symmetric Theory of Gravity”. In: *Phys. Rev. Lett.* 42 (7 1979), pp. 417–421. DOI: 10.1103/PhysRevLett.42.417. URL: https://link.aps.org/doi/10.1103/PhysRevLett.42.417.

[65] Ted Jacobson. “Thermodynamics of space-time: The Einstein equation of state”. In: *Phys. Rev. Lett.* 75 (1995), pp. 1260–1263. DOI: 10.1103/PhysRevLett.75.1260. arXiv: gr-qc/9504004.

[66] T. Padmanabhan. “Gravity and/is Thermodynamics”. In: *Curr. Sci.* 109 (2015), pp. 2236–2242. DOI: 10.18520/v109/i12/2236-2242. arXiv: 1512.06546 [gr-qc].

[67] David Mattingly. “Modern tests of Lorentz invariance”. In: *Living Rev. Rel.* 8 (2005), p. 5. DOI: 10.12942/lrr-2005-5. arXiv: gr-qc/0502097.

[68] V. Berzi and V. Gorini. “Reciprocity principle and the lorentz transformations”. In: *J. Math. Phys.* 10 (1969), pp. 1518–1524. DOI: 10.1063/1.1665000.

[69] Rafael D. Sorkin. “Forks in the road, on the way to quantum gravity”. In: *Int. J. Theor. Phys.* 36 (1997), pp. 2759–2781. DOI: 10.1007/BF02435709. arXiv: gr-qc/9706002.

[70] Djamel Dou. “Causal sets, a possible interpretation for the black hole entropy and related topics”. PhD thesis. SISSA, Trieste, 1999. arXiv: gr-qc/0106024.

[71] Robert M. Wald. *General Relativity*. Chicago, USA: Chicago Univ. Pr., 1984. DOI: 10.7208/chicago/9780226870373.001.0001.

[72] Samuel Fletcher. “Relativistic Spacetime Structure”. In: (Aug. 2019).

[73] Fay Dowker. “Introduction to causal sets and their phenomenology”. In: *Gen. Rel. Grav.* 45.9 (2013), pp. 1651–1667. DOI: 10.1007/s10714-013-1569-y.

[74] Rafael D. Sorkin. “Causal sets: Discrete gravity”. In: *School on Quantum Gravity*. Sept. 2003, pp. 305–327. DOI: 10.1007/0-387-24992-3_7. arXiv: gr-qc/0309009.

[75] Luca Bombelli et al. “Space-time as a causal set”. In: *Phys. Rev. Lett.* 59 (5 1987), pp. 521–524. DOI: 10.1103/PhysRevLett.59.521. URL: https://link.aps.org/doi/10.1103/PhysRevLett.59.521.

[76] D.P. Rideout and R.D. Sorkin. “A Classical sequential growth dynamics for causal sets”. In: *Phys. Rev. D* 61 (2000), p. 024002. DOI: 10.1103/PhysRevD.61.024002. arXiv: gr-qc/9904062.

[77] Steven Weinberg and Edward Witten. “Limits on Massless Particles”. In: *Phys. Lett. B* 96 (1980), pp. 59–62. DOI: 10.1016/0370-2693(80)90212-9.

[78] Xavier Bekaert and Nicolas Boulanger. “The Unitary representations of the Poincare group in any spacetime dimension”. In: *2nd Modave Summer School in Theoretical Physics*. Nov. 2006. arXiv: hep-th/0611263.

[79] Eugene P. Wigner. “On Unitary Representations of the Inhomogeneous Lorentz Group”. In: *Annals Math.* 40 (1939). Ed. by Y.S. Kim and W.W. Zachary, pp. 149–204. DOI: 10.2307/1968551.

[80] W.K. Tung. *Group Theory in Physics*. World Scientific, 1985. ISBN: 9789971966560. URL: https://books.google.hr/books?id=O89tgpOBO04C.

[81] Steven Weinberg. *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, June 2005. ISBN: 978-0-521-67053-1, 978-0-511-25204-4.

[82] Jan B. Gutowski. *Lecture notes on symmetries and particle physics*. 2007. URL: http://personal.maths.surrey.ac.uk/st/jg0033/Resources/lectnotes(master).pdf.

[83] Katherine Brading and Harvey R. Brown. “Noether’s theorems and gauge symmetries”. In: (Aug. 2000). arXiv: hep-th/0009058.

[84] Domenico Giulini. "Laue's theorem revisited: Energy–momentum tensors, symmetries, and the habitat of globally conserved quantities". In: *Int. J. Geom. Meth. Mod. Phys.* 15.supp01 (2018), p. 1850182. DOI: 10.1142/S0219887818501827. arXiv: 1808.09320 [math-ph].

[85] Emmy Noether. "Invariant variation problems". In: *Transport Theory and Statistical Physics* 1.3 (1971), 186–207. ISSN: 1532-2424. DOI: 10.1080/00411457108231446. URL: http://dx.doi.org/10.1080/00411457108231446.

[86] Katherine A. Brading. "Which symmetry? Noether, Weyl, and conservation of electric charge". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 33.1 (2002), pp. 3 – 22. ISSN: 1355-2198. DOI: https://doi.org/10.1016/S1355-2198(01)00033-8. URL: http://www.sciencedirect.com/science/article/pii/S1355219801000338.

[87] Máximo Bañados and Ignacio A. Reyes. "A short review on Noether's theorems, gauge symmetries and boundary terms". In: *Int. J. Mod. Phys. D* 25.10 (2016), p. 1630021. DOI: 10.1142/S0218271816300214. arXiv: 1601.03616 [hep-th].

[88] Michael Forger and Hartmann Romer. "Currents and the energy momentum tensor in classical field theory: A Fresh look at an old problem". In: *Annals Phys.* 309 (2004), pp. 306–389. DOI: 10.1016/j.aop.2003.08.011. arXiv: hep-th/0307199.

[89] Steven Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. New York: John Wiley and Sons, 1972. ISBN: 978-0-471-92567-5, 978-0-471-92567-5.

[90] S. Weinberg. "Photons and Gravitons in S-Matrix Theory: Derivation of Charge Conservation and Equality of Gravitational and Inertial Mass". In: *Physical Review* 135 (1964), pp. 338–345.

[91] L.D. Landau and E.M. Lifschits. *The Classical Theory of Fields*. Vol. Volume 2. Course of Theoretical Physics. Oxford: Pergamon Press, 1975. ISBN: 978-0-08-018176-9.

[92] P.A.M. Dirac. *General Theory of Relativity*. General Theory of Relativity. Princeton University Press, 1996. ISBN: 9780691011462. URL: https://books.google.hr/books?id=Jtu-n8AzJhMC.

[93] John F. Donoghue, Mikhail M. Ivanov, and Andrey Shkerin. "EPFL Lectures on General Relativity as a Quantum Field Theory". In: (Feb. 2017). arXiv: 1702.00319 [hep-th].

[94] Ivica Smolić. "Spacetimes dressed with stealth electromagnetic fields". In: *Phys. Rev. D* 97.8 (2018), p. 084041. DOI: 10.1103/PhysRevD.97.084041. arXiv: 1711.07490 [gr-qc].

[95] Florian Loebbert. "The Weinberg-Wirten theorem on massless particles: An essay". In: *Annalen der Physik* 17 (Sept. 2008), pp. 803 –829. DOI: 10.1002/andp.200810305.

[96] Alejandro Jenkins. "Constraints on emergent gravity". In: *Int. J. Mod. Phys. D* 18 (2009), pp. 2249–2255. DOI: 10.1142/S0218271809015941. arXiv: 0904.0453 [gr-qc].

[97] Alejandro Jenkins. "Topics in theoretical particle physics and cosmology beyond the standard model". Other thesis. May 2006. arXiv: hep-th/0607239.

[98] Robert M. Wald. "The thermodynamics of black holes". In: *Living Rev. Rel.* 4 (2001), p. 6. DOI: 10.12942/lrr-2001-6. arXiv: gr-qc/9912119.

[99] Erik Curiel. "Singularities and Black Holes". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2021. Metaphysics Research Lab, Stanford University, 2021.

[100] C. DeWitt et al. *Black Holes*. Houches Ecole d'ete de Physique Theorique. Gordon and Breach, 1973. ISBN: 9780677156101. URL: https://books.google.hr/books?id=sUr-EVqZLckC.

[101] Eric Poisson. *A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics*. Cambridge University Press, Dec. 2009. DOI: 10.1017/CBO9780511606601.

[102] Geoffrey Compere. "An introduction to the mechanics of black holes". In: *2nd Modave Summer School in Theoretical Physics*. Nov. 2006. arXiv: gr-qc/0611129.

[103] S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1973. DOI: 10.1017/CBO9780511524646.

[104] J. Earman. *Bangs, crunches, whimpers, and shrieks: Singularities and acausalities in relativistic space-times*. 1995.

[105] Aron C. Wall. "The Generalized Second Law implies a Quantum Singularity Theorem". In: *Class. Quant. Grav.* 30 (2013). [Erratum: Class.Quant.Grav. 30, 199501 (2013)], p. 165003. DOI: 10.1088/0264-9381/30/19/199501. arXiv: 1010.5513 [gr-qc].

[106] Bryce S. DeWitt. "Quantum Theory of Gravity. I. The Canonical Theory". In: *Phys. Rev.* 160 (5 1967), pp. 1113–1148. DOI: 10.1103/PhysRev.160.1113. URL: https://link.aps.org/doi/10.1103/PhysRev.160.1113.

[107] S. Droz, W. Israel, and S. M. Morsink. "Black holes: The inside story". In: *Phys. World* 9N1 (1996), pp. 34–37.

[108] By Wahiba Toubal. "No-Hair Theorems and introduction to Hairy Black Holes". In: 2010.

[109] N. E. Mavromatos. "Eluding the no hair conjecture for black holes". In: *5th Hellenic School and Workshops on Elementary Particle Physics*. Sept. 1995. arXiv: gr-qc/9606008.

[110] Sudipta Sarkar. "Black Hole Thermodynamics: General Relativity and Beyond". In: *Gen. Rel. Grav.* 51.5 (2019), p. 63. DOI: 10.1007/s10714-019-2545-y. arXiv: 1905.04466 [hep-th].

[111] Sean M. Carroll. *Spacetime and Geometry*. Cambridge University Press, July 2019. ISBN: 978-0-8053-8732-2, 978-1-108-48839-6, 978-1-108-77555-7.

[112] Istvan Racz and Robert M. Wald. "Global extensions of space-times describing asymptotic final states of black holes". In: *Class. Quant. Grav.* 13 (1996), pp. 539–553. DOI: 10.1088/0264-9381/13/3/017. arXiv: gr-qc/9507055.

[113] James Lucietti and Harvey S. Reall. "Gravitational instability of an extreme Kerr black hole". In: *Phys. Rev. D* 86 (2012), p. 104030. DOI: 10.1103/PhysRevD.86.104030. arXiv: 1208.1437 [gr-qc].

[114] Sergio Dain and Ivan Gentile de Austria. "On the linear stability of the extreme Kerr black hole under axially symmetric perturbations". In: *Class. Quant. Grav.* 31.19 (2014), p. 195009. DOI: 10.1088/0264-9381/31/19/195009. arXiv: 1402.2848 [gr-qc].

[115] Robert M. Wald. *Quantum Field Theory in Curved Space-Time and Black Hole Thermodynamics*. Chicago Lectures in Physics. Chicago, IL: University of Chicago Press, 1995. ISBN: 978-0-226-87027-4.

[116] Robert M. Wald. "Black hole entropy is the Noether charge". In: *Phys. Rev. D* 48.8 (1993), pp. 3427–3431. DOI: 10.1103/PhysRevD.48.R3427. arXiv: gr-qc/9307038.

[117] S. W. Hawking and J. B. Hartle. "Energy and angular momentum flow into a black hole". In: *Commun. Math. Phys.* 27 (1972), pp. 283–290. DOI: 10.1007/BF01645515.

[118] S. W. Hawking. "Black holes in general relativity". In: *Commun. Math. Phys.* 25 (1972), pp. 152–166. DOI: 10.1007/BF01877517.

[119] P. K. Townsend. "Black holes: Lecture notes". In: (July 1997). arXiv: gr-qc/9707012.

[120] W. Israel. "Third Law of Black-Hole Dynamics: A Formulation and Proof". In: *Phys. Rev. Lett.* 57 (4 1986), pp. 397–399. DOI: 10.1103/PhysRevLett.57.397. URL: https://link.aps.org/doi/10.1103/PhysRevLett.57.397.

[121] B.G. Kyle. *Chemical and Process Thermodynamics*. Chemical and Process Thermodynamics. Prentice Hall PTR, 1999. ISBN: 9780130812445. URL: https://books.google.hr/books?id=hCApAAAAYAAJ.

[122] Robert F. Sekerka. "4 - Third Law of Thermodynamics". In: *Thermal Physics*. Ed. by Robert F. Sekerka. Amsterdam: Elsevier, 2015, pp. 49–52. ISBN: 978-0-12-803304-3. DOI: https://doi.org/10.1016/B978-0-12-803304-3.00004-1. URL: https://www.sciencedirect.com/science/article/pii/B9780128033043000041.

[123] S. W. Hawking. "Particle Creation by Black Holes". In: *Commun. Math. Phys.* 43 (1975). Ed. by G. W. Gibbons and S. W. Hawking. [Erratum: Commun.Math.Phys. 46, 206 (1976)], pp. 199–220. DOI: 10.1007/BF02345020.

[124] Jennie H. Traschen. “An Introduction to black hole evaporation”. In: *1999 Londrona Winter School on Mathematical Methods in Physics*. Aug. 1999. arXiv: gr-qc/0010055.

[125] Maulik K. Parikh and Frank Wilczek. “Hawking radiation as tunneling”. In: *Phys. Rev. Lett.* 85 (2000), pp. 5042–5045. DOI: 10.1103/PhysRevLett.85.5042. arXiv: hep-th/9907001.

[126] Maulik K. Parikh. “A Secret tunnel through the horizon”. In: *Gen. Rel. Grav.* 36 (2004), pp. 2419–2422. DOI: 10.1142/S0218271804006498. arXiv: hep-th/0405160.

[127] Per Kraus and Frank Wilczek. “Selfinteraction correction to black hole radiance”. In: *Nucl. Phys. B* 433 (1995), pp. 403–420. DOI: 10.1016/0550-3213(94)00411-7. arXiv: gr-qc/9408003.

[128] L. Vanzo, G. Acquaviva, and R. Di Criscienzo. “Tunnelling Methods and Hawking’s radiation: achievements and prospects”. In: *Class. Quant. Grav.* 28 (2011), p. 183001. DOI: 10.1088/0264-9381/28/18/183001. arXiv: 1106.4153 [gr-qc].

[129] Ryan Kerner and Robert B. Mann. “Tunnelling, temperature and Taub-NUT black holes”. In: *Phys. Rev. D* 73 (2006), p. 104010. DOI: 10.1103/PhysRevD.73.104010. arXiv: gr-qc/0603019.

[130] S. M. Christensen and S. A. Fulling. “Trace Anomalies and the Hawking Effect”. In: *Phys. Rev. D* 15 (1977), pp. 2088–2104. DOI: 10.1103/PhysRevD.15.2088.

[131] Sean P. Robinson and Frank Wilczek. “A Relationship between Hawking radiation and gravitational anomalies”. In: *Phys. Rev. Lett.* 95 (2005), p. 011303. DOI: 10.1103/PhysRevLett.95.011303. arXiv: gr-qc/0502074.

[132] L. Bonora and M. Cvitan. “Hawking radiation, W-infinity algebra and trace anomalies”. In: *JHEP* 05 (2008), p. 071. DOI: 10.1088/1126-6708/2008/05/071. arXiv: 0804.0198 [hep-th].

[133] A. Zee. *Einstein Gravity in a Nutshell*. New Jersey: Princeton University Press, May 2013. ISBN: 978-0-691-14558-7.

[134] G. Oliveira-Neto. "Identifying conical singularities". In: *J. Math. Phys.* 37 (1996), pp. 4716–4723. DOI: 10.1063/1.531650.

[135] Jacob D. Bekenstein. "On Page's examples challenging the entropy bound". In: (June 2000). arXiv: gr-qc/0006003.

[136] John Dougherty and Craig Callender. "Black Hole Thermodynamics: More Than an Analogy?"

[137] David Wallace. "The case for black hole thermodynamics, Part I: phenomenological thermodynamics". In: *Stud. Hist. Phil. Sci. B* 64 (2018), pp. 52–67. DOI: 10.1016/j.shpsb.2018.05.002. arXiv: 1710.02724 [gr-qc].

[138] T. Padmanabhan. "Thermality of the Rindler horizon: A simple derivation from the structure of the inertial propagator". In: *Phys. Rev. D* 100.4 (2019), p. 045024. DOI: 10.1103/PhysRevD.100.045024. arXiv: 1905.08263 [gr-qc].

[139] Chethan Krishnan. "Quantum Field Theory, Black Holes and Holography". In: *Croatian Black Hole School*. Nov. 2010. arXiv: 1011.5875 [hep-th].

[140] Ted Jacobson and Renaud Parentani. "Horizon entropy". In: *Found. Phys.* 33 (2003), pp. 323–348. DOI: 10.1023/A:1023785123428. arXiv: gr-qc/0302099.

[141] Raf Guedens, Ted Jacobson, and Sudipta Sarkar. "Horizon entropy and higher curvature equations of state". In: *Phys. Rev. D* 85 (2012), p. 064017. DOI: 10.1103/PhysRevD.85.064017. arXiv: 1112.6215 [gr-qc].

[142] Christopher Eling, Raf Guedens, and Ted Jacobson. "Non-equilibrium thermodynamics of spacetime". In: *Phys. Rev. Lett.* 96 (2006), p. 121301. DOI: 10.1103/PhysRevLett.96.121301. arXiv: gr-qc/0602001.

[143] W. Greiner et al. *Field Quantization*. Springer, 1996. ISBN: 9783540591795. URL: https://books.google.hr/books?id=VvBAvf0wSrIC.

[144] Merced Montesinos and Ernesto Flores. "Symmetric energy-momentum tensor in Maxwell, Yang-Mills, and proca theories obtained using only Noether's theorem". In: *Rev. Mex. Fis.* 52 (2006), pp. 29–36. arXiv: hep-th/0602190.

[145] Daniel N. Blaschke et al. “The energy–momentum tensor(s) in classical gauge theories”. In: *Nucl. Phys. B* 912 (2016), pp. 192–223. DOI: 10.1016/j.nuclphysb.2016.07.001. arXiv: 1605.01121 [hep-th].

[146] M. Leclerc. “Canonical and gravitational stress-energy tensors”. In: *Int. J. Mod. Phys. D* 15 (2006), pp. 959–990. DOI: 10.1142/S0218271806008693. arXiv: gr-qc/0510044.

[147] John M. Lee. *Introduction to Smooth Manifolds*. Springer Science Business Media, 2003. ISBN: 0387954481, 9780387954486.

[148] Alex B. Nielsen. “Black holes and black hole thermodynamics without event horizons”. In: *Gen. Rel. Grav.* 41 (2009), pp. 1539–1584. DOI: 10.1007/s10714-008-0739-9. arXiv: 0809.3850 [hep-th].