Energy conditions in semiclassical and quantum gravity

Sergi Sirera Lahoz

Supervised by
Professor Toby Wiseman

Imperial College London
Department of Physics, Theory Group

Submitted in partial fulfilment of the requirements for the degree of
Master of Science of Imperial College London
MSc Quantum Fields and Fundamental Forces
2nd October 2020

Abstract

We explore the state of *energy conditions* in semiclassical and quantum gravity. Our review introduces energy conditions with a philosophical discussion about their nature. Energy conditions in classical gravity are reviewed together with their main applications. It is shown, among other results, how one can obtain *singularity theorems* from such conditions. Classical and quantum physical systems violating the conditions are studied. We find that quantum fields entail negative energy densities that violate the classical conditions. The character of these violations motivates the *quantum interest conjecture*, which in turn guides the design of quantum energy conditions. The two best candidates, the AANEC and the QNEC, are examined in detail and the main applications revisited from their new perspectives. Several proofs for interacting field theories and/or curved spacetimes are demonstrated for both conditions using techniques from microcausality, holography, quantum information, and other modern ideas. Finally, the AANEC and the QNEC will be shown to be equivalent in $\mathbb{M}^4$.

Acknowledgements

First of all, I would like to express my gratitude to Prof. Toby Wiseman for supervising this dissertation. Despite the challenges of not being able to meet in person, your guidance and expertise have helped me greatly to stay on track. The simple and clear way in which you communicate ideas has stimulated my interest in the topic and deepened my comprehension. Besides, thanks to your kindness and empathy, I left every meeting with renewed energy and motivation, which has made writing this dissertation a much easier and enjoyable activity.

Second, I feel extremely lucky for the friends I have made at Imperial this year. I appreciate the support that we all have given each other to not only endure this MSc, but also have some fun along the way. I would also like to thank the rest of the theory group at Imperial for creating a nice and supportive learning environment.

Third, to all the amazing friends I have made in London since I first came here. I am thankful for all the unforgettable experiences that we have lived together and the ones still to come.

Fourth, to my friends and family back in Barcelona. Thank you for all the emotional support and encouragement. I am very lucky to have you in my life.

A special thanks goes to everyone who has volunteered to proofread the dissertation through its different stages and to anyone who will read it in the future.

Above all, I would like to thank my mom, my dad and my sister. Thank you for letting me follow my dreams. Without your unconditional love and support, this would not have been possible. For that I will always be deeply thankful.

But, it is similar to a building, one wing of which is made of fine marble, but the other wing of which is built of low grade wood.

Albert Einstein, 1936
about the equation $G_{ab} = 8\pi T_{ab}$

Conventions and Abbreviations

EFEs	Einstein Field Equations
ECs	Energy Conditions
$\mathbb{M}^4$	Minkowski spacetime in 4 dimensions
(A)dS	(Anti) de Sitter spacetime
NEC	**Null Energy Condition**
WEC	Weak Energy Condition
DEC	Dominant Energy Condition
SEC	Strong Energy Condition
ANEC	**Averaged Null Energy Condition**
AANEC	**Achronal Averaged Null Energy Condition**
AWEC	Averaged Weak Energy Condition
ASEC	Averaged Strong Energy Condition
ADEC	Averaged Dominant Energy Condition
QNEC	**Quantum Null Energy Condition**
QWEC	Quantum Weak Energy Condition
QSEC	Quantum Strong Energy Condition
GSL	Generalized Second Law
BCC	Boundary Causality Condition
QFC	Quantum Focussing Conjecture

Natural units	$G = c = \hbar = 1$
Metric signature convention	$(-,+,+,+)$
Fourier transform	$\hat{f}(k) = \int d^n x e^{ik\cdot x} f(x)$
D'Alembert operator	$\Box := \nabla^a \nabla_a$

Contents

List of Figures

Introduction

Throughout the history of science, there is a very small number of concepts that have endured all the different ages until today. Most of them have either died or been replaced by stronger ideas. *Energy* is probably the utmost example of such everlasting concepts. Its first appearance was in the 4th century BC due to Aristotle, who coined the term ἐνέργεια (*energeia*), defined as a sort of "actuality".[1] Actuality refers to a realised potential contrasted to mere unfulfilled potentiality. In other words, *energeia* concerns "what is" rather than "what could be". Hence, it attempts to capture the nature of the ontological components of our universe, what exists physically. The idea that something like energy is conserved can be traced even further back in history to the presocratic philosophers living in the 5th and 6th century BC. For instance, in Empedocles' worldview about the natural world, which was based on the four elements (earth, water, fire and air) "*nothing comes to be or perishes*", which can be identified as a precursor of the conservation law. We see that energy, then, is one of the first concepts humans came up with when they began to explain nature in a scientific manner.

Since its birth, the concept of energy has undergone a continuous metamorphosis, but the main intuition behind it has persevered. Around 2,000 years later, in the 17th century, Leibniz introduced the *vis viva* (living force), which would eventually evolve to become what we now call *kinetic energy*. In the early 19th century, the word "energy" was first used to describe this *vis viva* by Thomas Young, and the use of energy physics, mainly by a group of British scientists and engineers, revolutionised both industry and academia.

[1] There is currently some debate between Aristotle scholars as to whether *energeia* should be understood as *actuality* or *activity*.

Nowadays, energy is still a central concept in our understanding of the universe. In fact, only three years ago Wall wrote: "*perhaps the most important and powerful concept in physics is 'energy'*" [131]. Presently, energy is fundamentally understood both in General Relativity (GR) and Quantum Field Theory (QFT) via a mathematical object called the *stress-energy tensor*, represented as the rank-2 tensor T_{ab}. This tensor, which takes a different form for different classical and quantum fields, is intimately related to the geometry of spacetime through Einstein field equations (EFEs). However, it lacks the mathematical structure that the geometric tensors comprise. This leads to problematic scenarios where one can find very exotic and controversial stress tensors and spacetimes which, despite being highly unphysical, still solve EFEs.

The topic of this dissertation are *energy conditions*, which were originally designed in the mid 20th century to fill in this gap. Relating to the ancient Aristotelian definition, energy conditions aim to differentiate what is *actually* the case in our universe from what could in principle be the case given EFEs. This way, energy conditions can be used to safely rule out the exotic phenomena.

Outline

We will begin in chapter 1 with a discussion about the nature of such conditions. In it, we will define them properly, expand on their original motivations and present their different formulations. Following this discussion, an important point will be made about our lack of understanding in regard to how energy conditions fit in the theoretical framework of General Relativity (GR).

In chapter 2, we introduce the classical energy conditions, which are both local and positive, meaning that a specific quantity is imposed to be positive at every spacetime point. Of these, the most relevant one will be the *null energy condition* (NEC). A central use of the conditions is to prove a list of classical theorems and results. These will be exposed, focusing mainly on singularity theorems, which are the context where energy conditions were first introduced. However, violations of the classical conditions can be found within classical gravity, and so we will show how some of these violations arise.

In chapter 3, we will carry on with the discussion about violations of energy conditions now from a quantum perspective. In short, quantum fields allow negative energies, which are ruled out by the classical conditions. This entails that there is a need for new quantum energy conditions so that such quantum fields can be included in our theory. Nonetheless, the quantum violations seem to follow a trend which can be formulated through the *quantum interest conjecture*. The conjecture tells us that such negative energies are either short-lived and/or small in magnitude. Hence, one can use this idea as a guide to propose new enhanced conditions. In the recent literature, there have mainly been two perspectives to tackle this challenge.

The first one, which is studied in chapter 4, involves sacrificing locality and retaining positivity, and gives rise to energy conditions *averaged* along causal geodesics. In this chapter, we will begin by presenting a list of the main averaged conditions and revisit the consequences that they imply. The rest of the chapter is centered around the *achronal averaged null energy condition* (AANEC), which is arguably the best version of a universally valid energy condition up to date. We will display the arguments behind four of the most interesting proofs of the AANEC that connect it to very different branches of current research activity.

The second perspective one can take to design quantum conditions is to sacrifice positivity while retaining locality. This strategy, which leads to *quantum energy inequalities*, is studied in detail in chapter 5. Similarly, we start off the discussion by introducing the main energy inequalities and revisiting their consequences in this new light. From then, we centre our attention to the *quantum null energy condition* (QNEC), which is an inequality that has recently received much research interest. We will illustrate two different proofs of the QNEC that connect it to interesting active ideas of research and use methods very similar to those used in the AANEC proofs. In the last section, we will consider an additional proof of the QNEC that is specially interesting because it connects it directly to the AANEC. This will allow us to make some comments about the relationship between the two.

Finally, we will close this dissertation by summarizing the principal conclusions that we have obtained and suggesting some directions in which future research could be pursued.

This dissertation is written within the theoretical framework of GR in 4 dimensions. In order to introduce quantum effects, chapters 4 and 5 will assume semiclassical gravity, where the stress-energy tensor will be quantized and treated as a Quantum Field Theory (QFT) object. Some of the arguments in these chapters will also involve ideas from quantum gravity.

Before beginning our journey, we want to stress the relevance of this topic in current and future research. Understanding energy once and for all would be a major scientific deed. More concretely, we believe that the results shown in this work demonstrate that the AANEC and the QNEC tell us interesting and non-trivial facts about the sweet intersection where GR and QFT should merge into a unified theoretical description, a theory that we currently lack and that is driving a considerable amount of research effort. These energy conditions tie in together neatly considerations about gravity with features of quantum fields. Hence, learning more about energy conditions would most certainly bring us one step closer to a complete theory of *quantum gravity*. We conjecture that, once the place these conditions occupy within GR and QFT is clarified and their implications fully developed, they may provide nice tests to decide between different quantum gravity candidates.

How to read this dissertation?

Ideally, from the first to the final page. However, we understand that different readers have different backgrounds and expectations. The first paragraph of each chapter provides a flavour of the main points being discussed. That being said, if one is familiar with the background theory and the relevant challenges, and is only interested in the latest developments, then reading chapters 4 and 5 will suffice. If, alternatively, one is specially interested in the applications of energy conditions, we suggest reading chapters 2 and 3. Finally, for those interested in a general picture of the topic and a philosophical and conceptual discussion, chapter 1 is the one to read, although we would strongly encourage everyone to read this chapter, as it offers an approachable yet essential introduction to the topic.

Chapter 1

The Nature of Energy Conditions

This chapter is devoted to the understanding of energy conditions from a conceptual perspective before jumping into any idiosyncrasies of specific conditions. Energy conditions are deep and non-trivial statements about our universe. Their use, consequences and violations are constantly being widely studied, but their conceptual meaning and fundamentality status remain somehow mysterious, thus making this discussion necessary. Our conceptual examination here is not restricted to the classical pointwise energy conditions introduced in chapter 2 but also applies to the quantum versions studied in chapters 4 and 5.

1.1 What is an Energy Condition?

Energy Conditions (ECs) are usually defined based on the motivations that originally prompted their proposal, but the truth is that they are multifaceted objects with differing interpretations. It is not clear to us yet whether any of these is fundamentally preeminent. This is why it is more prudent to first give the most general features that all energy conditions share. However, due to its historical weight and continued domination in the literature, in the next section (1.2) we will present the traditional depiction and motivation of ECs.

Energy Conditions are several things simultaneously:

- **constraints** on the mathematical potential of one of the three rank-2 tensors T_{ab}, R_{ab}

and/or G_{ab},

- **coordinate-independent inequalities**, therefore concealing fundamental facts about the geometry and/or physical components of our universe, and
- **assumptions** about "physically reasonable" geometries and/or sources of matter.

1.2 Marble and wood: the GR asymmetry G_{ab} vs T_{ab}

Energy conditions were originally proposed by relativists. Accordingly, it is important to see where they structurally fit in the theory of GR and where the different components playing a role in ECs come from. In order to do that, we start with the action for GR

$$S_{GR}[g,\phi] = S_{EH}[g] - S_M[g,\phi], \tag{1.1}$$

where g_{ab} is a Lorentzian metric and ϕ represents the collection of any other matter fields we wish to include in the theory. To obtain the relevant field equations we will invoke the action principle and require the variation of the action to vanish.

The Einstein-Hilbert action is

$$S_{EH}[g] = \frac{1}{16\pi}\int d^4x\sqrt{-g}R, \tag{1.2}$$

where R, the *Ricci scalar*, is the simplest scalar one can form with the metric and first and second derivatives of the metric.[1] The $\sqrt{-g}$ factor ensures the measure is Lorentz invariant. Varying S_{EH} with respect to g_{ab} yields Einstein Field Equations (EFEs) in the absence of matter, i.e. $G_{ab} = 0$, where G_{ab} is the *Einstein tensor* defined as $G_{ab} = R_{ab} - \frac{1}{2}g_{ab}R$, and R_{ab} is the *Ricci tensor*. These tensors encode information about the geometry of the universe. To include matter, we use the action for matter fields

$$S_M[g,\phi] = \int d^4x\sqrt{-g}\mathcal{L}_\phi[g,\phi], \tag{1.3}$$

[1] Other scalars in 4 dimensions involve higher derivatives of the metric.

where $\mathcal{L}_\phi$ is the Lagrangian density for the different fields. Varying with respect to those fields ϕ will produce equations of motions for the fields coupling them to g_{ab}. Varying with respect to the metric will then give us the matter content appearing in EFEs. Because the variation has to be a scalar, this will be represented by a symmetric rank 2 tensor which we call the *stress-energy tensor* T_{ab}. Hence, the variation

$$\delta_g S_M = \frac{1}{2} \int d^4x \sqrt{-g} T_{ab} \delta g^{ab} \tag{1.4}$$

is used to define T_{ab} as

$$T_{ab} := \frac{2}{\sqrt{-g}} \frac{\delta_g S_M}{\delta g^{ab}}. \tag{1.5}$$

Combining the variations with respect to the metric of the two components of S_{GR} gives

$$\delta_g S_{GR} = \int d^4x \sqrt{-g} \left(\frac{1}{16\pi} G_{ab} - \frac{1}{2} T_{ab} \right) \delta g^{ab}, \tag{1.6}$$

and requiring this variation to vanish finally yields the full Einstein Field Equations in the presence of matter

$$G_{ab} = 8\pi T_{ab}. \tag{1.7}$$

At this point, it is convenient to define GR as a theoretical framework. In the literature, GR is broadly defined as a spacetime together with Einstein Field Equations (e.g. [71]). A spacetime is a pair $(\mathcal{M}, g)$, where $\mathcal{M}$ is a connected 4-dimensional Hausdorff C^∞ manifold equipped with a Lorentzian metric g [71]. Note that according to this definition, the components of GR refer only to the LHS of EFEs, i.e. the Einstein tensor G_{ab}. It is then understood that a GR solution, that is, a solution of the EFEs for a specific chosen metric g_{ab}, will give us information about the matter content component T_{ab}.

This was the predominant fashion in which EFEs were tackled around the 50s and 60s. One would first choose any metric g_{ab}, calculate the respective G_{ab} and then find the corresponding stress tensor as $T_{ab} = \frac{1}{8\pi} G_{ab}$. In this fashion, T_{ab} is treated as a trivial rank 2 tensor and whatever physical meaning it might have is not taken into account. As a result, one can easily construct exotic GR solutions where one can, for instance, travel almost instantaneously between spacetime points with spacelike separation or travel backwards in time (wormholes and time travel). These exotic phenomena not only do not seem to take place in our universe, but their existence would lead to deep causal paradoxes. That is why we do not want them in our theory and thus require formal mechanisms to rule them out.

This is where energy conditions naturally come into play as constraints on the stress tensor T_{ab}, characterizing assumptions about what we believe any "physically reasonable" source of matter and energy would have to look like. Energy conditions then prescribe a new fashion to solve EFEs. First, we think about what characterizes the stress tensor of "physically reasonable" sources of matter and design an energy condition depicting that assumption. Then, using EFEs, we see what curvature we get from such a class of stress tensors. This is the traditional and predominant understanding of what energy conditions are, and they were originally proposed, arguably *ad hoc*, as means of proving some central theorems in GR which we will see in the next chapter.

The two contrasting styles show that one can find solutions to EFEs either starting with considerations about g_{ab} or T_{ab}. It is important to realize that EFEs do not show any preference between the two in terms of their fundamentality status within the theory . Thus, it is preferable to define GR as the combination of three ingredients: GR $= G_{ab} + T_{ab} +$ EFEs, where G_{ab} is given by a spacetime $(\mathcal{M}, g)$.

Under this definition, GR is symmetric between the two tensors appearing in EFEs, which conflicts with the treatment that they were usually given in the literature, where the geometry side G_{ab} is normally assigned a more elemental and noble status, and the "physical" side is only assigned a secondary role. In fact, in 1936 Einstein himself likened EFEs to "*a building, one wing of which is made of fine marble, but the other wing of which is made of low grade wood*" [29], and there is no question as to which wing represents which side of the equation. In that paper, Einstein carries on arguing that "*the phenomenological representation of matter is, in fact, only a crude substitute for a representation which would correspond to all known properties of matter*". In a 1926 letter to Besso he wrote:

> "I wonder if the equation $G_{ab} = 8\pi T_{ab}$ still has any reality left within itself, especially when facing quanta. I doubt it, strongly. However, the left hand side surely contains a deeper truth. If the equations $R_{ab} = 0$ really determine the behaviour of the singularities, then the law governing this behaviour would be rooted in a much deeper reason than the former equation, which is not unified and of only a phenomenological kind." (Einstein, 11 August 1926)

What does he mean by a phenomelogical stress tensor? It means that T_{ab} only gives us information about some idealised properties of matter but it tells us nothing about the actual nature of the matter under consideration. Energy conditions, in their traditional presentation, can be thought to upgrade the the right wing of the GR building by giving more "structure" to the stress-energy tensor.

1.3 A twofold job

Energy conditions have a difficult job which involves balancing two tasks that may sometimes conflict with each other. This double job, although considerably intricate, can be understood via a simple analogy. Energy conditions act like the security guard of a very exclusive club. Their job is both to guarantee entrance into the club to the people who are considered "reasonable" and also to forbid entry to anyone who is considered "unreasonable". The challenge of balancing these two duties is the same reason we end up with a large zoo of energy conditions. Violations are found which motivate the proposal of weaker conditions. However, these cannot be weak enough to let in all the unreasonable people. Finding the perfect balance of these two opposites and thus finding the ultimate energy condition has proven to be a difficult enterprise involving deep and non-trivial statements both in GR and QFT.

The analogy can even be further developed to illustrate the main problems with classical pointwise ECs and the need to introduce quantum versions. Imagine that the restriction to enter our exclusive club is for your favourite number to satisfy the inequality $n \geq 0$. After all, who has a negative favourite number? This generally works, but in some cases it does not, since it leaves some reasonable people out of the club (mainly quantum fields). In other words, our condition is failing at characterising faithfully the collective of reasonable people. However, as will be studied thoroughly in chapter 3, these violations follow a trend. Namely, the number of reasonable people with negative favourite numbers and the magnitude of such numbers are small in size.

This trend guides us to propose newly updated conditions. There have been mainly two types of proposals. The first strategy, which we study in chapter 4, is to ask that the average of the favourite numbers is zero or positive. We no longer ask any condition from specific individuals but instead ask a condition from a bigger collection of people. This way, people can have negative favourite numbers as long as they are overcompensated by other people's positive numbers. The second one, which is studied in chapter 5, is simply to lower the bar. We keep asking every person to satisfy a condition but we allow them a small range of negative favourite numbers.

1.4 A threefold formulation

We have seen that ECs were originally motivated as statements about the stress tensor, and are still broadly understood as concerning aspects about the matter and energy content of the universe, hence their name. We call this the *physical formulation* of ECs. However, EFEs tell us that T_{ab} and G_{ab} are equivalent up to a constant scaling factor. Therefore, we can use them to write the conditions in terms of the geometry tensors G_{ab} or R_{ab}. We call this the *geometric formulation*. Hence, EFEs bridge the physical and geometrical formulations, making them equivalent and interchangeable.

In addition, we can define a new formulation by specifying the form of the stress tensor and finding algebraic expressions for its components. This is only a special case of the physical formulation, but it comes in handy when interpreting the physical formulation, as it connects it to the physical quantities of energy density and pressure. We call this the *effective formulation*. Hawking and Ellis give a classification of stress-energy tensors [71]. In order to write the *effective formulation* it is conventional to use the *Hawking-Ellis Type I* stress tensors, which can be written as

$$T^{ab} = \left(\begin{array}{c|ccc} \rho & 0 & 0 & 0 \\ \hline 0 & p_1 & 0 & 0 \\ 0 & 0 & p_2 & 0 \\ 0 & 0 & 0 & p_3 \end{array}\right) \tag{1.8}$$

where there is one timelike and three spacelike components representing the energy density and the pressure in the three spatial directions. This is the form that the majority of classical fields can take. One can also find in the literature the stress-energy tensor of a perfect fluid being used for this purpose, i.e. $T^{ab} = (\rho + p)v^a v^b + pg^{ab}$ where v^a is the vector field of the velocity of the fluid [79], but we will use the *Type I* expression when we study the classical pointwise energy conditions in chapter 2.

1.5 Fundamentality of Energy Conditions

Energy conditions have a very peculiar character. Unlike most of the other statements in GR, we cannot classify ECs as either postulates or consequences from the postulates [23]. They are seemingly innocent and independent assumptions[2] that nonetheless have a far-reaching impact on many aspects of the theory.

Because of this, it is specially challenging to assess their fundamentality status within the theory. A noteworthy trait of many of the classical consequences of the ECs is that they are independent of EFEs. For instance, one could assume the negation of EFEs and still be able to prove Hawking's Area Theorem using the geometric formulation of ECs. The only role that the EFEs could have here is to bridge the geometric to the physical formulations, thus providing a physical interpretation of what the energy condition in use amounts to.

Curiel uses this and other considerations to attempt to place ECs in a list of different structures in GR from more to less fundamental [23]. According to him, ECs are less fundamental than the

[2]It must be mentioned that these assumptions concerning what counts as "physically reasonable" are largely based on our physical intuition, something which we have seen recurrently discredited in the history of physics.

event structure, *causal structure*, *conformal structure* and *differential structure* since, without the latter we could not even write a tensor; and without the others we could not differentiate between null, timelike and spacelike vectors.[3] Also, the averaged conditions (chapter 4) will require some *affine structure* to be defined. On the other hand, ECs appear to be independent of *temporal orientability*, *topological structure*, *metric structure* and the EFEs.

From this discussion, Curiel claims that ECs do not fit nicely anywhere on his list [23]. He argues that this might suggests that ECs should not after all be considered as components of a spacetime theory, but rather as running parallel to it. However, there is another complication. Despite EFEs being less fundamental than ECs, it is question-begging that if we assume EFEs, the ECs immediately become more powerful, as if they were trying to tell us than GR, where EFEs hold, is more special than other modified spacetime theories.

Finally, the various difficulties which arise in assessing the fundamentality of these conditions as well as the inability of deriving them from any first principles should prevent us from favouring one formulation over the other, i.e. the physical over the geometric one. In fact, most of the classical results use only the geometric formulation. With all that we know at present, it could well be that ECs are actually statements about the geometrical tensors.

[3] In fact, many recent results on quantum energy conditions connect them tightly to causality considerations.

Chapter 2

Energy conditions in classical gravity

In this chapter we review the main pointwise energy conditions used within the context of classical gravity. This chapter is included both for historical interest as well as for grounding and motivating our subsequent discussion of quantum energy conditions. After presenting the 4 main classical conditions and their interpretations we will see how they can be used to prove useful results, giving special attention to the singularity theorems; and finally consider their key violations.

2.1 Survey of pointwise Energy Conditions

Pointwise energy conditions restrict the form of the stress-energy tensor at every spacetime point. They do so by requiring that a specific contraction of the given rank 2 tensor with two vectors is never negative. Different conditions use either different linear combinations of T_{ab} or different vectors, i.e. timelike or null. As mentioned in the previous section, EFEs build a bridge from the physical to the geometric formulation. Further, using the Hawking-Ellis Type I stress tensor allows one to write the conditions in the effective formulation. Now we present the 4 main pointwise energy conditions: Weak, Strong, Dominant and Null; stated in the different formulations and discuss their respective interpretations.

Weak Energy Condition (WEC)

physical	$T_{ab}t^a t^b \geq 0$	$\forall$ timelike future pointing vector t^a
geometric	$G_{ab}t^a t^b \geq 0$	$\forall$ timelike future pointing vector t^a
effective	$\rho \geq 0$ and $\rho + p_i \geq 0$	$i = (1,2,3)$

In this case, the effective formulation provides a straightforward interpretation of the physical form. It tells us that an observer travelling on a timelike curve always measures the energy density from all the matter fields in the universe to be non-negative.[1] Also, there is no restriction on the sign of the pressure but, if negative, its magnitude cannot be bigger than that of the energy density. It is, however, much more challenging to find an interpretation of the geometric form. This is because there is no intuitive understanding of the Einstein tensor in geometric terms. Unlike the Riemann tensor, which is easily associated with geodesic deviation, there is no similar interpretation that captures the geometric significance of G_{ab}.

Strong Energy Condition (SEC)

physical	$\left(T_{ab} - \frac{1}{2}T g_{ab}\right) t^a t^b \geq 0$	$\forall$ timelike future pointing vector t^a
geometric	$R_{ab}t^a t^b \geq 0$	$\forall$ timelike future pointing vector t^a
effective	$\rho + p_i \geq 0$ and $\rho + \sum p_i \geq 0$	

Here, the effective formulation tells us that in case of having negative pressures, the energy density has to be larger than the magnitude of the momentum flux in each of the three spacelike directions. It also tells us that the sum of the pressures in the three directions cannot negatively dominate the energy density at one specific spacetime point.

To find an interpretation for the physical formulation of the SEC, we can choose the vector t^a to be $t^a = (1,0,0,0)$. Hence, the condition tells us that in Minkowski spacetime,

$$T_{00} + \frac{1}{2}T \geq 0. \tag{2.1}$$

This quantity has received the name of *effective energy density* [16], although it was originally

[1] This can also be understood as implying that the gravitational interaction is never repulsive.

named *effective density of gravitational mass* [132]. We can understand the physical formulation of the SEC as imposing that an effective energy density is non-negative as measured by an observer following a timelike worldline.

To obtain the geometric form from the physical one, we have taken EFEs and multiplied both sides by the metric, thus finding the expression $-R = 8\pi T$. Then, one can rewrite EFEs as

$$R_{ab} = 8\pi \left(T_{ab} - \frac{1}{2} g_{ab} T \right), \tag{2.2}$$

and we use this to bridge the physical and geometric formulations of the SEC.

The geometric SEC can be given an interpretation via the geodesic deviation equation

$$R^{c}{}_{abd} t^a \xi^b t^d = t^a \nabla_a (t^b \nabla_b \xi^c) := \frac{D^2}{d\tau^2} \xi^c. \tag{2.3}$$

Here, ξ^a is a vector field along the affinely parametrised geodesic γ defined by the integral curves of the timelike future-directed vector field t^a with affine parameter τ. ξ^a is such that the contraction $\xi^a t_a = 0$ and $\mathcal{L}_t \xi^a = 0$ at one spacetime point. Hence, ξ^a has to be spacelike and can be imagined as a *connecting vector* from the geodesic γ to an infinitesimally nearby geodesic also defined by the integral curves of t^a. The right hand side of equation 2.3 can then be thought of as the relative acceleration between the two nearby geodesics. Taking the radial component we can define the *average radial acceleration* [23]

$$A_r := -\frac{1}{3} \xi_r t^a \nabla_a (t^b \nabla_b \xi^r). \tag{2.4}$$

Being a scalar, this quantity is coordinate independent and therefore refers to an inherent geometric notion. Using the geodesic deviation equation 2.3 we can rewrite A_r as

$$A_r = -\frac{1}{3} R_{ab} t^a t^b. \tag{2.5}$$

Then, if the SEC is satisfied, the RHS of equation 2.5 will be negative, so $A_r \leq 0$, meaning that the two nearby geodesics will never diverge from each other. Hence, the geometric interpretation of the SEC is that congruences of timelike geodesics converge (or at least do not diverge).

Dominant Energy Condition (DEC)

physical	$T_{ab}t^a\xi^b \geq 0$	$\forall$ co-oriented timelike future pointing vectors t^a and ξ^a
geometric	$G_{ab}t^a\xi^b \geq 0$	$\forall$ co-oriented timelike future pointing vectors t^a and ξ^a
effective	$\rho \geq 0$ and $\lvert p_i \rvert \leq \rho$	

One can also alternatively formulate the WEC with only one timelike vector t^a satisfying the two conditions $T_{ab}t^at^b \geq 0$ and $|T^{ab}t_a| \leq 0 \;\; \forall\, t^a$, the latter implying that $T^{ab}t_a$ is a non-spacelike (or causal) vector.

The physical and effective formulations imply that a timelike observer measures the energy density to be non-negative and the momentum flux to be causal and with the same orientation as the proper time of the observer. One can then think of the DEC as ruling out superluminal travel of any source of the stress-energy tensor. The geometric formulation suffers from the same difficulties as the WEC and therefore lacks a plain interpretation.

Null Energy Condition (NEC)

physical	$T_{ab}k^ak^b \geq 0$	$\forall$ null vectors k^a
geometric	$R_{ab}k^ak^b \geq 0$	$\forall$ null vectors k^a
effective	$\rho + p_i \geq 0$	

The first point to notice here is that in this case EFEs bridge the stress-energy tensor T_{ab} directly to the Ricci tensor R_{ab}. The reason for this being that null vectors satisfy $k^2 = k^ak^bg_{ab} = 0$ and so

$$8\pi T_{ab}k^ak^b = \left(R_{ab} - \frac{1}{2}Rg_{ab}\right)k^ak^b = R_{ab}k^ak^b, \tag{2.6}$$

implying that

$$T_{ab}k^ak^b \geq 0 \qquad \Rightarrow \qquad R_{ab}k^ak^b \geq 0. \tag{2.7}$$

The physical interpretation of the NEC is similar to the WEC one. It tells us that the sum of energy density and pressure in each of the three orthogonal spatial directions is never negative.

It is, however, less clear to imagine situations where an observer travels along a null geodesic since physical observers are normally taken to be moving along timelike geodesics. Thus, it is not altogether clear what the physical significance of the NEC amounts to.

To give an interpretation of the geometric form of the NEC we follow the exact same process as we did for the SEC with one specific difference. Instead of having a timelike vector t^a defining the geodesics, we now have them defined by the null vector k^a. Then, the corresponding connecting vector ξ^a is not spacelike but also null (ξ^a can be proportional to k^a). The average radial acceleration is the same, with the single difference of having a factor of 2 rather than 3 in the denominator

$$A_r = -\frac{1}{2}\xi_r k^a \nabla_a (k^b \nabla_b \xi^r) = -\frac{1}{2} R_{ab} k^a k^b \tag{2.8}$$

and so similarly the NEC entails $A_r \leq 0$, implying the convergence (or at least non-divergence) of null geodesic congruences.

The zoo of classical ECs is much more plethoric than what we have presented here. There are many other less interesting conditions such as the TEC (Trace Energy Condition), SDEC (Strong Dominant Energy Condition), FEC (Flux Energy Condition) and others that have sporadically appeared in the literature. We do not include them because they either have been found to be hopeless or they have not been followed up by quantum versions.

2.2 Implications

ECs are all linked together in a net of logical relations. Here we outline all the implications between the four main classical conditions. After presenting the quantum versions, we will build up this net to include them.

- **DEC $\Rightarrow$ WEC**

It is easily seen that the WEC is only a special case of the DEC. If we require $-T^a_{\ b} t^b$ to be a causal and future-oriented vector field, then the DEC becomes logically equivalent to the WEC.

- **WEC ⇒ NEC**

The first one to prove this implication was Tipler [119]. He proved that if $T_{ab}t^a t^b \geq -b$ is a finite lower bound for all timelike unit vectors t^a, then $T_{ab}k^a k^b \geq 0$ is satisfied for all null k^a. Note that restricting ourselves to consider only normalized vectors for the WEC implies that the lower bound can be negative (i.e. b does not necessarily have to equal zero).[2] This result can be obtained, as shown by Kontou and Saunders [79], by noticing that any unit timelike vector can be rewritten as a linear combination of two null vectors $t^a = \frac{1}{r}k^a + rn^a$ where $r > 0$ and $k^a n_a = -\frac{1}{2}$. Then the WEC becomes

$$-b \leq T_{ab}t^a t^b = \frac{1}{r^2}T_{ab}k^a k^b + (T_{ab} + T_{ba})k^a n^b + r^2 T_{ab}n^a n^b. \tag{2.9}$$

Kontou and Sanders then show that multiplying both sides by r^2 and taking the limit $r \to 0^+$ entails the NEC, i.e. $T_{ab}k^a k^b \geq 0$. In addition, because of our previous result, this also shows that DEC ⇒ NEC.

- **SEC ⇒ NEC**

In his proof of the last implication, Tipler originally assumed T_{ab} to be a *Hawking-Ellis Type I* stress-energy tensor. However, Kountou and Saunders' proof does not require any previous assumption for T_{ab} and it can in fact be replaced by any rank-2 tensor. Hence, the same argument also shows that

$$R_{ab}t^a t^b \geq 0 \qquad \Rightarrow \qquad R_{ab}k^a k^b \geq 0. \tag{2.10}$$

for all timelike t^a (geometric SEC) and null k^a (geometric NEC).

As a side remark, it is noteworthy that we have used both physical and geometric formulations of the conditions to prove implications between them. One might have hoped that if all implications were much more easily proven using only one of the two formulations, this might suggest that that formulation should be regarded as more fundamental. However, we are again

[2] The same applies to the rest of ECs using timelike vectors.

let down and the fundamentality mystery persists.

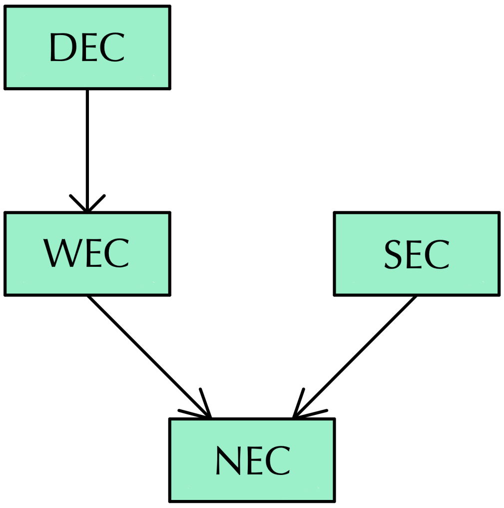


Figure 2.1: Logical relations between classical ECs. For instance, the line from the WEC to NEC means that if the WEC is true, then the NEC is also true. Roughly, the strongest conditions are displayed at the top and the weakest at the bottom. This diagram will be extended in the following chapters.

The NEC is the weakest condition since it is implied by all the rest. Also, contrary to what their names might suggest, we note that the SEC does not imply the WEC. These implication relations come in handy when discussing consequences and violations of the energy conditions, since they reduce the amount of work to be done. On one hand, proving that a physical system obeys a strong condition implies that it will also obey a weaker condition, given that the latter is implied by the former. On the other hand, proving that a system violates the weaker condition immediately entails that it will also violate the more demanding condition.

2.3 Applications

Here we review a non-exhaustive list of the main results one can derive using the different energy conditions. Showing that a specific result can be derived from the NEC, the weakest condition, will immediately entail that the same result is satisfied by the rest of the conditions. Similarly, results satisfied by the WEC will also be satisfied by the DEC. For most of these consequences, EFEs are assumed so that one can bridge the two formulations. This allows us to give a physical interpretation to the geometrical results. However, the results can be obtained independently of EFEs.

2.3.1 Singularity Theorems

Spacetime singularities are pathologies in the geometry of spacetime, breakdowns of GR which are almost as old as the theory itself. Singularity theorems are a collection of arguments showing that GR predicts the existence of singularities in the universe. When singularities were first found, there was debate concerning their physical reality. It was thought that they could just be artifacts resulting from the strong symmetries of the models where they were found (Schwarzschild and Friedman-Robertson-Walker), and that they did not actually exist. Finally, singularity theorems closed this debate and forced us to accept singularities whether we like them or not.

What is a singularity?

Let us attempt here to understand the notion of a *singular spacetime*. Singularities, despite being ubiquitous in the GR literature, are very difficult to pin down in a general and precise mathematical definition.

The first and easiest example where we can see a singularity is in the Schwarzschild metric, discovered in 1916 [111] only one year after the publication of GR and given by the line element

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2, \tag{2.11}$$

where $d\Omega^2$ is the 2-dimensional solid angle. This metric describes the geometry outside a spherical source of matter like a star. The components of the metric diverge at two points: when $r = 2M$ and $r = 0$. The first one is only a *coordinate singularity*, meaning that it is an artifact of our coordinate choice carrying no physical significance which can be easily removed via a coordinate transformation. The second one, however, is a true *curvature singularity* signaling a genuine pathological behaviour of the geometry.

Hence, a first attempt to define singularities is as the "places" where the metric components are not smooth or where its determinant vanishes. But this definition quickly fails, as it is obviously problematic to define a singularity as "place", i.e. an event in spacetime, if the metric is not well-defined around it [125]. In our definition of a spacetime as the pair $(\mathcal{M}, g)$ it is implied that the metric g is defined everywhere on $\mathcal{M}$. Therefore, since the metric is not well-defined in a singularity, these are not included in the spacetime. This makes it impossible to refer to singularities as events in spacetime.[3]

A natural next attempt to characterize singularities is as boundaries of a manifold. The spacetime then would be the manifold without the boundary, but we could still talk about singularities as the "places" where spacetime ends, the edges of the universe, with the only difference that these edges can be reached with a finite distance rather than infinite. However, it turns out that this can only be done for some strongly symmetric models like Schwarzschild or Friedman-Robertson-Walker [125].

Our next attempt is also a natural thing to do. Scalars are coordinate-invariant quantities, so we know that if they diverge in some coordinates, they will do so in every other coordinate system. If we can find a scalar quantity quantifying the geometrical curvature we could define singularities as the domain where that quantity diverges. This can be done for Schwarzschild, for which the *Kretschmann scalar* is

$$K := R^{abcd}R_{abcd} = \frac{48M^2}{r^2}, \tag{2.12}$$

which clearly diverges as we approach the singularity at $r = 0$. This too, however, cannot be generalized to any spacetime model. The easiest example to see this is with a *conical singularity* with the metric[4]

$$ds^2 = dr^2 + (\lambda r)^2 d\phi^2, \tag{2.13}$$

with variable domains $0 < r < \infty$, $0 \leq \phi < 2\pi$ and periodicity $\phi \sim \phi + 2\pi$. If $\lambda = 1$, we can transform back to $ds^2 = dx^2 + dy^2$ ($\mathbb{R}^2$) and we then see that $r = 0$ is just a coordinate singularity removable with a shift. However, for $\lambda \neq 1$ things turn out differently. Defining $\phi' = \lambda\phi$ then $ds^2 = dr^2 + r^2 d\phi'^2$. This is clearly locally isometric to $\mathbb{R}^2$, but fails to be globally isometric to

[3] For a discussion on different definitions of singularities see Wald [125] pages 212-216 and the Stanford Encyclopedia of Philosophy article on *spacetime singularities* [24] (chapter 1).

[4] Or in 2-dimensional Minkowski.

$\mathbb{R}^2$ since the period of ϕ' is now $2\pi\lambda$. Consider a small circle with infinitesimal radius $r = \epsilon$. Then the ratio of its circumference c over its radius is $\frac{c}{r} = \frac{2\pi\lambda\epsilon}{\epsilon} = 2\pi\lambda$, independent of ϵ. This ratio then stays constant as $\epsilon \to 0$, i.e. in the conical singularity. We can visualise this with an origami exercise. We take a sheet of paper (i.e. $\mathbb{R}^2$) and introduce the coordinates r and ϕ. We then select a range of ϕ which we cut and remove. Finally, we stick the two ends together and we make a cone [123].

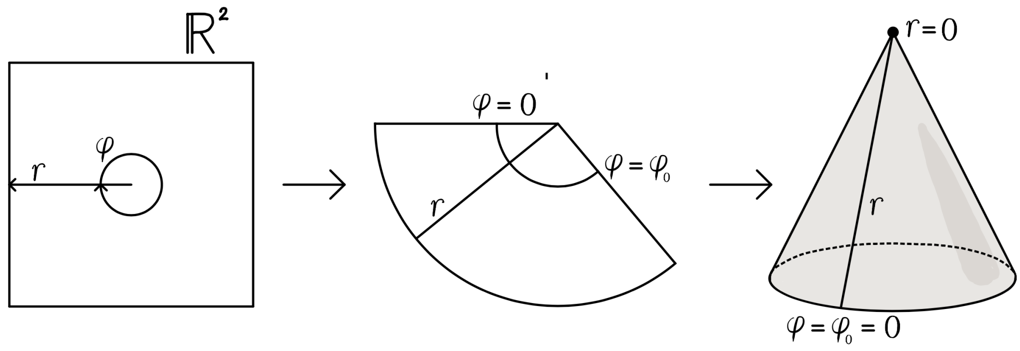


Figure 2.2: Conical singularity at $r = 0$.

If we zoom in on the surface of the cone and look at an infinitesimally small patch, we will see something very similar to our original piece of paper. This will be true for all points except one, our conical singularity at $(r = 0, \phi = 0)$. Yet, we cannot characterize this singularity with any curvature scalar since, being flat space, the Riemann tensor will vanish everywhere.

Quite ironically, the best characterization of singularities to date is based not on what they are but on what they entail. This is done using the notion of *geodesic completeness*. The idea is that if singularities are something like holes in spacetime, then geodesics which run into them surely will have a finite length in their parameter and will not be able to be extended beyond the singularity.

To make this idea more explicit we need to give some definitions.

- A *curve* γ is a smooth map from an open interval $(a, b) \subset \mathbb{R}$ to $\mathcal{M}$

$$\gamma : \qquad (a, b) \longrightarrow \mathcal{M}. \tag{2.14}$$

- Then, a point $p \in \mathcal{M}$ is a *future endpoint* of a future directed causal curve γ if for any neighbourhood Θ of p there exists $t_0 \in (a, b)$ such that $\forall\ t > t_0$, $\gamma(t) \in \Theta$. Namely, the curve stops at p.

- A curve γ is *future-inextendible* if it has no future endpoint. Similarly, we can get analogous definitions for *past endpoints* and *past-inextendible* curves.
- Then, we define an *inextendible curve* as a curve which is both future and past-inextendible.
- A geodesic is *complete* if there exists an *affine parameter* λ for it that extends to the full range $\lambda \in (-\infty, \infty)$.
- Finally, a spacetime is *geodesically complete* if all inextendible causal geodesics are complete.

All this technology allows us to define a *singular spacetime* as a spacetime which is *geodesically incomplete* and *inextendible*.

An example of a non-singular spacetime is Minkowski $\mathbb{M}^4$ while an example of a singular spacetime is Schwarzschild. This definition is not free of problems but is definitely the one that works best and it suffices in the examples we consider here.[5]

Raychaudhuri's equation

An important element in singularity theorems concerns the study of the evolution of timelike or null congruences of geodesics. The rate of separation between geodesics in these congruences will depend on the curvature of the geometry where they live. For simplicity, here we will only consider congruences of null curves but a similar analysis with a couple of differences can be done for timelike curves. We will follow a mixture of the notations used by Hawking & Ellis [71] and Wald [125] in our analysis of the null case.

We define k^a to be the tangent vector along one null curve and η^a to be the orthogonal vector from one geodesic in the congruence to an infinitesimally nearby geodesic. One can prove that the timelike separation between light signals (signals travelling along null curves) coming from the same source remains constant (page 87 of [71]). Therefore, it is more interesting to study the spacelike separation between such geodesics and that is why we take η^a to be spacelike.

[5] For a comment on the problems of this definition see Wald [125] page 215.

They satisfy the relation

$$\mathcal{L}_k \eta^a = 0, \tag{2.15}$$

from which we get

$$\frac{D}{d\lambda}\eta^a = \eta^b \nabla_b k^a \qquad \frac{D^2}{d\lambda^2}\eta^a = -R^a{}_{bcd}\eta^c k^b k^d, \tag{2.16}$$

where λ is an affine parameter for the null curve. The second equation is the analogous to equation 2.3 for null tangent vectors. By choosing *pseudo-orthonormal* coordinates one can show that only two of the components of the first equation are non-trivial [71], so

$$\frac{D}{d\lambda}\bar{\eta}^m = \bar{\eta}^n \nabla_n \bar{k}^m, \tag{2.17}$$

where $m, n = 1, 2$ and so the new vectors live in a 2-dimensional space. The tensor field $\nabla_m k_n$ can be decomposed in an anti-symmetric, a traceless-symmetric and a trace part as

$$\nabla_m k_n = \frac{1}{2}\theta h_{mn} + \sigma_{mn} + \omega_{mn} \tag{2.18}$$

where we have defined

Expansion	$\theta = h^{mn}\nabla_m k_n$	trace	(2.19)
Shear	$\sigma_{mn} = \nabla_{(m}k_{n)} - \frac{1}{2}\theta h_{mn}$	traceless-symmetric	(2.20)
Vorticity	$\omega_{mn} = \nabla_{[m}k_{n]}$	anti-symmetric	(2.21)

and where h_{mn} is the corresponding metric in the 2-dimensional space. The definition of this induced metric from the original 4-dimensional g_{ab} is more subtle than in the timelike case.[6] Here, the space where the deviation vectors live is 2-dimensional, and we call it $\hat{V}_p$ (with a hat) following Wald [125]. This space is defined from the 3-dimensional tangent space $\tilde{V}_p$ (with a tilde) of vectors orthogonal to k^a by identifying as equivalent any two vectors $x^a, y^a \in \tilde{V}_p$ which satisfy $x^a - y^a = ck^a$, c being a constant $c \in \mathbb{R}$ [125]. $\tilde{V}_p$ is in turn a subspace of V_p, the tangent

[6]For the timelike case the analogous h_{ab} is simply the induced metric on the spatial slices. It is 3-dimensional and given by $h_{ab} = g_{ab} + t_a t_b$.

space at a point $p \in \mathcal{M}$. This is useful because while there is no natural vector living in $\tilde{V}_p$ that one can obtain from a vector in V_p, it turns out that there is one for $\hat{V}_p$. For dual vectors, one finds that there is a natural $\tilde{\omega}_a \in (\tilde{V}_p)^*$ which acts on vectors on $\tilde{V}_p$ obtained from a dual $\omega_a \in V_p^*$. In turn, one can obtain a dual $\hat{\omega}_a \in (\hat{V}_p)^*$ from $\tilde{\omega}_a$ by imposing $\hat{\omega}_a k^a = \omega_a k^a = 0$. Now, a general tensor $\hat{T}^{a_1 \cdots a_n}{}_{b_1 \cdots b_m}$ in $\hat{V}_p$ will be naturally obtained from a tensor $T^{a_1 \cdots a_n}{}_{b_1 \cdots b_m}$ in V_p iff contracting one of the indices (upper or lower) of the latter tensor with either k^a or k_a and contracting the rest of them with vectors or covectors that will themselves lead to vectors or covectors in the spaces $\hat{V}_p$ and $(\hat{V}_p)^*$ gives zero as a result. The 2-dimensional metric h_{mn} is the tensor living in $\hat{V}_p$ that we obtain by implementing this process on the spacetime metric g_{ab} and, respectively, h^{mn} is equally obtained from g^{mn}.[7]

The *expansion*, *shear* and *vorticity* are combined in the *Raychaudhuri equation* for null geodesics [71, 125]

$$\frac{d\theta}{d\lambda} = -R_{ab} k^a k^b + 2\omega^2 - 2\sigma^2 - \frac{1}{2}\theta^2. \tag{2.22}$$

The sign of the RHS will determine whether geodesics in the congruence diverge or converge as we travel along them, i.e. whether the orthogonal distance between them increases or decreases. Vorticity will contribute to divergence[8] while shear will contribute to convergence.

We can find connections to energy conditions from this equation. In fact, textbooks usually present energy conditions after studying the Raychaudhuri equation [71, 121, 125], and Liu and Rebouças even claim to derive the SEC and NEC from purely geometrical considerations (i.e. the Raychaudhuri equation) with the only assumption that gravity is attractive [84]. However, it is easy to see that this argument fails [23]. We shall see here the NEC case but the SEC case is completely analogous.

One can use Frobenius Theorem to show that vorticity vanishes for a hypersurface orthogonal congruence [23], so $\omega_{ab} = 0$ for the cases we are interested in. In fact, we can also rightly assume

[7] We have avoided writing the hats in the induced metric for the sake of simplicity.
[8] Vorticity (or twist) can be understood as an analogous to a centrifugal force.

our congruence to be *vorticity-free* by looking at the differential equation for ω_{mn} [71]

$$\frac{d\omega_{mn}}{d\lambda} = -\theta\omega_{mn} + 2\omega_{p[m}\sigma_{n]p}, \tag{2.23}$$

which is linear and homogeneous in ω_{mn}. What this means is that if we start with a congruence of null geodesics which is initially vorticity-free , they will remain vorticity-free, i.e. $\omega_{mn} = 0$. But one might ask, is it an impossible condition to have a congruence of vorticity-free null geodesics in the first place? If so, none of this would apply. However, it happens to always be possible to arrange the null congruence to be initially vorticity-free. Hence, we can omit the second term in equation 2.22.

The last two terms in equation 2.22 are explicitly negative since the shear is a "spatial" tensor so $\sigma^2 := \sigma^{mn}\sigma_{mn} \geq 0$ and the expansion is a scalar so $\theta^2 \geq 0$. The requirement of gravity being attractive is interpreted as demanding the sign of the RHS of 2.22 is negative.[9] Liu and Rebouças [84] claim that this implies the geometrical NEC $R_{ab}k^a k^b \geq 0$, but the reality is that it only implies the weaker statement

$$R_{ab}k^a k^b \geq -\frac{1}{2}\theta^2 - 2\sigma^2. \tag{2.24}$$

Still, we see that the geometrical study of the evolution of geodesic congruences connects in some way with the geometrical forms of energy conditions.

General recipe for singularity theorems

The first singularity theorem for a general spacetime was published by Penrose in 1965 [99] and in the next years new theorems were proved and improved by Hawking and Penrose [68, 69, 72].[10] Such theorems require a list of necessary ingredients in order to cook up singularities. One of these ingredients are energy conditions. In fact, the earliest appearances of energy conditions

[9] There is disagreement concerning this interpretative equivalence between gravity being attractive and geodesics converging [23].

[10] Previous theorems proving the existence of singularities were restricted to stress-energy tensors for perfect fluids [78, 103].

were in these theorems, where they were suspiciously introduced *ad hoc* as means to obtain the desired results. Kontou and Sanders [79] conveniently divide the ingredients in three categories.

- Boundary conditions
- Energy conditions
- Causality conditions

With different combinations of them one can obtain spacetimes with at least one incomplete causal geodesic, that is, a singular spacetime. We can get a taste of how the theorems work by understanding the job each of these conditions has. In short, the boundary condition implies that geodesics start focusing, the energy condition ensures that the focusing continues and the causality condition assures that there are no conjugate points along the geodesics. This will inevitably lead to geodesic incompleteness and thus to singularities. We now expand more on what each of these conditions entails.

Boundary conditions involve the presence of a *closed trapped surface*. The intuition behind trapped surfaces is finely explained in Hawking & Ellis using an idea by Penrose [71, 99]. Imagine a point on the surface of a sphere which emits regular pulses of light spreading radially from the source. One wavefront will point into the centre of the sphere and an opposite wavefront will point radially away from it. These two points will define two new spheres which will be either bigger or smaller than the original one. Normally, one would expect the sphere defined by the outwards wavefront to be bigger than the original and the one defined by the inwards wavefront to be smaller, as it is shown in figure 2.3.

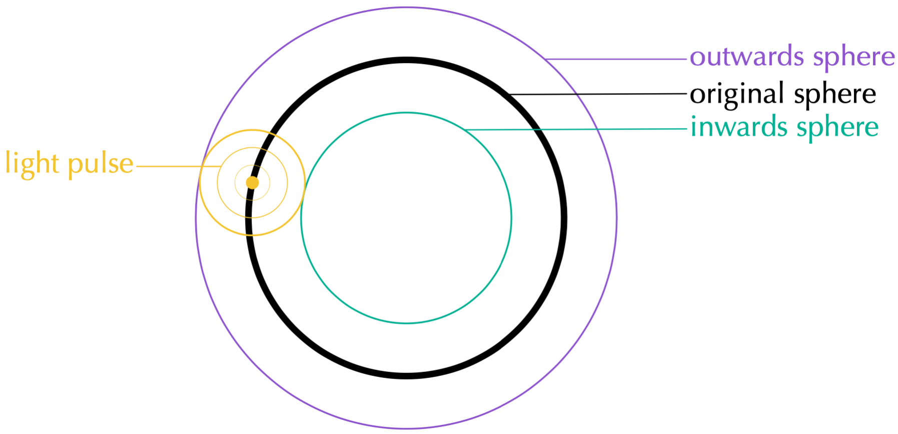


Figure 2.3: Trapped Surface (diagram idea taken from [71])

Our original sphere will be a closed trapped surface if the both spheres defined by the wavefronts have a smaller area than the initial one. More rigorously, a trapped surface is defined as a co-dimension 2 spacelike submanifold which has two null normals with negative expansion [79].

There are two main cases in our universe where we expect trapped surfaces to exist. The first one is when a sufficiently massive star collapses and forms a black hole. After rewriting the Schwarzschild metric in Eddington-Finkelstein coordinates, one can realise that we have trapped surfaces inside the event horizon where outgoing and ingoing light signals both point into the the centre at $r = 0$ (the event horizon itself being an asymptotic trapped surface). The second one is the initial cosmological singularity in the past.

The role of energy conditions in the theorems, as we have seen from our analysis of the Raychaudhuri equation, is the convergence of null vorticity-free geodesic congruences

$$\frac{d\theta}{d\lambda} + \frac{\theta^2}{2} \leq 0 \tag{2.25}$$

which can be solved to give the inequality

$$\theta \leq \frac{2}{\lambda + 2\theta^{-1}|_{\Sigma}}, \tag{2.26}$$

where $\theta^{-1}|_{\Sigma}$ is the initial value of the expansion at the surface Σ. If this value is negative, then we see that the expansion will diverge for some $\lambda \leq 2|\theta^{-1}|_{\Sigma}|$. This phenomenon is known as *geodesic focusing*.

The divergence of the expansion is important because it indicates the existence of *conjugate points* (Proposition 4.4.1 in Hawking & Ellis [71]). A conjugate point p is a point for which infinitesimally neighbouring geodesics intersect through q. In order to have a conjugate point, the geodesics need to be extendible enough along their parameter to reach that point. What Raychaudhuri's equation tells us, then, is that every null *complete* geodesic will have a conjugate point.

Finally, the causality condition ensures causality is not violated by closed timelike curves. This is done with the presence of a *Cauchy surface* which, assuming the null geodesics are complete, entails the existence of inextendible null geodesics containing no conjugate points. This clearly contradicts the result from Raychaudhuri's equation. The tension is resolved by settling that the null geodesics are incomplete. Hence, the spacetime is singular.

The proper statement of the theorem is [71]

Theorem 2.1 (**Null Singularity Theorem**). *The spacetime* $(\mathcal{M}, g)$ *is future null geodesically incomplete (singular) if:*

1. *There is a closed trapped surface;*
2. $R_{ab}k^a k^b \geq 0 \ \forall$ *null vectors* k^a*; and*
3. *There is a non-compact Cauchy surface.*

Other singularity theorems (e.g. for the timelike case) can be proven following a similar logic. They can also be proven in a way that does not involve the Raychaudhuri equation [94], but the presence of energy conditions is still essential. We have seen a singularity theorem proven by the NEC, which is the weakest of the 4 main pointwise conditions in classical gravity. Other theorems were initially proven by the SEC, which quickly became problematic due to its many violations that we will see in the next section. This threatened the validity of the theorems, urging physicists to prove the theorems from weaker conditions. An example is Tipler's proof that the WEC was enough to prove one of the theorems which originally used the SEC [119].

The discussion of proving singularity theorems with the more recent and weaker versions of energy conditions which take into account quantum effects is paramount. It is still not clear yet whether some of these are strong enough to produce the theorems. Fewster and Kontou confirm to be working on a proof of a singularity theorem using a quantum energy inequality so we can expect new results on this to come out soon [79]. We will comment more on this in chapters 4 and 5 after presenting the relevant quantum energy conditions.

2.3.2 Cosmic Censorship Conjecture

The main idea behind cosmic censorship is to hypothesise the non-existence of *naked singularities*, motivated by the association of naked singularities with a loss of predictability. Empirically, we do not see such singularities so it makes sense to speculate that a theoretical proof can be found validating the conjecture. In fact, the existence of black holes originating in gravitational collapse in a generic situation requires this conjecture to be true [73]. There are two inequivalent versions:

- **Weak Cosmic Censorship**: *Singularities are required to be clothed by an event horizon from an observer on the future null infinite region* $\mathscr{J}^+$.
- **Strong Cosmic Censorship**: *The classical fate of all observers should be predictable from non-singular initial data on a Cauchy surface.*

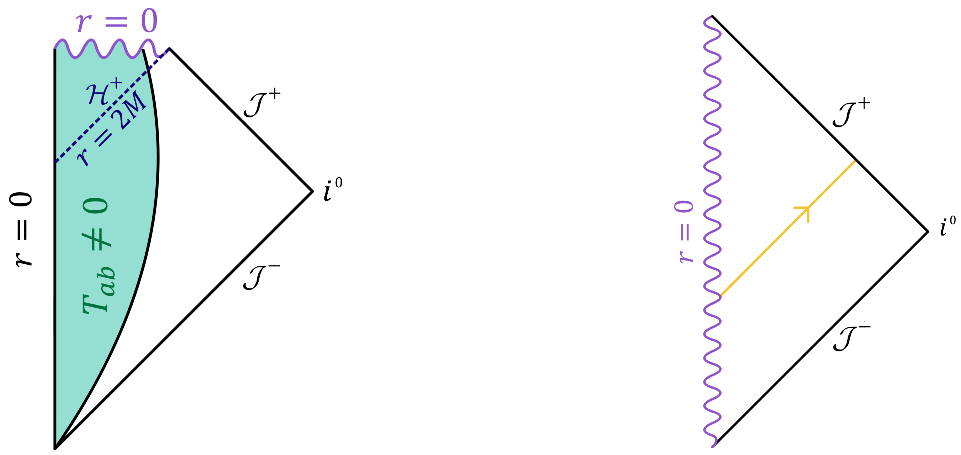
The strong version relates to the impossibility of naked singularities because these would lead to an inability to fully predict the future fate of an observer. Both formulations have problems, as we can find specific spacetimes which satisfy the weak conjecture while violating the strong one; as well as spacetimes for which the strong one is true but the weak one is false. It is still an open problem to find a proper formulation of the conjecture. There are two main reasons why this is a difficult task. The first one is that it is easy to come up with counterexamples. The second one is that cosmological singularities (e.g. the Big Bang) are excluded.

But what is a naked singularity? As we saw in section 2.3.1, the singularity at $r = 0$ in Schwarzschild is enclosed by an event horizon. However, this assumes that $M > 0$ in the metric equation 2.11. For illustrative purposes, let us now imagine that $M < 0$, which still solves EFEs. The metric 2.11 can then be rewritten as:

$$ds^2 = -\left(1 + \frac{2|M|}{r}\right) dt^2 + \left(1 + \frac{2|M|}{r}\right)^{-1} dr^2 + r^2 d\Omega^2. \tag{2.27}$$

The corresponding *Carter-Penrose diagrams* for each of these two cases are shown in figure 2.4.

In subfigure 2.4b, the yellow line represents an outgoing null geodesic travelling from the singularity at $r = 0$ to $\mathscr{J}^+$. This is an example of a naked singularity. The situation gets worse if we consider the more physical scenario of the gravitational collapse of a star forming a black hole while keeping our condition $M < 0$. The corresponding diagram is given in the following figure



(a) $M > 0$ (gravitational collapse) (b) $M < 0$

Figure 2.4: Carter-Penrose diagrams for Schwarzschild with $M > 0$ and $M < 0$. $\mathcal{J}^+$ and $\mathcal{J}^-$ are the future and past null infinity respectively. i^0 is the spatial infinity. $\mathcal{H}^+$ is the future event horizon.

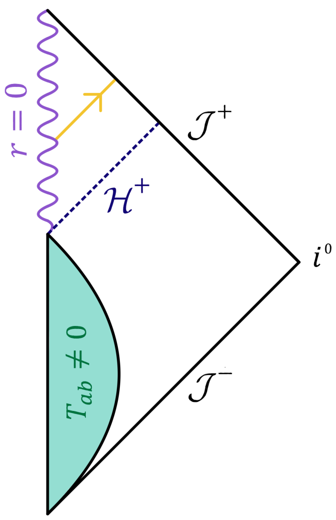


Figure 2.5: Carter-Penrose diagram for the gravitational collapse of Schwarzschild $M < 0$

After the collapse, spacetime is described by our $M < 0$ Schwarzschild where $r = 0$ is singular. Before that, however, $r = 0$ is just the origin of our polar coordinates. We do not know the spacetime metric in the region inside the star where $T_{ab} \neq 0$ but we do not need to. We just know that $r = 0$ there is non-singular. Hence, this collapse generates a naked singularity. Nonetheless, this is ruled out by the *Positive Mass Theorem* proven by Schoen and Yau in 1979 [109, 110] and two years later by Witten in the context of supersymmetry algebra in supergravity [135]. The theorem has also been proven by Penrose, Sorkin and Woolgar using Borde's convergence theorems [100, 115] and by Page, Surya and Woolgar from holography causality [96]. The theorem is stated as follows:

Theorem 2.2 (Positive Mass Theorem). *An isolated, asymptotically flat physical system with T_{ab} satisfying the DEC has a total energy, including contributions from both matter and gravitation which is non-negative.*

Hence, this is a case where the DEC is used[11] to rule out solutions to EFEs with naked singularities.[12] However, this is still far from qualifying as a proof of the censorship conjecture. This is mainly because the censorship conjecture is, after all, a conjecture rather than a theorem that can be proven. A proof would at least require a proper formulation for the conjecture, which we still lack.

However, it seems reasonable that an energy condition should be a key element in the conjecture. Ultimately, what we want is to restrict the possible solutions to EFEs by not allowing naked singularities, and this is the kind of job energy conditions were initially designed to perform.

2.3.3 Other results

There are a number of other important results that one can only derive using energy conditions. One of them is the Hawking Area Theorem for Black Holes. The Area Theorem, also known as the second law of black hole mechanics,[13] states that the area A of the event horizon of a black hole does not decrease in time, i.e. $\delta A \geq 0$ [6, 126]. This theorem was proved by Hawking from the geometric NEC [67, 70].

However, the same problem arises here as it did for the singularity theorems. As we will see in section 2.4, all four pointwise energy conditions are violated. As a result, the original proof by Hawking using the NEC is no longer valid, and a need for proving the theorem with weaker conditions arises. This has been done by Lesourd, who has strengthened the Area Theorem by weakening the energy conditions needed to prove it [82, 83].

Another example of the use of energy conditions in cosmology is the *Cosmic No-Hair Conjec-*

[11] Note that the proof of the theorem from holography causality [96] does not depend on the DEC or any other pointwise energy condition.

[12] See page 305 of Wald [125].

[13] Alluding to the second law of thermodynamics by the analogous roles played by the entropy and the area of a black hole.

ture, which proposes that generic spacetimes with positive cosmological constant $\Lambda > 0$ will locally be described by the de Sitter metric asymptotically at late times. This conjecture, however, requires an energy condition to control the range of possible energy tensors. The conjecture was proved by Wald from a simple argument using the DEC and SEC and the time-like Raychaudhuri equation, which has to be supplemented with a contribution from Λ [124]. This conjecture explains how anisotropies in the early universe can be diluted so that today we see an isotropic universe. It does not, however, resolve the *horizon problem*, which refers to the lack of explanation as to how causally disconnected regions of the universe can be equally homogeneous and isotropic.

2.4 Violations

Some solutions of EFEs violate energy conditions. One might think that such violations are always undesired results that we should avoid. This is sometimes true, but not always. Recall that in section 1.3 we described the job of energy conditions as twofold. On one hand, they are used to prove useful results as the ones we have just seen. On the other hand, they are also useful to rule out unphysical exotic phenomena. This is done by showing that such phenomena violate the conditions. Here we will first consider this positive application of the conditions and later deal with the unwelcome classical violations.[14] The violations involving quantum effects will be left for consideration in chapter 3.

2.4.1 Traversable wormholes & time machines

Exotic matter is defined by Visser [121] as matter which violates the NEC, and consequently the rest (SEC, WEC, DEC).[15] These exotic phenomena are not believed to be actual features of our universe, so energy conditions are introduced to rule them out. The existence of exotic matter would lead to alarming violations of causality and other paradoxes.

[14]We have restricted our attention to a small number of violations but there are other interesting phenomena violating the classical conditions such as a setup of accelerating mirrors or the Hawking radiation.

[15]Note that in [92] Morris and Thorne opt to define exotic matter as that which violates the WEC.

An example of exotic matter are *traversable wormholes.*[16] Before 1988, all wormhole solutions to EFEs were believed to be doomed to nothing more than science fiction. These include the Einstein-Rosen bridge, the Kerr black hole, Wheeler wormholes and wormholes from naked singularities [121].

The radical work in 1988 by Morris and Thorne [92] demonstrated that traversable wormholes constructions are, in principle, in sympathy with our laws of physics. A wormhole, which is roughly a shortcut between two spacelike separated regions in the universe (see figure 2.6), is considered to be traversable if a person can travel through it and come back alive within a reasonable time. Such a wormhole could then be used for interstellar and time travel. For a thorough transverse review of wormhole physics we refer to Visser [121]. We follow the conventions used there.

The metric proposed by Morris and Visser for traversable wormholes is

$$ds^2 = -e^{2\phi(l)}dt^2 + dl^2 + r^2(l)d\Omega^2. \tag{2.28}$$

The l-coordinate ranges from $l \to -\infty$ (one asymptotically flat region) to $l \to \infty$ (the other asymptotically flat region). The radius of the throat of the wormhole r_0 is the minimum value of $r(l)$ which is chosen to be placed at $l = 0$ without loss of generality (see figure 2.6). To make the wormhole traversable, they assume that there is no event horizon (where $e^{2\phi} \to 0$), which implies that $\phi(l)$ has a finite value everywhere.

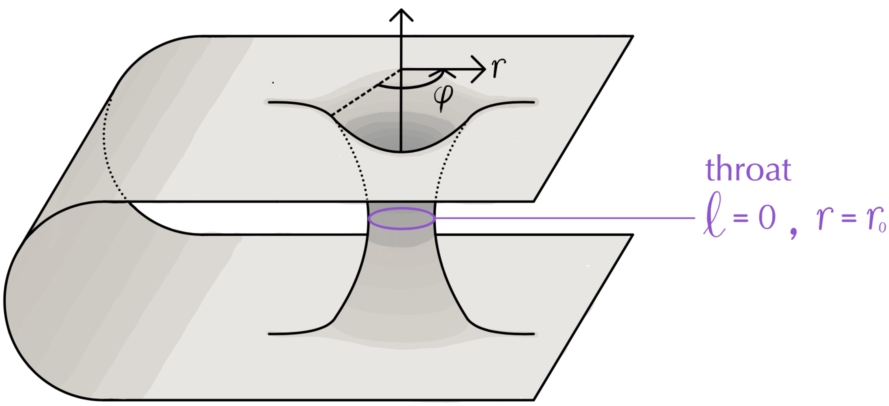


Figure 2.6: Wormhole spacetime. Slice with $\theta = \frac{\pi}{2}$ and $t = \text{const}$.

[16] Another example which we will not consider here are *warp drive spacetimes.*

However, the results are neater if we go to Schwarzschild coordinates (t, r, θ, ϕ)

$$ds^2 = -e^{2\phi_\pm(r)}dt^2 + \left(1 - \frac{b_\pm(r)}{r}\right)^{-1} dr^2 + r^2 d\Omega^2, \tag{2.29}$$

where now we have four functions $\phi_\pm(r)$ (the redshift function) and $b_\pm(r)$ (the shape function) determining the metric instead of the two original ones $\phi(l)$ and $r(l)$. This is because in Schwarzschild coordinates we need two patches each covering one side of the universe and connecting in the throat $[r_0, \infty)$, so the functions double in number.

Our goal is to show that this solution violates the energy conditions, and the procedure to do so is quite straightforward although tedious. We need to arduously compute the Einstein tensor ($g_{ab} \to \Gamma^a_{cb} \to R^a{}_{cbd} \to R_{ab} \to R \to G_{ab}$). Then, we can use EFEs to get T_{ab} and show that it violates the conditions.

The only non-vanishing components of G_{ab} are found to be the following

$$G_{tt} = \frac{b'}{r^2}, \tag{2.30}$$

$$G_{rr} = -\frac{b}{r^3} + 2\left(1 - \frac{b}{r}\right)\frac{\phi'}{r}, \tag{2.31}$$

$$G_{\theta\theta} = G_{\phi\phi} = \left(1 - \frac{b}{r}\right)\left[\phi'' + \phi'\left(\phi' + \frac{1}{r}\right)\right] - \frac{1}{2r^2}(b'r - b)\left(\phi' + \frac{1}{r}\right), \tag{2.32}$$

where the prime denotes an r-derivative. Note that we have not distinguished between $\pm$ functions. This is allowed from continuity conditions at the throat. The corresponding non-vanishing components of the stress-energy tensor are given distinct names

$$\rho := T_{tt} = \frac{b'}{8^2}, \tag{2.33}$$

$$\tau := -T_{rr} = \left[\frac{b}{r^3} - 2\left(1 - \frac{b}{r}\right)\frac{\phi'}{r}\right], \tag{2.34}$$

$$p := T_{\theta\theta} = T_{\phi\phi} = \frac{1}{8\pi}\left[\left(1 - \frac{b}{r}\right)\left[\phi'' + \phi'\left(\phi' + \frac{1}{r}\right)\right] - \frac{1}{2r^2}(b'r - b)\left(\phi' + \frac{1}{r}\right)\right] \tag{2.35}$$

where ρ is the *energy density*, τ is the *radial tension* and p is the *transverse pressure*.

Using the first two equations (2.33 and 2.34) and the special conditions at the throat one can deduce the inequality (see section 11.4 of Visser for details [121]):

$$\exists \tilde{r}_* \Big| \forall r \in (r_0, \tilde{r}_*), \qquad (\rho - \tau) < 0. \tag{2.36}$$

This encodes the fact that $b(r) < r \quad \forall\, r > r_0$. At the throat, where $b(r_0) = r_0$ we have the inequality

$$\rho(r_0) - \tau(r_0) \leq 0. \tag{2.37}$$

These inequalities clearly violate the effective formulation of the NEC in the region close to the throat of the wormhole $(r_0, \tilde{r}_*)$.

It is not very difficult to realise that with traversable wormholes one can effectively achieve superluminal and time travel. Imagine two regions, A and B, connected by a traversable wormhole. First, one can effectively travel faster than light between A and B by taking the wormhole shortcut without needing to surpass the speed of light locally. Secondly, traversable wormholes come inevitably with *closed timelike curves* [57, 66], which allows for the manufacture of time machines [93]. Imagine you are in region A and decide to go to region B via the usual non-wormhole way, which would take a long time. Then, once in region B, you could decide to quickly go back to A through the wormhole and... ta da! You would find yourself before you actually left, thus violating any sensible notion of causality.

To avoid such paradoxical scenarios and "*make the universe safe for historians*", Hawking proposed the *Chronology Protection Conjecture* [66], the proof of which seems to require some sort of energy condition.

2.4.2 Negative cosmological constant Λ

Einstein Field Equations can be modified by adding a cosmological constant term which explains the observational accelerated expansion of the universe:

$$G_{ab} + \Lambda g_{ab} = 8\pi T_{ab}. \tag{2.38}$$

This equation can be rearranged by moving the cosmological term to the RHS and defining the stress tensor $T^{\Lambda}_{ab} = -\frac{\Lambda}{8\pi} g_{ab}$. With it we can show that the SEC is violated with a positive Λ (the one driving the accelerated expansion) [17, 18, 25, 71]. Also, the WEC is violated by a negative cosmological constant [71, 121], which in turn implies the violation of the DEC. The NEC is the only one which remains valid in the presence of Λ:

$$\checkmark \quad \text{NEC} \quad T^{\Lambda}_{ab} k^a k^b = -\frac{\Lambda}{8\pi} k^2 = 0 \qquad \forall \text{ null } k^a, \tag{2.39}$$

$$\times \quad \text{WEC} \quad T^{\Lambda}_{ab} t^a t^b = -\frac{\Lambda}{8\pi} t^2 < 0 \qquad \forall \text{ timelike } t^a \ (t^2 < 0) \text{ and } \Lambda < 0, \tag{2.40}$$

$$\times \quad \text{SEC} \quad \left(T^{\Lambda}_{ab} - \frac{1}{2} T^{\Lambda} g_{ab} \right) t^a t^b = \frac{\Lambda}{4\pi} t^2 < 0 \quad \forall \text{ timelike } t^a \ (t^2 < 0) \text{ and } \Lambda > 0. \tag{2.41}$$

It is noteworthy mentioning that, by taking the cosmological "constant" to be time-dependent, one can make it obey the WEC and SEC. This has been done by Maleknejad and Sheikh-Jabbari in the context of an inflationary model for which the anisotropies of the universe are allowed to increase but within a bounded amount [89].

2.4.3 Classical scalar field ϕ

The classical scalar field produces one of the most plain violations of classical energy conditions. The *non-minimally coupled scalar field* has Lagrangian density [79, 121]:

$$\mathcal{L} = \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2} m^2 \phi^2 + \frac{1}{2} \xi R \phi^2, \tag{2.42}$$

where ξ is the gravity-scalar coupling constant.

The special case where $\xi = 0$ is called the *minimally coupled scalar*, and it can be easily proved to violate the SEC. Its stress tensor is given by [71, 79, 121]

$$T_{ab} = (\nabla_a \phi)(\nabla_b \phi) - \frac{1}{2} g_{ab}((\nabla\phi)^2 + m^2 \phi^2), \tag{2.43}$$

so, after taking the trace, we can write for any timelike vector t^a

$$\left(T_{ab} - \frac{1}{2}T g_{ab}\right) t^a t^b = (t^a \nabla_a \phi)^2 - \frac{1}{2}m^2\phi^2. \tag{2.44}$$

We can easily find a situation where the RHS is negative and thus the SEC is violated. Take, for example, $m > 0$ with $t^a = (1, 0, 0, 0)$ and ϕ being time-independent at an arbitrary point in that frame (see page 120 of Visser [121]). We cannot have a vanishing time derivative of the scalar everywhere because then it would not have any dynamics, and if the scalar is non-zero then by the mass term we know that it must have some oscillatory dynamics. However, to prove the SEC is violated we only need this to be satisfied at a point, which is unproblematic. Another physical example would be a scalar field constant over some spatial region but a function of time. This would effectively be described as a harmonic oscillator in time, where in the extremes of the potential all the derivatives, including time, vanish.

The rest of the pointwise energy conditions are satisfied by the minimally coupled scalar [121]

✓	NEC	$T_{ab}k^a k^b = (k^a \nabla_a \phi)^2 \geq 0$	$\forall$ null k^a,	(2.45)
✓	WEC	$T_{ab}t^a t^b = (t^a \nabla_a \phi)2 + \frac{1}{2}((\nabla\phi)^2 + m^2\phi^2) \geq 0$	$\forall$ timelike t^a,	(2.46)
✓	DEC	$\lvert T^{ab}t_a\rvert^2 = -(t^a \nabla_a \phi)^2 m^2\phi^2 - \frac{1}{4}((\nabla\phi)^2 + m^2\phi^2)^2 \leq 0$	$\forall$ timelike t^a.	(2.47)

However, one can show that all of them are violated by the non-minimally coupled scalar (i.e. $\xi \neq 0$). It is sufficient to show that it is violated by the NEC, and the other violations will follow automatically. The stress tensor for the non-minimally coupled scalar is the same as in equation 2.43, which we now denote $T_{ab}^{(\xi=0)}$ with a few extra terms [79]

$$T_{ab}^{(\xi)} = T_{ab}^{(\xi=0)} + \xi(g_{ab}\Box_g - \nabla_a\nabla_b - G_{ab})\phi^2 \tag{2.48}$$

The two last extra terms can be made negative so that, after contracting with null vectors, the NEC is violated. An example provided by Kontou and Sanders is when $\nabla\phi = 0$ and the terms with second derivative dominate over the terms which are proportional to ϕ^2 [79].

Is this violation physically relevant? Should we count it as evidence against the validity of the pointwise ECs? The answer depends on the empirical reality of scalar fields. That is, are scalar fields real components of our universe? If they are, then surely this invalidates the violated conditions. Until the first decade of this century, there were slight concerns regarding the physical reality of scalar fields. Despite being used ubiquitously by theoreticians as useful and pedagogical toy models, particles described by scalar fields were either shown to be not fundamental or lacking experimental evidence [5].

Before the experimental discovery of the Higgs boson in 2012 [1, 22], the only physical evidence for scalar fields were the mesons such as the pion π. Mesons, however, are known to be composed by quark-antiquark pairs so they are not fundamental particles and their description as scalar fields starts to fail at high energies [5].

Another particle described as a scalar field is the *axion*, a hypothetical elementary particle which was proposed in 1977 by Peccei and Quinn [97, 98] to provide a solution to the strong CP problem in Quantum Chromodynamics. They are also cold dark matter candidates. In fact, some very recent (June 2020) anomalous results from the XENON1T experiment suggest axions as their most probable explanation [2, 74, 118].[17]

There are other scalar field candidates such as the *Brans-Dicke scalar* or the *dilaton* of membrane theories [5]. Aside of these, the experimental detection of the Higgs particle seems to close the debate on the physical reality of scalars. Usually we think of the Higgs boson as a minimally coupled scalar, which would entail only the violation of the SEC. In the recent years, however, there has been some interest in using the Higgs as the scalar driving inflation, i.e. a *Higgs inflaton*, which would need to be non-minimally coupled thus violating all the classical conditions [7]. To keep the dispute alive one would have to argue that the Higgs is not a fundamental particle and that it is made up of other elementary components. Hence, it is most prudent to take the violation of ECs by scalar fields as a serious issue that has to be addressed.

[17] See [28] for a counterargument to that interpretation.

Chapter 3

Road to Quantum Energy Conditions: the Quantum Interest Conjecture

All pointwise classical energy conditions are manifestly violated by quantum effects both experimental and theoretically [31]. In this chapter we review the main quantum violations of the classical ECs that led to the proposal of new versions (studied in chapters 4 and 5). These violations are very peculiar. They seem to follow a trend of being small in magnitude and/or duration. These tendencies are articulated with the *Quantum Interest Conjecture*, which we will state after looking at the violations. The conjecture not only motivates the proposal of quantum energy conditions but also guides us in this profound endeavor. In this chapter we will work in Minkowski spacetime $\mathbb{M}^4$, since it only takes a violation in flat space to show that an energy condition is false.

3.1 The Casimir Effect

The *Casimir effect* refers to the presence of local negative energy densities induced by a pair of parallel conducting plates in vacuum [19].[1] Hence, it clearly violates all the classical pointwise ECs. As a result of the negative energy density, the plates will experience an attractive force

[1]For experimental confirmation of the Casimir effect see [14, 112, 116, 117]

between them that can be measured. To demonstrate the violations, we have to construct the corresponding stress-energy tensor entirely from symmetry considerations. The physical system is depicted in the following diagram

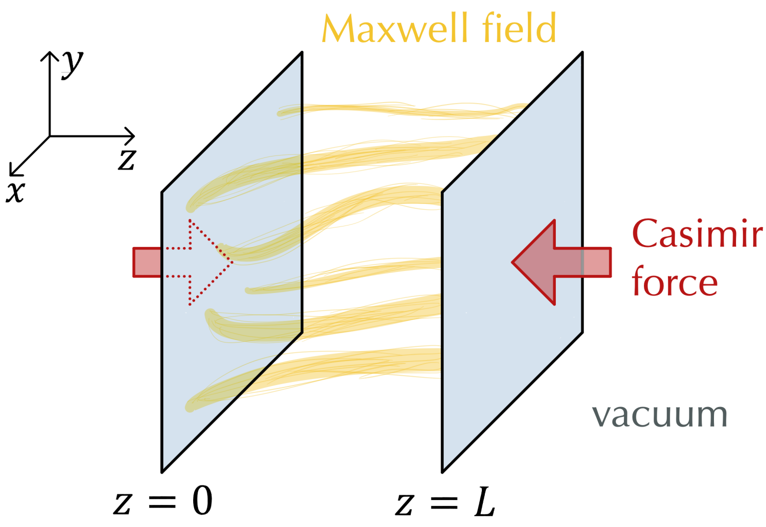


Figure 3.1: Casimir setting

Because of how we placed the plates, the only possible boundary condition on the electrodynamic field will be in the z-direction and will depend on the distance between the plates L. The stress tensor T^{CAS}_{ab} for the Casimir Maxwell field is symmetric in its two indices, so we need symmetric terms to construct it. The only options available are the $\mathbb{M}^4$ metric $\eta_{ab} = \text{diag}(-1, 1, 1, 1)$ and the combination $z_a z_b$, where z^a is the normalized vector in the z-direction, i.e. $z^a = (0, 0, 0, 1)$ (see page 121 of [121] for further details). We have to include functions multiplying each term for generality which can only depend on L and z. We choose to write the tensor as

$$T^{\text{CAS}}_{ab} = \frac{1}{L^4}\left(f_1\left(\frac{z}{L}\right)\eta_{ab} + f_2\left(\frac{z}{L}\right)z_a z_b\right), \tag{3.1}$$

where the factors are placed taking into account dimensional analysis and making the functions dimensionless. The two functions are actually just constants proportional to each other. This is obtained from the fact that T^{CAS}_{ab} is both conserved (i.e. $\partial_a T^{ab}_{\text{CAS}} = 0$) and traceless from conformal invariance of the electromagnetic field (i.e. $T^{\text{CAS}} = T^{\text{CAS}}_{ab}\eta^{ab} = 0$). We get $f_1 = C$ and $f_2 = -4f_1 = -4C$. Hence,

$$T^{\text{CAS}}_{ab} = \frac{C}{L^4}\left(\eta_{ab} - 4z_a z_b\right). \tag{3.2}$$

The constant C can be determined in different ways [19, 52, 102] and it is found to be $C = \frac{\pi^2}{720}$. Then the final expression for the stress tensor is:

$$T_{ab}^{\text{CAS}} = \frac{\pi^2}{720L^4}\left(\eta_{ab} - 4z_a z_b\right) = \frac{\pi^2}{720L^4}\text{diag}(-1, 1, 1, -3) \tag{3.3}$$

The T_{00}^{CAS} component gives the energy density and the T_{33}^{CAS} gives the pressure in the z direction

$$\rho = -\frac{\pi^2}{720L^4} \qquad\qquad p = -3\frac{\pi^2}{720L^4} \tag{3.4}$$

which violate all the classical ECs. The negative energy density is not just a product of the idealisation and symmetrical impositions of the physical system in our analysis since it was found that negative energies still dominate in physically realizable systems [114]. There are two striking features of this effect. First, that it is negative. Second, that it is tiny in magnitude. This can be seen after reinstating factors of $\hbar$ and c, which gives the magnitude

$$|\rho| = \left|-\frac{\hbar c\pi^2}{720L^4}\right| \ll 1. \tag{3.5}$$

3.1.1 Casimir wormholes

Traversable wormholes violate the energy conditions, as they have to obey the inequality 2.37. We have just now seen that such negative energy densities can be physically realised by the Casimir effect. This begs the question: can we construct traversable wormholes with the Casimir effect?

In another 1988 paper, Morris, Thorne and Yurtsever [93] briefly considered this possibility, but they rapidly ruled it out by arguing that such construction would require the distance between the plates L to be smaller than the Compton wavelength of the electron, thus making it physically unfeasible.

A more detailed analysis is provided by Visser [121] where he proposes a more realistic stress-energy tensor that takes into account the mass density σ of the metallic plates

$$T_{ab}^{\text{CAS}} = \sigma \hat{t}_a \hat{t}_b \left[\delta(z) + \delta(z-a)\right] + \Theta(z)\Theta(a-z)\frac{\pi^2}{720L^4}\left(\eta_{ab} - 4\hat{z}_a\hat{z}_b\right), \tag{3.6}$$

where $\hat{t}^a$ and $\hat{z}^a$ are normalized vectors in the t and z-directions. It turns out that it is practically impossible to build a wormhole with this stress tensor, as the mass of the plates will eclipse the Casimir energy.

On a different note, equations 3.4 can be thought to imply the equation of state $p = \omega\rho$ with $\omega = 3$. In a very recent paper, using this equation of state and performing an analysis on the shape and redshift functions $b(r)$ and $\phi(r)$ of the wormhole metric (see section 2.4.1), Garattini demonstrates that casimir energy can really source a traversable wormhole [56]. From a general analysis of $b(r)$ and $\phi(r)$ he finds the following inequality implications:

$$\omega r_1^2 - r_0^2 > 0 \quad \Rightarrow \quad \text{Black Hole,} \tag{3.7}$$

$$\omega r_1^2 - r_0^2 = 0 \quad \Rightarrow \quad \text{Traversable Wormhole,} \qquad r_1^2 = \frac{\pi^3 l_p^3}{r_0 90} \tag{3.8}$$

$$\omega r_1^2 - r_0^2 < 0 \quad \Rightarrow \quad \text{Singularity,} \tag{3.9}$$

where l_p is the Planck length and r_0 is the radius of the throat as defined in section 2.4.1. Focusing on the wormhole case, Garattini finds that it corresponds to the stress tensor

$$T_{ab} = \frac{r_0^2}{8\pi\omega r^4}\Big[\text{diag}(-1,-\omega,1,1) + (\omega_t(r)-1)\text{diag}(0,0,1,1)\Big] \tag{3.10}$$

$$= T_{ab}^{\text{CAS}} + \frac{r_0^2}{8\pi\omega r^4}(\omega_t(r)-1)\text{diag}(0,0,1,1). \tag{3.11}$$

where, to go from the first to the second line, we have connected the first term to the Casimir stress tensor in equation 3.3. To do so, we specified $\omega = 3$, made the identification $\frac{r_0^2}{8\pi\omega} = \frac{\pi^2}{720}$ and changed from radial to cartesian spatial coordinates $(t, r, \theta, \phi) \to (t, x, y, z)$, which results in

$$T_{ab}^{\text{rad CAS}} := \frac{r_0^2}{8\pi\omega r^4}\text{diag}(-1,-\omega,1,1)^{\text{radial}} = \frac{\pi^2}{720L^4}\text{diag}(-1,1,1,-3)^{\text{cartesian}} = T_{ab}^{\text{CAS}}. \tag{3.12}$$

The additional term in 3.10 comes from the fact that here our prefactor in T_{ab} is an r-dependent function rather than just a constant as in 3.3. This wormhole is of Planckian size [56].

In another recent paper, Maldacena, Milekhin and Popov [86] claim to have found a wormhole solution of Einstein Maxwell theory with a charged massless Weyl fermion coupled to the $U(1)$ gauge field with action given by

$$S_{EM\chi} = \int d^4x\sqrt{g}\left(\frac{1}{16\pi}R - \frac{1}{4g^2}F^2 + i\overline{\chi}(\not\nabla - i\not A)\chi\right). \tag{3.13}$$

The wormhole they find is a *long* traversable wormhole, where long refers to the fact that the path through the interior of the wormhole is longer than the path outside it. Hence, this wormhole does not entail causality violations. Maldacena et al. [86] divide the spacetime in the 3 regions shown in the diagram 3.2. The patches were two of the regions overlap have coinciding metrics.

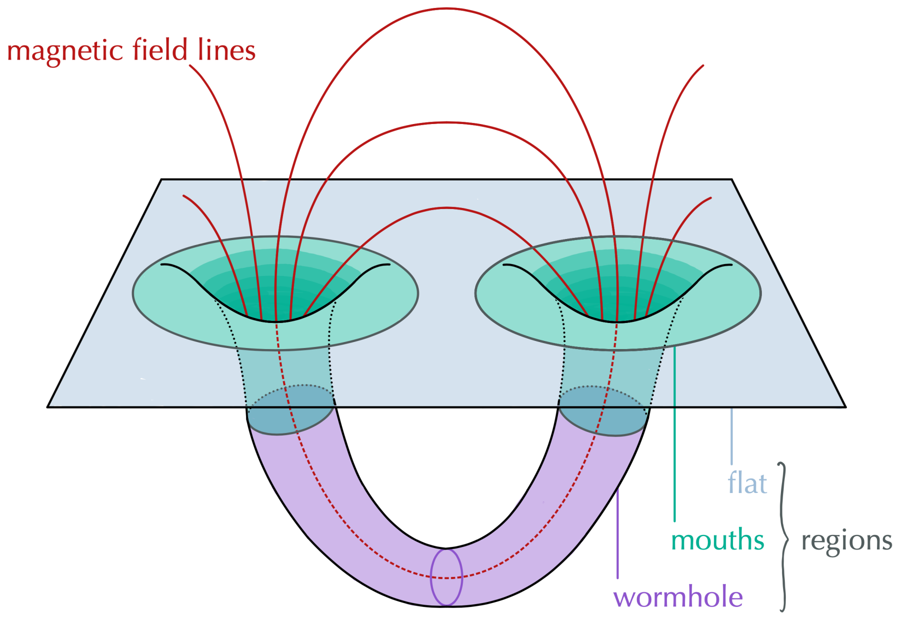


Figure 3.2: Long traversable wormhole: 3 regions and magnetic field lines.

The nearly flat space region (greyish blue) describes the two wormhole mouths as point-like charged masses. The spacetime geometry in each of the two mouths (green) is that of a black hole. Finally, the geometry of the wormhole region (purple) is approximately $AdS_2 \times S^2$ and is given by the line element

$$ds^2 = r_e^2 \left(-(\rho^2 + 1)d\tau^2 + \frac{d\rho^2}{\rho^2 + 1} + d\Omega^2 \right). \tag{3.14}$$

This solution includes a magnetic field living in S^2 also shown in the diagram. We have seen that wormholes require negative energy. In this case, the Casimir-like energy comes from the massless fermion travelling along the circular magnetic field lines in and out of the wormhole with radius L. From the action, we see that we have a single 4-dimensional Dirac fermion. One of the key results of this paper is that in the throat of the wormhole the 4-dimensional fermion effectively behaves as a number q of 2-dimensional Dirac fermions moving independently of each other localised on *Landau levels*. The number q is related to the magnetic fields of the black holes that join together to form the wormhole. The 2-dimensional stress tensor for the q complex fermions has components

$$\hat{T}_{tt} = \hat{T}_{xx} = -\frac{q}{8\pi l^2} \frac{1}{4\pi r_e^2}, \tag{3.15}$$

where the hat means that the trace part has been extracted. The 4-dimensional stress tensor will come from these 2-dimensional fermionic contributions as well as the 4-dimensional contribution of the magnetic field. For large q this setting provides us with the negative energies needed to construct the traversable wormhole.

In short, we have seen several examples of how the Casimir effect and traversable wormholes are intimately connected. This exciting topic would suffice for an entire dissertation on its own and we hope to see more results in the near future.

3.2 Quantum Coherence Effects

It is very easy to cook up quantum states that violate the NEC by superposing the vacuum with particle states where the negative signs can be traced back to the effect of *normal ordering*.

First, let us show an example, taken from [37], that shows how negative signs arise when we give a quantum description of quantities that would classically always be positive. Consider the state

$$|\psi\rangle = \cos\epsilon\,|0\rangle + \sin\epsilon\ {:}\Phi^2{:}\ (f)\,|0\rangle\,, \tag{3.16}$$

where ${:}\Phi^2{:}\ (f)$ is defined by

$${:}\Phi^2{:}\ (f) := \int d^4x\ {:}\Phi^2{:}\ (x) f(x), \tag{3.17}$$

where $\Phi(x)$ is the real scalar field and $f \in C_0^\infty(\mathbb{R}^4)$. The scalar field can be written as the integral

$$\Phi(x) = \int \frac{d^3\mathbf{k}}{(2\pi)^3\sqrt{2w}} \left(e^{-ik\cdot x} a(\mathbf{k}) + e^{ik\cdot x} a^*(\mathbf{k})\right), \qquad k^\mu = (w, \mathbf{k}), \quad w = \sqrt{|\mathbf{k}|^2 + m^2} \tag{3.18}$$

where $a^*(\mathbf{k})$ and $a(\mathbf{k})$ are the usual creation and annihilation obeying the commutation relations

$$[a(\mathbf{k}), a^*(\mathbf{k})] = 0 \qquad\qquad [a(\mathbf{k}), a^*(\mathbf{k})] = (2\pi)^3\delta^{(3)}(\mathbf{k} - \mathbf{k}')\mathbb{1}. \tag{3.19}$$

Bracketing ${:}\Phi^2{:}\ (f)$ with the vacuum and Fourier transforming the function $f(x)$ we can calculate the *vacuum expectation value* [37], which is found to vanish due to the normal order $\langle 0|\ {:}\Phi^2{:}\ (f)\,|0\rangle = 0$. To calculate this *vev*, one has to use the commutation relations and the fact that all the $a(\mathbf{k})$ operators annihilate the vacuum ket and all the $a^*(\mathbf{k})$ annihilate the vacuum bra. However, the norm of the vector ${:}\Phi^2{:}\ (f)\,|0\rangle$ is

$$\left|{:}\Phi^2{:}\ (f)\,|0\rangle\right|^2 = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{d^3\mathbf{k}'}{(2\pi)^3} \frac{\left|\hat{f}(k + k')\right|^2}{2ww'}. \tag{3.20}$$

Hence, from this we see that ${:}\Phi^2{:}\ (f)$ does not annihilate the vacuum.[2] These two facts combined lead us to the conclusion that ${:}\Phi^2{:}\ (f)$ has to have some negative values.

[2]The norm is zero only when f=0. [37]

Choosing f so that $|{:}\Phi^2{:}\,(f)\,|0\rangle|^2 = 1$ the expectation value of our cooked up state is [37]

$$\begin{aligned}\langle\psi|\,{:}\Phi^2{:}\,(f)\,|\psi\rangle &= \cos^2\epsilon\,\langle 0|\,{:}\Phi^2{:}\,(f)\,|0\rangle + \cos\epsilon\sin\epsilon\,\langle 0|\,({:}\Phi^2{:}\,(f))^2\,|0\rangle \\ &\quad + \cos\epsilon\sin\epsilon\,\langle 0|\,({:}\Phi^2{:}\,(f))^2\,|0\rangle + \sin^2\epsilon\,\langle 0|\,({:}\Phi^2{:}\,(f))^3\,|0\rangle \\ &= 2\cos\epsilon\sin\epsilon + \sin^2\epsilon\,\langle 0|\,({:}\Phi^2{:}\,(f))^3\,|0\rangle \\ &= \sin 2\epsilon + \sin^2\epsilon\,\langle 0|\,({:}\Phi^2{:}\,(f))^3\,|0\rangle \\ &\approx 2\epsilon + \mathcal{O}(\epsilon^2) \qquad \text{for small } \epsilon. \end{aligned} \tag{3.21}$$

Hence, with a negative small ϵ we can get negative expectation values. This example shows that the positivity of some quantities is a classical feature that becomes more ambiguous and obscure once we move to a quantum context.

Now, let us see another simple example of a cooked up bosonic state that, using the same type of argument, will give us negative energy density. Consider the state [46]

$$|\psi\rangle = N(|0\rangle + \epsilon\,|2\rangle) \tag{3.22}$$

The expectation value of the normal ordered time-time component of the stress-energy tensor will give us the local energy density

$$\rho = \langle\psi|\,{:}T_{tt}{:}\,|\psi\rangle = N^2\left[2\mathrm{Re}(\epsilon\,\langle 0|\,{:}T_{tt}{:}\,|2\rangle) + |\epsilon|^2\,\langle 2|\,{:}T_{tt}{:}\,|2\rangle\right], \tag{3.23}$$

where the vacuum expectation value has also vanished for the same reason as in our previous example. Likewise, provided the first term does not vanish, it will dominate for small ϵ; and the sign of ϵ can then be chosen to make ρ negative.

In fact, these are just two examples of a much more general and profound fact about positivity in the quantum formalism. In the original paper that proved that quantum fields lead to negative energies and thus violations of energy conditions, Epstein, Glasser and Jaffe proposed the following argument [31]. Let A be a positive local observable with vanishing vacuum expectation value. Then,

$$\left|A^{\frac{1}{2}}\left|0\right\rangle\right|^2 = \left\langle 0\right| A \left|0\right\rangle = 0 \qquad \Rightarrow \qquad A^{\frac{1}{2}}\left|0\right\rangle = 0 \qquad \Rightarrow \qquad A\left|0\right\rangle = 0. \tag{3.24}$$

A measurements in the vacuum will produce random results. The fact that the vacuum expectation value of A is zero can either mean two things [37, 79]

1. Outcome of A is always zero $A\left|0\right\rangle = 0$ with probability 1,
2. Outcome spectrum of A prolongs to negative values.

However, if it is the case that $A\left|0\right\rangle = 0$, then from Reeh-Schlieder theorem [62] we know that A is not a positive operator in the sense that $\left\langle\psi\right| A \left|\psi\right\rangle \geq 0, \quad \forall \left|\psi\right\rangle$. The theorem tells us that no local positive operators can annihilate the vacuum. Or, in other words, if $A\left|0\right\rangle = 0$ then $\left\langle 0\right| A \left|0\right\rangle = 0$. Hence, the positivity condition of A is necessarily lost.

It is important to note that these examples, contrary to the Casimir effect, do not set a lower bound on the value of energy densities. That is, the local negative values can be made arbitrarily large [42, 43]. However, there are different type of non-local constraints on negative energies. One of them is that the total energy integrated over all space in QFT is positive, even if some local energy can go negative. If instead of integrating over all space one chooses to integrate only along a timelike geodesic one obtains a *quantum inequality* [42, 43, 46] which generally looks like

$$\int_{-\infty}^{\infty} d\tau \left\langle\psi\right| {:}T_{tt}{:} \left|\psi\right\rangle(\tau) f(\tau) \geq -\frac{c}{\tau_0^4}, \tag{3.25}$$

where $\left\langle\psi\right| {:}T_{tt}{:} \left|\psi\right\rangle(\tau)$ represents the local energy density and τ is the proper time along the geodesic. c is just a dimensionless constant and τ_0 represents the width of the function $f(\tau)$ [46]. This inequality is very revealing, as it tells us that negative energies cannot be long-lived. The larger the value of τ_0 is, the smaller the magnitude of the energy will be. Hence, even if local energy can be arbitrarily large, the same energy density integrated along a causal curve has to be constrained in some way.

3.3 The Quantum Interest Conjecture

Broadly speaking, we have seen that negative energy densities produced by quantum effects are bounded either in their magnitude or duration. These are good news, because it implies that classical pointwise ECs, although surely violated in QFT, can be thought of as being approximately true.

These tendencies are captured by the *Quantum Interest Conjecture*, which receives its name from an analogy with the financial term [51]. The conjecture states that negative energies must be overcompensated by positive energies where by overcompensation we mean that positive energy has to dominate over the negative. The amount of overcompensation required will be determined depending on the magnitude and duration of the negative energies. In other words, you are allowed to borrow negative energies from the universe. But be careful, because you will have to repay that energy with an interest which will grow both as you borrow more negative energy and as you keep that negative energy for longer. The interest conjecture is mathematically expressed in the form of quantum inequalities as equation 3.25.

3.4 Foreword to chapters 4 and 5

All our discussion so far has led us to a crossroad. We know that classical pointwise energy conditions are simply false in QFT. Hence, because energy conditions are key pieces in our understanding of the world around us, we find ourselves in need of new weaker versions in agreement with quantum effects.

Since this hurdle was first discovered, there have been mainly two strategies, both of them inspired by the ideas behind the quantum interest conjecture. Locality and positivity are two features of the classical ECs that immediately become ambiguous in the quantum context. Hence, we are presented with two options if we want to design valid quantum energy conditions. We can either sacrifice locality or we can sacrifice positivity.

The first strategy, which we study in chapter 4, sacrifices locality by imposing energy conditions on a larger set of points rather than at every spacetime point. This allows us to have unbounded local negative energies as long as the energy density integrated over our chosen set of points, which will be a geodesic worldline, is non-negative. Such conditions are called *average energy conditions*.

The second strategy, which we consider in chapter 5, attempts to maintain the idea of imposing conditions locally but instead sacrifices positivity by lowering the bar to allow a finite range of negative energies. This bounds the magnitude and duration of negative energies or fluxes of negative energies locally.

The two alternatives are not incompatible. In fact, the study of the relations and implications between the two types of conditions is an interesting and active field of research, and we will see an example of their connection in section 5.4.

Our assessment of the merit of any new energy condition will be based on how well it performs in the following tasks:

- Keep producing the original classical results (e.g. singularity theorems),
- Being obeyed by all "physically reasonable" systems (specially quantum fields) and violated by all "physically unreasonable" systems (e.g. traversable wormholes), and
- Being recovered or proven from general considerations in semiclassical or quantum gravity.

Ideally, the condition should be valid for any physical state in any physical Lorentzian spacetime. From now, one has to work at least in the framework of *semiclassical gravity*, where we quantize the stress-energy tensor in EFEs while keeping the geometry side classical. EFEs become

$$G_{ab} = 8\pi \langle T_{ab}^{ren} \rangle_w \,, \tag{3.26}$$

where $\langle T_{ab}^{ren} \rangle_w$ is the expectation value of the renormalized stress-energy tensor and w represents a large class of physical states known as Hadamard states.[3] There are several philosophical and technical nuisances associated with semiclassical gravity that we will not delve into. An

[3] For details on Hadamard states see [37, 39]

example are quantum fluctuations of spacetime. Even if spacetime is treated as a classical object, by virtue of EFEs the fluctuations in the stress-energy tensor will incite fluctuations of the spacetime geometry. As a result, our initial assumption that spacetime is a classical object is not entirely true. However, it is believed that semiclassical gravity is nonetheless a good approximation to any quantum gravity theory.

Our mission is to find an energy condition that is universally true. For this reason, we are interested in proofs in curved spacetimes and interacting theories since these are the most general arenas available to us. Proofs in flat spacetime and free theories were the first ones to appear, but we will dedicate our attention to the much more interesting and recent proofs involving interacting fields in curved spacetime.

Chapter 4

Averaged Energy Conditions

In this chapter we explore the first strategy of proposing quantum energy conditions. The main idea is to integrate the expectation value of classical pointwise energy conditions over causal geodesics and require the averaged value to be non-negative. First, we will briefly review the family of different *averaged conditions* that have been discussed in the literature. For the rest of the chapter, we will mainly focus on the AANEC, the *achronal averaged null energy condition*, which is currently the energy condition in best shape. We will revisit the implications, consequences and violations for the AANEC, which will showcase what makes this condition so appealing. The main part of this chapter will be devoted to present a number of different proofs of the AANEC in semiclassical and quantum gravity involving different ideas of current research activity. The AANEC will be revisited at the end of next chapter (section 5.4) where it will be compared to the QNEC.

4.1 The Achronal Averaged Null Energy Condition and its siblings

The classical violations considered in 2.4.2 and 2.4.3 have shown that the NEC is in a better shape than the rest of pointwise energy conditions. Thus, it is reasonable to think that we should depart from the NEC to design an averaged condition. This leads us to the

Averaged Null Energy Condition (ANEC)

$$\int_{-\infty}^{\infty} d\lambda T_{ab} k^a k^b \geq 0 \qquad \forall \text{ null } k^a \text{ tangent to a null geodesic } \gamma(\lambda), \tag{4.1}$$

We can classify null geodesics in the two categories *achronal* and *chronal*. A null geodesic is achronal iff no two points on it can be joined by a timelike curve [71]. In other words, timelike curves cannot cross achronal geodesics more than one time. Complete achronal null geodesics are also commonly known as *null lines* [127]. We can use this to create a finer condition:

Achronal Averaged Null Energy Condition (AANEC)

$$\int_{-\infty}^{\infty} d\lambda\ T_{ab} k^a k^b \geq 0 \qquad \forall \text{ null } k^a \text{ tangent to a null line } \gamma(\lambda), \tag{4.2}$$

For reasons that will become apparent in section 4.2, this turns out to be our best shot at writing a universally true energy condition. We have reasons to believe that the AANEC is a fundamental property of all gravitationally interacting matter in semiclassical gravity which, needless to say, presents an exciting opportunity to test quantum gravity ideas.

Some siblings of the ANEC that have also been considered in the literature include:

Averaged Weak Energy Condition (AWEC)

$$\int_{-\infty}^{\infty} d\tau\ T_{ab} t^a t^b \geq 0 \qquad \forall \text{ timelike } t^a \text{ tangent to a timelike geodesic } \gamma(\tau), \tag{4.3}$$

Averaged Strong Energy Condition (ASEC)

$$\int_{-\infty}^{\infty} d\tau \left(T_{ab} - \frac{1}{2} T g_{ab}\right) t^a t^b \geq 0 \quad \forall \text{ timelike } t^a \text{ tangent to a timelike geodesic } \gamma(\tau), \tag{4.4}$$

Averaged Dominant Energy Condition (ADEC)

$$\int_{-\infty}^{\infty} d\tau\ T_{ab} t^a \xi^b \geq 0 \quad \forall \text{ co-oriented timelike } t^a \text{ and } \xi^a \text{ tangent to timelike geodesics } \gamma(\tau), \tag{4.5}$$

Note that here we have only presented the physical formulation of the averaged conditions. This is because it is the predominant formulation used in proofs and applications in the recent literature. However, nothing stops us from writing the geometrical and effective formulations exactly in the same way we did for the pointwise conditions, and our worry regarding to the actual fundamental meaning of energy conditions does not fade away. An example is the *averaged null convergence condition* (ANCC), which is the ANEC in its geometrical formulation.

We can simplify these statements with a shorthand notation. Let us do it just for the AANEC. We define $T_{kk} = T_{ab}k^a k^b$ where k^a is a null vector tangent to a null line $\gamma(\lambda)$. Then, the AANEC becomes the statement that $\int_{-\infty}^{\infty} d\lambda T_{kk} \geq 0$. Also, note that the NEC is the statement $T_{kk} \geq 0$.

4.2 Applications revisited for the AANEC

All these averaged conditions are clearly weaker than their corresponding classical pointwise ECs. More specifically, we have that NEC $\Rightarrow$ ANEC $\Rightarrow$ AANEC, and similarly for the others. This can be easily obtained by integrating the pointwise conditions along the corresponding geodesics. In addition, we also have the implications ADEC $\Rightarrow$ AWEC $\Rightarrow$ ANEC and ASEC $\Rightarrow$ ANEC that follow from the classical implications.[1] With these we can upgrade the diagram for the logical relations between ECs.

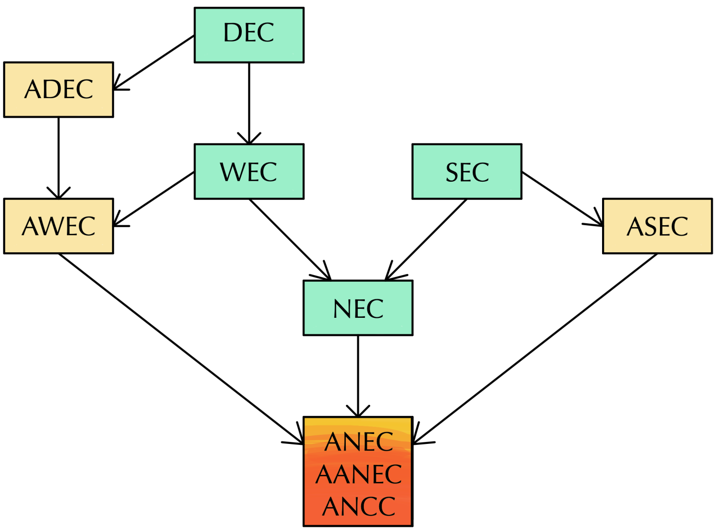


Figure 4.1: Logical relations of classical and averaged ECs.

[1] Whether the ADEC, AWEC and ASEC imply the NEC is less interesting and not as clear, so we leave it out of the discussion.

Because these new conditions are weaker, one has to show that they can still be used to produce the classical results. For instance, singularity theorems have been proven from averaged conditions in a classical setting using different methods [10, 11, 54, 104, 105, 119]. In particular, the arguments in [54, 104, 105] show that singularities are present in solutions where the ANEC is obeyed. Another example is Hawking's Area Theorem for black holes, which has been obtained from a *damped* version of the ANEC by Lesourd [82, 83]. In general, it has been shown that the averaged conditions and, in particular the ANEC, do a good job at reproducing the classical results.

However, these conditions are not free of violations either. Even the ANEC is found to be violated in some settings. For instance, if we take $\mathbb{M}^4$ and we compactify it in one spatial dimension then the quantum scalar field violates the ANEC. To get a Casimir-like effect we could compactify the z direction by associating the plane $z = L$ with the plane $z = 0$. Then, a geodesic for which the value of z changes will not obey the ANEC. Note that such geodesics are *chronal*. Another example where a quantum scalar field violates the ANEC is in the spacetime around a black hole, i.e. Schwarzschild, where all the complete null geodesics are also *chronal* [122]. Finally, in the paper we considered in section 3.1.1 that proposed a traversable wormhole sourced by a charged massless fermion [86] the authors show that the ANEC is violated in a null geodesic in a cylindrical spacetime, which implies that the geodesic is *chronal*. Nonetheless, the ANEC is an improvement over the NEC, since it is obeyed by Casimir-like energies [41, 59, 58] where the NEC will be violated.

These three violations share the common feature of *chronality*. This is what motivates the use of the *achronal* ANEC. In a short yet powerful paper, Graham and Olum show that the AANEC suffices to produce all the classical results[2] while avoiding the non-desired violations of the ANEC [60]. The key point is that requiring achronality ha no effect on proving the theorems since usually it is with achronal geodesics that we prove them. The authors conjecture that the AANEC is obeyed in all *self-consistent*[3] solutions of semiclassical gravity.

[2] In the paper they consider topological censorship, closed timelike curves, positive mass theorems, singularity theorems and superluminal signals.

[3] For a discussion on what self-consistency entails we refer to their paper [60].

4.3 Proving the AANEC

The ANEC has been checked to be obeyed in Minkowski for arbitrary QFTs [34, 48, 44, 76, 85, 120]. While this gives good confidence in the soundness of the condition, it does not tell us whether or not it is a generally valid statement. This is why here we decide to focus our attention on the more recent proofs which involve interacting CFTs and/or curved spacetimes. The AANEC has recently received a significant amount of research interest and it has been shown from very different angles that it is actually a fundamental property of QFT. Here, we will present what we believe are the four most interesting proofs of the AANEC. In order, they connect the AANEC with the Generalized 2$^{\text{nd}}$ Law (GSL), causality, holography and quantum information. The following table summarizes the general validity of these proofs.[4]

	GR	**QFT**
GSL [127]	arbitrary curved spacetime	any minimally coupled quantum fields
Causality [65]	$\mathbb{M}^4$, $d > 2$	interacting CFTs
Holography [75]	asymptotic AdS	interacting CFTs
Quantum Information [35, 106]	$\mathbb{M}^4$, dS, AdS and extremal horizons	interacting CFTs

Table 4.1: General validity of AANEC proofs.

4.3.1 AANEC from the generalized 2$^{\text{nd}}$ Law

The first proof we consider is based on a 2010 paper by Wall [127]. The main idea behind the proof is that if the *generalized 2nd law of black hole thermodynamics* is true, then this must entail some kind of restriction on the possible form of the stress-energy tensor. This restriction, under certain conditions, turns out to be the ANEC on null lines, i.e. the AANEC.

The generalized 2$^{\text{nd}}$ law states that the generalized entropy $S_{gen} = \left(\frac{A}{4} + S_{out}\right)$ cannot decrease:

$$\frac{dS_{gen}}{dt} = \frac{d}{dt}\left(\frac{A}{4} + S_{out}\right) \geq 0. \tag{4.6}$$

Here, A is usually taken to refer to the area of the event horizon of a black hole. However, in [127]

[4]Note that in the discussion of the proofs below we sometimes use ANEC and AANEC interchangeably.

the proof only assumes that A is the expectation value of the area of the intersection of a *future causal horizon* H_{fut} and a complete time slice T, i.e. $H_{fut} \cap T$, where for any future pointing timelike curve W_{fut}, H_{fut} is defined to be the boundary of its past, i.e. $H_{fut} := \partial \mathcal{I}^{-}(W_{fut})$. S_{out} refers to the entropy outside the horizon region, i.e. $\mathcal{I}^{-}(W_{fut}) \cap T$ [127].

How is this statement connected to the stress-energy tensor? Imagine there was no constrain on the stress tensor that stopped us from sending a flux of negative energy into a black hole. As a result of this, the area of the event horizon A would get smaller, but this change would not need to be connected to any change in the entropy outside the black hole S_{out}. Hence, we would be violating the generalized 2^{nd} law. This suggests that, if this law is valid, then there must be some sort of restriction on what the stress tensor can be [127].

The proof makes use of CPT symmetry, which in this case is equivalent to time reversal symmetry. This allows us to establish the anti-generalised 2^{nd} law, defined in the same manner as the original law but using past causal horizons H_{past} instead.

The next step is to use the *null splitting theorem* proved by Galloway [55], which entails that a null line lives in a constant achronal horizon H which is simultaneously a past and a future horizon. The horizon can be used to divide the rest of the spacetime in three regions: the chronological future $F := \mathcal{I}^{+}(H)$, the chronological past $P := \mathcal{I}^{-}(H)$ and a region O with spacelike separation with respect to the horizon. The first two are open regions while the last is closed.

Now, the stress-energy tensor due to quantum fields can be inserted, generating a perturbation in the spacetime. This will be translated into the creation of past and future perturbed horizons H_{fut} and H_{past} which, given they satisfy the generalized and anti-generalized law respectively, will imply the ANEC on null lines for all quantum fields. This can be seen by performing a perturbation expansion of the metric and the stress tensor in the parameter $\hbar$[5] and using EFEs in each order. The zeroth order equation simply describes classical fields obeying the NEC, which is saturated on H and where $H_{fut} = H_{past}$. There is a half order $\hbar^{\frac{1}{2}}$ contribution in the

[5]Hence, we will recover the parameter $\hbar$ for this section only and write it explicitely.

metric representing the graviton's effects in the metric fluctuations which is unaccompanied by a reciprocal half order term in the stress tensor. The first order term is the interesting one, as it is the one that will give us the AANEC. This equation reads

$$\langle G^{(1)}_{ab}(g^{(0)}_{ab}, g^{(\frac{1}{2})}_{ab}, g^{(1)}_{ab})\rangle = 8\pi \, \langle T^{(1)}_{ab}\rangle \,, \tag{4.7}$$

where the superscripts tell us the order of $\hbar$ in each term.

To prove the AANEC Wall introduces two complete time slices T_1 and T_2 of the spacetime where T_2 lies to the future of T_1 everywhere except where they intersect. The Penrose diagram for this spacetime configuration is shown in figure 4.2.[6]

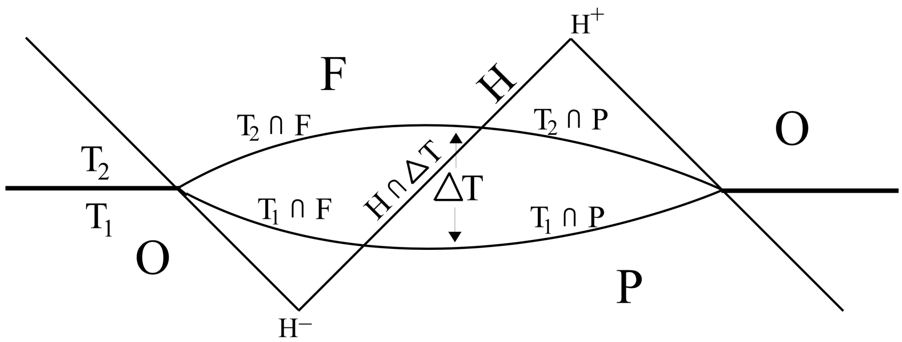


Figure 4.2: Penrose diagram for the unperturbed classical spacetime. ΔT quantifies the region between the two slices. (taken from [127])

For an arbitrary time slice T the lower order dynamic contribution of the generalized entropy will be the zeroth order $\hbar^0$ and this will depend on the zeroth order contribution of the entropy outside the horizon and the first order term for the expectation value of the area. Then, from the diagram we can obtain the following inferences for the two laws [127]

$$\textit{generalized } 2^{nd} \textit{ Law} \qquad \Delta\left[S^{(0)}_{out}(P\cap T) + \frac{A^{(1)}}{4\hbar}(H_{fut})\right] \geq 0, \tag{4.8}$$

$$\textit{anti-generalized } 2^{nd} \textit{ Law} \qquad \Delta\left[S^{(0)}_{out}(F\cap T) + \frac{A^{(1)}}{4\hbar}(H_{past})\right] \leq 0. \tag{4.9}$$

[6]Note that the spacetime boundary is not included in the diagram.

By subtracting the two we can combine them into a single inequality

$$\Delta\left[S^{(0)}_{out}(P\cap T) - S^{(0)}_{out}(F\cap T)\right] + \Delta\left[\frac{A^{(1)}}{4\hbar}(H_{fut}) - \frac{A^{(1)}}{4\hbar}(H_{past})\right] \geq 0. \tag{4.10}$$

Fundamental properties of QFT such as *unitarity* and *weak monotonicity* tell us that the outside entropies have to be such that the first term in 4.10 is negative (see pages 11-12 of [127] for details), which implies that the second term with the areas must be positive. In other words, the area of the future horizon always increases more rapidly than the area of the past horizon.

To finalise the proof, we have to invoke the Raychaudhuri equation and add a few extra assumptions. The Raychaudhuri equation for null geodesics is given by equation 2.22 which we restate here

$$\frac{d\theta}{d\lambda} = -R_{ab}k^a k^b + 2\omega^2 - 2\sigma^2 - \frac{1}{2}\theta^2. \tag{4.11}$$

Recall that for null vectors $8\pi T_{ab}k^a k^b = G_{ab}k^a k^b = R_{ab}k^a k^b$. In our present case, vorticity and shear squared are assumed to vanish, i.e. $\omega^2 = \sigma^2 = 0$. Because our horizon is constant, the value of the expansion parameter θ on H is zero. Finally, we are only interested in the first order contribution in $\hbar$ so the Raychaudhuri equation becomes

$$\frac{d\,\langle\theta^{(1)}\rangle}{d\lambda} = -8\pi\,\langle T^{(1)}_{ab}\rangle\,k^a k^b. \tag{4.12}$$

The conclusion we obtained from the (anti-)generalised 2^{nd} law was that the expectation value of the area of the future horizon will always be increasing faster than the one of the past horizon. This idea can be captured by the expansion parameter in the form of the inequality

$$\langle\theta^{(1)}_{fut} - \theta^{(1)}_{past}\rangle \geq 0. \tag{4.13}$$

In the limit where the perturbation is very small, the perturbed horizons should accord with the unperturbed classical horizon ($\hbar^0$ order) for which we found that the expansion vanishes. Hence, we have that $\langle\theta^{(1)}_{fut}\rangle\,|_{\lambda=\infty} = 0$ and $\langle\theta^{(1)}_{past}\rangle\,|_{\lambda=-\infty} = 0$. Finally, using these two facts one can solve 4.13 to find that

$$\int_{-\infty}^{\infty} d\lambda \, \langle T_{ab}^{(1)} \rangle \, k^a k^b \geq 0 \tag{4.14}$$

which is the ANEC on null lines with tangent vector k^a. □

It must be mentioned that there is another core assumption without which the proof would be invalid. It has been assumed that there exists a nice renormalization scheme for the quantity S_{out}, otherly known as von Neumann entropy, which normally diverges in QFT. In the paper, Wall does not complete this renormalization but rather shows that one needs only to assume a number of reasonable properties for the renormalized entropy to ensure that the proof still holds [127].[7]

4.3.2 AANEC from causality in QFT

The second proof of the ANEC we will consider comes from a 2016 paper by Hartman, Kundu and Tajdini [65]. This is a proof from QFT in $\mathbb{M}^4$ that works by imposing causality constraints on *4-point correlating functions*. Energy conditions were originally introduced from a gravity perspective, but the authors in this paper show that they also are non-trivial and interesting statements in QFT, or at least the ANEC is. The proof uses *null coordinates* $(u, v, \vec{x}_\perp)$ which are defined by

$$u = t - y \qquad\qquad v = t + y \tag{4.15}$$

where the usual coordinates are $(t, y, \vec{x}_\perp)$ and $\vec{x}_\perp$ is 2-dimensional. The Minkowski metric then becomes

$$ds^2 = -du dv + d\vec{x}_\perp^2. \tag{4.16}$$

Let us start by stating the ANEC as it will be proven. We define the null energy operator $\mathcal{E}$ as [65]

$$\mathcal{E} := \int_{\infty}^{\infty} du \; T_{uu}(u, v = 0, \vec{x}_\perp = 0), \tag{4.17}$$

[7]For attempts at renormalizing the generalized entropy see [26, 27, 53, 113, 134].

where we have used the shorthand notation $T_{uu} := T_{ab}k^a k^b$ with null vetor k^a. The NEC is the statement that $T_{uu} \geq 0$ and the ANEC is the statement that

$$\langle\psi| \mathcal{E} |\psi\rangle \geq 0 \qquad \forall |\psi\rangle \, . \tag{4.18}$$

In other words, $\mathcal{E}$ is a positive operator. Although the same can be done for higher point functions, 4-point functions are useful because they can be thought of as causality tests. Using the null coordinates we can draw a spacetime diagram as shown in figure 4.3. B and C can be

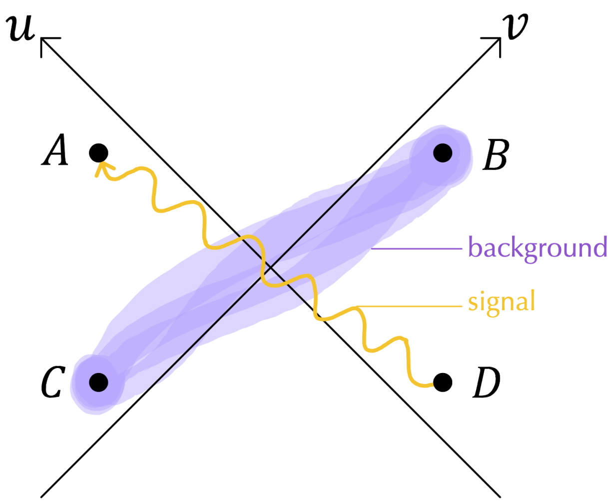


Figure 4.3: Spacetime diagram of a 4-point correlating function causality test with two background operators (B,C) and two probe operators (A,D).

thought of as creating a background that a signal from D to A has to travel through bending out of the lightcone.

Operator Product Expansion in the Lightcone limit

The *operator product expansion* (OPE) captures the idea that if we have two operators in two spacetime points x_1 and x_2 we should in principle be able to express their product as a linear combination of local operators which we can arbitrarily choose to place at x_1

$$\psi(x_1)\psi(x_2) \simeq \sum_\theta C^\theta_{\psi\psi}(x_{12})\mathcal{O}(x_1). \tag{4.19}$$

The approximation works best in the limit where x_1 and x_2 are close to each other. From dimensional analysis, we know that the LHS has dimension $2\Delta_\psi$ where Δ_ψ is the dimension of

ψ. Then, we can build the RHS in a way that respects dimensional analysis. We get

$$\text{RHS} = \sum_{\theta} \mathcal{O}_{a_1 \ldots a_l}(x_1 - x_2)^{a_1} \ldots (x_1 - x_2)^{a_l} |x_1 - x_2|^{-2\Delta_\psi + \Delta_\mathcal{O} - l_\mathcal{O}} C^\theta_{\psi\psi}(\mathcal{M}(x_1 - x_2)). \tag{4.20}$$

The first term in the powers of the absolute value of the distance, i.e. $-2\Delta_\psi$, is the one that gives us the correct dimension and the other two powers cancel out the dimensions given by the operator $\mathcal{O}$ and the distances that are contracted with it.

The proof works in a *conformal field theory* (CFT). A nice consequence of CFTs is that the term $C^\theta_{\psi\psi}(\mathcal{M}(x_1 - x_2))$, which generally is a function, becomes just a constant. Another nice aspect about CFTs in Lorentzian signature is that the OPE can be used as a good asymptotic series to approximate the operators, despite not converging.

It is illustrative to first look at the *Euclidean OPE limit*. This is the limit where $u, v \to 0$ simultaneously, represented by the green arrows in figure 4.4. In this limit, all the position dependent terms ($(x_1 - x_2)^{a_r}$, $|x_1 - x_2|^{-2\Delta_\psi + \Delta_\mathcal{O} - l_\mathcal{O}}$) scale to zero at the same rate, which renders it easy to see the dominant terms in the OPE. The factors $|x_1 - x_2|^{-l_\mathcal{O}}$ cancel with the terms $(x_1 - x_2)^{a_r}$ and we are left with the dominant terms $|x_1 - x_2|^{-2\Delta_\psi + \Delta_\mathcal{O}}$. Moreover, we see that the terms with lowest $\Delta_\mathcal{O}$ dominate, that is, the lowest dimension operators dominate. The *lightcone limit* (or null limit) is slightly different. Here we take the limit $v \to 0$ while keeping u fixed. This is represented by the blue arrows in the same figure 4.4. Note that in lightcone coordinates we can make the substitution $(x_1 - x_2)^2 = -uv$ so the position dependent term becomes $(-uv)^{-\Delta_\psi + \frac{\Delta_\mathcal{O} - l_\mathcal{O}}{2}}$.

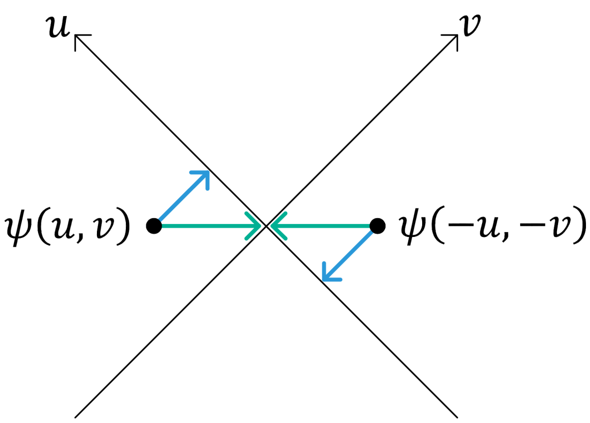


Figure 4.4: Euclidean (green) and Lightcone (blue) OPE limit.

Now we have to ask ourselves which operator components are important in each limit. In the Euclidean limit all the operator components come with the same order, but in the lightcone limit this is no longer true. For the sake of simplicity we can assume that $\vec{x}_\perp = 0$ without loss of generality. The operator indices are contracted with the position dependent terms in the following way

$$\mathcal{O}_{\underbrace{u\ldots u}_{n}\underbrace{v\ldots v}_{\bar{n}}}\underbrace{(x_1-x_2)^u\ldots(x_1-x_2)^u}_{n}\underbrace{(x_1-x_2)^v\ldots(x_1-x_2)^v}_{\bar{n}}. \tag{4.21}$$

However, in the null limit we take $v \to 0$ so all the operators with v components will be suppressed. Another difference with respect to the Euclidean limit is that here the expansion is no longer controlled by the dimension of the operator $\Delta_{\mathcal{O}}$ but rather by the combination of the scaling dimension minus the spin $\tau := \Delta_{\mathcal{O}} - l_{\mathcal{O}}$ which receives the name of *twist* [65]. As a result, we can have high dimension dominant operators in the null limit as long as they have low twist, i.e. as long as they also have high spin.

Hence, we have to find the low twist operators in a CFT. The minimum twist will be given by *unitary bounds* which will prevent us from having negative normed states. In d spacetime dimensions we have the following bounds

$$\text{scalars} \qquad \tau \geq \frac{d-2}{2} \qquad \tau \geq 1 \text{ in } d=4 \tag{4.22}$$

$$\text{tensors} \qquad \tau \geq d-2 \qquad \tau \geq 2 \text{ in } d=4, \tag{4.23}$$

where the first inequality is saturated for free scalars and the second one for conserved currents. Examples of low twist operators are the stress-energy tensor, which exists in any CFT and is always conserved, and its derivatives

$$T_{ab} \qquad l=2 \qquad \Delta_T = d = 4 \qquad \tau = 2 \tag{4.24}$$

$$\partial_c T_{ab} \qquad l=3 \qquad \Delta_{\partial T} = 5 \qquad \tau = 2 \tag{4.25}$$

$$\partial_c\partial_d T_{ab} \qquad l=4 \qquad \Delta_{\partial\partial T} = 6 \qquad \tau = 2 \tag{4.26}$$

$$\ldots$$

and this process can be repeated indefinitely.

Scalars, which have $l = 0$, may have lower twist than these tensors but they will be ignored as they do not affect any of the calculations. Then, in an interacting theory the dominant terms will be given by the stress tensor, its derivatives and spin-one currents. The latter, however, just like the scalars, will not affect the calculations and thus can be safely ignored. Hence, we can finally conclude that the dominant terms in the lightcone OPE are the ones with low twist and with no v-components: T_{uu}, $\partial_u T_{uu}$, $\partial_u^2 T_{uu}$, etc.

Note that this would not be possible in a free theory. For instance, for the free scalar ϕ, its equation of motion allows us to construct arbitrarily high spin conserved currents such as $\phi\partial_a\partial_b\partial_c\partial_d\phi$ which would appear at the same order as the stress tensor in the lightcone OPE. In interacting theories, however, these conserved currents all fade away as their twist grows.

In addition, another important caveat is that this method is problematic in a CFT with $d = 2$ since the twist vanishes, i.e. $\tau_T = d - 2 = 0$. This results in an infinite number of other tensors with twist zero where we have a product of the stress tensor with any other tensor.

Taking into account all these considerations we can write the normalized OPE

$$\frac{\psi(u,v)\psi(-u,-v)}{\langle\psi(u,v)\psi(-u,-v)\rangle} \simeq \mathbb{1} + C_{\psi\psi T}(-uv)^{\frac{d-2}{2}}u^2\sum_{m=0}^{\infty} b_m(u\partial_u)^m T_{uu}(0) \tag{4.27}$$

$$= \mathbb{1} + C_{\psi\psi T}vu^3\sum_{m=0}^{\infty} b_m(u\partial_u)^m T_{uu}(0) \qquad \text{for } d = 4. \tag{4.28}$$

The identity $\mathbb{1}$ is the trivial leading term with twist $\tau_{\mathbb{1}} = 0$, $C_{\psi\psi T}$ is the OPE coefficient and b_m are the coefficients that have to be determined. We can do so by plugging this expression into a 3-point function [65]. The main use of OPEs is to convert complicated higher-point functions to combinations of lower-point functions. The expressions for the following correlators are fixed by conformal invariance and have been obtained in [95]

$$\langle\psi\psi T\rangle := \langle\psi(u,v)\psi(-u,-v)T_{uu}(u_3,v_3)\rangle = \frac{C_{\psi\psi T}(2u)^2(-2uv)^{1-\Delta_\psi}}{(u^2-u_3^2)^3(\underbrace{\cancel{v^2}}_{v\to 0}-v_3^2)}, \tag{4.29}$$

$$\langle TT\rangle := \langle T_{uu}(u,0)T_{uu}(u_3,v_3)\rangle = \frac{C_T}{(u-u_3)^6 v_3^3}, \tag{4.30}$$

where the coefficient C_T roughly represents the number of degrees of freedom, or fields, in the theory. To determine the coefficients b_m we just have to take the OPE equation (4.28), multiply from the right with the stress tensor and, using the correlators above, take the vacuum expectation value. Doing this one finds that the coefficients b_m are completely fixed by conformal invariance and are given by

$$b_m = \frac{240}{C_T(m+3)(m+5)(m+1)!}. \tag{4.31}$$

Also, Ward's identity helps us to identify $C_{\psi\psi T} = -\frac{2\Delta_\psi}{3\pi^2}$ which is a sign indicating a relation between the stress tensor and the generator of conformal transformations. It is useful to rewrite the OPE 4.28 in an integral form

$$\frac{\psi(u,v)\psi(-u,-v)}{\langle\psi(u,v)\psi(-u,-v)\rangle} \overset{v\to 0}{\simeq} \mathbb{1} + \frac{10\Delta_\psi}{\pi^2 C_T} v u^2 \int_{-u}^{u} d\tilde{u} \left(1-\frac{\tilde{u}^2}{u^2}\right)^2 T_{uu}(\tilde{u}, v=0, \vec{x}_\perp = 0). \tag{4.32}$$

That this edit is allowed can be checked by taking the expression above, Taylor expanding T_{uu} in derivatives around the origin, performing the trivial integrals order by order and then recovering the initial expression for the OPE.

So far, we have taken the limit $v \to 0$. Now we incorporate the second limit $u \to \infty$ that is carried out only after the first limit (see figure 4.5). This can also be expressed with the statement $|v| << |u^{-1}| << 1$.

Hence, the final result of taking this double limit is

$$\frac{\psi(u,v)\psi(-u,-v)}{\langle\psi(u,v)\psi(-u,-v)\rangle} \overset{v\to 0, u\to\infty}{\simeq} \mathbb{1} + \frac{10\Delta_\psi}{\pi^2 C_T} v u^2 \int_{-u}^{u} d\tilde{u} T_{uu}(\tilde{u}, v=0, \vec{x}_\perp = 0), \tag{4.33}$$

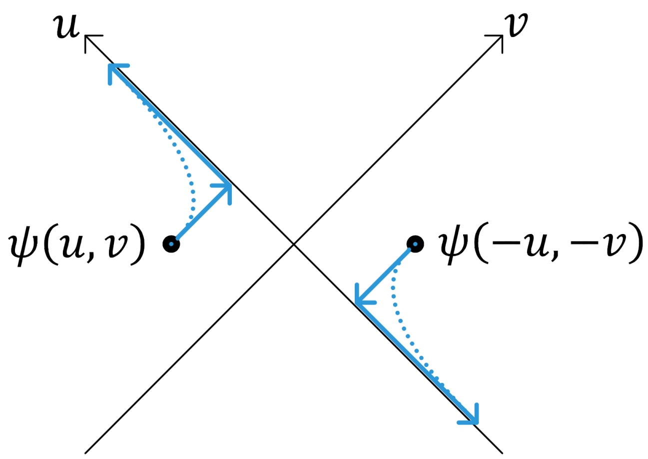


Figure 4.5: Lightcone OPE double limit. First $v \to 0$, then $u \to \infty$.

which we can express more conveniently as

$$\frac{\psi(u,v)\psi(-u,-v)}{\langle\psi(u,v)\psi(-u,-v)\rangle} \overset{v\to 0, u\to\infty}{\simeq} \mathbb{1} + \lambda_T v u^2 \mathcal{E} \tag{4.34}$$

where we have defined $\lambda_T = \frac{10\Delta_\psi}{\pi^2 C_T}$ and where $\mathcal{E}$ is the null energy operator. Hence, we recover the null energy operator as the dominant contribution of the OPE in the lightcone limit.

Causality in QFT

In quantum field theories, microcausality is the statement that operators commute at spacelike separation:

$$[\mathcal{O}_1(x_1), \mathcal{O}_2(x_2)] = 0 \qquad \forall\ (x_1 - x_2) > 0. \tag{4.35}$$

Roughly, what this commutator is expressing is the possibility of sending a signal from x_1 to x_2. Since the vanishing commutator is an operator expression, it can be inserted inside a correlator.

Note that $[\mathcal{O}_1(x_1), \mathcal{O}_2(x_2)] = 0$ is a statement made in Lorentzian signature. However, QFTs can also be defined in Euclidean signature, where there are no such things as commutators of local operators or, more precisely, all commutators vanish. Another difference between the two signatures is that, in Lorentzian, the order of the operators in a correlator matters, and conventionally we choose them to be normal ordered. In Euclidean, however, correlators are single-valued, which means that all the orderings produce the same correlator. We can convert from one signature to the other with the relation $\tau = it$ where $\tau \in \mathbb{C}$ (purely imaginary) is the time coordinate in Euclidean and $t \in \mathbb{R}$ is the time component in Lorentzian. This conversion requires *analytic continuation* of the correlator function, which was originally defined for real τ. Now we want to analytically continue τ off the real axis. Sometimes there is an ambiguity in the analytic continuation which means we can continue the function in different ways. This is a *branch cut* and it is associated with Lorentzian commutators and the fact that the operator ordering matters when we translate to Lorentzian.

In order to see all this explicitely, we can first look at the easier example of the 2-point function given by

$$G_E(\tau, x) = \langle \mathcal{O}_1(\tau, x)\mathcal{O}_2(0)\rangle = (\tau^2 + x^2)^{-\Delta} \tag{4.36}$$

where the subscript E refers to the fact that it is defined in Euclidean. This expression is single-valued $\forall\ \tau$. Going to Lorentzian we get the correlator

$$\langle \mathcal{O}_1(t, x)\mathcal{O}_2(0)\rangle = (-t^2 + x^2)^{-\Delta} \tag{4.37}$$

which is no longer single-valued $\forall\ t$. In fact, when $t = x$ (or $\tau = ix$) we encounter a branch cut after which the function cannot be arbitrarily analytically continued. Basically, we can either choose to get past the branch point from the right or from the left as shown in figure 4.6, each corresponding to a different ordering in Lorentzian signature and hence to different correlators distinguished by a phase.

$$\langle \mathcal{O}_1(t, x)\mathcal{O}_2(0)\rangle_R = e^{-i\pi\Delta}(t^2 + x^2)^{-\Delta} \tag{4.38}$$

$$\langle \mathcal{O}_2(0)\mathcal{O}_1(t, x)\rangle_L = e^{+i\pi\Delta}(t^2 + x^2)^{-\Delta}. \tag{4.39}$$

Hence, in Lorentzian signature, the branch point is at the exact point where the commutator becomes non-vanishing, which lies on the lightcone whose centre is at the origin where $\mathcal{O}_1(0)$ is situated. A useful strategy to differentiate the two paths or orderings is to implement the $i\epsilon$ *prescription*, which allows us to analytically continue along the imaginary axis by infinitesimally

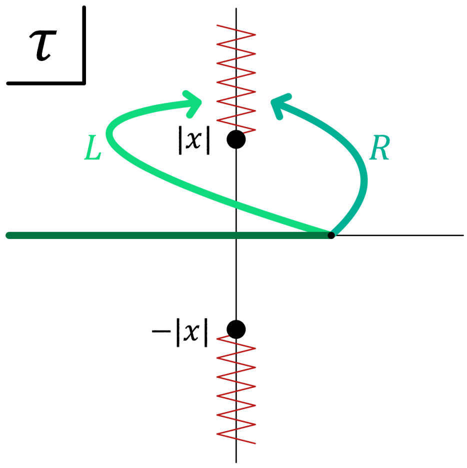


Figure 4.6: Branch cut in analytic continuation of τ.

translating the branch points. In this prescription we can write the correlators as

$$\langle \mathcal{O}_1(t,x)\mathcal{O}_2(0)\rangle_R = G_E(i(t-i\epsilon),x) \tag{4.40}$$

$$\langle \mathcal{O}_2(0)\mathcal{O}_1(t,x)\rangle_L = G_E(i(t+i\epsilon),x). \tag{4.41}$$

Now let us look at the more relevant 4-point function. We can place the operators in the arrangement shown in figure 4.7. The operator at x_2 will be our probe operator that we will

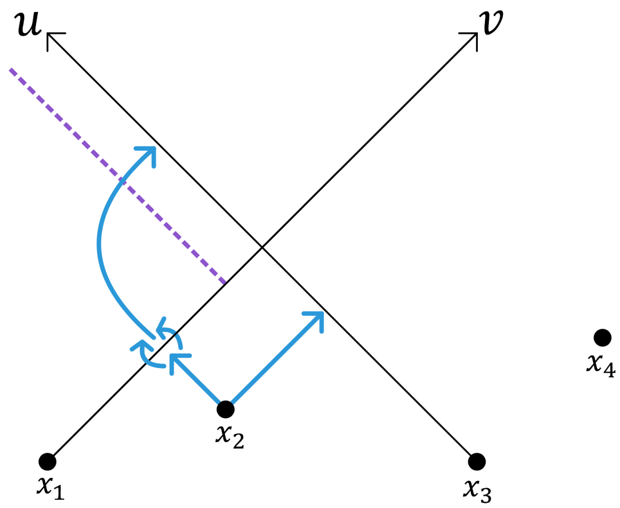


Figure 4.7: Operators of a 4-point function. x_2 is evolved across the two lightcones created by x_1 and x_3. Causality entails the second branch cut will not occur before the second lightcone.

move across the lightcones defined by the fixed operators at points x_1 and x_3. Time evolving the operator at x_2 we first encounter a branch cut when crossing the lightcone defined by x_1. This is the same situation we had with the 2-point function and we can get past the lightcone by choosing one of the two possible orderings for points x_1 and x_2. We keep evolving the operator until we encounter the lightcone defined by point x_3. However, now it is no longer certain that the branch cut will occur on the lightcone. We need extra information to tell us where we will find the cut. This extra non-trivial information is provided by causality, which essentially states that it cannot be found before the lightcone since, if it were, then causality would be violated. In other words, the background created by the operator at x_3 would find its way outside the lightcone. In the literature, this second branch cut is often referred to as the *second sheet*. In the correlator language this is translated to the statement that correlators must be analytic on the second sheet as we approach the second lightcone. Hence, we find that there is a deep connection between causality and the analiticity of correlators.

Rindler positivity

In Euclidean signature, to ensure that when we translate to Lorentzian signature the theory is causal we need to impose *reflection positivity*. In the Lorentzian signature this positivity of correlators is manifested in a slightly different way called *Rindler positivity* [20, 87]. Here we present a sketch of what these statements imply and how they are used in the proof of the ANEC.

In the Euclidean plane defined by τ and y, we can choose different lines under which we can impose reflection positivity. For instance, if we have 3 operators with timelike separation and we reflect them with a horizontal line with τ constant then the positivity condition tells us that the inner product of the state with its reflection, which is a norm, must be positive. This implies that the Lagrangian is real. If instead we have 2 spacelike separated operators and we reflect them with a line with constant y this can no longer be thought of as a norm but rather as a correlator. Then, what reflection positivity is telling us is that in Lorentzian the correlator has to be positive.

To find what the reflection looks like in Lorentzian we take a slightly tortuous yet logical journey. First we evolve one of these points in Lorentzian, say in the right Rindler wedge (figure 4.8). Then translate to Euclidean and reflect the evolved point using reflection positivity. Finally, we take the reflected evolved point in Euclidean and we translate back to Lorentzian. The final point obtained in the left Rindler wedge, that is, the Rindler reflected point is given by

$$\overline{x} = \overline{(t, y, \vec{x}_{\perp})} = (-t^*, -y^*, \vec{x}_{\perp}), \tag{4.42}$$

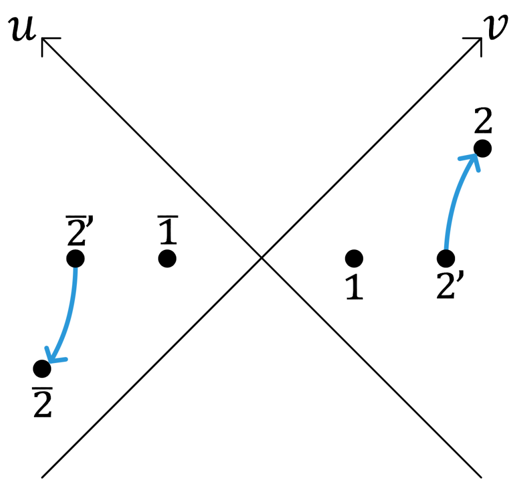


Figure 4.8: Rindler reflection in Lorentzian.

Hence, for a 4-point function, Rindler positivity implies that

$$G_+ := \langle \mathcal{O}_1 \mathcal{O}_2 \overline{\mathcal{O}_1 \mathcal{O}_2} \rangle > 0, \tag{4.43}$$

where the spacetime points and their reflections are represented as in figure 4.8.

From the diagram we also see that operators at points 1 and 2 (and $\overline{1}$ and $\overline{2}$ respectively) are, or at least can be, timelike separated. Hence, different orderings between these operators produces the same correlator. This is why we write $\overline{\mathcal{O}_1 \mathcal{O}_2}$ with just one bar, as $\overline{\mathcal{O}_1 \mathcal{O}_2} = \overline{\mathcal{O}_2 \mathcal{O}_1}$ [65]. Different orderings between the reflected and non-reflected points, however, would indeed give different results which need not be positive. Still, these are bounded as follows

$$|G_{disordered}| \leq G_+, \tag{4.44}$$

where an example of $G_{disordered}$ could be the correlator $\langle \mathcal{O}_1 \overline{\mathcal{O}_1} \mathcal{O}_2 \overline{\mathcal{O}_2} \rangle$.

Final argument

Now let us finally put all these pieces together and reproduce the proof of the ANEC by causality proposed by Hartman, Kundu and Tajdini in [65]. Let us start with a simplified example to show the main intuition behind the proof.

Our toy theory has a branch cut starting at the null line where $v = 0$. This encodes the causality connotation that signals can be sent along the null line. In the correlator expression this is represented by a prefactor of $\frac{1}{v^\Delta}$. We can easily shift this branch point with the parameter a so that the prefactor reads $\frac{1}{(v-a)^\Delta}$. This shift will represent either a time delay or a time advance depending on the sign of a. To avoid signals being bent outside of the lightcone, the causal structure then demands $a > 0$. For small parameters, we may expand in series as

$$\frac{1}{(v-a)^\Delta} \simeq \frac{1}{v^\Delta}\left(1 + a\frac{\Delta}{v} + \mathcal{O}(a^2)\right), \tag{4.45}$$

where we have written explicitely the first order correction and ignored the higher order ones. Note that this correction is small in a but also growing as one approaches the lightcone where $v \to 0$. Such small corrections but growing corrections, as we have seen, come with sign constraints like $a > 0$ [63, 64, 87]. This sign constraint is where the ANEC will appear.

Having this intuition, we can now consider the setting that will actually give us the ANEC which is represented in figure 4.9.

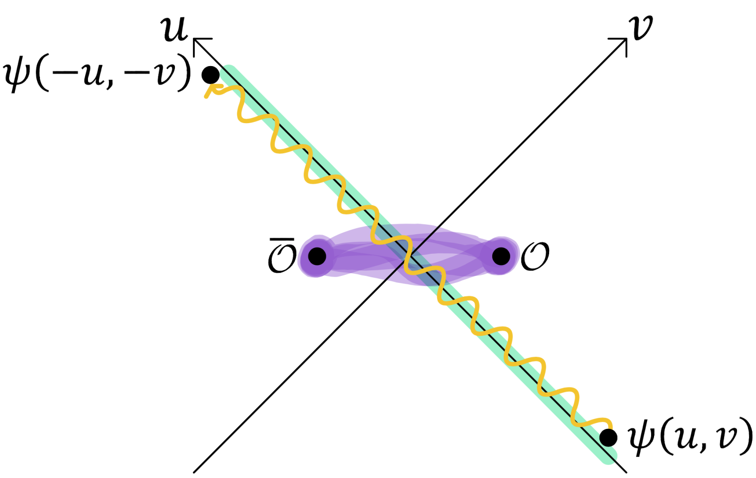


Figure 4.9: ANEC setting. Operators $\mathcal{O}$'s create a background that a signal from the probe operators has to go through. The ANE $\mathcal{E}$ is integrated along the null green line.

Here, $\mathcal{O}$ are general operators creating some background state and ψ are the probe operators in the sense that we want to see if a signal from one *psi* can reach the other ψ. The ANEC will be proven by considering the normalized correlator [65]

$$G_{\times}(u,v) = \frac{\langle\overline{\mathcal{O}(\delta)}\psi(u,v)\psi(-u,-v)\mathcal{O}(\delta)\rangle}{\langle\overline{\mathcal{O}(\delta)}\mathcal{O}(\delta)\rangle\,\langle\psi(u,v)\psi(-u,-v)\rangle}. \tag{4.46}$$

Using the OPE equation (4.34) we get

$$G_{\times} \simeq 1 + \frac{\lambda_T}{N_\delta}vu^2\,\langle\overline{\mathcal{O}(\delta)}\mathcal{E}\mathcal{O}(\delta)\rangle \tag{4.47}$$

$$= 1 + \frac{\lambda_T}{N_\delta}vu^2\left\langle\overline{\mathcal{O}(\delta)}\int_{-\infty-i\epsilon}^{\infty+i\epsilon} duT_{uu}(u)\mathcal{O}(\delta)\right\rangle \tag{4.48}$$

where $N_\delta := \langle\mathcal{O}(\delta)\overline{\mathcal{O}(\delta)}\rangle > 0$ from Rindler positivity. If we switch the order of some of the operators in the correlator the $i\epsilon$ prescription gives a different result. For instance, the Rindler positive correlator

$$G_{+} := \langle\overline{\mathcal{O}\psi}\mathcal{O}\psi\rangle \simeq 1 + \frac{\lambda_T}{N_\delta}vu^2\left\langle\overline{\mathcal{O}(\delta)}\cancelto{0}{\int_{-\infty-i\epsilon}^{\infty-i\epsilon} duT_{uu}(u)\mathcal{O}(\delta)}\right\rangle, \tag{4.49}$$

where we have used the definitions $\psi := \psi(-u,-v)$ and $\overline{\psi} := \psi(u,v)$ and the small but growing contribution to the correlator vanishes.[8] Applying equation 4.44 we conclude that $\langle\overline{\mathcal{O}\psi}\psi\mathcal{O}\rangle$ (our $G_\times$) is bounded by $\langle\overline{\mathcal{O}\psi}\mathcal{O}\psi\rangle$.

The proof uses the relabeled coordinates $v = -\eta\sigma$ and $u = \frac{1}{\sigma}$ where $0 < \eta \ll 1$. The causality considerations will take the form of $G(\sigma)$ being analytic in the upper half of the complex σ plane. By *Residues theorem* we know that analytic functions integrated over closed contours which enclose no poles have to equal zero.

$$\oint d\sigma\; G = 0, \tag{4.50}$$

which, for the contour in figure 4.10, we can separate the contribution from the arc and the

[8] Roughly, it vanishes because the contour can be closed without encountering any branch cuts.

one horizontal line as

$$\text{Re}\left(\oint d\sigma\ (G-1)\right) = 0. \tag{4.51}$$

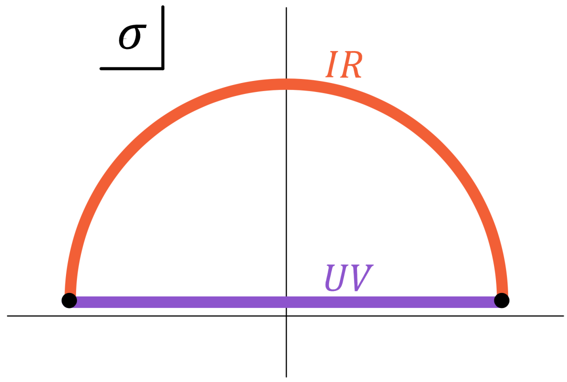


Figure 4.10: Sigma contour. The radius of the disk is (mention radius of the disk is $\eta \ll 1$.

The horizontal line and the arc have been drawn in different colours because they capture quite different physics. On one hand, the red integral along the arc where $|v| \ll |u^{-1}| \ll 1$ so we can use the lightcone OPE which is dominated by low twist operators. This regime of the correlator probes the *infrared* (IR) region of the theory, which is known. On the other hand, in the horizontal purple line we have $|v|, |u^{-1}| \ll 1$ but without demanding any fraction between the two, which implies that the lightcone OPE can no longer be used here. This section of the contour probes the *ultraviolet* (UV) region of the theory, which is unknown. Equation 4.10 then relates known parts of the theory where we can perform calculations to unknown parts of the theory where we cannot. The IR part of the contour then can be calculated using the lightcone OPE, and the UV contour can be bounded by unitarity as [65]

$$\text{Re}\ G_\times \leq |G_\times| \leq |G_+| = \left(\langle \overline{\mathcal{O}\psi}\mathcal{O}\psi\rangle \langle \overline{\psi\mathcal{O}}\psi\mathcal{O}\rangle\right)^{\frac{1}{2}} \sim \langle \overline{\mathcal{O}}\mathcal{O}\rangle \langle \overline{\psi}\psi\rangle + \mathcal{E}, \tag{4.52}$$

where the relations $\langle \overline{\mathcal{O}\psi}\mathcal{O}\psi\rangle \sim \langle \overline{\psi\mathcal{O}}\psi\mathcal{O}\rangle \sim \langle \overline{\mathcal{O}}\mathcal{O}\rangle \langle \overline{\psi}\psi\rangle$ have been used [65]. This entails that for $\sigma \in \mathbb{R}$, $|\sigma| < 1$ we have

$$\text{Re}\ G \leq 1 + \mathcal{E}. \tag{4.53}$$

Hence, putting all the pieces together we find that equation 4.44 can be rewritten as [65]

$$i\,\langle \overline{\mathcal{O}}\mathcal{E}\mathcal{O}\rangle = \frac{N_\delta}{\pi\eta\lambda_T}\int_{-R}^{R} d\sigma\ \text{Re}\ (1 - G(\sigma)), \tag{4.54}$$

where from equation 4.53 we know that the RHS is positive, i.e. $i\left\langle\overline{\mathcal{O}(y=\delta)}\mathcal{E}\mathcal{O}(y=\delta)\right\rangle \geq 0$. Rotating with R in the Euclidean τ-y plane by $\frac{\pi}{2}$ and defining $\mathcal{O}_R := R\cdot\mathcal{O}$ we get

$$\langle\mathcal{O}_R^\dagger(t=i\delta)\mathcal{E}\mathcal{O}_R(t=-i\delta)\rangle \geq 0, \tag{4.55}$$

where $\mathcal{O}_R$ is still an arbitrary operator. In a CFT, this is sufficient to entail that the expectation value of the averaged null energy $\mathcal{E}$ in any state is greater than zero, i.e. $\langle\mathcal{E}\rangle_\psi \geq 0\ \forall$ states $|\psi\rangle$, hence proving the ANEC. □

Summing up, we have found that the null energy operator controls the lightcone OPE. Then, some causality constraints, which are equivalent to analiticity have allowed us to relate the unknown physics in the UV with the known physics in the IR; and this relation has been used to prove the ANEC.

4.3.3 AANEC from holography

The proof we consider in this section was proposed by Kelly and Wall in 2014 [75] and makes use of the AdS-CFT correspondence (*anti-de Sitter/conformal field theory*). The generality of this proof is challenged by the fact that not all CFTs have holographic duals. However, this proof also uses the lightcone limit presented in the last proof (section 4.3.2), where all CFTs are found to be holographic.

The AdS/CFT correspondence and the holographic dictionary

The AdS spacetime can be represented as a cylinder, with the time dimension spreading to both infinities while the spatial dimensions are bounded. The spacetime geometry in the *bulk*, i.e. inside the cylinder, is dynamical, but is fixed on the *boundary*. Fascinatingly, the boundary provides a fundamental description of the whole spacetime. That is, the information of a d-dimensional AdS spacetime theory is contained in its $(d-1)$-dimensional boundary, as if it were a *hologram*. In a way, the gravity theory living in the bulk can be thought to emerge from the information on the boundary. Yet, because the spacetime on the boundary is fixed, this fundamental description will most certainly not include any dynamical gravity. This idea was remarkably specified by Maldacena in a groundbreaking paper in 1997 [88], where he described the boundary of a particular gravity theory in the bulk as a $\mathcal{N} = 4$ super yang-mills theory. This influential paper initiated the study of the AdS/CFT correspondence, which is the conjecture that the boundary of an AdS spacetime is described by a CFT, and has not yet been mathematically proven. However, because of the extensive list of results and connections that the conjecture has provided, it would be very surprising if it turned out to be not true.

One of the useful consequences of the correspondence is that it provides a *holographic dictionary* that allows one to translate quantities and calculations from one side of the duality to the other. The proof we present here makes use of the AdS/CFT dictionary [61] to translate the stress tensor in the CFT to a component of the spacetime metric in the bulk.[9]

Sketch of the proof

The proof in [75] has a couple of similarities with the previous proof (section 4.3.2). Both proofs rely on the lightcone OPE. In this case, holography will be shown to be the counterpart to the lightcone OPE. In addition, the proof from holography also uses causality arguments. However, instead of causality in QFT, this proof will make use of ideas of causality in quantum gravity. But what is causality in quantum gravity anyway? This is still an unanswered question since, as of today, we have no local criteria stating what constitutes a violation of causality in quantum gravity.[10] Nonetheless, using the AdS-CFT correspondence, we can at least claim that the causal structure should be respected on the boundary of AdS.

Imagine that a signal is sent from a point in the boundary into the bulk and landing on a different point in the boundary. This signal, whose endpoints both lie on the boundary, should better not violate causality in the CFT living on the boundary. This statement is captured by the requirement that the final endpoint of the signal must lie on the future of the original source point of the signal. This can be seen in figure 4.11, where the constraint is given by the

[9] An example of this is the Brown-York Stress Tensor.

[10] One specific challenge for any account of causality in quantum gravity is the *Black Hole information paradox*, which seems to suggest that the full evaporation of a black hole involves some sort of violation of causality.

positivity of the *Shapiro*-like delay $\Delta t \geq 0$.

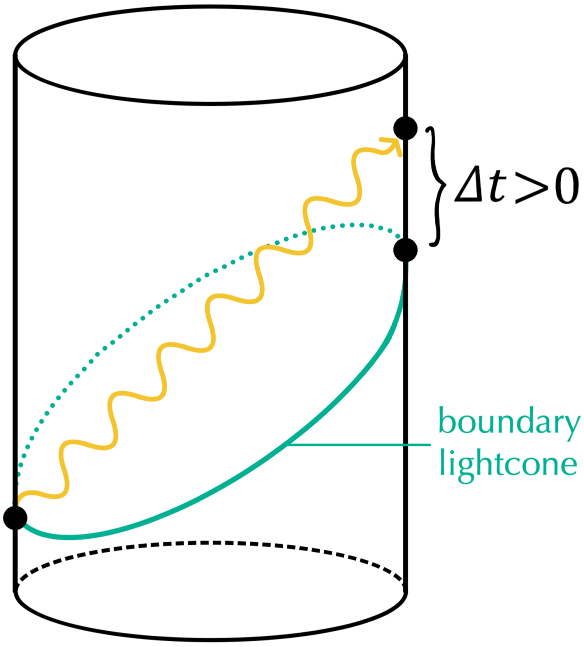


Figure 4.11: Cylinder representing the boundary of AdS. A signal is sent from a point on the boundary to another point on the boundary. Causality in the CFT entails $\Delta t \geq 0$.

The trajectory that will be considered, which is not a geodesic,[11] has fixed $z = z_0$, $\vec{x}_\perp = 0$ and it travels near the boundary. The AdS_{d+1} line interval in lightcone coordinates is given by

$$ds^2 = \underbrace{\frac{1}{z^2}(-dudv + dz^2 + d\vec{x}_\perp^2)}_{\text{empty AdS}} + \underbrace{h_{ab}dx^a dx^b}_{\text{matter content}}, \tag{4.56}$$

where the AdS asymptotics are captured in the expression for the metric

$$h_{ab} = z^{d-2} t_{ab}(u, v, \vec{x}_\perp) + \mathcal{O}(z^d). \tag{4.57}$$

Note that this perturbative analysis is not allowed because of the smallness of h_{ab} (a small deviation from AdS) but rather because of the smallness of z (close to the boundary). Hence, the higher order terms $\mathcal{O}(z^d)$ are only small near the boundary where z is small. This means that the metric in the bulk can be something completely different from AdS. In other words, the bulk is filled with stars, black holes and other sources of energy and matter. But as one approaches the boundary, the metric becomes asymptotically AdS. This is why we choose our

[11]This is not a problem at all since causality should apply to all causal curves and not be restricted to geodesics.

trajectory to be close to the boundary. Hence, the curve $v(u)$ will start in the boundary, marginally enter into the bulk and then travel parallel to the boundary in the null direction u until it hits the boundary again (see figure 4.12).[12]

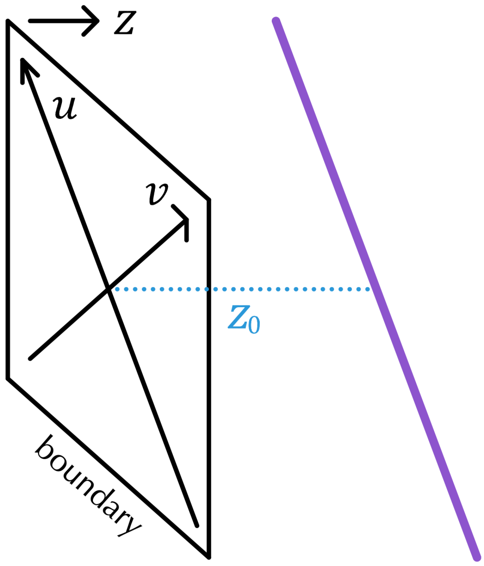


Figure 4.12: Poincaré patch showing trajectory running parallel and close to boundary

In empty AdS, this curve will remain at $v = 0$, but this will not be the case anymore when there is some matter and there will be some shift in v. Nonetheless, causality implies that the shift is positive when we project it on the boundary to ensure that the signal does not propagate faster than light. Along the null trajectory, we have

$$0 = ds^2 = \frac{1}{z_0^2}(-dudv) + h_{uu}du^2 + \mathcal{O}(h^2). \tag{4.58}$$

From this we obtain the differential equation $v'(u) = z_0^2 h_{uu}$ where the prime denotes a u derivative. Another reason why this trajectory is specially interesting is that it minimizes the time duration between the two boundary endpoints and it maximizes the v deflection. Solving the equation we obtain the causality constraint for the total *Shapiro*-like time delay

$$\Delta v = z_0^2 \int_{-\infty}^{\infty} du\ h_{uu}(u, v = 0, \vec{x}_\perp = 0) \geq 0. \tag{4.59}$$

This inequality, when translated to the CFT language, will give us the ANEC. The AdS-CFT

[12] The points where the trajectory hits the boundary are where $u = \pm$, which might be challenging to picture from the figure.

dictionary tells us that h_{ab} (i.e. the *graviton*) in the AdS geometry is dual to a stress tensor T_{ab}^{CFT} evaluated in a CFT state [61]

$$\langle\Phi| T_{ab}^{CFT} |\Phi\rangle = \frac{d}{16\pi} t_{ab}, \tag{4.60}$$

where t_{ab} is given by equation 4.57 (leading contribution to h_{ab} near the boundary). Writing the causal constraint on the *Shapiro*-like delay in CFT language reads

$$\Delta v \sim z_0^2 \int_{-\infty}^{\infty} du \ \langle\Phi| T_{ab}^{CFT}(u, v=0, \vec{x}_\perp = 0) |\Phi\rangle = \mathcal{E} \geq 0. \tag{4.61}$$

and hence the ANEC is proven. □

This argument shows that the ANEC ensures *boundary causality* is satisfied, that is, causally disconnected points in the boundary must also be causally disconnected through the bulk. However, it turns out that there is a weaker condition that preserves this property which was found by Engelhardt and Fischett [30] and named the *averaged light cone tilt condition* or *boundary causality condition* (BCC).

4.3.4 AANEC from quantum information

The next set of proofs we consider were first introduced by Faulkner, Leigh, Parrikar and Wang in 2016 [35] in Minkowski spacetime and very recently extended (May 2020) by Rosso for de Sitter, Anti de Sitter and extremal horizons [106]. The argument is based on the use of entanglement inequalities. It does not require holography and it works equally for free and interacting theories, making it more general than the previous proofs.

Suppose we have two density matrices σ and ρ, the first one being the vacuum $|0\rangle\langle 0|$ and the second one being a general pure state $|\Psi\rangle\langle\Psi|$. We will use the vacuum as our reference state to calculate the relative entropy

$$S(\rho|\sigma) = \mathrm{tr}(\rho \ln \rho) - \mathrm{tr}(\rho \ln \sigma) \tag{4.62}$$

$$= \langle K^{\sigma} \rangle_{\rho} - [S(\rho) - S(\sigma)], \tag{4.63}$$

where in the second line we have rewritten the expression using the expectation value in state ρ of the *modular Hamiltonian* defined as $K^{\sigma} = -\ln \sigma - S(\sigma)$. Here, $S(\rho)$ is the usual *Von Neumann* entropy (or entanglement entropy) $S(\rho) := \mathrm{tr}(\rho \ln \rho)$. Note that if $\sigma = \rho$ then the relative entropy vanishes. Hence, this quantity in a way captures the idea of *distinguishability* in quantum information theory or, in other words, it tells us how different the two states are.

The relative entropy satisfies a number of inequalities, the most basic one being that it is a positive number $S(\rho|\sigma) \geq 0$. The entanglement inequality that will give us the ANEC comes from *monotonicity* or, more specifically, *monotonicity under partial trace* (MPT) applied to the full modular Hamiltonian $\hat{K}_A = K_A - K_{A^c}$, where A is a subregion of A_0 (i.e. $A \subset A_0$) and A^c is the complementary subregion [8, 9]. The intuition behind the inequality is that as the subregion A shrinks, it becomes more difficult to distinguish between the two states.

$$S(\rho_A|\sigma_A) \leq S(\rho_{A_0}|\sigma_{A_0}). \tag{4.64}$$

Written in terms of the modular Hamiltonians, this inequality reads,

$$\langle K^{\sigma}_{A_0} \rangle_{\rho} - \langle K^{\sigma}_{A} \rangle_{\rho} - \cancel{S(\rho_{A_0})} + \cancel{S(\sigma_{A_0})} + \cancel{S(\rho_A)} - \cancel{S(\sigma_A)} \geq 0, \tag{4.65}$$

where the entropy terms cancel because we are considering a pure state. Similarly, for the complementary subregion we have

$$\langle K^{\sigma}_{A_0^c} \rangle_{\rho} - \langle K^{\sigma}_{A^c} \rangle_{\rho} - \cancel{S(\rho_{A_0^c})} + \cancel{S(\sigma_{A_0^c})} + \cancel{S(\rho^c_A)} - \cancel{S(\sigma_{A^c})} \geq 0. \tag{4.66}$$

Adding the two together we can write this in terms of the full modular Hamiltonian

$$\langle \hat{K}^{\sigma}_{A} \rangle_{\rho} - \langle \hat{K}^{\sigma}_{A_0} \rangle_{\rho} \geq 0. \tag{4.67}$$

Since we have not specified the state in ρ at any time, this result remains valid for any state. Hence, we conclude that

$$\hat{K}^{\sigma}_{A} - \hat{K}^{\sigma}_{A_0} \geq 0, \tag{4.68}$$

i.e. it is a positive operator. What remains to be done now is to relate the positivity of this operator to the positivity of the ANEC. It turns out that we can do this for Minkowski [35], dS, AdS and spacetimes near extremal horizon [106] while retaining generality in our QFT. The method is parallel for each of these cases so here we will only present a sketch for extremal horizon spacetimes. The idea is to take the *half-space* of a *Cauchy surface* and study the shape-dependence of entanglement entropy as we deform the half-space.[13] The result is that the first order contribution to the operator $\hat{K}^{\sigma}_{A} - \hat{K}^{\sigma}_{A_0}$ in the deformation is the null energy operator $\mathcal{E}$ multiplied by other positive factors. Then, the positivity of the operator entails the positivity of $\mathcal{E}$, thus proving the ANEC.

Spacetime near extremal black hole horizons

In Einstein-Maxwell theory an extremal black hole is described by the line element:

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega_2^2 \qquad\qquad f(r) = \left(\frac{r - r_h}{r}\right)^2, \tag{4.69}$$

where r_h is the radius of the horizon. In the limit near the horizon, this can be approximated as

$$ds^2 \simeq \frac{-d\sigma^2 + r_h^2 d\theta^2}{\sin^2\theta} + r_h^2 d\Omega_2^2, \tag{4.70}$$

which is the geometry of $AdS_2 \times S^2$.[14]

[13]The study of these dependencies uses perturbative techniques that were originally introduced in CFTs [33].
[14]To get this result we have done two coordinate transformations and then imposed the near horizon limit. First to the tortoise coordinates (t, r^*) and then to (σ, θ). For the specific details see pages 6-7 of Rosso [106].

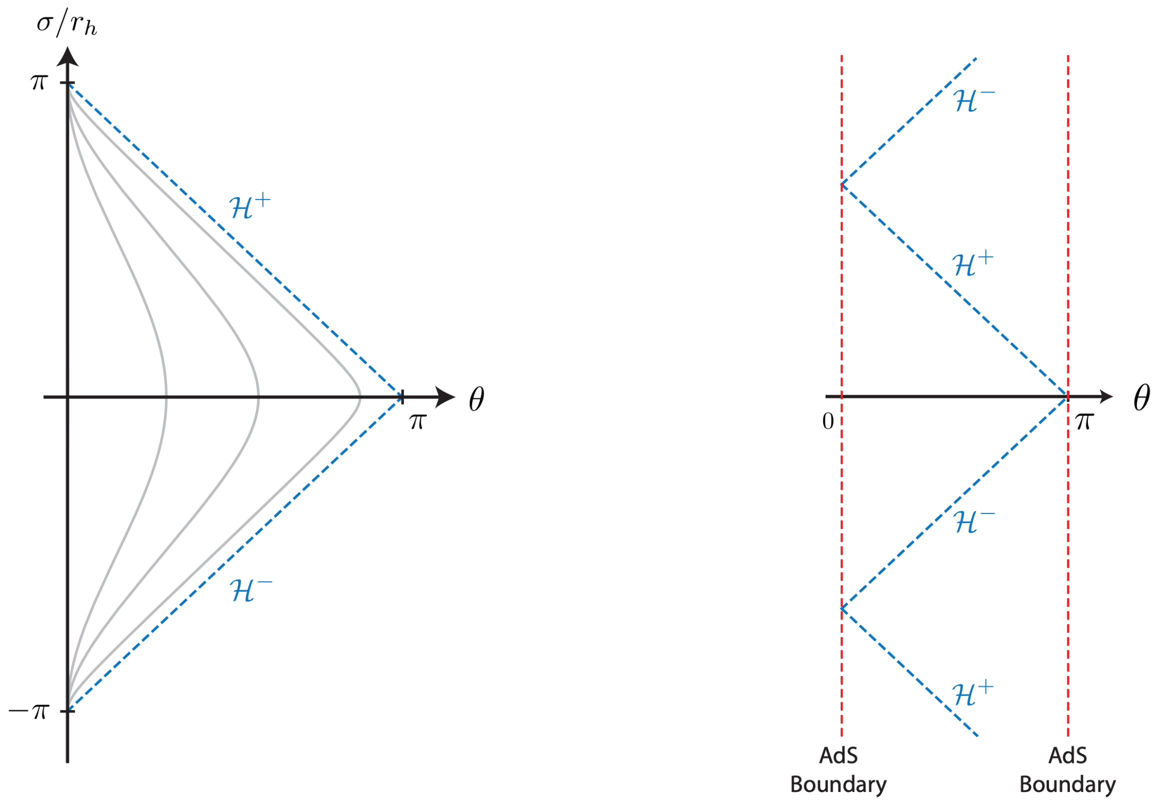


Figure 4.13: Carter-Penrose diagram of the spacetime near the horizon (exterior region of extremal black hole) and its maximal extension. (taken from [106])

We consider the Cauchy surface defined by $\sigma = 0$ and $0 \leq \theta \leq \pi$. Then, the half space A_0 is the region with $\sigma = 0$ and $0 \leq \theta \leq \frac{\pi}{2}$ and, using the null coordinates $\theta_\pm$, its causal domain is

$$\mathcal{D}(A_0) = \left\{ (\sigma, \theta, \vec{v}) \in \mathbb{R} \times (0, \pi) \times \mathbb{R}^2 : \quad \theta_+ < \frac{\pi}{2}, \quad \theta_- < \frac{\pi}{2} \right\}.^{15} \tag{4.71}$$

Now we deform this region in the null direction θ_- using the vector $\zeta^a = \zeta^-(\vec{v})\delta^a_-$ and we obtain the deformed region (see figure 4.14)

$$\mathcal{D}(A) = \left\{ (\sigma, \theta, \vec{v}) \in \mathbb{R} \times (0, \pi) \times \mathbb{R}^2 : \quad \theta_+ < \frac{\pi}{2}, \quad \theta_- < \frac{\pi}{2} + \zeta^-(\vec{v}) \right\}. \tag{4.72}$$

[15] Note that $\vec{v}$ are the coordinates in the sphere.

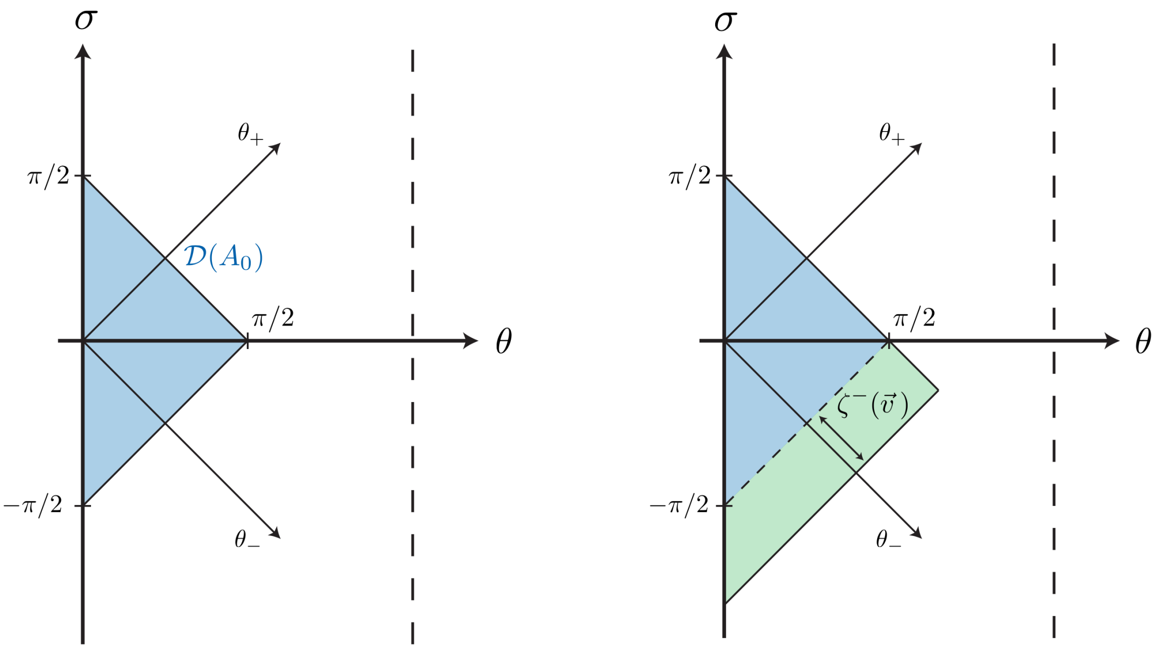


Figure 4.14: (Left) Causal domain $\mathcal{D}(A_0)$. (Right) Deformed causal domain $\mathcal{D}(A)$. (taken from [106])

The main technical result of the paper by Rosso [106] is to show that the modular Hamiltonians are given by

$$K_A - K_{A_0} = \pi r_h^2 \int_{S^2} d\Omega(\vec{x}_\perp)\zeta^-(\vec{x}_\perp) \int_0^\infty d\lambda T_{\lambda\lambda}(\lambda, \vec{x}_\perp) + \mathcal{O}(\zeta^a)^2. \tag{4.73}$$

Subtracting the analogous expression for the complementary subregion A_0 we get the expression for the full modular Hamiltonian

$$\hat{K}_A - \hat{K}_{A_0} = \pi r_h^2 \int_{S^2} d\Omega(\vec{x}_\perp)\zeta^-(\vec{x}_\perp) \int_{-\infty}^\infty d\lambda T_{\lambda\lambda}(\lambda, \vec{x}_\perp) + \mathcal{O}(\zeta^a)^2 \tag{4.74}$$

$$= \pi r_h^2 \int_{S^2} d\Omega(\vec{x}_\perp)\zeta^-(\vec{x}_\perp)\mathcal{E}(\vec{x}_\perp) + \mathcal{O}(\zeta^a)^2. \tag{4.75}$$

Assuming that $\zeta^-(\vec{x}_\perp) \geq 0$, which implies that $\mathcal{D}(A_0) \subseteq \mathcal{D}(A)$, the monotonicity condition of equation 4.68 implies that $\mathcal{E} \geq 0$, which constitutes the proof of the ANEC. □

The proofs for the maximally symmetric spacetimes dS and AdS use the exact same logic. This generalizes the previous result by the same author proving the ANEC for dS and AdS in CFTs [107].

Chapter 5

Quantum Energy Inequalities

We have seen that the classical local energy conditions are violated by negative quantum energies, and that we can feasibly solve this issue by averaging the conditions along geodesics. However, are all local conditions completely ruled out? The second strategy we consider to design quantum energy conditions is to retain locality and lower the bound so that the negative energies are admitted. *Quantum energy inequalities* (QEIs) are attempts at doing that. In this chapter, we will introduce these inequalities and focus specially on the *quantum null energy conditio*n (QNEC), arguably the most interesting of the batch. We will look at the virtues and particularly the challenges that these conditions face in terms of recovering the classical theorems. We will also sketch two different proofs of the QNEC and, finally, the relation between the QNEC and the ANEC will be studied.

5.1 The Quantum Null Energy Condition and its siblings

The first historical occurrence of a QEI dates back to 1978 and is due to Ford [45]. Because negative energies supposed a threat for the second law of thermodynamics, Ford found a lower bound on the magnitude and duration of fluxes of negative energies so that the law was not violated macroscopically.

In [79], Kontou and Sanders give a generic model for the form of any QEI

$$\langle \rho(f) \rangle_w \geq - \langle \mathfrak{Q}(f) \rangle_w . \tag{5.1}$$

Here, ρ represents our desired bounded quantity such as the energy density or $T^{ren}_{ab} k^a k^b$ for null k^a, among other possibilities. The states w, as we mentioned before, are the Hadamard states, and we refer to [37, 39] for details on these states. In a way, they represent physically reasonable states. f is a function in spacetime satisfying $f \geq 0$ and, finally, $\mathfrak{Q}(f)$ is an unbounded operator. The QEI is non-trivial only if the inequality $\langle \rho(f) \rangle_w \leq c + c' |\langle \mathfrak{Q}(f) \rangle_w|$ is not valid for any state and for any constants c and c'. A notable aspect of the operators $\mathfrak{Q}(f)$ is that they are normally state-dependent. In addition, the stress tensor is not an easily generalizable object in QFTs. These facts difficult our goal of finding a universally valid inequality.

If the support of the function f is on a timelike curve $\gamma(\tau)$, the LHS of 5.1 becomes [79]

$$\langle \rho \circ \gamma \rangle_w (f^2) := \int_{-\infty}^{\infty} d\tau f^2(\tau) \langle \rho \rangle_w (\gamma(\tau)) \qquad \text{with } f \in C_0^\infty(\mathbb{R}, \mathbb{R}). \tag{5.2}$$

Let us look now at the main examples of QEIs.

Quantum Weak Energy Condition (QWEC)

The QWEC for the minimally coupled scalar ϕ ($\xi = 0$) can be formulated as [38, 47, 48, 50, 79]

$$\langle :\rho: \circ \gamma \rangle_w (f^2) \geq -\frac{1}{16\pi^2} \int_{-\infty}^{\infty} dt |f''(t)|^2 \tag{5.3}$$

with $:\rho: = T^{ren}_{00} - \langle T^{ren}_{00} \rangle_{w_0}$ where w_0 refers to the vacuum in Minkowski. With a rescaling of f, it can be shown that RHS $\propto -\frac{1}{t_0^4}$. This relation reveals the physical content of the inequality by connecting it to some of our previous results. In the limit $t_0 \to \infty$, the RHS goes to zero, relating to the averaged conditions. In the other limit $t_0 \to 0$ the negative energies have no lower bound, agreeing with the fact that some local negative energies can be made arbitrarily large.

Quantum Strong Energy Condition (QSEC)

For the non-minimally coupled scalar with coupling $\xi \in [0, \xi_c]$ where ξ_c is the conformal coupling, the QSEC is defined as [16, 40]

$$\langle{:}\rho_t{:} \circ \gamma\rangle_w (f^2) \geq -\Big[\mathfrak{Q}_A(f)\mathbb{1} + \langle{:}\phi^2{:} \circ\gamma\rangle_w (\mathfrak{Q}_B(f) + \mathfrak{Q}_C(f))\Big], \tag{5.4}$$

with

$$\mathfrak{Q}_A(f) = \int_0^\infty \frac{d\alpha}{\pi} \left(\phi^*(\hat{\rho}_1, W_0)(\overline{f}_\alpha, f_\alpha) + 2\xi\alpha^2\phi^* W_0(\overline{f}_\alpha, f_\alpha)\right), \tag{5.5}$$

$$\mathfrak{Q}_B[f](\tau) = \frac{1-2\xi}{2} m^2 f^2(\tau) + 2\xi(f'(\tau))^2, \tag{5.6}$$

$$\mathfrak{Q}_C[f](\tau) = f^2(\tau)\xi\mathcal{R}_\xi(\tau), \tag{5.7}$$

where $\rho_t := T_{ab}t^a t^b - \frac{T}{2}$ is the term appearing in the classical SEC and $\mathcal{R}_\xi(\tau) = \left(R_{ab}t^a t^b - \frac{2\xi}{2}R\right)(\tau)$.[1] For details on how these terms are obtained and what their components mean we refer to the original papers [16, 40]. Here, it will suffice to mention that $\mathfrak{Q}_A(f)$ is a state-independent term, $\mathfrak{Q}_B(f)$ is a state-dependent term and $\mathfrak{Q}_C(f)$ is a state-dependent curvature term.

Quantum Null Energy Condition (QNEC)

$$\langle T_{kk}\rangle \geq \frac{1}{2\pi\mathcal{A}} S''_{out}[\Sigma]. \tag{5.8}$$

This is evaluated at a point p for a null vector k^a. Contrary to the QWEC and QSEC as we have expressed them, which only hold for the minimally and non-minimally coupled scalars respectively, this condition is supposed to hold for any theory, free and interacting. In this sense, it is a more interesting energy condition due to its more general status. On the LFS we have the expectation value of the null-null component of the stress-energy tensor. On the RHS, $S_{out}[\Sigma] := -\mathrm{tr}(\rho \ln \rho)$ is the Von Neumann entropy appearing in the expression for the generalised entropy 4.6,[2] and Σ is a codimension-2 surface bounding the spatial region

[1] Note that all these quantities have been originally defined for arbitrary spacetime dimensions but we have written them specifically for $d = 4$.

[2] Note that in this section we have defined the Von Neumann entropy including a minus sign, which we did

containing all the quantum fields that source the entropy. This boundary has the null vector k^a as its normal at point p and $\mathcal{A}$ is defined to be the area of an infinitesimal neighbourhood of p in that boundary. In addition, the boundary Σ divides a Cauchy surface in two. The subscript "out" in the entropy means that we consider it only in one of the two Cauchy sides. Which of the sides we choose is not relevant since the QNEC remains invariant under the swapping of $k^a \leftrightarrow -k^a$. Finally, the quantity that appears in the QNEC is the second functional derivative of the Von Neumann entropy with respect to the affine parameter λ of the null geodesics perpendicular to Σ. The boundary Σ can be deformed along these null geodesics, and the family of thusly deformed boundaries $\Sigma[\lambda]$ will preserve the Cauchy-splitting property [128].

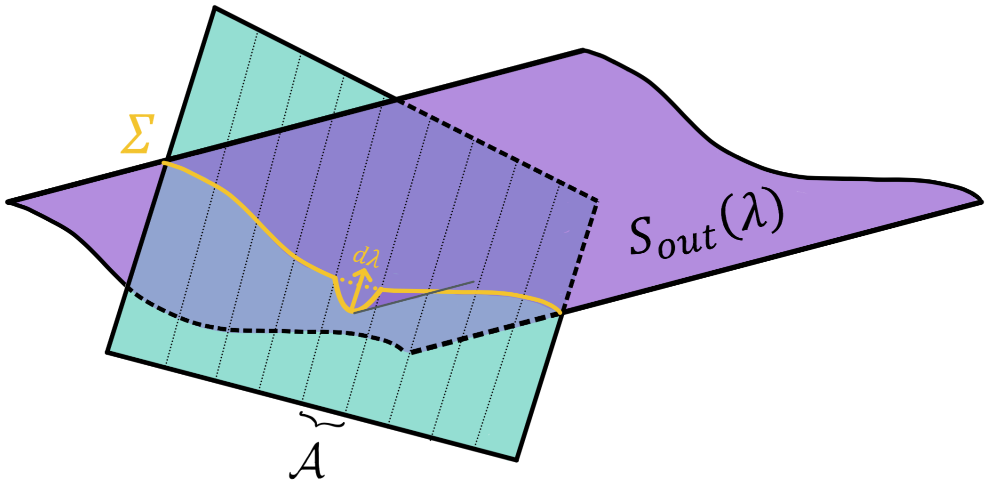


Figure 5.1: Cauchy-splitting spatial surface Σ. The entropy is evaluated in one of the sides of the Cauchy surface (purple). $d\lambda$ represents a small variation in the null direction (diagram idea taken from [12]).

The main difference between this condition and the NEC is that $S''_{out}[\Sigma]$ can be negative, so the QNEC welcomes cases which would otherwise be ruled out by the NEC. Note that while the entropy S_{out} diverges and thus would require a renormalization procedure, its second functional derivative is a finite quantity [12]. The QNEC has received significant attention recently and in section 5.3 we will briefly review some of the recent proofs.

The QNEC can be easily connected to the Quamtum Focussing Conjecture (QFC) proposed in [13] via the generalized 2nd law 4.6. The QFC can be stated as the condition $S''_{gen} \leq 0$ where the primes denote functional derivatives in the null direction.[3] Using this and the generalized

not include in section 4.3.4 where we proved the ANEC from relative entropy.

[3]In the original paper [13], the QFC is defined as the statement $\frac{d\Theta}{d\lambda} \leq 0$ where Θ is the quantum expansion,

2nd law one can recover the QNEC. Roughly stated, this result is obtained by realising that the focussing of the area part in the generalized entropy is proportional to the expectation value of the null-null component of the stress tensor, i.e. $A'' \propto \langle T_{kk} \rangle$.

5.2 Applications revisited for QEIs

Ideally, we would like these conditions to be capable of producing the classical results such as the singularity theorems; be obeyed by all the "physically reasonable" systems; and violated by the "unreasonable" ones. This task presents evident challenges due to the negative local form of these inequalities. In addition, QEIs are statements in QFT with no gravity, so to obtain the classical theorems in general relativity we will somehow have to make a connection with gravity.

For instance, because QEIs impose local and negative constraints, it is at least not obvious how they can be used to prove any singularity theorem. In fact, no such proof has been found yet. However, Brown, Fewster and Kontou have made some progress in hypothesising a potential singularity theorem from the QSEC as stated above (5.4) [15, 36, 40]. Departing from the QSEC and utilizing the semiclassical EFEs they propose the geometric energy condition

$$\int_{-\infty}^{\infty} R_{ab} t^a t^b f^2(\tau) \geq -Q_m(|f^{(m)}|^2 + \tilde{Q}_m |f|^2). \tag{5.9}$$

Rather oddly, this condition only works for even dimensions ($m = \frac{n}{2}$). In addition, the condition requires the Hadamard states w to have a finite maximum magnitude $|(:\Phi:^2\ \gamma)_w| \leq \phi_*^2$. The two Q's are constants which will depend on m, ϕ_* and the mass of the field. Then, with an enhanced initial contraction, they hypothesise that one can construct a singularity theorem. In [79], the same group of researchers claims to be currently working on an actual proof of the singularity theorem, so hopefully we will be able to see some results soon enough. However, as we have seen, such a theorem would require a significant amount of non-trivial assumptions. The QNEC presents even more difficulties which we shall not consider here.

---

a quantity analogous to the classical expansion θ that was defined in equation 2.19 to prove the singularity theorems.

Concerning the task of ruling out exotic spacetimes, it has generally been shown that QEIs allow exotic spacetimes but only where the "exotic part" involves *Planck lengths* [32, 49, 101]. Hence, without a theory of quantum gravity one cannot explore these curiosities any further in the context of semiclassical gravity.

5.3 Proving the QNEC

The QNEC has received a lot of attention int he recent literature and has been proven from different angles, some of them parallel to the strategies we have seen used to prove the ANEC. It was obtained in 2012 for conformal vacuua in $1 + 1$ dimensions from considerations of the generalised 2nd law [129]. In 2016, the QNEC was proven for free and superrenormalizable QFTs [12] and for holographic CFTs [77]. One year later, in 2017, the proof was extended to general interacting CFTs [4]. More recently, in 2018, the QNEC has been obtained from the ANEC [21]. The latest proof of the QNEC uses the Rényi entrioy and is from July 2020 [91].

Here, we will only examine the more consequential proof for interacting conformal theory [4] which utilizes plenty of the technology that we presented when proving the ANEC from causality and from relative entropy (sections 4.3.2 and 4.3.4). In addition, despite being only valid in free theory, we will also consider the latest proof of the ANEC from Rényi entropy [91] because it is peculiarly different from all the arguments previously shown, and because it might point to some interesting directions for future research. The paper relating the QNEC with the ANEC [21] will be reviewed in the next and final section. The following table summarizes the general validity of these proofs.

	GR	**QFT**
Causality+information [4]	$\mathbb{M}^4$	interacting CFTs
Rényi [91]	$\mathbb{M}^4$	free and superrenormalizable field theories
QNEC & ANEC [21]	$\mathbb{M}^4$	algebraic QFT

Table 5.1: General validity of QNEC proofs.

5.3.1 QNEC from causality + quantum information + quantum chaos

This proof from 2017 was proposed by Balakrishnan, Faulkner, Khandker and Wang and applies to all interacting CFTs in flat spacetime [4]. It combines two of the procedures that we have studied to proof the ANEC. First, we will require a more precise notion of microcausality that will be used in the same fashion as in [65] (section 4.3.2). Secondly, we will make use of the evolution by the full modular Hamiltonians introduced in the ANEC proof [35] (section 4.3.4).

We start with the inclusion property for regions A and B, which in QFTs happens at the level of causal domains (causal diamonds and causal wedges). Hence, we have that $\mathcal{D}(B) \subset \mathcal{D}(A)$ where B is a small deformation of A in the null direction as seen in left side of figure 5.2. Taking $\bar{A}$, the complement of A, we obtain the two causally disconnected regions represented in the right side of figure 5.2.

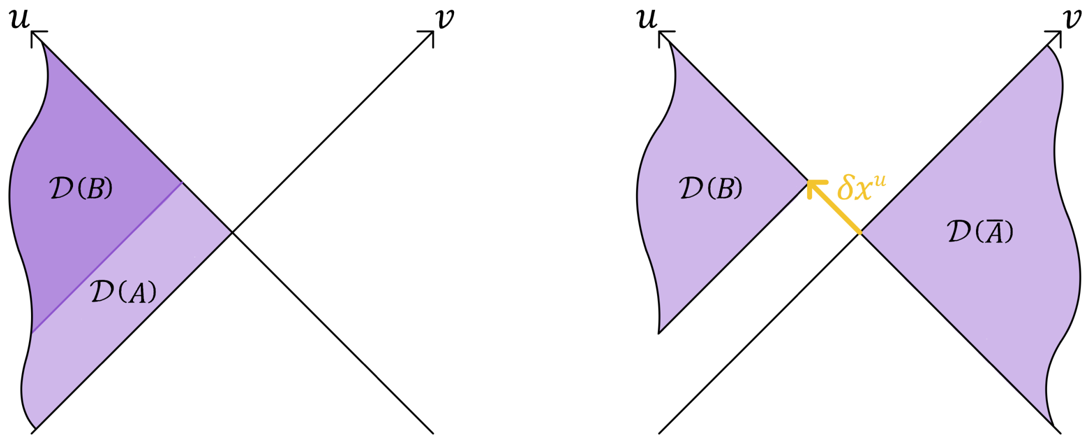


Figure 5.2: (Left) Inclusion property of causal domains $\mathcal{D}(B) \subset \mathcal{D}(A)$. (Right) Causal domains of the two causally disconnected regions, $\mathcal{D}(B)$ and $\mathcal{D}(\bar{A})$. δx^u is the null separation between the two. (diagram idea taken from [4])

The casual disconnection is expressed by the vanishing commutator

$$[\mathcal{D}(B), \mathcal{D}(\bar{A})] = 0. \tag{5.10}$$

As in the ANEC proof, the QNEC will be obtained by answering the question: can we send a signal from region $\mathcal{D}(\bar{A})$ to region $\mathcal{D}(B)$? That is to say, we want to probe the presence of negative energy in the null region between the two causal domains. Because they are spacelike separated, from microcausality we know that any two operators one in each region will have to commute, i.e. $[M_B, M_{\bar{A}}] = 0$. Using this expression for local operators is how we obtained the ANEC. Now, to obtain the QNEC, we need to impose a stronger condition which will make use of the modular Hamiltonians. We will consider time evolution by the modular Hamiltonian, which we get with the reduced density matrix of the state excited by the insertion of the two probe operators $\mathcal{O}_B$ and $\mathcal{O}_{\bar{A}}$.

$$\rho_B = e^{-2\pi K_B^{\psi}} \qquad \rho_{\bar{A}} = e^{-2\pi K_{\bar{A}}^{\psi}}. \tag{5.11}$$

This means that, instead of local operators, we are now dealing with non-local ones, i.e. $M_B(s) = \rho_B^{is}\mathcal{O}_B\rho_B^{-is}$ and $M_{\bar{A}}(s) = \rho_{\bar{A}}^{is}\mathcal{O}_B\rho_{\bar{A}}^{-is}$.

The correlator under consideration is

$$f(s) = \frac{\langle\psi|\, \mathcal{O}_B e^{-isK_B} e^{isK_{\bar{A}}} \mathcal{O}_{\bar{A}}\, |\psi\rangle}{\langle\Omega|\, \mathcal{O}_B e^{-isK_B^0} e^{isK_{\bar{A}}^0} \mathcal{O}_{\bar{A}}\, |\Omega\rangle}. \tag{5.12}$$

This correlator has all the desired properties that a *chaos function* should have, as can be shown from algebraic QFT [4]. However, due to the non-local nature of these operators, the correlator is a quantity highly challenging to compute. This is why we again go to the lightcone limit to compute it, where everything gets significantly simplified. In this limit and in the null coordinates, following an analogous procedure as [65] did for the ANEC, now with a defect OPE and using the replica trick[4] the authors find the perturbative expansion

$$f(s) \simeq 1 - e^s z^d \frac{16\pi\Delta_{\mathcal{O}}}{d(-\Delta v)} \mathcal{Q}(A, B; y) \qquad z^2 := -\frac{\Delta v(\Delta u - \delta x^u)}{4}, \tag{5.13}$$

where

$$\mathcal{Q} := \int_{\partial A}^{\partial B} du\, \langle T_{uu}(u, v = 0, y)\rangle_\psi - \frac{1}{2\pi}\left(\frac{\delta S(B)}{\delta x^u(y)} - \frac{\delta S(A)}{\delta x^u(y)}\right). \tag{5.14}$$

We see that the small correction is a more local version of the null energy operator $\mathcal{E}$ (equation 4.17). Instead of integrating over a complete achronal null geodesic we are now just integrating over the null gap between the two spatially separated regions. Moreover, the derivatives of

[4] For more details on these techniques we refer to their paper [4].

the entanglement entropies at the endpoints are included. As we take the null gap to be infinitesimally small, the stress tensor will be evaluated at a point and the other term will become the second derivative of the entanglement entropy. Hence, the positivity of this quantity will entail the validity of the QNEC. □

To prove that this quantity is positive we again rely on the connection between causality and analiticity of the function $f(s)$ in the s-complex plane. The function will be analytic in the strip of width $\beta = 2\pi$ in the imaginary axis (figure 5.3).

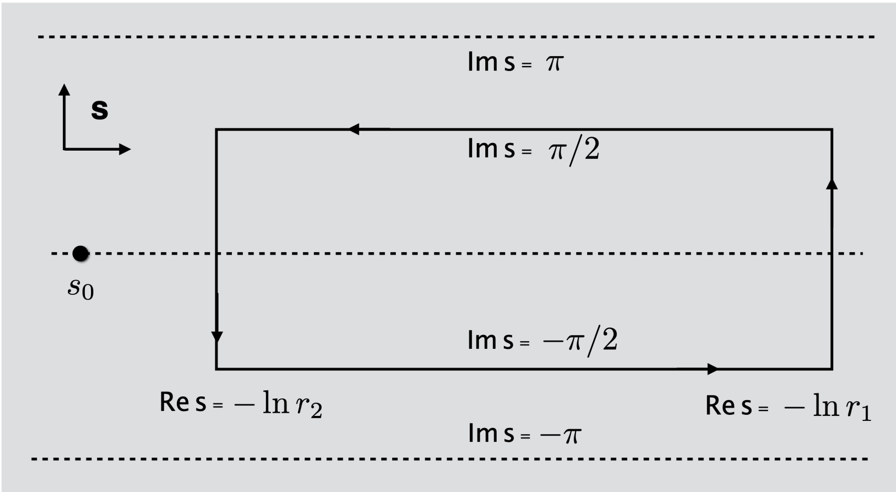


Figure 5.3: Analiticity strip of $f(s)$. The contour in the image, which is analogous to the contour we chose when proving the ANEC, is used in [4] to prove the QNEC. (taken from [4])

For the ANEC, the equivalent strip width was $\beta = \pi$ and here for the QNEC it is extended to 2π. The endlines of the strip can be bounded via a Cauchy-Schwarz logic and using a contour deformation argument it can be shown that the quantity $\mathcal{Q}$ is positive, i.e. $\mathcal{Q} \geq 0$. This is exactly the same idea that is used in the chaos bound by Maldacena et. al. [87]. If the QNEC was to be violated, that is, if $\mathcal{Q} < 0$ then the operators would get shifted in a way that would violate causality.

Nonetheless, note that this argument proves only an integrated version of the QNEC which also includes strong subaddativity. In addition, there are various hints that suggest that the QNEC is actually saturated. But in free QFT one can compute the QNEC and find that it is not saturated. It is not clear yet how to consistently reconcile these considerations.

Finally, it is worth mentioning that this proof involves almost 100 pages of dense and subtle calculations while, as we saw when the QNEC was introduced, by assuming the QFC it can be proved in not much more than 2 lines.

5.3.2 QNEC for Rényi entropy

The next proof of the QNEC we consider was proposed in July 2020 by Moosa, Rath and Su [91] and it shows that a *Rény generalisation* of the QNEC is valid for free and superrenormalizable theories in $d > 2$ spacetime dimensions and $n > 1$ *Rényi parameter*. Basically, the authors claim that the QNEC is a statement about an energy bound which can notably be expressed purely in terms of quantum information via the same relative entropy that we considered to prove the ANEC in section 4.3.4. In fact, we will use the same rewriting in terms of the modular hamiltonian

$$
\begin{aligned}
S(\rho_R|\sigma_R) &= \mathrm{tr}(\rho_R \ln \rho_R) - \mathrm{tr}(\rho_R \ln \sigma) && (5.15)\\
&= \langle K_R^\sigma \rangle_\rho - [S(\rho_R) - S(\sigma_R)] && (5.16)\\
&= \Delta \langle K_R^\sigma \rangle - \Delta S(R), && (5.17)
\end{aligned}
$$

where the delta refers to an evaluation of a quantity in state ρ with respect to the state σ and R is the subregion where the states and consequently the relative entropy are defined. Recall that the meaning that we associate to this quantity is a measure of the distinguishability between the two states. Taking σ to be the vacuum, the modular Hamiltonian can be nicely written as

$$
K_R^\sigma = -\ln \sigma_R = \int_{V(\vec{y})}^{\infty} d\vec{y} d\lambda (\lambda - V(\vec{y})) T_{\lambda\lambda}(\lambda, \vec{y}), \tag{5.18}
$$

where now we are using $\vec{y}$ for the transverse direction and $V(\vec{y})$ is the entangling surface which divides the Cauchy surface Σ (figure 5.4). This also means that the *Hilbert space* is factorizable into pencils (as shown in figure 5.5).

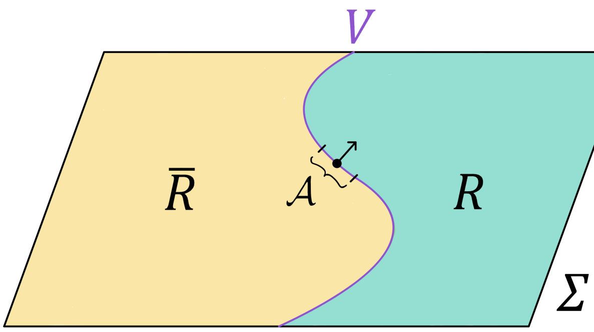


Figure 5.4: The entangling surface $V(y)$ divides the cauchy surface Σ into regions R and $\bar{R}$. The deformation of width $\mathcal{A}$ points towards region R and thus makes it smaller.

With the modular Hamiltonian at hand one can show that

$$S''_{rel} \sim T_{\lambda\lambda} - S''(\rho), \tag{5.19}$$

where again the double prime refers to a second derivative in the null direction evaluated in the same neighbourhood of point p with width $\mathcal{A}$. The null deformation is such that makes R smaller. Then, after taking $\mathcal{A} \to 0$ one can show that the positivity of the quantity in equation 5.19 entails the QNEC. That is to say, the QNEC can be generally restated as the quantum information statement constraining the relative entropy $S''_{rel} \geq 0$ as shown in [81].

This paraphrasing of the QNEC allows for a natural Renyi generalisation that was first conjectured by Lashkari in 2018 [80]. Lashkari could not find any counterexamples to this statement when he was constraining the structure of correlator functions using the monotonicity property of the relative entropy. The Renyi QNEC then reads

$$\frac{1}{\mathcal{A}} S''_n(\rho_A|\sigma_A) \geq 0. \tag{5.20}$$

where $\mathcal{A} \to 0$. $S_n(\rho_A|\sigma_A)$ is the *Sandwhiced Renyi Divergence* (SRD) given by

$$S_n(\rho_A|\sigma_A) = \frac{1}{n-1} \ln \operatorname{tr} \left(\sigma^{\frac{1-n}{2n}} \rho \sigma^{\frac{1-n}{2n}} \right)^n \tag{5.21}$$

and n is the so called Renyi parameter. In order to prove the QNEC one just needs to show that the second null derivative of this quantity is positive. The paper we consider in this section showed that this generalised Renyi QNEC is satisfied for arbitrary states ρ in free and superrenormalizable theories [91]. The technique they use, namely, *null quantization*, is one which is very special of free theories with $d > 2$ and would not work if we add in interactions. The basic idea is that the reduced state ρ_R in the subregion R is unitarily equivalent to the state on the null plane $\mathcal{N}$ in the future of R together with some data at future null infinity $\mathcal{I}^+$ as shown in the causal diamond in left side of figure 5.5 [91].

In free theories, the state in the null plane $\mathcal{N}$ has the nice property of *ultralocality* which allows one to break up the null plane into a tensor product of *pencils*. A pencil is a strip in the null plane with infinitesimal width $\mathcal{A}$ which runs all the way along the affine parameter λ from $-\infty$ to ∞ (right side of figure 5.5).

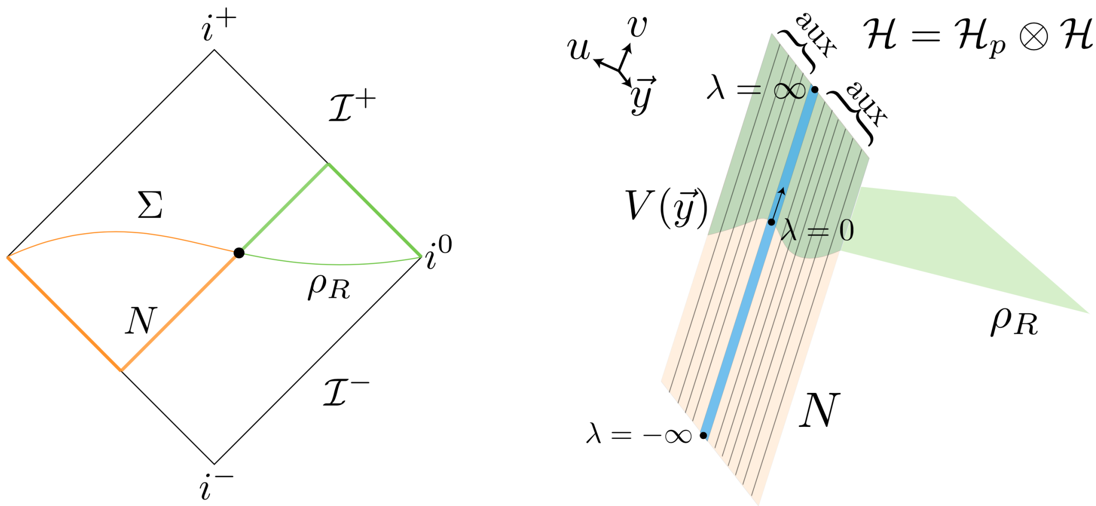


Figure 5.5: (Left) Carter-Penrose diagram showing the spatial region R where ρ is defined (green). R is unitarily equivalent to the two null regions on its future (green). (Right) Pencil decomposition, or null quantization. (taken from [91])

Null Quantization. The reduced state R on a region R (light green) is unitarily equivalent to part of N along with part of null infinity (bold green). On the right, we show the null quantization of N, featuring transverse pencils of transverse area A. Deformations of V (y) around the point p are equivalent to derivatives along the pencil p. The auxiliary system includes both, the other pencils and relevant portions of null infinity.

In other words, light trajectories will remain parallel to each other in a null plane in free theories. If interactions are added in then this would no longer be the case. The pencil under consideration containing the point p where the derivatives of the Rényi entropy are evaluated will contain most of the information we are interested in while the auxiliary region, i.e. all the other pencils will be of no interest to us. As we take the limit $\mathcal{A} \to 0$ we can write the null derivatives appearing in the QNEC as λ-derivatives and, also as a result of that limit, the state of the pencil will get closer to the vacuum with the first order correction being a 1-particle state [91].

$$\rho(\lambda) = \rho^{(0)} + \mathcal{A}^{\frac{1}{2}}\rho^{(1)}(\lambda) + \mathcal{O}(\mathcal{A}) \tag{5.22}$$

$$\hookrightarrow \rho^{(0)} = \underbrace{\sigma_p}_{\text{vacuum pencil}} \otimes \rho^{(0)}_{\text{aux}}, \tag{5.23}$$

$$\hookrightarrow \rho^{(1)}(\lambda) = \underbrace{\sum_i \left(\sigma_p \int dr d\theta f_i(r,\theta)\partial\Phi(re^{i\theta} - \lambda)\right)}_{1-\text{particle pencil}} \otimes \rho_{\text{aux}i}. \tag{5.24}$$

With the expressions for the relevant states at hand, one can then evaluate the SRD relative to the vacuum $S(\rho^{(0)} + \mathcal{A}^{\frac{1}{2}}\rho^{(1)}|\sigma)$. However, using the results from Hijano and May [90] one is allowed to replace the vacuum for nearby states. That is, we can consider the SRD of a particular state plus a small perturbation with respect to that same state $S(\tilde{\rho}^{(0)} + \mathcal{A}^{\frac{1}{2}}\tilde{\rho}^{(1)}|\tilde{\rho}^{(0)})$. To prove that this quantity is positive it is helpful to define the quantity

$$\hat{Z}_n(\rho|\sigma) := \text{tr}\left(\sigma^{\frac{1-n}{2n}}\rho\sigma^{\frac{1-n}{2n}}\right)^n \tag{5.25}$$

so that the SRD becomes

$$S_n(\rho|\sigma) = \frac{1}{n-1}\ln \hat{Z}_n(\rho|\sigma). \tag{5.26}$$

With the perturbative expansion of the state with small parameter $\mathcal{A}$ one can write [91]

$$\hat{Z}_n(\rho(\lambda)|\sigma) = \hat{Z}_n^{(0)} + \mathcal{A}^{\frac{1}{2}}\hat{Z}_n^{(1)} + \mathcal{A}\hat{Z}_n^{(1,1)}(\lambda) + \mathcal{O}(\mathcal{A}, \mathcal{A}^{\frac{3}{2}}), \tag{5.27}$$

where the first order term that is dependent on λ is $\hat{Z}_n^{(1,1)}$. Hence, with the redefinition

$$Z_n^{(1,1)} = \frac{1}{n-1}\hat{Z}_n^{(1,1)} \tag{5.28}$$

the QNEC becomes the statement

$$\frac{d^2}{d\lambda^2} Z_n^{(1,1)} \geq 0. \tag{5.29}$$

Now, using the pencil state we can take a double λ-derivative of this quantity and, after applying some contour manipulations the authors in [91] find the expression[5]

$$\ddot{Z}_n^{(1,1)} = \frac{2n}{n-1} \sum_{k=1}^{n-1} \sin^2\left(\frac{\pi k}{n}\right) \langle \bar{\Psi}_k \Psi_k \rangle \,, \tag{5.30}$$

where the double dot represents the double λ-derivative. We will not deal with the specific expressions for the Ψ's operators here but we note that by reflection positivity the two-point function is positive, i.e. $\langle \bar{\Psi}_k \Psi_k \rangle > 0$. Then, we see that for integer $n > 1$ every term in equation 5.30 is manifestly positive, thus proving the validity of the Rényi QNEC for these states.

The more general proof for arbitrary $n > 1$ is a bit more subtle. The explicit calculation of the second null derivative of the SRD gives the result

$$\ddot{Z}_n^{(1,1)} = \frac{1}{4} \sum_{\alpha\beta} e^{-\pi(K_\alpha + K_\beta)} \int_{-\infty}^{\infty} d\omega e^{-2\pi\omega} \hat{F}_n(\omega) Q(v_{\alpha\beta} - \omega) |M_{\alpha\beta}(\omega)|^2 \tag{5.31}$$

where again we are only interested in the positivity of the terms and they all turn out to be manifestly positive, hence proving the Rényi QNEC for $n > 1$ free field theories. □

Importantly, the authors in [91] note that taking the limit $n \to 1^+$, their expression 5.31 reproduces the original result in [12] which proved the QNEC for free fields. Finally, they also found a specific state that violated the Rényi QNEC for $n < 1$.

5.4 QNEC & ANEC

The relation between the QNEC and the ANEC is an appealing topic of research. The QNEC is a weaker condition that indeed implies the ANEC by integrating it along null geodesics [4].

[5] The calculation is dense and subtle. Here we have only highlighted its main points but for a complete analysis we refer to [91].

However, the last proof of the QNEC we consider will show that their relation seems to be tighter.

This recent proof of the QNEC was proposed by Ceyhan and Faulkner in 2018 and uses algebraic QFT arguments [21]. We include a brief discussion of this paper for a couple of reasons. First, it introduces new methods and perspectives to obtain the QNEC. Second, it not only relates interestingly to other QNEC proofs, but it also connects to the ANEC, which is a striking feature on its own.

The paper shows that the QNEC is logically equivalent to the ANEC for a specific set of states in Minkowski. The way they achieve this result is by deriving a new way of calculating the derivative of relative entropy. As we saw in the last proof (section 5.3.2) the QNEC can be expressed as a statement about derivatives of the relative entropy. The authors in [21] were able to relate the ANEC to this new expression for the derivative of the relative entropy, showing that in this case if the ANEC holds, then the QNEC holds, and vice versa.

In algebraic QFT, we associate specific algebras to spacetime regions. A simple example are *Rindler cuts* where the operators in region $\mathcal{A}$ commute with the operators in the complement region $\mathcal{A}'$. These operators generate Von Neumann algebras in their respective regions. Using the vacuum state $|\Omega\rangle \in \mathcal{H}$, which is cyclic and separating, we can construct the states

$$J_\Omega \Delta_\Omega^{\frac{1}{2}} \alpha \left|\Omega\right\rangle = \alpha^\dagger \left|\Omega\right\rangle \qquad \alpha \in \mathcal{A}, \tag{5.32}$$

which allow us to define the *modular operators* Δ_Ω. J_Ω represents a given anti-unitary operator. We can use a Hilbert space factorisation with the two causally separated spacetime regions to write the modular operators as a tensor product of density matrices, i.e. $\Delta_\Omega = (-\rho')^{-1} \otimes (\rho)$. The action on the operators on these algebras will be determined by the *modular flow*, which is defined as $\sigma_s(\alpha) = \Delta_\Omega^{is} \alpha \Delta_\Omega^{-is}$.[6] The algebras in the null cuts in 2 and higher dimensions can

[6]In the case of null Rindler cuts in vacuum, such action takes the form of a boost.

be translated and boosted:

$$\text{translation} \qquad U_{-a}\mathcal{A}_b U_a = \mathcal{A}_{b+a}, \qquad \text{where } U_a = e^{-iaP}, \tag{5.33}$$

$$\text{boost} \qquad \Delta_\Omega^{is}\mathcal{A}_a\Delta_\Omega^{-is} = \mathcal{A}_{ae^{2\pi s}}. \tag{5.34}$$

In the proof, the authors use *half-sided modular inclusion* [3, 133] to calculate the difference between boosts

$$-\ln\Delta_{\Omega;a_2} + \ln\Delta_{\Omega;a_1} = (a_1 - a_2)P,^7 \tag{5.35}$$

where P is the averged null energy operator (ANE), previously called $\mathcal{E}$.

$$P := \int d\vec{y}\int_{-\infty}^{\infty} dx^+ A(\vec{y})T_{++} \geq 0. \tag{5.36}$$

The positivity of this quantity comes from general properties of modular operators under inclusion, as we saw in section 4.3.4, and it implies the positivity of the ANEC. In other words, the ANEC is equivalent to $P \geq 0$ in this context.

Now we have to relate this operator to the QNEC. As we saw in the last section 5.3.2, the QNEC can be restated as the positivity of the second null derivative of the relative entropy

$$\partial^2 S(a) \geq 0 \tag{5.37}$$

where we write the relative entropy as $S(a) := S_{\text{rel}}(\psi|\Omega;\mathcal{A}_a)$ with the parameter a labeling the cut under consideration. These entropies satisfy an important sum rule

$$-(S(a_2) - S(a_1)) + ((\bar{S}(a_2) - \bar{S}(a_1)) = (a_2 - a_1)P_\psi \tag{5.38}$$

where $P_\psi = \langle\mathcal{E}\rangle_\psi$ is called a *purification*. Taking $a_1 \to a_2$ this becomes

$$-\partial S(a) + \partial\bar{S}(a) = P_\psi. \tag{5.39}$$

[7] This expression is roughly equivalent to equation 4.73.

Assuming $S_{\rm rel}(\phi|\Omega;\mathcal{A}_a) < \infty$ where $\phi = u'\psi$ and $u' \in \mathcal{A}'_a$ the authors in [21] define the function

$$M(a) := \inf \langle\phi|\, P\,|\phi\rangle\,, \tag{5.40}$$

where "inf" refers to the *infimum*, or the greatest lower bound, in this case of the quantity $\langle\phi|\, P\,|\phi\rangle$.

Although slightly in a different form, Wall conjectured that $M(a) = -\partial S(a)$, which gives a new way of calculating the derivative of the relative entropy. This is what the authors in [21] prove. The QNEC follows because the minimization function satisfies $M(a_1) \geq M(a_2)$ for $a_1 < a_2$ which entails $\partial^2 S(a) \geq 0$.

For any purification P_ϕ we have that $\partial\bar{S}_\phi(a) \geq 0$ so from equation 5.39 we obtain the inequality $M(a) \geq -\partial S(a)$. The challenge now is to show that this inequality can actually be saturated. In other words, we want to find a minimising state ϕ_m for which

$$P_\phi = -\partial S(a) \quad\Rightarrow\quad -\partial S(a) \leq M(a) \leq P_{\phi_m} = -\partial S(a) \quad\Rightarrow\quad M(a) = -\partial S(a). \tag{5.41}$$

The minimising states found in [21] are the states $|\psi_s\rangle = u'_s\,|\psi\rangle$ constructed via *relative modular flow* or, more technically, via *Connes cocycles*. The paper contains a detailed study of the entropies of these states and, with the help of two lemmas proved in the paper they are able to relate it to the ANE as

$$-\frac{1}{2\pi}\frac{\delta S_{rel}(\psi|\Omega;x^+(\Delta))}{\delta x^+(y)} = \inf_\phi \langle\phi|\,\mathcal{E}_+(y)\,|\phi\rangle = \inf_{s\in\mathbb{R}} \langle\psi_s|\,\mathcal{E}_+(y)\,|\psi_s\rangle\,, \tag{5.42}$$

$$\mathcal{E}_+(y) := \int_{-\infty}^{\infty} dx^+ T_{++}(x^+,y). \tag{5.43}$$

This shows the main claim of the paper, i.e. that for such minimising states the QNEC reduces to the ANEC. In other words, the positivity of the ANEC in these states just becomes the QNEC implying that the two conditions are logically equivalent in Minkowski.

Conclusion

In this dissertation, we have explored the idea of energy conditions. We have included the main developments from the first appearance of an energy condition to the latest result. We began by reviewing the classical energy conditions, their applications in proving some theorems and the physical systems, classical and quantum, which violate them. From this we concluded that energy conditions were pivotal objects in our comprehension of the universe, and we identified the challenge that the non-inclusion of quantum fields posed, demanding the proposal of more modern quantum energy conditions. We have presented the two strongest candidates, i.e. the AANEC and the QNEC, to tackle this issue, and exposed their most relevant connections to various ideas of current research interest. Despite the significant amount of progress in the recent years, there is still a lot more to be understood about energy conditions.

In terms of proving the AANEC and QNEC, we suggest the following straightforward tracks:

- Extend the AANEC proof from causality in QFT from flat to general spacetimes.

This could facilitate extending the QNEC proofs to some curved spacetimes, which is yet to be done. In addtion, more specifically, we have seen the QNEC proof for general interacting CFTs [4] made use of the results and techniques developed in the AANEC proof in $\mathbb{M}^4$ using modular hamiltonians. We have also seen that the same techniques were used recently to prove the AANEC for dS, AdS and extremal horizons [106]. Hence, another suggestion is

- Integrate the results by Rosso [106, 108] in proving the AANEC to extend the argument made in [4] and [21] to prove the QNEC in (A)dS and potentially other curved spacetimes.

On another note, comparing their performance in recovering classical theorems clearly renders the AANEC in a better shape than the QNEC. This can be seen, for instance, by the lack of any singularity theorem proven from the QNEC up to date. However, the QNEC is still a promising condition, as our discussion in the last chapter showed that it appears to be a fundamental feature of any QFT in $\mathbb{M}^4$. In addition, we also demonstrated how in some special circumstances the ANNEC and the QNEC can be made equivalent. This already suggests some potential directions for future research:

- Prove singularity theorems from the QNEC. This work, as we had explained in section 5.2, is already under construction.
- Explore deeper the connection between the AANEC and QNEC. In particular, see in which way it could be extended for curved spacetimes. There are hints that the same argument that was used to show the connection in Minkowski [21] could be extended to (A)dS spacetimes [108].

Finally, we want to stress again that our conceptual understanding of energy conditions and their fundamental place in the corresponding theoretical framework is still blurred. Because interpretations of ECs depend on their formulation, this results in our ignorance as to whether or not there is a fundamental interpretation that governs all the others and, if so, which one is it. We believe this question should be taken seriously since energy conditions seem to play a key role in the way we understand the universe. Hence, we also suggest as a direction for research:

- Resolve fundamentality status of energy conditions and find their rightful interpretation.

This may be clarified once we assess the general validity of the connection between the AANEC and QNEC, and when we fully develop their applications.

We finish this dissertation by updating our diagram for the logical relations between ECs, where we also include the relations to the boundary causality condition (BCC), the generalized 2$^{\text{nd}}$ law (GSL) and the quantum focusing condition (QFC), all of which have been properly defined in this work. The green line between the ANEC and QNEC refers to the last argument shown in section 5.4. That the BCC imlpies the GSL was shown in a paper by Wall [130], and that through the Gao-Wald theorem it implies the ANEC was shown in [30].

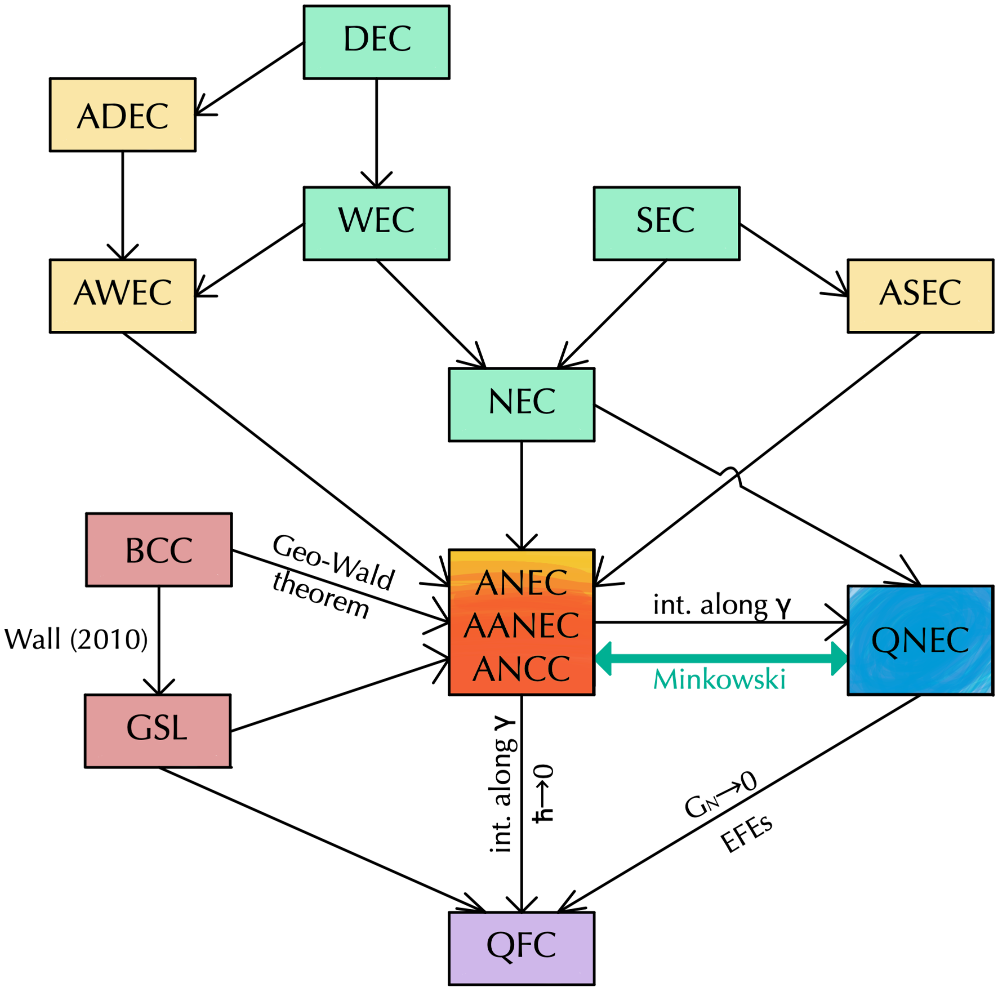


Figure 5.6: Logical relations between the relevant energy conditions.

We believe this is a promising area of research that will produce interesting and relevant contributions to our understanding of the universe in the forthcoming years. In particular, we believe a better understanding of energy conditions will bring us one step closer to a theory of quantum gravity. Hence, we encourage anyone to pursue research in any of the directions suggested here.

Bibliography

[1] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1–29, 2012.

[2] E. Aprile et al. Observation of Excess Electronic Recoil Events in XENON1T. 6 2020.

[3] H. Araki and L. Zsido. Extension of the structure theorem of Borchers and its application to half-sided modular inclusions. *Rev. Math. Phys.*, 17:491–543, 2005.

[4] S. Balakrishnan, T. Faulkner, Z. U. Khandker, and H. Wang. A general proof of the quantum null energy condition. 2017.

[5] C. Barceló, M. Visser, and D. V. Ahluwalia. Twilight for the Energy Conditions? *International Journal of Modern Physics D*, 11(10):1553–1560, Jan. 2002.

[6] J. M. Bardeen, B. Carter, and S. Hawking. The Four laws of black hole mechanics. *Commun. Math. Phys.*, 31:161–170, 1973.

[7] A. Barvinsky, A. Kamenshchik, and A. Starobinsky. Inflation scenario via the Standard Model Higgs boson and LHC. *JCAP*, 11:021, 2008.

[8] D. D. Blanco and H. Casini. Localization of Negative Energy and the Bekenstein Bound. *Phys. Rev. Lett.*, 111(22):221601, 2013.

[9] H. J. Borchers. On the use of modular groups in quantum field theory. *Annales de l'I.H.P. Physique théorique*, 63(4):331–382, 1995.

[10] A. Borde. Geodesic focusing, energy conditions and singularities. *Class. Quant. Grav.*, 4:343–356, 1987.

[11] A. Borde. The existence of singularities in general relativity despite isolated failures of geodesic focusing. *Journal of Mathematical Physics*, 28(11):2683–2684, Nov. 1987.

[12] R. Bousso, Z. Fisher, J. Koeller, S. Leichenauer, and A. C. Wall. Proof of the Quantum Null Energy Condition. *Phys. Rev. D*, 93(2):024017, 2016.

[13] R. Bousso, Z. Fisher, S. Leichenauer, and A. C. Wall. Quantum focusing conjecture. *Phys. Rev. D*, 93(6):064044, 2016.

[14] G. Bressi, G. Carugno, R. Onofrio, and G. Ruoso. Measurement of the Casimir force between parallel metallic surfaces. *Phys. Rev. Lett.*, 88:041804, 2002.

[15] P. Brown, C. Fewster, and E. Kontou. Classical and quantum strong energy inequalities and the Hawking singularity theorem. In *15th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories*, 3 2019.

[16] P. J. Brown, C. J. Fewster, and E.-A. Kontou. A singularity theorem for Einstein–Klein–Gordon theory. *Gen. Rel. Grav.*, 50(10):121, 2018.

[17] R. Caldwell. A Phantom menace? *Phys. Lett. B*, 545:23–29, 2002.

[18] R. R. Caldwell, M. Kamionkowski, and N. N. Weinberg. Phantom energy and cosmic doomsday. *Phys. Rev. Lett.*, 91:071301, 2003.

[19] H. Casimir. On the Attraction Between Two Perfectly Conducting Plates. *Indag. Math.*, 10:261–263, 1948.

[20] H. Casini. Wedge reflection positivity. *J. Phys. A*, 44:435202, 2011.

[21] F. Ceyhan and T. Faulkner. Recovering the QNEC from the ANEC. *Commun. Math. Phys.*, 377(2):999–1045, 2020.

[22] S. Chatrchyan et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B*, 716:30–61, 2012.

[23] E. Curiel. *A Primer on Energy Conditions*, volume 13, pages 43–104. 2017.

[24] E. Curiel. Singularities and black holes. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2020 edition, 2020.

[25] M. P. Dabrowski, T. Stachowiak, and M. Szydlowski. Phantom cosmologies. *Phys. Rev. D*, 68:103519, 2003.

[26] S. de Alwis and N. Ohta. Thermodynamics of quantum fields in black hole backgrounds. *Phys. Rev. D*, 52:3529–3542, 1995.

[27] J.-G. Demers, R. Lafrance, and R. C. Myers. Black hole entropy and renormalization. *Fields Inst. Commun.*, 15:227–232, 2001.

[28] L. Di Luzio, M. Fedele, M. Giannotti, F. Mescia, and E. Nardi. Solar axions cannot explain the XENON1T excess. 6 2020.

[29] A. Einstein. Physics and reality. *Journal of The Franklin Institute*, 221:349–382, Mar. 1936.

[30] N. Engelhardt and S. Fischetti. The Gravity Dual of Boundary Causality. *Class. Quant. Grav.*, 33(17):175004, 2016.

[31] H. Epstein, V. Glaser, and A. Jaffe. Nonpositivity of the energy density in quantized field theories. *Il Nuovo Cimento*, 36(3):1016–1022, Apr. 1965.

[32] A. E. Everett and T. A. Roman. A Superluminal subway: The Krasnikov tube. *Phys. Rev. D*, 56:2100–2108, 1997.

[33] T. Faulkner, R. G. Leigh, and O. Parrikar. Shape Dependence of Entanglement Entropy in Conformal Field Theories. *JHEP*, 04:088, 2016.

[34] T. Faulkner, R. G. Leigh, O. Parrikar, and H. Wang. Modular Hamiltonians for Deformed Half-Spaces and the Averaged Null Energy Condition. *JHEP*, 09:038, 2016.

[35] T. Faulkner, R. G. Leigh, O. Parrikar, and H. Wang. Modular Hamiltonians for Deformed Half-Spaces and the Averaged Null Energy Condition. *JHEP*, 09:038, 2016.

[36] C. Fewster and E. Kontou. A new derivation of singularity theorems with weakened energy hypotheses. *Classical and Quantum Gravity*, 37(6), Mar. 2020. © 2020 The Author(s). Published by IOP Publishing Ltd.

[37] C. J. Fewster. Lectures on quantum energy inequalities. 8 2012.

[38] C. J. Fewster and S. Eveson. Bounds on negative energy densities in flat space-time. *Phys. Rev. D*, 58:084010, 1998.

[39] C. J. Fewster and E.-A. Kontou. Quantum strong energy inequalities. *Phys. Rev. D*, 99(4):045001, 2019.

[40] C. J. Fewster and E.-A. Kontou. Quantum strong energy inequalities. *Phys. Rev. D*, 99(4):045001, 2019.

[41] C. J. Fewster, K. D. Olum, and M. J. Pfenning. Averaged null energy condition in spacetimes with boundaries. *Phys. Rev. D*, 75:025007, 2007.

[42] C. J. Fewster and T. A. Roman. Null energy conditions in quantum field theory. *Phys. Rev. D*, 67:044003, 2003. [Erratum: Phys.Rev.D 80, 069903 (2009)].

[43] C. J. Fewster and T. A. Roman. Erratum: Null energy conditions in quantum field theory [Phys. Rev. D 67, 044003 (2003)]. , 80(6):069903, Sept. 2009.

[44] A. Folacci. Averaged null energy condition for electromagnetism in Minkowski space-time. *Phys. Rev. D*, 46:2726–2729, 1992.

[45] L. Ford. Quantum Coherence Effects and the Second Law of Thermodynamics. *Proc. Roy. Soc. Lond. A*, 364:227–236, 1978.

[46] L. Ford. The Classical singularity theorems and their quantum loop holes. *Int. J. Theor. Phys.*, 42:1219–1227, 2003.

[47] L. Ford, M. J. Pfenning, and T. A. Roman. Quantum inequalities and singular negative energy densities. *Phys. Rev. D*, 57:4839–4846, 1998.

[48] L. Ford and T. A. Roman. Averaged energy conditions and quantum inequalities. *Phys. Rev. D*, 51:4277–4286, 1995.

[49] L. Ford and T. A. Roman. Quantum field theory constrains traversable wormhole geometries. *Phys. Rev. D*, 53:5496–5507, 1996.

[50] L. Ford and T. A. Roman. Restrictions on negative energy density in flat space-time. *Phys. Rev. D*, 55:2082–2089, 1997.

[51] L. Ford and T. A. Roman. The Quantum interest conjecture. *Phys. Rev. D*, 60:104018, 1999.

[52] S. Fulling. *Aspects of Quantum Field Theory in Curved Space-time*, volume 17. 1989.

[53] D. V. Fursaev and S. N. Solodukhin. On one loop renormalization of black hole entropy. *Phys. Lett. B*, 365:51–55, 1996.

[54] G. J. Galloway. Some results on the occurrence of compact minimal submanifolds. *Manuscr. Math.*, 35:209–219, 1981.

[55] G. J. Galloway. Maximum principles for null hypersurfaces and null splitting theorems. *Annales Henri Poincare*, 1:543–567, 2000.

[56] R. Garattini. Casimir Wormholes. *Eur. Phys. J. C*, 79(11):951, 2019.

[57] R. P. Geroch. Topology in general relativity. *J. Math. Phys.*, 8:782–786, 1967.

[58] N. Graham and K. D. Olum. Negative energy densities in quantum field theory with a background potential. *Phys. Rev. D*, 67:085014, 2003. [Erratum: Phys.Rev.D 69, 109901 (2004)].

[59] N. Graham and K. D. Olum. Plate with a hole obeys the averaged null energy condition. *Phys. Rev. D*, 72:025013, 2005.

[60] N. Graham and K. D. Olum. Achronal averaged null energy condition. *Phys. Rev. D*, 76:064001, 2007.

[61] S. Gubser, I. R. Klebanov, and A. M. Polyakov. Gauge theory correlators from noncritical string theory. *Phys. Lett. B*, 428:105–114, 1998.

[62] R. Haag. *Local quantum physics: Fields, particles, algebras.* 9 1992.

[63] T. Hartman, S. Jain, and S. Kundu. A New Spin on Causality Constraints. *JHEP*, 10:141, 2016.

[64] T. Hartman, S. Jain, and S. Kundu. Causality Constraints in Conformal Field Theory. *JHEP*, 05:099, 2016.

[65] T. Hartman, S. Kundu, and A. Tajdini. Averaged Null Energy Condition from Causality. *JHEP*, 07:066, 2017.

[66] S. Hawking. The Chronology protection conjecture. *Phys. Rev. D*, 46:603–611, 1992.

[67] S. W. Hawking. Gravitational radiation from colliding black holes. *Phys. Rev. Lett. 26: No. 21, 1344-6(24 May 1971).*

[68] S. W. Hawking. The occurrence of singularities in cosmology. ii. *Proc. Roy. Soc. (London), Ser. A, 295: 490-3(Dec. 20, 1966).*, 1 1966.

[69] S. W. Hawking. The occurrence of singularities in cosmology. iii. causality and singularities. *Proc. Roy. Soc. (London), Ser. A, 300: 187-201(Aug. 30, 1967).*, 1 1967.

[70] S. W. Hawking. Black holes in general relativity. *Comm. Math. Phys.*, 25(2):152–166, 1972.

[71] S. W. Hawking and G. F. Ellis. *The large scale structure of space-time [by] S. W. Hawking and G. F. R. Ellis.* University Press Cambridge [Eng.], 1973.

[72] S. W. Hawking and R. Penrose. The singularities of gravitational collapse and cosmology. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 314(1519):529–548, 1970.

[73] P. S. Joshi. *Spacetime Singularities*, pages 409–436. 2014.

[74] K. Kannike, M. Raidal, H. Veermäe, A. Strumia, and D. Teresi. Dark Matter and the XENON1T electron recoil excess. 6 2020.

[75] W. R. Kelly and A. C. Wall. Holographic proof of the averaged null energy condition. *Phys. Rev. D*, 90(10):106003, 2014. [Erratum: Phys.Rev.D 91, 069902 (2015)].

[76] G. Klinkhammer. Averaged energy conditions for free scalar fields in flat space-times. *Phys. Rev. D*, 43:2542–2548, 1991.

[77] J. Koeller and S. Leichenauer. Holographic Proof of the Quantum Null Energy Condition. *Phys. Rev. D*, 94(2):024026, 2016.

[78] A. Komar. Necessity of singularities in the solution of the field equations of general relativity. *Phys. Rev.*, 104:544–546, Oct 1956.

[79] E.-A. Kontou and K. Sanders. Energy conditions in general relativity and quantum field theory. 3 2020.

[80] N. Lashkari. Constraining Quantum Fields using Modular Theory. *JHEP*, 01:059, 2019.

[81] S. Leichenauer, A. Levine, and A. Shahbazi-Moghaddam. Energy density from second shape variations of the von neumann entropy. *Phys. Rev. D*, 98:086013, Oct 2018.

[82] M. Lesourd. A remark on the energy conditions for Hawkingarea theorem. *Gen. Rel. Grav.*, 50(6):61, 2018.

[83] M. Lesourd. A remark on the energy conditions for hawking's area theorem. *General Relativity and Gravitation*, 50, 06 2018.

[84] D. Liu and M. Reboucas. Energy conditions bounds on f(T) gravity. *Phys. Rev. D*, 86:083515, 2012.

[85] R. Longo. Entropy distribution of localised states. *Commun. Math. Phys.*, 373(2):473–505, 2019.

[86] J. Maldacena, A. Milekhin, and F. Popov. Traversable wormholes in four dimensions. 7 2018.

[87] J. Maldacena, S. H. Shenker, and D. Stanford. A bound on chaos. *JHEP*, 08:106, 2016.

[88] J. M. Maldacena. The Large N limit of superconformal field theories and supergravity. *Int. J. Theor. Phys.*, 38:1113–1133, 1999.

[89] A. Maleknejad and M. Sheikh-Jabbari. Revisiting Cosmic No-Hair Theorem for Inflationary Settings. *Phys. Rev. D*, 85:123508, 2012.

[90] A. May and E. Hijano. The holographic entropy zoo. *JHEP*, 10:036, 2018.

[91] M. Moosa, P. Rath, and V. P. Su. A Renyi Quantum Null Energy Condition: Proof for Free Field Theories. 7 2020.

[92] M. Morris and K. Thorne. Wormholes in space-time and their use for interstellar travel: A tool for teaching general relativity. *Am. J. Phys.*, 56:395–412, 1988.

[93] M. Morris, K. Thorne, and U. Yurtsever. Wormholes, Time Machines, and the Weak Energy Condition. *Phys. Rev. Lett.*, 61:1446–1449, 1988.

[94] B. O'Neill. *Semi-Riemannian Geometry With Applications to Relativity*. 1983.

[95] H. Osborn and A. Petkou. Implications of conformal invariance in field theories for general dimensions. *Annals Phys.*, 231:311–362, 1994.

[96] D. N. Page, S. Surya, and E. Woolgar. Positive mass from holographic causality. *Phys. Rev. Lett.*, 89:121301, 2002.

[97] R. D. Peccei and H. R. Quinn. Constraints imposed by CP conservation in the presence of pseudoparticles. *Phys. Rev. D*, 16:1791–1797, Sep 1977.

[98] R. D. Peccei and H. R. Quinn. CP conservation in the presence of pseudoparticles. *Phys. Rev. Lett.*, 38:1440–1443, Jun 1977.

[99] R. Penrose. Gravitational collapse and space-time singularities. *Phys. Rev. Letters*, Vol: 14(57-59), 1 1965.

[100] R. Penrose, R. Sorkin, and E. Woolgar. A Positive mass theorem based on the focusing and retardation of null geodesics. 2 1993.

[101] M. J. Pfenning and L. Ford. The Unphysical nature of 'warp drive'. *Class. Quant. Grav.*, 14:1743–1751, 1997.

[102] G. Plunien, B. Muller, and W. Greiner. The Casimir Effect. *Phys. Rept.*, 134:87–193, 1986.

[103] A. Raychaudhuri. Relativistic cosmology. i. *Phys. Rev.*, 98:1123–1126, May 1955.

[104] T. Roman. Quantum Stress Energy Tensors and the Weak Energy Condition. *Phys. Rev. D*, 33:3526–3533, 1986.

[105] T. Roman. On the 'Averaged Weak Energy Condition' and Penrose's Singularity Theorem. *Phys. Rev. D*, 37:546–548, 1988.

[106] F. Rosso. Achronal averaged null energy condition for extremal horizons and (A)dS. *JHEP*, 07:023, 2020.

[107] F. Rosso. Global aspects of conformal symmetry and the ANEC in dS and AdS. *JHEP*, 03:186, 2020.

[108] F. Rosso. Global aspects of conformal symmetry and the ANEC in dS and AdS. *JHEP*, 03:186, 2020.

[109] R. Schoen and S. T. Yau. On the proof of the positive mass conjecture in general relativity. *Comm. Math. Phys.*, 65(1):45–76, 1979.

[110] R. Schoen and S. T. Yau. Proof of the positive mass theorem. ii. *Comm. Math. Phys.*, 79(2):231–260, 1981.

[111] K. Schwarzschild. On the gravitational field of a mass point according to Einstein's theory. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1916:189–196, 1916.

[112] M. Setare and A. Seyedzahedi. Casimir Energy between a Sinusoidally Corrugated Sphere and a Plate Using Proximity Force Approximation. *Indian J. Phys.*, 90(5):583–588, 2016.

[113] S. N. Solodukhin. One loop renormalization of black hole entropy due to nonminimally coupled matter. *Phys. Rev. D*, 52:7046–7052, 1995.

[114] V. Sopova and L. Ford. The Energy density in the Casimir effect. *Phys. Rev. D*, 66:045026, 2002.

[115] R. D. Sorkin and E. Woolgar. Causal structure and the positivity of mass. In *Marcel Grossmann Meeting on General Relativity*, page 754, Jan. 1992.

[116] M. J. Sparnaay. Attractive Forces between Flat Plates. , 180(4581):334–335, Aug. 1957.

[117] M. J. Sparnaay. Measurements of attractive forces between flat plates. *Physica*, 24(6):751–764, Jan. 1958.

[118] F. Takahashi, M. Yamada, and W. Yin. XENON1T anomaly from anomaly-free ALP dark matter and its implications for stellar cooling anomaly. 6 2020.

[119] F. J. Tipler. Energy conditions and spacetime singularities. *Phys. Rev. D*, 17:2521–2528, 1978.

[120] R. Verch. The Averaged Null energy condition for general quantum field theories in two-dimensions. *J. Math. Phys.*, 41:206–217, 2000.

[121] M. Visser. *Lorentzian wormholes: From Einstein to Hawking.* 1995.

[122] M. Visser. Gravitational vacuum polarization. ii. energy conditions in the boulware vacuum. *Phys. Rev. D*, 54:5116–5122, Oct 1996.

[123] R. M. Wald. Asymptotic behavior of homogeneous cosmological models in the presence of a positive cosmological constant. *Phys. Rev. D*, 28:2118–2120, 1983.

[124] R. M. Wald. Asymptotic behavior of homogeneous cosmological models in the presence of a positive cosmological constant. *Phys. Rev. D*, 28:2118–2120, 1983.

[125] R. M. Wald. *General Relativity.* Chicago Univ. Pr., Chicago, USA, 1984.

[126] R. M. Wald. *Quantum field theory in curved spacetime and black hole thermodynamics.* 1994.

[127] A. C. Wall. Proving the Achronal Averaged Null Energy Condition from the Generalized Second Law. *Phys. Rev. D*, 81:024038, 2010.

[128] A. C. Wall. A proof of the generalized second law for rapidly changing fields and arbitrary horizon slices. *Phys. Rev. D*, 85:104049, 2012. [Erratum: Phys.Rev.D 87, 069904 (2013)].

[129] A. C. Wall. Testing the Generalized Second Law in 1+1 dimensional Conformal Vacua: An Argument for the Causal Horizon. *Phys. Rev. D*, 85:024015, 2012.

[130] A. C. Wall. The Generalized Second Law implies a Quantum Singularity Theorem. *Class. Quant. Grav.*, 30:165003, 2013. [Erratum: Class.Quant.Grav. 30, 199501 (2013)].

[131] A. C. Wall. Lower Bound on the Energy Density in Classical and Quantum Field Theories. *Phys. Rev. Lett.*, 118(15):151601, 2017.

[132] E. T. Whittaker. On gauss' theorem and the concept of mass in general relativity. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 149(867):384–395, 1935.

[133] H. Wiesbrock. Half-sided modular inclusions of algebras in quantum field theory. In *11th International Conference on Mathematical Physics (ICMP-11) (Satellite colloquia: New Problems in the General Theory of Fields and Particles, Paris, France, 25-28 Jul 1994)*, pages 359–365, 1994.

[134] E. Winstanley. Renormalized black hole entropy via the 'brick wall' method. *Phys. Rev. D*, 63:084013, 2001.

[135] E. Witten. A new proof of the positive energy theorem. *Comm. Math. Phys.*, 80(3):381–402, 1981.