To Infinity and Beyond

An Introduction to BMS Symmetries

Níckolas de Aguiar Alves

ABSTRACT: These are the lecture notes for a minicourse discussing the Bondi–Metzner–Sachs (BMS) symmetries, which are the symmetries at the conformal null infinity of asymptotically flat spacetimes. The course is part of the I São Paulo School on Gravitational Physics and it is intended to be self-contained, working mostly upon the assumption of previous knowledge about general relativity at undergraduate level (including some notions of differential geometry). We discuss the basic elements of group theory required to understand the BMS group, the construction of asymptotic null infinity and the definition of asymptotically flat spacetimes, some other general relativistic prerequisites such as Lie derivatives and conformal transformations, the BMS group itself, and some modern applications of BMS symmetries to fundamental physics.

KEYWORDS: BMS group, general relativity, asymptotic symmetries, conformal infinity.

VERSION: July 19, 2024.

Contents

1 Introduction

The notion of symmetry is ubiquitous throughout physics. It is at the core of every physical theory and one could argue it is a necessity for one to even do physics at all. Symmetries allow us to greatly simplify the difficult problems we encounter in the Universe and approach them in a feasible manner, and they often reflect fundamental insights about the structure of nature.

When studying symmetries, the concept of group becomes particularly useful (see, *e.g.*, Zee 2016). Groups are abstract mathematical structures similar to, but simpler than, vector spaces.

The "collection" of symmetries of a certain kind of a given physical system often, if not always, constitutes a group. Hence, knowing the basic parlance of group theory is a powerful manner of studying symmetries and their consequences.

The Bondi–Metzner–Sachs (BMS) group is the group of symmetries at infinity in asymptotically flat spacetimes. In other words, consider the following scenario in general relativity. One has a given distribution of matter in spacetime arranged in such a manner that, far away from all of this matter, spacetime is nearly flat in a suitable sense (which we shall define). Far from all these sources, spacetime resembles in a suitable sense Minkowski spacetime, so we would expect that the symmetries at infinity should somehow resemble the symmetries of Minkowski spacetime—the latter form the so-called Poincaré group. The BMS group is this group of symmetries at infinity, and it was a surprise when Bondi, Metzner, and Sachs discovered it was actually a group much bigger than the Poincaré group (Bondi, Van der Burg, and Metzner 1962; Sachs 1962b).

This simple fact hides deep physical truths about general relativity. Namely, it is a statement that general relativity *does not* reduce to special relativity at large distances. Rather, it reduces to a much more complex structure that arises only at infinity. This rich structure at infinity can be exploited, as done by Dappiaggi, Moretti, and Pinamonti (2017), to construct physically interesting states for quantum fields evolving on asymptotically flat backgrounds. There are also other physical implications. For example, as reviewed by Strominger (2018), the BMS symmetries imply correlations between S-matrix elements that yield the so-called soft graviton theorem (Weinberg 1965, 1995), which is a fundamental ingredient for understanding the infrared structure of cross sections in scattering experiments. BMS transformations can also be related to the so-called memory effect—originally discovered by Zel'dovich and Polnarev (1974)—which predicts that after the passage of a gravitational wave two nearby detectors will be permanently displaced. It is expected that this effect should be measurable in the near future (Favata 2010; Grant and Nichols 2023).

In these lecture notes, we provide a simple introduction to the BMS group and some of its physical consequences. We also discuss some prerequisites that are essential to understanding the discussion, such as group theory and the concept of conformal infinity. It is assumed that the reader is familiar with general relativity at an undergraduate level, and some applications might require previous knowledge of quantum field theory to be fully appreciated (although the author tried his best to keep the necessary knowledge to a minimum). Through familiarity with general relativity we also assume some familiarity with differential geometry. We use the same notation and conventions employed in the textbook by Wald (1984). This includes abstract index notation, metric signature $-+++$, and geometric units with $G = c = 1$. Latin indices $a, b, \ldots$ represent abstract indices, Greek indices $\mu, \nu, \ldots$ represent spacetime coordinate indices.

2 Symmetries and Groups

We begin by discussing what is the mathematical structure of the symmetries of a physical system. This will naturally lead us to the notion of group, which is a well-established and well-studied concept in mathematics. We shall then understand some of the basic ideas of group theory and in particular exploit them to investigate properties of general relativistic spacetimes.

2.1 Rotations

Perhaps the paradigmatic example of a symmetry is rotational symmetry. This means that the physical system does not change when we rotate it about a specific point and is present in many interesting examples in fundamental physics. For example, the hydrogen atom model in quantum mechanics presents spherical symmetry. So does the Schwarzschild solution in general relativity and many star solutions of interest.

Let us for a moment ponder about what is it that *defines* a rotation as a rotation. To be concrete, we are thinking right now about rotations as operations on three-dimensional real vectors, although the discussion is easily generalized to higher dimensions. We know that a generic rotation should be a linear transformation $R\colon \mathbb{R}^3 \to \mathbb{R}^3$. Indeed, $R(\alpha\vec{v}) = \alpha R\vec{v}$, for rotations should not "see" the length of a vector: they only care about its direction. Furthermore, $R(\vec{u} + \vec{v}) = R\vec{u} + R\vec{v}$, because the rotation occurs in a "solid" manner: both elements of the sum and the sum itself should all be rotated in precisely the same way. Hence, we conclude a general rotation should be a linear transformation.

This, however, is surely not enough to characterize what we mean by a rotation. There are far too many linear transformations! For example, given the canonical basis $\{\vec{e}_i\}$, we could pick the transformation defined by $T\vec{e}_i = (1 + \delta_{i2})\vec{e}_i$, where δ_{ij} is the Kronecker delta. This transformation simply stretches one of the coordinate axes, while keeping the remaining ones constant. This is surely a linear transformation, but it looks nothing like a translation.

The extra property that characterizes rotations is that they preserve angles and norms. Hence, they do not impart any sort of stretching into the vectors they act on and they do not change the relative angle between vectors. Mathematically this is expressed in terms of the scalar product in $\mathbb{R}^3$ as

$$R\vec{u} \cdot R\vec{v} = \vec{u} \cdot \vec{v}. \tag{1}$$

This can be shown to be equivalent to the requirement that

$$R^{\intercal}R = \mathbb{1}, \tag{2}$$

where $R^{\intercal}$ is the transpose of R and $\mathbb{1}$ is the identity matrix. Matrices that satisfy Eq. (2) are said to be orthogonal. This invites us to define

$$\mathrm{O}(3) = \left\{R \in \mathbb{M}_3; R^{\intercal}R = \mathbb{1}\right\}, \tag{3}$$

where $\mathbb{M}_n$ is the space of $n \times n$ real matrices and the "O" in O(3) stands for "orthogonal". O(3) is the so-called orthogonal group in three dimensions. Notice that we can analogously define the orthogonal group in n dimensions through

$$\mathrm{O}(n) = \left\{R \in \mathbb{M}_n; R^{\intercal}R = \mathbb{1}\right\}. \tag{4}$$

This is the collection (or rather the group) of rotations in n dimensions.

We would like to know which sort of structure represents collections of symmetries in general, so it is interesting for us to study some properties of $\mathrm{O}(n)$ to figure out an abstract definition. $R^{\intercal}R = \mathbb{1}$ seems too strict, since this is specific to rotations, and not every symmetry is a rotation. In fact, even linearity might be a stretch, since complicated symmetries could be nonlinear in principle.

One of the most basic facts we can notice about $O(n)$ is that it is endowed with a product. Namely, given two rotations R and S, we can also define RS, which stands for "apply the rotation S, and then the rotation R". This constitutes a rotation, for one can show through direct calculation that $RS \in O(n)$. This is our first axiom for a group: a group should be a set G endowed with a product $\cdot\colon G \times G \to G$. Indeed, notice that if we apply two symmetry transformations in sequence, the result should also be a symmetry. This is due to the fact that a symmetry related two physically equivalent configurations, and thus applying two symmetries in sequence yields a configuration that is physically equivalent to the original one.

Next, we notice that the product in $O(n)$ is associative. Indeed, given rotations R, S, and T, we have that $(RS)T = R(ST)$. We will also impose associativity as one of the axioms for a group. This is due both to the fact that it is mathematically interesting to do so—the resulting theory is quite rich—and due to the fact that it is even difficult to conceive how one could apply transformations to a physical system in a manner which is not associative. If we think of the transformation $(RS)T$ in terms of two consecutive transformations (T and then RS), it is not obvious that it is equivalent to $R(ST)$ (ST and then R). However, it seems intuitive that these two transformation processes should be equivalent to the "physical realization" T, then S, then R, which we would denote as RST (without parentheses). Hence, it seems these transformations should compose in an associative manner. A second argument could be that each transformation should be viewed as a function between configurations of a physical system, and the composition of functions is always associative.

We then notice that "doing nothing" should be considered a symmetry. Indeed, a physical configuration is physically indistinguishable from itself. Hence, we expect there to be an identity in a group. And, correctly, we have that $\mathbb{1} \in O(n)$, as one can promptly check.

Finally, undoing a symmetry transformation also constitutes a symmetry. If we perform a transformation in the system in such a way that the configuration after the transformation is indistinguishable from the configuration before the transformation, it is surely true that the statement also holds the other way around. Thus, the group should be populated with the inverses of every group element. As expected, for every $R \in O(n)$, $R^{-1} = R^{\intercal} \in O(n)$.

We thus arrive at the following definition.

Definition 1 [Group]:

A *group* is a pair $(G, \cdot)$ where G is a set and $\cdot\colon G \times G \to G$ is a function satisfying the following conditions:

i. $\cdot$ is associative, so $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$ for every $g_1, g_2, g_3 \in G$,

ii. $\cdot$ has a neutral element e, *i.e.*, there is an element $e \in G$ with $e \cdot g = g \cdot e = g$ for all $g \in G$,

iii. all elements in G have inverses, meaning that for each $g \in G$ there is some $g^{-1} \in G$ with $g \cdot g^{-1} = g^{-1} \cdot g = e$.

Groups which are also commutative ($g_1 \cdot g_2 = g_2 \cdot g_1$ for all $g_1, g_2 \in G$) are said to be *Abelian*. We often omit $\cdot$ when writing products when the product operation is clear: $g_1 g_2 = g_1 \cdot g_2$. ♠

Our interest in groups is similar to one's interest in vector spaces. These are useful mathematical concepts that have been well-developed by mathematicians and that occur once and again in physical contexts. As such, it is useful to know a thing or two about groups because this helps us understanding physical phenomena better. Zee (2016) discusses applications of group theory to physics. In the following, we will restrict our attention to groups that are of interest for our discussions.

Just as one has the notion of a subspace of a vector space, we also have the notion of a subgroup of a group.

Definition 2 [Subgroup]:
Let $(G, \cdot)$ be a group and $H \subseteq G$. We say $(H, \cdot)$ is a *subgroup* of $(G, \cdot)$ if $(H, \cdot)$ is a group. Notice the product has to be inherited from the "mother" group. ♠

As an example, consider

$$\mathrm{SO}(n) = \left\{R \in \mathbb{M}_n; R^{\intercal} R = \mathbb{1}, \det R = +1\right\}. \tag{5}$$

Notice that $\mathrm{SO}(n) \subseteq \mathrm{O}(n)$. In fact, every element $R \in \mathrm{O}(n)$ either lies in $\mathrm{SO}(n)$ or has $\det R = -1$. One can show that $\mathrm{SO}(n)$ is a subgroup of $\mathrm{O}(n)$, and it is $\mathrm{SO}(n)$ that is often referred to as the group of rotations. The reason is that $\mathrm{O}(n)$ is essentially $\mathrm{SO}(n)$ with the possibility of composing an element of $\mathrm{SO}(n)$ with a reflection, so $\mathrm{SO}(n)$ restricts our attention to the rotations that involve no reflections. The "S" in $\mathrm{SO}(n)$ stands for "special", and refers to the fact that the matrices have unit determinant.

In order to understand further the structure of $\mathrm{SO}(n)$ and $\mathrm{O}(n)$, it is useful to define a Lie group.

Definition 3 [Lie Group]:
A *Lie group* is a group $(G, \cdot)$ endowed with a smooth manifold structure in such a manner that the maps $\mu \colon G \times G \to G$ and $\iota \colon G \to G$ given by $\mu(g_1, g_2) = g_1 \cdot g_2$ and $\iota(g) = g^{-1}$ are smooth. ♠

Hence, a Lie group is a smooth group. This is useful because often in physics we encounter groups that have infinitely many elements, where "infinitely many" is meant in a continuous (or, more appropriately, smooth) way. This is the case of $\mathrm{O}(n)$ and $\mathrm{SO}(n)$. We can define coordinates on these groups by employing the so-called Euler angles (see, for example, Goldstein 1980). $\mathrm{SO}(n)$ is composed of a connected manifold, which means it is a "unique continuous piece". $\mathrm{O}(n)$, on the other hand, has two connected components, meaning there is a continuous piece with $\det R = +1$ and another with $\det R = -1$—one cannot smoothly change the signal of the determinant from $+1$ to -1.

2.2 Lorentz Group

Let us now move on to a more complicated example. Let us consider Minkowski spacetime $(\mathbb{R}^4, \eta_{ab})$. We would like to investigate the linear isometries of this spacetime. Isometries are metric-preserving diffeomorphisms, which means that they are transformations that preserve both the underlying space $\mathbb{R}^4$ and the metric η_{ab}. Hence, they can be understood as the fundamental symmetries of the spacetime. Linear isometries are a particular class of isometries that has the form

$$x^{\mu} \to x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} \tag{6}$$

for some matrix $\Lambda^{\mu}{}_{\nu}$.

For a diffeomorphism to preserve the metric it should satisfy

$$\eta_{\mu\nu} \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x'^{\nu}}{\partial x^{\sigma}} = \eta_{\rho\sigma} . \tag{7}$$

For a linear transformation, we can write this simply as

$$\eta_{\mu\nu} \Lambda^{\mu}{}_{\rho} \Lambda^{\nu}{}_{\sigma} = \eta_{\rho\sigma} . \tag{8}$$

In Cartesian coordinates, we can write η as the matrix

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{9}$$

and the expression ccan be written simply as

$$\Lambda^{\intercal} \eta \Lambda = \eta . \tag{10}$$

We thus define the group

$$\mathrm{O}(3,1) = \left\{ \Lambda \in \mathbb{M}_4 ; \Lambda^{\intercal} \eta \Lambda = \eta \right\}, \tag{11}$$

known as the $(3,1)$-pseudo-orthogonal group. “3” refers to the three positive signs in the Minkowski metric and “1” refers to the remaining negative sign. Notice this is a generalization of the orthogonal groups: $\mathrm{O}(n)$ could be defined as the group with matrices satisfying $R^{\intercal} \mathbb{1} R = \mathbb{1}$, which is similar to the condition defining $\mathrm{O}(3,1)$. We could also define $\mathrm{O}(p,q)$ in more generality, but these groups are not of interest for us.

$\mathrm{O}(3,1)$ is known as the Lorentz group. It is composed of rotations, Lorentz boosts, and both spatial and time reflections. Notice that $\mathrm{O}(3)$ is a subgroup of $\mathrm{O}(3,1)$ formed by the elements with

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & R & \\ 0 & & & \end{pmatrix}, \tag{12}$$

where R is an orthogonal matrix.

One could argue, correctly, that technically (12) is not an element of $\mathrm{O}(3)$. Indeed, $\mathrm{SO}(3)$ is formed by 3×3 matrices and (12) is obviously a 4×4 matrix. Yet, it seems odd to not consider (12) an element of $\mathrm{O}(3)$, given that matrices of that form behave precisely as elements of $\mathrm{O}(3)$. Let us then give meaning to this.

Definition 4 [Homomorphism]:

Let G and H be groups. We say a function $\phi \colon G \to H$ is a *homomorphism* if, and only if, $\phi(g_1)\phi(g_2) = \phi(g_1 g_2)$ for every $g_1, g_2 \in G$. ♠

Notice we wrote simply G rather than the tuple $(G, \cdot)$. This is common in group theory, and perhaps in all of mathematics. Notice also that the definition of a homomorphism resembles that of a linear transformation between vector spaces: both of them are mappings between two algebraic structures of the same kind (either groups or vector spaces) that preserve the algebraic structure. Hence, they are natural maps between these sorts of structures.

Just as in linear algebra, we can now define an isomorphism.

Definition 5 [Isomorphism]:

Let G and H be groups and $\phi\colon G \to H$ be a homomorphism. If ϕ is bijective, it is said to be an isomorphism. ♠

The inverse of a homomorphism, when it exists, is automatically a homomorphism. Isomorphisms between groups are analogous to isomorphims between vector spaces: they mean the two groups (or vector spaces) are algebraically "equal" as far as the structure they preserve is concerned. Two isomorphic groups can be understood as two copies of the same group.

Now we can give meaning to our previous idea: there is an injective homomorphism from $\mathrm{O}(3)$ into $\mathrm{O}(3,1)$. Or, alternatively, $\mathrm{O}(3,1)$ has a subgroup that is isomorphic to $\mathrm{O}(3)$. We often just make these statements briefly by saying $\mathrm{O}(3)$ is a subgroup of $\mathrm{O}(3,1)$, where the necessary homomorphisms are understood.

Just as we defined $\mathrm{SO}(n)$ to get rid of the reflections of $\mathrm{O}(n)$, we can define $\mathrm{SO}(3,1)$ to focus on a smaller subgroup of $\mathrm{O}(3,1)$. There is, however, a caveat: while $\mathrm{O}(n)$ has two connected components, $\mathrm{O}(3,1)$ has four. This is because in addition to spatial reflections we also have the independent time reflections. If we reverse time and apply a spatial reflection, we get a transformation with positive determinant, but which certainly is not a composition of "pure" rotations and boosts. Thus, we provide the following definitions.

$$\mathrm{SO}(3,1) = \left\{\Lambda \in \mathbb{M}_4; \Lambda^\intercal \eta \Lambda = \eta, \det \Lambda = +1\right\} \tag{13}$$

and

$$\mathrm{SO}^+(3,1) = \left\{\Lambda \in \mathbb{M}_4; \Lambda^\intercal \eta \Lambda = \eta, \det \Lambda = +1, \Lambda_{00} > 0\right\}. \tag{14}$$

$\Lambda_{00} > 0$ prevents time reflections, and when combined with $\det \Lambda = 1$ we end up ruling out spatial reflections too. Hence, $\mathrm{SO}^+(3,1)$ is connected.

Restricting our attention to $\mathrm{SO}^+(3,1)$ also has physical motivation. It is well-known that the weak interactions explicitly break parity (spatial reflection) and time reversal symmetries (see, *e.g.*, Schwartz 2014). Hence, even in flat spacetime, $\mathrm{O}(3,1)$ is not a fundamental symmetry of nature, but $\mathrm{SO}^+(3,1)$ is. Of course, the group reflecting the spacetime symmetries is $\mathrm{O}(3,1)$, but mathematical simplicity also allows us to focus on $\mathrm{SO}^+(3,1)$. This is what we will do for the rest of these notes.

Finally, a comment about nomenclature is in place. $\mathrm{SO}(3,1)$ is known as the proper Lorentz group. The group

$$\mathrm{O}^+(3,1) = \left\{\Lambda \in \mathbb{M}_4; \Lambda^\intercal \eta \Lambda = \eta, \Lambda_{00} > 0\right\} \tag{15}$$

is known as the orthochronous Lorentz group, since it forbids time reversals. Finally, $\mathrm{SO}^+(3,1) = \mathrm{SO}(3,1) \cap \mathrm{O}^+(3,1)$ is the proper orthochronous Lorentz group, also called the restricted Lorentz group.

2.3 Poincaré Group

When discussing the Lorentz group we restricted our attention to linear isometries. Apart from mathematical simplicity, there is no reason to make this restriction. Hence, we now consider the full group of isometries of Minkowski spacetime. These are the transformations $x^\mu \to x'^\mu$ such that

$$\eta_{\mu\nu} \frac{\partial x'^\mu}{\partial x^\rho} \frac{\partial x'^\nu}{\partial x^\sigma} = \eta_{\rho\sigma} . \tag{16}$$

Notice this equation means the Jacobian $\frac{\partial x'^\mu}{\partial x^\rho}$ must be a Lorentz transformation. Hence, we have the differential equation

$$\frac{\partial x'^\mu}{\partial x^\rho} = \Lambda^\mu{}_\rho . \tag{17}$$

Integrating this equation leads us to

$$x'^\mu = \Lambda^\mu{}_\rho x^\rho + a^\mu , \tag{18}$$

for an arbitrary vector $a^a \in \mathbb{R}^4$ which appears as an integration constant. If we omit the indices, we can write the resulting transformation as

$$x \to x' = \Lambda x + a. \tag{19}$$

This is known as a Poincaré transformation, and we denote (19) as (Λ, a). Once we restrict attention to $\Lambda \in \mathrm{SO}^+(3,1)$, we get the group

$$\mathrm{ISO}^+(3,1) = \left\{(\Lambda, a); \Lambda \in \mathrm{SO}^+(3,1), a \in \mathbb{R}^4\right\}. \tag{20}$$

This is known as the proper orthochronous Poincaré group, or restricted Poincaré group.

To find the group product of the restricted Poincaré group we perform two Poincaré transformations in sequence. Define

$$\begin{cases} x \to x' = \Lambda x + a, \\ x' \to x'' = \Lambda' x' + a'. \end{cases} \tag{21}$$

Then

$$x \to x'' = \Lambda' \Lambda x + \Lambda' a + a'. \tag{22}$$

Hence,

$$(\Lambda', a') \odot (\Lambda, a) = (\Lambda' \Lambda, \Lambda' a + a'), \tag{23}$$

where $\odot$ denotes the group product.

We have two interesting subgroups of $\mathrm{ISO}^+(3,1)$. Firstly, notice that

$$\left\{(\Lambda, 0); \Lambda \in \mathrm{SO}^+(3,1)\right\} \tag{24}$$

is isomorphic to th restricted Lorentz group $\mathrm{SO}^+(3,1)$. Furthermore,

$$\left\{(\mathbb{1}, a); a \in \mathbb{R}^4\right\} \tag{25}$$

is isomorphic to $(\mathbb{R}^4, +)$, the group of translations. There are some interesting ideas connected to these two subgroups.

Firstly, notice that given any $(\Lambda, b) \in \mathrm{ISO}^+(3,1)$ and $a \in \mathbb{R}^4$ we have that

$$\begin{aligned}(\Lambda^{-1}, -\Lambda^{-1}b) \odot (\mathbb{1}, a) \odot (\Lambda, b) &= (\Lambda^{-1}, -\Lambda^{-1}b) \odot (\Lambda, a + b), &(26a)\\ &= (\Lambda^{-1}\Lambda, \Lambda^{-1}a + \Lambda^{-1}b - \Lambda^{-1}b), &(26b)\\ &= (\mathbb{1}, \Lambda^{-1}a) \in \mathbb{R}^4, &(26c)\end{aligned}$$

where $(\Lambda^{-1}, -\Lambda^{-1}b) = (\Lambda, b)^{-1}$ and the inclusion in the last line implicitly assumes an isomorphism between (25) and $(\mathbb{R}^4, +)$. This means that R^4 is not only a subgroup of $\mathrm{ISO}^+(3,1)$, but a normal subgroup.

Definition 6 [Normal Subgroup]:

Let G be a group and $N \subseteq G$ be a subgroup. We say N is a *normal* subgroup, and write $N \lhd G$, if it holds that for any $n \in N$ and any $g \in G$ we have $g^{-1}ng \in N$. ♠

Normal subgroups are a particularly important class of subgroups, as discussed by Geroch (1985), for example. For our purposes, it is useful to notice the following. Let $g \in G$ (G a group) and $H \subseteq G$ be a subgroup. Then define

$$g^{-1}Hg = \{g^{-1}hg; h \in H\}. \tag{27}$$

This is in general a new subgroup of G that is isomorphic to H. Hence, we get many copies of H inside G. For a normal subgroup, we have a uniqueness property in the sense that all of these copies coincide, for $g^{-1}Ng \subseteq N$ due to the very definition of normal subgroup.

Notice that $\mathrm{SO}^+(3,1)$ is not a normal subgroup. Indeed,

$$\begin{aligned}(\Lambda'^{-1}, -\Lambda'^{-1}a) \odot (\Lambda, 0) \odot (\Lambda', a) &= (\Lambda'^{-1}, -\Lambda'^{-1}a) \odot (\Lambda\Lambda', \Lambda a), &(28a)\\ &= (\Lambda'^{-1}\Lambda\Lambda', \Lambda'^{-1}\Lambda a - \Lambda'^{-1}a), &(28b)\end{aligned}$$

and this is not in general an element of (24).

It is also interesting to notice that if we understand $\mathrm{SO}^+(3,1)$ and $\mathbb{R}^4$ as the subgroups given on (24) and (25), then $\mathrm{SO}^+(3,1) \cap \mathbb{R}^4 = \{(\mathbb{1}, 0)\}$, which is the subgroup composed solely by the neutral element. Furthermore, notice that any element of $\mathrm{ISO}^+(3,1)$ can be written in the form

$$(\Lambda, a) = (\Lambda, 0) \odot (\mathbb{1}, \Lambda^{-1}a). \tag{29}$$

We thus write $\mathrm{ISO}^+(3,1) = \mathrm{SO}^+(3,1)\mathbb{R}^4$, meaning that every Poincaré transformation is a translation followed by Lorentz transformation. Due to $\mathbb{R}^4$ being normal, this can also be rewritten as $\mathrm{ISO}^+(3,1) = \mathbb{R}^4 \odot \mathrm{SO}^+(3,1)$ with

$$(\Lambda, a) = (\mathbb{1}, a) \odot (\Lambda, 0). \tag{30}$$

This sort of structure has a particular name in group theory.

Definition 7 [Semidirect Product]:

Let G be a group with neutral element e, $H, N \subseteq G$ subgroups, and $N \lhd G$. If

$$G = HN = \{hn; h \in H, n \in N\} \tag{31}$$

and $H \cap N = \{e\}$, then we say G is the semidirect product of N and H and write $G = H \ltimes N$. ♠

The notation $H \ltimes N$ makes reference to the facts that, as a set, $G = H \times N$ and, as a group, $N \lhd G$.

As discussed by Robinson (1996), for example, semidirect products can also be defined for two given groups and used as a way of building a third new group. For us, this internal definition will be enough.

3 Symmetries in Curved Spacetimes

Now that we know how to characterize symmetries in terms of groups, the next step should be to adapt this discussion to curved spacetimes. This will require us to develop some new geometric language in order to appropriately formulate what constitutes a symmetry in a curved spacetime. Furthermore, we will be able to provide a geometric, coordinate-free construction of the isometry groups we previously discussed. Parts of this section are inspired by the book by Wald (1984, App. C).

This section will rely considerably on a basic understanding of differential geometry as required for general relativity. Examples of textbooks on differential geometry are the ones by Lee (2012) and Tu (2011), but basic differential geometry is also reviewed in general relativity books such as the ones by Hawking and Ellis (1973) and Wald (1984).

3.1 Pullbacks and Pushforwards

We now take an active point of view on diffeomorphisms. Rather than thinking about them as mere changes of coordinates, we will think of them as active transformations that take one manifold into another one and discuss how the structures on these manifolds transform. To understand this, we will have to discuss some properties about smooth mappings between manifolds.

Let M and N be manifolds. Given a point $p \in M$, we denote the space of tangent vectors to M at p by T_pM. Suppose we are given a smooth mapping $\phi \colon M \to N$. This mapping gives us some structure to relate M and N. Which sorts of relations can we derive from this?

Suppose first that we are given a smooth function $f \colon N \to \mathbb{R}$. We can "pullback" f using ϕ by defining a new map $\phi_* f \colon M \to \mathbb{R}$ through $\phi_* f = f \circ \phi$. This can be depicted in the diagram

$$\begin{array}{ccc} M & \xrightarrow{\phi} & N \\ & \underset{\phi_* f}{\searrow} & \downarrow f \\ & & \mathbb{R} \end{array}$$

We call $\phi_* f$ the pullback of f through ϕ. This name comes from the fact that ϕ is "pulling back" the function f from N to M.

Next suppose we have a vector $v \in T_pM$ (we sometimes omit the abstract indices to avoid cluttering the notation in what follows). While we could pull back a function, we can "push forward" this vector. Recall that a vector at p is a linear map $v \colon \mathscr{C}^\infty(p) \to \mathbb{R}$ ($\mathscr{C}^\infty(p)$ being the space of functions which are smooth at p) satisfying certain conditions so that it behaves as a directional derivative (see, *e.g.*, Lee 2012; Tu 2011). We can define a new vector $\phi^* v \in T_{\phi(p)}N$ by imposing that

$$\phi^* v(f) = v(\phi_* f) = v(f \circ \phi) \tag{32}$$

for every $f \in \mathscr{C}^\infty(\phi(p))$. Diagrammatically this can be written as

$$\begin{array}{ccc} \mathscr{C}^\infty(p) & \xleftarrow{\ \phi_*\ } & \mathscr{C}^\infty(\phi(p)) \\ {\scriptstyle v}\big\downarrow & \swarrow {\scriptstyle \phi^* v} & \\ \mathbb{R} & & \end{array}$$

where ϕ_* is the pullback map defined as $\phi_* f = f \circ \phi$.

Notice that the pushforward ϕ^* is a linear map $\phi^* : T_pM \to T_{\phi(p)}N$ and can thus be interpreted as the derivative of ϕ at p in the usual sense of multivariable calculus (although it is now generalized to manifolds). The pushforward of a smooth mapping is also known as the differential or the tangent map.

This process does not stop at the level of vectors. Let $\omega \in T^*_{\phi(p)}N$ be a covector at $\phi(p) \in N$. Recall that a covector in $T^*_{\phi(p)}N$ is a linear map $\omega\colon T_{\phi(p)}N \to \mathbb{R}$. We can pull back this covector to a new covector $\phi_*\omega \in T^*_pM$ by defining

$$\phi * \omega(v) = \omega(\phi^* v) \tag{33}$$

for every $v \in T_pM$. We can write this in the diagram

$$\begin{array}{ccc} T_pM & \xrightarrow{\ \phi^*\ } & T_{\phi(p)}N \\ & {\scriptstyle \phi_*\omega} \searrow & \big\downarrow {\scriptstyle \omega} \\ & & \mathbb{R} \end{array}$$

where $\phi^* : T_pM \to T_{\phi(p)}N$ is the pushforward.

Given a tensor T of type $(0, l)$ at p, we can also pull-it-back by defining ϕ_*T through

$$\phi_*T(v_1, \dots, v_l) = T(\phi^* v_1, \dots, \phi^* v_l). \tag{34}$$

Analogously, for a tensor T of type $(k, 0)$ we define the pushforward through

$$\phi^*T(\omega_1, \dots, \omega_k) = T(\phi_*\omega_1, \dots, \phi_*\omega_k). \tag{35}$$

This is consistent with our previous definition of pushforward of a vector due to the definition of pullback of a covector.

Notice we thus get a number of relations between different structures defined on M and N. Nevertheless, we have quite some restrictions. We cannot push forward nor pull back a tensor of mixed type, since we do not know how to pull back contravariant tensors and we do not know how to push forward covariant tensors.

Suppose, however, that $\phi\colon M \to N$ is not only smooth, but actually a diffeomorphism. We recall this means that ϕ is smooth, bijective, and has a smooth inverse. In particular this implies that $\dim M = \dim N$. In this case, we have the inverse mapping ϕ^{-1} and we can exploit it to define the pullbacks and pushforwards that ϕ cannot handle. Given a tensor T of type (k, l) at $p \in M$, we define

$$\phi^*T(\omega_1, \dots, \omega_k, v_1, \dots, v_l) = T(\phi_*\omega_1, \dots, \phi_*\omega_k, (\phi^{-1})_*v_1, \dots, (\phi^{-1})_*v_l). \tag{36}$$

The pullback ϕ_* is defined analogously, but since it holds that $\phi_* = (\phi^{-1})^*$ it suffices to work with the pushforward.

If we consider a diffeomorphism $\phi \colon M \to M$, we get an interesting structure. If we are giving a tensor field T, we can now consider the pushforward of the entire field, $\phi^* T$. This leads us to a new tensor field that can be compared to the original tensor field T. In general, these two tensor fields will not be the same, even though they are related by a diffeomorphism. In coordinate-parlance, the coordinate components of T are preserved, but they are taken to $\phi(p)$ instead of being kept at p. This is an ambiguity in the description, and hence it is understood as a gauge symmetry. We get physical symmetries in the particular case in which $\phi^* T = T$, so that acting on the spacetime with a diffeomorphism keeps everything unchanged—this is not a mere redundancy, but an actual symmetry of the spacetime. The diffeomorphisms that keep the metric invariant, *i.e.*, the diffeomorphisms with $\phi^* g = g$, are called isometries.

3.2 Lie Derivatives

An interesting structure we can consider when dealing with diffeomorphisms is a one-parameter group of diffeomorphisms. This is a map $\phi \colon \mathbb{R} \times M \to M$ such that $\phi_t \colon M \to M$ is a diffeomorphism for every $t \in \mathbb{R}$and the map $t \to \phi_t$ is such that $\phi_t \circ \phi_s = \phi_{t+s}$. Notice then that $\{\phi_t\}_{t \in \mathbb{R}}$ then has a natural group structure.

Take a point $p \in M$ and consider a one-parameter group of diffeomorphisms in M. Notice that $\gamma_p(t) = \phi_t(p)$ defines a curve in M—this is called an orbit of ϕ_t. If we differentiate γ_p at $t = 0$, we get a vector $v_p \in T_p M$. Through this process we get a vector field v associated to ϕ_t which is everywhere parallel to the orbits of ϕ_t. Inversely, given a smooth vector field v, it is always possible to find a one-parameter group of diffeomorphisms in M whose orbits are all parallel to v. This is said to be the flow of v. For details, see the books by Lee (2012, Chap. 9), Tu (2011, Chap. 14), and Wald (1984, Sec. 2.2). We should mention that the one-parameter group of diffeomorphisms might not be defined for all parameter values $t \in \mathbb{R}$, but rather be defined only on a smaller interval. In any case, it is always possible to find a local flow.

Suppose now we want to consider how a given tensor field T changes in spacetime. It is, of course, useful to have a derivative of T to analyze this variation. We would like to compute something like

$$\left.\frac{\mathrm{d}T}{\mathrm{d}t}\right|_p \sim \lim_{t \to 0} \frac{T(p+t) - T(p)}{t}. \tag{37}$$

Of course, this equation makes no sense. There is no vector space structure on the manifold for $p + t$ to make sense and we cannot compare tensors at different points of the manifold. However, if we are given a one-parameter group of diffeomorphisms $\{\phi_t\}$ with tangent vector field v (or, equivalently, if we are given the vector field and consider the one-parameter group of diffeomorphisms) we can define

$$\pounds_v T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} = \lim_{t \to 0} \frac{\phi^*_{-t} T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} - T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l}}{t}. \tag{38}$$

Notice that $\phi^*_{-t} T = \phi_{t*} T$ picks the tensor field at "$p + t$" and pulls it back to p. Hence, Eq. (38) gives a precise mathematical meaning to Eq. (37). $\pounds_v$ is known as the Lie derivative with respect to v^a. Notice it is a linear map that takes smooth (k, l)-tensor fields to smooth (k, l)-tensor fields. The Lie derivative preserves contractions and it can be shown that the Lie derivative obeys the Leibnitz rule on tensor products,

$$\pounds_v (S \otimes T) = (\pounds_v S) \otimes T + S \otimes (\pounds_v T). \tag{39}$$

Notice that $£_v T = 0$ everywhere if, and only if, ϕ_t is a symmetry of T for all T, *i.e.*, $\phi_t^* T = T$ for all t.

At a point p, notice that

$$\left(£_v T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l}\right)_p = \lim_{t \to 0} \frac{\left(\phi_{-t}^* T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l}\right)_p - \left(T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l}\right)_p}{t} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\phi_{-t}^* T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l}\right)_p. \tag{40}$$

In particular, consider the case of a scalar function. Then

$$(£_v f)_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\phi_{-t}^* f)_p, \tag{41a}$$

$$= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\phi_{t*} f)_p, \tag{41b}$$

$$= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (f \circ \phi_t)_p, \tag{41c}$$

$$= v(f)_p, \tag{41d}$$

where the last step can be performed by choosing a coordinate system and employing the chain rule. We thus learn that

$$£_v f = v(f). \tag{42}$$

Now that we know how the Lie derivative acts on functions, the next simplest step is to learn how it acts on vector fields. To find that, introduce a coordinate system such that the parameter t along the integral lines of v^a is one of the coordinates, x^1. In this manner, we have that

$$v^a = \left(\frac{\partial}{\partial x^1}\right)^a. \tag{43}$$

This corresponds to choosing the function x^1 such that $v(x^1) = 1$, which clearly can always be done in neighborhoods in which v^a does not vanish.

In such a coordinate system, acting with ϕ_{-t} is equivalent to performing the coordinate transformation $x^1 \to x^1 + t$ while holding the remaining coordinates fixed. The matrix components of the pushforward $\phi_{-t}^* : T_p M \to T_{\phi_{-t}(p)} M$ are then given in the coordinate basis by

$$(\phi_{-t}^*)^\mu{}_\nu = \delta^\mu{}_\nu. \tag{44}$$

Hence, this means that

$$(\phi_{-t}^* T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l})(x^1, \ldots, x^n) = T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l}(x^1 + t, \ldots, x^n). \tag{45}$$

Using this in the expression for the Lie derivative, Eq. (38) on the previous page, we conclude that in this coordinate system

$$£_v T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} = \frac{\partial T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l}}{\partial x^1}. \tag{46}$$

Notice that using this expression it is particularly simple to prove the validity of the Leibniz rule.

Using this general expression, we see that given a vector field w^a we have, in this coordinate system adapted to v^a,

$$\pounds_v w^\mu = \frac{\partial w^\mu}{\partial x^1}. \tag{47}$$

Similarly, we can compute the commutator[1] $[v, w]^\mu$ and find that

$$[v, w]^\mu = \frac{\partial w^\mu}{\partial x^1} \tag{48}$$

too. Hence, we conclude that in this coordinate system we have

$$\pounds_v w^\mu = [v, w]^\mu. \tag{49}$$

Since this equation is covariant and both quantities are defined in a coordinate-independent manner, we can conclude that

$$\pounds_v w^a = [v, w]^a \tag{50}$$

as a tensor equality.

Having the expressions for the Lie derivatives of scalars and vector fields, we can derive all other cases. These formulae are more easily expressed in terms of derivative operators, so that is what we will do. As an example, let us consider the Lie derivative for one-forms. Given some one-form μ_a and a vector field w^a, we have that

$$\pounds_v (\mu_a w^a) = v(\mu_a w^a), \tag{51a}$$
$$= v^b w^a \nabla_b \mu_a + v^b \mu_a \nabla_b w^a. \tag{51b}$$

This follows from using the expression of the Lie derivative for scalar functions. However, if we instead used the Leibnitz rule and the expression for the Lie derivative of a vector field we would have found

$$\pounds_v (\mu_a w^a) = w^a \pounds_v \mu_a + \mu_a \pounds_v w^a, \tag{52a}$$
$$= w^a \pounds_v \mu_a + \mu_a [v, w]^a, \tag{52b}$$
$$= w^a \pounds_v \mu_a + \mu_a v^b \nabla_b w^a - \mu_a w^b \nabla_b v^a. \tag{52c}$$

Bringing everything together and solving for $w^a \pounds_v \mu_a$ we find that

$$w^a \pounds_v \mu_a = v^b w^a \nabla_b \mu_a + \mu_a w^b \nabla_b v^a \tag{53}$$

for all w^a, and therefore it follows that

$$\pounds_v \mu_a = v^b \nabla_b \mu_a + \mu_b \nabla_a v^b. \tag{54}$$

The general expression for an arbitrary tensor field can be obtained by induction. One finds that

$$\pounds_v T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} = v^c \nabla_c T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} - \sum_{i=1}^{k} T^{a_1 \cdots c \cdots a_k}{}_{b_1 \cdots b_l} \nabla_c v^{a_i} + \sum_{j=1}^{l} T^{a_1 \cdots a_k}{}_{b_1 \cdots c \cdots b_l} \nabla_{b_j} v^c. \tag{55}$$

This expression holds for *any* derivative operator (see Wald 1984, Sec. 3.1), not only for the Levi-Civita connection.

[1] Recall that the commutator of two vector fields v^a and w^a is the vector field $[v, w]^a$ such that $[v, w](f) = v(w(f)) - w(v(f))$. Using a derivative operator, we can write $[v, w]^a = v^b \nabla_b w^a - w^b \nabla_b v^a$.

3.3 Conformal Killing Vector Fields

At this stage, we are ready to discuss an important class of concepts: Killing vector fields and conformal Killing vector fields.

We begin by giving meaning to the term "conformal". A conformal transformation is a transformation that, in some sense, preserves angles. Within differential geometry and general relativity, a conformal transformation is a smooth map such that $\phi_* g_{ab} = \Omega^2 g_{ab} = \tilde{g}_{ab}$ for some (smooth) function $\Omega > 0$. Notice that angles are preserved in a specific sense: $g_{ab} k^a k^b = 0$ if, and only if, $\tilde{g}_{ab} k^a k^b = 0$. Therefore, under a conformal transformation of the metric, null vectors are preserved and, as a consequence, so are the lightcones. This means therefore that the causal structure of spacetime is preserved.

We previously mentioned that a diffeomorphism $\phi\colon M \to M$ such that $\phi^* g_{ab} = 0$ is called an isometry. There is a second interesting class of diffeomorphisms that "almost preserve" the metric, known as conformal isometries. There are the diffeomorphisms such that $\phi^* g_{ab} = \Omega^2 g_{ab}$ for some function $\Omega > 0$. Hence, a conformal isometry is a diffeomorphism that also happens to be a conformal transformation.

Suppose now we are given a one-parameter group of conformal isometries (regular isometries then become a particular case). We surely must have that

$$\pounds_v g_{ab} = \lambda g_{ab} \tag{56}$$

for some function λ yet to be determined (and that should vanish for a regular isometry). Using Eq. (55) on the preceding page for a derivative operator with $\nabla_a g_{bc} = 0$, we find that

$$\pounds_v g_{ab} = \nabla_a v_b + \nabla_b v_a \tag{57}$$

and Eq. (56) becomes

$$\nabla_a v_b + \nabla_b v_a = \lambda g_{ab}. \tag{58}$$

For general λ, this is known as the conformal Killing equation. For $\lambda = 0$ (corresponding to regular isometries rather than conformal isometries) this is known as the Killing equation.

Assuming spacetime is n-dimensional and contracting both sides of the conformal Killing equation with g^{ab} leads to

$$2\nabla_a v^a = n\lambda, \tag{59}$$

which establishes the value of λ. Hence, the conformal Killing equation becomes

$$\nabla_a v_b + \nabla_b v_a = \frac{2}{n}(\nabla_c v^c) g_{ab}. \tag{60}$$

Notice that Killing vector fields are infinitesimal generators of isometries, so they capture—in an infinitesimal sense—what are the symmetries of the metric. Conformal Killing vector fields are a bit more generous and consider the conformal symmetries of the metric.

In a curved spacetime, symmetries correspond to the integral lines of complete Killing vector fields, where "complete" means that their integral lines (*i.e.*, their flow) is defined for all times. The symmetry group of the spacetime is a subgroup of the group composed by all diffeomorphisms, Diff(M) (the group product is the composition of maps). As an example, we say that a spacetime is spherically symmetric when the group of isometries has a subgroup isomorphic to SO(3).

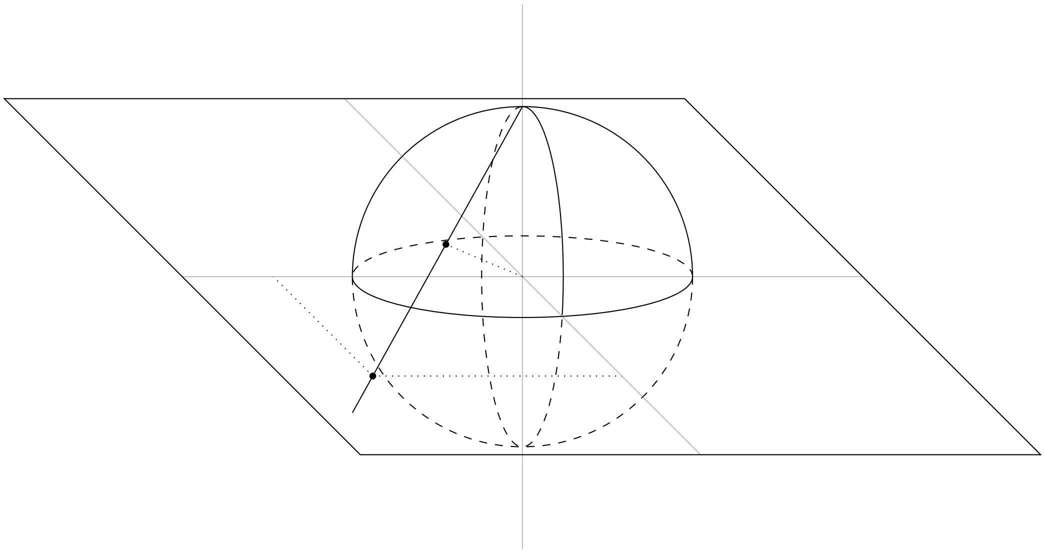


Figure 3.1: *Diagram for construction of stereographic coordinates. One picks a point in the sphere and connects it to the north pole. One follows the line determined in this manner until it intercepts the xy-plane, at which point one records the coordinates x and y of the intersection. The coordinates attributed to the point in the sphere are then* $\zeta = x + iy$ *and its conjugate* $\bar{\zeta}$*. This corresponds to* $\zeta = e^{i\phi}\cot\left(\frac{\theta}{2}\right)$.

3.4 Case Study: Conformal Isometries on a Sphere

As a case study, let us consider what are the isometries and the conformal isometries on the two-sphere $\mathbb{S}^2$. This will be a convenient discussion for later.

To perform the calculations, it will be useful to employ some coordinate system. Spherical coordinates[2] (θ, ϕ) are one of the options, but we will employ a different choice. Namely, we choose to work with stereographic coordinates. This corresponds to defining a complex coordinate $\zeta = e^{i\phi}\cot\left(\frac{\theta}{2}\right)$. Geometrically this construction is illustrated on Fig. 3.1. One draws the sphere as the unit sphere in three dimensions and traces lines from the north pole to different points of the sphere. Given a point on the sphere, the line from it to the north pole can be extended untill it crosses the xy-plane, marking the coordinate $\zeta = x + iy$ at the crossing point. In this manner, each point in the sphere is mapped to a point in the plane, and the north pole is mapped to infinity. It is worth mentioning that this is a way of viewing the sphere as the Riemann sphere $\mathbb{C} \cup \{\infty\}$.

The round metric on the sphere is given by

$$\mathrm{d}\mathbb{S}^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2 . \tag{61}$$

In stereographic coordinates, we get

$$\mathrm{d}\mathbb{S}^2 = 2\gamma_{\zeta\bar{\zeta}}\,\mathrm{d}\zeta\,\mathrm{d}\bar{\zeta}, \tag{62a}$$

[2] In our conventions, θ is the polar angle and ϕ is the azimuthal angle.

with

$$\gamma_{\zeta\bar{\zeta}} = \frac{2}{(1+\zeta\bar{\zeta})^2}. \tag{62b}$$

The Christoffel symbols for this metric in stereographic coordinates are

$$\Gamma^{\zeta}{}_{\zeta\zeta} = -\frac{2\bar{\zeta}}{1+\zeta\bar{\zeta}} \quad \text{and} \quad \Gamma^{\bar{\zeta}}{}_{\bar{\zeta}\bar{\zeta}} = -\frac{2\zeta}{1+\zeta\bar{\zeta}}. \tag{63}$$

Notice that for a vector field Y^a to be real, it is necessary that $Y^{\bar{\zeta}} = \overline{Y^{\zeta}}$.

Let us then consider the conformal Killing equation. A vector field Y^a is a conformal Killing vector field on the sphere if it satisfies

$$\nabla_a Y_b + \nabla_b Y_a = \nabla_c Y^c \gamma_{ab}, \tag{64}$$

where γ_{ab} is the round metric. It is convenient to write this in the form

$$\gamma_{bc} \nabla_a Y^c + \gamma_{ac} \nabla_b Y^c = \nabla_c Y^c \gamma_{ab}. \tag{65}$$

In components, this expression is given by

$$\partial_{\zeta} Y^{\bar{\zeta}} = 0, \tag{66a}$$

$$\partial_{\bar{\zeta}} Y^{\zeta} = 0, \tag{66b}$$

$$\partial_{\bar{\zeta}} Y^{\bar{\zeta}} + \Gamma^{\bar{\zeta}}{}_{\bar{\zeta}\bar{\zeta}} Y^{\bar{\zeta}} + \partial_{\zeta} Y^{\zeta} + \Gamma^{\zeta}{}_{\zeta\zeta} Y^{\zeta} = \nabla_c Y^c. \tag{66c}$$

Eq. (66c) is trivial. The derivatives in Eqs. (66a) and (66b) should be understood as Wirtinger derivatives, which for $\zeta = x + iy$ are defined as the linear operators

$$\frac{\partial}{\partial \zeta} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \quad \text{and} \quad \frac{\partial}{\partial \bar{\zeta}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right). \tag{67}$$

As a consequence, the equation

$$\frac{\partial f}{\partial \bar{\zeta}} = 0 \tag{68}$$

is equivalent to the Cauchy–Riemann equations, and thus it means that f is an analytic function of ζ. Hence, Eqs. (66a) and (66b) imply that Y^{ζ} ($Y^{\bar{\zeta}}$) is an analytic function of ζ ($\bar{\zeta}$).

Hence, the local Killing fields are characterized by linear combinations of fields with the form $Y^{\zeta} = \alpha \zeta^n$ for $\alpha \in \mathbb{C}$ and $n \geq 0$ an integer. There is, however, a caveat: some of these fields may not be global Killing vector fields (as we would need for a complete Killing vector field). This is because the stereographic coordinate we are working with is only defined away from the north pole, since $\zeta = \infty$ *at* the north pole. To avoid this, we define a new stereographic coordinate ξ through $\xi = -e^{i\phi} \tan\left(\frac{\theta}{2}\right)$. ξ is then the coordinate antipodally related to ζ. These two coordinate systems are related by

$$\xi = -\frac{1}{\bar{\zeta}}, \tag{69}$$

with an analogous expression for the conjugates. The line element in the ξ coordinate is identical to the expression in terms of the ζ coordinate, and hence so are the results of the Killing equation: Y^ξ must be an analytic function of ξ.

Consider now the vector field defined by $Y^\zeta = \alpha\zeta^n$. When we change coordinates from ζ to ξ, we find that

$$Y^\xi = \frac{(-1)^n\bar{\alpha}}{\xi^{n-2}}. \tag{70}$$

Since $n \geq 0$, this is only an analytic function of ξ for $n \leq 2$.

Therefore, we find that there is a six-dimensional real space of possible conformal Killing vector fields, which is spanned by $Y^\zeta = 1, \zeta, \zeta^2, i, i\zeta, i\zeta^2$.

We then have a new question: are these vector fields complete?

Let us consider some curve that is everywhere parallel to the most general vector field with

$$Y^\zeta = \alpha + \beta\zeta + \gamma\zeta^2 \tag{71}$$

for $\alpha, \beta, \gamma \in \mathbb{C}$. This curve is characterized by the differential equation

$$\frac{\mathrm{d}\zeta}{\mathrm{d}t} = \alpha + \beta\zeta + \gamma\zeta^2 \tag{72}$$

The full solution has the form

$$\zeta(t) = \frac{a(t)\zeta(0) + b(t)}{c(t)\zeta(0) + d(t)}, \tag{73}$$

where the specific functional form of $a(t), b(t), c(t)$, and $d(t)$ depends on the values of α, β, γ, and the general case is particularly complicated. One can choose their normalization such that $ad - bc = 1$. Such a transformation is known as a Möbius transformation, and these transformations correspond precisely to the conformal transformations on the Riemann sphere.

Notice that a Möbius transformation

$$\zeta \to \frac{a\zeta + b}{c\zeta + d} \tag{74}$$

with $ad - bc = 1$ has precisely one pole at $\zeta = -\frac{d}{c}$ and one zero at $\zeta = -\frac{b}{c}$. They are analytic in the complex plane and are a diffeomorphism of the Riemann sphere onto itself.

The (non-conformal) Killing vector fields on the sphere are the conformal Killing vector fields which satisfy the extra condition $\nabla_a Y^a = 0$. Imposing this condition on a vector field with the form (71) we conclude that a (non-conformal) Killing vector field has the form (71) with the extra conditions that $\beta + \bar{\beta} = 0$ and $\gamma = \bar{\alpha}$. These are three real constraints, so we get a three-dimensional real space of Killing vector fields.

What is the group of conformal isometries on the sphere? To answer this question, we must study how two Möbius transformations compose. Consider two consecutive Möbius transformations

$$\zeta \to \zeta' = \frac{a\zeta + b}{c\zeta + d} \quad \text{and} \quad \zeta' \to \zeta'' = \frac{a'\zeta' + b'}{c'\zeta' + d'}. \tag{75}$$

We want to express the coefficients of the resulting transformation

$$\zeta \to \zeta'' = \frac{a''\zeta + b''}{c''\zeta + d''} \tag{76}$$

in terms of the original coefficients. We assume that

$$ad - bc = a'd' - b'c' = 1. \tag{77}$$

One can show that

$$\begin{pmatrix} a'' & b'' \\ c'' & d'' \end{pmatrix} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{78}$$

with

$$a''d'' - b''c'' = 1. \tag{79}$$

This is the product of two 2×2 complex matrices with unit determinant. Furthermore, notice that the Möbius transformation corresponding to the matrix A is the same Möbius transformation corresponding to the matrix $-A$. Hence, the group is the group $SL(2, \mathbb{C})/\mathbb{Z}_2$: the group of 2×2 complex matrices with unit determinant up to sign. This may seem complicated, but it turns out that

$$SL(2, \mathbb{C})/\mathbb{Z}_2 \cong SO^+(3, 1). \tag{80}$$

Hence, the group of conformal isometries on the two-sphere is just the restricted Lorentz group.

From our analysis it is difficult to see this, but it turns out that the group of isometries on the two-sphere is SO(3), as one would expect. This can be more easily seem through the language of Lie algebras.

3.5 Case Study: Killing Isometries on Minkowski Spacetime

It is instructive to consider as a second case study the isometries in Minkowski spacetime. We should, of course, recover the Poincaré transformations. While this time we are dealing only with isometries, we are now in a four-dimensional manifold (as compared to a two-dimensional manifold), so the system of differential equations becomes more difficult to solve.

The Killing equation is

$$\nabla_a \xi_b + \nabla_b \xi_a = 0. \tag{81}$$

In globally inertial Cartesian coordinates this expression becomes

$$\partial_\mu \xi_\nu + \partial_\nu \xi_\mu = 0. \tag{82}$$

Consider Eq. (82) for $\mu = \nu = t$. This yields

$$\partial_t \xi_t = 0, \tag{83}$$

and thus ξ_t is time-independent. Picking $\mu = t$ and $\nu = x$, for example, leads to

$$\partial_t \xi_x = -\partial_x \xi_t . \tag{84}$$

Since the right-hand side is time-independent, so must be the left-hand side. Hence, ξ_x must be at most linear in t. More generally, ξ_μ can depend at most linearly on x^ν and it is independent of x^μ. Can there be crossed terms, *i.e.*, terms of the form yz, for example?

Differentiate Eq. (82) on the previous page and apply it a few times to see that

$$\begin{aligned}
\partial_\rho \partial_\mu \xi_\nu &= -\partial_\rho \partial_\nu \xi_\mu , && (85a)\\
&= -\partial_\nu \partial_\rho \xi_\mu , && (85b)\\
&= +\partial_\nu \partial_\mu \xi_\rho , && (85c)\\
&= +\partial_\mu \partial_\nu \xi_\rho , && (85d)\\
&= -\partial_\mu \partial_\rho \xi_\nu , && (85e)\\
&= -\partial_\rho \partial_\mu \xi_\nu , && (85f)
\end{aligned}$$

and thus

$$\partial_\rho \partial_\mu \xi_\nu = 0, \tag{86}$$

meaning crossed terms are not allowed.

Taking all of this into consideration, we have that

$$\xi_t = a + bx + cy + dz, \tag{87}$$

for $a, b, c, d \in \mathbb{R}$, with similar expressions for the other components. Imposing Eq. (82) on the preceding page for this ansatz leads to relations between the constant coefficients of the different components ξ_μ. At the end of the say, one gets the ten-dimensional vector space spanned by the Killing vectors given by

$$P_t^a = \left(\frac{\partial}{\partial t}\right)^a, \quad P_x^a = \left(\frac{\partial}{\partial x}\right)^a, \quad P_y^a = \left(\frac{\partial}{\partial y}\right)^a, \quad P_z^a = \left(\frac{\partial}{\partial z}\right)^a, \tag{88a}$$

$$J_x^a = -z\left(\frac{\partial}{\partial y}\right)^a + y\left(\frac{\partial}{\partial z}\right)^a, \quad J_y^a = -x\left(\frac{\partial}{\partial z}\right)^a + z\left(\frac{\partial}{\partial x}\right)^a, \quad J_z^a = -y\left(\frac{\partial}{\partial x}\right)^a + x\left(\frac{\partial}{\partial y}\right)^a, \tag{88b}$$

$$K_x^a = x\left(\frac{\partial}{\partial t}\right)^a + t\left(\frac{\partial}{\partial x}\right)^a, \quad K_y^a = y\left(\frac{\partial}{\partial t}\right)^a + t\left(\frac{\partial}{\partial y}\right)^a, \quad K_z^a = z\left(\frac{\partial}{\partial t}\right)^a + t\left(\frac{\partial}{\partial z}\right)^a. \tag{88c}$$

Using methods we do not cover in these notes (namely, the theory of Lie algebras) one can show that the diffeomorphisms generated by these ten vector fields constitute precisely the Poincaré group $\mathrm{ISO}^+(3,1)$. The vectors P_μ^a generate translations, J_i^a generate rotations, and K_i^a generate boosts.

4 Asymptotically Flat Spacetimes

Our goal in these lectures is to understand the asymptotic symmetries of asymptotically flat spacetimes. For that end, we must have a good understanding of what we mean by "asymptotic" and what we mean by "asymptotically flat". The goal of this section is to give attention to these terms and discuss how they are currently understood in general relativity.

4.1 Infinity in Minkowski Spacetime

We shall begin by giving meaning to "infinity" in Minkowski spacetime. Our approach is similar in nature to the one by Wald (1984, Chap. 11).

To begin our discussion, let us choose a system of coordinates to better visualize what we are doing. Spherical coordinates turn out to be convenient for our discussion, so we write the Minkowski line element as

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}r^2 + r^2\,\mathrm{d}\mathbb{S}^2\,, \tag{89}$$

where $\mathrm{d}\mathbb{S}^2$ stands for the round metric on the sphere.

When we talk about what happens "at infinity" in Minkowski spacetime, there are three possibilities we could be considering. Namely, we could be thinking about taking the limit $t \to \pm\infty$ or taking the limit $r \to +\infty$. This, however, is not a clear enough picture of what infinity is in Minkowski spacetime. For instance, we do not know if taking $r \to +\infty$ toward different directions will lead to the same result. Furthermore, we could take $r \to +\infty$ while holding t constant or while holding $u = t - r$ constant, for example, and this turns out to lead to very different results. Therefore, we must be more careful.

For concreteness, let us assume we want to take the limit $r \to +\infty$ at constant $u = t - r$. In this case, it makes sense to change coordinates so we use u rather than t when writing the line element. Doing so leads us to the line element

$$\mathrm{d}s^2 = -\,\mathrm{d}u^2 - 2\,\mathrm{d}u\,\mathrm{d}r + r^2\,\mathrm{d}\mathbb{S}^2\,. \tag{90}$$

At this point we could attempt to take the limit $r \to +\infty$ in order to understand what is the geometry of infinity in Minkowski spacetime. Nevertheless, there is a difficulty in doing so: this limit leads to a divergence in the metric. This divergence is not due to a bad choice of coordinates, because it has geometrical (and hence physical) implications. Namely, the area of a sphere with radius r diverges as we take $r \to +\infty$. Hence, it is not possible to take this limit for this metric.

Despite this difficulty, we would still like to understand the structure of infinity. To do so, we will make a new change of coordinates. We define a new coordinate $l = \frac{1}{r}$, which turns the line element into

$$\mathrm{d}s^2 = -\,\mathrm{d}u^2 + 2l^{-2}\,\mathrm{d}u\,\mathrm{d}l + l^{-2}\,\mathrm{d}\mathbb{S}^2\,. \tag{91}$$

A few comments are in order. Firstly, notice that the physical region of Minkowski spacetime corresponds to $l > 0$—the limiting case $l = 0$ would in principle be infinity (in the sense that $r \to +\infty$ as $l \to 0^+$) and $l < 0$ is non-physical. Secondly, this coordinate patch does not cover the points with $r = 0$, since l diverges there. Thirdly, notice that this was merely a coordinate change, and taking the limit $l \to 0^+$ is not any more possible than taking the limit $r \to +\infty$ was before.

Nevertheless, this new coordinate system seems to be "better behaved" at infinity. Namely, we know infinity sits at $l = 0$, which we can barely grasp. The only problem we face now is the presence of the l^{-2} factors in Eq. (91). To deal with them, we will take an apparently *ad hoc* approach. We will define a new, unphysical metric $\tilde{g}_{ab}$ through

$$\tilde{g}_{ab} = l^2 \eta_{ab}\,, \tag{92}$$

where η_{ab} is the Minkowski metric. The line element for this new unphysical metric is now

$$\mathrm{d}\tilde{s}^2 = -l^2\,\mathrm{d}u^2 + 2\,\mathrm{d}u\,\mathrm{d}l + \mathrm{d}\mathbb{S}^2\,. \tag{93}$$

This new unphysical metric has an advantage: it is well-behaved in the $l \to 0^+$ limit. We can thus describe the geometry of infinity by plugging in $l = 0$ and getting the induced metric

$$\mathrm{d}\tilde{\sigma}^2 \,|_{l=0} = 0\,\mathrm{d}u^2 + \mathrm{d}\mathbb{S}^2 \,, \tag{94}$$

where we write $0\,\mathrm{d}u^2$ explicitly as a reminder that this is a metric in a three-dimensional manifold. We write $\mathrm{d}\sigma^2$ instead of $\mathrm{d}s^2$ to make explicit that this is the induced metric. We can now consider the behavior of tensor fields at infinity by perfoming suitable transformations of tensor fields on Minkowski spacetime in order to map them from the original spacetime with metric η_{ab} to the unphysical spacetime with metric $\tilde{g}_{ab}$ and subsequently evaluating them at $l = 0$. In this sense, infinity becomes a set of regular points.

Let us summarize what is the trick we just employed and notice some of the ideas that went in performing it. We had an issue with the metric η_{ab} because it diverged in the region we wanted to consider—namely, infinity. Nevertheless, we found a function l that conveniently vanishes precisely at the region in which the metric diverges. We then noticed that the metric $\tilde{g}_{ab} = l^2\eta_{ab}$ is finite in the region we wanted to study. Hence, we decided to let go of η_{ab} and work with $\tilde{g}_{ab}$ instead, at least while we are considering the behavior at infinity. By making appropriate transformations of tensor fields so they can be understood as being defined in the spacetime given by $\tilde{g}_{ab}$, we can analyze their behavior at infinity by simply evaluating them at infinity, which became a set of regular points in the new unphysical spacetime.

It is important to emphasize that $\tilde{g}_{ab}$ is an unphysical metric. It is not the metric that describes the spacetime geometry as measured in experiments. It is defined as a mathematical construction that allows us to "bring in infinity" and consider it as a regular point. In general, $\tilde{g}_{ab}$ will not even satisfy the Einstein field equations, so it is not an accurate model of the bulk spacetime[3] we are trying to describe. Nevertheless, the price we pay to study infinity is that we are forced to work with $\tilde{g}_{ab}$ rather than with the physical metric η_{ab}.

This procedure is called a conformal compactification. It is conformal because we have a transformation of the form $\eta_{ab} \to \Omega^2\eta_{ab}$ for some function Ω which is positive in the bulk spacetime. It is a compactification because it transforms a spacetime that is not compact into a spacetime that is compact. There are, however, better choices of compactification. In our construction, we defined a function $\Omega = l$ which was not defined at some points of spacetime—namely those with $r = 0$. We can, however, define a conformal compactification that takes proper account of all points in Minkowski spacetime.

To that end, we will make some more changes of coordinates. We begin by defining the advanced time coordinate $v = t + r$. This is in contrast to the retarded time coordinate $u = t - r$. Both of these coordinates are depicted on Fig. 4.1 on the following page. In terms of retarded and advanced times and angular coordinates the Minkowski line element is written as

$$\mathrm{d}s^2 = -\,\mathrm{d}u\,\mathrm{d}v + \frac{1}{4}(v-u)^2\,\mathrm{d}\mathbb{S}^2 \,. \tag{95}$$

It is convenient to work with null coordinates—*i.e.*, coordinates with

$$\eta_{ab}\left(\frac{\partial}{\partial u}\right)^a\left(\frac{\partial}{\partial u}\right)^b = 0 \tag{96}$$

[3]By "bulk" we mean the interior of the spacetime, or the set of all "finite" points. This contrasts with infinity, which would be the "boundary" of spacetime.

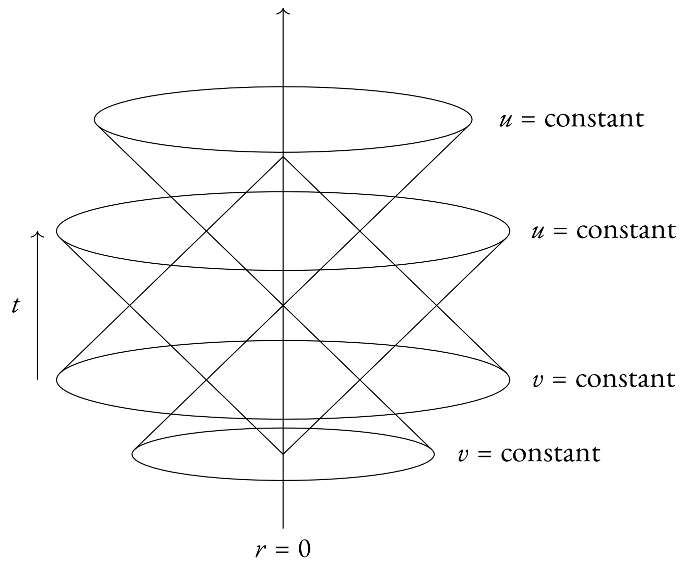


Figure 4.1: *Illustration, with one dimension suppressed, of the physical meaning of the null coordinates u and v. Surfaces of constant u are outgoing spherical "waves", while surfaces of constant v are their incoming analogues. The figure is based on Hawking and Ellis 1973, Fig. 12.i.*

(and similarly for v)—because we can compactify each of these coordinates separately and still get null coordinates. This will allow us at the end to draw a diagram in which all (radial) null geodesics are at 45° angles, which is convenient for reading the causal structure of spacetime.

We will now compactify the coordinates of Eq. (95) on the previous page. Our goal is to take the coordinates u and v which obey

$$-\infty < u \leq v < +\infty \tag{97}$$

and map them to coordinates with range in a finite interval. This can be done by means of a function $f\colon (a, b) \to \mathbb{R}$ where $-\infty < a < b < +\infty$ and such that f is injective. Two examples are the functions

$$\operatorname{artanh}\colon (-1, 1) \to \mathbb{R} \tag{98}$$

and

$$\tan\colon \left(-\frac{\pi}{2}, +\frac{\pi}{2}\right) \to \mathbb{R}, \tag{99}$$

but one could also work with other options. We will use the tangent function. We define new compactified null coordinates U and V through

$$u = \tan U \quad \text{and} \quad v = \tan V. \tag{100}$$

Notice that this maps the infinite range of u and v tp a finite range of U and V. Namely, U and V satisfy

$$-\frac{\pi}{2} < U \leq V < +\frac{\pi}{2}. \tag{101}$$

The middle inequality holds because tan is crescent and $u \leq v$.

In our new choice of coordinates, we have the line element

$$\mathrm{d}s^2 = \sec^2 U \sec^2 V \left[-\mathrm{d}U\,\mathrm{d}V + \frac{1}{4}\sin^2(V-U)\,\mathrm{d}\mathbb{S}^2\right]. \tag{102}$$

At this stage, our coordinates are already compactified. We just need to get rid of the divergences of the metric at infinity. To do so, we perform a conformal transformation. As Eq. (102) suggests, we take

$$\Omega = 2\cos U \cos V \tag{103}$$

so that we define $\tilde{g}_{ab} = \Omega^2 \eta_{ab}$ and get

$$\mathrm{d}\tilde{s}^2 = -4\,\mathrm{d}U\,\mathrm{d}V + \sin^2(V-U)\,\mathrm{d}\mathbb{S}^2. \tag{104}$$

For convenience, we finally introduce coordinates T and R through

$$T = U + V \quad \text{and} \quad R = V - U \tag{105}$$

which mimics (up to normalization) $u = t - r$ and $v = t + r$. This finally leads to the unphysical metric

$$\mathrm{d}\tilde{s}^2 = -\,\mathrm{d}T^2 + \mathrm{d}R^2 + \sin^2 R\,\mathrm{d}\mathbb{S}^2. \tag{106}$$

The physical region corresponds to the limits

$$-\pi < T - R \leq T + R < +\pi, \tag{107}$$

which in particular implies $R \geq 0$.

Eq. (106) is the line element for the Einstein static universe (Choquet-Bruhat 2015, Sec. VII.2.1). This spacetime has topology $\mathbb{R} \times \mathbb{S}^3$ and it is a solution to the Einstein field equations, but for a perfect fluid stress tensor with a cosmological constant component. This is, of course, a very different scenario from the Minkowski metric, which is a vacuum solution. This greatly exemplifies that the unphysical metric is really unphysical.

We can make a plot of the region given on Eq. (107). Plotting R on the horizontal axis and T on the vertical axis, we get the diagram shown in Fig. 4.2 on the next page. This is known as the Penrose diagram for Minkowski spacetime. It is a finite drawing of all of Minkowski spacetime, with each point representing a sphere $\mathbb{S}^2$. We can "double" the diagram by letting each point be a hemisphere, and this version of the diagram is interesting because it can be "wrapped around the Einstein static universe", as shown in Fig. 4.3 on page 26.

The region $T + R = +\pi$ of the Penrose diagram corresponds to the limits $r \to +\infty$ at constant u. Each direction leads to a different point and each different value of u leads to a different point as well. In total, we get a three-dimensional manifold $\mathscr{I}^+$ with topology $\mathbb{R} \times \mathbb{S}^2$: $\mathbb{R}$ corresponds to the values of u and $\mathbb{S}^2$ to the possible directions. Similarly, $T - R = -\pi$ is the region $\mathscr{I}^-$ of limits $r \to +\infty$ under constant v and it also has topology $\mathbb{R} \times \mathbb{S}^2$ for analogous reasons. $\mathscr{I}^+$ is known as the future null infinity and the symbol $\mathscr{I}^+$ is pronounced "scri plus". $\mathscr{I}^-$ ("scri minus") is the past null infinity.

The point $T = 0$ with $R = \pi$ is at the boundary between $\mathscr{I}^+$ and $\mathscr{I}^-$. Notice this is indeed a point: for $R = \pi$ we are at one of the poles of the three-sphere, which is thus a single point. This

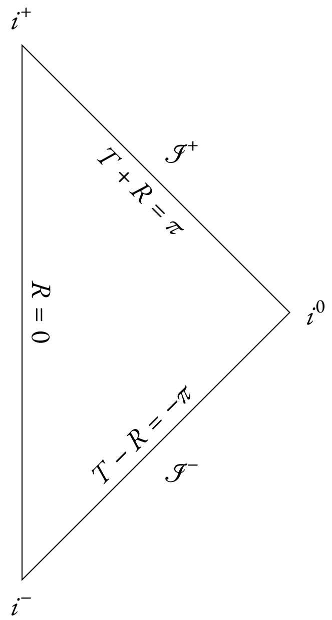


Figure 4.2: *Penrose diagram of Minkowski spacetime.*

point, denoted i^0, is known as spatial infinity and it corresponds to the limits $r \to +\infty$ at constant t. Notice that, since this is a single point, limits in different directions all coincide.

We still have two interesting points. $T = \pi$ with $R = 0$ is at the future of $\mathscr{I}^+$. This is located at the other pole of the three-sphere and thus is also corresponds to a single point, denoted i^+ and known as future timelike infinity. This is the limit $t \to +\infty$ at constant r, and there is no direction-dependence either. Similar comments are in order for the past timelike infinity i^-, located at $T = -\pi$ with $R = 0$.

All timelike curves start at i^- and end at i^+. All null curves start at $\mathscr{I}^-$ and end at $\mathscr{I}^+$. All spacelike curves start and end at i^0. Fig. 4.4 on page 27 illustrates some curves of interest on the Penrose diagram for Minkowski spacetime.

Conformal compactifications were introduced in general relativity by Penrose (1963, 1965), and we shall extensively use them in the rest of these notes.

We should mention an important property of the conformal transformations we are employing to find an unphysical spacetime. While it is true that $\tilde{g}_{ab}$ is unphysical it is important to notice that a vector k^a will satisfy $\tilde{g}_{ab} k^a k^b = 0$ for $\tilde{g}_{ab} = \Omega^2 g_{ab}$ if, and only if, $g_{ab} k^a k^b = 0$. As a consequence, conformal transformations preserve lightcones, and thus they preserve the causal structure of spacetime. This means that Penrose diagrams are useful tools for analyzing the causal structure of spacetime.

Notice also that in the Penrose diagram for Minkowski spacetime all radial light rays move at 45° angles. This allows us to read the causal structure of Minkowski spacetime directly from its Penrose diagram. This is the reason why these sorts of diagrams are popular in general relativity.

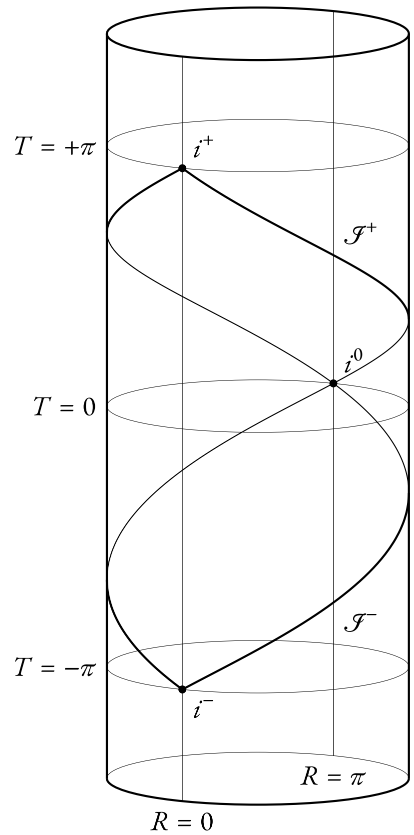


Figure 4.3: *Minkowski spacetime embedded in the Einstein static universe. Since the Einstein static universe has the topology* $\mathbb{R} \times \mathbb{S}^3$ *we represent it as a cylinder. The Penrose diagram for Minkowski spacetime is wrapped around the Einstein cylinder.*

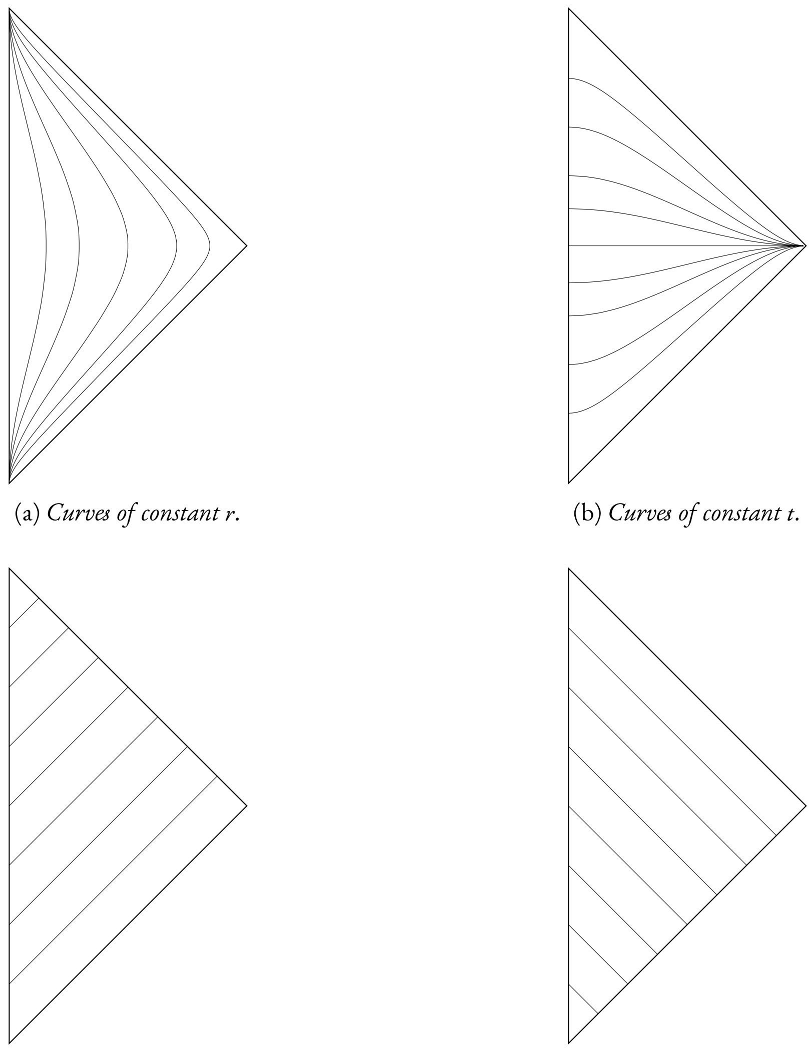
4.2 Asymptotically Flat Spacetimes

At this point we would like to adapt our discussion of infinity to more general spacetimes. This is related to defining what is an asymptotically flat spacetime.

In a general spacetime, the structure of infinity could be much different from that of Minkowski spacetime. Some spacetimes might not even get to infinity. Consider a dust-filled closed Friedman–Lemaître–Robertson–Walker universe, for example. Such a universe begins at a Big Bang, ends at a Big Crunch, and has compact spatial sections (it is spatially a sphere). Therefore, we never really get to infinity. We either reach the north or south pole of the spatial spheres or we reach the singularities at the beginning or end of spacetime.

We want to consider spacetimes that asymptotically look like Minkowski spacetime. Hence, their behavior at infinity should somehow resemble that of Minkowski spacetime. We know Minkowski spacetime has five different infinities, so there are five different meanings we can give to "asymptotically flat", plus combinations of them.

What are reasonable conditions we can impose? It makes sense to consider spacetimes in which the gravitational field falls off away from a central distribution of matter, so we have an



(a) *Curves of constant r.*

(b) *Curves of constant t.*

(c) *Curves of constant u.*

(d) *Curves of constant v.*

Figure 4.4: *Different curves represented in the Penrose diagram for Minkowski spacetime. Vertical curves are curves with constant r. Horizontal curves are curves with constant t. Diagonal lines have constant u (outgoing) or v (incoming).*

asymptotically flat spacetime at null and spatial infinities. There is less interest in spacetimes which are asymptotically flat at timelike infinities, though, because this means the spacetime becomes flat at late or early times, so the matter content has vanished. While there are situations in which this can be interesting, these are more specific spacetimes that often will not be what we are looking for. Hence, one often defines asymptotic flatness at spatial and null infinities.

Our definition will actually be even simpler than that. For our purposes, a single null infinity will typically be enough. Hence, we will discuss asymptotically flat spacetimes at future null infinity (the past case is analogous). We follow the book by Dappiaggi, Moretti, and Pinamonti (2017, Chap. 2), which also discusses the timelike case. For the more complete case which involves spatial infinity, see the book by Wald (1984, Chap. 11).

We will first discuss in a handwaving way each of the axioms that we will employ and then state the final definition. The goal is to motivate the definition. We will find a definition by mimicking and generalizing the construction of conformal infinity we performed for Minkowski spacetime.

We start with a spacetime (M, g_{ab}). We want to enforce conditions on this spacetime so that it can be considered asymptotically flat. Our construction should be coordinate-independent in nature, so that it actually captures the physical aspects of spacetime rather than spurious coordinate behaviors.

Our first demand will be that there is an unphysical spacetime $(\tilde{M}, \tilde{g}_{ab})$ which conformally extends (M, g_{ab}). This unphysical spacetime will play the same role that the Einstein static universe played for Minkowski spacetime.

We need a way of injecting M into $\tilde{M}$. Before, we had a natural way of viewing Minkowski spacetime as a submanifold of the Einstein static universe, which was given by the coordinate restrictions on the Einstein universe. In our more general case, this role is played by a function $\psi\colon M \to \tilde{M}$ that should satisfy some convenient properties. In practice, it is convenient to choose ψ to be an embedding. Let us define this by following Tu (2011).

Definition 8 [Immersion]:

Let $\psi\colon M \to \tilde{M}$ be a smooth map between the manifolds M and $\tilde{M}$. We say ψ is an immersion if, and only if, $\psi^*\colon T_pM \to T_{\psi(p)}\tilde{M}$ is injective for all $p \in M$. ♠

Definition 9 [Embedding]:

Let M and $\tilde{M}$ be smooth manifolds. An immersion $\psi\colon M \to \tilde{M}$ is called an embedding if, and only if, it is injective and $\psi\colon M \to \psi(M)$ is a homeomorphism, where $\psi(M)$ is endowed with the subspace topology. ♠

Definition 10 [Embedded Submanifold]:

Let M and $\tilde{M}$ be smooth manifolds. If $\psi\colon M \to \tilde{M}$ is an embedding, we say that $\psi(M)$ is an embedded submanifold of $\tilde{M}$. ♠

It holds that embeddings are diffeomorphisms onto their images. Hence, we can think of an embedding as being a generalization of diffeomorphism that does not need to be surjective. It makes a copy of M in $\tilde{M}$.

Hence, we have the next condition we want to impose on (M, g_{ab}): there is an embedding $\psi\colon M \to \tilde{M}$, and we also require as a technical condition that $\psi(M)$ is an open subset of $\tilde{M}$.

Having a smooth embedding, we have a way of connecting the two manifolds M and $\tilde{M}$. Nevertheless, we still have not imposed any restrictions on the metrics g_{ab} and $\tilde{g}_{ab}$. To do so, we notice that $\psi: M \to \psi(M)$ is a diffeomorphism, and hence we can use it to perform both pullbacks and pushforwards of tensor fields in M and $\psi(M)$. We will then ask that there is a function $\Omega \in \mathscr{C}^\infty(\psi(M))$ such that $\Omega > 0$ and

$$\tilde{g}_{ab}\,|_{\psi(M)} = \Omega^2 \psi^* g_{ab}. \tag{108}$$

This ensures that ψ represents a conformal transformation between the spacetimes (M, g_{ab}) and $(\tilde{M}, \tilde{g}_{ab})$.

Given these objects, we define future null infinity as the boundary $\mathscr{I}^+ = \partial\psi(M)$. $\mathscr{I}^+$ is an embedded submanifold of $\tilde{M}$ and it is not to the past of any points of $\psi(M)$. In symbols, we write

$$\mathscr{I}^+ \cap J^-\left(\psi(M); \tilde{M}\right) = \emptyset. \tag{109}$$

The causal past of a set S, $J^-(S)$, is defined as follows.

Definition 11 [Causal Past]:

Let (M, g_{ab}) be a spacetime. Consider a set $S \subseteq M$. The causal past of S in M, denoted $J^-(S; M)$ or simply $J^-(S)$, is defined as the set of all points in M from which one can reach S by means of a future-directed causal curve. See the book by Wald (1984, Chap. 8) for further details. ♠

Hence, $\mathscr{I}^+$ is defined in this general construction as a submanifold of $\tilde{M}$ satisfying a couple of special properties. In particular, it is a future boundary in the sense that is is not in the past of any points in the spacetime.

Near infinity we expect the spacetime to be reasonably well-behaved, so that it actually looks like Minkowski spacetime. Thus, we demand it enjoys reasonably good causal properties. We cannot allow for closed timelike curves near infinity, for example. We enforce this behavior by demanding that strong causality holds in a neighborhood of infinity.

Definition 12 [Strong Causality]:

Consider a spacetime (M, g_{ab}). We say (M, g_{ab}) is strongly causal if for any point $p \in M$ and every neighborhood U of p there is a neighborhood V of p with $p \in V \subseteq O$ such that no causal curve intersects V more than once. ♠

As discussed by Wald (1984, Chap. 8), strong causality essentially means that no causal curve comes arbitrarily close to intersecting itself. This prevents causality violations upon perturbations of the metric in an arbitrarily small neighborhood of a given point. There are stronger causality impositions one could consider, but we shall assume only strong causality.

We need to impose that Ω brings infinity in. This is done by imposing that Ω vanishes on $\mathscr{I}^+$. However, we defined Ω only on $\psi(M)$. Hence, we ask that Ω extends to a function $\Omega \in \mathscr{C}^\infty(\tilde{M})$ such that

$$\Omega|_{\mathscr{I}^+} = 0. \tag{110}$$

Furthermore, we want $\mathscr{I}^+$ to be precisely the submanifold with $\Omega = 0$, so we must also require that

$$\tilde{\nabla}_a \Omega|_{\mathscr{I}^+} \neq 0. \tag{111}$$

This ensures $\Omega = 0$ is indeed a hypersurface and ensures there is a non-vanishing vector normal to the surface. Hence, we get a single "sheet" with $\Omega = 0$ rather than a plateau.

We want to obtain the full range of limits $r \to +\infty$ at constant u's. To ensure we get all points at infinity, we define the vector field $n^a = \tilde{g}^{ab}\tilde{\nabla}_b \Omega$. We then demand the existence of a function $\omega \in \mathscr{C}^\infty(\tilde{M})$ with $\omega > 0$ on $\psi(M) \cup \mathscr{I}^+$ such that

$$\tilde{\nabla}_a \left(\omega^4 n^a\right)\big|_{\mathscr{I}^+} = 0 \tag{112}$$

and such that the integral curves of $\omega^{-1} n^a$ are complete on $\mathscr{I}^+$. This technical condition is meant to ensure that $\mathscr{I}^+ \cong \mathbb{R} \times \mathbb{S}^2$.

Finally, we need for the Riemann tensor to fall off sufficiently quickly near infinity. Thus, we ask that (M, g_{ab}) satisfies the vacuum Einstein field equations on some neighborhood of the boundary of $\psi(M)$. This can be weakened to requiring that the vacuum Einstein field equations hold asymptotically (Wald 1984, Chap. 11). The point is that at infinity we should get a vacuum solution, so that we are far away from any matter sources.

We now bring everything together to get to the following definition.

Definition 13 [Asymptotically Flat Spacetime at Future Null Infinity]:
Consider a spacetime (M, g_{ab}). Suppose that there are

i. an unphysical spacetime $(\tilde{M}, \tilde{g}_{ab})$,

ii. a smooth embedding $\psi\colon M \to \tilde{M}$ such that $\psi(M)$ is open in $\tilde{M}$,

iii. and a smooth function $\Omega\colon \psi(M) \to \mathbb{R}$ with $\Omega > 0$ and

$$\tilde{g}_{ab} = \Omega^2 \psi^* g_{ab}. \tag{113}$$

Furthermore, suppose that these objects are such that the following conditions are met:

i. $\psi(M)$ is a manifold with boundary $\mathscr{I}^+ = \partial\psi(M)$, where $\mathscr{I}^+$ is an embedded three-manifold of $\tilde{M}$ and it holds that $\mathscr{I}^+ \cap J^-\left(\psi(M); \tilde{M}\right) = \emptyset$.

ii. Strong causality holds in $(\tilde{M}, \tilde{g}_{ab})$ at a neighborhood of $\mathscr{I}^+$.

iii. Ω can be extended to a smooth function $\Omega\colon \tilde{M} \to \mathbb{R}$ with $\Omega|_{\mathscr{I}^+} = 0$ and $\tilde{\nabla}_a \Omega|_{\mathscr{I}^+} \neq 0$.

iv. Denoting $n^a = \tilde{g}^{ab}\tilde{\nabla}_b \Omega$, there is a smooth function $\omega\colon \tilde{M} \to \mathbb{R}$ with $\omega > 0$ such that $\tilde{\nabla}_a \left(\omega^4 n^a\right)|_{\mathscr{I}^+} = 0$ and such that the integral lines of $\omega^{-1} n^a$ are complete.

v. The vacuum Einstein field equations hold for (M, g_{ab}) on a neighborhood of infinity, or at least asymptotically as one approaches infinity.

If all of these conditions are met, we say that (M, g_{ab}) is asymptotically flat at future null infinity. ♠

4.3 Case Study: Schwarzschild Spacetime

As a case study, let us show that Schwarzschild spacetime is asymptotically flat at future null infinity (and actually at past null infinity too). We follow the discussion by Schmidt and Walker (1983) and we will not make an effort to prove all of the conditions of asymptotic flatness, but rather show the conformal compactification procedure.

We want to get to future null infinity, so it makes sense to employ a retarded time coordinate instead of the more usual Schwarzschild time coordinate. In this scenario, u is known as the retarded Eddington–Finkelstein coordinate. The Schwarzschild metric is written as

$$\mathrm{d}s^2 = -\left(1 - \frac{2M}{r}\right)\mathrm{d}u^2 - 2\,\mathrm{d}u\,\mathrm{d}r + r^2\,\mathrm{d}\mathbb{S}^2\,. \tag{114}$$

M is the black hole's mass, and the coordinate ranges are $u \in \mathbb{R}$ and $r > 0$. Notice that $r = 0$ is *not* a part of spacetime, since it is a physical singularity.

We proceed as with Minkowski spacetime. Define a new coordinate $l = \frac{1}{r}$, which is now defined on the entire spacetime, since $r > 0$ strictly. Using this new coordinate we get the line element

$$\mathrm{d}s^2 = -(1 - 2Ml)\,\mathrm{d}u^2 + 2l^{-2}\,\mathrm{d}u\,\mathrm{d}l + l^{-2}\,\mathrm{d}\mathbb{S}^2\,. \tag{115}$$

As with the Minkowski spacetime, we have a divergence at $l \to 0^+$ because the area of the spheres tends to infinity. We solve this by multiplying the metric by l^2 (which means our conformal factor is $\Omega = l$) to get to

$$\mathrm{d}\tilde{s}^2 = -l^2(1 - 2Ml)\,\mathrm{d}u^2 + 2\,\mathrm{d}u\,\mathrm{d}l + \mathrm{d}\mathbb{S}^2\,. \tag{116}$$

This unphysical metric can now be extended so that $l \in \mathbb{R}$. Through this procedure we get the unphysical spacetime $(\tilde{M}, \tilde{g}_{ab})$.

5 The BMS Group

We are almost ready to discuss the Bondi–Metzner–Sachs (BMS) group. The last pre-requisite we should discuss is the basic theory of Carrollian manifolds, which is the sort of structure present at null infinity. This language will provide the natural arena for us to discuss the BMS group.

5.1 Carrollian Structures

By going back to the examples we provided of Minkowski and Schwarzschild spacetimes, one can notice that the line element for the induced metric at $\mathscr{I}^+$ appears to have the form given in Eq. (94) on page 22,

$$\mathrm{d}\tilde{\sigma}^2\,\big|_{l=0} = 0\,\mathrm{d}u^2 + \mathrm{d}\mathbb{S}^2\,. \tag{117}$$

Notice this is not a pseudo-Riemannian metric, because it is degenerate. If we denote this metric by $\tilde{h}_{ab}$, we know there is a vector $n^a \neq 0$ such that

$$\tilde{h}_{ab}\,n^a = 0. \tag{118}$$

Namely, in these coordinates we have

$$n^a = \left(\frac{\partial}{\partial u}\right)^a. \tag{119}$$

These sorts of spacetimes can be understood as limiting case of Lorentzian spacetimes. Namely, consider the Minkowski metric

$$ds^2 = -c^2\, dt^2 + dx^2 + dy^2 + dz^2\,, \tag{120}$$

where we intentionally wrote the c factors explicitly. Notice now that in the unusual limit $c \to 0$ we get

$$ds^2 = -0\, dt^2 + dx^2 + dy^2 + dz^2\,, \tag{121}$$

which is a degenerate metric. Hence, the sort of manifold we are interested in can be understood as the $c \to 0$ limit of a Lorentzian manifold.

These sorts of limits were originally considered by Lévy-Leblond (1965), who was interested in studying the $c \to 0$ limit of the Poincaré group. Notice that in this limit the lightcones "close" around the time axis, and there is no propagation. This resembles the strange causality features present in the books by Lewis Carroll, in which Alice now and again finds herself in unexpected situations. This prompted the name "Carrollian manifold" for manifolds endowed with a metric that is degenerate along a specific direction.

The work of Lévy-Leblond (1965) was mostly focused on the Carroll group, which is the $c \to 0$ limit of the Poincaré group. Carrollian structures in differential geometry were defined and studied by Duval, Gibbons, and Horvathy (2014a,b) and Duval, Gibbons, Horvathy, and Zhang (2014). We now review the basic ideas and definitions necessary for the study of BMS symmetries.

When describing a pseudo-Riemannian manifold, we typically provide a pair (M, g_{ab}). M is the underlying smooth manifold on which we consider a smooth metric tensor g_{ab}. By analogy, one would in principle expect that for Carrollian manifolds we should specify a pair $(\mathscr{I}^+, \tilde{h}_{ab})$, where $\mathscr{I}^+$ is the underlying smooth manifold and $\tilde{h}_{ab}$ is the degenerate metric tensor field. This, however, is insufficient. Due to $\tilde{h}_{ab}$ being degenerate, we also need to specify the non-vanishing vector n^a such that $\tilde{h}_{ab} n^a = 0$. This vector has the role of defining the kernel of the metric tensor. A triple $(\mathscr{I}^+, \tilde{h}_{ab}, n^a)$ is known as a weak Carrollian structure, or simply a Carrollian structure.

The adjective "weak" makes reference to the fact that one can strengthen the definition of a Carrollian structure by further specifying a covariant derivative. To understand why, recall that in a pseudo-Riemannian manifold the metric singles out a particular choice of covariant derivative by imposing that the Christoffel symbols be given by

$$\Gamma^\rho{}_{\mu\nu} = \frac{1}{2} g^{\rho\sigma}\left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}\right). \tag{122}$$

Nevertheless, this is not applicable for a Carrollian metric. Since Carrollian metrics are degenerate, the inverse metric tensor g^{ab} does not exist. Hence, imposing that the metric is parallelly transported by the covariant derivative does not single out a covariant derivative. Duval, Gibbons, Horvathy, and Zhang (2014) go further and ask that n^a is too parallelly transported, but this is also not enough to specify a single covariant derivative. Thus, one has to specify a covariant derivative in addition to a weak Carrollian structure. A quadruple $(\mathscr{I}^+, \tilde{h}_{ab}, n^a, \tilde{\nabla}_a)$ is known as a strong Carrollian structure.

Definition 14 [Carrollian Structures]:

A weak Carrollian structure, or simply a Carrollian structure, is a triple $(\mathscr{I}^+, \tilde{h}_{ab}, n^a)$ where $\mathscr{I}^+$ is a smooth manifold with dimension d, $\tilde{h}_{ab}$ is a positive semi-definite symmetric tensor with matrix rank $d-1$, and n^a is a non-vanishing vector with $\tilde{h}_{ab} n^a = 0$.

A strong Carrollian structure is a quadruple $(\mathscr{I}^+, \tilde{h}_{ab}, n^a, \tilde{\nabla}_a)$ where $(\mathscr{I}^+, \tilde{h}_{ab}, n^a)$ is a weak Carrollian structure and $\tilde{\nabla}_a$ is a covariant derivative with $\tilde{\nabla}_a \tilde{h}_{bc} = 0$ and $\tilde{\nabla}_a n^b = 0$. ♠

Now suppose we are given an asymptotically flat spacetime (M, g_{ab}) with unphysical conformal extension $(\tilde{M}, \tilde{g}_{ab})$ and conformal factor Ω. Then future null infinity has a natural Carrollian structure $(\mathscr{I}^+, \tilde{h}_{ab}, n^a)$. Namely, $\mathscr{I}^+$ is just future null infinity itself, $\tilde{h}_{ab}$ is the metric induced on $\mathscr{I}^+$ by $\tilde{g}_{ab}$, and $n^a = \tilde{g}^{ab}\tilde{\nabla}_b \Omega$.

To proceed, notice that there is a gauge freedom in the definition of future null infinity. Namely, if (M, g_{ab}) is asymptotically flat with conformal factor Ω, then the conformal factor $\omega\Omega$ for $\omega \in \mathscr{C}^\infty(\tilde{M})$ and $\omega > 0$ would work just as well. Using this freedom, the Carrollian structure of null infinity can be made into a strong Carrollian structure. As discussed by Wald (1984, Chap. 11), one can use this gauge freedom to impose that $\tilde{\nabla}_a n^b = 0$. Notice that this condition is equivalent to imposing that $\tilde{\nabla}_a \tilde{\nabla}_b \Omega = 0$, which does not need to hold in general. However, one can always find a choice of Ω for which this holds.

The strong Carrollian structure ends up being too strong for our purposes. Hence, we will work mostly with the weak Carrollian structure of null infinity and only comment later on the role the strong Carrollian structure can play.

5.2 Symmetries at Null Infinity

We are now ready to discuss the symmetry group of future null infinity. To do so, consider the gauge freedom mentioned at the end of the last section, which consists of exchanging the conformal factor Ω according to $\Omega \to \omega\Omega$ for smooth $\omega > 0$. Using that $\tilde{g}_{ab} = \Omega^2 g_{ab}$ (we are omitting the pushforward for simplicity), that $\tilde{h}_{ab}$ is the metric induced by $\tilde{g}_{ab}$ on $\mathscr{I}^+$, and that $n^a = \tilde{g}^{ab}\tilde{\nabla}_b \Omega$ we get that under $\Omega \to \omega\Omega$ the Carrollian structure at null infinity transforms according to

$$\mathscr{I}^+ \to \mathscr{I}^+, \quad \tilde{h}_{ab} \to \omega^2 \tilde{h}_{ab}, \quad \text{and} \quad n^a \to \omega^{-1} n^a. \tag{123}$$

There is no *a priori* preference on whether we should choose Ω or $\omega\Omega$ when performing the conformal compactification. After all, this procedure is unphysical. Therefore, we must consider the two scenarios as being physically equivalent. We thus get an equivalence relation being Carrollian structures

$$(\mathscr{I}^+, \tilde{h}_{ab}, n^a) \sim (\mathscr{I}^+, \omega^2 \tilde{h}_{ab}, \omega^{-1} n^a). \tag{124}$$

The transformations that take $(\mathscr{I}^+, \tilde{h}_{ab}, n^a)$ to an equivalent structure $(\mathscr{I}^+, \omega^2 \tilde{h}_{ab}, \omega^{-1} n^a)$ should be regarded as symmetries of future null infinity. The transformations that preserve $\mathscr{I}^+$ are diffeomorphisms, the transformations that preserve $(\mathscr{I}^+, \tilde{h}_{ab})$ (up to gauge transformation) are conformal isometries, and the transformations that preserve $(\mathscr{I}^+, \tilde{h}_{ab}, n^a)$ are a subclass of conformal isometries that preserve the so-called strong conformal geometry (see, *e.g.*, Penrose 1974; Penrose and Rindler 1986).

We will begin our discussion by taking a coordinate approach. Using the gauge freedom in the definition of future null infinity one can choose a coordinate system near $\mathscr{I}^+$ in which (see, *e.g.*, Wald 1984, Chap. 11)

$$d\tilde{s}^2 = 2\,d\Omega\,du + d\mathbb{S}^2 + \mathcal{O}(\Omega), \tag{125}$$

where Ω is the conformal factor used in the definition of null infinity. The induced metric $\tilde{h}_{ab}$ then has the line element

$$d\tilde{\sigma}^2 = d\mathbb{S}^2 \tag{126}$$

while the normal vector n^a is given by

$$n^a = \left(\frac{\partial}{\partial u}\right)^a. \tag{127}$$

Notice this choice of coordinates does not restrict the generality of our analysis. It merely provides a "standard" triple

$$(\mathscr{I}^+, \tilde{h}_{ab}, n^a) = \left(\mathbb{R} \times \mathbb{S}^2, d\mathbb{S}^2, \left(\frac{\partial}{\partial u}\right)^a\right). \tag{128}$$

This can be thought of as a canonical choice of null infinity which we use to compare with other possible choices. Notice that these other possible choices are just as physical as this canonical choice.

We will find the conformal isometries that preserve the strong conformal structure—*i.e.*, the transformations that preserve the Carrollian structure—by finding the transformations that preserve $\mathscr{I}^+$, then restrict them to the transformations that preserve $(\mathscr{I}^+, \tilde{h}_{ab})$ (up to gauge transformation), and finally restricting these to the transformations that preserve $(\mathscr{I}^+, \tilde{h}_{ab}, n^a)$ (up to gauge transformation).

The transformations that preserve $\mathscr{I}^+$ are diffeomorphisms. This is still extremely general. We can restrict this large group further by imposing that $(\mathscr{I}^+, \tilde{h}_{ab})$ is preserved up to gauge transformation. Since the gauge transformations map $(\mathscr{I}^+, \tilde{h}_{ab})$ to $(\mathscr{I}^+, \omega^2 \tilde{h}_{ab})$ we see that this is the group of conformal isometries of $(\mathscr{I}^+, \tilde{h}_{ab})$. In the canonical triplet, we have

$$(\mathscr{I}^+, \tilde{h}_{ab}) = (\mathscr{I}^+, d\mathbb{S}^2). \tag{129}$$

Hence, we want transformations that are conformal isometries of the sphere (since we have the metric of a sphere) and change u in an arbitrary way. We know from Section 3.4 that the conformal isometries on the sphere are the Möbius transformations, which in stereographic coordinates can be written as

$$\zeta \to \frac{a\zeta + b}{c\zeta + d} \tag{130}$$

with

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C})/\mathbb{Z}_2 \cong \mathrm{SO}^+(3,1). \tag{131}$$

Furthermore, we have to assign a general transformation law for u. Since u does not occur in the metric, we can preserve the conformal structure of the metric by perfoming any transformation of the form

$$u \to F(u, \zeta, \bar{\zeta}) \tag{132}$$

with $\frac{\partial F}{\partial u} > 0$ (a condition which is meant to keep the coordinates well-defined).

We thus get the general coordinate transformations

$$\begin{cases} \zeta \to \dfrac{a\zeta + b}{c\zeta + d}, \\ u \to F(u, \zeta, \bar{\zeta}). \end{cases} \tag{133}$$

This collection of transformations form a group known as the Newman–Unti group (Newman and Unti 1962). As a set, the Newman–Unti group is given by $\mathrm{SO}^+(3,1) \times \mathscr{C}^\infty(\mathbb{R} \times \mathbb{S}^2)$.

For completeness, we mention that under the Möbius transformation

$$\zeta \to \zeta' = \frac{a\zeta + b}{c\zeta + d} \tag{134}$$

the metric of the sphere transforms according to (recall Eq. (62) on page 16)

$$2\gamma_{\zeta'\bar{\zeta}'}\,\mathrm{d}\zeta\,\mathrm{d}\bar{\zeta}' = 2K(\zeta,\bar{\zeta})^2\gamma_{\zeta\bar{\zeta}}\,\mathrm{d}\zeta\,\mathrm{d}\bar{\zeta} \tag{135}$$

with

$$K(\zeta,\bar{\zeta}) = \frac{1 + \zeta\bar{\zeta}}{(a\zeta + b)(\bar{a}\bar{\zeta} + \bar{b}) + (c\zeta + d)(\bar{c}\bar{\zeta} + \bar{d})}. \tag{136}$$

Now we impose that the Carrollian structure is preserved as a whole. This means we also must have that $n^a \to \omega^{-1} n^a$. We know already that $\omega = K(\zeta,\bar{\zeta})$ because we have already restricted the set of possible transformations to the Newman–Unti group. To further restrict the transformations, we notice that the u parameter is defined by the equation

$$n^a \bar{\nabla}_a u = 1, \tag{137}$$

and this equation is gauge-independent. Hence, since the normal vector transforms as $n^a \to \omega^{-1} n^a$, it follows that we must have

$$\bar{\nabla}_a u \to \omega \bar{\nabla}_a u. \tag{138}$$

This is better expressed in the language of differential forms, in which case we find that the transformation for u must be such that

$$\mathrm{d}u \to \mathrm{d}u' = K(\zeta,\bar{\zeta})\,\mathrm{d}u\,, \tag{139}$$

where we used $\omega = K(\zeta,\bar{\zeta})$. Integrating Eq. (139) allows us to conclude that the allowed transformations for u are those with the form

$$u \to u' = K(\zeta,\bar{\zeta})(u + f(\zeta,\bar{\zeta})), \tag{140}$$

where $f \in \mathscr{C}^\infty(\mathbb{S}^2)$ is an arbitrary function that occurs as an integration constant. We assume it to be smooth because we are working in a smooth manifold, so the chart transition maps should be smooth.

The transformations that preserve the (weak) Carrollian structure are thus given by

$$\begin{cases} \zeta \to \dfrac{a\zeta + b}{c\zeta + d}, \\ u \to K(\zeta,\bar{\zeta})(u + f(\zeta,\bar{\zeta})), \end{cases} \tag{141}$$

where K is given by Eq. (136) on the previous page. These are known as Bondi–Metzner–Sachs (BMS) transformations and they form the (restricted) BMS group (Bondi, Van der Burg, and Metzner 1962; Sachs 1962b), which we will denote as G_{BMS}. The adjective "restricted" refers to the fact that we are ignoring time and space reflections.

Notice that a BMS transformation is characterized by a pair $(\Lambda, f) \in \text{SO}^+(3,1) \times \mathscr{C}^\infty(\mathbb{S}^2)$, where Λ relates to the coefficients a, b, c, and d in Eq. (141) on the preceding page by

$$\Pi^{-1}(\Lambda) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{142}$$

with $\Pi\colon \text{SL}(2,\mathbb{C})/\mathbb{Z}_2 \to \text{SO}^+(3,1)$ being the isomorphism between the two groups. We denote the function K associated to Λ by K_Λ. Using this notation, the group product $\odot$ is given by

$$(\Lambda', f') \odot (\Lambda, f) = (\Lambda'\Lambda, f + (K_{\Lambda^{-1}} \circ \Lambda) \cdot (f' \circ \Lambda)), \tag{143}$$

where $\circ$ denotes composition of mappings and $\cdot$ denotes the pointwise product of functions.

Notice that $\text{SO}^+(3,1)$ is isomorphic to the subgroup

$$\mathscr{L} = \left\{(\Lambda, 0); \Lambda \in \text{SO}^+(3,1)\right\} \tag{144}$$

of the BMS group. Similarly, $\mathscr{C}^\infty(\mathbb{S}^2)$ is isomorphic to

$$T = \left\{(\mathbb{1}, f); f \in \mathscr{C}^\infty(\mathbb{S}^2)\right\}. \tag{145}$$

Eq. (143) allows us to conclude that T is a normal subgroup of the BMS group. $\mathscr{L}$ is not a normal subgroup of the BMS group.

Notice that these two subgroups satisfy

$$\mathscr{L} \cap T = \{(\mathbb{1}, 0)\}, \tag{146}$$

which means their intersection is given by the BMS group's identity. Furthermore, using Eq. (143) we can see that any element (Λ, f) of the BMS group can be written in the form

$$(\Lambda, f) = (\Lambda, 0) \odot (\mathbb{1}, f). \tag{147}$$

Hence, we can write the BMS group as $\mathscr{L}T$. These results allow us to conclude that the BMS group is the semidirect product of $\mathscr{L}$ and T. Hence, the BMS group is given by $\mathscr{L} \ltimes T$. Using that $\mathscr{L} \cong \text{SO}^+(3,1)$ and $T \cong \mathscr{C}^\infty(\mathbb{S}^2)$, we can write that the BMS group is given by the semidirect product $\text{SO}^+(3,1) \ltimes \mathscr{C}^\infty(\mathbb{S}^2)$.

5.3 Alternative Derivation with Conformal Killing Vector Fields

Let us now derive the BMS group once again, but this time employing vector field methods. In other words, we want to find which vector fields generate the BMS transformations. This will make it easier for us to compare the results with those of the Poincaré group.

We know we want to consider conformal isometries. Furthermore, these conformal isometries should also preserve (in a conformal sense) the normal vector n^a. We can state these conditions mathematically by stating that we are looking for diffeomorphisms generated by vector fields ξ^a such that

$$\pounds_\xi \tilde{h}_{ab} = \lambda \tilde{h}_{ab} \quad \text{and} \quad \pounds_\xi n^a = -\frac{\lambda}{2} n^a, \tag{148}$$

where the second condition follows from the fact that n^a transforms as $n^a \to \omega^{-1} n^a$ while $\tilde{h}_{ab}$ transforms as $\tilde{h}_{ab} \to \omega^2 \tilde{h}_{ab}$.

Eq. (148) establishes a system of differential equations for the vector field ξ^a. Eq. (55) on page 14 allows us to express this system of differential equations in terms of *any* derivative operator on $\mathscr{I}^+$. We choose to do so with the derivative operator D_a, which is defined as having the same Christoffel symbols as the Levi-Civita connection on the round sphere, with the remaining Christoffel symbols vanishing. This choice of connection turns $(\mathscr{I}^+, \tilde{h}_{ab}, n^a, D_a)$ into a strong Carrollian structure.

Using the condition imposed on the metric, we find that

$$\lambda \tilde{h}_{ab} = \pounds_\xi \tilde{h}_{ab}, \tag{149a}$$

$$= \xi^c D_c \tilde{h}_{ab} + \tilde{h}_{cb} D_a \xi^c + \tilde{h}_{ac} D_b \xi^c, \tag{149b}$$

$$= D_a (\tilde{h}_{cb} \xi^c) + D_b (\tilde{h}_{ac} \xi^c). \tag{149c}$$

The equation

$$D_a (\tilde{h}_{cb} \xi^c) + D_b (\tilde{h}_{ac} \xi^c) = \lambda \tilde{h}_{ab} \tag{150}$$

is the conformal Killing equation for a vector field $\tilde{h}_{ab} \xi^b$ defined on the sphere. Therefore, we can conclude that the projection of ξ^a on the sphere is a conformal Killing vector field on the sphere. If we denote this conformal Killing vector field as Y^a we can thus write

$$\xi^a = Y^a + F n^a, \tag{151}$$

for some smooth function F which we assume is such that $\frac{\partial F}{\partial u} > 0$. This family of vector fields generate the Newman–Unti transformations.

At this point, it is useful to notice that Eqs. (150) and (151) imply

$$D_a Y_b + D_b Y_a = \lambda \tilde{h}_{ab}. \tag{152}$$

This can be understood as an equation on the two-sphere, in which we have an inverse metric available. Contracting the expression with the inverse metric allows us to conclude that

$$D_a Y^a = \lambda. \tag{153}$$

This last equation can be understood as both on the sphere or on $\mathscr{I}^+$, as it is the same statement in both cases.

As in our previous derivation, we get from the Newman–Unti group to the BMS group by considering the behavior of the normal vector n^a. We have

$$-\frac{\lambda}{2} n^a = \pounds_\xi n^a, \tag{154a}$$

$$= \xi^b D_b n^a - n^b D_b \xi^a, \tag{154b}$$

$$= -n^b D_b \xi^a, \tag{154c}$$

where we used the fact that $D_a n^b = 0$. In terms of D_a, the differential equation implied by the behavior of n^a is

$$n^b D_b \xi^a = \frac{\lambda}{2} n^a, \tag{155}$$

where λ is given by Eq. (153) on the preceding page. We can use Eq. (151) on the previous page to find that

$$n^b D_b \xi^a = n^b D_b Y^a + n^b D_b (F n^a), \tag{156a}$$
$$= (n^b D_b F) n^a. \tag{156b}$$

This last expression can be obtained by using $D_a n^b = 0$ and $n^b D_b Y^a = 0$—the latter can be obtained by expressing the equation in terms of Christoffel symbols.

Eqs. (155) and (156) can be combined to yield

$$(n^b D_b F) n^a = \frac{\lambda}{2} n^a. \tag{157}$$

This implies

$$n^a D_a F = \frac{\lambda}{2}, \tag{158}$$

which together with Eq. (153) on the previous page yields

$$n^a D_a F = \frac{D_a Y^a}{2}. \tag{159}$$

This is a differential equation that must be respected by the function F if we want n^a to be preserved up to a gauge transformation.

To solve this differential equation, introduce the coordinate u through $n^a D_a u = 1$. Since $D_a Y^a$ can be understood as defined on the sphere, it bears no dependence on the parameter u. We can then integrate the differential equation to get

$$F(u, \zeta, \bar{\zeta}) = \frac{D_a Y^a}{2} u + f(\zeta, \bar{\zeta}), \tag{160}$$

where $f \in \mathscr{C}^\infty(\mathbb{S}^2)$ arises as an integration "constant".

Bringing everything together we find that the generic vector field generating a BMS transformation is given by

$$\xi^a = Y^a + \left(\frac{D_b Y^b}{2} u + f(\zeta, \bar{\zeta}) \right) n^a, \tag{161}$$

with Y^a being some conformal Killing vector field on the two-sphere, $n^b D_b u = 1$, and $f \in \mathscr{C}^\infty(\mathbb{S}^2)$.

5.4 Poincaré Group as a BMS Subgroup

Now that we know what the BMS group is, it is interesting understand how the Poincaré group fits within it.

We begin by noticing that the BMS group is given by $G_{\text{BMS}} = \text{SO}^+(3,1) \ltimes \mathscr{C}^\infty(\mathbb{S}^2)$, while the Poincaré group is $\text{ISO}^+(3,1) = \text{SO}^+(3,1) \ltimes \mathbb{R}^4$. The semidirect product structure is certainly similar, but we still need to understand how and whether the Poincaré group is a subgroup of the BMS group.

The physical reason we expect $\mathrm{ISO}^+(3,1)$ to fit inside G_{BMS} is because G_{BMS} is the symmetry group at future null infinity, while $\mathrm{ISO}^+(3,1)$ is the symmetry group in the bulk of Minkowski spacetime. It seems reasonable that bulk symmetries should extend to boundary symmetries, so it is reasonable to expect that there is a relation between these two groups. In fact, before the analysis by Bondi, Van der Burg, and Metzner (1962) and Sachs (1962b), it was thought that the symmetry group at infinity should simply be the Poincaré group. Hence, a byproduct of our discussion should be to understand what are the extra symmetries present in the BMS group.

We start by considering Lorentz transformations. These are the vector fields J_i^a and K_i^a given on Eq. (88) on page 20. These vectors induce vectors on $\mathscr{I}^+$. These induced vectors can be computed by considering the expressions on Eq. (88) on page 20, changing coordinates from (t,x,y,z) to $(u,r,\zeta,\bar{\zeta})$, dropping the r-component of the resulting vectors (which is not defined intrinsically on $\mathscr{I}^+$) and then taking the $r \to +\infty$. Through this procedure, the generators of Lorentz transformations take the general form

$$\xi^a = Y^a + \frac{u}{2} D_b Y^b \left(\frac{\partial}{\partial u}\right)^a, \tag{162}$$

where Y^a is some conformal Killing field on the sphere. Comparing this equation with Eq. (161) on the preceding page lets us notice that Lorentz transformations correspond precisely to the $\mathrm{SO}^+(3,1)$ contributions to the BMS group. This is not surprising, since $\mathrm{SO}^+(3,1)$ is indeed the Lorentz group.

We still have to discuss the role of translations. These are the fields P_μ^a given on Eq. (88) on page 20. We can compute the induced vectors on future null infinity. One finds the general form

$$\xi^a = \tilde{f}(\zeta,\bar{\zeta}) \left(\frac{\partial}{\partial u}\right)^a, \tag{163}$$

where $\tilde{f}$ is a linear combination of spherical harmonics with $l \leq 1$. If we compare this result with Eq. (161) on the preceding page, we find that the translations are inside the BMS group by means of the $\mathscr{C}^\infty(\mathbb{S}^2)$ factor. In fact, this $\mathscr{C}^\infty(\mathbb{S}^2)$ factor generalizes the translations by allowing transformations of the form

$$\xi^a = f(\zeta,\bar{\zeta}) \left(\frac{\partial}{\partial u}\right)^a, \tag{164}$$

where $f \in \mathscr{C}^\infty(\mathbb{S}^2)$, meaning f is now a linear combination of spherical harmonics with any value for l. This thus leads us to the concept of generalized translations, or supertranslations.

From this discussion, one may be prompted to conclude that the Poincaré group is a subgroup of the BMS group. While this is correct, there is an important caveat: there is not a preferred choice of Poincaré subgroup.

To be more clear, let us work with the Lorentz group. This is motivated by the fact that there is actually a preferred choice of translations, but there is no preferred choice of Lorentz subgroup of the BMS group. The "natural" choice of Lorentz subgroup of G_{BMS} is given by

$$\mathscr{L} = \left\{(\Lambda, 0) \in G_{\mathrm{BMS}}; \Lambda \in \mathrm{SO}^+(3,1)\right\}. \tag{165}$$

It is correct to say that $\mathscr{L} \cong SO^+(3,1)$. Hence, the Lorentz group is a subgroup of the BMS group. Nevertheless, let $g \in \mathscr{C}^\infty(\mathbb{S}^2)$. Noticing first that $(\mathbb{1}, g)^{-1} = (\mathbb{1}, -g)$, we point out that

$$(\mathbb{1}, -g) \odot (\Lambda, 0) \odot (\mathbb{1}, g) = (\Lambda, -g + (K_{\Lambda^{-1}} \circ \Lambda) \cdot (g \circ \Lambda)), \tag{166}$$

which in general is not an element of $\mathscr{L}$. Hence, $\mathscr{L}$ is not a normal subgroup. This means that while $\mathscr{L}$ is a copy of the Lorentz group inside the BMS group, so is $g^{-1}\mathscr{L}g$.

In some cases, this is not a deep statement. For example, consider Minkowski spacetime. We have a bulk Lorentz group, which induces symmetries at infinity. The induced group can then be chosen to be the correct Lorentz group. However, in the general case of an asymptotically flat spacetime we might not have bulk symmetries to choose which is the correct Lorentz group. Hence, both $\mathscr{L}$ and $g^{-1}\mathscr{L}g$ should be considered as possible Lorentz groups. It is impossible to prefer one over the other. This leads to difficulties in defining angular momentum at null infinity (Winicour 1980).

It is interesting to point out that the analogous problem does not occur for translations. Sachs 1962a has shown that the translations are in fact the unique four-dimensional normal subgroup of the BMS group. Therefore, although there is not a preferred definition of Lorentz transformations within the BMS group, there is a preferred definition of translations.

5.5 Possible Criticism of the Derivation of the BMS Group

Since our derivation of the BMS group has led us to an infinite-dimensional group, one could challenge some assumptions we made during the derivation. The purpose of this section is to consider two possible critiques of the derivation of the BMS group and argue that the results are, in fact, correct. Hence, we will try to point out possible problems with the arguments and show that these "problems" are in fact necessary to obtain the Poincaré group at infinity.

We begin by noticing that an appropriate definition of asymptotic symmetries at null infinity should reproduce *at least* the Poincaré group. Whatever symmetry group we have at infinity, it should have the Poincaré group as a subgroup. This is due to the fact that at null infinity we have "effectively" Minkowski spacetime, and the Poincaré group is the symmetry group of Minkowski spacetime. All Poincaré transformations should have an asymptotic version at null infinity, and thus should belong to the group of asymptotic symmetries. Hence, an asymptotic symmetry group that is smaller than the Poincaré group should be considered inadmissible on physical grounds.

The first possible criticism one could make to the derivations we provided is that we considered that $(\mathscr{I}^+, \tilde{h}_{ab}, n^a)$ was preserved only conformally. Hence, instead of using the equivalence relation

$$(\mathscr{I}^+, \tilde{h}_{ab}, n^a) \sim (\mathscr{I}^+, \omega^2 \tilde{h}_{ab}, \omega^{-1} n^a), \tag{167}$$

we should have used

$$(\mathscr{I}^+, \tilde{h}_{ab}, n^a) \sim (\mathscr{I}^+, \tilde{h}_{ab}, n^a), \tag{168}$$

which means we should consider only isometries, not conformal isometries. In other words, we should always work with $\omega = 1$.

If one goes back to our derivation an considers this new imposition, one will find this does not rule out the supertranslations. Rather, it only affects Lorentz transformations. Boosts, to be more precise. The only transformations with $\omega \neq 1$ are the Lorentz boosts, which become conformal transformations on the sphere at infinity. Hence, excluding conformal transformations seems to be too strong of a requirement, because it leads to a asymptotic symmetry group that does not have the Poincaré group as a subgroup. We conclude that we must, in fact, consider the conformal transformations, precisely as we did.

The second criticism one could make is that we imposed that the weak Carrollian structure should be preserved, but we could have imposed conservation of the strong Carrollian structure. As shown in Appendix A, it turns out that preserving the strong Carrollian structure rules out supertranslations, but it also rules out spatial translations. Hence, it is too strong of a requirement, and we cannot ask for the strong Carrollian structure to be preserved by the asymptotic symmetries without losing the Poincaré group in the process.

We conclude, therefore, that supertranslations seem to be a necessity if we want to enjoy the Poincaré group at infinity. They are a feature, not a bug.

6 Physical Consequences and Applications

Now that we know what the BMS group is we are ready to discuss some of its physical implications and applications. We have already mentioned the difficulty in defining angular momentum at null infinity in the previous section, but now we dive deeper in the physical consequences of BMS symmetries.

6.1 Physical Realization of Supertranslation

So far, BMS transformations, and supertranslations especially, may seem to be purely theoretical. After all, supertranslations are symmetries that are only available at infinity, and infinity is always far away. Now we discuss how a supertranslation can be physically realized, what what it means to do so.

To get some intuition, we begin by discussing boosts. Like supertranslations, boosts are diffeomorphisms. As a consequence, two spacetimes that differ by a boost (or by a supertranslation) are to be considered physically the same. There is no experiment that can distinguish between these two possibilities. Nevertheless, it is possible to physically realize a boost. In other words, one can "induce" a boost by means of a physical process.

Consider a spacetime comprised of a single massive particle of mass M. By Birkhoff's theorem one knows this spacetime is given by Schwarzschild spacetime. This is independent of whether the particle is "at motion" or "at rest": both solutions are physically equivalent, because they are diffeomorphic.

Consider, however, a spacetime as follows. We start with a single massive particle of mass M, but at advanced time v_0 a photon with energy E is sent in from infinity *en route* to a collision with the particle of mass M. Upon the collision, the massive particle absorbs the photon. Hence, for advanced time $v < v_0$ the solution is the Schwarzschild solution with mass M, but for $v > v_0$ the solution is the Schwarzschild solution with mass $M + E$. Furthermore, these two solutions are boosted relative to each other, because after the collision the massive particle has energy $M + E$ and also is moving relative to its initial state. In other words, it was accelerated at the instant of the collision. Notice that in this example the global spacetime is not Schwarzschild spacetime.

Similarly, one can induce a supertranslation by throwing in matter. Hawking, Perry, and Strominger (2017) considered throwing in a null shockwave with the stress-energy-momentum tensor having the form

$$T_{vv} = \frac{T(\zeta, \bar{\zeta})}{4\pi r^2} + \cdots, \tag{169}$$

where $T(\zeta, \bar{\zeta})$ is some function on the sphere and the dots indicate corrections that are necessary for the stress tensor to be conserved. The $l = 0$ component of T induces a change in mass, the $l = 1$ components induce a change in momentum, and the $l \geq 2$ components induce supertranslations.

6.2 Construction of Hadamard Vacua

There are four basic forces in our universe. Gravity, electromagnetism, and the strong and weak nuclear forces. While gravity is well-described by general relativity, the standard model of particle physics provides a good description of the remaining three (Schwartz 2014). However, while general relativity is a classical field theory, the standard model is a quantum field theory constructed on flat spacetime. Thus, in principle, general relativity cannot explain how the fields of the standard model fall, and the standard model knows nothing about gravity.

While a full description of gravity at the quantum level would require a knowledge of a full theory of quantum gravity, which at the present is only a dream, we can make do with an approximation. Quantum field theory in curved spacetime is the description of how quantum fields evolve upon a curved background spacetime. In this approximation, the background spacetime is assumed to be fixed, so that the quantum fields to not affect the background geometry.

An interesting approach to quantum field theory that is particularly useful in the curved spacetime formulation is known as the algebraic approach (see, *e.g.*, Aguiar Alves 2023; Dappiaggi, Moretti, and Pinamonti 2017; Fewster and Rejzner 2020; Hollands and Wald 2015; Wald 1994). In this approach, one takes a dual view of the theory. One part of the theory is described by the algebra of observables. This is the space of all observables to be considered in the theory and, as the name suggests, it has the structure of an algebra, which is a vector space with some extra special properties. On the other hand, we have the space of states, which are positive linear normalized functionals on the algebra of observables. It is usually easy to construct the algebra of observables, at least in the case of a non-interacting theory. However, even in the case of a non-interacting theory it is very hard to construct physically meaningful states.

By "physically meaningful" we mean that the state should allow for the expectation value of the stress-energy-momentum tensor to be well-defined. This is a non-trivial condition known as the Hadamard condition. It essentially states that in the ultraviolet limit any "physically meaningful" state should resemble the usual vacuum in Minkowski spacetime.

Dappiaggi, Moretti, and Pinamonti (2017) obtained a method to construct physically meaningful states by exploiting the BMS group. Namely, one does as follows. Assume a spacetime which is asymptotically flat at future null and timelike infinities (notice this requires the existence of the timelike infinity). Then one can show that the algebra describing observables in the bulk of the spacetime can be fitted in a one-to-one manner into the algebra describing observables at the boundary of spacetime, *i.e.*, at future null and timelike infinities. Hence, any observable in the bulk of the spacetime can be described at the boundary. One can then define a state at the boundary by imposing that it is a BMS-invariant state. Since the BMS group is infinite-dimensional, this is a strong restriction that singles out a state. This state can then be used to define an induced state in the bulk algebra, which turns out to the a Hadamard state. In other words, this construction induces a physically meaningful state.

The elementary idea is that one exploits the infinite-dimensional symmetry group at the boundary to define a state, which can then be pulled-back to the bulk. This construction is reviewed in detail by Dappiaggi, Moretti, and Pinamonti (2017).

6.3 Weinberg's Soft Graviton Theorem

Consider a scattering process in some collider experiment. One typically will collide n particles and produce m particles as a consequence. This is an $n \to m$ scattering process. A typical task in quantum field theory is to compute the cross section for such a scattering. However, there is an important remark. In the out state, there might be particles that will not be measured by the detectors of the experiment. For example, there could be gravitons or low energy photons. One should somehow take this particles into account. This is done by considering the so-called inclusive cross sections.

In order to prove that these inclusive cross sections are well-defined (and, more specifically, that they do not suffer of infrared divergences) one uses the so-called soft theorems. Our particular interest is in the Weinberg soft graviton theorem (Weinberg 1965). This theorem relates the scattering amplitude of the $n \to m$ process with the scattering amplitude of the $n \to m+k$ process that involves the emission of k soft gravitons, *i.e.*, of k gravitons with vanishingly small energy. The result is that the two amplitudes are proportional to each other, with the proportionality factor being known as the soft factor.

Relations between scattering amplitudes are often consequences of symmetries, and so is the case in this scenario. More specifically, this is the consequence of a conserved charge. Strominger (2014) argued that for the scattering problem to be well-defined in general relativity one should have a specific relation between the BMS transformations at past and future null infinity. This relation induces the conservation of infinitely many charges at infinity, each charge being related to a different supertranslation. Imposing this conservation of charges at the quantum level leads to a Ward identity—a dynamical consequence of the charge conservation. In the case of the BMS conserved charges, this Ward identity leads precisely to Weinberg's soft graviton theorem.

This application is reviewed in detail by Strominger (2018). Weinberg (1965, 1995) discusses the soft theorems and their applications to infrared divergences and inclusive cross sections.

6.4 Gravitational Memory Effect

The gravitational memory effect was originally discovered in linearized gravity by Zel'dovich and Polnarev (1974). It basically consists on the prediction that the passage of a gravitational wave permanently displaces the relative positions of two nearby inertial detectors. It turns out that this effect is too a reflection of BMS transformations.

A gravitational wave can be interpreted as a null shockwave such as the one on Eq. (169) on the previous page. Hence, the passage of a gravitational wave induces a supertranslation in the spacetime. We thus can understand the memory effect as being due to a supertranslation (Strominger and Zhiboedov 2016).

This application is interesting because the memory effect is particularly physical, and it is expected to be measurable by future gravitational wave detectors (Favata 2010; Grant and Nichols 2023). Hence, we might soon have an experimental test of supertranslations.

The relation between supertranslations and the memory effect is briefly discussed by Strominger (2018). The memory effect is reviewed, for example, by Bieri and Polnarev (2024).

6.5 Soft Hair on Black Holes

Finally, we discuss the results obtained by Hawking, Perry, and Strominger (2016, 2017) concerning black holes.

A known prediction in black hole thermodynamics is the information loss puzzle. Hawking (1974, 1975) originally noticed that the behavior of quantum fields on a spacetime containing a black hole formed by gravitational collapse leads the black hole to emit radiation. This radiation, now known as Hawking radiation, is thermal with temperature inversely proportional to the black hole's mass (in the case of a Schwarzschild black hole). This means that the black hole loses mass to the radiation bath, and the temperature rises as the black hole becomes smaller. This means the black hole loses mass even faster, until it eventually completely evaporates.

The complete evaporation argument assumes the calculations—which are performed using quantum field theory in curved spacetimes—to hold up to the last instants of evaporation. Hence, the argument cannot be fully trusted. Yet, it is interesting to consider its consequences.

The main puzzle is what is now known as the information loss puzzle, or information loss paradox. It was noticed by Hawking (1976) that if the quantum field starts at a pure state at very early times, it will be in a mixed state after the black hole evaporates. The basic idea is that the mixed state for the fields outside the black hole is a partial trace of the full pure state. However, once the black hole evaporates, the outer part is all that is left, and this is a mixed state. This means the time evolution of the quantum field state is not unitary.

This conclusion is seen by many physicists as troublesome, but there is nothing problematic with it in principle (Unruh and Wald 2017). It does involve problematic argument, though, due to the details of the operator algebras describing the quantum fields, but we will not dive into these details here.

What is mainly of interest for us is noticing that there are a few conserved quantities in the stellar collapse spacetime. In principle, one would expect all ten Poincaré charges and the total electric charge to be conserved. Hence, we have eleven "hairs" on the black hole characterizing what it was formed from. Nevertheless, this is still not enough information to capture everything that was going on in the spacetime at early times.

Hawking, Perry, and Strominger (2016, 2017) noticed, however, that supertranslations also induce conserved charges in the spacetime. The conservation of these supertranslation charges means there are infinitely many conserved charges at null infinity. This means much more information is preserved in the spacetime than previously thought.

It should be pointed out, though, that this does not mean all the information is preserved nor that the evolution is unitary. Supertranslations provide an important piece to understand the puzzle, but do not solve it completely.

A Preserving the Strong Carrollian Structure

In this appendix, we consider how preserving the strong Carrollian structure can affect the BMS group. For simplicity, we will from the start ignore the contributions due to Lorentz transformations, since we are interested in knowing whether preserving the strong Carrollian structure is enough to get rid of the supertranslations. This considerably simplifies the calculations, but will still allow us to conclude that the strong Carrollian structure is too strong to define asymptotic symmetry groups. This is due to the fact that, as we shall see, we end up losing the spatial translations with this definition.

Since we are working with only supertranslations, we are working from the start with isometries (not conformal isometries) that preserve the strong conformal geometry. Hence, we have that the vector fields ξ^a generating the supertranslations are such that the induced metric $\tilde{h}_{ab}$ and the normal vector n^a respect

$$£_\xi \tilde{h}_{ab} = 0 \quad \text{and} \quad £_\xi n^a = 0. \tag{170}$$

To impose the invariance of the strong Carrollian structure, we will also impose that ξ^a should be such that

$$£_\xi \tilde{\nabla}_a = 0. \tag{171}$$

If we can provide meaning to this equation (which can be done), it will be a natural condition to impose if we want the strong Carrollian structure to be preserved.

Here is how we will define the Lie derivative of the covariant derivative. Firstly, we notice that the Lie derivative can be thought of as a difference between two geometrical objects at a point. The difference between two covariant derivatives is a $(1,2)$-tensor, so the Lie derivative of the covariant derivative will be a $(1,2)$-tensor. That being said, we will define $£_\xi \tilde{\nabla}_a$ by demanding that

$$£_\xi(\tilde{\nabla}_a \psi^b) = (£_\xi \tilde{\nabla}_a)\psi^b + \tilde{\nabla}_a (£_\xi \psi^b). \tag{172}$$

Using Eq. (55) on page 14 and solving for $(£_\xi \tilde{\nabla}_a)\psi^b$ leads us to

$$(£_\xi \tilde{\nabla}_a)\psi^b = \xi^c(\tilde{\nabla}_c \tilde{\nabla}_a - \tilde{\nabla}_a \tilde{\nabla}_c)\psi^b + \psi^d \tilde{\nabla}_a \tilde{\nabla}_d \xi^b, \tag{173a}$$

$$= R_{acd}{}^b \xi^c \psi^d + \psi^d \tilde{\nabla}_a \tilde{\nabla}_d \xi^b, \tag{173b}$$

where the second step is merely an use of the definition of the Riemann tensor along with some simplification. Notice that $R_{abc}{}^d$ is the Riemann tensor associated with $\tilde{\nabla}_a$. We should mention that this calculation assumes implicitly that the covariant derivative is torsionless, as otherwise the torsion tensor would play a role in the equations above.

We want to impose Eq. (171). Thus, we can impose that $(£_\xi \tilde{\nabla}_a)\psi^b = 0$ for all vectors ψ^a. This leads to

$$R_{acd}{}^b \xi^c \psi^d + \psi^d \tilde{\nabla}_a \tilde{\nabla}_d \xi^b = 0 \tag{174}$$

for all vectors ψ^a. Hence, it must hold that

$$R_{acd}{}^b \xi^c + \tilde{\nabla}_a \tilde{\nabla}_d \xi^b = 0. \tag{175}$$

To solve this equation, it is convenient to pick $\tilde{\nabla}_a = D_a$, the Levi-Civita connection on the sphere. This implies that $R_{abc}{}^d$ contracted with n^a in any of its indices will vanish. Recalling that the general expression for a vector field generating a supertranslation is (Eq. (161) on page 38)

$$\xi^a = f n^a, \tag{176}$$

where $f \in \mathscr{C}^\infty(\mathbb{S}^2)$, we get that

$$f R_{acd}{}^b n^c + D_a D_d (f n^b) = 0. \tag{177}$$

It then follows that

$$D_a D_b f = 0. \tag{178}$$

Using stereographic coordinates, we conclude using Eq. (63) on page 17 that this equation corresponds to imposing that

$$\partial_\zeta \partial_\zeta f = -\frac{2\bar{\zeta}}{1+\zeta\bar{\zeta}} \partial_\zeta f, \tag{179a}$$

$$\partial_{\bar{\zeta}} \partial_{\bar{\zeta}} f = -\frac{2\zeta}{1+\zeta\bar{\zeta}} \partial_{\bar{\zeta}} f, \tag{179b}$$

$$\partial_\zeta \partial_{\bar{\zeta}} f = 0, \tag{179c}$$

$$\partial_{\bar{\zeta}} \partial_\zeta f = 0. \tag{179d}$$

Denote $f_\zeta = \partial_\zeta f$ and $f_{\bar{\zeta}} = \partial_{\bar{\zeta}} f$. Then Eq. (179d) tells us that f_ζ is an analytic function of ζ, and, equivalently, bears no dependence on $\bar{\zeta}$. Thus, the left-hand side of Eq. (179a) does not depend on $\bar{\zeta}$, but the right-hand side does. This is only possible if both of them vanish. Hence, $f_\zeta = 0$. Similarly, $f_{\bar{\zeta}} = 0$. We thus get to the system of differential equations

$$\begin{cases} \partial_\zeta f = 0, \\ \partial_{\bar{\zeta}} f = 0, \end{cases} \tag{180}$$

which is solved by $f = \text{constant}$.

Therefore, imposing the preservation of the strong Carrollian structure imposes that the function f is a constant. On the one hand, this does rule out the supertranslations. On the other hand, it also rules out spatial translations. We thus conclude the strong Carrollian structure is way too strong to be imposed on the asymptotic symmetry group.

References

Aguiar Alves, Níckolas de (2023). "Nonperturbative Aspects of Quantum Field Theory in Curved Spacetime". MSc thesis. Santo André, Brazil: Federal University of ABC. xxiv, 152 pp. arXiv: 2305.17453 [gr-qc] (cit. on p. 42).

Bieri, Lydia and Alexander Polnarev (2024). "Gravitational Wave Displacement and Velocity Memory Effects". *Classical and Quantum Gravity* 41, 135012. DOI: 10.1088/1361-6382/ad4dfe. arXiv: 2402.02594 [gr-qc] (cit. on p. 44).

Bondi, Hermann, M. G. J. Van der Burg, and A. W. K. Metzner (1962). “Gravitational Waves in General Relativity, VII. Waves from Axi-Symmetric Isolated System”. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 269.1336, pp. 21–52. DOI: 10.1098/rspa.1962.0161 (cit. on pp. 2, 36, 39).

Choquet-Bruhat, Yvonne (2015). *Introduction to General Relativity, Black Holes, and Cosmology*. Oxford: Oxford University Press (cit. on p. 24).

Dappiaggi, Claudio, Valter Moretti, and Nicola Pinamonti (2017). *Hadamard States From Light-like Hypersurfaces*. SpringerBriefs in Mathematical Physics 25. Cham: Springer. DOI: 10.1007/978-3-319-64343-4. arXiv: 1706.09666 [math-ph] (cit. on pp. 2, 28, 42, 43).

Duval, C., G. W. Gibbons, and P. A. Horvathy (2014a). “Conformal Carroll Groups and BMS Symmetry”. *Classical and Quantum Gravity* 31.9, 092001. DOI: 10.1088/0264-9381/31/9/092001. arXiv: 1402.5894 [gr-qc] (cit. on p. 32).

——— (2014b). “Conformal Carroll Groups”. *Journal of Physics A: Mathematical and Theoretical* 47.33, 335204. DOI: 10.1088/1751-8113/47/33/335204. arXiv: 1403.4213 [hep-th] (cit. on p. 32).

Duval, C., G. W. Gibbons, P. A. Horvathy, and P. M. Zhang (2014). “Carroll versus Newton and Galilei: Two Dual Non-Einsteinian Concepts of Time”. *Classical and Quantum Gravity* 31.8, 085016. DOI: 10.1088/0264-9381/31/8/085016. arXiv: 1402.0657 [gr-qc] (cit. on p. 32).

Favata, Marc (2010). “The Gravitational-Wave Memory Effect”. *Classical and Quantum Gravity* 27.8, 084036. DOI: 10.1088/0264-9381/27/8/084036. arXiv: 1003.3486 [gr-qc] (cit. on pp. 2, 44).

Fewster, Christopher J. and Kasia Rejzner (2020). “Algebraic Quantum Field Theory: An Introduction”. In: *Progress and Visions in Quantum Theory in View of Gravity: Bridging Foundations of Physics and Mathematics*. Ed. by Felix Finster et al. Cham: Springer, pp. 1–61. DOI: 10.1007/978-3-030-38941-3_1. arXiv: 1904.04051 [hep-th] (cit. on p. 42).

Geroch, Robert (1985). *Mathematical Physics*. Chicago Lectures in Physics. Chicago: The University of Chicago Press (cit. on p. 9).

Goldstein, Herbert (1980). *Classical Mechanics*. 2nd ed. Addison-Wesley Series in Physics. Reading, Massachussets: Addison-Wesley (cit. on p. 5).

Grant, Alexander M. and David A. Nichols (2023). “Outlook for Detecting the Gravitational-Wave Displacement and Spin Memory Effects with Current and Future Gravitational-Wave Detectors”. *Physical Review D* 107.6, 064056. DOI: 10.1103/PhysRevD.107.064056. arXiv: 2210.16266 [gr-qc] (cit. on pp. 2, 44). Erratum: “Erratum: Outlook for Detecting the Gravitational-Wave Displacement and Spin Memory Effects with Current and Future Gravitational-Wave Detectors [Phys. Rev. D 107 , 064056 (2023)]”. *Physical Review D* 108.2, 029901 (2023). DOI: 10.1103/PhysRevD.108.029901.

Hawking, Stephen W. (1974). “Black Hole Explosions?” *Nature* 248, pp. 30–31. DOI: 10.1038/248030a0 (cit. on p. 44).

——— (1975). “Particle Creation by Black Holes”. *Communications In Mathematical Physics* 43.3, pp. 199–220. DOI: 10.1007/BF02345020 (cit. on p. 44).

——— (1976). “Breakdown of Predictability in Gravitational Collapse”. *Physical Review D* 14.10, pp. 2460–2473. DOI: 10.1103/PhysRevD.14.2460 (cit. on p. 44).

Hawking, Stephen W. and George F. R. Ellis (1973). *The Large Scale Structure of Spacetime*. Cambridge Monographs on Mathematical Physics. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511524646 (cit. on pp. 10, 23).

Hawking, Stephen W., Malcolm J. Perry, and Andrew Strominger (2016). "Soft Hair on Black Holes". *Physical Review Letters* 116.23, 231301. DOI: 10.1103/PhysRevLett.116.231301. arXiv: 1601.00921 [hep-th] (cit. on p. 44).

——— (2017). "Superrotation Charge and Supertranslation Hair on Black Holes". *Journal of High Energy Physics* 2017.5, 161. DOI: 10.1007/JHEP05(2017)161. arXiv: 1611.09175 [hep-th] (cit. on pp. 42, 44).

Hollands, Stefan and Robert M. Wald (2015). "Quantum Fields in Curved Spacetime". *Physics Reports* 574, pp. 1–35. DOI: 10.1016/j.physrep.2015.02.001. arXiv: 1401.2026 [gr-qc] (cit. on p. 42).

Lee, John (2012). *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics 218. New York, NY: Springer. DOI: 10.1007/978-1-4419-9982-5 (cit. on pp. 10, 12).

Lévy-Leblond, Jean-Marc (1965). "Une Nouvelle Limite Non-Relativiste Du Groupe de Poincaré". *Annales de l'institut Henri Poincaré. Section A, Physique Théorique* 3.1, pp. 1–12 (cit. on p. 32).

Newman, Ezra T. and Theodore W. J. Unti (1962). "Behavior of Asymptotically Flat Empty Spaces". *Journal of Mathematical Physics* 3.5, pp. 891–901. DOI: 10.1063/1.1724303 (cit. on p. 35).

Penrose, Roger (1963). "Asymptotic Properties of Fields and Space-Times". *Physical Review Letters* 10.2, pp. 66–68. DOI: 10.1103/PhysRevLett.10.66 (cit. on p. 25).

——— (1965). "Zero Rest-Mass Fields Including Gravitation: Asymptotic Behaviour". *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 284.1397, pp. 159–203. DOI: 10.1098/rspa.1965.0058 (cit. on p. 25).

——— (1974). "Relativistic Symmetry Groups". In: *Group Theory in Non-Linear Problems*. Ed. by A. O. Barut. Nato Advanced Study Institutes Series 7. Dordrecht: Springer, pp. 1–58. DOI: 10.1007/978-94-010-2144-9_1 (cit. on p. 33).

Penrose, Roger and Wolfgang Rindler (1986). *Spinors and Space-Time*. Vol. 2: *Spinor and Twistor Methods in Space-Time Geometry*. Cambridge Monographs on Mathematical Physics. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511524486 (cit. on p. 33).

Robinson, Derek J. S. (1996). *A Course in the Theory of Groups*. Graduate Texts in Mathematics 80. New York, NY: Springer. DOI: 10.1007/978-1-4419-8594-1 (cit. on p. 10).

Sachs, R. K. (1962a). "Asymptotic Symmetries in Gravitational Theory". *Physical Review* 128.6, pp. 2851–2864. DOI: 10.1103/PhysRev.128.2851 (cit. on p. 40).

——— (1962b). "Gravitational Waves in General Relativity VIII. Waves in Asymptotically Flat Space-Time". *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 270.1340, pp. 103–126. DOI: 10.1098/rspa.1962.0206 (cit. on pp. 2, 36, 39).

Schmidt, B. and M. Walker (1983). "Analytic Conformal Extensions of Asymptotically Flat Space-Times". *Journal of Physics A: Mathematical and General* 16.10, pp. 2187–2190. DOI: 10.1088/0305-4470/16/10/015 (cit. on p. 31).

Schwartz, Matthew D. (2014). *Quantum Field Theory and the Standard Model*. Cambridge: Cambridge University Press. DOI: 10.1017/9781139540940 (cit. on pp. 7, 42).

Strominger, Andrew (2014). “On BMS Invariance of Gravitational Scattering”. *Journal of High Energy Physics* 2014.7, 152. DOI: 10.1007/JHEP07(2014)152. arXiv: 1312.2229 [hep-th] (cit. on p. 43).

——— (2018). *Lectures on the Infrared Structure of Gravity and Gauge Theory*. Princeton: Princeton University Press. arXiv: 1703.05448 [hep-th] (cit. on pp. 2, 43, 44).

Strominger, Andrew and Alexander Zhiboedov (2016). “Gravitational Memory, BMS Supertranslations and Soft Theorems”. *Journal of High Energy Physics* 2016.1, 86. DOI: 10.1007/JHEP01(2016)086. arXiv: 1411.5745 [hep-th] (cit. on p. 43).

Tu, Loring W. (2011). *An Introduction to Manifolds*. 2nd ed. Universitext. New York, NY: Springer. DOI: 10.1007/978-1-4419-7400-6 (cit. on pp. 10, 12, 28).

Unruh, William G. and Robert M. Wald (2017). “Information Loss”. *Reports on Progress in Physics* 80.9, 092002. DOI: 10.1088/1361-6633/aa778e. arXiv: 1703.02140 [hep-th] (cit. on p. 44).

Wald, Robert M. (1984). *General Relativity*. Chicago: The University of Chicago Press (cit. on pp. 2, 10, 12, 14, 20, 28–30, 33).

——— (1994). *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. Chicago Lectures in Physics. Chicago: The University of Chicago Press (cit. on p. 42).

Weinberg, Steven (1965). “Infrared Photons and Gravitons”. *Physical Review* 140 (2B), B516–B524. DOI: 10.1103/PhysRev.140.B516 (cit. on pp. 2, 43).

——— (1995). *The Quantum Theory of Fields*. Vol. I: *Foundations*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139644167 (cit. on pp. 2, 43).

Winicour, Jeffrey (1980). “Angular Momentum in General Relativity”. In: *General Relativity and Gravitation: One Hundred Years After the Birth of Albert Einstein*. Ed. by Alan Held. Vol. 2. New York, NY: Plenum Press, pp. 71–96 (cit. on p. 40).

Zee, Anthony (2016). *Group Theory in a Nutshell for Physicists*. Princeton: Princeton University Press (cit. on pp. 1, 5).

Zel’dovich, Yakov B. and A. G. Polnarev (1974). “Radiation of Gravitational Waves by a Cluster of Superdense Stars”. *Soviet Astronomy* 18, p. 17 (cit. on pp. 2, 43).