Decay in Outgoing Null Directions of Solutions of the Massive Dirac Equation in certain Asymptotically Flat, Static Spacetimes

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr. rer. nat.)
der Fakultät für Mathematik
der Universität Regensburg

vorgelegt von

Jan-Hendrik Treude

aus Reutlingen

Dezember 2014

Promotionsgesuch eingereicht am 01.12.2014

Die Arbeit wurde angeleitet von Prof. Dr. Felix Finster.

Prüfungsausschuss:	Vorsitzender:	Prof. Dr. Bernd Ammann
	1. Gutachter:	Prof. Dr. Felix Finster
	2. Gutachter:	Prof. Dr. Jean-Philippe Nicolas (Université Brest, France)
	weiterer Prüfer:	Prof. Dr. Stefan Friedl

Kurzzusammenfassung

Gegenstand der vorliegenden Arbeit ist die massive Dirac Gleichung auf gewissen statischen, asymptotisch flachen Lorentz-Mannigfaltigkeiten. Mittels analytischer Methoden für hyperbolische Gleichungen wird untersucht, wie sich Lösungen dieser Gleichung im Unendlichen der Mannigfaltigkeit, genauer im lichtartigen Unendlichen, verhalten. Spezieller geht es um die Fragestellung, inwiefern die Amplitude einer Lösung im lichtartigen Unendlichen gegen Null konvergiert.

Die genaue Analyse verläuft unter der Einschränkung auf glatte Lösungen der Dirac Gleichung, deren Träger räumlich kompakt ist, sowie der Annahme, dass die zugrunde liegende Raumzeit statisch und asymptotisch flach ist. Außerdem ist es essentiell, dass die Masse in der Dirac-Gleichung ungleich Null ist. Für solche Lösungen wird dann bewiesen, dass ihre Amplitude schneller als jede Potenz der auslaufenden lichtartigen Koordinate gegen Null konvergiert wenn diese Koordinate gegen unendlich strebt.

Eine spezielle Herangehensweise in dieser Arbeit ist, die Dirac Gleichung direkt in angepassten lichtartigen Koordinaten zu untersuchen. Dies führt unmittelbar zu Methoden, die mit dem charakteristischen Anfangswertproblem oder auch Goursat-Problem der Dirac Gleichung verwandt sind, sowie zu Energieabschätzungen in Gebieten mit lichtartigen Rändern.

Insgesamt gesehen ist die Herangehensweise eine störungstheoretische, nahegelegt durch die asymptotische Flachheit der zugrunde liegenden Raumzeit: Zunächst wird die entsprechende Fragestellung im flachen Minkowskiraum bearbeitet, wo sich die Dirac Gleichung mittels Greenscher Funktionen lösen lässt. Anschließend wird darauf aufbauend die eigentliche Problemstellung mittels Störungstheorie behandelt. Dabei spielen Energieabschätzungen, speziell die Ausnutzung der sogenannten Stromerhaltung der Dirac Gleichung, eine wesentliche Rolle.

Abstract

The topic of the thesis at hand is the massive Dirac equation on certain asymptotically flat, static spacetimes. Using analytical methods for hyperbolic equations, the behaviour of solutions of this equation at infinity of the manifold is analyzed, more precisely at so-called lightlike infinity. Specifically, it is studied to what extent solutions decay to zero at lightlike infinity.

The concrete analysis is restricted to smooth and spatially compactly supported solutions, and is always under the general assumption that the underlying spacetime is static and asymptotically flat. Moreover, it is absolutely essential that the mass in the Dirac equation is nonzero. It is then proved that such solutions decay to zero faster than any inverse power of the outgoing null coordinate as that coordinate tends to infinity.

A particular approach taken in this thesis is to analyze the Dirac equation directly in adapted lightlike coordinates. This immediately leads to methods related to the characteristic initial value problem or Goursat problem for the Dirac equation, and to the use of energy estimates in domains with lightlike boundaries.

The general approach is to proceed by perturbation theoretic arguments, as is suggested by the asymptotic flatness of the underlying spacetime: First the question of decay is addressed in flat Minkowski spacetime, for which the Dirac equation can be solved rather explicitly by Green's function methods. Building on this, the actual equation is then treated by a perturbation argument. In this argument, energy estimates play a crucial role, especially in form of the so-called conserved current of the Dirac equation.

Acknowledgements

In the following I would like to express my gratitude to people who played an important role during the time I worked on this thesis.

Most of all I want to thank my doctoral advisor Professor Felix Finster for his constant encouragement and for taking much time helping me throughout the various phases of my doctoral studies. I greatly benefited from his experience in many ways, not only concerning technical aspects, and our conversations about (mathematical) research in general, life as a researcher, and the high and low side of doing research were very valuable to me. I also want to thank him for being the gentle and kind person he is, this made me feel very welcome in Regensburg and in his research group from the first day on.

Next, I want to thank Professor Jean-Philippe Nicolas for his interest in my work and the many useful inputs he gave me during his stay in Regensburg and the other occasions we met. I especially want to thank him for inviting me to Brest and taking the time to work with me and to show me the beautiful surroundings. Being able to discuss and exchange ideas with him was a big motivation and inspiration for me.

Then there are of course many different people in Regensburg who helped me during the last years and whom I want to thank.

I want to thank my secondary advisor Professor Bernd Ammann for his comments about my work, and generally for being there as another experienced person to talk to.

Special thanks go to my very pleasant office mate Olaf Müller for helping me with various questions about Lorentzian geometry, spin geometry etc., for proof-reading parts of this thesis, and of course for our many interesting discussions.

I am grateful to Johannes Kleiner for being an enrichment in many different ways, and for his constant reminder of the human side of doing science. I also want to thank him for proof-reading various parts of this thesis.

To my spin geometry learning companion Nikolai Nowaczyk I want to say thank you for exploring parts of these intriguing mathematical objects together with me.

I also want to thank Nicolas Ginoux for always being willing to answer my various "beginners questions" about spin geometry over the years.

Moreover, I want to thank all other present and past members of the Mathematical Physics working group whom I got to know over the years – Andreas Grotz, Simone Murro, Moritz Reintjes, Christian Röken, and Daniela Schiefeneder – for their input and for giving our group seminars and other activities such a sympathetic atmosphere.

Maybe even more importantly, let me thank all the people who have prevented me from going crazy at times while doing mathematics (or trying to). Above all Rebekka.

Last but not least, I gratefully acknowledge the financial and academic support by the DFG-Graduiertenkolleg GRK 1692 "Curvature, Cycles, Cohomology" during the last three years.

To true science,

and all people pursuing honest research interests.

Contents

Introduction

The main topic of this thesis is asymptotic behaviour of solutions of the massive Dirac equation on a Lorentzian manifold. More specifically, for a certain class of asymptotically flat spacetimes with a well-defined notion of outgoing null geodesics, the following question is studied:

How do solutions of the massive Dirac equation behave (decay) along these outgoing null geodesics as one moves out to infinity?

This introduction contains an overview of the thesis, as well as some information about the motivation behind studying the above question and relations to other topics. We begin by describing the main results of the thesis.

1. What are the Main Results of this Thesis?

The core part of the thesis consists in studying decay properties of solutions of the massive Dirac equation on spacetimes (M, g) which have the form

$$\begin{cases} M = \mathbb{R}_t \times (r_0, \infty)_r \times N \\ g = (1 + A(r))^2\big[-\mathrm{d}t^2 + \mathrm{d}r^2\big] + R(r)^2 g_N \end{cases} . \tag{$*$}$$

Here (N, g_N) is a compact Riemannian spin manifold and $A, R \in C^\infty(r_0, \infty)$, for which we assume that $R, 1 + A > 0$. One may notice that any such spacetime is a warped product over the 1+1 dimensional base $Q = \mathbb{R}_t \times (r_0, \infty)_r$, $g_Q = (1 + A(r))^2[-\mathrm{d}t^2 + \mathrm{d}r^2]$, with Riemannian fibre (N, g_N), and warping function R. The "dynamically relevant" information is contained in (Q, g_Q) and R.

One can visualize such a spacetime as in figure 1 by a Penrose diagram for the Q-part. Notice to this end that (Q, g_Q) is conformally equivalent to a vertical strip $\mathbb{R}_t \times (r_0, \infty)_r$ of Minkowski spacetime, and therefore has the same Penrose diagram.[1] As the picture correctly suggests, (M, g) is globally hyperbolic if and only if $r_0 = -\infty$. For the purpose of summarizing the results we focus on this case since it simplifies the statements while keeping the main point.

Besides being of this special form, we further assume that for some $r_m > r_0$ our spacetimes (M, g) satisfy the following two asymptotic boundedness and decay conditions:

i.) For any $k \in \mathbb{N}$ we have

$$\|A(r)\|_{C^k(r_m,\infty)}\,, \|R^{-1}\|_{C^k(r_m,\infty)} < \infty\,.$$

ii.) There exists a constant $C > 0$ and some $\alpha > 0$ such that

$$|A(r)|, |A'(r)|, |R^{-1}(r)|, |(R^{-1})'(r)| \leq \frac{C}{(1+r)^\alpha} \qquad \forall r > r_m\,.$$

[1] To be precise, we claim nothing about regularity of the conformally attached boundary. We only use the picture for the purpose of illustration, in particular of the null geodesics and the causal structure.

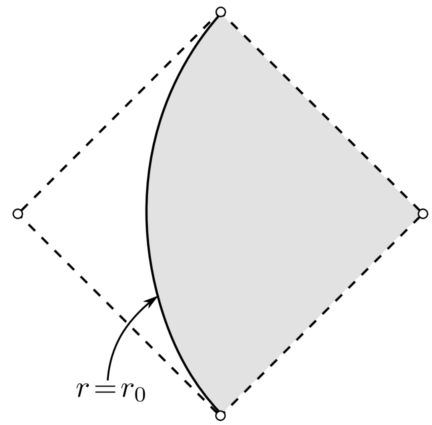

FIGURE 1. Visualization of a spacetime as considered in this thesis.

A well-known example of a spacetime which satisfies all these conditions is the *exterior Schwarzschild spacetime* (cf. Example 3.2.2).

Concerning spinors, there is a correspondence between spin structures on M and on N such that the spinor bundle $\mathbb{S}M$ of M can be identified with the (pullback to M of the) direct sum $\mathbb{S}N \oplus \mathbb{S}N$ of two copies of the spinor bundle of N. The Dirac operator then has the "block form"

$$\mathcal{D}_M = \mathrm{i}\mathrm{e}^{-a} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \partial_t + \mathrm{i}\mathrm{e}^{-a} \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \left(\partial_r + \frac{a'}{2} + \frac{n-1}{2} \frac{R'}{R} \right) + \frac{\mathrm{i}}{R} \begin{pmatrix} 0 & \mathcal{D}_N \\ \mathcal{D}_N & 0 \end{pmatrix}.$$

Here $a \in C^\infty(r_0, \infty)$ is defined by the identity $\mathrm{e}^{a(r)} = 1 + A(r)$. Using this formula it is possible to "split off" the explicit N-dependence by a separation of variables argument, which allows to restrict ourselves to the 1+1 dimensional part Q (with some "potential" reflecting the N-part).

The main result stated below is concerned with decay of solutions of the massive Dirac equation. Since these are sections of a vector bundle one cannot directly speak of them being small or large. To do so, we use the following inner product: First, the spinor bundle $\mathbb{S}M$ is equipped with a non-degenerate, but indefinite Hermitian inner product $\prec \cdot, \cdot \succ_{\mathbb{S}M}$ in the fibers. Being indefinite this inner product is not well suited to capture the "size" of spinors. However, making use of the future-pointing unit-length timelike vector field $T = \mathrm{e}^{-a}\partial_t$ on M, the inner product

$$\langle \cdot, \cdot \rangle_T := \prec \cdot, \gamma(T) \cdot \succ_{\mathbb{S}M}$$

is positive definite. Here $\gamma(T)$ denotes Clifford multiplication by T. We use this inner product to measure the size of spinors.

Finally, in the analysis of the resulting equation on Q (which takes up most of the work), we make crucial use of the *null coordinates*

$$v = t + r \quad \text{and} \quad u = t - r\,.$$

The main result is concerned with decay of solutions of the massive Dirac equation as the coordinate v tends to infinity, while the coordinate u is restricted to some finite range (concerning the N-dependence everything is uniform). Graphically speaking, we are thus making an estimate in a "null strip" as illustrated in figure 2. The precise result, in a version which is simple to understand, is as follows (cf. Corollary 4.7.5):

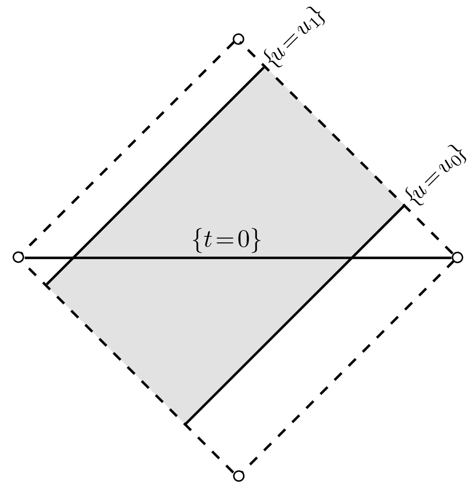

FIGURE 2. Illustration of a null strip in which the estimate takes place.

Theorem. (Superpolynomial decay in outgoing null directions) *Let (M, g) be of the form $(*)$ with $r_0 = -\infty$, and assume that it satisfies the asymptotic conditions i.) and ii.). Fix a nonzero mass $m \neq 0$. Then for every $k \in \mathbb{N}$ and every $u_0 < u_1$ there exists a constant $C_k > 0$ and a number $s(k) \in \mathbb{N}$ such that the following holds: Let $\psi \in \Gamma^\infty_{sc}(\mathbb{S}M)$ satisfy $(\mathcal{D}-m)\psi = 0$ and $\operatorname{supp} \psi|_{t=0} \subset (-u_1, -u_0) \times N$. Then it holds that*

$$|\psi(v, u, \omega)|_T \leq \frac{C_k}{(1+v)^k} \, \|\psi|_{t=0}\|_{H^{s(k)}(\Sigma)} \tag{$**$}$$

for all $v > 0$, $u \in [u_0, u_1]$, $\omega \in N$. Both C_k and $s(k)$ depend on k, the constants in i.) and ii.), and C_k also depends on $|u_0 - u_1|$ and m.

Expressed in words, this result states that smooth, spatially compactly supported solutions of the massive Dirac equation decay as fast as any inverse power of v along outgoing null geodesics.

Let us say a few words about the methods which are used to obtain this result. The basic strategy is a *perturbative* one: First we explicitly solve a model equation derived from our actual equation of interest (the "free part"), whose solutions can easily be shown to decay as claimed in $(**)$. Afterwards, in the second, technically more demanding step we show that the difference between the model equation and the actual equation (the "perturbation" or "error terms") does not affect the behaviour $(**)$. More concretely, to carry out these steps the following general methods are used:

- For the free part, we use *Green's function methods* to derive an explicit integral representation formula for solutions of the model equation in terms of their (characteristic) initial data.
- To control the perturbation, we use *energy estimates* in domains with lightlike boundaries, using the so-called "conserved current" of the Dirac equation. Here we also need to make use of the boundedness and decay assumptions of the metric stated earlier.
- To tie together the previous two parts, we use the *Lippmann-Schwinger equation* (or *Duhamel formula*), see Section 4.6.1 and 4.6.2 for an explanation.

A more detailed outline is given below on page xvii of this introduction in the summary of the fourth chapter.

2. What is the Motivation behind this Question?

The starting point for the work on this thesis was the article [**MN04**] by L. Mason and J.-P. Nicolas on *conformal scattering* in asymptotically simple spacetimes, which again builds upon ideas going back at least as far as to Penrose's article [**Pen65**]. To explain this motivation, we start with a brief summary of [**MN04**].[2]

Conformal scattering in asymptotically simple spacetimes. In [**MN04**] the authors concentrate on globally hyperbolic spacetimes (M, g) with the following property: One assumes that there exists an open, bounded, conformal embedding $(M, g) \hookrightarrow (\widetilde{M}, \widetilde{g})$ into a larger globally hyperbolic spacetime $(\widetilde{M}, \widetilde{g})$ such that the following properties hold:

(AS1) Viewing M as subset of $\widetilde{M}$, its topological boundary has the structure
$$\partial M = \{\iota^+\} \dot\cup \mathscr{I}^+ \dot\cup \{\iota^0\} \dot\cup \mathscr{I}^- \dot\cup \{\iota^-\}\,,$$
where $\iota^+, \iota^-, \iota^0 \in \widetilde{M}$ such that
  (a) $\partial \widetilde{I}^-(\iota^+) \cap \partial \widetilde{I}^+(\iota^-) = \{\iota^0\}$, where $\widetilde{I}^\pm$ denote the causal future and past in $(\widetilde{M}, \widetilde{g})$,
  (b) $\mathscr{I}^+ := \partial \widetilde{I}^-(\iota^+) \setminus \{\iota^+, \iota^0\}$ and $\mathscr{I}^- := \partial \widetilde{I}^+(\iota^-) \setminus \{\iota^-, \iota^0\}$ are two smooth null hypersurfaces of $(\widetilde{M}, \widetilde{g})$.

(AS2) The conformal factor Ω, i.e. the smooth positive function on M with $\widetilde{g} = \Omega^2 g$, extends smoothly to all of $\widetilde{M}$. Moreover, $\Omega|_{\partial M} = 0$ and $\mathrm{d}\Omega|_{\mathscr{I}^+ \cup \mathscr{I}^-}$ vanishes nowhere, hence Ω is a boundary defining function for the smooth part of ∂M.

(AS3) Every future-directed, inextendible null geodesic in (M, g) acquires a future endpoint on $\mathscr{I}^+$ and a past endpoint on $\mathscr{I}^-$.

The points ι^+ and ι^- are called *future timelike infinity* and *past timelike infinity*, the point ι^0 is called *spacelike infinity*, and the null hypersurfaces $\mathscr{I}^+$ and $\mathscr{I}^-$ are called *future null infinity* and *past null infinity*. Spacetimes (M, g) satisfying (A1)–(A3) are called *(smoothly) asymptotically simple*, and one should think of ∂M as a conformally attached boundary at infinity to such a spacetime.

The basic example of a spacetime which satisfies these conditions is Minkowski spacetime $\mathbb{R}^{1,n}$, for which the so-called *Penrose conformal compactification* provides the desired conformal embedding (see Appendix A.2). For this particular example the structure of the conformal embedding is illustrated in figure 3, where the angular coordinates are suppressed so that each point in the diagram should be thought of as representing an $(n-1)$–sphere. For a general asymptotically simple spacetime one should keep the same picture in mind. Besides Minkowski spacetime, other examples have been constructed by work of Friedrich, Corvino, Schoen, Chrusciel, and Delay (see [**CD02**] and references therein).

Now we come to the main point of [**MN04**]. Namely, if (M, g) satisfies the conditions (A1)–(A3), one can use this to study the asymptotic behaviour of *conformally invariant equations* (or better, their solutions) towards the conformal boundary ∂M in the following way: First of all, recall that a linear equation[3] $P_g \phi = 0$ on a vector bundle E over a

[2]For the sake of readability and in order not to distract from the main point to be made, the presentation in the following is a bit simplified concerning some technical aspects. In case of interest we refer the reader to the original article [**MN04**].

[3]Linearity is not crucial here, and conformal scattering constructions for nonlinear equations were carried out for instance in [**BSZ90**] and [**Jou10, Jou12**]. The analysis of course becomes more difficult then.

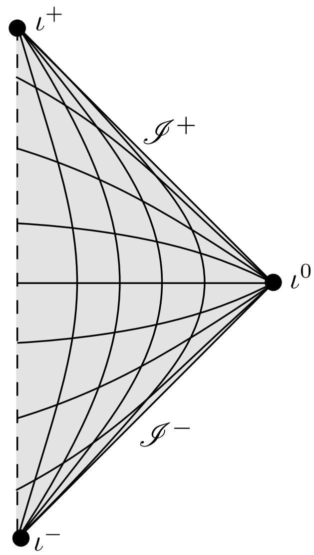

FIGURE 3. The schematic picture of an asymptotically simple spacetime, thought of as being conformally embedded into $1+1$ dimensional Minkowski spacetime. The horizontal lines represent Cauchy surfaces of "constant time", the vertical lines hypersurfaces of "constant radius". Since the embedding is conformal, (radial) null geodesics are still given by straight lines at 45 degree inclination angle.

Lorentzian, or more generally pseudo-Riemannian manifold (M, g) is said to be *conformally invariant* if for every conformal factor $\Omega \in C^\infty(M)$ there exists a transformation T_Ω of the fields $\phi \in \Gamma^\infty(E)$ such that

$$P_g(\phi) = 0 \quad \Longleftrightarrow \quad P_{\Omega^2 g}(T_\Omega \phi) = 0\,.$$

In other words, T_Ω identifies the spaces of solutions of the two equations $P_g\phi = 0$ and $P_{\Omega^2 g}\psi = 0$. In most examples $T_\Omega\phi = \Omega^{-\alpha}\phi$ for some $\alpha > 0$. Some examples of conformally invariant equations for which this is the case are the Maxwell equations, the massless Dirac equation, and the conformally coupled wave equation. These are precisely the three equations considered in [**MN04**], so in the following we let $P_g\phi = 0$ denote either of these equations.

Besides conformal invariance, another key property of these three equations is that they have a well-posed Cauchy problem if (M, g) is globally hyperbolic. Furthermore, they obey finite propagation speed according to the causality of g. If (M, g) satisfies (A1)–(A3), then one can combine these properties with conformal invariance of the equation $P_g\phi = 0$ to smoothly extend any (rescaled) solutions of the equation $P_g\phi = 0$ to the larger spacetime $(\widetilde{M}, \widetilde{g})$ in the following way:

Let $\phi \in \Gamma^\infty_{sc}(E)$ be a smooth, spatially compactly supported solution of $P_g\phi = 0$. Then $T_\Omega\phi = \Omega^{-\alpha}\phi$ satisfies $P_{\widetilde{g}}(T_\Omega\phi) = 0$ since $\widetilde{g}|_M = \Omega^2 g$. Next, pick some Cauchy surface $\widetilde{\Sigma} \subset \widetilde{M}$ which intersects M and extend $(T_\Omega\phi)|_{\widetilde{\Sigma}\cap M}$ smoothly to all of $\widetilde{\Sigma}$ (this is possible since ϕ has spatially compact support). By global hyperbolicity of $\widetilde{M}$, one can solve the Cauchy problem of the equation $P_{\widetilde{g}}\widetilde{\phi} = 0$ in $\widetilde{M}$, taking as initial data the extension of $(T_\Omega\phi)|_{\widetilde{\Sigma}\cap M}$ to $\widetilde{\Sigma}$. Let us call the corresponding solution $\widetilde{T_\Omega\phi} \in \Gamma^\infty(\widetilde{E})$.

Inside the original spacetime M it then holds that $P_{\widetilde{g}}T_\Omega\phi = 0$ and $P_{\widetilde{g}}\widetilde{T_\Omega\phi} = 0$, and since $T_\Omega\phi$ and $\widetilde{T_\Omega\phi}$ coincide on $\widetilde{\Sigma}\cap M$ it follows from uniqueness of the Cauchy problem and finite propagation speed they two must actually coincide in M. Hence $\widetilde{T_\Omega\phi}$ is indeed a smooth extension of $T_\Omega\phi$.

This seemingly simple observation has an interesting consequence concerning the asymptotic behaviour of spatially compactly supported solutions of $P_g\phi = 0$ towards the conformal boundary ∂M. Namely, one just needs to combine the following two pieces of information:

i.) $T_\Omega\phi = \Omega^{-\alpha}\phi$ extends smoothly to ∂M,

ii.) Ω vanishes on ∂M by (AS2), so $\Omega^{-\alpha}$ diverges towards ∂M (since $\alpha > 0$).

As an immediate consequence it follows that ϕ must decay to zero towards ∂M at least as fast as Ω^α.

Actually even more can be said, at least for the massless Dirac equation and the Maxwell equation. For these equations, in a further step the authors of [**MN04**] identify the space of spatially compactly solutions of the equation $P_g\phi = 0$ with their "traces" on the boundary ∂M. More precisely, after fixing a Cauchy surface $\Sigma \subset M$, consider the maps

$$\mathfrak{T}^\pm : \Gamma^\infty_c(E|_\Sigma) \longrightarrow \Gamma^\infty(\widetilde{E}|_{\mathscr{I}^\pm})\,, \quad \phi_0 \longmapsto \widetilde{T_\Omega\phi}|_{\mathscr{I}^\pm}\,,$$

where $\phi \in \Gamma^\infty_{sc}(E)$ denotes the solution of the Cauchy problem for $P_g\phi = 0$ with initial data $\phi|_\Sigma = \phi_0$. Then, and this where most of the actual analysis is contained, it is shown that for certain naturally associated L^2-Sobolev inner products on $\Gamma^\infty_c(E|_\Sigma)$ and $\Gamma^\infty_c(\widetilde{E}|_{\mathscr{I}^\pm})$, the maps $\mathfrak{T}^\pm$ are continuous linear *isomorphisms* between the spaces H and $H^\pm_{\mathscr{I}}$ obtained by taking the closure of $\Gamma^\infty_c(E|_\Sigma)$ and $\Gamma^\infty_c(\widetilde{E}|_{\mathscr{I}^\pm})$ in the respective norm.

Notice that injectivity of the maps $\mathfrak{T}^\pm$ strengthens the previous decay result in the following way: Previously we only knew that any spatially compactly supported solution ϕ of $P_g\phi = 0$ must decay to zero at least as fast as Ω^α as one approaches ∂M. Now we can say that ϕ must actually decay *precisely* as Ω^α (unless $\phi = 0$). Namely, otherwise $(\Omega^{-\alpha}\phi)|_{\partial M} = 0$, which by injectivity of $\mathscr{I}^\pm$ can only hold if $\phi = 0$.

Besides being interesting simply as far as understanding asymptotic properties of solutions is concerned, there are also other interesting applications. For instance, one may define the *scattering operator*

$$S := \mathfrak{T}^+ \circ \mathfrak{T}^- : H^-_{\mathscr{I}} \longrightarrow H^+_{\mathscr{I}}\,,$$

which maps an "incoming state" at past null infinity $\mathscr{I}^-$ to the corresponding "outgoing state" at future null infinity $\mathscr{I}^+$ which results from the time evolution of the equation. In this way one obtains a connection to *scattering theory*, which is the reason why the whole approach described up to this point is called "conformal scattering theory".

It is basically a similar construction on which the principle of "holography" and the *AdS/CFT-conjecture* build, and conformal scattering theory can also be used to construct so-called *Hadamard states* in the framework of algebraic quantum field theory on curved spacetimes (for both, see the introduction of [**DMP06**] and further references therein). The next part of this introduction contains more about the relation to usual scattering theory. For now, let us simply agree that the method of conformal scattering is of interest, and continue towards the actual topic of this thesis.

Conformal scattering for conformally non-invariant equations. One limitation of the conformal approach to scattering theory is of course its restriction to equations which are conformally invariant. This for instance excludes the arguably most prototypical wave equation on a Lorentzian manifold, the usual scalar wave equation $\Box_g \phi = 0$ (i.e., the *minimally* instead of conformally coupled one). It also excludes all *massive equations* such as the Klein-Gordon equation $(\Box_g - m^2)\phi = 0$ or the massive Dirac equation $(\mathcal{D} - m)\psi = 0$, which are crucial for Quantum Field Theory. Here $m > 0$ is a parameter referred to as *mass*.

This limitation, or rather the question

> *Can the methods of conformal scattering theory be extended in some way to conformally non-invariant equations?*,

was the original motivation for this thesis. Clearly for an arbitrary equation this seems unreasonable. But one might hope that if conformal invariance is broken only in some controllable way, then maybe the methods can still be suitably adapted. One natural class of equations for which this happens are "massive perturbations" of conformally invariant equations.[4]

For this reason, in this thesis we study the *massive Dirac equation* $(\mathcal{D} - m)\psi = 0$. Recall that the Dirac operator $\mathcal{D}$ is conformally invariant in the following way: If M is spin with $\dim M = n+1$, and if $\widetilde{g} = \Omega^2 g$ are two conformally related Lorentzian metrics on M, then after suitably identifying the two spinor bundles the two corresponding Dirac operators $\mathcal{D}$ and $\widetilde{\mathcal{D}}$ are related by

$$\widetilde{\mathcal{D}}\big(\Omega^{-\frac{n}{2}}\psi\big) = \Omega^{-\frac{n}{2}-1}\,\mathcal{D}\psi\,.$$

Therefore, if ψ is a solution of the massive Dirac equation $\mathcal{D}\psi = m\psi$, then the rescaled spinor field $\widetilde{\psi} := \Omega^{-\frac{n}{2}}\psi$ satisfies the equation

$$\widetilde{\mathcal{D}}\widetilde{\psi} = \frac{1}{\Omega} m \widetilde{\psi}\,. \tag{$*$}$$

So one sees explicitly that for $m \neq 0$ conformal invariance is broken by the fact that the mass term picks up a factor Ω^{-1}.

Returning to the setup of asymptotically simple spacetimes, a crucial observation is that by (A2) the conformal factor Ω *vanishes* as one approaches the conformal boundary $\partial M \subset \widetilde{M}$. Consequently, the term $\Omega^{-1}m$ blows up, illustrating that the equation $(*)$ for $\widetilde{\psi}$ has in some sense a "singularity" on ∂M. Since ∂M is precisely the place where we want to understand the behaviour of a solution $\widetilde{\psi}$, the question is therefore:

> *How does the singularity of Ω^{-1} on ∂M influence the behaviour of a solution $\widetilde{\psi}$ of $(*)$, defined on $\widetilde{M} \setminus \partial M$, as one approaches ∂M?*

In the following we are first going to present a rough analogue to $(*)$ which will suggest a particular type of behaviour of the solutions as one approaches the singularity. Afterwards, however, we will argue why this analogy is perhaps not so good in the current situation and hence might be misleading. That it is indeed not a correct analogy is also confirmed by the results of this thesis.

First we describe the analogue to the Dirac equation. By assumption (A2) the conformal factor Ω is a boundary defining function for the smooth part $\mathscr{I}^{\pm}$ of the boundary. Therefore it can be used as a transverse coordinate $x^0 = \Omega$ to $\mathscr{I}^{\pm}$, i.e. it can be suitably completed to a set of coordinates $(x^0, x^1, \dots, x^n)$ such that the part of $\mathscr{I}^{\pm}$ lying in the corresponding coordinate domain corresponds precisely to the points p with

[4] The word "perturbation" is not quite sensible here since (as is seen in this thesis) introducing a mass term may substantially alter the properties of interest here.

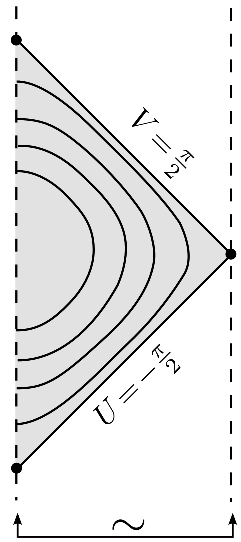

FIGURE 4. Level sets of the conformal factor $\Omega = 2\cos U \cos V$ for the Penrose compactification of Minkowski spacetime into the Einstein cylinder. As the level sets illustrate, $\mathrm{d}\Omega$ becomes lightlike towards $\mathscr{I}^{\pm}$.

$x^0(p) = \Omega(p) = 0$. Concerning the analogue to equation $(*)$, we write $x = x^0$, simply ignore all the other coordinates $(x^1, \ldots, x^n)$ and also the vector-valuedness of the Dirac equation, and consider the ordinary differential equation

$$\mathrm{i}f'(x) = \frac{m}{x} f(x)\,. \tag{$**$}$$

So in a sense we have replaced the Dirac operator $\widetilde{D} = \mathrm{i}\widetilde{\gamma}^{\mu}\widetilde{\nabla}_{\mu}$ by $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}$. Equation $(**)$ can of course be explicitly solved, and its solutions have the form

$$f(x) = c\, x^{-\mathrm{i}m} = c\,\mathrm{e}^{-\mathrm{i}m\log x}\,.$$

Notice in particular that $\lim_{x\to 0}(f(x)\cdot \mathrm{e}^{\mathrm{i}m\log x}) = c$ exists and may be viewed as a sort of "initial value of f at $x = 0$" since one can recover f from the knowledge of this limit.

Going back to the rescaled Dirac equation $(*)$, this may be taken as a suggestion that perhaps a solution $\widetilde{\psi}$ of $(*)$ behaves like $\widetilde{\psi} \sim \Omega^{-\mathrm{i}m}$ as one approaches $\mathscr{I}^{\pm}$. In other words, the suggestion is that $\widetilde{\psi}$ will simply oscillates faster and faster. If this were true, one could define "initial values" of $\widetilde{\psi}$ along $\mathscr{I}^{\pm}$ by first dividing by this oscillating factor and then taking the limit of $\Omega^{\mathrm{i}m}\widetilde{\psi}$ towards $\mathscr{I}^{\pm}$.

Let us now give an argument for why it does not quite work like this in asymptotically simple spacetimes. Basically there is one simple reason:

> The coordinate $x^0 = \Omega$ is in a sense not a good "time direction" with respect to which the (massive) Dirac equation $(*)$ is hyperbolic.

Maybe the simplest way to make this statement more sensible is by pointing out that $\mathrm{d}\Omega$ becomes lightlike along $\mathscr{I}^{\pm} = \Omega^{-1}(\{0\})$, which are lightlike hypersurfaces (of $\widetilde{M}$). For Minkowski spacetime this is illustrated by figure 4, where level sets of Ω are sketched. For this reason, i.e. since they are not spacelike, $\mathscr{I}^{\pm}$ are no natural hypersurfaces where one might pose initial values. To be more accurate, this is not quite true since sometimes one can actually solve so-called characteristic initial value problems for hyperbolic equations, for which initial values are prescribed on a characteristic hypersurface instead of a spacelike one. For instance this is actually done in [**MN04**], where solving such a characteristic initial value problem is a key part of the conformal scattering construction. However, as far as the equation $(*)$ is concerned, the additional fact that $m\Omega^{-1}$ has a *singularity on* $\mathscr{I}^{\pm}$ (for $m \neq 0$) provides a rather convincing intuitive argument for why matters are different here:

For this argument, one views $m\Omega^{-1}$ as an infinitely high "potential wall" at $\mathscr{I}^{\pm}$, which prevents a solution of $(*)$ from reaching $\mathscr{I}^{\pm}$. Vividly speaking, at each spacetime point the solution can propagate along the causal directions into the future (timelike and lightlike ones). The lightlike directions are those which lead onto $\mathscr{I}^{+}$, whereas the timelike ones are those leading into ι^{+}. Now the intuitive idea is that for $m \neq 0$ the infinitely high potential wall $m\Omega^{-1}$ prevents the solution from reaching $\mathscr{I}^{+}$, and rather "pushes it into ι^{+}". Consequently, most of the wave should propagate into ι^{+} if $m \neq 0$. In the massless case ($m = 0$) on the other hand, the solution may propagate "unhindered" onto $\mathscr{I}^{+}$.[5]

Of course this is only an intuitive account of what might be going on around $\mathscr{I}^{+}$, and one should try to investigate to which extent this picture is actually correct.

The real starting point for this thesis. Making the previously outlined intuitive description more precise may be considered as the "real" starting point of this thesis and a way to understand its contribution. Namely, the question addressed in this thesis is precisely how $\widetilde{\psi}$ does in fact behave at $\mathscr{I}^{+}$.

The first attempt was to deduce asymptotic behaviour of solutions similar as for the analogue equation presented above. More precisely, the idea consisted in trying to rescale a solution in such a way that the rescaled solution satisfies (in suitable coordinates) some equation which can be continued to the boundary. As this more and more did not seem to work, the focus shifted towards studying decay properties by the usual methods for hyperbolic equations, i.e. by energy estimates. This then lead to the actual contents of the main body of this thesis, in particular to the result that a (spatially compactly supported) solution $\widetilde{\psi}$ decays "superpolynomially fast" towards $\mathscr{I}^{+}$ (see the summary of the results on page iii above for more details).

This indeed indicates that $\widetilde{\psi}$ does "not reach" $\mathscr{I}^{+}$ in the sense that it decays very rapidly. In a second step, it would of course be interesting to determine what precisely happens around the point ι^{+}. This is still open for future research, some indications are made in Chapter 5.

3. What are some Relations to Other Topics?

In the following we describe some similar research topics and directions. Of course this is not meant to be an exhaustive overview over any particular field or fields. The aim is rather to provide some coarse orientation on the integration of the results into a larger context by elaborating on various related topics.

Scattering theory. As already mentioned, the results of [**MN04**] do not only yield precise decay properties of certain fields at null infinity, but also contain the construction of the "scattering operator" $S : H^{-}_{\mathscr{I}} \to H^{+}_{\mathscr{I}}$ which maps between "asymptotic in- and out-states". In the following, we use the example of scattering theory to give some further physical motivation.

[5] The fact that, at infinity, the massless Dirac equation is basically a transport equation along null geodesics guarantees that it indeed does so, see [**MN04**, p. 213].

Scattering theory can be used in many different specific situations, but the connecting underlying theme is always that one wants to understand the asymptotic behaviour of some dynamical system (for large times) in terms of a related, simpler dynamical system. Such a system could describe the motion of particles, or the propagation of waves (as in this thesis), but also something rather different such as population dynamics. Scattering theory can be used to study such a system whenever there exists another, typically more simple dynamical system which approximates the first one in some asymptotic regime. The point is that this second, simpler system should be explicitly computable (in some sense) and thus allow to say something about the actual system of interest, which typically will be too complicated to solve explicitly.

To stay within the range of this thesis, one may imagine some sort of waves propagating through a part of the universe (spacetime) which contains an isolated gravitating object such as a star, a collection of stars, a black hole, or similar. The waves could for instance be lightwaves described by Maxwell's equations, or quantum mechanical waves (electrons) described by the Dirac equation. In any case, the dynamical system of interest is an equation $P\phi = 0$ which describes the propagation of these waves in the given background spacetime. Now the idea in the present setup is that far away from the isolated object, the background spacetime should resemble flat Minkowski spacetime more and more, so the same should hold for the propagation of waves. Mathematically this means that if $P_0\phi = 0$ describes the propagation of the same type of waves in flat Minkowski spacetime, then

$$P\phi = P_0\phi + R\phi\,,$$

where $R\phi$ becomes negligible at large distances from the isolated object (it might contain terms proportional to r^{-1}, r^{-2}, ... where r is the distance to the object). Let us call the part where R is not negligible the "region of interaction".

Moreover, certain solutions of the so-called "free equation" $P_0\phi = 0$ usually have a clear physical interpretation ("plane waves"), and it is often possible to analyze or decompose an arbitrary solution of $P_0\phi = 0$ in terms of these special ones. This then leads to the following pictures: One images a wave entering from far away into the region of spacetime where the isolated object influences strongly the propagation of waves. While still far away, this wave can be interpreted in terms of these special solutions of the free equation. After a while the wave enters into the region, where the isolated gravitating object influences strongly its propagation. Here it may not be possible to give a clear interpretation of the wave. But after yet another while (some part of) the wave will move away from the isolated object, and after some more time be so far away from it that it can again be interpreted in terms of the special solutions of the free equation. Now the idea behind the "scattering map" is that it simply assigns to each wave which enters the region of interaction the wave which leaves this region after some time. If one knows the action of this map on *all* possible incoming waves, then one understands effectively the way in which the isolated object influences wave propagation. Typically this means that the scattering map $S : \{\text{incoming states}\} \to \{\text{outgoing states}\}$ should be a bijection.

So much for the physical picture behind scattering theory. In practice, what one really has to show of course depends very much on the particular problem and how it is formulated. The setup in which scattering theory is most developed is where the dynamics (for instance the propagation of waves) can be described in terms of a 1-parameter group of unitary operators on a Hilbert space. Such a dynamics is always generated by a self-adjoint operator H, called *Hamiltonian*, and, roughly speaking, the mathematics of scattering theory in this framework consists in showing that H is unitarily equivalent to another, more simple self-adjoint operator H_0. This operator is typically derived from H by neglecting certain terms which are believed to be "small" in some sense, and describes the free dynamics. There exists a large body of literature on this particular approach, some well-known textbooks include [**DG97**], [**Yaf92, Yaf10**], or [**RS79**]. Formulated in this way, scattering theory may be considered as a part of spectral theory for operators, so let us call it the "spectral approach to scattering theory".

The spectral approach to scattering theory has also been used in the context of wave equations on Lorentzian manifolds by various people in the present and past, and it has lead to very interesting results (see for instance the introduction of [**MN04**] for a few references). While being very powerful, the spectral approach to scattering theory suffers from one drawback: It basically works only for equations which do not explicitly depend on time.[6] In the setting of wave equations on Lorentzian manifolds this means that the underlying spacetime should admit a timelike Killing vector field. However, similarly as argued in the introduction of [**MN04**], from the standpoint of partial differential equations the asymptotic behaviour of waves for large times, i.e. for instance the structure of the solutions close to $\mathscr{I}^+$ (whenever such a notion is available), should mostly depend on the asymptotic structure of the metric in that part of infinity and not so much on whether it depends on time or not (at least the parts of the waves which travel out to infinity). The reason for this is that differential operators are *local*. Therefore it might seem reasonable to believe that there should exist other methods besides spectral theory by which one can extract the asymptotic behaviour of these waves, and which continue to work also for spacetimes without a timelike Killing field.

To be clear, what one *cannot* give up is the assumption that the underlying spacetime should have some well-defined "asymptotic structure at infinity" which yields a sort of "simpler comparison dynamics" in the sense of scattering theory as explained in the beginning. However, as the notion of asymptotically simple spacetimes from above shows, this is not necessarily tied to the existence of a timelike Killing field.[7] Therefore, the proposal in [**MN04**] is that the conformal approach to scattering theory, as explained before, could provide an alternative to the spectral approach in these spacetimes, at least for certain equations. It was also shown in [**MN04**] that in case the underlying spacetime is taken to be Minkowski spacetime, the conformal and the spectral theoretic approach lead to the same scattering operator. This provides further affirmation that the conformal approach may indeed be a "correct way" to study scattering also on spacetimes without timelike Killing vectors. Since their paper [**MN04**], the authors have further pursued the conformal approach in [**MN09**], [**MN12**], [**Nic13**], and conformal scattering has also been studied for a nonlinear equation in [**Jou10, Jou12**] by a student of one of these authors. Of course, with the conformal approach one is restricted to conformally invariant equations, or one has to try to extend this approach in some way to other equations, which brings us back to one of the original motivations behind this thesis.

Analysis on manifolds with a structure at infinity. It is a common theme in the analysis of (partial) differential equations to study an equation in some "asymptotic regime" in which the equation simplifies. In this regime it is then often possible to make

[6] This is not completely correct, but it seems that the explicitly time-dependent situations which can be covered are rather limited.

[7] One could imagine that at least asymptotically there should exist a timelike "almost-Killing field", which could re-open the door for spectral methods.

more specific statements about solutions of the equation. For instance, scattering theory as just explained fits into this scheme, the asymptotic regime being the region at large distance to the "interaction region" in this case.

One particular setup in which this general theme is used, and which relates very closely to this thesis, is the analysis on manifolds with a well-defined "asymptotic structure at infinity". Indeed, the particular spacetimes on which the Dirac equation is studied in this thesis have such a structure at infinity (see page iii of this introduction).

In the context of Riemannian geometry on *noncompact* manifolds, a common theme is to study the spectrum (or other properties) of the usual differential operators like the Laplace or Dirac operator on manifolds which have some sort of asymptotic structure at infinity: Typical examples are *asymptotically flat* manifolds, *hyperbolic* manifolds, manifolds with *cylindrical ends*, manifolds with *conical ends*, and others. There exists a large amount of articles in this context, a number of references can be found in the introduction of the article [**ALN04**].

One basic idea of the analytic machinery used in these situations, which by now is rather well developed, is the introduction of an appropriate functional framework (Sobolev spaces, pseudodifferential operator calculi) which is adapted to the specific structure at infinity. A basic exposition of these ideas can be found in the textbook [**Mel95**] by Melrose under the keywords "scattering calculus" and "b-calculus". This approach has been generalized by Ammann, Lauter, and Nistor in [**ALN04, ALN07**] to so-called Riemannian manifolds with a "Lie structure at infinity".

Studying spectral properties of various geometric differential operators on such manifolds is in fact closely related to the propagation of waves for certain wave equations on corresponding *static* spacetimes (cf. [**Mel95**]). This relation has been studied by various people, for instance by Melrose, Vasy, and Wunsch (cf. [**Vas08, MVW08**]).

If one wants to study wave equations on *non-static* (or non-stationary) spacetimes, it is a natural question whether perhaps one can make a good definition of a space*time* with an asymptotic structure at infinity which then allows to develop an adequate functional framework to study asymptotic properties of solutions of wave equations on such spacetimes as well. For the wave equation, some very interesting recent results in this direction were obtained by Vasy for asymptotically de Sitter like, asymptotically anti de Sitter like, and asymptotically hyperbolic and Kerr-de Sitter like spacetimes ([**Vas10, Vas12, Vas13**]), by Melrose, Sá Barreto, and Vasy for asymptotically de Sitter-Schwarschild like spacetimes ([**MSBV13**]), and by Baskin, Vasy, and Wunsch for asymptotically Minkowski spacetimes ([**BVW12**]). In these works the authors derive very precise asymptotic expansions and decay rates for solutions of the wave equation. The approach is always in the spirit of introducing an appropriate (microlocal) functional framework adapted to the asymptotic geometry at infinity, and a basic paradigm is to study *Fredholm properties* of (non-elliptic) operators such as the wave operator on suitable functional spaces. Let us also mention the very recent work of Hintz and Vasy which uses similar ideas to study local and global existence for semilinear wave equations on asymptotically de Sitter, Kerr-de Sitter and Minkowski spacetimes ([**HV13**]), and for quasilinar wave equations on asymptotically de Sitter and Kerr-de Sitter spacetimes ([**Hin13, HV14**]).

Returning to the beginning of the second part of this introduction, also the definition of an asymptotically simple spacetime provides a definition for an asymptotic structure at infinity. As explained, it uses ideas from conformal geometry and allows to study asymptotics of solutions of conformally invariant wave equations. In a similar spirit, i.e. using conformal geometry (but in a more elaborate way), Gover and Waldron have recently developed a "boundary calculus" for equations on conformally compact pseudo-Riemannian manifolds of arbitrary signature ([**GW14**]). Similar to the ideas explained in the second part of the introduction, here one tries to relate solutions of equations in the original manifold to certain "boundary values" on the boundary of the conformal compactification. As also mentioned before, this has relations to the AdS/CFT-conjecture, see the introduction of [**GW14**]. Trying to use the approach of Gover and Waldron to study asymptotic (decay) properties of solutions of wave equation in the Lorentzian case is recent ongoing work.

Decay of linear waves. For linear hyperbolic equations, for which existence and uniqueness of solutions is well-understood in many cases, decay or more generally asymptotic behaviour of solutions is one of the remaining general questions about its solutions one can always attempt to study. For various different equations this has been and still is continuously studied by many people for different reasons. Besides being interested in decay properties for their own sake, one commonly encountered motivation is that one wants to use decay properties of solutions of linear equations for studying nonlinear perturbations of these equations (see for instance [**Tao06**, Ch. 3]).

Instead of only restricting to nonlinear perturbations of linear equations, one may of course also start from a genuinely nonlinear equation and try to approach it by studying various linearizations. Here the paradigmatic example which fits into the general relativistic context of this thesis are the Einstein equations. For these, one of the current aims is to prove stability of certain special solutions, most notably the Kerr black hole spacetime, under perturbations of their initial values. Here "stability" means stability of their global geometric properties such as geometric (in-)completeness or asymptotic falloff properties of the metric (similar to those of asymptotic simplicity). The approach of first studying decay properties of solutions of linear equations (on these specific spacetimes) has been successful for Minkowski spacetime (cf. [**CK93**]). We refer to the review article [**CGP10**, Sec. 6] for more details. In this spirit, one may view any decay result for a wave equation on a curved spacetime which can be coupled to the Einstein equations as a sort of naive "linear stability result" for the coupled system of that wave equation and the Einstein equation.

Coming back to decay estimates for linear equations and to the massive Dirac equation in particular, let us mention the work of Finster, Kamran, Smoller, and Yau on the massive Dirac equation on certain black hole spacetimes like the Kerr-Newman spacetime ([**FKSY02, FKSY03, FKSY09**] or the Reissner-Nordström spacetime ([**FKSY00**]). Here the original question was whether there can exist "bound states" of the massive Dirac equation on these spacetimes, i.e. solutions which remain localized in some bounded domain. These would correspond to the quantum mechanical analogue of classical planetary orbits. To answer this question, the authors studied the decay of local energy as time tends to infinity. The existence of bound states would mean that local energy does not decay. The methods used by these authors are quite different compared to the ones described before in that they derive an explicit integral representation of the Dirac propagator, i.e. the time evolution operator defined by the Dirac equation. This leads to certain Fourier-integrals which the authors estimate by the "saddle-point method" in the limit where time tends to infinity. In this way the authors manage to show that local energy and actually also the solution itself (at any point in space) decay like $t^{-\frac{5}{6}}$ as $t \to \infty$. The interpretation of this is that the wave propagates into the black hole and/or escapes to infinity. Compared to the decay results in this thesis, one may note that the limit $t \to \infty$ (at fixed x) corresponds to approaching timelike infinity ι^+. The slower decay in timelike directions compared to the superpolynomial decay in lightlike directions obtained in this thesis reinforces the picture that the solution propagates mainly into ι^+ instead of $\mathscr{I}^+$. It is interesting to ask whether one can somehow describe these two decay properties together in a more uniform way.

4. Brief Summaries of the Individual Chapters

To finish up this introduction, we present a brief summary of the content of the individual chapters of the thesis. Basically the first two chapters contain background information to make the whole thesis more self-contained and to put the content into a larger context. In practice they also contain material, especially about the Dirac equation, which seems to be commonly known but is hard (or maybe not at all) to find in the existing literature. Nevertheless, the main new contributions of this thesis are contained in the third and fourth chapter. In particular Chapter 4 contains the whole analytical treatment of the problem.

Summary of Chapter 1. The basic objects of interest in this thesis are *spinors* on Lorentzian manifolds, a concept which is geometrically somewhat involved. The first chapter serves as brief introduction into the definition of spinors in order to make the thesis more self-contained, and also to give readers without background knowledge about spinors at least a small impression of the general algebraic and geometric underpinnings of the concept of a spinor.[8]

In practice there exist different, equivalent approaches to spinors (see the Introduction of Chapter 1). In this thesis we follow the representation-theoretic approach using Clifford algebras, spin groups, and the spin representation on the algebraic side, and spin structures as well as the machinery of principal bundles and associated vector bundles on the geometric side. This approach has the advantage that it works in any dimension (and signature). All necessary concepts will be touched upon in the first chapter, although sometimes of course only very briefly. The hope is to provide a readable, sometimes intuitive (often quite dense) summary, which is understandable for readers having no background knowledge about spinors, and is a nice repetition for readers who are already familiar with spinors.

Summary of Chapter 2. In the second background chapter some general methods regarding the analysis of hyperbolic equations on Lorentzian manifolds are presented. The focus is on the Dirac equation, and the methods explained here in a general context are used later in the thesis in specific situations. In short, the methods presented are:

- *Energy estimates* in a geometric fashion, involving the *conserved current* of the Dirac equation, the *energy-momentum tensor formalism*, and related ideas,
- *Symmetric hyperbolic systems*, a general type of hyperbolic equations which applies to the Dirac equation and can be used to treat the Cauchy problem and finite propagation speed,
- *Representation formulas* for solutions of (some) hyperbolic equations in terms of their initial data, which can be obtained by *Green's function methods.*

[8] It should be pointed out that the actual analysis in Chapter 4 starts with a concrete equation, and one should be able to follow it without any (or at least not too much) knowledge about spinors.

Many of the methods summarized in this chapter are used later in Chapter 4. Although Chapter 4 is written sufficiently detailed to be readable on its own, the presentation in the second chapter embeds the methods used there into a larger context. The hope is that in this way the approach taken in this thesis is given a more methodical touch. In particular, Chapter 2 could be informative for potential readers with little background knowledge about hyperbolic equations by helping them understand the methods used later in the thesis in a somewhat more general context.

Of course, most of what is written in Chapter 2 is standard material about hyperbolic equations and can be found in most textbooks of the subject, as cited in Chapter 2. However, applied specifically to the Dirac equation it is difficult (or not possible) to trace back some of the parts of this chapter to the existing literature even if these parts seem to be common knowledge. For this reason the chapter contains a number of detailed, although elementary computations for the Dirac equation, for instance concerning energy-momentum tensors.

Summary of Chapter 3. With this chapter the actually new contributions of this thesis start. In a sense, the third chapter prepares the stage for the analytical investigations of Chapter 4, which make up the largest part of the thesis. Namely, Chapter 3 is concerned with the structure of the particular spacetimes on which we analyze the Dirac equation, and the form of the Dirac equation on these spacetimes.

Concretely, in the first part of Chapter 3 we introduce the class of spacetimes we consider, and describe their causal properties as well as the asymptotic conditions at infinity which we impose. Alongside we also describe some examples from the literature which fit into this class of spacetimes.

In the second part of Chapter 3 we study spinors on these spacetimes, deriving for instance a convenient expression for the Dirac operator in "block form" (see page iv of this introduction). Moreover, we derive various explicit formulas for Clifford multiplication, the spin connection, and inner products on the spinor bundle.

In the last part of this chapter we first show how the Dirac equation can be further simplified by making a conformal rescaling of spinor fields. Afterwards we introduce certain pointwise, L^2, and Sobolev inner products on spinors, which are going to be used in the analysis of the fourth chapter. At the end of the chapter we also study how the current of a solution of the Dirac equation behaves in presence of the "interior boundary" at $r = r_0$ of these spacetimes (see figure 1), i.e. when it is conserved or not.

Summary of Chapter 4. This is the core chapter of the whole thesis, which contains all the analytic work. Let us therefore give a brief overview over the different parts of this chapter.

Section 4.2 starts with the previously mentioned Dirac operator in "block form" (and after the conformal rescaling mentioned above), and rewrites this further. The main step consists in separating off the N-dependence in the equation (recall that $M = \mathbb{R}_t \times (r_0, \infty)_r \times N$). Concretely this is done by projecting the Dirac equation $(\mathcal{D} - m)\psi = 0$ onto an $L^2(\mathbb{S}N)$-orthonormal basis of eigenspinors of the Dirac operator $\mathcal{D}_N$ of N, which works very conveniently using the block form of the Dirac operator.[9] Each projected equation has the structure of a Dirac equation in flat 1+1-dimensional Minkowski spacetime with an "external potential" reflecting the presence of N and the mass $m > 0$. As such, each equation is a pair of two transport equations which are coupled by the "potential". In

[9]That this can be done in the first place relies on the fact that the Dirac operator on a compact Riemannian manifold is elliptic and (essentially) self-adjoint on L^2 with discrete spectrum.

order to study these equations analytically, we transform to null coordinates v and u as introduced on page iv of this introduction. Expressed in these coordinates the Dirac equation takes a particular form which we refer to as the *Dirac null system*.

In the next step, in Section 4.3 we decompose the Dirac null system into two parts: One part contains the (as it turns out) relevant terms for our purposes but at the same time is sufficiently simple such that it can be solved explicitly. It is called the "free part". The second part, the "perturbation", contains all other terms which in the end turn out to be "small" compared to the free part. Let us point out that the free part also contains the mass $m \neq 0$, which is crucial for the decay properties.

Having made this splitting, in Section 4.4 we explicitly solve the free part. Here the key observation is that the free part is in a sense equivalent to the *Klein-Gordon equation* in flat 1+1 dimensional flat Minkowski spacetime. For this reason it is possible to derive a representation for its solution in terms of a convolution integral of their initial data with the *Green's function* of the Klein-Gordon equation. Using this formula and the specific form of the Green's function, it is possible to directly estimate the decay of solutions as v tends to infinity. Here it is crucial to have a *nonzero mass*, otherwise the decay properties would not be as strong as obtained here (which is superpolynomially fast).

The following Sections 4.5 and 4.6 consist of showing, by a perturbative argument based on the *Lippmann-Schwinger equation* (see Sec. 4.6.1), that the perturbation terms do not alter the decay properties for $v \to \infty$. This is achieved by various *energy estimates*, which in detail are rather complicated but in principle simply combine *current conservation* of the Dirac equation as useful *a priori estimate* with the decay properties of the metric coefficients at infinity. How these estimates can then be combined with the decay properties of the free part by the Lippmann-Schwinger equation is outlined in Section 4.6.2 and carried out explicitly in Section 4.6.4.

The previous steps yield decay of solutions of the projected parts of the full Dirac equation of M. Finally, in Section 4.7 we combine these decay estimates into an analogous estimate for solutions of the full Dirac equation on M. As indicated before, here it is crucial that we projected onto an L^2-orthonormal eigenbasis of $\mathcal{D}_N$. Namely, this together with having explicit control of the separation constants (the eigenvalues of $\mathcal{D}_N$) in the estimates for the projected equations allows to "sum up" these individual estimates using Parseval's theorem. This then finally establishes the main result obtained in this thesis.

Notation and Conventions

In the following one finds a list of some general notations and notational conventions used throughout this thesis. Mostly we follow standard conventions of the respective fields.

General conventions

- The signature of a Lorentzian metric is $(- + \cdots +)$.
- The Clifford relations are defined with sign convention $vw + vw = -2\langle v, w\rangle$.
- Both index-free and abstract index notation are used throughout this thesis. For instance, the (covariant) divergence of a symmetric 2-tensor $T = T_{\mu\nu}$ is denoted by $\operatorname{div} T = \nabla^\mu T_{\mu\nu}$.
- The Einstein summation convention is used. Greek indices $\alpha, \beta, \mu, \nu, \ldots$ usually run from 0 to n, and Roman indices $i, j, k, \ell, \ldots$ run from 1 to n.
- We sometimes write X_x for a set X whose points are denoted by x to stress the notation for the points. For instance, we commonly write $\mathbb{R}_t \times \Sigma_x$.
- In estimates, numerical constants may change from line to line without this being explicitly stated.

Notation related to general differential and pseudo-Riemannian geometry

$(M^{p,q}, g)$	Pseudo-Riemannian manifold of signature (p, q)
(M, g)	Pseudo-Riemannian manifold of some signature
$\mathsf{SO}^+(M)$	Bundle of oriented and time-oriented orthonormal frames of (M, g)
$\mathrm{d}\mu_g$	Volume measure of a pseudo-Riemannian manifold (M, g)
$I^\pm, J^\pm$	Timelike and causal future and past (in Lorentzian signature)
$\mathrm{d}\mu_\Sigma$	Induced volume measure on nondegenerate submanifold $\Sigma \subset (M, g)$
ν_Σ	Unit-normal to nondegenerate hypersurface $\Sigma \subset (M, g)$
$\Gamma^\infty(E)$	Smooth sections of a vector bundle E
$\Gamma^\infty_c(E)$	Compactly supported smooth sections
$\Gamma^\infty_{sc}(E)$	Spatially compactly supported smooth sections
$\Gamma_{L^2}(E)$	L^2-sections (w.r.t. some inner product and volume measure)
$\Gamma_{\mathcal{D}'}(E)$	Distributional sections

Notation related to Clifford algebras and spin groups

$\mathbb{R}^{p,q}$	$\mathbb{R}^{p+q}$ equipped with canonical inner product of signature (p, q)
$\mathsf{SO}^+(p, q)$	Time- and space-orientation preserving orthogonal maps of $\mathbb{R}^{p,q}$
$\mathrm{Cl}(p, q)$	Real Clifford algebra in signature (p, q)
$\mathbb{C}\mathrm{l}(n)$	Complex Clifford algebra in dimensions n
$\{\cdot, \cdot\}$	Anticommutator of two endomorphisms
$\mathsf{Spin}^+(p, q)$	Proper spin group in signature (p, q)
$\vartheta_{p,q}$	Spin covering of $\mathsf{SO}^+(p, q)$ by $\mathsf{Spin}^+(p, q)$

$\mathbb{S}_{p,q}$	Space of spinors in signature (p,q)
$\rho_{p,q}$	Spin representation of $\mathsf{Spin}^+(p,q)$ on $\mathbb{S}_{p,q}$
$\gamma_{\mathbb{S}_{p,q}}$	Clifford multiplication of $\mathbb{R}^{p,q}$ on $\mathbb{S}_{p,q}$
$\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$	$\mathsf{Spin}^+(p,q)$-invariant Hermitian inner product on $\mathbb{S}_{p,q}$.

Notation related to spinors

$\mathsf{Spin}^+(M)$	Spin structure for (M,g)
$\mathbb{S}M$	Classical spinor bundle of (M,g) (w.r.t. some spin structure)
$\prec\cdot,\cdot\succ_{\mathbb{S}M}$	Natural Hermitian inner product on $\mathbb{S}M$
$\gamma(X)$	Clifford multiplication by $X \in TM$ on $\mathbb{S}M$
$\nabla^{\mathbb{S}}, \nabla^{\mathbb{S}M}, \nabla$	Spin connection on $\mathbb{S}M$
$\mathcal{D}_M, \mathcal{D}$	Classical Dirac operator on $\mathbb{S}M$

Notation related to analysis

$\lVert\cdot\rVert_{L^p_x(\Omega)}$	L^p-norm on a space Ω whose points are denoted by x
$\lVert\cdot\rVert_{H^k_x(\Omega)}$	Sobolev-type norm of order k on a set Ω_x
$\lVert\cdot\rVert_{C^k_x(\Omega)}$	C^k-norm on a set Ω_x

CHAPTER 1

Geometric and Algebraic Concepts behind the Dirac Equation

In this chapter we briefly summarize the necessary algebraic and geometric background material which is needed in order to introduce the Dirac equation on a Lorentzian manifold. To be fair, there exist some slightly different approaches to the Dirac equation, such as the orthonormal frame or representation-theoretic approach going back to Cartan (cf. [**Car13**] or [**Car66**]) which is mostly used in modern differential (spin) geometry, the 2-spinor approach of Penrose (cf. [**PR86**]) which is restricted to $1+3$ dimensions and is popular in the General Relativity community, the $U(2,2)$-gauge theory approach of Finster (cf. [**Fin98**]), and maybe others. In this thesis we are going to use the first approach, and the background material summarized in the following is tailored to this end.

Before starting, let us point out that everything in the following is textbook material, and can in more detail be found for instance in [**LM89**] or [**Fri00**] for the Riemannian case, and in [**Bau81**] for the general pseudo-Riemannian case. As far as only the algebraic material is concerned, [**Har90**] and [**Gre78**, Ch. 10–11] are good, exhaustive treatments.

1.1. Clifford Algebras, Spin Groups, and Spin Spaces

The theory of spinors and the Dirac equation on Lorentzian, or more generally pseudo-Riemannian manifolds builds on two basic algebraic ingredients: the so-called *spin covering* $\vartheta_{p,q} : \mathsf{Spin}^+(p,q) \to \mathsf{SO}^+(p,q)$ and the *spin representation* $\rho_{p,q} : \mathsf{Spin}^+(p,q) \to \mathsf{GL}(\mathbb{S}_{p,q})$. Here (p,q) denotes the signature of the underlying pseudo-Riemannian manifold or inner product space. In the following we briefly describe both of these, starting from the theory of *Clifford algebras* and with a focus on the Lorentzian signature case.

1.1.1. A brief review of Clifford algebras. If $(V, \langle\cdot,\cdot\rangle)$ is a real inner product space[1], then the *Clifford algebra* $\mathrm{Cl}(V, \langle\cdot,\cdot\rangle)$ of $(V, \langle\cdot,\cdot\rangle)$ is the real, unital algebra generated by V subject to the relations

$$\{v, w\} := v \cdot w + w \cdot v = -2\,\langle v, w\rangle \qquad \forall v, w \in V\,. \tag{1.1.1}$$

The bracket $\{\cdot,\cdot\}$ is referred to as *anti-commutator*, and the relations (1.1.1) are known as *Clifford relations*. It is not difficult to show that the Clifford algebra is unique up to isomorphism. Similar to the exterior algebra it can be realized as a quotient of the tensor algebra of V by the ideal generated by the Clifford relations (1.1.1). As we will see, it can also be realized concretely in terms of matrices, see the classification table 1 and the concrete examples in Section 1.1.3.

We will usually denote multiplication in the Clifford algebra by simple juxtaposition of elements, and the explicit reference to the inner product of V will be omitted from

[1]Throughout this thesis, an *inner product* is a symmetric, nondegenerate, but not necessarily positive definite bilinear form.

$(q-p) \bmod 8 = \dots$	$\mathrm{Cl}(p,q) \cong \dots$
$0, 6$	$\mathsf{M}(N, \mathbb{R})$
$1, 5$	$\mathsf{M}(N, \mathbb{C})$
$2, 4$	$\mathsf{M}(N, \mathbb{H})$
3	$\mathsf{M}(N, \mathbb{H}) \oplus \mathsf{M}(N, \mathbb{H})$
7	$\mathsf{M}(N, \mathbb{R}) \oplus \mathsf{M}(N, \mathbb{R})$

TABLE 1. The classification of real Clifford algebras. The number $N \in \mathbb{N}$ on the right-hand side can be determined from $\dim_{\mathbb{R}} \mathrm{Cl}(p,q) = 2^{N+1}$.

notation. In this thesis, only the cases where V is either Minkowski spacetime $\mathbb{R}^{1,n}$ or Euclidean space $\mathbb{R}^n$ are relevant in the end. In these cases we denote the Clifford algebra by $\mathrm{Cl}(1,n)$ and $\mathrm{Cl}(n)$, respectively. More generally, we denote the Clifford algebra of the inner product space $\mathbb{R}^{p,q}$ of signature (p,q) by $\mathrm{Cl}(p,q)$.[2]

Note that, basically by definition, there is always an injection $V \hookrightarrow \mathrm{Cl}(V)$. Moreover, if $e_1, \dots, e_n \in V$ is a basis, then it follows from (1.1.1) that

$$\{1\} \cup \{e_\mu \mid 1 \le \mu \le n\} \cup \{e_\mu e_\nu \mid 1 \le \mu < \nu \le n\} \cup \ldots \cup \{e_1 \cdots e_n\} \tag{1.1.2}$$

is a basis for $\mathrm{Cl}(V)$. In particular, this implies that $\dim \mathrm{Cl}(V) = 2^{\dim V + 1}$. This appears very similar as for the exterior algebra, and the two are indeed isomorphic as vector spaces, although *not* as algebras (cf. [**Har90**, Prop. 9.11]).

Clifford algebras can be completely classified by observing that for any $p, q \ge 0$, there exist isomorphisms (cf. [**Har90**, Lemma 11.17])

$$\mathrm{Cl}(p,q) \otimes \mathrm{Cl}(0,2) \cong \mathrm{Cl}(q, p+2)\,, \tag{1.1.3}$$

$$\mathrm{Cl}(p,q) \otimes \mathrm{Cl}(2,0) \cong \mathrm{Cl}(q+2, p)\,, \tag{1.1.4}$$

$$\mathrm{Cl}(p,q) \otimes \mathrm{Cl}(1,1) \cong \mathrm{Cl}(p+1, q+1)\,. \tag{1.1.5}$$

This shows that one can iteratively construct any Clifford algebra from $\mathrm{Cl}(1,0)$, $\mathrm{Cl}(0,1)$, $\mathrm{Cl}(1,1)$, $\mathrm{Cl}(2,0)$, and $\mathrm{Cl}(0,2)$. Some examples will be explicitly computed in Section 1.1.3 below.

The classification of real Clifford algebras can be summarized more compactly by combining (1.1.3)– (1.1.5) with the fact that for any $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$, there exist isomorphisms between the following real algebras (cf. [**Har90**, Lemma 11.24]):

$$\mathsf{M}(n, \mathbb{R}) \otimes_{\mathbb{R}} \mathsf{M}(n, \mathbb{R}) \cong \mathsf{M}(nm, \mathbb{R})\,, \tag{1.1.6}$$

$$\mathsf{M}(n, \mathbb{R}) \otimes_{\mathbb{R}} \mathbb{K} \cong \mathsf{M}(n, \mathbb{K})\,, \tag{1.1.7}$$

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{C} \oplus \mathbb{C}\,, \tag{1.1.8}$$

$$\mathbb{H} \otimes_{\mathbb{R}} \mathbb{H} \cong \mathsf{M}(4, \mathbb{R})\,, \tag{1.1.9}$$

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{H} \cong \mathsf{M}(2, \mathbb{C})\,. \tag{1.1.10}$$

This allows to neatly fit all real Clifford algebras into the table 1.

Remark 1.1.1. For Lorentzian signature $(1,3)$, one has $(p-q) \bmod 8 = -2 \bmod 8 = 6$, hence $\mathrm{Cl}(1,3) \cong \mathsf{M}(4, \mathbb{R})$. If, however, one chooses signature $(3,1)$ (and does not implement an additional sign in the Clifford relation), then $\mathrm{Cl}(3,1) \cong \mathsf{M}(2, \mathbb{H})$. For complex Clifford algebras we will see that this distinction disappears.

[2]In the convention used here, p is the number of minus signs in the inner product.

1.1.2. Complex Clifford algebras. Clifford algebras can just as well be defined over complex inner product spaces.[3] Since in complex spaces one can always multiply by i and $\mathrm{i}^2 = -1$, the notion of signature makes no sense in the complex case, and there is only one ($\mathbb{C}$-bilinear) inner product on $\mathbb{C}^n$. Thus there is also a unique complex Clifford algebra associated to $\mathbb{C}^n$ (up to equivalence), which we denote by $\mathbb{C}\mathrm{l}(n)$.

As is not difficult to show, if V is a real inner product space, then the complexification of $\mathrm{Cl}(V)$ is isomorphic to the Clifford algebra over the complexification of V. Therefore it follows from (1.1.3) that

$$\mathbb{C}\mathrm{l}(n+2) \cong \mathbb{C}\mathrm{l}(n) \otimes_{\mathbb{C}} \mathbb{C}\mathrm{l}(2)\,. \tag{1.1.11}$$

From this and the fact that $\mathbb{C}\mathrm{l}(1) \cong \mathbb{C} \oplus \mathbb{C}$ and $\mathbb{C}\mathrm{l}(2) \cong \mathsf{M}_2(\mathbb{C})$ (cf. the following section) it follows that

$$\boxed{\begin{aligned} \mathbb{C}\mathrm{l}(2m) &\cong \mathsf{M}(2^m,\mathbb{C}) \\ \mathbb{C}\mathrm{l}(2m+1) &\cong \mathsf{M}(2^m,\mathbb{C}) \oplus \mathsf{M}(2^m,\mathbb{C}) \end{aligned}} \tag{1.1.12}$$

1.1.3. Clifford algebras in low dimensions. Let us now consider some explicit low-dimensional examples to lighten up the somewhat dry presentation so far. In each example, one may compare to the classification table 1.

Example 1.1.2. The simplest example is certainly one-dimensional Euclidean space $\mathbb{R}^1$. Let $e_1 = 1 \in \mathbb{R}$, then a basis for $\mathrm{Cl}(1)$ is given by $\{1_{\mathrm{Cl}}, e_1\}$. By the Clifford relations we must have $e_1^2 = -\langle e_1, e_1\rangle = -1$, from which it follows that $\mathrm{Cl}(1)$ is isomorphic to $\mathbb{C}$ as a *real* algebra if we identify 1_{Cl} with $1 \in \mathbb{C}$ and e_1 with $\mathrm{i} \in \mathbb{C}$. As one easily sees, the complex Clifford algebra $\mathbb{C}\mathrm{l}(1)$ is isomorphic to $\mathbb{C} \oplus \mathbb{C}$ as a complex algebra.

For the following examples, it is useful to make use of the famous *Pauli matrices*

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.1.13}$$

As one easily verifies, the Pauli matrices satisfy the following relations

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}\mathbb{1}_{2\times 2}\,, \qquad \sigma_i\sigma_j = \mathrm{i}\sum_{k=1}^{3} \epsilon_{ijk}\sigma_k + \delta_{ij}\mathbb{1}_{2\times 2}\,. \tag{1.1.14}$$

Example 1.1.3. Next we consider two-dimensional Euclidean space $\mathbb{R}^2$, and denote by $e_1, e_2 \in \mathbb{R}^2$ the standard basis. We embed $\mathbb{R}^2$ linearly into $\mathsf{M}(2,\mathbb{C})$ by setting

$$e_1 \mapsto \gamma_1 := \mathrm{i}\sigma_2\,, \quad e_2 \mapsto \gamma_2 := \mathrm{i}\sigma_1\,. \tag{1.1.15}$$

By (1.1.14) the matrices γ_1, γ_2 satisfy the Clifford relations. Since $\gamma_1\gamma_2 = \mathrm{i}\sigma_3$ is linearly independent of γ_1, γ_2, it follows from (1.1.2) that $\mathrm{Cl}(2)$ is isomorphic to the real subspace of $\mathsf{M}(2,\mathbb{C})$ spanned by $\mathbb{1}_2, \mathrm{i}\sigma_1, \mathrm{i}\sigma_2, \mathrm{i}\sigma_3$. Since these four matrices are in fact a $\mathbb{C}$-basis for $\mathsf{M}(2,\mathbb{C})$, this also shows that the complex Clifford algebra $\mathbb{C}\mathrm{l}(2)$ is isomorphic to $\mathsf{M}(2,\mathbb{C})$.

The Clifford algebra of $\mathbb{R}^2$ can also be related to the quaternions: Denoting the basis elements of $\mathrm{Cl}(2)$ by

$$\mathbf{1} := 1_{\mathrm{Cl}}\,, \quad \mathbf{i} := e_1\,, \quad \mathbf{j} := e_2\,, \quad \mathbf{k} := e_1e_2\,, \tag{1.1.16}$$

one easily sees that they satisfy the relations

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1\,, \quad \mathbf{ij} = \mathbf{k}\,, \quad \mathbf{jk} = \mathbf{i}\,, \quad \mathbf{ki} = \mathbf{j}\,. \tag{1.1.17}$$

[3] With the inner product being still symmetric and bilinear and *not* hermitian.

These are precisely the same relations which are also satisfied by the usual basis elements of the *quaternions* $\mathbb{H}$ (which are also commonly denotey by $\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}$). This shows that $\mathrm{Cl}(2)$ is isomorphic to the quaternions $\mathbb{H}$ as a *real* algebra (compare to table 1).

Example 1.1.4. Let us consider $\mathbb{R}^{3,0}$, i.e. $\mathbb{R}^3$ equipped with the *negative definite* inner product $\langle v, w\rangle_- := -v_1w_1 - v_2w_2 - v_3w_3$. Denoting once more by $e_1, e_2, e_3 \in \mathbb{R}^3$ the standard basis, we embed $\mathbb{R}^3$ linearly into $\mathsf{M}(2,\mathbb{C})$ via

$$e_1 \mapsto \sigma_1\,, \quad e_2 \mapsto \sigma_2\,, \quad e_3 \mapsto \sigma_3\,. \tag{1.1.18}$$

Again by (1.1.14) these matrices satisfy the correct Clifford relations and, as one easily verifies, the map (1.1.18) actually induces an isomorphism between $\mathrm{Cl}(3,0)$ and the real subspace of $\mathsf{M}(2,\mathbb{C})$ spanned by the eight $\mathbb{R}$-linearly independent matrices $\mathbb{1}_2$, $\mathrm{i}\mathbb{1}_2$, σ_1, σ_2, σ_3, $\mathrm{i}\sigma_1$, $\mathrm{i}\sigma_2$, $\mathrm{i}\sigma_3$. Since this subspace is all of $\mathsf{M}(2,\mathbb{C})$ this shows that $\mathrm{Cl}(3,0) \cong \mathsf{M}(2,\mathbb{C})$ and $\mathbb{C}\mathrm{l}(3,0) \cong \mathsf{M}(2,\mathbb{C}) \oplus \mathsf{M}(2,\mathbb{C})$.

Since $\mathrm{Cl}(3,0)$ is generated by the Pauli matrices, it is also known as the *Pauli algebra*. It is commonly used in physics, for instance in relation to spin in non-relativistic quantum mechanics. Since there one is mostly only interested in the complex Clifford algebra, the choice of the negative definite inner product (which may seem strange at first sight) makes no difference. Note from table 1 that $\mathrm{Cl}(3) = \mathrm{Cl}(0,3) \cong \mathbb{H} \oplus \mathbb{H}$, which makes it a bit less convenient than $\mathrm{Cl}(3,0)$.

Now we come to some Lorentzian examples.

Example 1.1.5. Consider $\mathbb{R}^{1,1}$, and denote by $e_0, e_1 \in \mathbb{R}^{1,1}$ the standard basis. We linearly embed $\mathbb{R}^{1,1}$ into $\mathsf{M}(2,\mathbb{R})$ via

$$e_0 \mapsto \gamma_0 := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\,, \quad e_1 \mapsto \gamma_1 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,. \tag{1.1.19}$$

One can relate these matrices back to Pauli matrices, but that is not particularly helpful. As one easily verifies, these matrices satisfy the Clifford relations of $\mathbb{R}^{1,1}$ and we have

$$\gamma_0\gamma_1 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\,.$$

Since this is linearly independent of γ_0 and γ_1, the Clifford algebra $\mathrm{Cl}(1,1)$ is isomorphic to the span of $\mathbb{1}_2, \gamma_0, \gamma_1, \gamma_0\gamma_1$ in $\mathsf{M}(2,\mathbb{R})$, which is just all $\mathsf{M}(2,\mathbb{R})$. Hence $\mathrm{Cl}(1,1) \cong \mathsf{M}(2,\mathbb{R})$, and $\mathbb{C}\mathrm{l}(1,1) \cong \mathsf{M}(2,\mathbb{C})$.

Example 1.1.6. Consider $\mathbb{R}^{1,2}$, and denote by $e_0, e_1, e_2 \in \mathbb{R}^{1,2}$ the standard basis. We embed $\mathbb{R}^{1,2}$ linearly into $\mathsf{M}(2,\mathbb{C})$ via

$$e_0 \mapsto \gamma_0 := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\,, \quad e_1 \mapsto \gamma_1 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,, \quad e_2 \mapsto \gamma_2 := \begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}\,. \tag{1.1.20}$$

As one easily verifies, these matrices satisfy the Clifford relations of $\mathbb{R}^{1,2}$. Moreover, one may check that the eight matrices $\mathbb{1}_2, \gamma_0, \gamma_1, \gamma_2, \gamma_0\gamma_1, \gamma_0\gamma_2, \gamma_1\gamma_2, \gamma_0\gamma_1\gamma_2$ form a basis of $\mathsf{M}(2,\mathbb{C})$ as a real vector space. This explicitly shows that $\mathrm{Cl}(1,2) \cong \mathsf{M}(2,\mathbb{C})$ as a real algebra. It follows that $\mathbb{C}\mathrm{l}(1,2) \cong \mathsf{M}(2,\mathbb{C}) \oplus \mathsf{M}(2,\mathbb{C})$.

Example 1.1.7. As last example, we consider $\mathbb{R}^{1,3}$. Here one finds various different explicit representations of $\mathrm{Cl}(1,3)$ in the physics literature on Quantum Field Theory. Two widely used ones are the so-called *Dirac* and *Weyl representation*: Set

$$\gamma_0^D := \begin{pmatrix} -\mathbb{1}_2 & 0 \\ 0 & \mathbb{1}_2 \end{pmatrix}\,, \quad \gamma_0^W := \begin{pmatrix} 0 & -\mathbb{1}_2 \\ -\mathbb{1}_2 & 0 \end{pmatrix}\,, \quad \gamma_j^D = \gamma_j^W := \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}\,, \tag{1.1.21}$$

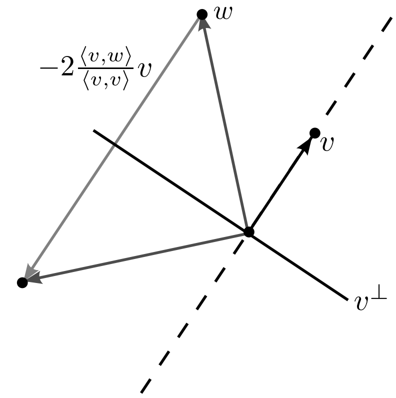

FIGURE 1.1. Illustration of formula (1.1.23) for a reflection across $v^\perp$.

where $j = 1, 2, 3$, and where "D" stands for Dirac and "W" for Weyl. Either of the sets of matrices $\{\gamma_\mu^D\}_{\mu=0,\dots,3}$ and $\{\gamma_\mu^W\}_{\mu=0,\dots,3}$ satisfies the Clifford relations of $\mathbb{R}^{1,3}$ and one can show that in either case the real subalgebra of $\mathsf{M}(4,\mathbb{C})$ generated by these matrices is isomorphic to $\mathrm{Cl}(1,3)$. Furthermore, as a direct inspection shows, the complex subalgebra generated by either of the two set is all of $\mathsf{M}(4,\mathbb{C})$, which shows that $\mathbb{C}\mathrm{l}(1,3) \cong \mathsf{M}(4,\mathbb{C})$.

1.1.4. Spin groups and spin covering. One of the reasons for studying Clifford algebras is that they allow to construct a special $2\!:\!1$ covering map of $\mathsf{SO}^+(p,q)$, which in some cases also turns out to be the universal covering.[4]

Denote by $\mathrm{Cl}^*(p,q) \subset \mathrm{Cl}(p,q)$ the group of invertible elements. Then $\mathrm{Cl}^*(p,q)$ acts on $\mathrm{Cl}(p,q)$ by the *twisted adjoint representation*

$$\mathcal{A}\mathrm{d}_\varphi(z) = \chi(\varphi) z \varphi^{-1}, \quad \varphi \in \mathrm{Cl}^*(p,q), z \in \mathrm{Cl}(p,q). \tag{1.1.22}$$

Here $\chi \in \mathsf{GL}(\mathrm{Cl}(p,q))$ is the unique linear map which satisfies $\chi(v_1 \cdots v_k) = (-1)^k v_1 \cdots v_k$ for all $v_1, \dots, v_k \in \mathbb{R}^{p,q}$.[5] The reason why $\mathcal{A}\mathrm{d}$ is interesting, and why one implements χ, is that for $v, w \in \mathbb{R}^{p,q}$ with $\langle v, v\rangle \neq 0$ it follows from the Clifford relations that

$$\mathcal{A}\mathrm{d}_v(w) = w - 2\frac{\langle v, w\rangle}{\langle v, v\rangle} v\,. \tag{1.1.23}$$

This is precisely the reflection of w across the hyperplane orthogonal to v, see figure 1.1 for an illustration. So by restricting $\mathcal{A}\mathrm{d}_v$ to $\mathbb{R}^{p,q}$ we obtain an orthogonal transformation.

For a general element $\varphi \in \mathrm{Cl}^*(p,q)$, the map $\mathcal{A}\mathrm{d}_\varphi$ need of course not preserve the subspace $\mathbb{R}^{p,q} \subset \mathrm{Cl}(p,q)$. But if we denote by $\Gamma \subset \mathrm{Cl}^*(p,q)$ the subgroup consisting of precisely all those elements that *do* preserve $\mathbb{R}^{p,q}$, then one can show that actually $\mathcal{A}\mathrm{d}_\varphi \in \mathsf{O}(p,q)$ for all $\varphi \in \Gamma$. Hence we obtain an orthogonal representation $\mathcal{A}\mathrm{d} : \Gamma \to \mathsf{O}(p,q)$. Even more, since by the classical theorem of *Cartan-Dieudonné* any orthogonal map can be written as composition of reflections (cf. [**Har90**, Thm. 4.23]), it follows from (1.1.23) that $\mathcal{A}\mathrm{d} : \Gamma \to \mathsf{O}(p,q)$ is *surjective*. Furthermore, one can show that the kernel of $\mathcal{A}\mathrm{d}$ consists precisely of all multiples of the identity.

[4] This is useful, e.g., when studying projective representations of $\mathsf{SO}(3)$ or $\mathsf{SO}^+(1,3)$, which show up in quantum mechanics. This gives an abstract "explanation" for the appearance of spin in quantum physics from a representation-theoretic point of view, see [**Wei05**, Sec. 2.7] or [**Sch95**, Sec. III.4]. We are only interested in the spin covering as an ingredient for the definition of spinors on Lorentzian manifolds.

[5] χ is the related to the natural grading of $\mathrm{Cl}(p,q)$ into even and odd elements, cf. [**Har90**, p. 183].

Instead of Γ, only the following subgroups of it are mostly of interest:

$$\mathsf{Pin}(p,q) := \{v_1 \cdots v_k \mid v_i \in \mathbb{R}^{p,q}, \langle v_i, v_i\rangle = \pm 1, \}\,, \tag{1.1.24}$$
$$\mathsf{Spin}(p,q) := \{v_1 \cdots v_{2k} \mid v_i \in \mathbb{R}^{p,q}, \langle v_i, v_i\rangle = \pm 1\}\,, \tag{1.1.25}$$
$$\mathsf{Pin}^+(p,q) := \{v_1 \cdots v_k \mid v_i \in \mathbb{R}^{p,q}, \langle v_i, v_i\rangle = \pm 1,\ \#(\text{timelike } v_i) = \text{even}\}\,, \tag{1.1.26}$$
$$\mathsf{Pin}_+(p,q) := \{v_1 \cdots v_k \mid v_i \in \mathbb{R}^{p,q}, \langle v_i, v_i\rangle = \pm 1,\ \#(\text{spacelike } v_i) = \text{even}\}\,, \tag{1.1.27}$$
$$\mathsf{Spin}^+(p,q) := \mathsf{Spin}(p,q) \cap \mathsf{Pin}^+(p,q) = \mathsf{Spin}(p,q) \cap \mathsf{Pin}_+(p,q)\,. \tag{1.1.28}$$

In their order of appearance, these subgroups are called *pin group*, *spin group*, *orthochronous pin group*, *proper pin group*, and *proper (or reduced) spin group*. Being (closed) subgroups of the group of invertible elements of $\mathrm{Cl}(p,q)$, which is an open subset of the finite-dimensional algebra $\mathrm{Cl}(p,q)$ and hence a Lie group, all these groups have a natural Lie group structure. Note that for a definite inner product, i.e. either $p = 0$ or $q = 0$, there are only the pin and the spin group since in these case all vectors are either spacelike or timelike.

The importance of these subgroups lies in the fact that the restriction of $\mathcal{A}\mathrm{d}$ to either one yields a 2 : 1 covering of one of the following well-known subgroups of $\mathsf{O}(p,q)$, listed to match the ordering of the (s)pin groups (cf. [**Har90**, Prop. 10.33]):

$$\mathsf{O}(p,q) \tag{1.1.29}$$
$$\mathsf{SO}(p,q) := \{A \in \mathsf{O}(p,q) \mid \det A = 1\} \tag{1.1.30}$$
$$\mathsf{O}^+(p,q) := \{A \in \mathsf{O}(p,q) \mid A \text{ preserves time-orientation}\} \tag{1.1.31}$$
$$\mathsf{O}_+(p,q) := \{A \in \mathsf{O}(p,q) \mid A \text{ preserves space-orientation}\} \tag{1.1.32}$$
$$\mathsf{SO}^+(p,q) := \mathsf{SO}(p,q) \cap \mathsf{O}^+(p,q) = \mathsf{SO}(p,q) \cap \mathsf{O}_+(p,q) = \mathsf{O}^+(p,q) \cap \mathsf{O}_+(p,q)\,. \tag{1.1.33}$$

For details on the definition of these groups, see [**Har90**, Ch. 4]. For our purposes only the last case is of interest, and we denote the 2 : 1 covering map of $\mathsf{SO}^+(p,q)$ obtained in this way by

$$\vartheta_{p,q} : \mathsf{Spin}^+(p,q) \to \mathsf{SO}^+(p,q)\,. \tag{1.1.34}$$

This covering is called the *spin covering* of $\mathsf{SO}^+(p,q)$.

Remark 1.1.8. The spin covering is universal precisely either for definite signatures $(n,0), (0,n)$ with $n \geq 3$, or for Lorentzian and "anti-Lorentzian" signatures $(1,n), (n,1)$ with $n \geq 3$, see [**Bau81**, Folgerung 1.2]. In particular it is universal for the "physically relevant" cases $\mathbb{R}^3$ and $\mathbb{R}^{1,3}$.

1.1.5. Description of the spin covering at the Lie algebra level. For later purposes it is of interest to describe the Lie algebra homomorphism $(\vartheta_{p,q})_* : \mathfrak{spin}^+(p,q) \to \mathfrak{so}^+(p,q)$ which is induced by the spin covering $\vartheta_{p,q} : \mathsf{Spin}^+(p,q) \to \mathsf{SO}^+(p,q)$. To this end, we first give a brief description of the Lie algebras $\mathfrak{spin}^+(p,q)$ of $\mathsf{Spin}^+(p,q)$ and $\mathfrak{so}^+(p,q)$ of $\mathsf{SO}^+(p,q)$.

Due to the inclusion $\mathsf{Spin}^+(p,q) \subset \mathrm{Cl}(p,q)$, one can identify $\mathfrak{spin}^+(p,q)$ with a linear subspace of the Clifford algebra. To determine this subspace one can differentiate suitable curves in $\mathsf{Spin}^+(p,q) \subset \mathrm{Cl}(p,q)$ passing through the identity element of $\mathsf{Spin}^+(p,q)$. Concretely, for any $1 \leq \mu < \nu \leq n$, consider the curve $\gamma_{\mu\nu} : \mathbb{R} \to \mathsf{Spin}^+(p,q) \subset \mathrm{Cl}(p,q)$

given by

$$\gamma_{\mu\nu}(t) := \begin{cases} (\cos(t)e_\mu - \sin(t)e_\nu)\cdot e_\mu & \langle e_\mu, e_\mu\rangle = -1, \langle e_\nu, e_\nu\rangle = -1 \\ (\cos(t)e_\mu + \sin(t)e_\nu)\cdot(-e_\mu) & \langle e_\mu, e_\mu\rangle = +1, \langle e_\nu, e_\nu\rangle = +1 \\ (\cosh(t)e_\mu - \sinh(t)e_\nu)\cdot e_\mu & \langle e_\mu, e_\mu\rangle = -1, \langle e_\nu, e_\nu\rangle = +1 \end{cases},$$

where $e_1, \dots, e_n \in \mathbb{R}^{p,q}$ is an orthonormal basis. Each of these curves satisfies $\gamma_{\mu\nu}(0) = 1$ and $\dot\gamma_{\mu\nu}(0) = e_\mu e_\nu$, and from this it follows that

$$\mathfrak{spin}^+(p,q) = \operatorname{span}\{e_\mu e_\nu \mid 1 \le \mu < \nu \le p+q\} \subset \mathrm{Cl}(p,q)\,. \tag{1.1.35}$$

The Lie algebra of $\mathsf{SO}^+(p,q)$ can also be determined by differentiating suitable curves in $\mathsf{SO}^+(p,q)$ passing through the identity element. Concretely one can take "rotation matrices", and one finds that $\mathfrak{so}^+(p,q)$ is spanned by the matrices $\{E_{\mu\nu}\}_{1\le\mu<\nu\le n}$, where

$$E_{\mu\nu} = \begin{array}{c} \\ \\ \end{array}\begin{array}{cc} \begin{pmatrix} & \vdots & & \vdots & \\ \cdots & 0 & \cdots & -\langle e_\nu, e_\nu\rangle & \cdots \\ & \vdots & & \vdots & \\ \cdots & \langle e_\mu, e_\mu\rangle & \cdots & 0 & \cdots \\ & \vdots & & \vdots & \end{pmatrix} & \begin{array}{l} \\ \leftarrow \mu\text{-th row} \\ \\ \leftarrow \nu\text{-th row} \\ \\ \end{array} \\ \begin{array}{ccc} \uparrow & \qquad & \uparrow \\ \mu\text{-th column} & & \nu\text{-th column} \end{array} & \end{array} \tag{1.1.36}$$

Finally, to compute the differential of the spin covering one can check that for any $1 \le \mu < \nu \le n$ and any $\alpha = 1, \dots, n$ it holds that

$$(\vartheta_{p,q})_*(e_\mu e_\nu)e_\alpha = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \vartheta_{p,q}(\gamma_{\mu\nu}(t))e_\alpha = \begin{cases} 0 & \alpha \ne \mu,\nu \\ 2\langle e_\mu, e_\mu\rangle e_\nu & \alpha = \mu \\ -2\langle e_\nu, e_\nu\rangle e_\mu & \alpha = \nu \end{cases}.$$

From this we can read off that

$$\boxed{(\vartheta_{p,q})_*(e_\mu e_\nu) = 2E_{\mu\nu}}\,. \tag{1.1.37}$$

1.1.6. Spin groups in low dimensions. Let us revitalize the presentation by computing some examples of spin groups. We are going to cover the same examples for which we also computed the Clifford algebras in Section 1.1.3.

Example 1.1.9. In the case of $\mathbb{R}$ with positive definite inner product, we have $e_1^2 = -1 \in \mathrm{Cl}(1)$. Therefore it follows immediately from (1.1.25) that $\mathsf{Spin}(1) = \{\pm 1\} = \mathbb{Z}_2$. Since $\mathsf{SO}(1) = \{1\}$, the spin covering is simply $\vartheta_1(\pm 1) = 1$.

Example 1.1.10. For the case of $\mathbb{R}^2$, we use the description of $\mathrm{Cl}(2)$ as given by Example 1.1.3. In this form, a unit vector $(\cos\alpha, \sin\alpha) \in \mathbb{R}^2$ is represented in $\mathrm{Cl}(2)$ by the matrix

$$\begin{pmatrix} 0 & \cos\alpha + \mathrm{i}\sin\alpha \\ -\cos\alpha + \mathrm{i}\sin\alpha & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathrm{e}^{\mathrm{i}\alpha} \\ -\mathrm{e}^{-\mathrm{i}\alpha} & 0 \end{pmatrix}.$$

To compute $\mathsf{Spin}(2)$, according to its definition (1.1.25) we need to compute products of an even number of unit vectors. For the product of two unit vectors, we find

$$\begin{pmatrix} 0 & \mathrm{e}^{\mathrm{i}\alpha} \\ -\mathrm{e}^{-\mathrm{i}\alpha} & 0 \end{pmatrix}\begin{pmatrix} 0 & \mathrm{e}^{\mathrm{i}\beta} \\ -\mathrm{e}^{-\mathrm{i}\beta} & 0 \end{pmatrix} = \begin{pmatrix} -\mathrm{e}^{\mathrm{i}(\alpha-\beta)} & 0 \\ 0 & -\mathrm{e}^{-\mathrm{i}(\alpha-\beta)} \end{pmatrix}.$$

It follows by induction that

$$\mathsf{Spin}(2) = \left\{ \begin{pmatrix} \mathrm{e}^{\mathrm{i}\theta} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\theta} \end{pmatrix} \,\middle|\, \theta \in \mathbb{R} \right\} \cong \mathsf{U}(1) \cong S^1 \cong \mathsf{SO}(2)\,. \tag{1.1.38}$$

The spin covering $\vartheta_2 : \mathsf{Spin}(2) \to \mathsf{SO}(2)$ can be easily computed by writing out explicitly the action of $\mathsf{Spin}(2)$ on $\mathbb{R}^2$. Identifying both $\mathsf{Spin}(2)$ and $\mathsf{SO}(2)$ with $U(1) \cong S^1$, one finds that it is given by $\vartheta_2(z) = z^2$.

Example 1.1.11. In order to compute $\mathsf{Spin}(3)$ we use two tricks. Firstly, we use that $\mathsf{Spin}(3) = \mathsf{Spin}(0,3) \cong \mathsf{Spin}(3,0)$. While it generally holds that $\mathsf{Spin}(p,q) = \mathsf{Spin}(q,p)$, here this can also be seen from the fact that $\mathsf{SO}(3) = \mathsf{SO}(0,3) = \mathsf{SO}(3,0)$ together with the fact that in the definite, three-dimensional case the spin covering is the universal covering, and the universal covering is unique. Secondly, instead of directly determining $\mathsf{Spin}(3,0)$ as we did in the previous examples, we rather determine its Lie algebra $\mathfrak{spin}(3,0)$ using (1.1.35). To this end we use the representation of $\mathrm{Cl}(3,0)$ constructed in terms of Pauli matrices in Example 1.1.4. From this and the second identity for the Pauli matrices in (1.1.14) it follows that

$$\mathfrak{spin}(3,0) = \mathrm{span}_{\mathbb{R}}\{\mathrm{i}\sigma_1, \mathrm{i}\sigma_2, \mathrm{i}\sigma_3\}\,. \tag{1.1.39}$$

Finally, since $\mathsf{Spin}(3,0)$ is simply connected (the spin covering is universal), it follows that[6]

$$\mathsf{Spin}(3,0) = \left\{ \exp\left(-\tfrac{\mathrm{i}}{2}\vec{\alpha}\cdot\vec{\sigma}\right) \middle| \vec{\alpha} \in \mathbb{R}^3 \right\} = \mathsf{SU}(2)\,. \tag{1.1.40}$$

The negative sign here is just convention, whereas the factor of $\frac{1}{2}$ has the practical consequence that $\vartheta_3(\exp(-\frac{\mathrm{i}}{2}\vec{\alpha}\cdot\vec{\sigma})) \in \mathsf{SO}(3)$ is precisely the rotation around the axis $\vec{\alpha} \in \mathbb{R}^3$ by the angle $|\vec{\alpha}|$. This can be seen from equation (1.1.37), and already describes the spin covering $\vartheta_3 : \mathsf{Spin}(3) \to \mathsf{SO}(3)$.

There are of course other ways to see all this, in particular one can find many more elementary approaches relying on nice explicit computations in the physics literature, see for instance [**Sak94**, Ch. 3].

Now we turn to Lorentzian examples.

Example 1.1.12. The case $\mathbb{R}^{1,1}$ can be worked out similarly to $\mathbb{R}^2$. However, instead of using the representation of $\mathrm{Cl}(1,1)$ presented in Example 1.1.5, it is more advantageous to use a representation where the basis $e_0, e_1 \in \mathbb{R}^{1,1}$ corresponds to the matrices

$$\gamma_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \quad \gamma_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,. \tag{1.1.41}$$

In this way one also obtains an isomorphism $\mathrm{Cl}(1,1) \cong \mathsf{M}(2,\mathbb{R})$. Now any unit vector $v \in \mathbb{R}^{1,1}$ can be written as

$$v = \begin{cases} e_0 \cosh\theta + e_1 \sinh\theta & \text{if } \langle v, v\rangle = -1 \\ e_0 \sinh\theta + e_1 \cosh\theta & \text{if } \langle v, v\rangle = 1 \end{cases}$$

[6]Here we use arrows to denote vectors in $\mathbb{R}^3$ and write $\vec{\alpha}\cdot\vec{\sigma} = \alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3$ for $\vec{\alpha} \in \mathbb{R}^3$.

for a unique $\theta \in \mathbb{R}$, called the *hyperbolic angle* of v. Therefore, as element of $\mathrm{Cl}(1,1) \cong \mathsf{M}(2,\mathbb{R})$ with the isomorphism as just described, v is given by the matrix

$$v = \begin{cases} \begin{pmatrix} 0 & \mathrm{e}^{-\theta} \\ \mathrm{e}^{\theta} & 0 \end{pmatrix} & \langle v, v \rangle = -1 \\ \begin{pmatrix} 0 & -\mathrm{e}^{-\theta} \\ \mathrm{e}^{\theta} & 0 \end{pmatrix} & \langle v, v \rangle = 1 \end{cases}$$

This simple form is what makes the choice (1.1.41) more convenient here. As before, in order to determine $\mathsf{Spin}(1,1)$, we have to compute products of even numbers of unit vectors. As one easily checks, for the product of two unit vectors $v, w \in \mathbb{R}^{1,1}$ there are four possibilities depending on the causal characters of v and w:

(1) Both v and w are timelike with hyperbolic angles $\theta, \varphi \in \mathbb{R}$:

$$v \cdot w = \begin{pmatrix} \mathrm{e}^{-\theta+\varphi} & 0 \\ 0 & \mathrm{e}^{\theta-\varphi} \end{pmatrix}.$$

(2) Both v and w are spacelike with hyperbolic angles $\theta, \varphi \in \mathbb{R}$:

$$v \cdot w = -\begin{pmatrix} \mathrm{e}^{-\theta+\varphi} & 0 \\ 0 & \mathrm{e}^{\theta-\varphi} \end{pmatrix}.$$

(3) v is timelike and w is spacelike with hyperbolic angles $\theta, \varphi \in \mathbb{R}$:

$$v \cdot w = \begin{pmatrix} \mathrm{e}^{-\theta+\varphi} & 0 \\ 0 & -\mathrm{e}^{\theta-\varphi} \end{pmatrix}.$$

(3) v is spacelike and w is timelike with hyperbolic angles $\theta, \varphi \in \mathbb{R}$:

$$v \cdot w = -\begin{pmatrix} \mathrm{e}^{-\theta+\varphi} & 0 \\ 0 & -\mathrm{e}^{\theta-\varphi} \end{pmatrix}.$$

From this one easily shows by induction that

$$\mathsf{Spin}(1,1) = \left\{ \pm \begin{pmatrix} \mathrm{e}^{\theta} & 0 \\ 0 & \mathrm{e}^{-\theta} \end{pmatrix}, \pm \begin{pmatrix} \mathrm{e}^{\theta} & 0 \\ 0 & -\mathrm{e}^{-\theta} \end{pmatrix} \middle| \; \theta \in \mathbb{R} \right\}, \tag{1.1.42}$$

$$\mathsf{Spin}^+(1,1) = \left\{ \pm \begin{pmatrix} \mathrm{e}^{\theta} & 0 \\ 0 & \mathrm{e}^{-\theta} \end{pmatrix} \middle| \; \theta \in \mathbb{R} \right\}. \tag{1.1.43}$$

Notice that we have an isomorphism $\mathsf{Spin}^+(1,1) \cong \mathsf{GL}(1,\mathbb{R}) \cong \mathbb{R} \setminus \{0\}$ given by

$$\mathbb{R} \setminus \{0\} \ni x \longmapsto \begin{pmatrix} x & 0 \\ 0 & \frac{1}{x} \end{pmatrix} \in \mathsf{Spin}^+(1,1). \tag{1.1.44}$$

Finally, computing explicitly the action of $\mathsf{Spin}(1,1)$ on $\mathbb{R}^{1,1}$, one finds that the spin covering $\vartheta_{1,1} : \mathsf{Spin}^+(1,1) \to \mathsf{SO}^+(1,1)$ is given by

$$\vartheta_{1,1}\left(\pm \begin{pmatrix} \mathrm{e}^{\theta} & 0 \\ 0 & \mathrm{e}^{-\theta} \end{pmatrix} \right) = \begin{pmatrix} \cosh 2\theta & -\sinh 2\theta \\ -\sinh 2\theta & \cosh 2\theta \end{pmatrix}. \tag{1.1.45}$$

For 1+2 and 1+3 dimensions, it is actually easier to first cover the higher-dimensional case since here one can again exploit that the spin covering is the universal covering.

Example 1.1.13. For $\mathsf{Spin}^+(1,3)$ one can again exploit that the spin covering is the universal covering of $\mathsf{SO}^+(1,3)$, and so $\mathsf{Spin}^+(1,3)$ is simply connected. In this way one finds that $\mathsf{Spin}^+(1,3) \cong SL(2,\mathbb{C})$ as a real Lie group, since this is a simply connected Lie group with the same Lie algebra as $\mathsf{SO}^+(1,3)$. The spin covering can be obtained explicitly by integrating any Lie algebra isomorphism between $\mathfrak{sl}(2,\mathbb{C})$ and $\mathfrak{so}^+(1,3)$, similar as in Example 1.1.11. Explicit computations can be found in many physics textbooks, or also in [**HT94**, Ch. 3].

Example 1.1.14. To determine $\mathsf{Spin}^+(1,2)$, one possible way is to embed $\mathbb{R}^{1,2}$ into $\mathbb{R}^{1,3}$, and construct corresponding embeddings of the Clifford algebra and spin group. Therefore one obtains an embedding $\mathsf{Spin}^+(1,2) \subset \mathsf{Spin}^+(1,3) \cong \mathsf{SL}(2,\mathbb{C})$. If done in the correct way, one can simply read off that $\mathsf{Spin}^+(1,2) \cong \mathsf{SL}(2,\mathbb{R})$. To do this in detail, one has to work out the $1+3$-dimensional case first, which we did not do explicitly.

1.1.7. The spin representation and spinors. Now we finally come to spinors, which are vectors of a certain representation space of the (complex) Clifford algebra, or rather of the spin group, called (complex) spin representation.[7]

First of all, we recall from (1.1.12) that any complex Clifford algebra $\mathbb{C}\mathrm{l}(p,q)$ is either isomorphic to $\mathsf{M}(N,\mathbb{C})$ in case that $p+q$ is even, or it is isomorphic to $\mathsf{M}(N,\mathbb{C}) \oplus \mathsf{M}(N,\mathbb{C})$ in case that $p+q$ is odd. To continue from here, we need one general piece of information about representations of algebras, more specifically of matrix algebras.

If $\mathcal{A}$ is a complex unital algebra, then by a *representation* of $\mathcal{A}$ on a complex vector space V we understand a homomorphism of *complex* unital algebras $\rho : \mathcal{A} \to \mathsf{End}(V)$.[8] Further, a representation $\rho : \mathcal{A} \to \mathsf{End}(V)$ is called *irreducible* if there does not exist any nontrivial linear subspace $U \subset V$ such that $\rho(a)U \subset U$ for all $a \in \mathcal{A}$. Finally, two representations $\rho_1 : \mathcal{A} \to \mathsf{End}(V_1)$ and $\rho_2 : \mathcal{A} \to \mathsf{End}(V_2)$ are called *equivalent* if there exists an isomorphism $\phi : V_1 \to V_2$ of complex vector spaces such that

$$\rho_1(a) = \phi^{-1} \circ \rho_2(a) \circ \phi \qquad \forall a \in \mathcal{A}. \tag{1.1.46}$$

Having recalled the necessary concepts about representations, the basic result we need is the following:

Proposition 1.1.15. (Irreducible representations of matrix algebras)

i.) Up to equivalence, the complex algebra $\mathsf{M}(N,\mathbb{C})$ *has precisely one irreducible representation on a complex vector space, namely the standard one by matrix multiplication on* $\mathbb{C}^N$.

ii.) Up to equivalence, the complex algebra $\mathsf{M}(N,\mathbb{C}) \oplus \mathsf{M}(N,\mathbb{C})$ *has precisely two inequivalent irreducible representations on a complex vector space. These are the actions on* $\mathbb{C}^N$ *where* (A_1, A_2) *acts by* A_1 *or by* A_2.

The proof of this statement is not very difficult and, although for some reason not mentioned there explicitly, can easily be derived from the contents of [**Har90**, Ch. 8]. The key observation is that matrix algebras are *simple*, i.e. possess no nontrivial two-sided ideals.

Returning to Clifford algebras, for $p+q$ *even* it follows from Proposition 1.1.15 and the isomorphism $\mathbb{C}\mathrm{l}(p,q) \cong \mathsf{M}(N,\mathbb{C})$ that $\mathbb{C}\mathrm{l}(p,q)$ possesses precisely one irreducible

[7] As a word of warning, the general theory of representations of Clifford algebras is somewhat involved since there are many different structures around which depend on signature and dimension. For instance, one constantly has $\mathbb{R}, \mathbb{C}$, and $\mathbb{H}$ lurking around at the same time, which forces one to be quite careful about which maps are linear over which of these fields. Therefore looking up the word "spin representation" in a book on spin geometry or Clifford algebras such as [**Har90**] or [**LM89**] can be a little overwhelming. Here we will be as simple as possible for our purposes.

[8] The reason why we stress the word "complex" here is that often, in particular in the literature about representations of Clifford algebras, representations are only assumed to be $\mathbb{R}$-linear. While this is more convenient for certain purposes, here it would lead to unnecessary complications.

representation $\mathbb{C}\mathrm{l}(p,q) \to \mathsf{End}(\mathbb{S}_{p,q})$ on a complex vector space $\mathbb{S}_{p,q}$, called *spin space*. Restricting this representation to $\mathsf{Spin}^+(p,q) \subset \mathbb{C}\mathrm{l}(p,q)$ we obtain a complex representation of the spin group. If $p+q$ is *odd*, then it follows from Proposition 1.1.15 and the isomorphism $\mathbb{C}\mathrm{l}(p,q) \cong \mathsf{M}(N,\mathbb{C}) \oplus \mathsf{M}(N,\mathbb{C})$ that $\mathbb{C}\mathrm{l}(p,q)$ has precisely two (inequivalent) representations. As it turns out, however, their restrictions to the spin group are equivalent representations for the following reason:

Let $\mathbb{C}\mathrm{l}^0(p,q) \subset \mathbb{C}\mathrm{l}(p,q)$ be the subalgebra of *even elements*, i.e. the linear subspace spanned by all elements of the form $v_1 \cdots v_{2k} \in \mathbb{C}\mathrm{l}(p,q)$ for $v_1, \dots, v_{2k} \in \mathbb{R}^{p,q}$. Notice that this is indeed a subalgebra, and that by (1.1.25) we have $\mathsf{Spin}^+(p,q) \subset \mathbb{C}\mathrm{l}^0(p,q)$. Next, it is not difficult to see that $\mathbb{C}\mathrm{l}^0(p,q)$ is isomorphic to $\mathbb{C}\mathrm{l}(p-1,q)$ (cf. [**Har90**, Thm. 9.38]), which is isomorphic to $\mathsf{M}(N,\mathbb{C})$. Therefore $\mathbb{C}\mathrm{l}^0(p,q)$ is a simple (unital) algebra, from which it follows that the restriction of either of the two irreducible representations of $\mathbb{C}\mathrm{l}(p,q)$ to $\mathbb{C}\mathrm{l}^0(p,q)$ is injective since the kernel (as linear map) is always a two-sided ideal. For dimensional reasons these restrictions must therefore be isomorphisms, hence irreducible in particular. But then, again since $\mathbb{C}\mathrm{l}^0(p,q) \cong \mathsf{M}(N,\mathbb{C})$, it follows from Proposition 1.1.15 that they must be equivalent as representations of $\mathbb{C}\mathrm{l}^0(p,q)$, hence also as representations of $\mathsf{Spin}^+(p,q) \subset \mathbb{C}\mathrm{l}^0(p,q)$.

So we have seen that no matter whether $p+q$ is odd or even, we always end up with a uniquely determined representation of $\mathsf{Spin}^+(p,q)$.

Definition 1.1.16. The *(complex) spin representation*

$$\rho_{p,q} : \mathsf{Spin}^+(p,q) \to \mathsf{GL}(\mathbb{S}_{p,q}) \tag{1.1.47}$$

is defined to be the restriction of any of the irreducible representations of $\mathbb{C}\mathrm{l}(p,q)$ to $\mathsf{Spin}^+(p,q) \subset \mathbb{C}\mathrm{l}(p,q)$. The space $\mathbb{S}_{p,q}$ is called *spin space* or *space of spinors.*

There exist two further important algebraic structures on the space of spinors. Firstly, notice that by construction we have a linear map

$$\gamma_{\mathbb{S}_{p,q}} : \mathbb{R}^{p,q} \to \mathsf{End}(\mathbb{S}_{p,q}) \tag{1.1.48}$$

by letting each vector in $\mathbb{R}^{p,q} \subset \mathbb{C}\mathrm{l}(p,q)$ act on $\mathbb{S}_{p,q}$ via the action of $\mathbb{C}\mathrm{l}(p,q)$. This map is called *Clifford multiplication.* It satisfies the Clifford relations

$$\{\gamma_{\mathbb{S}_{p,q}}(v), \gamma_{\mathbb{S}_{p,q}}(w)\} = -2\,\langle v,w\rangle\,\mathbb{1}_{\mathbb{S}_{p,q}} \qquad \forall v,w \in \mathbb{R}^{p,q}\,, \tag{1.1.49}$$

where $\{\cdot,\cdot\}$ again denotes the anticommutator of two endomorphisms. variance-property with respect to the spin representation $\rho_{p,q}$ and the spin covering $\vartheta_{p,q} : \mathsf{Spin}^+(p,q) \to \mathsf{SO}^+(p,q)$.[9]

Lemma 1.1.17. *For any $v \in \mathbb{R}^{p,q}$ and $g \in \mathsf{Spin}^+(p,q)$ we have*

$$\gamma_{\mathbb{S}_{p,q}}(\vartheta_{p,q}(g)v) \circ \rho_{p,q}(g) = \rho_{p,q}(g) \circ \gamma_{\mathbb{S}_{p,q}}(v)\,, \tag{1.1.50}$$

where $\vartheta_{p,q} : \mathsf{Spin}^+(p,q) \to \mathsf{SO}^+(p,q)$ is the spin covering.

Proof. In the following, we also denote the action of $\mathbb{C}\mathrm{l}(p,q)$ on $\mathbb{S}_{p,q}$ by $\rho_{p,q}$. Then we simply compute

$$\begin{aligned}\rho_{p,q}(g) \circ \gamma_{\mathbb{S}_{p,q}}(v) &= \rho_{p,q}(g)\rho_{p,q}(v)\\ &= \rho_{p,q}(g)\rho_{p,q}(v)\rho_{p,q}(g^{-1})\rho_{p,q}(g)\\ &= \rho_{p,q}(gvg^{-1})\rho_{p,q}(g)\end{aligned}$$

[9] Such equivariance properties are crucial if we later want to "lift" these algebraic structures to the spinor bundle.

$$
\begin{aligned}
&= \rho_{p,q}(\vartheta_{p,q}(g)v)\rho_{p,q}(g) \\
&= \gamma_{\mathbb{S}_{p,q}}(\vartheta_{p,q}(g)v) \circ \rho_{p,q}(q)\,.
\end{aligned}
$$

Here we used the definition of $\vartheta_{p,q}$ in terms of the twisted adjoint action (1.1.22), and that any $g \in \mathsf{Spin}^+(p,q)$ is even, i.e. $\chi(g) = g$. □

Secondly, $\mathbb{S}_{p,q}$ admit some useful invariant inner products. Optimally, one would like to have a positive definite, $\mathsf{Spin}^+(p,q)$-invariant inner product. Unfortunately, or interestingly, this is only possible in the case of definite signature $(p,0)$ or $(0,q)$.

Lemma 1.1.18. *Let $p,q \neq 0$. Then $\mathbb{S}_{p,q}$ does not admit any positive definite, $\mathsf{Spin}^+(p,q)$-invariant inner product.*

Proof. Suppose for the sake of contradiction that $(\cdot,\cdot)$ were such an inner product. Since $p,q \neq 0$, there exists a pair of orthogonal vectors $v,w \in \mathbb{R}^{p,q}$ with $\langle v,v\rangle = -1$ and $\langle w,w\rangle = 1$. As we have seen in Section 1.1.5, we have $vw \in \mathfrak{spin}^+(p,q)$. Therefore, since $(\cdot,\cdot)$ is $\mathsf{Spin}^+(p,q)$-invariant, $vw \in \mathrm{Cl}(p,q)$ must act skew-symmetrically. But then it follows from positivity of $(\cdot,\cdot)$ that for any $\psi \in \mathbb{S}_{p,q}$ we have

$$
0 \leq (\rho_{p,q}(vw)\psi, \rho_{p,q}(vw)\psi) = -(\psi, \rho_{p,q}(vw)\rho_{p,q}(vw)\psi) = -(\psi,\psi)\,. \tag{$*$}
$$

Here we used that due to the Clifford relations we have

$$
\rho_{p,q}(vw)\rho_{p,q}(vw) = \rho_{p,q}(vwvw) = -\rho_{p,q}(vvww) = 1\,.
$$

Now $(*)$ is obviously a contradiction to the assumption that $(\cdot,\cdot)$ is positive definite. □

Nevertheless, there do exist *indefinite* $\mathsf{Spin}^+(p,q)$-invariant inner products. To see this, we begin by fixing some arbitrarily chosen positive definite Hermitian inner product $(\cdot,\cdot)$ on $\mathbb{S}_{p,q}$ and an orthonormal basis $e_1,\dots,e_{p+q} \in \mathbb{R}^{p,q}$ where the first p vectors of this basis are timelike. It follows from the Clifford relations that the subgroup $G \subset \mathrm{Cl}(p,q)$ which is generated by $\pm 1, e_1, \dots, e_{p+q}$ is a finite group. Therefore we can define another positive definite inner product $(\cdot,\cdot)'$ by averaging over this group, i.e. we set

$$
(\psi,\phi)' := \sum_{g\in G} (\rho_{p,q}(g)\psi, \rho_{p,q}(g)\phi)\,. \tag{1.1.51}
$$

As one easily verifies, this new inner product is G-invariant, so in particular we have

$$
(\gamma_{\mathbb{S}_{p,q}}(e_\mu)\psi, \gamma_{\mathbb{S}_{p,q}}(e_\mu)\phi)' = (\psi,\phi)' \qquad \forall \mu = 1,\dots,p+q \quad \forall \psi,\phi \in \mathbb{S}_{p,q}\,.
$$

Therefore, for all timelike vectors of the basis we have

$$
\begin{aligned}
\forall \mu = 1,\dots,p: \qquad (\gamma_{\mathbb{S}_{p,q}}(e_\mu)\psi, \phi)' &= (\underbrace{\gamma_{\mathbb{S}_{p,q}}(e_\mu)\gamma_{\mathbb{S}_{p,q}}(e_\mu)}_{=1}\psi, \gamma_{\mathbb{S}_{p,q}}(e_\mu)\phi)' \\
&= (\psi, \gamma_{\mathbb{S}_{p,q}}(e_\mu)\phi)'\,,
\end{aligned} \tag{1.1.52}
$$

whereas for all spacelike vectors of the basis we have

$$
\begin{aligned}
\forall \nu = p+1,\dots,p+q: \qquad (\gamma_{\mathbb{S}_{p,q}}(e_\nu)\psi, \phi)' &= (\underbrace{\gamma_{\mathbb{S}_{p,q}}(e_\nu)\gamma_{\mathbb{S}_{p,q}}(e_\nu)}_{=-1}\psi, \gamma_{\mathbb{S}_{p,q}}(e_\nu)\phi)' \\
&= -(\psi, \gamma_{\mathbb{S}_{p,q}}(e_\nu)\phi)'\,.
\end{aligned} \tag{1.1.53}
$$

In the case of definite signature this (positive definite) inner product is already Spin-invariant.

Proposition 1.1.19. *Suppose that either $p = 0$ or $q = 0$ and set $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}} := (\cdot,\cdot)'$ with $(\cdot,\cdot)'$ defined by* (1.1.51)*. Then Clifford multiplication is skew-symmetric with respect to $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$ if $p = 0$, and symmetric if $q = 0$. Consequently, in both cases $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$ is $\mathsf{Spin}(p,q)$-invariant.*

Proof. First let $p = 0$. Then it follows directly from (1.1.53) that actually Clifford multiplication by *any* vector in $\mathbb{R}^{0,q}$ is skew-symmetric. In particular, this implies that any unit vector $v \in \mathbb{R}^{0,q}$ acts orthogonally since

$$(\gamma_{\mathbb{S}_{0,q}}(v)\psi, \gamma_{\mathbb{S}_{0,q}}(v)\phi)' = -(\psi, \gamma_{\mathbb{S}_{0,q}}(v)^2\phi)' = (\psi,\phi)' .$$

Since every element of the spin group is a product of an even number of unit vectors, $\mathsf{Spin}(0,q)$-invariance follows immediately. For $q = 0$ one can proceed along the same arguments. □

In the case of indefinite signature $p, q \neq 0$ the positive definite inner product (1.1.51) cannot be $\mathsf{Spin}^+(p,q)$-invariant due to Lemma 1.1.18. Here the trick is to redefine the inner product once more by setting

$$\prec\psi,\phi\succ_{\mathbb{S}_{p,q}} := \mathrm{i}^{\frac{p(p-1)}{2}} (\gamma_{\mathbb{S}_{p,q}}(e_1)\cdots\gamma_{\mathbb{S}_{p,q}}(e_p)\psi,\phi)' \qquad \forall \psi,\phi \in \mathbb{S}_{p,q} . \tag{1.1.54}$$

Proposition 1.1.20. *Suppose that $p, q \neq 0$. Then* (1.1.54) *defines an indefinite Hermitian inner product on $\mathbb{S}_{p,q}$. Furthermore, Clifford multiplication is symmetric with respect to $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$ if p is odd, and skew-symmetric if p is even. As a consequence, in either case $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$ is $\mathsf{Spin}^+(p,q)$-invariant.*

Proof. That $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$ is Hermitian can easily be verified using (1.1.52) and the Clifford relations. That it must be indefinite follows from Lemma 1.1.18 once we have shown $\mathsf{Spin}^+(p,q)$-invariance.

Concerning (skew-)symmetry of Clifford multiplication, due to linearity it suffices to show this for the basis vectors $e_1, \dots, e_{p+q}$. For these it follows easily from (1.1.52), (1.1.53) and the Clifford relations.

Concerning the $\mathsf{Spin}^+(p,q)$-invariance, notice first that for any $v \in \mathbb{R}^{p,q}$ we have

$$\prec\gamma_{\mathbb{S}_{p,q}}(v)\psi, \gamma_{\mathbb{S}_{p,q}}(v)\phi\succ_{\mathbb{S}_{p,q}} = (-1)^{p+1} \prec\psi, \gamma_{\mathbb{S}_{p,q}}(v)^2\phi\succ_{\mathbb{S}_{p,q}} = (-1)^{p+1}\langle v,v\rangle \prec\psi,\phi\succ_{\mathbb{S}_{p,q}} .$$

From this the $\mathsf{Spin}^+(p,q)$-invariance follows by noting once more that every element of $\mathsf{Spin}^+(p,q)$ is the product of an even number of timelike unit vectors and an even number of spacelike unit vectors. □

Let us sum up the discussion about invariant inner products so far in a corollary.

Corollary 1.1.21. (Invariant inner products on spinors) *For every signature (p,q), there exists a $\mathsf{Spin}^+(p,q)$-invariant Hermitian inner product $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$ on $\mathbb{S}_{p,q}$. It is positive definite if and only if $p = 0$ or $q = 0$.*

Let us also note that in the Lorentzian case, i.e. $p = 1$, formula (1.1.54) reduces to

$$\prec\psi,\phi\succ_{\mathbb{S}_{1,n}} := (\gamma_{\mathbb{S}_{p,q}}(e_0)\psi,\phi)' \qquad \forall \psi,\phi \in \mathbb{S}_{p,q} . \tag{1.1.55}$$

As usual in the Lorentzian situation, here we have shifted the enumeration of basis vectors to $e_0, e_1, \dots, e_n$ with e_0 timelike and $e_1, \dots, e_n$ spacelike.

Remark 1.1.22. In the case of definite signature $(p,0)$ or $(0,q)$, the hermitian inner product (1.1.51) on the spin space is positive definite, see Proposition 1.1.19. In the case of indefinite signature (p,q), i.e. $p, q \neq 0$, the signature of $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$ is always split, i.e. $(\frac{N}{2}, \frac{N}{2})$, where $N = \dim \mathbb{S}_{p,q}$ is always even (cf. [**Bau81**, Satz 1.12] or [**Har90**, Thm. 13.1]). For signature $(1,1)$, $(1,2)$, and $(1,3)$, for which we have explicitly computed the Clifford algebras and spin groups in Section 1.1.3 and Section 1.1.6, one can verify this by hand.

Let us make one last, important observation about inner products on spin space in Lorentzian signature. For any vector $v \in \mathbb{R}^{1,n}$ we consider the inner product

$$\langle \cdot, \cdot \rangle_v := \prec \cdot, \gamma_{\mathbb{S}_{1,n}}(v) \cdot \succ_{\mathbb{S}_{1,n}} , \tag{1.1.56}$$

where $\prec \cdot, \cdot \succ_{\mathbb{S}_{1,n}}$ is the $\mathsf{Spin}^+(1,n)$-invariant inner product on $\mathbb{S}_{1,n}$. Notice that this is still Hermitian since $\gamma_{\mathbb{S}_{1,n}}(v)$ is symmetric with respect to $\prec \cdot, \cdot \succ_{\mathbb{S}_{1,n}}$. The important observation is that this inner product is *positive definite* if v is timelike and future-pointing. Although it is not $\mathsf{Spin}^+(1,n)$-invariant it is nevertheless very important for the Dirac equation on globally hyperbolic Lorentzian manifolds (see Section 2.1).

Lemma 1.1.23. *Consider the case of Lorentzian signature, and let $v \in \mathbb{R}^{1,n}$ be timelike and future-pointing. Then the inner product $\langle \cdot, \cdot \rangle_v := \prec \cdot, \gamma_{\mathbb{S}_{1,n}}(v) \succ_{\mathbb{S}_{1,n}}$ is positive definite.*

Proof. Let $e_0, \ldots, e_n \in \mathbb{R}^{1,n}$ be the standard basis and set $\gamma_\mu := \gamma_{\mathbb{S}_{1,n}}(e_\mu)$. By (1.1.55) we have $\prec \cdot, \cdot \succ_{\mathbb{S}_{1,n}} = (\cdot, \gamma_0 \cdot)'$, where $(\cdot, \cdot)'$ is positive definite. We decompose v as $v = v^\mu e_\mu$ and note that $v^0 > 0$ since v is future-pointing. Next, we consider the continuous curve of future-pointing timelike vectors given by $v(t) = v^0 e_0 + t v^j e_j$, where $t \in [0,1]$. Here we use that the set of future-pointing timelike vectors is convex. We have

$$\prec \cdot, \gamma_{\mathbb{S}_{1,n}}(v(t)) \cdot \succ_{\mathbb{S}_{1,n}} = \prec \cdot, v^0 \gamma_0 + t v^j \gamma_j \cdot \succ_{\mathbb{S}_{1,n}} = \big(\cdot, (v^0 \mathbb{1} + t v^j \gamma_0 \gamma_j) \cdot \big)' .$$

If we can show that all eigenvalues of $v^0 \mathbb{1} + t v^j \gamma_0 \gamma_j$ are strictly positive for all $t \in [0,1]$, then we are done. To this end, observe first that for $t = 0$ the eigenvalues are all just $v^0 > 0$. Secondly, by the Clifford relations we have $\gamma_{\mathbb{S}_{1,n}}(v(t))^2 = -\langle v(t), v(t) \rangle > 0$, and since also $\gamma_0^2 = 1$ it follows that $v^0 \mathbb{1} + t v^j \gamma_0 \gamma_j = \gamma_0 \gamma_{\mathbb{S}_{1,n}}(v(t))$ is invertible, hence all of its eigenvalues are nonzero. But then, since the eigenvalues depend continuously on t they must remain positive for all $t \in [0,1]$. This concludes the proof. $\square$

1.2. Spin Structures, Spinor Bundles, and the Dirac Operator

Now we turn to geometry and explain how one can introduce spinors and the Dirac equation on curved spacetimes, using the algebraic machinery summarized previously. Since the whole construction is rather abstract, let us start however with some words about the (possible) "meaning" of a spinor.

1.2.1. What is a spinor? In Section 1.1.7 we have defined spinors as vectors in a representation space of the Clifford algebra. While this is of course a perfectly valid mathematical definition, it does not shed much light on any possible meaning of a spinor. In the following we will try to make the concept of a spinor more "concrete" through various elaborations.

Let us start with physics and Dirac's discovery of spinors. In his attempt to find a relativistic analogue of the Schrödinger equation, Dirac was lead to taking the "square root" of the Klein-Gordon equation, i.e. the equation

$$(\Box - m^2)\phi = 0 , \tag{1.2.1}$$

where $\phi : \mathbb{R}^{1,3} \to \mathbb{C}$ is a scalar function on Minkowski spacetime and $\Box = -\partial_t^2 + \Delta_{\mathbb{R}^3}$ is the wave operator. The reason was that although (1.2.1) is Lorentz-invariant, it does not allow an interpretation of a general solution ϕ along the usual lines of quantum mechanics, i.e. as a "probability amplitude" (cf. [**Dir28**] or [**Sak67**, Sec. 3.1]). Since the problems were at least partially related to the fact that (1.2.1) is a second-order equation, Dirac's idea was to find a first-order differential operator $\mathcal{D}$ satisfying

$$\mathcal{D}^2 = \Box \, . \tag{1.2.2}$$

Then it would follow that

$$(\mathcal{D}+m)(\mathcal{D}-m) = \Box - m^2 \, ,$$

so instead of studying all solutions of (1.2.1) one could study more special solutions which are already solutions of the *massive Dirac equation*

$$\mathcal{D}\psi = m\psi \, , \tag{1.2.3}$$

which in fact allows an interpretation of ψ by the usual rules of quantum mechanics as a probability amplitude (cf. [**Sak67**, Sec. 3.1]).

Making the general Ansatz $\mathcal{D} = \mathrm{i}\gamma^\mu \partial_\mu$ for a first-order operator, a quick computation reveals that $\mathcal{D}$satisfies (1.2.2) if and only if the coefficients γ^μ satisfy the *Clifford relations*

$$\{\gamma^\mu, \gamma^\nu\} = -2\eta^{\mu\nu} \qquad \forall \mu, \nu = 0, \ldots, 3 \, , \tag{1.2.4}$$

where $\eta^{\mu\nu}$ denotes the Minkowski metric. Now as is not difficult to see, these relations cannot be satisfied if the γ^μ are taken to be complex numbers, but it works if one allows them to be matrices. This means that ψ in (1.2.3) cannot be just a scalar function, but must instead be a vector-valued function $\psi : \mathbb{R}^{1,3} \to \mathbb{C}^N$ for some suitable $N > 1$. As we of course already know from the previous section, choosing the γ^μ such that (1.2.4) holds means nothing but that we choose a representation of the Clifford algebra $\mathrm{Cl}(1, 3)$ on $\mathbb{C}^N$, and consequently also obtain a representation of the spin group on $\mathbb{C}^N$.

This connects the abstract definition of a spinor as vector in a representation space of the Clifford algebra to Dirac's original idea. Still it does not really give spinors a "meaning". Concerning this question, I want to challenge the reader to briefly reflect on which kind of answer he or she would accept at all to this. In the following I will touch on two further aspects which might be helpful.

First, most of us probably would like to have some intuitive understanding of the abstract concept of a spinor, at best anchored somehow in the "real world" (at least as a physics-inclined person). For instance, the abstract notion of a manifold seems graspable because we perceive manifold-like objects in our everyday life (like balls, donuts etc.). Equally, the abstract concept of a tangent vector feels familiar because we seem to have a feeling for what a "direction" in space is (like a road sign). Now this is of course a somewhat vague way of "intuitively understanding" abstract concepts, and it breaks down much earlier before concepts such as a spinor.[10] Nevertheless it would of course still be nice to have some "rooting of abstract concepts in the real world", at least as far as they appear in the formulation of physical theories.

Since spinors are nothing which we seem to directly perceive in classical physics (on first sight, but see below), let us turn to quantum physics. Here one can "answer" the question about the meaning of spinors by just taking notice that the fundamental objects of our reality (call them "particles" or something else), according to theory but more importantly according to "real-world experiments", seem to possess a basic property which one calls *spin*, and which we can describe in our theories precisely by using the abstract

[10] For instance, do you have a similar intuitive feeling about the concept of *spacetime* which does not just reduce to thinking about it as space? At least for the author spacetime is already quite difficult to really grasp intuitively.

mathematical spinors encountered previously. This property behaves like an additional degree of freedom besides the properties of "being somewhere in space(time)" and "moving somewhere in space(time)". The latter two giving the concept of manifold and tangent vectors a "real-world anchoring", one could simply say that this spin-property should (or at least could) do the same for the abstract concept of a spinor. Of course, this still seems rather unintuitive, but the same objection would simply apply to anything except for that which we already know. One possibility to make spinors and spin more graspable is to study precisely how it *does* manifest in the real world by studying real-world physical experiments involving spin, as they are described in many physics textbooks. Then, maybe some day the concept of a spinor will not seem so abstract and alien anymore.

After this excursion to physics, let us come back to our everyday world. Also here it is possible to get some insight into the "meaning" of the concept of a spinor by its relation to rotations. Namely, there exist numerous experiments one can find on the internet and perform oneself, such as for instance Dirac's famous "belt trick", which illustrate the following geometric property of a spinor (in three Euclidean dimensions): A spinor is an object, for which a rotation about 2π is *not* equivalent to making no rotation at all, but a rotation about 4π is. While this seems strange on first encounter, the fact that there really are actual experiments one can perform can make this "strange fact" more familiar. Mathematically it corresponds to the property that the group of rotations in three dimensions is not simply connected, but its universal (double) cover the spin group is. This lies at the core of the mathematical concept of a spinor as outlined previously. For some related insights into possible geometric meanings of spinors, there is a nice recent talk by Atiyah ([**Ati13**]).

1.2.2. Spin structures. Now we start with the construction of spinors on curved spacetimes or more generally on pseudo-Riemannian manifolds. The general construction uses the machinery of principal bundles and associated vector bundles. While some of these concepts will be recalled very briefly along the way, we refer readers unfamiliar with these notions to either the quick recapitulation in [**LM89**, App. G, Ch. II. 3] or to more detailed expositions such as [**Bau09**], [**Ble81**], or [**Hus94**].

Let (M, g) be an oriented and time-oriented pseudo-Riemannian manifold of signature (p, q).[11] To state it clearly, the goal is to construct a complex vector bundle $\mathbb{S}M \to M$ whose fiber is isomorphic to the spin space $\mathbb{S}_{p,q}$, and which has analogue structures as those of $\mathbb{S}_{p,q}$. Most importantly we want a Clifford multiplication of elements of $\mathbb{S}M$ by tangent vectors, which ties $\mathbb{S}_{p,q}$ to the metric g via the Clifford relations (1.1.49). In the following I try to give an intuitive description of the whole construction. The reader already familiar with this can jump straight ahead to Definition 1.2.1.

To see how our goal could be achieved, let us start locally. We fix some local oriented and time-oriented orthonormal frame $E_1, \ldots, E_n \in \Gamma^\infty(TM|_U)$ over an open subset $U \subset M$. Decomposing every tangent vector over U in this frame yields a local trivialization $TM|_U \cong U \times \mathbb{R}^{p,q}$. To achieve our goal locally over U, we simply set $\mathbb{S}M|_U := U \times \mathbb{S}_{p,q}$ and define Clifford multiplication of an element of $\mathbb{S}M|_U$ with a vector in $TM|_U$ (at the same base point of U) simply via the Clifford multiplication of $\mathbb{R}^{p,q}$ on $\mathbb{S}_{p,q}$ as defined in Section 1.1.7. For $v \in T_xM|_U$, we denote Clifford multiplication by $\gamma(v) \in \mathsf{End}(\mathbb{S}_xM|_U)$. Other structures like the inner product $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$ can be carried over to $\mathbb{S}M|_U$ as well.

[11] These orientation assumptions are not strictly needed in order to introduce spinors in the pseudo-Riemannian case, but we shall assume them nevertheless. See also Remark 1.2.8.

In the next step, unless we can actually choose a global orthonormal frame over all of M we need to find a way to suitably "patch together" such local constructions. To anyone with experience in differential geometry it might be clear at this point that the key question is what the *transition functions* between two overlapping local constructions should be. This can be figured out by looking at Clifford multiplication. To understand this, let us first suppose that $e_1, \ldots, e_n$ and $e'_1, \ldots, e'_n$ are two oriented and time-oriented orthonormal bases at one point $x \in M$. Then these are always related by a proper orthochronous Lorentz transformation $A \in \mathsf{SO}^+(p,q)$, i.e. we have $e_\mu = Ae_\mu$ for all $\mu = 1, \ldots, n$. Therefore, if we pick $s \in \mathsf{Spin}^+(p,q)$ with $\vartheta_{p,q}(s) = A$, then we know from Lemma 1.1.17 that

$$\gamma(e_\mu) = \gamma(Ae'_\mu) = \gamma(\vartheta_{p,q}(s)e'_\mu) \overset{(1.2.5)}{=} \rho_{p,q}(s)\gamma(e'_\mu)\rho_{p,q}(s)^{-1} \,. \tag{1.2.5}$$

But this means nothing else but that the change of the orthonormal basis in T_xM described by $A \in \mathsf{SO}^+(p,q)$ should be accompanied by a change of basis in the spin space $\mathbb{S}_xM|_U = \mathbb{S}_{p,q}$ described by $\rho_{p,q}(s) \in \mathsf{GL}(\mathbb{S}_{p,q})$.

Returning from orthonormal bases at a point to local orthonormal frames, suppose that $E_1, \ldots, E_n$ and $E'_1, \ldots, E'_n$ are two local oriented and time-oriented orthonormal frames over a common open set $U \subset M$. Then there exists a smooth map $A : U \to \mathsf{SO}^+(p,q)$ such that $E_\mu = AE'_\mu$ for all $\mu = 1, \ldots, n$. Following the idea from before, at each point $x \in U$ we need to chose an element $s_x \in \mathsf{Spin}^+(p,q)$ such that $\vartheta_{p,q}(s_x) = A_x$. Moreover, if we want to have a chance to construct a smooth bundle from this in the end, we need to make these choices smoothly in x. In other words, we need to find a smooth map $\widetilde{A} : U \to \mathsf{Spin}^+(p,q)$ such that the following diagram commutes:

$$\begin{tikzcd} & \mathsf{Spin}^+(p,q) \arrow[d, "\vartheta_{p,q}"] \\ U \arrow[ur, "s"] \arrow[r, "A"'] & \mathsf{SO}^+(p,q) \end{tikzcd} \tag{1.2.6}$$

This means that we need to lift the map A to the $2:1$ covering $\vartheta_{p,q} : \mathsf{Spin}^+(p,q) \to \mathsf{SO}^+(p,q)$. If U is simply connected this is always possible and, since the covering is $2:1$, there are precisely two possible choices.

Coming back to our goal of patching together the locally defined spinor bundles, the idea is as now follows: We cover M by open subset $\{U_\alpha\}_{\alpha\in\Lambda}$ such that each intersection $U_\alpha \cap U_\beta$ is simply connected. Moreover, we assume that on each U_α we can choose an oriented and time-oriented orthonormal frame, and we make one such choice for each U_α. Over each U_α we then construct the local spinor bundle $\mathbb{S}M|_{U_\alpha} := U_\alpha \times \mathbb{S}_{p,q}$. Next, for any pair U_α, U_β with nonempty intersection, the locally defined bundles $\mathbb{S}M|_{U_\alpha}$ and $\mathbb{S}M|_{U_\beta}$ can be patched together to a bundle $\mathbb{S}M|_{U_\alpha\cup U_\beta}$ by choosing a lift $s_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathsf{Spin}^+(p,q)$ (1.2.6) of the transition function $A_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathsf{SO}^+(p,q)$ between the chosen orthonormal frames on U_α and U_β, and using $s_{\alpha\beta}$ as transition function between $\mathbb{S}M|_{U_\alpha}$ and $\mathbb{S}M|_{U_\beta}$. The interesting part starts if we encounter intersections of more than two open sets, say $U_\alpha \cap U_\beta \cap U_\gamma \neq \emptyset$. Then we have to make *consistent* choices of transition function $s_{\alpha\beta}, s_{\beta\gamma}, s_{\alpha\gamma}$ in the sense that these have to satisfy the *cocycle condition*

$$s_{\alpha\beta}s_{\beta\gamma} = s_{\alpha\gamma} \,.$$

Otherwise it would not be possible to patch the three locally defined bundles together to a bundle on $U_\alpha \cup U_\beta \cup U_\gamma$. Now in order to really patch together *all* the locally defined bundles into one global bundle $\mathbb{S}M$, we need to choose lifts $s_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathsf{Spin}^+(p,q)$ for *any* choice of $\alpha, \beta \in \Lambda$ with $U_\alpha \cap U_\beta \neq \emptyset$ in such a way that the cocycle condition holds for *all* triples $\alpha, \beta, \gamma \in \Lambda$ such that $U_\alpha \cap U_\beta \cap U_\gamma \neq \emptyset$.

That we can indeed construct a global bundle if we manage to make such consistent choices follows from the description of locally trivial bundles in terms of transition functions (cf. [**Hus94**, Ch. 5]). Whether or not making such consistent choices is possible at all can be described in cohomological terms and depends on the topology of M, see Theorem 1.2.2 and Remark 1.2.8 below.

After this long explanation, we come to the usual definition of a spin structure, which condenses all that was just said. The only adjustment one has to make in order to make the connection from the previous discussion to the following definition is that instead of patching together the locally defined spinor bundles with the choices of lifts (1.2.6) (in case this is possible), one rather patches together the locally defined $\mathsf{Spin}^+(p,q)$-principal bundles $U_\alpha \times \mathsf{Spin}^+(p,q)$ to a global $\mathsf{Spin}^+(p,q)$-principal bundle $\mathsf{Spin}^+(M) \to M$ using the same transition functions (1.2.6). This bundle should be thought of as a subset of the frame bundle of $\mathbb{S}M$, similarly to how the bundle of oriented and time-oriented orthonormal frames is a subbundle of the frame bundle of TM. If constructed in the way outlined above, it will have the properties described in the following definition.

Definition 1.2.1. Let (M,g) be an oriented and time-oriented pseudo-Riemannian manifold of signature (p,q), and let $\mathsf{SO}^+(M) \to M$ be the bundle of oriented and time-oriented orthonormal frames. A *spin structure* on (M,g) consists of a $\mathsf{Spin}^+(p,q)$-principal bundle $\mathsf{Spin}^+(M) \to M$ and a smooth map $\Theta : \mathsf{Spin}^+(M) \to \mathsf{SO}^+(M)$ such that the following diagram commutes:

$$\begin{tikzcd}
\mathsf{Spin}^+(M) \times \mathsf{Spin}^+(p,q) \arrow[r, "\cdot"] \arrow[d, "\Theta \times \vartheta_{p,q}"] & \mathsf{Spin}^+(M) \arrow[d, "\Theta"] \arrow[dr] & \\
 & & M \\
\mathsf{SO}^+(M) \times \mathsf{SO}^+(p,q) \arrow[r, "\cdot"] & \mathsf{SO}^+(M) \arrow[ur] &
\end{tikzcd} \tag{1.2.7}$$

Here $\vartheta_{p,q} : \mathsf{Spin}^+(p,q) \to \mathsf{SO}^+(p,q)$ is the spin covering, and the horizontal arrows denote right multiplication in the principal bundles.

An oriented and time-oriented pseudo-Riemannian manifold together with the choice of a spin structure is called a *spin manifold*.

As should be clear from the discussion before, a general oriented and time-oriented pseudo-Riemannian manifold (M,g) might not possess a spin structure at all (similarly to how a general manifold might not possess an orientation). Furthermore, in the case that (M,g) does possess a spin structure, there may exist several inequivalent ones.[12] Both the matter of existence and the matter of uniqueness of spin structures can be reformulated in cohomological terms, and the following holds:

Theorem 1.2.2. (Existence and uniqueness of spin structures) *Let (M,g) be an oriented and time-oriented pseudo-Riemannian manifold. Then the following hold:*

(1) (M,g) possesses a spin structure if and only if its second Stiefel-Whitney class vanishes, $w_2(TM) = 0 \in H^2(M;\mathbb{Z}_2)$.

[12] Here two spin structures $(\mathsf{Spin}^+(M), \Theta)$ and $(\mathsf{Spin}^+(M)', \Theta')$ are called equivalent if $\mathsf{Spin}^+(M)$ and $\mathsf{Spin}^+(M)'$ are equivalent as $\mathsf{Spin}^+(p,q)$-principal bundles in way compatible with maps Θ and Θ'.

(2) *If (M, g) possesses a spin structure, then the set of (inequivalent) spin structures of (M, g) is in bijection to $H^1(M; \mathbb{Z}_2)$.*

In particular, if (M, g) is simply connected then it admits at most one spin structure.

For a proof, as well as an explanation of the terms in this result, we refer to [**Bau81**, Ch. 2], the strictly Riemannian case can also be found in many other books such as [**LM89**] and [**Fri00**]. Here we only make the further remark that the vanishing of the first Stiefel-Whitney class $w_1(TM) \in H^1(M; \mathbb{Z}_2)$ is precisely the criterion for a manifold to be orientable. Hence existence of a spin structure may be viewed as a "higher orientability criterion".

Let us illustrate existence and uniqueness of spin structures by some examples.

Example 1.2.3. *(*Existence of spin structures on parallelizable manifolds) A manifold M is said to be *parallelizable* if there exists a smooth global frame $E_1, \dots, E_n$ of TM (for instance, any Lie group is parallelizable). In this case M is clearly orientable. If M additionally carries a metric g, then by the usual Gram-Schmidt procedure one always finds a smooth, global, orthonormal frame. Picking a global oriented and time-oriented frame $E = (E_1, \dots, E_n)$ one obtains a trivialization $\mathsf{SO}^+(M) \cong M \times \mathsf{SO}^+(p, q)$. Then one can simply define a spin structure on (M, g) by setting $\mathsf{Spin}^+(M) := M \times \mathsf{Spin}^+(p, q)$ and defining $\Theta : \mathsf{Spin}^+(M) \to \mathsf{SO}^+(M)$ as $\Theta = \mathsf{id}_M \times \vartheta_{p,q}$. Nevertheless, there might still exist other inequivalent spin structures, see Example 1.2.5.

Example 1.2.4. *(*Orientable 3-manifolds and globally hyperbolic spacetimes) Due to *Steenrod's theorem*, an orientable 3-manifold is necessarily also parallelizable. As such, any orientable 3-manifold possesses spin structures. Consequently, so does any 4-manifold of the form $M = \mathbb{R} \times N$ with N being an orientable 3-manifold. This includes all globally hyperbolic spacetimes in $1 + 3$ dimensions with orientable Cauchy surfaces.

Example 1.2.5. *(*Inequivalent spin structures on S^1) Take $S^1 = \mathbb{R}_\theta/2\pi\mathbb{Z}$ with its usual round metric $\mathrm{d}\theta^2$ and usual orientation. Then M is parallelizable by the vector field ∂_θ, and this choice induces an isomorphism $\mathsf{SO}(S^1) \cong S^1 \times \mathsf{SO}(1) = S^1$. Hence, as explained before, one possible spin structure on S^1 is given by $\mathsf{Spin}(S^1) = S^1 \times \mathsf{Spin}(1) = S^1 \times \mathbb{Z}^2$. This spin structure is called *trivial spin structure* of S^1.

On the other hand, set $\mathsf{Spin}(S^1)' = (\mathbb{R}_\theta \times \mathsf{Spin}(1))/\sim$, where the equivalence relation is generated by $(\theta, \pm 1) \sim (\theta + 2\pi, \mp 1)$. This is a $\mathsf{Spin}(1)$-principal bundle in the obvious way, and $\theta'([\theta, \pm 1]) = \partial_\theta$ makes it a spin structure, called *nontrivial spin structure* of S^1.

Notice that clearly these two spin structures cannot be equivalent since $\mathsf{Spin}(S^1)$ and $\mathsf{Spin}(S^1)'$ are not even homeomorphic. Since $H^1(S^1; \mathbb{Z}_2) \cong \mathbb{Z}_2$, these are the only two spin structures on S^1.

Example 1.2.6. *(*Inequivalent spin structures on T^n) Following up the previous example, one easily sees that the n dimensional torus $T^n = S^1 \times \cdots \times S^1$ admits 2^n inequivalent spins structures.

Example 1.2.7. *(*Products of spin manifolds) Suppose that (M, g) is an oriented and time-oriented pseudo-Riemannian manifold, and suppose that $M = Q \times N$ (the metric g need not be a product). Then we have $TM \cong \widetilde{TQ} \oplus \widetilde{TN}$ where $\widetilde{TQ}$ is the pullback of TQ under the projection $Q \times N \to Q$ and similar for N. Suppose that both Q and N are orientable and "spinnable", i.e $w_1(Q) = w_2(Q) = 0$ and also for N. Then it follows from additivity of Stiefel-Whitney classes (cf. [**MS74**, Ch. 4]) that

$$w_2(TM) = w_2(TQ) \cup w_0(TN) + w_1(TQ) \cup w_1(TN) + w_0(TQ) \cup w_2(TN) = 0\,,$$

hence (M, g) admits a spin structure.

More examples can be found for instance in [**Bau81**, Ch. 2].

Remark 1.2.8. *(*Spin structures in the non-orientable cases) If (M, g) is a pseudo - Riemannian manifold, then the metric g induces an orthogonal splitting $TM = Z \oplus R$ where $g|_Z$ is negative definite and $g|_R$ is positive definite. One calls (M, g) *time-orientable* if the vector bundle Z is orientable, and *space-orientable* if the vector bundle R is orientable. As remarked above, these two conditions can be reformulated as $w_1(Z) = 0$ and $w_1(R) = 0$, respectively. In particular it follows from the additivity of Stiefel-Whitney classes (cf. [**MS74**, Ch. 4]), i.e.

$$w_1(TM) = w_1(Z) \cup w_0(R) + w_0(Z) \cup w_1(R)\,,$$

that time- and space-orientability together imply orientability. Depending on whether Z, R, and TM are orientable or not, one can reduce the structure group of the orthonormal frame bundle $\mathsf{O}(M)$ from $\mathsf{O}(p, q)$ to one of the subgroups $\mathsf{O}^+(p, q)$, $\mathsf{O}_+(p, q)$, $\mathsf{SO}(p, q)$, or $\mathsf{SO}^+(p, q)$.

Concerning spin structures, one can generalize Definition 1.2.1 to each degree of orientation in the obvious way by replacing $\mathsf{SO}^+(p, q)$ and $\mathsf{Spin}^+(p, q)$ by the respective other subgroups of $\mathsf{O}(p, q)$ and $\mathsf{Pin}(p, q)$. Then the condition for the existence of a spin structure is that

$$w_2(TM) = w_1(Z) \cup w_1(R) \qquad \text{in } H^2(M; \mathbb{Z}_2)\,, \tag{1.2.8}$$

see [**Bau81**, Satz 2.2]. If (M, g) is time-orientable (or space-orientable), this reduces to the criterion $w_2(TM) = 0$ of Theorem 1.2.2 *(1)*. If (M, g) is Riemannian, then $Z = 0$ is zero-dimensional, so also here the condition reduces to $w_2(TM) = 0$. What is interesting in the indefinite case is that the splitting $TM = Z \oplus R$ depends on the metric g, so that the existence of a spin structure also depends on the metric. In the Riemannian situation, on the other hand, the existence of spin structures is a purely topological condition.

1.2.3. The spinor bundle. Given a spin structure, one can define the spinor bundle as an associated vector bundle to it using the spin representation. To fix notation, recall that for a G-principal bundle $P \to M$ and a representation $\rho : G \to \mathsf{GL}(V)$, the *associated vector bundle* is defined as

$$P \times_\rho V := (P \times V)/\sim\,,$$

where the equivalence relation $\sim$ is generated by

$$(p, v) \sim (p \cdot g, \rho(g)v) \quad \forall g \in G \qquad \forall (p, v) \in P \times V\,.$$

Here $p \cdot g$ denotes the right-action of G on P. The projection $P \to M$ and the vector space structure of V turn $P \times_\rho V$ into a vector bundle in the obvious way. As usual for quotients, the elements of $P \times_\rho V$ are denoted by $[p, v]$ with $p \in P$ and $v \in V$. More details on this construction, and on (principal) bundles in general, can be found in [**Bau09**], [**Ble81**], or [**Hus94**].

Definition 1.2.9. Let (M, g) be an oriented and time-oriented pseudo-Riemannian manifold of signature (p, q), and suppose that $(\mathsf{Spin}^+(M), \Theta)$ is a spin structure. Then the corresponding *(classical) spinor bundle* is the associated vector bundle

$$\mathbb{S}M := \mathsf{Spin}^+(M) \times_{\rho_{p,q}} \mathbb{S}_{p,q} \longrightarrow M\,, \tag{1.2.9}$$

where $\rho_{p,q} : \mathsf{Spin}^+(p, q) \to \mathsf{GL}(\mathbb{S}_{p,q})$ is the complex spin representation.

Before we introduce additional structures on the spinor bundle, let us consider some examples first.

Example 1.2.10. *(*Spinor bundles on parallelizable manifolds) Suppose that (M, g) is parallelizable and $E_1, \ldots, E_n \in \Gamma^\infty(TM)$ is a global oriented and time-oriented orthonormal frame. Let $\mathsf{SO}^+(M) \cong M \times \mathsf{SO}^+(p,q)$ be the induced isomorphism, and let $\mathsf{Spin}^+(M) = M \times \mathsf{Spin}^+(p,q)$ be the corresponding trivial spin structure (cf. Example 1.2.3). Then also the corresponding spinor bundle is trivial, $\mathbb{S}M \cong M \times \mathbb{S}_{p,q}$, and sections of $\mathbb{S}M$ are simply functions $\psi : M \to \mathbb{S}_{p,q}$.

Let us point out once more that every $1+3$ dimensional oriented globally hyperbolic spacetime (M, g) is of this form. However, let us also stress that even in this case there need not be a unique spinor bundle since there need not be a unique spin structure. In particular, the identification of sections of the spinor bundle with $\mathbb{S}_{p,q}$-valued smooth functions on M will in general only be true for the trivial spin structure described above.[13]

That different spin structures can indeed lead to different spinor bundles can once more be illustrated by the example of S^1.

Example 1.2.11. *(*Spinor bundles on S^1) Consider $S^1 = \mathbb{R}_\theta/2\pi\mathbb{Z}$ with the usual metric $\mathrm{d}\theta^2$. First let $\mathsf{Spin}(S^1) = S^1 \times \mathbb{Z}^2$ be the trivial spin structure on S^1. Since $\mathsf{Spin}(S^1)$ is a trivial bundle, also $\mathbb{S}S^1 = S^1 \times \mathbb{S}_1 = S^1 \times \mathbb{C}$ is a trivial vector bundle. Sections of this bundle are simply (complex-valued) functions on S^1, which may be identified with 2π-periodic functions on $\mathbb{R}$.

On the other hand, let $\mathsf{Spin}(S^1)'$ be the nontrivial spin structure as introduced in Example 1.2.5. Then $\mathbb{S}'S^1 \cong (\mathbb{R}_\theta \times \mathbb{C})/\sim$, where $(\theta, z) \cong (\theta + 2\pi, -z)$. This is something like a (complex) Möius strip, and sections of $\mathbb{S}'S^1$ can be identified with *anti-periodic* functions ψ on $\mathbb{R}$, i.e. satisfying $\psi(\theta + 2\pi) = -\psi(\theta)$.

A general principle for the construction of additional objects on associated vector bundles is the following: Let $E = P \times_\rho V$ be an associated vector bundle to some G-principal bundle P and a representation $\rho : G \to \mathsf{GL}(V)$. Then any "structure" on V which is "equivariant" in a suitable way for the G-action induces a corresponding structure on E. Concerning spinor bundles, two important structures that carry over from the linear algebraic regime to the nonlinear "curved" setting in this way are *Clifford multiplication* and the *inner product on spin space*.

In the following, (M, g) is an oriented and time-oriented pseudo-Riemannian manifold with a fixed spin structure $\mathsf{Spin}^+(M)$ and associated spinor bundle $\mathbb{S}M$.

First, for $v \in T_xM$ we want to define Clifford multiplication by v as a linear map $\gamma(v) \in \mathsf{End}(\mathbb{S}_xM)$. To this end we choose an oriented and time-oriented orthonormal basis $(e_1, \ldots, e_n) \in \mathsf{SO}^+_x(M)$, and we choose a "spin frame" $s \in \mathsf{Spin}^+_x(M)$ with $\Theta(s) = (e_1, \ldots, e_n)$. Then every element of $\mathbb{S}_xM$ has the form $[s, z]$ for some $z \in \mathbb{S}_{p,q}$, and writing $v = v^\mu e_\mu$ we define

$$\gamma(v)[s, \psi] := [s, \gamma_{\mathbb{S}_{p,q}}(v^1, \ldots, v^n)\psi] \qquad \forall \psi \in \mathbb{S}_{p,q}\,, \tag{1.2.10}$$

where $\gamma_{\mathbb{S}_{p,q}}$ is Clifford multiplication of $\mathbb{R}^{p,q}$ on $\mathbb{S}_{p,q}$ as defined in Section 1.1.7. Due to the equivariance property (1.2.5) of $\gamma_{\mathbb{S}_{p,q}}$ the above is well-defined, i.e. independent of the choices made.

In this way we obtain we obtain a linear map $\gamma : TM \to \mathsf{End}(\mathbb{S}M)$, which we again call *Clifford multiplication*. It is pretty much clear by definition that γ is actually a smooth vector bundle map. Moreover, it follows from (1.1.49) that it again satisfies the

[13] If (M, g) is additionally assumed to be simply connected, then the trivial spin structure is the unique spin structure, so that one does not have to worry about these ambiguities.

Clifford relations

$$\{\gamma(X), \gamma(Y)\} = -2g(X,Y)\mathbb{1}_{\mathbb{S}M} \qquad \forall X, Y \in \Gamma^\infty(TM)\,. \tag{1.2.11}$$

Secondly, we want to define a hermitian inner product on the fibers of $\mathbb{S}M$. To this end, let $x \in M$ and choose a "spin frame" $s \in \mathsf{Spin}^+_x(M)$ at x. Then any vector in $\mathbb{S}_xM$ can again be written as $[s,z]$ for some $z \in \mathbb{S}_{p,q}$, and we define

$$\prec [s,\psi], [s,\phi] \succ_{\mathbb{S}_xM} := \prec \psi, \phi \succ_{\mathbb{S}_{p,q}} \qquad \forall [s,z],[s,w] \in \mathbb{S}_xM\,. \tag{1.2.12}$$

Since $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$ is $\mathsf{Spin}^+(p,q)$-invariant, it follows that this is well-defined, i.e. independent of the choices made. It defines a hermitian inner product on $\mathbb{S}_xM$ which is always non-degenerate and has the same signature as $\prec\cdot,\cdot\succ_{\mathbb{S}_{p,q}}$. So it is positive definite in case that g is positive or negative definite, and it is indefinite (and has split signature) in case that g is indefinite. Moreover, it is also not difficult to see that (1.2.12) actually defines a smooth hermitian inner product on $\mathbb{S}M$. Notice also that every $s \in \mathsf{Spin}^+_x(M)$ can now be viewed as orthonormal basis for $\mathbb{S}_xM$ with respect to $\prec\cdot,\cdot\succ_{\mathbb{S}M}$ if we identify s with $[s,z_1],\ldots,[s,z_N]$ for some (once and for all fixed) orthonormal basis $z_1,\ldots,z_N \in \mathbb{S}_{p,q}$.

If (p,q) is the signature of the pseudo-Riemannian metric g, then it follows immediately from Proposition 1.1.20 that Clifford multiplication is skew-symmetric with respect to $\prec\cdot,\cdot\succ_{\mathbb{S}M}$ if p is even, and symmetric if p is odd. In particular, in the Riemannian case Clifford multiplication is symmetric, whereas in the Lorentzian case it is skew-symmetric.

In the Lorentzian case, any future-pointing timelike vector field $Z \in \Gamma^\infty(TM)$ allows to "positivize" $\prec\cdot,\cdot\succ_{\mathbb{S}M}$ by setting

$$\langle\cdot,\cdot\rangle_Z := \prec \cdot, \gamma(Z)\cdot \succ_{\mathbb{S}M}\ . \tag{1.2.13}$$

Since $\gamma(Z)$ acts symmetrically with respect to $\prec\cdot,\cdot\succ_{\mathbb{S}M}$ this is indeed again Hermitian. That it is positive definite follows from Lemma 1.1.23.

1.2.4. The spin connection. Having defined the important algebraic structures on the spinor bundle, we now come to the analytic structures. Here the first observation is that the Levi-Civita connection of g always lifts to a connection on $\mathbb{S}M$. Combining this connection with Clifford multiplication then leads to the Dirac operator as explained in the next section.

Explaining how the Levi-Civita connection lefts to a connection on the spinor bundle is done most conveniently by viewing the Levi-Civita connection as a connection on the frame bundle $\mathsf{SO}^+(M)$. Therefore we briefly recall this concept. For background about connections on principal bundles, we refer to [**LM89**, II. § 4] or [**Bau09**, Ch. 3].

Let $P \to M$ be a G-principal bundle and let $\mathfrak{g}$ be the Lie algebra of G. A *connection* on P is a $\mathfrak{g}$-valued 1-form $A \in \Omega^1(P;\mathfrak{g})$ with the two properties

i.) $A(X_\xi) = \xi$ for all $\xi \in \mathfrak{g}$,
ii.) $R_g^* A = \mathrm{Ad}_{g^{-1}} \circ A$ for all $g \in G$.

Here for $\xi \in \mathfrak{g}$ we denote by $X_\xi \in \Gamma^\infty(TP)$ the so-called *fundamental vector field*

$$X_\xi|_p := \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} p \cdot \exp(t\xi) \qquad \forall p \in P\,,$$

and for $g \in G$ we denote by R_g and Ad_g its right-action on P and adjoint action on $\mathfrak{g}$, respectively.

Next, for any associated vector bundle $E := P \times_\rho V$, a connection $A \in \Omega^1(TP;\mathfrak{g})$ induces a covariant derivative ∇^A on E. Locally, ∇^A is described as follows: If $s : U \to P$

is a smooth local section, then any $\psi \in \Gamma^\infty(E)$ is locally of the form $\psi|_U = [s, v]$ for some smooth function $v : U \to V$. Then, for any $X \in TM|_U$ one has

$$\nabla^A_X \psi = \left[s, \mathrm{d}v(X) + \rho_*(A(\mathrm{d}s(X)))v\right], \tag{1.2.14}$$

where $\rho_* : \mathfrak{g} \to \mathsf{End}(V)$ is the Lie algebra map induced by the homomorphism $\rho : G \to \mathsf{GL}(V)$. Notice that in (1.2.14) only the so-called *local connection 1-form* $s^*A = A \circ \mathrm{d}s \in \Omega^1(U; \mathfrak{g})$ enters, which is a 1-form on $U \subset M$ instead of P.

Returning to our actual situation of interest, the Levi-Civita connection ∇^{LC} on TM determines a connection $A^{\mathrm{LC}} \in \Omega^1(\mathsf{SO}^+(M); \mathfrak{so}^+(p, q))$. For an explicit description of how A^{LC} is determined from ∇^{LC} we refer to [**Bau09**, Bsp. 3.5], but see also (1.2.16) below. From A^{LC} one now obtains a connection $A^{\mathbb{S}} \in \Omega^1(\mathsf{Spin}^+(M); \mathfrak{spin}^+(p, q))$ as follows: Since the spin covering $\vartheta_{p,q} : \mathsf{Spin}^+(p, q) \to \mathsf{SO}^+(p, q)$ is a *covering*, the induced map $(\vartheta_{p,q})_* : \mathfrak{spin}^+(p, q) \to \mathfrak{so}^+(p, q)$ is an *isomorphism* of Lie algebras so that it can be inverted. Therefore, we can simply define

$$A^{\mathbb{S}} : T\,\mathsf{Spin}^+(M) \xrightarrow{T\Theta} T\,\mathsf{SO}^+(M) \xrightarrow{A^{\mathrm{LC}}} \mathfrak{so}^+(p, q) \xrightarrow{(\vartheta_{p,q})_*^{-1}} \mathfrak{spin}^+(p, q)\,. \tag{1.2.15}$$

Clearly this defines a 1-form on $\mathsf{Spin}^+(M)$, and the properties *i.)* and *ii.)* of a connection 1-form listed above follow easily from the equivariance properties of Θ (cf. Def. 1.2.1).

Definition 1.2.12. The connection $A^{\mathbb{S}} \in \Omega^1(\mathsf{Spin}^+(M); \mathfrak{spin}^+(p, q))$ defined by (1.2.15), and also the induced covariant derivative $\nabla^{\mathbb{S}}$ on $\mathbb{S}M$ are called *spin connection*.

On the practical side, it is of course of interest how one may explicitly compute (locally) the spin connection $\nabla^{\mathbb{S}}$ in terms of the Levi-Civita connection. To this end, we first recall a local formula relating ∇^{LC} and A^{LC}: Let $E = (E_1, \ldots, E_n) : U \to \mathsf{SO}^+(M)$ be a smooth oriented and time-oriented orthonormal local frame over some open set $U \subset M$. Then the local connection 1-form $E^*A^{\mathrm{LC}} = A^{\mathrm{LC}} \circ \mathrm{d}E \in \Omega^1(U; \mathfrak{so}^+(p, q))$, which conversely is all that is needed to express ∇^{LC} in terms of A^{LC} locally (cf. (1.2.14)), is explicitly given by

$$(E^*A^{\mathrm{LC}})(X) = \frac{1}{2} \sum_{\mu,\nu=1}^{n} \epsilon_\mu \epsilon_\nu \left\langle \nabla^{\mathrm{LC}}_X E_\mu, E_\nu \right\rangle E_{\mu\nu}\,, \tag{1.2.16}$$

where $\{E_{\mu\nu}\}_{1 \leq \mu < \mu \leq n}$ is the basis of $\mathfrak{so}^+(p, q)$ which already appeared in (1.1.36), and where $\epsilon_\mu = \langle e_\mu, e_\mu \rangle$ as usual. As for the proof of (1.2.16), this formula is usually how one actually defines the connection $A^{\mathrm{LC}} \in \Omega^1(\mathsf{SO}^+(M); \mathfrak{so}^+(p, q))$ in terms of ∇^{LC} in the first place, see [**Bau09**, Bsp. 3.5]. Using this formula, we can now give a convenient local description of $\nabla^{\mathbb{S}}$.

Lemma 1.2.13. (Local formula for spin connection) *Let $s : U \to \mathsf{Spin}^+(M)$ be a smooth local section, and let $E = (E_1, \ldots, E_n) := \Theta(s) : U \to \mathsf{SO}^+(M)$ be the corresponding smooth oriented and time-oriented local orthonormal frame. Let $z_1, \ldots, z_N \in \mathbb{S}_{p,q}$ be any basis and let $\zeta_1 := [s, z_1], \ldots, \zeta_N := [s, z_N] : U \to \mathbb{S}M$ be the corresponding local spinor frame. Then for any $\psi \in \Gamma^\infty(\mathbb{S}M)$ and $X \in \Gamma^\infty(TM)$ we have*

$$\boxed{\nabla^{\mathbb{S}}_X \psi|_U = \mathrm{d}\psi^\alpha(X)\zeta_\alpha + \frac{1}{4} \sum_{\mu,\nu=1}^{n} \epsilon_\mu \epsilon_\nu \left\langle \nabla^{\mathrm{LC}}_X E_\mu, E_\nu \right\rangle \gamma(E_\mu)\gamma(E_\nu)\psi}\,, \tag{1.2.17}$$

where $\psi|_U = \psi^\alpha \zeta_\alpha$ is the local decomposition of ψ in the frame $\zeta_1, \ldots, \zeta_N$.

PROOF. First one expresses $\nabla^{\mathbb{S}}$ locally by formula (1.2.14). Then one uses the definition (1.2.15) of $A^{\mathbb{S}}$ in terms of A^{LC} and $(\vartheta_{p,q})_*^{-1}$ and the explicit local expression (1.2.16) for A^{LC}. Finally, one uses that by (1.1.37) we have $(\vartheta_{p,q})_*(E_{\mu\nu}) = \frac{1}{2}\gamma_\mu\gamma_\nu$. From this (1.2.17) follows. □

Using this local formula for the spin connection, it is not difficult to verify explicitly that $\nabla^{\mathbb{S}}$ is actually metric for the hermitian inner product $\prec\cdot,\cdot\succ_{\mathbb{S}M}$, and that Clifford multiplication γ is parallel in the sense explained below. To be precise, being metric means that for any $\psi, \phi \in \Gamma^\infty(\mathbb{S}M)$ and any $X \in \Gamma^\infty(TM)$ we have

$$X \prec\psi,\phi\succ_{\mathbb{S}M} = \prec\nabla^{\mathbb{S}}_X\psi,\phi\succ_{\mathbb{S}M} + \prec\psi,\nabla^{\mathbb{S}}_X\phi\succ_{\mathbb{S}M} \qquad \forall\psi,\phi\in\Gamma^\infty(\mathbb{S}M)\,, \tag{1.2.18}$$

and saying that Clifford multiplication is parallel means that for any $X, Y \in \Gamma^\infty(TM)$ and any $\psi \in \Gamma^\infty(\mathbb{S}M)$ the Leibniz rule

$$\nabla^{\mathbb{S}}_X(\gamma(Y)\psi) = \gamma(\nabla^{\mathrm{LC}}_X Y)\psi + \gamma(Y)\nabla^{\mathbb{S}}_X\psi \tag{1.2.19}$$

is satisfied.

Remark 1.2.14. As for any connection, also the spin connection has its associated curvature. On the level of the spinor bundle, the curvature tensor $R^{\mathbb{S}} \in \Omega^2(M;\mathsf{End}(\mathbb{S}M))$ can be related to the curvature tensor of the Levi-Civita connection by the local formula

$$R^{\mathbb{S}}(X,Y) = \sum_{\mu,\nu=1}^{n} \epsilon_\mu\epsilon_\nu \frac{1}{4}\left\langle R^{\mathrm{LC}}(X,Y)E_\mu, E_\nu\right\rangle \gamma(E_\mu)\gamma(E_\mu)\,, \tag{1.2.20}$$

which holds in any (local) orthonormal frame $\{E_\mu\}$, see for instance [**LM89**, Thm. 4.15] for the Riemannian case (the other signatures work exactly analogously). In abstract index notation this can be written more compactly as $R^{\mathbb{S}}_{\mu\nu} = \frac{1}{4}R^{\mathrm{LC}}_{\mu\nu\alpha\beta}\gamma^\alpha\gamma^\beta$.

1.2.5. The Dirac operator. Finally we are in the position to define the famous Dirac operator.

Definition 1.2.15. (Dirac operator) Let (M,g) be an oriented and time-oriented pseudo-Riemannian spin manifold of signature (p,q) with spinor bundle $\mathbb{S}M$. The *Dirac operator* is the first-order differential operator $\mathcal{D}$ acting on sections $\psi \in \Gamma^\infty(\mathbb{S}M)$ by[14]

$$\boxed{\mathcal{D}\psi := \begin{cases} \mathrm{i}\,\mathsf{tr}_g(\gamma\otimes\nabla^{\mathbb{S}}\psi) = \mathrm{i}\gamma^\mu\nabla^{\mathbb{S}}_\mu\psi & \text{if } p \text{ is odd} \\ \mathsf{tr}_g(\gamma\otimes\nabla^{\mathbb{S}}\psi) = \gamma^\mu\nabla^{\mathbb{S}}_\mu\psi & \text{if } p \text{ is even}\end{cases}}\,. \tag{1.2.21}$$

The implementation of the imaginary factor for p odd is a useful convention to avoid the appearance of signs in integration by parts arguments since it balances the fact that Clifford multiplication is symmetric for p odd. For p even Clifford multiplication is *skew*-symmetric so that the imaginary factor is not needed to cancel negative signs.

Let us illustrate this definition by a few examples. In the first three examples the underlying manifold is just $\mathbb{R}^n$, so that there exists a unique spin structure.

Example 1.2.16. For $M = \mathbb{R}$ with the Euclidean metric we have $\mathbb{S}\mathbb{R} \cong \mathbb{R}\times\mathbb{C}$. Since $e_1 = \partial_x$ provides a global orthonormal frame and Clifford multiplication by e_1 is given by multiplication with i (cf. Example 1.1.2), the Dirac operator is given by $\mathcal{D}_{\mathbb{R}} = \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}$. Up to a sign this is the "momentum operator".

[14] Here $\gamma^\mu\nabla_\mu$ is to be understood in abstract index notation.

Example 1.2.17. For $M = \mathbb{R}^2$ with Euclidean metric, the vector fields ∂_x, ∂_y are a global orthonormal frame. The spinor bundle is of course the trivial bundle $\mathbb{S}M = \mathbb{R}^2 \times \mathbb{C}^2$ again, and it follows from Example 1.1.3 that the Dirac operator is given by

$$\mathcal{D}_{\mathbb{R}^2} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \partial_x + \begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \partial_y = \begin{pmatrix} 0 & \partial_x + \mathrm{i}\partial_y \\ -\partial_x + \mathrm{i}\partial_y & 0 \end{pmatrix} = \begin{pmatrix} 0 & \partial_z \\ -\partial_{\overline{z}} & 0 \end{pmatrix} .$$

This shows that solutions $\psi : \mathbb{R}^2 \to \mathbb{C}^2$ of $\mathcal{D}_{\mathbb{R}^2}\psi = 0$ are precisely the pairs $\psi = (f, g)$ where $f : \mathbb{C} \to \mathbb{C}$ is holomorphic and $g : \mathbb{C} \to \mathbb{C}$ is anti-holomorphic.

Example 1.2.18. For $M = \mathbb{R}^{1,2}$ with Minkowski metric, the vector fields ∂_t, ∂_x are a global orthonormal frame. Using the choice of gamma matrices (1.1.41), it follows that the Dirac operator on $\mathbb{S}\mathbb{R}^{1,1} = \mathbb{R}^{1,1} \times \mathbb{C}^2$ is explicitly given by

$$\mathcal{D}_{\mathbb{R}^{1,1}} = \mathrm{i} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \partial_t + \mathrm{i} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \partial_x = -\mathrm{i} \begin{pmatrix} 0 & \partial_t - \partial_x \\ \partial_t + \partial_x & 0 \end{pmatrix} . \tag{1.2.22}$$

Notice that for a section $\psi = (f, g)$, to satisfy $\mathcal{D}_{\mathbb{R}^{1,1}}\psi = 0$ means to satisfy the two (decoupled) transport equations

$$\partial_t f + \partial_x f = 0 , \qquad \partial_t g - \partial_x g = 0 ,$$

which simply say that f moves to the left and g moves to the right. If one instead considers the *massive* equation $\mathcal{D}_{\mathbb{R}^{1,1}}\psi = m\psi$, then one obtains a system of two *coupled* transport equations

$$\partial_t f + \partial_x f = \mathrm{i}mg , \qquad \partial_t g - \partial_x g = \mathrm{i}mf .$$

Example 1.2.19. For $M = \mathbb{R}^{1,3}$ with Minkowski metric, the vector fields $\partial_t, \partial_x, \partial_y, \partial_z$ are a global orthonormal frame. Hence $\mathbb{S}\mathbb{R}^{1,3} = \mathbb{R}^{1,3} \times \mathbb{C}^4$, and by Example 1.1.7 the Dirac operator (in the Dirac representation) is explicitly given by

$$\mathcal{D}_{\mathbb{R}^{1,3}} = \mathrm{i} \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & \mathbb{1}_2 \end{pmatrix} \partial_t + \mathrm{i} \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix} \partial_x + \mathrm{i} \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} \partial_y + \mathrm{i} \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix} \partial_z . \tag{1.2.23}$$

This is precisely the operator which Dirac introduced in [**Dir28**].

As last example, we again consider S^1 with its two inequivalent spin structures (cf. Example 1.2.5). This example shows explicitly that the choice of different spin structures can in fact lead to different analytic properties of the respective Dirac operators.

Example 1.2.20. Consider $S^1 = \mathbb{R}_\theta/2\pi\mathbb{Z}$ with its standard metric $\mathrm{d}\theta^2$, and trivialize $\mathsf{SO}(S^1) = S^1$ using the vector field ∂_θ.

First, we take the trivial spin structure $\mathsf{Spin}(S^1) = S^1 \times \mathbb{Z}_2$. Then we have already seen in Example 1.2.11 that $\mathbb{S}S^1 = S^1 \times \mathbb{C}$, so that sections of $\mathbb{S}S^1$ can be identified with 2π-periodic functions $\psi : \mathbb{R}_\theta \to \mathbb{C}$. It follows from the computation of Cl(1) in Example 1.1.2 that the Dirac operator is given by $\mathcal{D}_{S^1} = \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}\theta}$. For $k \in \mathbb{C}$, set $\psi_k(\theta) = \mathrm{e}^{-\mathrm{i}k\theta}$. Then ψ_k is 2π-periodic and $\mathcal{D}_{S^1}\psi_k = k\psi_k$ for all $k \in \mathbb{Z}$. Since $\{\psi_k\}_{k\in\mathbb{Z}}$ is an L^2-orthonormal basis on S^1, it follows that $\sigma(\mathcal{D}_{S^1}) = \mathbb{Z}$.

Let us now take the nontrivial spin structure $\mathsf{Spin}(S^1)'$ described in Example 1.2.5. As seen in Example 1.2.11, sections of the corresponding spinor bundle $\mathbb{S}'S^1$ can be identified with 2π-antiperiodic functions $\psi : \mathbb{R}_\theta \to \mathbb{C}$. As operator on these functions, the Dirac operator is again given by $\mathcal{D}'_{S^1} = \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}\theta}$. For any $k \in \mathbb{Z}$ set $\phi_k(\theta) = \mathrm{e}^{-\mathrm{i}(k+1/2)\theta}$. Then ϕ_k is a 2π-antiperiodic function on $\mathbb{R}_\theta$, and $\mathcal{D}'_{S^1}\phi_k = (k + \frac{1}{2})\phi_k$ for all $k \in \mathbb{Z}$. Further, the set $\{\phi_k\}_{k\in\mathbb{Z}}$ is an L^2-orthonormal basis of the closure in the $L^2(0, 2\pi)$-norm of the smooth antiperiodic functions on $\mathbb{R}$, which coincides with the L^2-space of sections of $\mathbb{S}'S^1$.[15] Therefore it follows that $\sigma(\mathcal{D}'_{S^1}) = \mathbb{Z} + \frac{1}{2}$.

Note that the two spectra $\sigma(\mathcal{D}_{S^1}) = \mathbb{Z}$ and $\sigma(\mathcal{D}'_{S^1}) = \mathbb{Z} + \frac{1}{2}$ differ. In particular, $\mathcal{D}_{S^1}$ has a nontrivial kernel, whereas $\mathcal{D}'_{S^1}$ has trivial kernel.

To conclude this first chapter, we come back to Dirac's original idea of taking the square-root of the Klein-Gordon equation. Namely, as the following famous and often used formula shows, the property which led to Dirac's discovery of his famous operator is still satisfied by the Dirac operator of a semi-Riemannian manifold.

Theorem 1.2.21. (Schrödinger–Lichnerowicz formula) *The Dirac operator of a semi-Riemannian spin manifold (M, g) of signature (p, q) satisfies the identity*

$$\mathcal{D}^2 = (-1)^{p+1}\left(\nabla^\mu\nabla_\mu + \frac{\mathrm{scal}}{4}\right), \tag{1.2.24}$$

where scal *is the scalar curvature of g.*

PROOF. First of all, the factor $(-1)^{p+1}$ is due to the fact that the definition of the Dirac operator contains an additional imaginary factor if p is odd (cf. (1.2.21)). In the following, we focus on the odd case where $\mathcal{D} = \mathrm{i}\gamma^\mu\nabla_\mu$, the even case just differs by a sign. First of all we have $\mathcal{D}^2\psi = -\gamma^\mu\gamma^\nu\nabla_\mu\nabla_\nu\psi$, and so it follows that

$$\begin{aligned}
2\,\mathcal{D}^2\psi &= -\gamma^\mu\gamma^\nu\nabla_\mu\nabla_\nu\psi - \gamma^\nu\gamma^\mu\nabla_\nu\nabla_\mu\psi \\
&= -\gamma^\mu\gamma^\nu\nabla_\mu\nabla_\nu\psi - \gamma^\nu\gamma^\mu\nabla_\mu\nabla_\nu\psi - \gamma^\nu\gamma^\mu R_{\nu\mu}\psi \\
&= -\{\gamma^\mu, \gamma^\nu\}\nabla_\mu\nabla_\nu\psi - \gamma^\nu\gamma^\mu R_{\nu\mu}\psi \\
&= 2g^{\mu\nu}\nabla_\mu\nabla_\nu\psi - \gamma^\nu\gamma^\mu R_{\nu\mu}\psi \\
&= 2\nabla^\mu\nabla_\mu\psi - \gamma^\nu\gamma^\mu R_{\nu\mu}\psi\,.
\end{aligned}$$

Here $R_{\mu\nu}$ denotes the spinorial curvature. It remains to show that $\gamma^\nu\gamma^\mu R_{\nu\mu}$ is the correct multiple of the scalar curvature. To this end one first expresses the spinorial curvature in terms of the curvature of the Levi-Civita connection via (1.2.20) and then makes use of (anti-)symmetry properties of the curvature tensors. The computation is not difficult but a little tricky, see for instance [**LM89**, p. 161]. $\square$

[15] To see this, simply observe that for ϕ antiperiodic, the function $\psi = \mathrm{e}^{-\mathrm{i}/2\theta}\phi$ is periodic and therefore it can be L^2-decomposed into the functions ψ_k. But then ϕ can be L^2-decomposed into the functions $\phi_k = \mathrm{e}^{\mathrm{i}/2\theta}\psi_k$.

CHAPTER 2

Some Analytical Tools for the Dirac Equation on Lorentzian Manifolds

The aim of this chapter is to summarize some general analytical methods one can use to study the Dirac equation on a Lorentzian manifold, thereby placing the methods used later in this thesis into a larger context.

The central analytical paradigm of this chapter is that the Dirac equation on a Lorentzian manifold is a *hyperbolic evolution equation*, more specifically a *symmetric hyperbolic system*. This makes it natural in the first place to think about its solutions as evolving from initial data, and hence to describe the space of all solutions in terms of an initial value problem. As far as concrete methods are concerned, one of the most important tools for hyperbolic equations are *energy estimates*, coming (often) from conserved or "almost-conserved" quantities. Specifically for the Dirac equation, a "fundamental conserved quantity" which is *always* available is the *conserved current*.

In the following these methods will be explained, as well some further methods such as the use of *Green's function*. Throughout this whole chapter, (M, g) denotes a Lorentzian spin manifold with spinor bundle $\mathbb{S}M$ and Dirac operator $\mathcal{D}$.

2.1. The Conserved Current

Let us start with what is arguably the most useful single tool for the analysis of the Dirac equation, the *conserved current*.

Definition 2.1.1. (Current) The *current vector field* or *associated vector field* of a spinor field $\psi \in \Gamma^\infty(\mathbb{S}M)$ is the vector field $J[\psi] \in \Gamma^\infty(TM)$ defined (in abstract index notation) as

$$J[\psi]^\mu := \prec \psi, \gamma^\mu \psi \succ_{\mathbb{S}M}, \tag{2.1.1}$$

where $\gamma^\mu : TM \to \mathsf{End}(\mathbb{S}M)$ denotes Clifford multiplication.[1]

Alternatively, one can also represent the current as the metrically equivalent 1-form $j[\psi]_\mu = \prec \psi, \gamma_\mu \psi \succ_{\mathbb{S}M}$, which in index-free notation can be written as

$$j[\psi](X) = \prec \psi, \gamma(X)\psi \succ_{\mathbb{S}M} \qquad \forall X \in TM\,. \tag{2.1.2}$$

The current satisfies an important positivity property: If $Z \in \Gamma^\infty(TM)$ is timelike and future-pointing, then

$$j[\psi](Z) = \prec \psi, \gamma(Z)\psi \succ_{\mathbb{S}M} \geq 0 \qquad \forall \psi \in \Gamma^\infty(\mathbb{S}M)\,. \tag{2.1.3}$$

This follows from the fact that $\prec \cdot, \gamma(Z) \cdot \succ_{\mathbb{S}M}$ is positive definite, see the end of Section 1.2.3 and Lemma 1.1.23.

Since $\mathcal{D} = \mathrm{i}\gamma^\mu \nabla_\mu$, it is almost self-evident that the Dirac operator will show up if we compute the divergence of $J[\psi]$.

[1]Since Clifford multiplication is symmetric, $J[\psi]$ is indeed a real vector field.

Lemma 2.1.2. *For any $\psi \in \Gamma^\infty(\mathbb{S}M)$ it holds that*

$$\operatorname{div} J[\psi] = \frac{1}{2}\,\mathsf{Im} \prec \mathcal{D}\psi, \psi \succ \ . \tag{2.1.4}$$

PROOF. Let $p \in M$, and let $\{e_\mu\}_{\mu=0,\dots,n}$ be a synchronous local orthonormal frame, i.e. such that $\nabla e_\mu|_p = 0$ (such local frames always exist, cf. [**Pet06**, Ch. 2, Ex. (5)]). The metric coframe $\{e^\mu\}$ of course has this property as well then. Denoting Clifford multiplication with e^μ by γ^μ and writing ∇_μ for ∇_{e_μ}, it follows that

$$\nabla_\mu(\gamma^\mu\psi)|_p = \gamma^\mu\nabla_\mu\psi|_p = -\mathrm{i}\,\mathcal{D}\psi|_p\,.$$

Here we also used that Clifford multiplication is parallel. Using once more that the chosen frame is synchronous at p, it now it follows that

$$\begin{aligned}
\operatorname{div} J[\psi](p) &= \nabla_\mu J[\psi]^\mu|_p \\
&= \nabla_\mu \prec \psi, \gamma^\mu\psi \succ_{\mathbb{S}M}|_p \\
&= \prec \nabla_\mu\psi, \gamma^\mu\psi \succ_{\mathbb{S}M}|_p + \prec \psi, \nabla_\mu\gamma^\mu\psi \succ_{\mathbb{S}M}|_p \\
&= \prec \gamma^\mu\nabla_\mu\psi, \psi \succ_{\mathbb{S}M}|_p + \prec \psi, \gamma^\mu\nabla_\mu\psi \succ_{\mathbb{S}M}|_p \\
&= \prec -\mathrm{i}\,\mathcal{D}\psi, \psi \succ_{\mathbb{S}M}|_p + \prec \psi, -\mathrm{i}\,\mathcal{D}\psi \succ_{\mathbb{S}M}|_p \\
&= \frac{1}{2}\,\mathsf{Im} \prec \mathcal{D}\psi, \psi \succ_{\mathbb{S}M}|_p\,.
\end{aligned}$$

Here we also used that Clifford multiplication is symmetric with respect to $\prec\cdot,\cdot\succ_{\mathbb{S}M}$. □

As an immediate consequence, $J[\psi]$ is divergence-free whenever $\prec \mathcal{D}\psi, \psi \succ_{\mathbb{S}M}$ is *real.*

Corollary 2.1.3. (Differential current conservation) *Let $\psi \in \Gamma^\infty(\mathbb{S}M)$. Suppose that $\mathcal{D}\psi = A\psi$ for some $A \in \Gamma^\infty(\mathsf{End}(\mathbb{S}M))$ which is (pointwise) symmetric with respect to $\prec\cdot,\cdot\succ_{\mathbb{S}M}$. Then* $\operatorname{div} J[\psi] = 0$. *In particular, this holds if $\mathcal{D}\psi = m\psi$ for some $m \in \mathbb{R}$.*

Integrating the divergence of the current over a bounded domain in spacetime and using *Gauß' law*, one can pass to the corresponding conservation law in integral form. The following is a special version of this.

Corollary 2.1.4. (Integrated current conservation) *Let (M, g) be a globally hyperbolic Lorentzian spin manifold, and let $\psi \in \Gamma^\infty_{sc}(\mathbb{S}M)$ be a solution of $(\mathcal{D}-m)\psi = 0$ with spatially compact support.*[2] *Let $\Sigma \subset M$ be a smooth, spacelike Cauchy surface with future-directed unit normal $\nu_\Sigma \in \Gamma^\infty(TM|_\Sigma)$ and induced volume element* $\mathrm{d}\mu_\Sigma$. *Set*

$$\mathfrak{C}[\psi] := \int_\Sigma \prec \psi, \gamma(\nu_\Sigma)\psi \succ_{\mathbb{S}M} \mathrm{d}\mu_\Sigma \in [0,\infty)\,. \tag{2.1.5}$$

Then this number is independent of the particular choice of Cauchy surface $\Sigma \subset M$.

PROOF. During this proof, for a Cauchy surface $\Sigma \subset M$ we denote the integral on the right-hand side of (2.1.5) by $\mathfrak{C}_\Sigma(\psi)$. Now let $\Sigma_1, \Sigma_2 \subset M$ be two Cauchy surfaces. The following neat argument is taken from [**BTW14**, p. 18], but it may be older.

Assume first that $\Sigma_2 \subset I^+(\Sigma_1)$, so that the domain $\Omega := I^+(\Sigma_1) \cap I^-(\Sigma_2)$ has a smooth boundary which is simply the union of Σ_1 and Σ_2, as illustrated in the left part of figure 2.1. Then a straight-forward application of Gauß' theorem to $0 = \int_\Omega \operatorname{div} J[\psi]\,\mathrm{d}\mu_g$ shows that $\mathfrak{C}_{\Sigma_1}(\psi) = \mathfrak{C}_{\Sigma_2}(\psi)$.

[2]Recall that having *spatially compact support* means that $\operatorname{supp}\psi \subset J(K)$ for some compact set $K \subset M$, where $J(K) = J^+(K) \cup J^-(K)$ is the union of the causal future and past of K. In particular this implies that $\operatorname{supp}\psi|_\Sigma \subset \Sigma$ is compact for any Cauchy surface $\Sigma \subset M$.

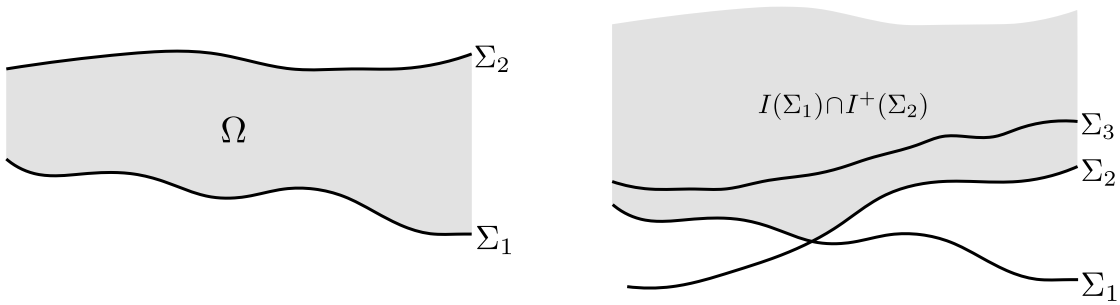

FIGURE 2.1. Illustration of the Cauchy surfaces in the proof of Corollary 2.1.4.

In case $\Sigma_1 \cap \Sigma_2 \neq \emptyset$, observe that $M' := I^+(\Sigma_1) \cap I^+(\Sigma_2)$ is again globally hyperbolic with respect to $g|_{M'}$. Thus, by [**BS03**, Thm. 1] there exists a smooth spacelike Cauchy surface $\Sigma_3 \subset M'$, see the right image in figure 2.1. It is not difficult to see that Σ_3 is also a Cauchy surface for M, and clearly $\Sigma_3 \subset I^+(\Sigma_1), I^+(\Sigma_2)$. Therefore the previous argument shows that $\mathfrak{C}_{\Sigma_1}(\psi) = \mathfrak{C}_{\Sigma_3}(\psi) = \mathfrak{C}_{\Sigma_2}(\psi)$. □

To see more clearly what this has to do with a conserved quantity in the usual sense, assume that $M \cong \mathbb{R}_t \times \Sigma_x$ is a foliation of M by Cauchy surfaces. Then the previous corollary states that if $\psi \in \Gamma^\infty_{sc}(\mathbb{S}M)$ satisfies $(\mathcal{D}-m)\psi = 0$, then

$$\mathfrak{C}(\psi) = \int_\Sigma \prec \psi(t,x), \gamma(\nu_{\Sigma_t})\psi(t,x) \succ_{\mathbb{S}M} \mathrm{d}\mu_{\Sigma_t}(x) \qquad \forall t \in \mathbb{R}\,, \tag{2.1.6}$$

i.e. the right-hand side is independent of t, and thus conserved under the time evolution given by the Dirac equation.

The right-hand side is of course precisely the L^2-norm of $\psi|_{\{t\}\times\Sigma}$ over the Cauchy hypersurface $\Sigma_t = \{t\} \times \Sigma$. Notice in this context that by (2.1.3) or Lemma 1.1.23, $\prec \cdot, \gamma(\nu_{\nu_{\Sigma_t}}) \cdot \succ_{\mathbb{S}M}$ is indeed a positive definite inner product on $\mathbb{S}M|_\Sigma$. However, it should be kept in mind that in the right-hand side not only ψ but also the inner product $\prec \cdot, \cdot \succ_{\mathbb{S}M}$, the unit-normal ν_{Σ_t}, and the volume element $\mathrm{d}\mu_{\Sigma_t}$ will in general depend explicitly on t. Only the specific combination in (2.1.6) is then independent of t (as long as $(\mathcal{D}-m)\psi = 0$). Of course, if the splitting $M \cong \mathbb{R}_t \times \Sigma$ is a *standard static splitting* (cf. Appendix A.1.2), then the t-dependence will be completely contained in ψ.

On the other hand, it should really be stressed once more that even in the *non-static* situation, the current (2.1.6) is always a *positive conserved quantity*. This is special, since many other equations, e.g. the scalar wave equation, admit conserved quantities only if the underlying spacetime has *symmetries* in the form of Killing fields (compare Section 2.3.3 or [**Ali10**, Ch. 4]). In this sense the Dirac equation is better behaved.

Remark 2.1.5. As the proof of Lemma 2.1.2 shows, for any $\psi, \phi \in \Gamma^\infty(\mathbb{S}M)$ it holds that

$$\nabla_\mu \prec \psi, \gamma^\mu \phi \succ_{\mathbb{S}M} = \mathrm{i}\left(\prec \mathcal{D}\psi, \phi \succ_{\mathbb{S}M} - \prec \psi, \mathcal{D}\phi \succ_{\mathbb{S}M}\right)\,. \tag{2.1.7}$$

This shows that $\mathcal{D}$ is formally self-adjoint with respect to the indefinite inner product

$$\langle \psi | \phi \rangle := \int_M \prec \psi, \phi \succ_{\mathbb{S}M} \mathrm{d}\mu_M \tag{2.1.8}$$

in the sense that

$$\langle \mathcal{D}\psi | \phi \rangle = \langle \psi | \, \mathcal{D}\phi \rangle \qquad \forall \psi \in \Gamma^\infty_c(\mathbb{S}M)\,, \phi \in \Gamma^\infty(\mathbb{S}M)\,. \tag{2.1.9}$$

For a Riemannian spin manifold (M, g), the inner product $\prec \cdot, \cdot \succ_{\mathbb{S}M}$ and thus also $\langle \cdot | \cdot \rangle$ is positive definite (cf. Corollary 1.1.21). Similar computations as above then show that the Dirac operator is a symmetric operator on $\Gamma_c^\infty(\mathbb{S}M) \subset \Gamma_{L^2}(\mathbb{S}M)$, the latter being defined as the closure of $\Gamma_c^\infty(\mathbb{S}M)$ with respect to the inner product $\langle \cdot | \cdot \rangle$. If (M, g) is complete, then $\mathcal{D}$ is even essentially self-adjoint on $\Gamma_c^\infty(\mathbb{S}M)$, see [**Fri00**, Ch. 4].

2.2. The Dirac Equation as Symmetric Hyperbolic System

As for any (differential) equation, two rather broad types of questions which one can ask about the Dirac equation are as follows:[3]

(A) Does the equation have solutions at all (existence)?
If yes, then how many solutions does it have (uniqueness)?
How can one conveniently describe or "parametrize" the set of all solutions?

(B) Which specific properties do general (or special) solutions have?

The main question of this thesis falls in category (B): We are interested in specific asymptotic properties of solutions of the massive Dirac equation. As is probably true for most type (B) questions this requires a variety of methods, some of which will be touched on later in this chapter (and will in special cases be applied in more detail later in the thesis).

The aim of this section is to give a brief answer to the questions in category (A). Here the fact that the Dirac equation is a hyperbolic equation allows to give a rather complete general answer in terms of the *Cauchy problem*, at least if the underlying spacetime is globally hyperbolic. In this context, let us recall that an equation is said to have a *well-posed* Cauchy problem for a certain class of initial data, if for each datum of this class there exists a unique solution. Usually one also demands that the solution depends continuously on the initial data in a suitable sense, but we ignore this in the following.

2.2.1. Symmetric hyperbolic systems and examples. Before we turn to the Dirac equation, we sketch some parts about the theory of symmetric hyperbolic systems. These are a particular type of hyperbolic equations which contain the Dirac equation. For more details and proofs we refer to standard textbooks such as [**Joh82**, Ch. 5], [**Eva10**, Ch. 7], [**Rin09**, Ch. 7], [**Fin02**], [**Ren08**], or [**Tay11**, Ch. 16] for the nonlinear case.

A general linear first order partial differential equation for a function $u : \mathbb{R}^{1,n} \to \mathbb{C}^N$ is always of the form

$$A^0(t,x)\partial_t u(t,x) + A^j(t,x)\partial_j u(t,x) + B(t,x)u(t,x) = w(t,x) \qquad (2.2.1)$$

for some given coefficients $A^0, A^1, \dots, A^n, B : \mathbb{R}^{1,n} \to \mathsf{M}(\mathbb{C}, N)$ and a given inhomogeneity $w : \mathbb{R}^{1,n} \to \mathbb{C}^N$. Such an equation is called *symmetric hyperbolic* if

(SH1) $A^\mu(t,x)$ is a Hermitian matrix for all $(t,x) \in \mathbb{R}^{1,n}$ and all $\mu = 0, \dots, n$.
(SH2) $A^0(t,x)$ is positive definite for all $(t,x) \in \mathbb{R}^{1,n}$, uniformly in (t,x).

Here being Hermitian and positive definite should be understood with respect to some (fixed) positive definite Hermitian inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{C}^N$.

Usually one further demands that the coefficients and the inhomogeneity possess some specified degree of regularity. For this brief exposition let us simply assume that they are smooth and all their derivatives are bounded.

[3] There are of course many other questions one might be interested in and which do not really fit into either of these two categories, for instance regarding (nonlinear) perturbations of an equation.

The aim in the following is to outline why these conditions allow to answer question (A) by the Cauchy problem. We first illustrate this by some concrete, simple examples.

Example 2.2.1. For $\lambda \in \mathbb{R}$, consider the simple scalar transport equation

$$\partial_t u - \lambda \partial_x u = 0\,,$$

which is symmetric hyperbolic. Its solutions are of the form

$$u(t,x) = u_0(\lambda t + x)\,. \tag{2.2.2}$$

The following points should be noticed: Firstly, u is *uniquely determined* by the *initial data* $u_0(x) = u(0,x)$, and u *depends continuously* on u_0 in any possible sense. Secondly, for any given initial data u_0 (which is sufficiently regular), u defined by (2.2.2) is a solution (*existence*). Thus the Cauchy problem is well-posed. Finally, notice that the value of u at any point $(t,x) \in \mathbb{R}$ depends only on the value of u_0 at the point $x + \lambda t$. One calls this effect *finite speed of propagation.* Actually, here the propagation speed is precisely given by $c = \lambda$.

As we will see later, these three properties, i.e.

(U) **U**nique determination by (and continuous dependence on) the initial data,
(E) **E**xistence of solutions for any given initial data,
(FPS) **F**inite **p**ropagation **s**peed,

are characteristic for (linear) hyperbolic evolution equations. Of course, in concrete situation one has to specify more explicitly what "any initial data" (and "continuous dependence") mean in order to turn these properties into clear mathematical statements. In the following, "any initial data" will be taken to mean any smooth initial data. Before we explain the analogue general results for symmetric hyperbolic systems, let us consider some more examples.

Example 2.2.2. If $A \in \mathsf{M}(N,\mathbb{C})$ is a Hermitian matrix, then the system

$$\partial_t u - A \partial_x u = 0$$

is symmetric hyperbolic. It can be explicitly solved, in a way similar to Example 2.2.1, by first computing the eigenvalues $\lambda_1, \ldots, \lambda_N \in \mathbb{R}$ and eigenvectors $e_1, \ldots, e_N \in \mathbb{C}^N$ of the matrix A. In terms of these, any solution of the equation has the form

$$u(t,x) = \sum_{j=1}^{N} \langle u_0(\lambda_j t + x), e_j \rangle_{\mathbb{C}^N}\, e_j\,, \tag{2.2.3}$$

where $u_0(x) = u(0,x)$ are again the initial values of the solution. Notice that the assumption of A being Hermitian is needed to guarantee that the eigenvalues $\lambda_1, \ldots, \lambda_N$ are *real.* If this were not the case and λ_j were a *complex* eigenvalue, then the expressions $u_0(\lambda_j t + x)$ in the right-hand side of (2.2.3) would not make sense. From (2.2.3) one immediately sees that also here the properties *i.)*, *ii.)*, and *iii.)* hold. Only now there are in general different propagation speeds associated to the different eigenvalues of A.

Example 2.2.3. One can of course also consider more than one spatial variable. For instance, for $\lambda_1, \lambda_2 \in \mathbb{R}$ the equation

$$\partial_t u - \lambda_1 \partial_x u - \lambda_2 \partial_y u = 0$$

is symmetric hyperbolic. Its solutions have the form

$$u(t,x,y) = u_0(\lambda_1 t + x, \lambda_2 t + y)\,,$$

where $u_0(x,y) = u(0,x,y)$ are the initial values again. As before, also here the properties *i.)*, *ii.)*, and *iii.)* hold.

Besides these special cases (and maybe a few others), one cannot solve equation (2.2.1) in a similar simple fashion.

As last example we consider an equation which is *not* symmetric hyperbolic, in order to have an example of an equation for which the Cauchy problem is not the correct way to study the equation.

Example 2.2.4. As a counter example, consider the equation

$$\partial_t u + \mathrm{i}\partial_x u = 0\,, \tag{2.2.4}$$

which is *not* symmetric hyperbolic since $\mathrm{i} \in \mathsf{M}(1,\mathbb{C})$ is of course not Hermitian. Writing the function u in the form $u = a + \mathrm{i}b$, we have

$$\partial_t u + \mathrm{i}\partial_x u = \big(\partial_t a - \partial_x b\big) + \mathrm{i}\big(\partial_x a + \partial_t b\big)\,.$$

This shows that if u satisfies the equation (2.2.4), then it actually satisfies the Cauchy-Riemann equations in the complex plane with coordinates t and x, i.e. it is analytic. Since analytic functions possess a convergent power series expansion around any point, it follows that for "initial values" $u_0(x)$ which are *not* real analytic, there cannot exist a solution u to (2.2.4) with $u(0,x) = u_0(x)$.

Even if one takes a real analytic function u_0 there need not exist a solution u of (2.2.4) defined on all of $\mathbb{C}$ with $u(0,x) = u_0(x)$. For instance, consider $u_0(x) = (x - \mathrm{i})^{-1}$, then the candidate for a solution is of course $u(t,x) = (x + \mathrm{i}t - \mathrm{i})^{-1}$, which however has a singularity at $(t,x) = (1,0)$. One can of course always construct a local solution on some small neighborhood of $\{t = 0\} \subset \mathbb{C}_{(t,x)}$ using local power series expansions of u_0. Notice also that *if* u is a solution of (2.2.4), then by the unique continuation property of holomorphic functions it is already uniquely determined by its "initial values" $u_0(x) = u(0,x)$.

What all this indicates is simply that the space of all solutions of (2.2.4) is not conveniently describable in terms of "initial values" at $t = 0$, so the Cauchy problem is not the correct way of looking at equation (2.2.4).

Next we outline the well-posedness of the Cauchy problem for symmetric hyperbolic systems. Of course we are not going to give a proof of this result here, but we will at least sketch some ideas.

2.2.2. From Fourier methods for simple systems to L^2-methods. One of the most fundamental observations about symmetric hyperbolic systems and wave equations in general is perhaps that most of their (linear) theory is an L^2-*theory* as will now be explained.

On a simple level, this can be revealed easily by Fourier analysis. To this end, let us consider once more the Cauchy problem for the scalar transport equation of Example 2.2.1, but this time on $\mathbb{R}_t \times S^1_\theta$ instead of $\mathbb{R}_t \times \mathbb{R}_x$ for simplicity. That is, we want to find a solution $u \in C^\infty(\mathbb{R}_t \times S^1_\theta)$ of the initial value problem

$$\begin{cases} \partial_t u - \lambda \partial_\theta u = 0 \\ \qquad u|_{t=0} = u_0 \end{cases}, \tag{2.2.5}$$

where $u_0 \in C^\infty(S^1)$ is given. Any smooth function $u \in C^\infty(\mathbb{R}_t \times S^1_\theta)$ can, for each fixed $t \in \mathbb{R}$, be decomposed in a Fourier series

$$u(t,\theta) = \sum_{k\in\mathbb{Z}} \widehat{u}_k(t)\mathrm{e}^{\mathrm{i}k\theta}\,,$$

where the Fourier coefficients $\widehat{u}_k(t)$ of course depend on t. Since u is smooth, this series converges pointwise and may be differentiated termwise. It also converges in $L^2(S^1)$. Further, it follows from standard properties of Fourier series that u satisfies the transport equation in (2.2.5) if and only if the Fourier coefficients satisfy the ordinary differential equations

$$\partial_t \widehat{u}_k(t) = -\mathrm{i}\lambda k \widehat{u}_k(t) \qquad \forall k \in \mathbb{Z}\,.$$

This simple equation is of course easily solved explicitly by

$$\widehat{u}_k(t) = \mathrm{e}^{-\mathrm{i}\lambda k t}\widehat{u}_k(0)\,,$$

where $\widehat{u}_k(0)$ are the Fourier coefficients of the initial value $u_0 \in C^\infty(S^1)$. Hence the candidate for the solution of the initial value problem (2.2.5) is [4]

$$u(t,\theta) = \sum_{k\in\mathbb{Z}} \widehat{u}_k(0)\mathrm{e}^{-\mathrm{i}\lambda k t}\mathrm{e}^{\mathrm{i}k\theta}\,. \tag{2.2.6}$$

Let us make some remarks about (2.2.6) which serve to illustrate the general theory:

(1) Since $|\mathrm{e}^{-\mathrm{i}\lambda k t}| = 1$, the Fourier series (2.2.6) converges just as good as the Fourier series of u_0.

(2) If u_0 is smooth, as we assumed above, then also u defined by (2.2.6) is smooth, both in θ and in t, and solves the initial value problem (2.2.5).

Even if we only assume $u_0 \in C^1(S^1)$, the function $u(t,\theta)$ defined by (2.2.6) is still once continuously differentiable and solves (2.2.5). Concerning differentiability, notice that formally taking either a t- or a θ-derivative of the right-hand side of (2.2.6) (by differentiating each term of the Fourier series), the k-th Fourier coefficient gets multiplied by k. The assumption $u_0 \in C^1$ guarantees that these resulting series still converge pointwise since one knows that that $u_0'(\theta) = \sum_{k\in\mathbb{Z}} \widehat{u}_k(0)\mathrm{i}k\mathrm{e}^{\mathrm{i}k\theta}$ for each $\theta \in S^1$. It is not difficult to show that they coincide with the t- and θ-derivatives of u, respectively.

(3) Moving towards the L^2-theory notice that even if we only assume $u_0 \in L^2(S^1)$, the Fourier series (2.2.6) still converges in L^2 (due to (1)). But now the function $u(t,\theta)$ defined by it need no longer be differentiable, but will in general only lie in the space $C(\mathbb{R}_t; L^2(S^1_\theta))$.[5] However, let us assume that $u_0 \in H^1(S^1)$, i.e. besides having the identity $u_0(\theta) = \sum_{k\in\mathbb{Z}} \widehat{u}_k(0)\mathrm{e}^{\mathrm{i}k\theta}$ in $L^2(S^1)$, we also assume that $\sum_{k\in\mathbb{Z}} \widehat{u}_k(0)k\mathrm{e}^{\mathrm{i}k\theta}$ converges in $L^2(S^1)$. Then it follows similar to the argument in (2) that u defined by (2.2.6) will lie in the space $C(\mathbb{R}_t; H^1(S^1_\theta)) \cap C^1(\mathbb{R}_t; L^2(S^1_\theta))$, and solve the initial value problem (2.2.6) if we interpret the θ-derivative as a weak derivative and interpret the equality in the transport equation (2.2.5) as an equality in the space $C(\mathbb{R}_t; L^2(S^1_\theta))$.

To make a long story short, the point is that the initial value problem (2.2.5) can be solved explicitly using Fourier series. Moreover, since the use of Fourier series (and also the Fourier transform) is at heart an “L^2-method” (being a unitary map between

[4] Notice that due to $\mathrm{e}^{-\mathrm{i}\lambda k t}\mathrm{e}^{\mathrm{i}k\theta} = \mathrm{e}^{\mathrm{i}k(\theta-\lambda t)}$ we have $u(t,\theta) = u_0(\theta - \lambda t)$ as in Example 2.2.1, although here we strictly speaking have to interpret $\theta - \lambda t$ as $(\theta - \lambda t) \bmod 2\pi$ since we are on S^1.

[5] This means that $t \mapsto u(t,\cdot)$ is a continuous curve in $L^2(S^1)$.

$L^2(S^1)$ and $\ell^2(\mathbb{Z})$ in our case), the natural setting to solve the Cauchy problem (2.2.5) are L^2-based Sobolev spaces as sketched under point (3) above.

Remark 2.2.5. It is instructive to try to solve the initial value problem for the "Cauchy-Riemann equations" (2.2.4) by Fourier methods. To make matters more simple we again take the spatial part to be S^1, and try to find a smooth solution $u \in C^\infty(\mathbb{R}_t \times S^1_\theta)$ of the equation

$$\partial_t u = -\mathrm{i}\partial_\theta u\,, \tag{$*$}$$

with prescribed initial condition $u|_{t=0} = u_0 \in C^\infty(S^1)$. Expanding $u(t,\cdot)$ in a Fourier series for each fixed t as before, one sees that in order for $(*)$ to hold, the Fourier coefficients of u would have to satisfy

$$\partial_t \widehat{u}_k(t) = k\widehat{u}_k(t) \qquad \forall k \in \mathbb{Z}\,.$$

The unique solution to this equation is of course given by $\widehat{u}_k(t) = \mathrm{e}^{kt}\widehat{u}_k(0)$. Formally then, the solution u is given by

$$u(t,\theta) = \sum_{k\in\mathbb{Z}} \widehat{u}_k(0)\mathrm{e}^{kt}\mathrm{e}^{\mathrm{i}k\theta}\,. \tag{$**$}$$

The problem with this expression is the following: For fixed $t > 0$ the function $k \mapsto \mathrm{e}^{kt}$ grows (exponentially) as $k \to \infty$. Therefore, if the right-hand side of $(**)$ is supposed to converge (for this fixed $t > 0$), the Fourier coefficients of u_0 would need to decay "extremely rapid" as $k \to \infty$. Moreover, since for increasing t the growth of e^{kt} becomes ever faster, we would need to assume ever faster decay of $\widehat{u}_k(0)$ to be able to construct a solution for large times in this way. Basically the only simple condition on u_0 for which it is clear that one obtains a solution defined on all of $\mathbb{R}_t \times S^1_\theta$ in this way is that only finitely many Fourier coefficients of u_0 are nonzero.

To be fair, this only shows that Fourier methods (and thus L^2-methods) are not a suitable way to go about solving the initial value problem for the equation (2.2.4). Of course we already know from Example 2.2.4 that the Cauchy problem is probably not a good way to look at this equation at all.

The message one should take home from these elementary examples is as follows: Fourier methods (or L^2-methods) might not work for general equations, but they *do* work very well for hyperbolic equations.

2.2.3. General symmetric hyperbolic systems and energy estimates. While the possibility to explicitly solve equations by the Fourier transform of course immediately breaks down if one allows *variable coefficients* in the equation, the power of the L^2-machinery, in particular Sobolev spaces and related estimates ("energy estimates") still holds up well. As said before, explaining this in details is not the scope here and is covered well in many textbooks. However, before we state some results about well-posedness of the Cauchy problem for general (linear) symmetric hyperbolic systems, let us at least take a brief look at the basic parts in the arguments.

So let us return to the Cauchy problem for a general (linear) symmetric hyperbolic system, i.e. to the problem of finding for prescribed initial values $u_0 : \mathbb{R}^n_x \to \mathbb{C}^N$ a solution $u : \mathbb{R}_t \times \mathbb{R}^n_x \to \mathbb{C}^N$ of

$$\begin{cases} A^0(t,x)\partial_t u(t,x) + A^j(t,x)\partial_j u(t,x) + B(t,x)u(t,x) = w(t,x) \\ \hfill u|_{t=0} = u_0 \end{cases}, \tag{2.2.7}$$

where $A^0, A^1, \dots, A^n, B : \mathbb{R}_t \times \mathbb{R}^n_x \to \mathsf{M}(N, \mathbb{C})$ and $w : \mathbb{R}_t \times \mathbb{R}^n_x \to \mathbb{C}^N$ are assumed to be smooth functions satisfying the conditions (SH1) and (SH2) of symmetric hyperbolicity. Moreover, for simplicity we shall again assume that all these functions and also their derivatives of any order are globally bounded, and further that $w(t, \cdot)$ has finite Sobolev norms (in x) of any order for every fixed t.

The key tool in the analysis of a symmetric hyperbolic system is the so-called *energy*

$$\mathcal{E}_u(t) := \int_{\mathbb{R}^n} \big\langle u(t,x), A^0(t,x)u(t,x)\big\rangle \,\mathrm{d}x\,. \tag{2.2.8}$$

Since A^0 is assumed to be uniformly positive definite by (SH2), $\mathcal{E}_u(t)$ is equivalent to the L^2_x-norm of $u(t, \cdot)$. The reason for using the energy instead of the L^2-norm is that the energy is nicely related to the equation (2.2.7). To see this, we consider the vector field

$$J^\mu(t,x) = \langle u(t,x), A^\mu(t,x)u(t,x)\rangle\ . \tag{2.2.9}$$

Computing its divergence, we find

$$\begin{aligned}
\partial_\mu J^\mu &= \langle \partial_\mu u, A^\mu u\rangle + \langle u, (\partial_\mu A^\mu)u\rangle + \langle u, A^\mu \partial_\mu u\rangle \\
&= 2\,\mathsf{Re}\langle u, A^\mu \partial_\mu u\rangle + \langle u, (\partial_\mu A^\mu)u\rangle \\
&= 2\,\mathsf{Re}\langle u, Bu\rangle + 2\,\mathsf{Re}\langle u, w\rangle + \langle u, (\partial_\mu A^\mu)u\rangle\ ,
\end{aligned}$$

where the last line of course only holds if u satisfies the equation.

All this looks very similar to the computations with the *current vector field* of the Dirac equation discussed in the previous Section, which indicates that symmetric hyperbolic systems are the correct analytical framework for the Dirac equation. However, whereas the current vector field of a solution of the Dirac equation is always *divergence-free* (or equivalently, the current is conserved), the vector field (2.2.9) associated to a solution of a general symmetric hyperbolic system will in general *not* be divergence-free. As a consequence, the energy (2.2.8) will in general not be conserved, but rather satisfies

$$\mathcal{E}_u(t) = \mathcal{E}_u(0) + \int_0^t \left(\int_{\mathbb{R}^n} (\partial_\mu J^\mu)(s,x)\,\mathrm{d}x \right) \mathrm{d}s\,. \tag{2.2.10}$$

Indeed, this just follows by integrating $\partial_\mu J^\mu$ over the domain $[0,t] \times \mathbb{R}^n \subset \mathbb{R}^{1,n}$ and using Gauß' theorem (assuming u is such that all integrals are well-defined), similar as in Corollary 2.1.4.

Nevertheless, notice that if u satisfies the equation, then by the computation from above $\partial_\mu J^\mu$ only contains zeroth order terms in u. Keeping in mind that all the coefficients of the equation and the inhomogeneity are assumed to be bounded, it might therefore be reasonable to believe that one can at least manage to bound $\partial_\mu J^\mu$ by $\big\langle u, A^0 u\big\rangle$ in such a way that one can apply a Grönwall estimate to (2.2.10). This is indeed the case (cf. [**Rin09**, Sec. 7.2]) and one can show that if u is a smooth solution of equation (2.2.8) for which the energy is finite, then

$$\mathcal{E}_u(t)^{\frac{1}{2}} \leq \mathrm{e}^{Ct} \left(\mathcal{E}_u(0)^{\frac{1}{2}} + C \int_0^t \|w(s,\cdot)\|_{L^2(\mathbb{R}_x)}\,\mathrm{d}s \right) \tag{2.2.11}$$

Similar estimates can also be established for the "higher order energies"

$$\mathcal{E}^{(k)}_u(t) := \frac{1}{2} \sum_{|\alpha| \leq k} \mathcal{E}_{\partial^\alpha u}(t) \tag{2.2.12}$$

by differentiating the equation (2.2.1). All these estimates (and similar estimates in general) are referred to as *energy estimates*.

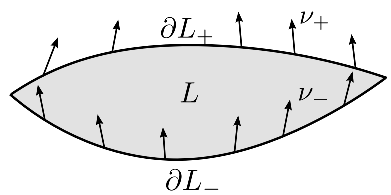

FIGURE 2.2. A lense-shaped domain $L \subset \mathbb{R}^{1+n}$, its boundary $\partial L = \partial L_+ \cup \partial L_-$ and the upward pointing unit-normals $\nu^\pm$ of the two boundary pieces.

Notice that (2.2.11) immediately implies uniqueness of smooth solutions of (2.2.7) whose energy is finite. Namely, in this case the right-hand side of (2.2.11) vanishes if applied to the difference of two solutions with the same initial data, and the thereby implied vanishing of the left-hand side implies that the two solutions must be equal.

One can remove the additional assumption of finite energy by using another key property of hyperbolic equations: *finite propagation speed*. For symmetric hyperbolic systems this manifests as follows: Consider a bounded domain $L \subset \mathbb{R}_t \times \mathbb{R}^n_x$ with piecewise smooth boundary as in figure 2.2, whose boundary decomposes into two smooth pieces $\partial L = \partial L_+ \cup \partial L_-$. The special shape of the domain L as sketched in figure 2.2 is meant to illustrate the following terminology: We say that L is *lense-shaped* if the boundary $\partial L \subset \mathbb{R}^{1+n}$ is spacelike with respect to the usual Minkowski metric on $\mathbb{R}^{1+n}$, and if moreover the matrix

$$A(p,\nu(p)) := A^\mu(t,x)\nu_\mu(t,x) = A^0(t,x)\nu_0(t,x) + A^j(t,x)\nu_j(t,x) \tag{2.2.13}$$

is positive definite for all $p = (t,x) \in \partial L$, where $\nu(p) \in \mathbb{R}^{1+n}$ is the upward pointing unit (co-)normal to ∂L in p, taken with respect to the Minkowski metric on $\mathbb{R}^{1+n}$. Notice that $A(p,\nu)$ is always positive definite if ν happens to be a (positive) multiple of ∂_t, since then $A(p,\nu)$ is just a (positive) multiple of $A^0(p)$, which is positive definite by (SH2). By continuity $A(p,\nu)$ will therefore still be positive definite if ν is sufficiently close to ∂_t.

Now suppose that $L \subset \mathbb{R} \times \mathbb{R}^n$ is a lense-shaped domain. If we integrate $\partial_\mu J^\mu$ over L and convert this integral into two boundary integrals using Gauß' theorem (with respect to the Minkowski metric), one over ∂L_+ and one over ∂L_-, then each of these integrals will be non-negative since L is lense-shaped. However, since we cannot say anything about the value of the interior integral of the divergence over L, this does not directly help to relate these boundary integrals in a useful way. But one can use the following little trick: Instead of integrating the divergence of J^μ, we take some $\lambda > 0$ and integrate the divergence of $\mathrm{e}^{-\lambda t} J^\mu$ instead. Then on the one hand, Gauß' law gives

$$\int_L \partial_\mu(\mathrm{e}^{-\lambda t} J^\mu)\,\mathrm{d}t\,\mathrm{d}x = \int_{\partial L_+} \mathrm{e}^{-\lambda t}\left\langle u, A^\mu \nu^+_\mu u\right\rangle \mathrm{d}\mu_+ - \int_{\partial L_-} \mathrm{e}^{-\lambda t}\left\langle u, A^\mu \nu^-_\mu u\right\rangle \mathrm{d}\mu_-\,, \tag{2.2.14}$$

where $\nu^\pm$ and $\mathrm{d}\mu_\pm$ are the future-pointing unit normal and the induced volume measure on $\partial L_\pm$ (all with respect to the Minkowski metric). On the other hand, and this is where the factor $\mathrm{e}^{-\lambda t}$ can be used, the divergence term in the integral on the left is simply

$$\partial_\mu(\mathrm{e}^{-\lambda t} J^\mu) = -\lambda \mathrm{e}^{-\lambda t} J^0 + \mathrm{e}^{-\lambda t}\partial_\mu J^\mu = \mathrm{e}^{-\lambda t}\big(-\lambda \left\langle u, A^0 u\right\rangle + \partial_\mu J^\mu\big)\,,$$

so that the left-hand side of (2.2.14) is given by

$$\int_L \partial_\mu(\mathrm{e}^{-\lambda t} J^\mu)\,\mathrm{d}t\,\mathrm{d}x = -\lambda \int_L \mathrm{e}^{-\lambda t}\left\langle u, A^0 u\right\rangle_{\mathbb{C}^N} \mathrm{d}t\,\mathrm{d}x + \int_L \mathrm{e}^{-\lambda t}\partial_\mu J^\mu\,\mathrm{d}t\,\mathrm{d}x\,.$$

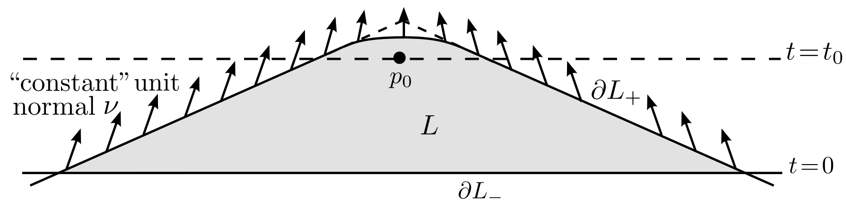

FIGURE 2.3. A "sufficiently flat (rounded off) cone" yields a lense-shaped region.

Since A^0 is positive definite, it follows that the left-hand side will be negative if only we choose $\lambda > 0$ sufficiently large, unless u vanishes everywhere on L (in which case both sides are just zero).

Returning to uniqueness of solutions, suppose that u_1, u_2 both satisfy the symmetric hyperbolic system (2.2.1) and coincide on ∂L_-. Then the difference $u := u_1 - u_2$, which also satisfies (2.2.1) (for $w = 0$) vanishes on ∂L_-, so that (2.2.14) reduces to

$$\int_L \partial_\mu(\mathrm{e}^{-\lambda t} J^\mu)\, \mathrm{d}t\, \mathrm{d}x = \int_{\partial L_+} \mathrm{e}^{-\lambda t} \left\langle u, A^\mu \nu_\mu^+ u \right\rangle \mathrm{d}\mu_+ \,.$$

Here the right-hand side is always non-negative since L is lense-shaped. On the other hand, unless $u|_L = 0$, the left-hand side will become negative if we choose $\lambda > 0$ sufficiently large. Therefore, to avoid a contradiction, we must have $u|_L = 0$ and hence $u_1|_L = u_2|_L$.

This argument can now be used to show uniqueness of smooth solutions of the initial value problem (2.2.7) without having to impose the additional restriction that the solutions have finite energy. To this end, suppose that $u_1, u_2 \in C^\infty(\mathbb{R} \times \mathbb{R}^n)$ both solve the initial value problem (2.2.7) for the same initial data. To show that $u_1 = u_2$, let $p_0 = (t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$. Since by assumption all coefficients of the symmetric hyperbolic system are bounded and A^0 is uniformly positive definite, there exists a fixed timelike vector $\nu \in \mathbb{R}^{1+n}$, close but *not* equal to e_0, such that the matrix $A^\mu(t,x)\nu_\mu$ is positive definite for *all* $(t,x) \in \mathbb{R} \times \mathbb{R}^n$. Next, as sketched in figure 2.3, take some cone in $\mathbb{R}^{1+n}$ containing p_0 whose boundary is spacelike and has ν (or some spatial rotations of ν) as upward pointing normal vector. Then by rounding off the tip of this cone we obtain a lense-shaped region L which contains p_0 and whose lower boundary ∂L_- is simply a piece of the hypersurface $\{t = 0\}$. And since u_1 and u_2 both coincide on $\{t = 0\}$ with u_0, the previous argument now shows that $u_1(p_0) = u_2(p_0)$. Since $p_0 \in \mathbb{R} \times \mathbb{R}^n$ was arbitrary, it follows that $u_1 = u_2$.[6]

Coming back to the notion of *finite propagation speed*, notice that any vector $\nu \in \mathbb{R}^{1+n}$ with the property that $A(t,x)^\mu \nu_\mu$ is positive definite for all $(t,x) \in \mathbb{R} \times \mathbb{R}^n$ gives an upper bound on the maximal speed of propagation. This can be seen for instance as illustrated in figure 2.4: Suppose that u solves the Cauchy problem (2.2.7) for initial data u_0. If we now change the data u_0 only in some finite region, then the solution can only change inside the grey-shaded cone for which ν (or suitable spatial rotations of ν) are normal to the boundary. This follows since any point outside this cone is contained in a lense-shaped region which does not intersect the part of the initial hypersurface $\{t = 0\}$ where

[6]One may notice that this argument does not require any boundedness assumption for B and w anymore since we just integrate over bounded domains. The matrices A^μ on the other hand need to be bounded to ensure the existence of suitable lense-shaped regions as used here.

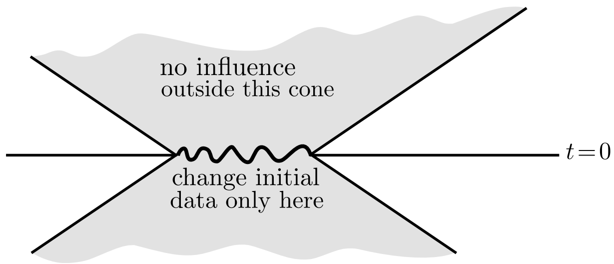

FIGURE 2.4. Illustration of finite speed of propagation.

we changed the initial data (like a cone as in figure 2.3). Notice in particular that if u_0 has *compact* support, then $u(t,\cdot)$ will still have compact support for any $t \in \mathbb{R}$. Here we of course always assume that the Cauchy problem actually has a solution.

So far we have outlined how energy estimates lead to uniqueness of smooth solutions and finite propagation speed. As for proving the *existence* of a solution of the Cauchy problem (2.2.7), one can use different methods: one possibility is to use functional analytic arguments to first show the existence of suitably defined weak solutions, followed by an appropriate regularity theory (cf. [**Rin09**, Sec. 7.4]); other methods include finite difference approximations (cf. [**Joh82**, Sec. 5.3]) or the vanishing viscosity method (cf. [**Eva10**, Sec. 7.3]). Since the actual arguments in all these methods are more lengthy than the simple arguments that lead to uniqueness, we do not discuss them at all but only mention that what is common to all these approaches is that energy estimates play a crucial role in each of them. In particular, here one also needs the "higher energies" (2.2.12). The following is the typical well-posedness result one obtains.

Theorem 2.2.6. **(H^k_{loc}-wellposedness for symmetric hyperbolic systems)** *Consider a symmetric hyperbolic system of the form* (2.2.1)*, and assume that all coefficients $A^0, \ldots, A^n, B$ and the inhomogeneity w are smooth and have bounded derivatives of any order. Then the following hold:*

i.) Let $k > \frac{n+3}{2}$. Then for any given initial data $u_0 \in H^k_{\mathrm{loc}}(\mathbb{R}^n_x; \mathbb{C}^N)$ there exists a unique global solution $u \in C(\mathbb{R}_t; H^k_{\mathrm{loc}}(\mathbb{R}^n_x; \mathbb{C}^N)) \cap C^1(\mathbb{R}_t; H^{k-1}_{\mathrm{loc}}(\mathbb{R}^n_x; \mathbb{C}^N))$ of the Cauchy problem (2.2.7)*.*

ii.) If $u_0 \in C^\infty(\mathbb{R}^n)$, then also $u \in C^\infty(\mathbb{R}_t \times \mathbb{R}^n_x)$.

As remarked before, the conditions on the coefficients and the inhomogeneity can be relaxed. For instance, due to finite propagation speed one actually does not need to assume any bounds on B and w (as long as they are smooth). For $A^0, \ldots, A^n$ one at least needs boundedness to guarantee (uniformly) finite propagation speed. It is also possible to relax the smoothness assumptions for the coefficients.

To conclude this discussion, let us stress once more the central role of energy estimates for symmetric hyperbolic systems. Since energy estimates are basically L^2-Sobolev space estimates, this illustrates again that the theory of linear symmetric hyperbolic systems is essentially an L^2-theory. In a broad sense these represent the adaption of "Fourier methods" to variable coefficient equations. Also later in this thesis energy estimates are one of the central tools we are going to use.

2.2.4. The Cauchy problem for the Dirac equation. Let us now return to the Dirac equation on a Lorentzian manifold. In order to have a well-posed Cauchy problem one has to make the assumption that the spacetime is *globally hyperbolic*, see Remark

2.2.7 below. Therefore in the following we assume that (M, g) is a globally hyperbolic Lorentzian spin manifold with spinor bundle $\mathbb{S}M$ and Dirac operator $\mathcal{D}$. In the following we will use some results from causal theory of Lorentzian manifolds, for which one may consult [**MS08**] or [**O'N83**, Ch. 14] as references.

Due to globally hyperbolicity, we may assume that the spacetime (M, g) has the form (cf. [**BS05**, Thm. 1.1])

$$\begin{cases} M = \mathbb{R}_t \times \Sigma_x \\ \quad g = -\beta(t,x)\,\mathrm{d}t^2 + h_t \end{cases} \tag{2.2.15}$$

where $\beta \in C^\infty(\mathbb{R} \times \Sigma)$ is a smooth, positive function, and h_t is a family of Riemannian metrics on Σ which depend smoothly on the parameter $t \in \mathbb{R}$. Moreover, in this decomposition each hypersurface $\Sigma_t := \{t\} \times \Sigma \subset M$ is a smooth spacelike Cauchy surface.

For given initial data $\psi_0 \in \Gamma^\infty(\mathbb{S}M|_{\Sigma_0})$, the Cauchy problem for the Dirac equation then consists in finding a solution $\psi \in \Gamma^\infty(\mathbb{S}M)$ to

$$\begin{cases} (\mathcal{D}{-}m)\psi = 0 \\ \qquad \psi|_{\Sigma_0} = \psi_0 \end{cases} . \tag{2.2.16}$$

In the following we will argue that there always exists a unique solution to this problem. The idea is to use the existence result for symmetric hyperbolic systems described in Theorem 2.2.6 to construct solutions locally, and then to patch these together.

To see the relation between the Dirac equation and symmetric hyperbolic systems, note that in local coordinates $(x^1, \ldots, x^n)$ on Σ and in a suitable local frame of the spinor bundle we have

$$\mathcal{D} = \mathrm{i}\gamma^\mu \nabla_\mu = -\mathrm{i}\beta(t,x)^{-\frac{1}{2}}\gamma_0 \partial_t + \mathrm{i}h^{ij}(t,x)\gamma_i\partial_j + \mathcal{B}(t,x)\,,$$

where $(\gamma_0, \gamma_1, \ldots, \gamma_n)$ is a set of γ-matrices in signature $(1, n)$, and $\mathcal{B}(t, x)$ contains all zero-order terms. Inverting the matrix in front of ∂_t (note that $\gamma_0^2 = 1$), it follows that ψ satisfies $(\mathcal{D}{-}m)\psi = 0$ if and only if

$$\partial_t\psi - \underbrace{\beta(t,x)^{\frac{1}{2}} h^{ij}(t,x)\gamma_0\gamma_i}_{=:A^j(t,x)} \partial_j\psi + B(t,x)\psi = 0\,, \tag{2.2.17}$$

where again $B(t, x)$ contains all zero-order terms. Due to the usual properties of γ-matrices we have

$$(\gamma_0\gamma_j)^\dagger = \gamma_j^\dagger\gamma_0^\dagger = -\gamma_j\gamma_0 = \gamma_0\gamma_j\,,$$

where $\dagger$ denotes taking the adjoint with respect to the standard inner product $\langle\cdot,\cdot\rangle$ on $\mathbb{C}^N$. To be more precise, this follows since by a suitable choice of spin frame the spinorial inner product $\prec\cdot,\cdot\succ_{\mathbb{S}M}$ is identified with $\langle\cdot, \gamma^0\cdot\rangle$. This shows that the matrices $A^j(t, x)$ are indeed Hermitian, and thus (2.2.17) has the form of a symmetric hyperbolic system.

Let us now argue for uniqueness first, since it is simpler and is also needed to patch together locally constructed solutions later. Similar as for symmetric hyperbolic systems before, uniqueness can be shown by a simple energy estimate. The only difference is that now we use the Dirac current defined in Definition 2.1.1, and the causal structure of (M, g) instead of the one of Minkowski spacetime. Since the current of a solution of the Dirac equation is *always* conserved, the argument is actually simpler than before. Namely, recall that if $\psi \in \Gamma^\infty(\mathbb{S}M)$ satisfies the Dirac equation $(\mathcal{D}{-}m)\psi = 0$, then the current $J^\mu = \prec\psi, \gamma^\mu\psi\succ_{\mathbb{S}M}$ is divergence-free (cf. Corollary 2.1.3). Using this one can show uniqueness of smooth solutions of (2.2.16) as follows:

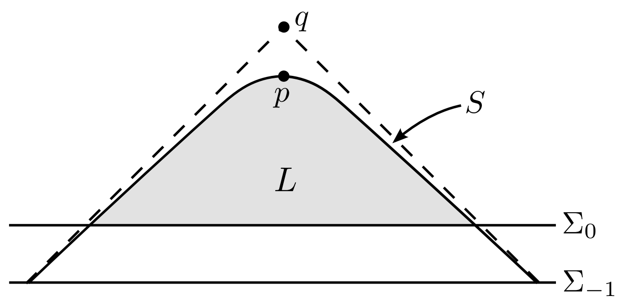

FIGURE 2.5. Choice of domain for the energy estimate to show uniqueness of solutions of the Cauchy problem in globally hyperbolic spacetimes.

Suppose that $\psi_1, \psi_2 \in \Gamma^\infty(\mathbb{S}M)$ are two solutions of (2.2.16) for the same Cauchy data $\psi_0 \in \Gamma^\infty(\mathbb{S}M|_{\Sigma_0})$. Set $\psi = \psi_1 - \psi_2$, and let $p \in M$ be any point, without loss $p \in I^+(\Sigma)$. As illustrated in figure 2.5, we first choose a point $q \in I^+(p)$, so that $p \in I^-(q) \cap I^+(\Sigma_{-1})$. One easily verifies that $I^-(q) \cap I^+(\Sigma_{-1})$ is still globally hyperbolic (with respect to the restricted metric), so there exists a smooth, spacelike Cauchy surface $S \subset I^-(q) \cap I^+(\Sigma_{-1})$ which contains p. Using that S is a Cauchy surface, it is not difficult to see that it must intersect Σ_0 in such a way that the two smooth spacelike hypersurfaces $\partial L_+ = S \cap J^+(\Sigma_0)$ and $\partial L_- = \Sigma_0 \cap J^-(S)$ bound an open bounded domain $L \subset M$, i.e. $\partial L = \partial L_+ \cup \partial L_-$. Therefore it follows from Corollary 2.1.3 and Gauß' theorem that

$$
\begin{aligned}
0 &= \int_L \nabla_\mu J^\mu \, \mathrm{d}\mu_g = \int_{\partial L_+} \prec\!\psi, \gamma(\nu_+)\psi\!\succ_{\mathbb{S}M} \mathrm{d}\mu_{\partial L_+} - \int_{\partial L_-} \prec\!\psi, \gamma(\nu_-)\psi\!\succ_{\mathbb{S}M} \mathrm{d}\mu_{\partial L_-} \\
&= \int_{\partial L_+} \prec\!\psi, \gamma(\nu_+)\psi\!\succ_{\mathbb{S}M} \mathrm{d}\mu_{\partial L_+} ,
\end{aligned}
$$

where $\nu_\pm$ are the future-directed unit normals of $\partial L_\pm$, and $\mathrm{d}\mu_{\partial L_\pm}$ are the induced volume measures. Here the integral over ∂L_- vanishes since $\psi|_{\Sigma_0} = \psi_1|_{\Sigma_0} - \psi_2|_{\Sigma_0} = 0$. Since ν_+ is timelike, we have $\prec\!\psi, \gamma(\nu_+)\psi\!\succ_{\mathbb{S}M} \geq 0$ (compare to Lemma 1.1.23). This now implies that $\psi|_S = 0$, so in particular $\psi(p) = 0$ and hence $\psi_1(p) = \psi_2(p)$. Since $p \in M$ was arbitrary, this shows that $\psi_1 = \psi_2$.

Next we argue that for any given $\psi_0 \in \Gamma^\infty(\mathbb{S}M|_{\Sigma_0})$ the Cauchy problem (2.2.16) in fact has a solution. As said before, to construct the solution one may patch together locally constructed solutions (constructed via Theorem 2.2.6), and use the already shown uniqueness. Roughly speaking, one proceeds as follows (for a precise argument, see [**Bär14**, Thm. 5.6]): Let $\psi_0 \in \Gamma^\infty(\mathbb{S}M|_{\Sigma_0})$ be given, and let $p \in M$. We want to construct a solution locally in some neighborhood of p. Without loss $p \in I^+(\Sigma_0)$. Take a point $q \in I^+(p)$, then by global hyperbolicity $I^-(q) \cap I^+(\Sigma_0)$ is a relatively compact open neighborhood of p, whose closure is $J^-(q) \cap J^+(\Sigma_0)$. We now construct a solution in $I^-(q) \cap I^+(\Sigma)$ in several steps. The construction is illustrated in figure 2.6.

First we cover the compact set $J^-(q) \cap \Sigma_0$ by finitely many, sufficiently small normal coordinate charts $(U_1, x_1), \ldots, (U_N, x_N)$ of the Cauchy surface Σ_0 such that the spinor bundle $\mathbb{S}M$ is trivializable over each Cauchy development $D(U_k) \subset M$.[7] Each of these Cauchy developments, being globally hyperbolic, has the form $D(U_k) \cong \mathbb{R} \times \mathbb{R}^n$, and the Dirac equation has the form (2.2.17) if expressed in the corresponding coordinates. Therefore we can apply Theorem 2.2.6 to get a (unique) solution ψ_k of the Dirac equation

[7]See [**O'N83**, Ch. 14] for the notion of Cauchy development.

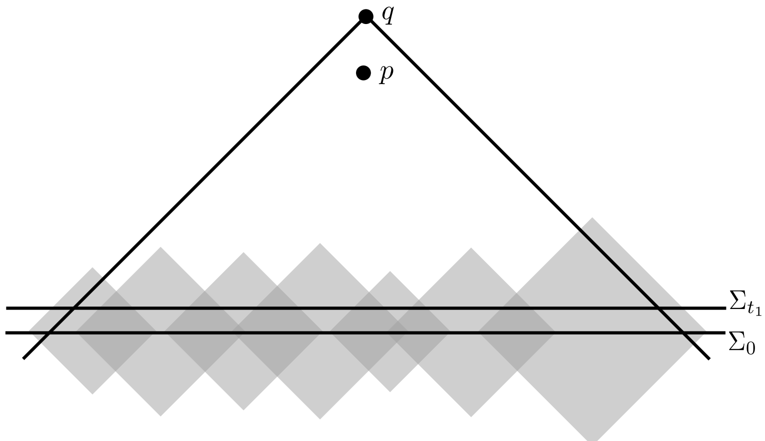

FIGURE 2.6. The first step in the construction of a solution of the Cauchy problem (2.2.16).

on $D(U_k)$ which on U_k coincides with the given initial data $\psi_0|_{U_k}$. Moreover, whenever $D(U_k) \cap D(U_\ell) \neq \emptyset$ the two solutions ψ_k and ψ_ℓ have to coincide on this intersection by the already established uniqueness. Note that $D(U_k) \cap D(U_\ell) = D(U_k \cap U_\ell)$ is again globally hyperbolic with Cauchy surface $U_k \cap U_\ell$. Therefore we can simply take the "union" of all these solutions, which yields a solution on some small "strip" $I^+(\Sigma_0) \cap I^-(\Sigma_{t_1}) \cap I^-(q)$, as sketched in figure 2.6.

In the next step we repeat this procedure, starting from Σ_{t_1} instead of Σ_0. This then yields a solution on $I^+(\Sigma_0) \cap I^-(\Sigma_{t_2}) \cap I^-(q)$ for some $t_2 > t_1$. If we want to continue like this, it is of course crucial to know whether we can actually reach the point p in this way. In other words, we have to ensure that the intermediate time steps $t_1, t_2, \dots$ do not converge to some $t_\infty < t(p)$. Here global hyperbolicity, which guarantees compactness of $J^-(q) \cap J^+(\Sigma_0)$ is again the key. Namely, one can show that by compactness of $J^-(q) \cap J^+(\Sigma_0)$ (which follows from global hyperbolicity) it is possible to choose the intermediate time steps *equidistant*, i.e. such that $t_j - t_{j-1} =: \Delta t$ is always the same. With this in mind it is now obvious that after finitely many steps we will have exhausted all of $I^-(q) \cap I^+(\Sigma_0)$.

In this way, for each $p \in I^+(\Sigma_0)$ we can construct a solution of the Dirac equation on some open neighborhood $I^-(q) \cap I^+(\Sigma_0)$ which on $I^-(q) \cap \Sigma_0$ coincides with ψ_0. Using once again the already established uniqueness, it is easy to see for two points $p_1, p_2 \in I^+(\Sigma_0)$ with corresponding solutions ψ_1, ψ_2, these two solutions must coincide on the intersection of their domains of definition. In this way we obtain a uniquely determined solution in $I^+(\Sigma_0)$, and for the past of Σ_0 one can proceed analogously.

This concludes the sketch of the proof of existence and uniqueness of solutions of the Cauchy problem (2.2.16) for smooth initial data. As a by-note, the same energy estimate that lead to uniqueness also establishes *finite propagation speed.* For instance, the solution ψ with initial data ψ_0 satisfies

$$\operatorname{supp} \psi \subset J(\operatorname{supp} \psi_0) = J^+(\operatorname{supp} \psi_0) \cup J^-(\operatorname{supp} \psi_0) . \tag{2.2.18}$$

As a consequence, if ψ_0 is compactly supported, the solution ψ will be spatially compactly supported. In particular $\psi|_{\Sigma_t}$ will have compact support for every $t \in \mathbb{R}$ in this case.

Let us conclude this section with some remark about the non-globally hyperbolic case, weak solutions, and more general symmetric hyperbolic systems on Lorentzian manifolds.

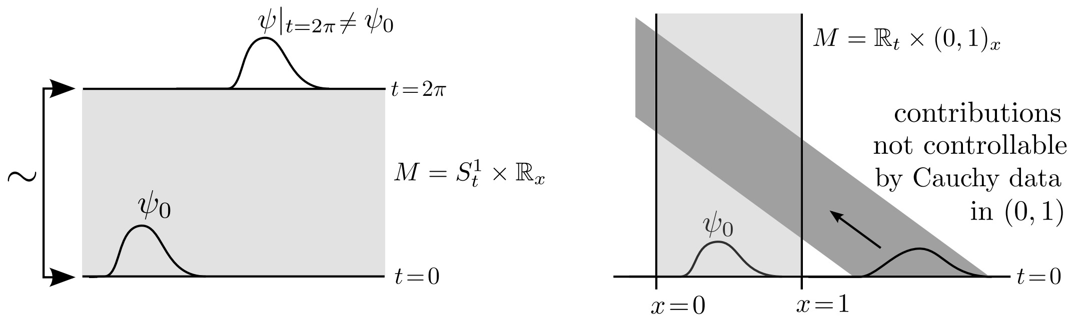

FIGURE 2.7. Possible failures of existence and uniqueness of the Cauchy problem. For explanations, see Remark 2.2.7.

Remark 2.2.7. For *non-* globally hyperbolic spacetimes there are two ways in which the Cauchy problem may fail to be well-posed: both existence and uniqueness of solutions for given initial conditions need not hold. This is illustrated by the following two examples, which are also sketched in figure 2.7: [8]

Non-existence: Start with any smooth solution $\psi : \mathbb{R}^{1,1} \to \mathbb{C}^2$ of the Dirac equation in $\mathbb{R}^{1,1}$ which is *not* 2π-periodic in t, i.e. $\psi(t+2\pi,x) \neq \psi(t,x)$. Now consider the quotient $M = \mathbb{R}^{1,1}/\sim$, where the equivalence relation is generated by $(t+2\pi,x) \sim (t,x)$. In the obvious way the metric, spin structure etc. descend from $\mathbb{R}^{1,1}$ to this quotient (note that M is not globally hyperbolic since there are closed timelike curves). Since ψ is not 2π-periodic in t, however, it does *not* descend to M. From this and local uniqueness of the Cauchy problem it follows that the Cauchy problem (2.2.16) on M with the same initial data $\psi_0 = \psi|_{t=0}$ cannot have a solution.

Non-uniqueness: Consider the finite strip $M = \mathbb{R}_t \times (0,1)_x \subset \mathbb{R}^{1,1}$, equipped with the Minkowski metric $g = -\,\mathrm{d}t^2 + \mathrm{d}x^2$. It is not globally hyperbolic since causal diamonds in (M,g) are not compact. For any compactly supported initial data $\psi_0 : (0,1)_x \to \mathbb{C}^2$ one can construct a solution of the Cauchy problem (2.2.16) by simply extending ψ_0 by zero to all of $\mathbb{R}_x$, solving the Cauchy problem in $\mathbb{R}^{1,1}$, and finally restricting the solution to M again. However, one can also extend ψ_0 in a completely different way (smoothly) to all $\mathbb{R}_x$, solve the Cauchy problem in $\mathbb{R}^{1,1}$ and restrict to M afterwards. Clearly this will in general result in a different solution, thus demonstrating non-uniqueness.

Remark 2.2.8. In this section we have only outlined how to construct smooth solutions starting from smooth initial data. However, pretty much the same arguments can also be used to show the existence of solutions in H^k_{loc} for data in H^k_{loc}. One can also study the persistence of intermediate or other types of regularity. For spacetimes which additionally are asymptotically flat, one finds a detailed analysis in [**Nic02**].

Remark 2.2.9. In the recent article [**Bär14**], a definition of symmetric hyperbolic systems on Lorentzian manifolds is made as follows ([**Bär14**, Def. 5.1]): If E is a vector bundle over a Lorentzian manifold (M,g) which is equipped with a nondegenerate (not necessarily positive definite) inner product $\langle\cdot,\cdot\rangle_E$, then a first order differential operator $P : \Gamma^\infty(E) \to \Gamma^\infty(E)$ is called *symmetric hyperbolic* if its principal symbol σ_P satisfies the following two conditions

[8] I owe the idea for the first example to Olaf Müller, see also [**MS13**, p. 161] or [**Mül14**].

i.) For every $\xi \in T_x^*M$, the principal symbol $\sigma_P(\xi) \in \mathsf{End}(E_x)$ is symmetric with respect to $\langle\cdot,\cdot\rangle$,

ii.) For every $\xi \in T_x^*M$ which is future-pointing and timelike, the bilinear form $\langle\cdot,\sigma_P(\xi)\cdot\rangle_E$ on E_x is positive definite.

As one can check, the Dirac operator satisfies these conditions. For such operators, one can make similar energy estimates as for the Dirac equation, using $\langle\cdot,\sigma_P(\xi)\cdot\rangle_E$ instead of the Dirac current.[9] The only difference is that this does not necessarily lead to a conserved "energy", and therefore one has to use Grönwall-type arguments as outlined in Section 2.2.3 for symmetric hyperbolic systems on $\mathbb{R}^n$. The upshot in the end is that on globally hyperbolic spacetimes such symmetric hyperbolic operators have a well-posed Cauchy problem, obey finite propagation speed as dictated by g (notice that condition *ii.)* ties the "causality" of the symmetric hyperbolic operator P to that of the spacetime), and have other useful properties. The arguments to show this are basically the same as the ones outlined above for the Dirac equation, compare [**Bär14**, Sec. 5].

2.3. A Collection of Further Methods

In the last part of this chapter we collect some further methods which can be used to analyze the Dirac equation (or also other hyperbolic equations), and which are used later in this thesis.

Before we start, let us recall that in this thesis the aim is to derive decay (and boundedness) estimates for solutions of the Dirac equation and the specific methods presented in the following are selected to that end. There exists of course a lot of literature about these and further methods for hyperbolic equations. Some textbooks which I found useful in that regard while working on this thesis, and which influenced the present text, were [**Ali10**], [**Fin02**], [**Hör97**], [**Kic96**], [**Lax06**], [**Sog08**], [**Tao06**].

2.3.1. Green's functions and representation formulas. If one wants to analyze a solution of some equation, the optimal situation one could wish for is to have an explicit expression for this solution. In most cases it is of course impossible to find exact expressions. Nevertheless it might still be possible to derive more abstract but still useful *representation formulas* (or identities) for a solution in terms of its initial values (or boundary values). One particular technique which is often useful to this end is that of *Green's functions*.

In the following we explain this technique for the Dirac equation, but also many other equations can be approached in the same spirit. We assume some familiarity with distributions (on manifolds), and we denote the action of a distribution T on a test function φ by $\langle T,\varphi\rangle$. A good basic reference for distributions is [**FJ98**], and concerning distributions on manifolds and vector bundles one may consult [**Wal12**, Ch. 1].

Let $p \in M$. A *Green's function at p* for the massive Dirac equation is a distribution $G_p \in \Gamma_{\mathcal{D}'}(\mathbb{S}M;\mathbb{S}_pM)$ which satisfies the equation

$$(\mathcal{D}-m)G_p = \delta_p\,. \tag{2.3.1}$$

Here $\delta_p \in \Gamma_{\mathcal{D}'}(\mathbb{S}M;\mathbb{S}_pM)$ is the Dirac delta distribution defined by $\langle\delta_p,\psi\rangle = \psi(p)$. While equation (2.3.1) has many solutions (one can for instance always add a solution of

[9] Note that the Dirac current precisely coincides with $\prec\cdot,\sigma_{\mathcal{D}}(\xi)\cdot\succ_{\mathbb{S}M}$ since the symbol of the Dirac operator is Clifford multiplication.

the homogeneous Dirac equation), it is a fundamental observation that there exist two particular solutions $G_p^{\pm}$ which are uniquely determined by the additional condition

$$\operatorname{supp} G_p^{\pm} \subset J^{\pm}(p)\,. \tag{2.3.2}$$

The Green's function G_p^+ is called *advanced Green's function at* p, and G_p^- is called *retarded Green's function at* p. They are basically constructed as follows (for technical details see [**Bär14**, Ch. 5] and [**BGP07**, Ch. 3]): Given $\varphi \in \Gamma_c^\infty(\mathbb{S}M)$, choose any $t_0 \in \mathbb{R}$ such that $\operatorname{supp}\varphi \subset I^+(\Sigma_{t_0})$, and let $\psi \in \Gamma^\infty(\mathbb{S}M)$ be the unique solution of $(\mathcal{D}-m)\psi = \varphi$ with $\psi|_{\Sigma_{t_0}} = 0$. Then one sets

$$\left\langle G_p^+, \varphi \right\rangle := \psi(p)\,. \tag{2.3.3}$$

Using uniqueness of solutions of the smooth Cauchy problem one can show that this is independent of the particular choice of t_0, and because of continuous dependence of the solution ψ on the inhomogeneity φ it follows that G_p^+ is indeed a distribution (i.e., is continuous). To see that G_p^+ satisfies (2.3.1), we first note that by the definition of how differential operators act on distributions we have

$$\left\langle (\mathcal{D}-m)G_p^+, \varphi \right\rangle = \left\langle G_p^+, (\mathcal{D}-m)^*\varphi \right\rangle = \left\langle G_p^+, (\mathcal{D}-m)\varphi \right\rangle \qquad \forall \varphi \in \Gamma_c^\infty(\mathbb{S}M)\,.$$

Here $(\mathcal{D}-m)^*$ is the formal adjoint of $\mathcal{D}-m$ with respect to the pairing $\int_M \prec \cdot,\cdot \succ_{\mathbb{S}M} \, d\mu_g$, which coincides with $\mathcal{D}-m$ by Remark 2.1.5. By definition of G_p^+ the right-hand side is the unique solution ψ of the Cauchy problem $(\mathcal{D}-m)\psi = (\mathcal{D}-m)\varphi$ which vanishes in the past of $\operatorname{supp}(\mathcal{D}-m)\varphi \subset \operatorname{supp}\varphi$. But obviously φ also has these properties, so it follows that $\psi = \varphi$, and thus

$$\left\langle (\mathcal{D}-m)G_p^+, \varphi \right\rangle = \varphi(p) = \langle \delta_p, \varphi \rangle \qquad \forall \varphi \in \Gamma_c^\infty(\mathbb{S}M)\,.$$

This shows (2.3.1). Notice furthermore that G_p^+ also satisfies

$$(\mathcal{D}_p - m)G_p^+ = \delta_p\,, \tag{2.3.4}$$

where the index p in $\mathcal{D}_p$ indicates that the Dirac operator acts on the p-variable.[10] This simply follows from the fact that

$$\left\langle (\mathcal{D}_p - m)G_p^+, \varphi \right\rangle = (\mathcal{D}_p - m)\left\langle G_p^+, \varphi \right\rangle = \varphi(p) = \langle \delta_p, \varphi \rangle\ ,$$

since $\left\langle G_p^+, \varphi \right\rangle$ is by construction a solution of $(\mathcal{D}-m)\left\langle G_p^+, \varphi \right\rangle = \varphi(p)$ (notice in particular that $\langle G_p, \varphi \rangle$ is always smooth in p). Finally, the support condition (2.3.2) can also be easily verified. For the construction of G_p^- one proceeds similarly.

By construction, if one can determine the Green's function explicitly, then they can be used to solve the inhomogeneous Dirac equation $(\mathcal{D}-m)\psi = \varphi$ by setting $\psi(p) := \left\langle G_p^{\pm}, \varphi \right\rangle$. In order to solve the *initial value problem* for the Dirac equation, which is what we are interested in, it turns out that their difference

$$K_p := G_p^+ - G_p^- \in \Gamma_{\mathcal{D}'}(\mathbb{S}M; \mathbb{S}_pM) \tag{2.3.5}$$

is very useful. It is called the *causal fundamental solution at* p of the massive Dirac equation, and it satisfies the *homogeneous* Dirac equation

$$(\mathcal{D}_q - m)K_p(q) = (\mathcal{D}_p - m)K_p(q) = 0\,, \tag{2.3.6}$$

[10] Formally one may view G^+ as a bidistribution $G_p^+(q) = G^+(p,q) \in \Gamma_{\mathcal{D}'}(\mathbb{S}M \boxtimes \mathbb{S}M)$ to understand the action of the Dirac operator on the p-variable.

where of course $K_p(q)$ is not really a function of q (and p), but rather a distribution. It also has the support property

$$\operatorname{supp} K_p \subset J(p) = J^+(p) \cup J^-(p)\,. \tag{2.3.7}$$

Now suppose that $\psi \in \Gamma^\infty_{sc}(\mathbb{S}M)$ is the solution of the Cauchy problem

$$\begin{cases} (\mathcal{D} - m)\psi = 0 \\ \quad \psi|_{t=t_0} = \psi_0 \end{cases}$$

for some $\psi_0 \in \Gamma^\infty_c(\mathbb{S}M|_{\Sigma_{t_0}})$. We claim that ψ can be written as

$$\boxed{\psi(p) = \frac{1}{\mathrm{i}} \int_{\Sigma_{t_0}} K_p(t_0, y)\gamma(\nu_{\Sigma_{t_0}})\psi_0(y)\,\mathrm{d}\mu_{\Sigma_{t_0}}(y)}\,. \tag{2.3.8}$$

Clearly the right-hand side needs some explanation. But first, notice that at least heuristically it solves the homogeneous Dirac equation since K_p does (as "function" of p). Therefore, as long as it is well-defined and happens to coincide with ψ_0 on Σ_{t_0}, by uniqueness of solutions of the Cauchy problem it must with ψ.

Let us now explain in which sense the right-hand side can be understood. First of all, the integral should of course be understood as the action of the distribution $K_p(t_0, \cdot)$ on the test function $\gamma(\nu_{\Sigma_{t_0}})\psi_0 \in \Gamma^\infty_c(\mathbb{S}M|_{\Sigma_{t_0}})$. Secondly, and this requires more attention, we have to justify why one can "restrict" the distribution K_p (which is a distribution on M) to the hypersurface Σ_{t_0}. Or in other words, we have to justify that one can "localize K_p in time". Before doing so, let us stress that in concrete situations, i.e. where K_p is explicitly known, it should simply be clear how the right-hand side of (2.3.8) is to be understood without the need for any of the abstract arguments given in the following.

Nevertheless, let us now give some general arguments. The key observation is that since K_p satisfies the *homogeneous* Dirac equation, it follows that it is in fact "smooth in t" in the following sense:[11] For fixed $\psi \in \Gamma^\infty_{sc}(\mathbb{S}M)$ we can define a distribution on $\mathbb{R}_t$ with values in $\mathbb{S}_pM$ by

$$C^\infty_c(\mathbb{R}_t) \ni \eta \longmapsto \langle K_p, \eta\psi\rangle \in \mathbb{S}_pM\,. \tag{2.3.9}$$

The above mentioned "smoothness of K_p in t" now refers to the claim that this distribution is actually given by a smooth function. That is, we claim that there exists a smooth function on $\mathbb{R}_t$, which we suggestively denote by $\langle K_p(t,\cdot), \psi(t,\cdot)\rangle$, such that

$$\langle K_p, \eta\psi\rangle = \int_{\mathbb{R}} \eta(t)\,\langle K_p(t,\cdot), \psi(t,\cdot)\rangle\,\mathrm{d}t\,. \tag{2.3.10}$$

We now argue why this is true. For brevity of notation, let us denote the distribution (2.3.9) by F. We will argue that F is a distribution of order zero and is thus given by the action of a (signed) measure, and that the same is true for $\partial_t^k F$ for any $k \in \mathbb{N}$. This then implies that F is in fact given by a smooth function.

To show that F has order zero, let $[-T, T] \subset \mathbb{R}$ be a compact interval, and set

$$A := ([-T, T] \times \Sigma) \cap \operatorname{supp}\psi \subset M\,.$$

Then A is compact, and using what it means for K_p to be a distribution (i.e. spelling out what continuity means for the map $K_p : \Gamma^\infty(\mathbb{S}M) \to \mathbb{S}_pM$), there exist $m \in \mathbb{N}$ and $C > 0$ (depending on T) such that

$$|\,\langle K_p, \varphi\rangle\,| \le C\,\|\varphi\|_{C^m(\mathbb{S}M)} \qquad \forall \varphi \in \Gamma^\infty_c(\mathbb{S}M)\,, \operatorname{supp}\varphi \subset A\,.$$

---

[11]The following argument is taken from [**Lax06**, Lemma 7.33], where one also finds more details.

For simplicity of the argument let us assume that $m = 1$, how the general case can be treated will be clear after we have given the argument. Let $\eta \in C_c^\infty(\mathbb{R})$ be given with $\operatorname{supp}\eta \subset [-T,T]$. Again for simplicity, let us assume that $\eta = \chi'$ for some $\chi \in C_c^\infty(\mathbb{R})$ with $\operatorname{supp}\chi \subset A$, the reduction of the general case to this is explained in [**Lax06**, p. 70]. Under these simplifying assumptions, we now have

$$\begin{aligned}
|\langle F,\eta\rangle| &= |\langle K_p, \eta\psi\rangle| \\
&= |\langle K_p, (\partial_t\chi)\psi\rangle| \\
&= |\langle K_p, \partial_t(\chi\psi) - \chi\partial_t\psi\rangle| \\
&\leq |\langle \partial_t K_p, \chi\psi\rangle| + |\langle K_p, \chi\partial_t\psi\rangle| \\
&\overset{(*)}{=} |\langle A(t,x,\partial_x)K_p, \chi\psi\rangle| + |\langle K_p, \chi\partial_t\psi\rangle| \\
&= |\langle K_p, \chi A(t,x,\partial_x)^*\psi\rangle| + |\langle K_p, \chi\partial_t\psi\rangle| \\
&\leq C\big(\|\chi A(t,x,\partial_x)\psi\|_{C^1(\mathbb{S}M)} + \|\chi\partial_t\psi\|_{C^1(\mathbb{S}M)} \big) \\
&\leq C' \|\chi\|_{C^1(-T,T)} \\
&\leq C' \|\eta\|_{C^0(-T,T)} .
\end{aligned}$$

In $(*)$ we used that K_p satisfies the homogeneous Dirac equation $(\mathcal{D}-m)K_p = 0$, which can be rewritten as

$$\partial_t K_p = A(t,x,\partial_x)K_p\,.$$

Here on the right-hand side $A(t,x,\partial_x)$ stands for a first-order differential operator which only differentiates into Σ-directions.

This estimate shows that, restricted to $C_c^\infty(-T,T)$, the distribution F has order zero. It implies that one can continuously extend F to all of $C_c(-T,T)$, so that by the Riesz representation theorem for the dual of $C_c(-T,T)$ this restriction of F is given by a (signed) measure. Concerning the simplifying assumption $m = 1$, one can basically repeat the previous computation also in the case of general $m \in \mathbb{N}$, the only difference being that then one has to move more than one t-derivative from χ onto ψ and F. The key is always that one can convert any t-derivative of K_p into spatial derivatives using the homogeneous Dirac equation. By the same argument one also shows that $\partial_t^k F$, restricted to $C_c^\infty(-T,T)$, has order zero for any $k \in \mathbb{N}$. This then shows that restricted to $C_c^\infty(-T,T)$, the distribution F is given by a smooth function. Finally one lets T tend to infinity to see that actually F is given by a smooth function on all of $C_c^\infty(\mathbb{R})$. Of course, the constants in the estimates before may blow up, but this simply reflects the fact that the smooth function representing F need not be integrable over all of $\mathbb{R}$. This completes our rough argument for the validity of (2.3.10).

Let us now turn back to the representation formula (2.3.8) which we are trying to prove. So far we have explained that the right-hand side should be understood as

$$\frac{1}{\mathrm{i}}\int_{\Sigma_{t_0}} K_p(t_0,y)\gamma(\nu_{\Sigma_{t_0}})\psi_0(y)\,\mathrm{d}\mu_{\Sigma_0}(y) = \frac{1}{\mathrm{i}}\,\langle K_p(t_0,\cdot), \gamma(\nu\Sigma_{t_0})\psi(t_0,\cdot)\rangle \tag{2.3.11}$$

It remains to argue why this actually coincides with $\psi(p)$. To this end, let $p = (t,x) \in M$ be given. Suppose that $p \in I^+(\Sigma_{t_0})$. Since K_p coincides with $-G_p^-$ in $I^-(p)$, we have to

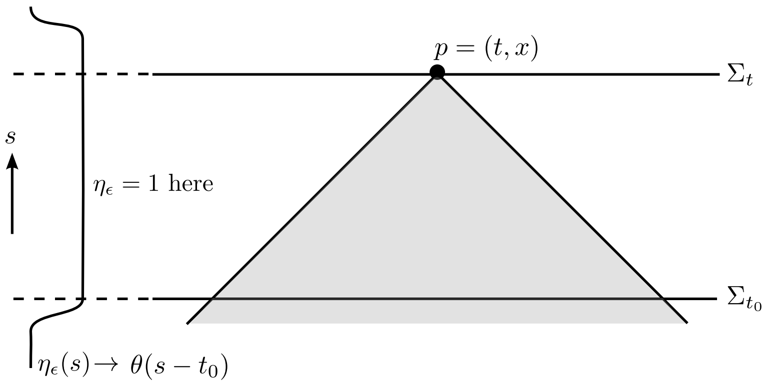

FIGURE 2.8. Illustration for the verification of the representation formula (2.3.8).

show that[12]

$$\psi(p) = -\frac{1}{\mathrm{i}} \left\langle G_p^-(t_0, \cdot), \gamma(\nu_{\Sigma_{t_0}})\psi(t_0, \cdot) \right\rangle = \mathrm{i} \left\langle G_p^-(t_0, \cdot), \gamma(\nu_{\Sigma_{t_0}})\psi(t_0, \cdot) \right\rangle .$$

Recall that by (2.3.4) we have

$$\varphi(p) = \left\langle G_p^-, (\mathcal{D}-m)\varphi \right\rangle \qquad \forall \varphi \in \Gamma_c^\infty(\mathbb{S}M) .$$

To use this, we take a cutoff $\eta \in C_c^\infty(\mathbb{R}_t)$ with $\eta(s) = 1$ for all $s \in [t_0, t]$, where we recall that $p = (t, x)$. See figure 2.8 for an illustration. Then we have

$$\psi(p) = (\eta\psi)(p) = \left\langle G_p^-, (\mathcal{D}-m)(\eta\psi) \right\rangle = \left\langle G_p^-, \mathrm{i}\gamma^0 \eta' \psi \right\rangle \overset{(*)}{=} \int_{\mathbb{R}} \eta'(s) \left\langle G_p^-(s, \cdot), \mathrm{i}\gamma^0 \psi(s, \cdot) \right\rangle \mathrm{d}s .$$

In $(*)$ we used that $\eta'(s) = 0$ for all s close to t, and that G_p^- satisfies the *homogeneous* Dirac equation outside of any open neighborhood of Σ_t. Therefore the integral formula is justified by the same arguments used before for K_p. Next, we choose a sequence of such cutoff functions $\eta_\epsilon \in C_c^\infty(\mathbb{R}_t)$ which converge to the Heaviside function at t_0 in $\mathcal{D}'(\mathbb{R}_t)$. Then $\eta'_\epsilon \to \delta_{t_0}$ in $\mathcal{D}'(\mathbb{R}_t)$ and therefore

$$\psi(p) \overset{\forall \epsilon > 0}{=} \int_{\mathbb{R}} \eta'_\epsilon(s) \left\langle G_p^-(s, \cdot), \mathrm{i}\gamma^0 \psi(s, \cdot) \right\rangle \longrightarrow \left\langle G_p^-(t_0, \cdot), \mathrm{i}\gamma^0 \psi(t_0, \cdot) \right\rangle .$$

This is precisely what we wanted to show. For $p \in I^-(\Sigma_{t_0})$ one can proceed similarly using G_p^+ instead of G_p^-. By continuity it follows that (2.3.8) also holds true for $p \in \Sigma_{t_0}$ simply because both sides are smooth.

If in the representation formula (2.3.8) one takes $p = (t_0, y)$ to be a point lying on the initial hypersurface Σ_{t_0}, then it follows that $K_{(t_0,x)}(t_0, y) = \mathrm{i}\gamma(\nu_{\Sigma_{t_0}})\delta_x(y)$. Using well-posedness of the Cauchy problem for distributional initial values (which can be shown using the arguments in [**Lax06**, Sec. 7.1]), this shows that one can also characterize the causal fundamental solution K_p for $p = (t, x)$ as the unique solution of the initial value problem

$$\begin{cases} (\mathcal{D}_q - m)K_{t,x}(q) = 0 \\ \qquad K_{t,x}(t, y) = \mathrm{i}\gamma(\nu_{\Sigma_{t_0}})\delta_x(y) \end{cases} . \tag{2.3.12}$$

[12]For clarity, note that G_p^- satisfies the *homogeneous* Dirac equation in $I^-(\Sigma_t)$, and therefore the right-hand side makes sense by the same arguments used for K_p before.

Except for the factor $\mathrm{i}\gamma(\nu_{\Sigma_{t_0}})$ in the initial values, this solution corresponds to what is sometimes called the *Riemann function* of the equation (cf. [**Lax06**, Def. 7.7]). So we see that in this context, the Riemann function is (up to a different "normalization") the same as what here is called causal fundamental solution.

Let us illustrate this abstract presentation by a concrete example.

Example 2.3.1. With the choice of γ-matrices as in Example 1.2.18, the Dirac operator in 1+1 dimensions has the form

$$\mathcal{D}_{\mathbb{R}^{1,1}} = \frac{1}{\mathrm{i}} \begin{pmatrix} 0 & \partial_t - \partial_x \\ \partial_t + \partial_x & 0 \end{pmatrix} ,$$

acting on spinorial wave functions $\psi : \mathbb{R}^{1,1} \to \mathbb{C}^2$. Using the characterization of the causal fundamental solution by the initial value problem (2.3.12), one may verify that for the *massless* Dirac equation (i.e. for $m = 0$), the causal fundamental solution is explicitly given by

$$K_{(t,x)}(s, y) = \frac{1}{\mathrm{i}} \begin{pmatrix} 0 & \delta\big((x-y) - (t-s)\big) \\ \delta\big((x-y) + (t-s)\big) & 0 \end{pmatrix} .$$

Notice that this may indeed be viewed as a distribution on $\mathbb{R}_y$ alone, which depends smoothly on the "parameters" t, x, s. The representation formula (2.3.8) for the solution ψ of the Cauchy problem with initial data $\psi_{t=0} = \psi_0 = (\psi_0^1, \psi_0^2)$ now becomes

$$\begin{aligned} \psi(t,x) &= \mathrm{i} \int_{\mathbb{R}} \frac{1}{\mathrm{i}} \begin{pmatrix} 0 & \delta((x-y) - t) \\ \delta((x-y)+t) & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_0^1(y) \\ \psi_0^2(y) \end{pmatrix} \mathrm{d}y \\ &= \int_{\mathbb{R}} \begin{pmatrix} \delta((x-y)-t) & 0 \\ 0 & \delta((x-y)+t) \end{pmatrix} \begin{pmatrix} \psi_0^1(y) \\ \psi_0^2(y) \end{pmatrix} \mathrm{d}y \\ &= \begin{pmatrix} \psi_0^1(x-t) \\ \psi_0^2(x+t) \end{pmatrix} . \end{aligned}$$

Noticing that the massless Dirac equation in 1+1 dimensions is just a pair of uncoupled transport equations, this formula indeed gives the correct solution.

In higher dimensions, computations with the Dirac equation are often complicated because spinors have increasingly many components. At least in the flat case it is therefore often helpful to use the Lichnerowicz formula (1.2.24), which here is simply

$$\mathcal{D}^2_{\mathbb{R}^{1,n}} = \square_{\mathbb{R}^{1,n}} = -\partial_t^2 + \Delta_{\mathbb{R}^n} \; . \tag{$*$}$$

Here the right-hand side acts componentwise, which is the reason why this formula simplifies many explicit computations (it "decouples" the components of the spinor field).[13] To illustrate the usefulness of $(*)$ regarding the computation of Green's functions and the causal fundamental solution, suppose we have already computed these for the Klein-Gordon equation. Denoting the Green's function for the Klein-Gordon equation by $S_p^\pm$, so that

$$(\square_{\mathbb{R}^{1,n}} - m^2) S_p^\pm = \delta_p \quad \text{and} \quad \operatorname{supp} S_p^\pm \subset J^\pm(p) \, ,$$

we set

$$G_p^\pm := (\mathcal{D} + m) S_p^\pm \, . \tag{2.3.13}$$

Then by $(*)$ it holds

$$(\mathcal{D}_{\mathbb{R}^{1,n}} - m) G_p^\pm = (\mathcal{D}_{\mathbb{R}^{1,n}} - m)(\mathcal{D}_{\mathbb{R}^{1,n}} + m) S_p^\pm = (\square_{\mathbb{R}^{1,n}} - m^2) S_p^\pm = \delta_p \, .$$

[13] Of course one can also use the Lichnerowicz formula in arbitrary spacetimes. But then $\mathcal{D}^2 = \nabla^\mu \nabla_\mu + \frac{\mathrm{scal}}{4}$ typically no longer acts diagonally, but still couples the different components of a spinor.

Since moreover $\operatorname{supp} G_p^\pm \subset \operatorname{supp} S_p^\pm \subset J^\pm(p)$, it follows by uniqueness of advanced and retarded Green's functions that $G_p^\pm$ must in fact be the advanced and retarded Green's functions for $\mathcal{D}_{\mathbb{R}^{1,n}} - m$. One can test (2.3.13) in 1+1 dimensions by computing explicitly the Green's function for the wave equation (i.e. for $m = 0$).

2.3.2. The geometric approach to energy estimates. It is arguable that just as analysis in general is the "art of making estimates", the analysis of (hyperbolic) evolution equations is the art of making *energy estimates*, i.e. estimates of various (L^2-based) norms of a solution $u(t,\cdot)$ in terms of the same norms of the initial data (and other quantities). This was already pointed out in Section 2.2 during the discussion of symmetric hyperbolic systems. From this perspective, given an equation, the first question is which estimates are in a sense natural for the equation one studies: Often the specific structure of the equation makes certain estimates more favorable, such as for instance current conservation for the Dirac equation.

For geometric equations on Lorentzian manifolds there is one key tool which one can use as guideline in the attempt to produce (good) energy estimates: *Stokes' or Gauß' theorem*. To this end, let (M,g) be an oriented Lorentzian manifold with volume form Ω_g. If $U \subset M$ is a bounded open subset with smooth boundary $\partial U \subset M$, then for any $X \in \Gamma^\infty(TM)$ we have[14]

$$\int_U (\mathrm{div} X)\Omega_g = \int_U \mathscr{L}_X \,\Omega_g = \int_U \big(\iota_X \underbrace{\mathrm{d}\Omega_g}_{=0} + \mathrm{d}(\iota_X\Omega_g)\big) \overset{\text{Stokes}}{=} \int_{\partial U} \iota_X \Omega_g \,. \tag{2.3.14}$$

Here we used the relation of the divergence to the Lie derivative of the volume form Ω_g (cf. [**Lee03**, p. 479]), and Cartan's "magic formula" for the Lie derivative (cf. [**Lee03**, Prop. 18.13]). If $\partial U \subset M$ is nondegenerate and has an outward-pointing unit-normal $\nu_{\partial U} \in \Gamma^\infty(TM|_{\partial U})$, we can decompose X along ∂U as

$$X|_{\partial U} = \langle \nu_{\partial U}, \nu_{\partial U}\rangle \,\langle X, \nu_{\partial U}\rangle\, \nu_{\partial U} + X^{\partial U}\,,$$

where $X^{\partial U}$ is the part of X that is tangential to ∂U. The factor $\langle \nu_{\partial U}, \nu_{\partial U}\rangle$ is always either $+1$ or -1, and takes into account the causal character of ∂U (it is of course constant on each component of ∂U). Since $(\iota_{X^{\partial U}}\Omega_g)|_{T\partial U} = 0$ (by antisymmetry of forms), it follows that in this case (2.3.14) reduces to the well-known *Gauß' theorem*

$$\int_U \mathrm{div}\, X \,\mathrm{d}\mu_g = \int_{\partial U} \langle \nu_{\partial U}, \nu_{\partial U}\rangle \,\langle X, \nu_{\partial U}\rangle \,\mathrm{d}\mu_{\partial U}\,, \tag{2.3.15}$$

where we have gone back to the notation $\mathrm{d}\mu_g$ for the volume measure of g, and where $\mathrm{d}\mu_{\partial U} = \iota_{\nu_{\partial U}}\Omega_g$ is the induced volume form on ∂U.

Before we explain how this can be used for energy estimates, let us make some remarks which are important for many practical applications:

(1) The requirement that the boundary $\partial U \subset M$ has to be smooth can be substantially relaxed (C^k-regularity and less, piecewise smooth, corner points etc.). This is covered in quite some detail in [**Sau06**, Ch. 1].
(2) One can also allow unbounded domains $U \subset M$, i.e. domains not contained in any compact subset. In this case one has to assume suitable integrability conditions of the vector field X, for instance that $\mathrm{div}\, X$ is integrable over U and that X decays sufficiently fast at infinity in order that potential "boundary

[14] Here we use integration of differential forms by writing Ω_g instead of using the notation $\mathrm{d}\mu_g$ for the volume measure as elsewhere in this thesis. This fits better with Stokes' theorem.

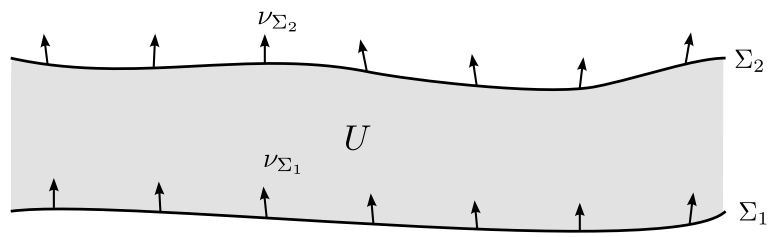

FIGURE 2.9. A domain U bounded by two Cauchy surfaces Σ_1 and Σ_2.

terms at infinity" vanish. Clearly, assuming that $(\operatorname{supp} X) \cap U$ is compact suffices to exclude all possible contributions at infinity.

(3) Also in the case of lightlike boundaries (i.e. with degenerate induced metric) one can convert (2.3.14) into a version similar to (2.3.15). However, since a normal vector to a lightlike submanifold is always tangential to this submanifold, this requires the choice of an additional perpendicular vector field to the boundary. For a simple case, this is carried out in Lemma 4.4.4.

(4) Whereas the assumption of orientability is necessary for Stokes' theorem, the Gauß' theorem version (2.3.15) continues to hold also if (M, g) is not orientable, see [**Lee03**, Thm. 14.34].

To see how the identity (2.3.15) might be useful for energy estimates, take a domain $U \subset M$ as in figure 2.9 whose boundary ∂U is the disjoint union of two Cauchy surfaces Σ_1 and Σ_2, say with $\Sigma_2 \subset I^+(\Sigma_1)$. Then for any vector field $X \in \Gamma^\infty(TM)$ which satisfies suitable integrability condition (for instance has spatially compact support), it follows from (2.3.15) that

$$\int_{\Sigma_2} \langle X, \nu_{\Sigma_2} \rangle \, \mathrm{d}\mu_{\Sigma_2} = \int_{\Sigma_1} \langle X, \nu_{\Sigma_1} \rangle \, \mathrm{d}\mu_{\Sigma_1} + \int_U \operatorname{div} X \, \mathrm{d}\mu_g \,. \tag{2.3.16}$$

As a special case, suppose that $M \cong \mathbb{R}_t \times \Sigma$ is a foliation of M by spacelike Cauchy hypersurfaces $\Sigma_t = \{t\} \times \Sigma$ with future-pointing unit normal $\nu \in \Gamma^\infty(TM)$, i.e. such that $\nu|_{\Sigma_t}$ is the future-pointing unit normal to Σ_t for any $t \in \mathbb{R}$. Then (2.3.16) describes how the quantity

$$\mathbf{E}_X(t) := \int_{\Sigma_t} \langle X, \nu \rangle \, \mathrm{d}\mu_{\Sigma_t} \tag{2.3.17}$$

changes in time. In particular, if $\operatorname{div} X = 0$, then (2.3.17) is *conserved* in time, i.e. independent of t.[15]

Suppose now that we want to derive estimates for solutions of an equation $P(u) = w$, where u is a section of some bundle E over the spacetime (M, g). The basic idea is then to construct vector fields $X[u] \in \Gamma^\infty(TM)$, depending on u and possibly other geometric objects on M, and apply the previous machinery to these vector fields. For any such vector field we then have the identity

$$\int_{\Sigma_2} \langle X[u], \nu_{\Sigma_2} \rangle \, \mathrm{d}\mu_{\Sigma_2} = \int_{\Sigma_1} \langle X[u], \nu_{\Sigma_1} \rangle \, \mathrm{d}\mu_{\Sigma_1} + \int_U \operatorname{div} X[u] \, \mathrm{d}\mu_g \,. \tag{2.3.18}$$

In most applications, in order to use this identity to obtain information about u, the vector field $X[u]$ should satisfy the following two basic criteria:

[15] Actually, if $\operatorname{div} X = 0$ then $\int_\Sigma \langle X, \nu_\Sigma \rangle \, \mathrm{d}\mu_\Sigma$ gives the same result for *any* choice of Cauchy surface $\Sigma \subset M$. To see this one can use the same argument as in Corollary 2.1.4.

(1) For a Cauchy surface $\Sigma \subset M$ with future-directed unit normal ν_Σ, the quantity $\langle X[u], \nu_\Sigma \rangle$ that one integrates over Σ has of course to be usable for some concrete purpose. Often one tries to use it to bound some integral norm $\|\cdot\|_\Sigma$ on Σ, i.e. one wants to have an estimate of the form

$$\|u\|_\Sigma^2 \leq C \int_\Sigma \langle X[u], \nu_\Sigma \rangle \, \mathrm{d}\mu_\Sigma \qquad \forall u \in \Gamma_c^\infty(E|_\Sigma)\,. \tag{2.3.19}$$

Notice that for this to work, a necessary condition is that $\langle X[u], \nu_\Sigma \rangle$ is *positive*.

(2) The divergence of $X[u]$ has to be controllable in some suitable sense in order to get some useful information about the "time evolution" of the quantity (2.3.17), i.e. how it changes between two Cauchy surfaces (for instance along a foliation of M by Cauchy surfaces). In the simplest case where $\operatorname{div} X[u] = 0$, one obtains a conservation law, i.e. the quantity of interest does not change at all.

To outline what can be done in the case where the divergence does not vanish, suppose that $M \cong \mathbb{R} \times \Sigma$ is a foliation by Cauchy surfaces with unit normal $\nu \in \Gamma^\infty(TM)$. A common idea is then to try to estimate $\operatorname{div} X[u]$ in terms of $\langle X[u], \nu \rangle$. If this can be done, Grönwall's inequality gives at least an (exponential) bound on the growth of (2.3.17). At first sight it might seem unreasonable that one might be able to estimate $\operatorname{div} X[u]$ in terms of $\langle X[u], \nu \rangle$, since $\operatorname{div} X[u]$ usually contains derivatives of u which are one order higher than those contained in $X[u]$. However, this is precisely the point where one can try to use that u satisfies the differential equation $P(u) = w$. Namely, making a clever choice for $X[u]$ in the first place one can try to arrange that all higher order derivatives of u which are contained in $\operatorname{div} X[u]$ (as compared to those contained in $X[u]$) appear in the form $P(u)$. Then these actually drop out, i.e. they can be replaced by the inhomogeneity w.

Of course, these two conditions severely restrict the construction of vector fields $X[u]$ from which one can successfully extract information about u in this way. From (2) it is also clear that the particular choices one has (if any at all) depend strongly on the structure of the equation one is studying. Fortunately, many equations from geometry and physics possess suitable vector (or tensor) fields. For instance, whenever an equation is the Euler-Lagrange equation of some functional, the *Noether theorem* can be used to systematically construct such quantities.

A concrete example was already given in Section 2.1, where the *conserved current* of the Dirac equation was discussed. The current vector field $J[\psi]$ satisfies both properties (1) and (2): It is divergence-free (if ψ satisfies the Dirac equation), and for any timelike vector ν one has $\langle J[\psi], \nu \rangle \geq 0$. The Dirac current controls the L^2-norm of a spinor field over a Cauchy surface. More generally, the same is true for symmetric hyperbolic systems on Lorentzian manifold as discussed in [**Bär14**, Sec. 5], cf. Remark 2.2.9, for which one has a similar quantity as the Dirac current. The corresponding current will in general not be conserved, but can always be estimated by a Grönwall argument as indicated above.

To have more flexibility, instead of constructing only vector fields $X[u] \in \Gamma^\infty(TM)$ one can also construct higher order tensor fields $T[u] \in \Gamma^\infty(T^*M^{\otimes r} \otimes TM^{\otimes s})$. For equations of higher order this can be more natural (consider for instance the energy-momentum tensor for the wave equation in Example 2.3.3 below). In order to use Gauß' theorem (2.3.15) for such a higher order tensor field $T[u]$, the idea is to first contract all but one of its indices ("slots") with other fields to get a vector field (or 1-form). For instance, set

$$T[u]^{\omega_1,\dots,\omega_{s-1}}_{X_1,\dots,X_r} := T[u](X_1,\dots,X_r,\omega_1,\dots,\omega_{s-1},\cdot) \in \Gamma^\infty(TM)$$

for some (cleverly chosen) $X_1,\dots,X_r \in \Gamma^\infty(TM)$ and $\omega_1,\dots,\omega_{s-1} \in \Gamma^\infty(T^*M)$. To the resulting vector field one can then apply Gauß' theorem to get identities of the form

$$\int_{\Sigma_2} \big\langle T[u]^{\omega_1,\dots,\omega_{s-1}}_{X_1,\dots,X_r}, \nu_{\Sigma_2} \big\rangle \,\mathrm{d}\mu_{\Sigma_2} = \int_{\Sigma_1} \big\langle T[u]^{\omega_1,\dots,\omega_{s-1}}_{X_1,\dots,X_r}, \nu_{\Sigma_1} \big\rangle \,\mathrm{d}\mu_{\Sigma_1} + \int_U \mathrm{div}\big(T[u]^{\omega_1,\dots,\omega_{s-1}}_{X_1,\dots,X_r}\big) \,\mathrm{d}\mu_g\,,$$

where as before U, Σ_1, Σ_2 are as sketched in figure 2.9.

Compared to vector field quantities such as the current of a symmetric hyperbolic system, using higher order tensor fields is more complicated (but also offers more flexibility) at least for the following two reasons:

(1) It is more difficult to check whether the quantity $\big\langle T[u]^{\omega_1,\dots,\omega_{s-1}}_{X_1,\dots,X_r}, \nu \big\rangle$ is useful at all (for instance controls some norm of u). Usually this will depend on the choice of the contracting fields $X_1,\dots,X_r,\omega_1,\dots,\omega_{s-1}$. On the other hand, the freedom to choose these contracting fields also offers more possibilities to arrange interesting quantities in this way.
(2) The "error term" $\mathrm{div}\, T[u]^{\omega_1,\dots,\omega_{s-1}}_{X_1,\dots,X_r}$ also contains derivatives of the contracting fields $X_1,\dots,X_r,\omega_1,\dots,\omega_{s-1}$. Therefore one either has to choose these fields in a clever way such that their derivatives simply drop out, or one has to be able to control these derivatives in some way.

Concerning the second point, if the tensor field is *symmetric* in all its indices (suppose it only has lower indices), then its divergence contains the derivatives of the contracting vector fields only in the form of Lie derivatives of the metric:

Lemma 2.3.2. *Let $T \in \Gamma^\infty(\mathrm{S}^{r+1}T^*M)$ be a symmetric tensor field, i.e. $T_{\mu_1\cdots\mu_r} = T_{(\mu_1\cdots\mu_r)}$ in abstract index notation. Then for any $X_1,\dots,X_r \in \Gamma^\infty(TM)$, we have*

$$\begin{aligned}\nabla^\mu(T_{\mu\nu_1\cdots\nu_r}X_1^{\nu_1}\cdots X_r^{\nu_r}) &= (\nabla^\mu T_{\mu\nu_2\cdots\nu_r})X_1^{\nu_1}\cdots X_r^{\nu_r} \\ &\quad + \frac{1}{2}\sum_{j=1}^r T_{\mu\nu_2\cdots\nu_r}X_1^{\nu_1}\cdots(\mathscr{L}_{X_j}g)^{\mu\nu_j}\cdots X_r^{\nu_r}\,.\end{aligned} \tag{2.3.20}$$

Proof. We compute in abstract index notation. First, by the Leibniz rule, we have

$$\begin{aligned}\nabla^\mu(T_{\mu\nu_1\cdots\nu_r}X_1^{\nu_1}\cdots X_r^{\nu_r}) &= (\nabla^\mu T_{\mu\nu_2\cdots\nu_r})X_1^{\nu_1}\cdots X_r^{\nu_r} \\ &\quad + \sum_{j=1}^r T_{\mu\nu_2\cdots\nu_r}X_1^{\nu_1}\cdots(\nabla^\mu X_j^{\nu_j})\cdots X_r^{\nu_r}\,.\end{aligned}$$

Next, since T is symmetric in all of its indices, for any $j = 1,\dots,r$ we can write

$$T_{\mu\nu_2\cdots\nu_r}X_1^{\nu_1}\cdots(\nabla^\mu X_j^{\nu_j})\cdots X_r^{\nu_r} = \frac{1}{2}T_{\mu\nu_2\cdots\nu_r}X_1^{\nu_1}\cdots(\nabla^\mu X_j^{\nu_j} + \nabla^{\nu_j}X_j^\mu)\cdots X_r^{\nu_r}\,.$$

Finally, using the Koszul formula for the Lie derivative one easily verifies that the symmetrized covariant derivative of X_j produces precisely the Lie derivative of g by X_j, i.e. $\nabla^\mu X_j^{\nu_j} + \nabla^{\nu_j}X_j^\mu = (\mathscr{L}_{X_j}g)^{\mu\nu_j}$. This concludes the proof. □

This computation shows that if one wants to use such higher order tensors to derive energy estimates it is useful to have *symmetries*, i.e. Killing fields X of the metric for which $\mathscr{L}_X\, g = 0$.

2.3.3. The energy-momentum tensor formalism. The energy-momentum tensor is a concrete example of a higher order tensor field of the type discussed at the end of the previous sections. It is available for any equation which is the Euler-Lagrange equation of some action functional on spacetime of the form discussed below. In the following we briefly explain the general idea behind the construction of the energy-momentum tensor, and then apply this idea to the Dirac equation.

The following outline (which follows [**Str04**, Sec. 2.3.3]) is more a heuristic argument involving some formal computations, which however in most concrete examples are easy to make precise. Making these computations precise in a general setting, on the other hand, requires the use of jet bundles (see for instance [**GMS97**, Sec. 3.5] or [**DF99**, Sec. 2.9]), which can easily obscure the simple basic idea. Here the aim is to explain the basic idea using (hopefully) just enough formality to keep the exposition simple but also making it possible to transfer the construction rigorously to specific examples. To follow the presentation it is certainly of advantage (although not strictly necessary) to have some acquaintance with Lagrangian field theory. Physics-inclined presentations may be found in [**HE73**, Sec. 3.3] and [**Str04**, Sec. 2.3].

Suppose now that we want to study an equation $P(u) = 0$ for sections of a vector bundle, of which we assume that it is the Euler-Lagrange equation (i.e., the equation of stationarity) of a functional

$$S[u] = \int_M \mathcal{L}(F(u,g),g)\,\Omega_g\,. \tag{2.3.21}$$

That is, we assume that $P(u) = 0$ is precisely the condition for u to satisfy

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} S[u+\epsilon v] = 0$$

for all (reasonable) v. Here the function $\mathcal{L}(F(u,g),g)$ should be thought of as some "expression" in u, derivatives of u, and g (see below for concrete examples). It should be local in the sense that for any point $p \in M$, the value $\mathcal{L}(F(u,g),g)(p)$ should only depend on u, derivatives of u, and g at the point p. Further, both $\mathcal{L}$ and F should be differentiable in a suitable sense.[16] Finally, the reason for writing (2.3.21) in the form $\mathcal{L}(F(u,g),g)$ is that this whole expression may depend on g in two slightly different ways: on the one hand it can depend on g "explicitly" in the sense that for instance certain contractions are being made. This refers to the explicit g-dependence of $\mathcal{L}$. On the other hand, g can enter in subtle "implicit" ways into $F(u,g)$ for instance if the Levi-Civita connection of g, or the Hodge-star operator of g appear (which depend on g in a subtle way). Of course this distinction is somewhat artificial, its purpose is solely to make one aware of the fact that the metric g could be "hidden" in $\mathcal{L}$ in more subtle ways than one might at first think (the dependence on g is crucial for what comes next).

The energy-momentum tensor arises from the "diffeomorphism invariance" or "geometric invariance" of the action. Namely, if $\phi : M \to M$ is a diffeomorphism then it follows from the transformation formula for integrals and $\phi(M) = M$ that

$$\int_M \phi^*(\mathcal{L}(F(u,g),g)\,\Omega_g) = \int_M \mathcal{L}(F(u,g),g)\,\Omega_g\,.$$

[16] These conditions are basically what one can formalize using jet bundles.

Suppose next that ϕ_ϵ is a 1-parameter family of diffeomorphisms, and assume moreover that[17]

$$\phi_\epsilon^*(\mathcal{L}(F(u,g),g)\,\Omega_g) = \mathcal{L}(F(\phi_\epsilon^* u, \phi_\epsilon^* g), \phi_\epsilon^* g)\Omega_{\phi_\epsilon^* g}\,.$$

In this case, by the Leibniz rule we have symbolically (one can of course write this out)

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_M \mathcal{L}(F(\phi_\epsilon^* u, \phi_\epsilon^* g), \phi_\epsilon^* g)\Omega_{\phi_\epsilon^* g} \\ &= \int_M \text{variation of } u \text{ by } \phi_\epsilon + \int_M \text{variation of } g \text{ by } \phi_\epsilon \end{aligned}$$

The first integral simply reproduces the Euler-Lagrange equations of the action functional S, and therefore vanishes if u satisfies these. The second integral contributes as

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_M \mathcal{L}(F(u, \phi_\epsilon^* g), \phi_\epsilon^* g) &= \int_M \left\{ \frac{\partial \mathcal{L}}{\partial F}\frac{\partial F}{\partial g_{\mu\nu}}(\mathscr{L}_X g)_{\mu\nu} + \frac{\partial \mathcal{L}}{\partial g_{\mu\nu}}(\mathscr{L}_X g)_{\mu\nu} \right\}\Omega_g \\ &\quad + \int_M \mathcal{L}\frac{\partial \Omega_g}{\partial g_{\mu\nu}}(\mathscr{L}_X g)_{\mu\nu}\,. \end{aligned}$$

Here X denotes the vector field which generates the 1-parameter group of diffeomorphisms ϕ_ϵ, so that

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} (\phi_\epsilon^* g)_{\mu\nu} = (\mathscr{L}_X g)_{\mu\nu}\,.$$

The variation of the volume form Ω_g again has to be a multiple of the volume form itself simply because the bundle of exterior forms of highest degree is a line bundle (i.e., has rank one). Using the local formula $\Omega_g = \sqrt{|g|}\,\mathrm{d}x^0 \wedge \cdots \wedge x^n$ one can show that in fact

$$\frac{\partial \Omega_g}{\partial g_{\mu\nu}} = \frac{1}{2} g^{\mu\nu}\Omega_g\,.$$

Therefore we can continue with

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_M \mathcal{L}(F(u, \phi_\epsilon^* g), \phi_\epsilon^* g) &= \int_M \underbrace{\left\{ \frac{\partial \mathcal{L}}{\partial F}\frac{\partial F}{\partial g_{\mu\nu}} + \frac{\partial \mathcal{L}}{\partial g_{\mu\nu}} + \frac{1}{2}\mathcal{L} g^{\mu\nu} \right\}}_{=:Q^{\mu\nu} \text{ is symmetric in } \mu,\nu} (\mathscr{L}_X g)_{\mu\nu}\Omega_g \\ &\overset{(2.3.20)}{=} 2\int_M \nabla_\mu(Q^{\mu\nu}X_\nu)\,\Omega_g - 2\int_M (\nabla_\mu Q^{\mu\nu})X_\nu\,\Omega_g \\ &= -2\int_M (\nabla_\mu Q^{\mu\nu})X_\nu\,\Omega_g\,. \end{aligned}$$

Here we introduced the tensor field

$$Q^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial F}\frac{\partial F}{\partial g_{\mu\nu}} + \frac{\partial \mathcal{L}}{\partial g_{\mu\nu}} + \frac{1}{2}\mathcal{L} g^{\mu\nu}\,,$$

which is symmetric and therefore allowed us to apply Lemma 2.3.2. In the end we also dropped the integral over the total divergence $\nabla_\mu(Q^{\mu\nu}X_\nu)$, which is unproblematic if we assume for instance that X is compactly supported since then possible boundary terms

[17] Notice that this is indeed an assumption on the functions $\mathcal{L}$ and F. Moreover, it has to be clear in the first place how sections u can be pulled back by diffeomorphisms. For general vector bundles (even for the spinor bundle) the latter is not a priori clear.

at "infinity" are suppressed. In total it follows that if u satisfies the Euler-Lagrange equations of the functional (2.3.21), then we have

$$0 = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_M \mathcal{L}(F(u, \phi_\epsilon^* g), \phi_\epsilon^* g) = \int_M (\nabla_\mu Q^{\mu\nu}) X_\nu \Omega_g \,.$$

Since this holds for *all* 1-parameter groups of diffeomorphisms ϕ_ϵ (say whose generating vector field X is compactly supported), this of course simply means that the tensor field $Q^{\mu\nu}$ must be *divergence-free.*

In practice, instead of $Q^{\mu\nu}$ it is more common to lower the indices and consider the tensor $T_{\mu\nu} = -Q_{\mu\nu} = -g_{\mu\alpha} g_{\mu\beta} Q^{\alpha\beta}$. This tensor field is the called *energy-momentum tensor* or *stress-energy-momentum tensor* or *stress-energy tensor.* It is explicitly given by the formula

$$\boxed{T_{\mu\nu} = \frac{\partial \mathcal{L}}{\partial F} \frac{\partial F}{\partial g^{\mu\nu}} + \frac{\partial \mathcal{L}}{\partial g^{\mu\nu}} - \frac{1}{2} \mathcal{L} g_{\mu\nu}} \,. \tag{2.3.22}$$

The last term is easily understood from the formula for $Q^{\alpha\beta}$ before. The first two terms and the different relative sign to the last term can be understood as follows: For a function $F(g^{\mu\nu}) = F(g_{\alpha\beta}(g^{\mu\nu}))$ which depends on the coefficients of the inverse g^{-1} implicitly through an explicit dependence on g, we have by the chain rule[18]

$$\frac{\partial F}{\partial g^{\mu\nu}} = \frac{\partial F}{\partial g_{\alpha\beta}} \frac{\partial g_{\alpha\beta}}{\partial g^{\mu\nu}} \,.$$

To determine the second factor on the right-hand side, notice first that due to $g_{\alpha\sigma} g^{\sigma\gamma} = \delta_\sigma^\gamma$ we have

$$0 = \frac{\partial}{\partial g^{\mu\nu}} \left(g_{\alpha\sigma} g^{\sigma\gamma} \right) = \frac{\partial g_{\alpha\sigma}}{\partial g^{\mu\nu}} g^{\sigma\gamma} + g_{\alpha\sigma} \delta_\mu^\sigma \delta_\nu^\gamma = \frac{\partial g_{\alpha\sigma}}{\partial g^{\mu\nu}} g^{\sigma\gamma} + g_{\alpha\mu} \delta_\nu^\gamma \,.$$

Using this it follows that

$$\frac{\partial g_{\alpha\beta}}{\partial g^{\mu\nu}} = \frac{\partial g_{\alpha\sigma}}{\partial g^{\mu\nu}} \delta_\beta^\sigma = \frac{\partial g_{\alpha\sigma}}{\partial g^{\mu\nu}} \underbrace{g^{\sigma\gamma} g_{\gamma\beta}}_{=\delta_\beta^\sigma} = -g_{\gamma\beta} g_{\alpha\mu} \delta_\nu^\gamma = -g_{\mu\alpha} g_{\nu\beta} \,.$$

This shows that

$$\frac{\partial F}{\partial g^{\mu\nu}} = -g_{\mu\alpha} g_{\nu\beta} \frac{\partial F}{\partial g_{\alpha\beta}} \,,$$

and the right-hand side is precisely the type of term which appears in $-g_{\mu\alpha} g_{\nu\beta} Q^{\alpha\beta}$.

Notice that $T_{\mu\nu}$ is a symmetric tensor field, and since it is metrically equivalent to $Q^{\mu\nu}$ and $Q^{\mu\nu}$ is divergence-free, the same is true for $T_{\mu\nu}$:

$$\nabla_\mu T^{\mu\nu} = 0 \,. \tag{2.3.23}$$

Therefore this tensor field is a natural candidate to use for energy estimates as explained in the previous section (given one can make the previous derivation rigorous for the specific equations of interest).

Let us now consider two simple and well-known examples. Afterwards we turn to the Dirac equation. Recall that besides the property of the energy-momentum tensor to be divergence-free (which follows from the formal arguments above, but is better verified explicitly again in concrete examples), one also has to keep the *positivity condition* for the resulting conserved quantities in mind, i.e. condition (2) of the previous section.

[18] The following computations are a bit indices-loaded. It is instructive to repeat them in one dimension, where all indices have only one value (i.e. where $g^{\mu\nu}$ and $g_{\mu\nu}$ are just reciprocal numbers).

Example 2.3.3. For $m \in \mathbb{R}$ fixed, consider the *Klein-Gordon equation* $(\Box - m^2)\phi = 0$ for a (possibly complex-valued) scalar fields $\phi \in C^\infty(M)$. It is not difficult to verify that the Klein-Gordon equation is the Euler-Lagrange equation of the functional (2.3.21) if one takes as Lagrangian

$$\mathcal{L}(\phi, \partial_\mu \phi) = \partial_\mu \phi\, \overline{\partial^\mu \phi} + \frac{m^2}{2}|\phi|^2 = g^{\mu\nu}\partial_\mu \phi\, \overline{\partial_\nu \phi} + \frac{m^2}{2}|\phi|^2 \,.$$

Evaluating the expression (2.3.22) for the energy-momentum tensor in this case yields

$$T[\phi]_{\mu\nu} = \mathsf{Re}(\partial_\mu \phi\, \overline{\partial_\nu \phi}) - \frac{1}{2}(\partial^\alpha \phi\, \overline{\partial_\alpha \phi} + m^2|\phi|^2) g_{\mu\nu} \,. \tag{2.3.24}$$

A direct computation shows that for any vector field X^ν one has the identity

$$\nabla^\mu\big(T[\phi]_{\mu\nu}X^\nu\big) = \mathsf{Re}\Big[(\Box\,\phi - m^2\phi)\overline{\mathrm{d}\phi(X)}\Big] + \frac{1}{2}T[\phi]_{\mu\nu}(\mathscr{L}_X g)^{\mu\nu} \,, \tag{2.3.25}$$

Comparison to (2.3.20) or an explicit computation show that $T[\phi]_{\mu\nu}$ is indeed divergence-free if ϕ satisfies the Klein-Gordon equation. As observed before, if X is a Killing field and u satisfies the Klein-Gordon equation, then one obtains a conserved quantity. Concerning positivity, notice that if Z is timelike and has unit-length, then

$$T_{\mu\nu}Z^\mu Z^\nu = |\nabla_Z \phi|^2 + \frac{1}{2}(-|\nabla_Z \phi|^2 + |D\phi|^2 + m^2|\phi|^2) = \frac{1}{2}(|\nabla_Z \phi|^2 + |D\phi|^2 + m^2|\phi|^2) \geq 0 \,,$$

where $|D\phi|^2 = \sum_{j=1}^n |\nabla_{E_i}\phi|^2$ for some $E_1, \ldots, E_n$ which are spacelike and complete Z to an orthonormal basis. If Z is the future-directed unit-normal to a foliation of M by Cauchy surfaces then this controls the first Sobolev norm of ϕ over these Cauchy surfaces (of course it will not be conserved unless Z is Killing). Using the Cauchy-Schwarz inequality, one can in fact show that $T_{\mu\nu}X^\mu Y^\mu \geq 0$ for any pair of future-directed causal vectors X^μ, Y^μ (cf. Lemma 2.3.9). See [**Ali10**, Ch. 4] for more about the usefulness of the energy-momentum tensor of the wave equation.

Example 2.3.4. Formulated in terms of the electromagnetic potential $A \in \Omega^1(M)$, the (source-free) Maxwell equations on a Lorentzian manifold (M, g) simply read $\delta\, \mathrm{d}A = 0$. Here $\delta = *\mathrm{d}*$ is the codifferential, and $*$ is the Hodge-star operator of g (see for instance [**Bau09**, Sec. 7.2]). Using that δ is the formal adjoint of d with respect to the pairing $\int_M g(\omega, \eta)\, \mathrm{d}\mu_g$ on differential forms (of the same degree), one easily checks that the Maxwell equations are the Euler-Lagrange equation of the functional (2.3.21) with Lagrange density

$$\mathcal{L}(A, \mathrm{d}A) = \langle \mathrm{d}A, \mathrm{d}A \rangle = g^{\mu\nu} g^{\alpha\beta} \partial_{[\mu} A_{\alpha]} \partial_{[\nu} A_{\beta]} \,,$$

where in the right-hand side the square-brackets refer to antisymmetrization in the usual abstract index notation. Evaluating equation (2.3.22) for the energy-momentum tensor, one finds

$$T_{\mu\nu} = \partial_{[\mu} A_{\gamma]} \partial_{[\nu} A^{\gamma]} - \frac{1}{2}\, \langle \mathrm{d}A, \mathrm{d}A \rangle\, g_{\mu\nu} \,. \tag{2.3.26}$$

The energy-momentum tensor of the Maxwell equation has the same positivity properties (cf. [**Ali10**, Ch. 4]).

Let us now consider the Dirac equation $(\mathcal{D} - m)\psi = 0$, which is the Euler-Lagrange equation of the functional

$$S[\psi] := \int_M \mathsf{Re} \prec\! \psi, (\mathcal{D} - m)\psi \!\succ_{\mathbb{S}M} \mathrm{d}\mu_g = \int_M \prec\! \psi, (\mathcal{D} - m)\psi \!\succ_{\mathbb{S}M} \mathrm{d}\mu_g \,. \tag{2.3.27}$$

Here the real part can be dropped since $\mathcal{D}-m$ is symmetric. However, the Lagrangian should be real-valued. To see that the Dirac equation is the Euler-Lagrange equation of this functional, we explicitly compute its first variation. Let $\psi \in \Gamma^\infty(\mathbb{S}M)$, then for any $\phi \in \Gamma_c^\infty(\mathbb{S}M)$ we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} S[\psi+\epsilon\phi] &= \int_M \big[\prec\psi, (\mathcal{D}-m)\phi\succ_{\mathbb{S}M} + \prec\phi, (\mathcal{D}-m)\psi\succ_{\mathbb{S}M} \big]\, \mathrm{d}\mu_g \\ &= \int_M \big[\prec(\mathcal{D}-m)\psi, \phi\succ_{\mathbb{S}M} + \overline{\prec(\mathcal{D}-m)\psi, \phi\succ_{\mathbb{S}M}} \,\big]\, \mathrm{d}\mu_g \\ &= 2\,\mathsf{Re}\int_M \prec(\mathcal{D}-m)\psi, \phi\succ_{\mathbb{S}M}\, \mathrm{d}\mu_g\,.\end{aligned}$$

Since this only vanishes for arbitrary ϕ if already $(\mathcal{D}-m)\psi = 0$, it follows that the Dirac equation is indeed the Euler-Lagrange equation of (2.3.27).

Unfortunately it is not so straight-forward to evaluate formula (2.3.22) for the energy-momentum tensor in this case. Namely, notice that for this we would need to compute

$$\frac{\partial}{\partial g^{\mu\nu}}\,\mathsf{Re}\prec(\mathcal{D}-m)\psi, \psi\succ_{\mathbb{S}M} = \frac{\partial}{\partial g^{\mu\nu}}\,\mathsf{Re}\prec(\mathrm{i}\gamma^\mu\nabla_\mu - m)\psi, \psi\succ_{\mathbb{S}M}\ .$$

The difficulty is that all spinorial objects like Clifford multiplication and the spin connection depend in a subtle way on the metric.[19] A possible rigorous approach for taking variations with respect to the metric can be found in [**BGM05**]. In particular, in Theorem 5.1 of this article one finds (in somewhat different language) that

$$\frac{\partial\,\mathcal{D}\psi}{\partial g^{\mu\nu}} = \mathrm{i}\gamma_\mu\nabla_\nu\psi + \mathrm{i}\gamma_\nu\nabla_\mu\psi + \mathrm{i}\gamma(\epsilon_{\mu\nu})\psi\,,$$

where $\epsilon_{\mu\nu}$ is some expression in the first derivatives of the variation of the metric, which does not matter for our purposes since it will drop out in the computation of the energy-momentum tensor. Namely, simply using this formula and writing out $\mathsf{Re}z = \frac{1}{2}(z+\overline{z})$, we find

$$\begin{aligned}&\frac{\partial}{\partial g^{\mu\nu}}\,\mathsf{Re}\prec(\mathcal{D}-m)\psi, \psi\succ_{\mathbb{S}M} \\ &\quad= \frac{1}{2}\Big[\prec\frac{\partial\,\mathcal{D}\psi}{\partial g^{\mu\nu}}, \psi\succ_{\mathbb{S}M} + \prec\psi, \frac{\partial\,\mathcal{D}\psi}{\partial g^{\mu\nu}}\succ_{\mathbb{S}M}\Big] \\ &\quad= \frac{1}{2\mathrm{i}}\Big[\prec\gamma_\mu\nabla_\nu\psi, \psi\succ_{\mathbb{S}M} + \prec\gamma_\nu\nabla_\mu\psi, \psi\succ_{\mathbb{S}M} + \prec\gamma(\epsilon_{\mu\nu})\psi, \psi\succ_{\mathbb{S}M} \\ &\qquad\quad - \prec\psi, \gamma_\mu\nabla_\nu\psi\succ_{\mathbb{S}M} - \prec\psi, \gamma_\mu\nabla_\mu\psi\succ_{\mathbb{S}M} - \prec\psi, \gamma(\epsilon_{\mu\nu})\psi\succ_{\mathbb{S}M}\Big] \\ &\quad= \mathsf{Im}\Big[\prec\gamma_\mu\nabla_\nu\psi, \psi\succ_{\mathbb{S}M} + \prec\gamma_\mu\nabla_\mu\psi, \psi\succ_{\mathbb{S}M}\Big]\,.\end{aligned}$$

Here we used in the end that Clifford multiplication is symmetric and therefore the third and sixth term cancel each other. Since the Lagrangian vanishes on solutions of $(\mathcal{D}-m)\psi = 0$, it follows from (2.3.22) that the energy-momentum tensor of the massive

[19] Actually, the whole spinor bundle depends in some sense on the metric via the choice of a spin structure, which refers to the metric.

Dirac equation is given by

$$T[\psi]_{\mu\nu} = \mathsf{Im}\Big[\prec \gamma_\mu \nabla_\nu \psi, \psi \succ_{\mathbb{S}M} + \prec \gamma_\mu \nabla_\mu \psi, \psi \succ_{\mathbb{S}M} \Big], \tag{2.3.28}$$

or in index-free notation

$$T[\psi](X,Y) = \mathsf{Im}\Big[\prec X \cdot \nabla_Y \psi, \psi \succ_{\mathbb{S}M} + \prec Y \cdot \nabla_X \psi, \psi \succ_{\mathbb{S}M} \Big]. \tag{2.3.29}$$

Since this derivation involved some good faith, let us verify directly that this tensor is indeed divergence-free if ψ satisfies the Dirac equation. This can be seen as a hint that (2.3.28) is indeed the "correct" expression for the energy-momentum tensor.

Lemma 2.3.5. *Let $\psi \in \Gamma^\infty(\mathbb{S}M)$ be a solution of the Dirac equation $(\mathcal{D}-m)\psi = 0$. Then the energy-momentum tensor* (2.3.28) *is divergence-free, i.e. $\nabla^\mu T[\psi]_{\mu\nu} = 0$.*

Proof. Let us first consider the tensor field

$$Q(X,Y) = \prec \psi, X \cdot \nabla_Y \psi \succ_{\mathbb{S}M} + \prec \psi, Y \cdot \nabla_X \psi \succ_{\mathbb{S}M} \ .$$

We compute the divergence of Q at a point $p \in M$ using a synchronous orthonormal frame $e_0, \ldots, e_n$ around p (i.e. with $\nabla e_\mu|_p = 0$). In the following all indices refer to this particular frame, and we always implicitly evaluate everything at p even though we do not explicitly write this. In particular, the properties of the frame imply that $\nabla_\mu \gamma_\nu = \gamma_\nu \nabla_\mu$ (cf. the proof of Lemma 2.1.2). Now we compute (at the point p)

$$\begin{aligned}
\nabla^\mu Q_{\mu\nu} &= \nabla^\mu\big(\prec \psi, \gamma_\mu \nabla_\nu \psi \succ_{\mathbb{S}M} + \prec \psi, \gamma_\nu \nabla_\mu \psi \succ_{\mathbb{S}M} \big) \\
&= \prec \nabla^\mu, \gamma_\mu \nabla_\nu \psi \succ_{\mathbb{S}M} + \prec \psi, \gamma_\mu \nabla^\mu \nabla_\nu \psi \succ_{\mathbb{S}M} \\
&\quad + \prec \nabla^\mu \psi, \gamma_\nu \nabla_\mu \psi \succ_{\mathbb{S}M} + \prec \psi, \gamma_\nu \nabla^\mu \nabla_\mu \psi \succ_{\mathbb{S}M} \\
&= \prec -\mathrm{i}\,\mathcal{D}\psi, \nabla_\nu \psi \succ_{\mathbb{S}M} + \prec \psi, \nabla_\nu(-\mathrm{i}\,\mathcal{D}\psi) \succ_{\mathbb{S}M} + \prec \psi, R^\mu_\nu \gamma_\mu \psi \succ_{\mathbb{S}M} \\
&\quad + \prec \nabla^\mu \psi, \gamma_\nu \nabla_\mu \psi \succ_{\mathbb{S}M} + \prec \psi, \gamma_\nu\, \mathcal{D}^2 \psi \succ_{\mathbb{S}M} + \tfrac{\mathrm{scal}}{4} \prec \psi, \gamma_\nu \psi \succ_{\mathbb{S}M} \\
&= \prec \psi, R^\mu_\nu \gamma_\mu \psi \succ_{\mathbb{S}M} + \prec \nabla^\mu \psi, \gamma_\nu \nabla_\mu \psi \succ_{\mathbb{S}M} + m^2 \prec \psi, \gamma_\nu \psi \succ_{\mathbb{S}M} + \tfrac{\mathrm{scal}}{4} \prec \psi, \gamma_\nu \psi \succ_{\mathbb{S}M} \ .
\end{aligned}$$

In the last step we used that $\mathcal{D}\psi = m\psi$. Previously we used that Clifford multiplication is symmetric, $\nabla^\mu \nabla_\nu = \nabla_\nu \nabla^\mu + R^\mu_\nu$ where R^μ_ν is the curvature tensor (of the spin connection), and the Lichnerowicz formula $\nabla^\mu \nabla_\mu = \mathcal{D}^2 + \frac{\mathrm{scal}}{4}$ (cf. Thm. (1.2.24)).

Since Clifford multiplication is symmetric, every term in the last line is real-valued. From this it follows that the divergence of the energy-momentum tensor $T^\psi_{\mu\nu} = -\frac{1}{2}\,\mathsf{Im} Q_{\mu\nu}$ vanishes in p. Since $p \in M$ was arbitrary, this concludes the proof. □

For further reassurance that (2.3.28) is the "correct" energy-momentum tensor, we compute its trace. We will see that the trace vanishes for all solutions of $(\mathcal{D}-m)\psi = 0$ if and only if $m = 0$. This is a further indication that (2.3.28) is the correct expression since the energy-momentum tensor of a field theory should be traceless if and only if the equation is conformally invariant (cf. [**FR04**, Sec. 5]), and the Dirac equation is conformally invariant if and only if $m = 0$ (cf. Remark 3.4.2 of the following chapter or page ix of the introduction).

Lemma 2.3.6. *Let $\psi \in \Gamma^\infty(\mathbb{S}M)$ be a solution of the Dirac equation $(\mathcal{D}-m)\psi = 0$. Then the trace of the energy-momentum tensor* (2.3.28) *is given by*

$$\mathsf{tr}_g\, T[\psi] = m \prec \psi, \psi \succ \ . \tag{2.3.30}$$

In particular the energy momentum tensor is traceless if ψ is massless, i.e. $m = 0$.

Proof. We compute

$$\begin{aligned}\operatorname{tr}_g T^\psi = T^{\psi\mu}_{\ \ \mu} &= -\frac{1}{2}\operatorname{Im}\big(\prec\psi, \gamma^\mu\nabla_\mu\psi\succ_{\mathbb{S}M} + \prec\psi, \gamma^\mu\nabla_\mu\psi\succ_{\mathbb{S}M} \big)\\ &= -\operatorname{Im}\prec\psi, -\mathrm{i}\,\mathcal{D}\psi\succ_{\mathbb{S}M} = \operatorname{Re}\prec\psi, \mathcal{D}\psi\succ_{\mathbb{S}M}\\ &= m \prec\psi,\psi\succ \ .\end{aligned}$$

This is precisely what we wanted to show. □

Up to this point the energy-momentum tensor (2.3.28) looks very promising. Unfortunately it lacks the desired *positivity property* described in Section 2.3.2, which renders it rather useless for energy estimates (at least in a straight-forward way). As we will see, this lack of positivity is related to the old "problem" of negative energy solutions of the Dirac equation.

Lemma 2.3.7. *Let (M, g) be a Lorentzian spin manifold and let $Z \in TM$ be a timelike vector. Then in general $T[\psi]_{\mu\nu}Z^\mu Z^\nu$ is unbounded from above and below, i.e. there can exist both $\psi \in \Gamma^\infty(\mathbb{S}M)$ for which this expression becomes arbitrarily large and others for which it becomes arbitrarily small (going to $-\infty$), even if one restricts to solutions of the Dirac equation $(\mathcal{D}-m)\psi = 0$. At least this is true in Minkowski spacetime for $Z = \partial_t$ being the standard timelike Killing field.*

Proof. In the following we focus on flat Minkowski spacetime $\mathbb{R}^{1,3}$, where we can solve the Dirac equation by spectral methods.[20] In Minkowski spacetime, the Dirac operator is simply given by $\mathcal{D}\psi = \mathrm{i}\gamma^\mu\partial_\mu\psi$ with constant γ-matrices $\gamma^0, \dots, \gamma^3$ (cf. Example 1.1.7), and it acts on functions $\psi : \mathbb{R}^{1,3} \to \mathbb{C}^4$. Moreover the spinorial inner product is just $\prec\cdot,\cdot\succ = \langle\cdot, \gamma^0\cdot\rangle_{\mathbb{C}^4}$, and thus

$$\begin{aligned}T_{00} &= -\operatorname{Im}\prec\psi, \gamma_0\partial_t\psi\succ = -\operatorname{Re}\prec\psi, \mathrm{i}\gamma^0\partial_t\psi\succ\\ &= -\operatorname{Re}\prec\psi, (-\mathrm{i}\gamma^j\partial_j + m)\psi\succ = -\operatorname{Re}\big\langle\psi, \gamma^0(-\mathrm{i}\gamma^j\partial_j + m)\psi\big\rangle =: -\operatorname{Re}\langle\psi, H\psi\rangle_{\mathbb{C}^4} ,\end{aligned}$$

where we used that $(\gamma^0)^2 = 1$ and where we have introduced the *Dirac Hamiltonian* $H = \gamma^0(-\mathrm{i}\gamma^j\partial_j + m)$. As is well-known (cf. [**Tha92**, Thm. 1.1]), the operator H is an unbounded self-adjoint operator on the domain $H^1(\mathbb{R}^3;\mathbb{C}^4) \subset L^2(\mathbb{R}^3;\mathbb{C}^4)$ with purely absolutely continuous spectrum $\sigma(H) = (-\infty, -m] \cup [m, \infty)$. Moreover, for every $k \in \mathbb{R}^3$ there exist $\phi_k^\pm \in C^\infty(\mathbb{R}^3;\mathbb{C}^4) \cap L^\infty(\mathbb{R}^3, \mathbb{C}^4)$, called *plane waves* or *generalized eigenfunctions* which satisfy (as differential equation, i.e. not in L^2)

$$H\phi_k^\pm = \pm\sqrt{|k|^2 + m^2}\,\phi_k^\pm ,$$

and $\big\langle\phi_k^\pm, \phi_k^\pm\big\rangle = 1$. Therefore we have

$$\big\langle\phi_k^\pm, H\phi_k^\pm\big\rangle = \pm\sqrt{|k|^2 + m^2}\,\big\langle\phi_k^\pm, \phi_k^\pm\big\rangle = \pm\sqrt{|k|^2 + m^2}\,.$$

Since this can obviously become arbitrarily large or small, this shows the claim. □

Due to this result, the energy-momentum tensor of the Dirac equation does not seem to be a very useful tool. Remembering the Lichnerowicz formula $\mathcal{D}^2 = \nabla^\mu\nabla_\mu + \frac{\mathrm{scal}}{4}$ (cf. Thm. 1.2.21) and its consequence

$$(\mathcal{D}-m)\psi = 0 \quad\Longrightarrow\quad 0 = (\mathcal{D}+m)(\mathcal{D}-m)\psi = \big(\nabla^\mu\nabla_\mu + \tfrac{\mathrm{scal}}{4} - m^2\big)\psi\,,$$

[20] Similar ideas should also work at least on static spacetimes.

it could be an alternative to use the energy-momentum tensor of the second order wave operator $\nabla^\mu\nabla_\mu$ instead. For this operator, a direct analogue to the energy-momentum tensor of the scalar wave equation (or Klein-Gordon equation in Example 2.3.3) is

$$T[\psi]_{\mu\nu} = \mathsf{Re} \prec \nabla_\mu\psi, \nabla_\nu\psi \succ_{\mathbb{S}M} - \frac{1}{2}\big(\prec \nabla_\alpha\psi, \nabla^\alpha\psi \succ_{\mathbb{S}M} + \prec \psi, \mathcal{B}\psi \succ_{\mathbb{S}M} \big) g_{\mu\nu} \,. \tag{2.3.31}$$

Here $\mathcal{B} \in \Gamma^\infty(\mathsf{End}(\mathbb{S}M))$ can be any endomorphism which is pointwise symmetric with respect to $\prec\cdot,\cdot\succ$. For instance it could just be multiplication by m^2 or it could contain also the scalar curvature term which appears in the Lichnerowicz formula. For now we leave it general. Let us note that one can also derive the energy-momentum tensor (2.3.31) from the general formula (2.3.22) using that the equation $(\nabla^\mu\nabla_\mu - \mathcal{B})\psi = 0$ is the Euler-Lagrange equation of the action functional (2.3.21) with Lagrangian

$$\mathcal{L}(\psi, \nabla_\mu\psi) = \prec \nabla_\mu\psi, \nabla^\mu\psi \succ_{\mathbb{S}M} + \prec \psi, \mathcal{B}\psi \succ_{\mathbb{S}M} \,. \tag{2.3.32}$$

To test the expression (2.3.31), let us compute its divergence.

Lemma 2.3.8. *For any $\psi \in \Gamma^\infty(\mathbb{S}M)$, the energy-momentum tensor* (2.3.31) *satisfies*

$$\begin{aligned} \nabla^\mu T[\psi]_{\mu\nu} &= \mathsf{Re} \prec (\nabla^\mu\nabla_\mu - \mathcal{B})\psi, \nabla_\nu\psi \succ_{\mathbb{S}M} \\ &\quad + \mathsf{Re} \prec R^\mu_\nu\psi, \nabla_\mu\psi \succ_{\mathbb{S}M} + \tfrac{1}{2} \prec \psi, (\nabla_\nu\mathcal{B})\psi \succ_{\mathbb{S}M} \,, \end{aligned} \tag{2.3.33}$$

where R^μ_ν is the (spinorial) curvature tensor.

Proof. We compute

$$\begin{aligned} \nabla^\mu T_{\mu\nu} &= \mathsf{Re} \prec \nabla^\mu\nabla_\mu\psi, \nabla_\nu\psi \succ_{\mathbb{S}M} + \mathsf{Re} \prec \nabla_\mu\psi, \nabla^\mu\nabla_\nu\psi \succ_{\mathbb{S}M} \\ &\quad - \tfrac{1}{2}\big(2\, \mathsf{Re} \prec \nabla_\alpha\psi, \nabla^\mu\nabla^\alpha\psi \succ_{\mathbb{S}M} + 2\, \mathsf{Re} \prec \nabla^\mu\psi, \mathcal{B}\psi \succ_{\mathbb{S}M} + \prec \psi, (\nabla^\mu\mathcal{B})\psi \succ_{\mathbb{S}M} \big) g_{\mu\nu} \\ &= \mathsf{Re} \prec (\nabla^\mu\nabla_\mu - \mathcal{B})\psi, \nabla_\nu\psi \succ_{\mathbb{S}M} + \mathsf{Re} \prec R^\mu_\nu\psi, \nabla_\mu\psi \succ - \tfrac{1}{2} \prec \psi, (\nabla_\nu\mathcal{B})\psi \succ_{\mathbb{S}M} \,. \end{aligned}$$

The curvature enters since in the second step we exchanged the covariant derivatives $\nabla^\mu\nabla_\nu$ (in the second term of the right-hand side in the first line). ☐

Notice that even if ψ satisfies $(\nabla^\mu\nabla_\mu - \mathcal{B})\psi = 0$ and $\nabla\mathcal{B} = 0$, the divergence still contains a curvature term. So in general one cannot hope for exactly conserved quantities (unless the connection ∇ is flat), which is a new feature as compared to the scalar Klein-Gordon equation. One may also notice that for $\mathcal{B} = m^2 - \frac{\text{scal}}{4}$, both terms in the second line of (2.3.33) become curvature terms. So in the case of the Dirac equation, where $\mathcal{B} = m^2 - \frac{\text{scal}}{4}$ is the natural choice, one would always have to deal with curvature-driven "error terms" if one wants to use the tensor field (2.3.31).

Unfortunately also the energy-momentum tensor (2.3.33) does not satisfy good positivity properties. This time the reason is that the inner product $\prec\cdot,\cdot\succ_{\mathbb{S}M}$ is indefinite. To see this, take again the example of Minkowski spacetime $\mathbb{R}^{1,n}$, where spinor fields are functions $\psi : \mathbb{R}^{1,n} \to \mathbb{C}^N$ and where $\prec\cdot,\cdot\succ_{\mathbb{S}\mathbb{R}^{1,n}} = \langle\cdot, \gamma^0\cdot\rangle_{\mathbb{C}^N}$. For $\mathcal{B} = m^2$, we then have for instance

$$\begin{aligned} T[\psi]_{00} &= T[\psi](\partial_t, \partial_t) \\ &= \big\langle \partial_t\psi, \gamma^0\partial_t\psi \big\rangle_{\mathbb{C}^N} + \frac{1}{2}\Big[- \big\langle \partial_t\psi, \gamma^0\partial_t\psi \big\rangle_{\mathbb{C}^N} + \sum_{j=1}^n \big\langle \partial_j\psi, \gamma^0\partial_j\psi \big\rangle_{\mathbb{C}^N} + m^2 \big\langle \psi, \gamma^0\psi \big\rangle_{\mathbb{C}^N} \Big] \\ &= \frac{1}{2}\Big[\big\langle \partial_t\psi, \gamma^0\partial_t\psi \big\rangle_{\mathbb{C}^N} + \sum_{j=1}^n \big\langle \partial_j\psi, \gamma^0\partial_j\psi \big\rangle_{\mathbb{C}^N} + m^2 \big\langle \psi, \gamma^0\psi \big\rangle_{\mathbb{C}^N} \Big] \,. \end{aligned}$$

Since the matrix $\gamma^0 \in \mathsf{M}(N,\mathbb{C})$ has both positive and negative eigenvalues, $T[\psi]_{00}$ will have no definite sign (i.e. for different ψ).

There is a way to circumvent the problem of non-positivity of the energy momentum tensor (2.3.31), but (of course) at yet a different cost. Namely, one can simply replace the indefinite inner product $\prec\cdot,\cdot\succ_{\mathbb{S}M}$ by some positive definite inner product $\langle\cdot,\cdot\rangle$ on $\mathbb{S}M$. That is, after choosing such a positive definite inner product $\langle\cdot,\cdot\rangle$ we consider the tensor field

$$T_{\langle\cdot,\cdot\rangle}[\psi]_{\mu\nu} := \langle\nabla_\mu\psi,\nabla_\nu\psi\rangle - \frac{1}{2}\big(\langle\nabla_\alpha\psi,\nabla^\alpha\psi\rangle + \langle\psi,\mathcal{B}\psi\rangle\big)g_{\mu\nu}\,. \tag{2.3.34}$$

This tensor field indeed has good positivity properties.

Lemma 2.3.9. *Let $\langle\cdot,\cdot\rangle$ be any positive definite Hermitian inner product on $\mathbb{S}M$, and suppose that $\langle\cdot,\mathcal{B}\cdot\rangle$ is also a positive definite Hermitian inner product. Let $\psi\in\Gamma^\infty(\mathbb{S}M)$. Then the tensor* (2.3.34) *satisfies $T_{\langle\cdot,\cdot\rangle}[\psi]_{\mu\nu}X^\mu Y^\nu \geq 0$ for any pair of future-pointing causal vectors $X,Y\in T_pM$.*

Proof. It is clear that the term containing $\mathcal{B}$ is positive, so we only need to worry about the other terms. For simplicity we thus set $\mathcal{B}=0$ in the following. We denote by $|\cdot|$ the norm induced by $\langle\cdot,\cdot\rangle$. Without loss we can assume that X has unit-length. Then we may complete X to an orthonormal basis $E_0 = X, E_1,\ldots,E_n$. With respect to this basis we decompose $Y = Y^0E_0 + Y^jE_j$, and since both X and Y are future-pointing, we have $Y^0>0$. Moreover, since Y is timelike we have $Y^0Y^0 < Y^jY_j$. Keeping this in mind and using the Cauchy-Schwarz inequality several times, we have

$$\begin{aligned}
T[\psi]_{\mu\nu}X^\mu Y^\nu &= Y^0|\nabla_0\psi|^2 + Y^j\langle\nabla_0\psi,\nabla_j\psi\rangle + \frac{Y^0}{2}\big(-|\nabla_0\psi|^2 + \big\langle\nabla^j\psi,\nabla_j\psi\big\rangle\big)\\
&\geq \frac{Y^0}{2}\big(|\nabla_0\psi|^2 + \big\langle\nabla^j\psi,\nabla_j\psi\big\rangle\big) - (Y^jY_j)^{\frac12}\bigg(\sum_{j=1}^n\langle\nabla_0\psi,\nabla_j\psi\rangle^2\bigg)^{\frac12}\\
&\geq \frac{Y^0}{2}\bigg(|\nabla_0\psi|^2 + \big\langle\nabla^j\psi,\nabla_j\psi\big\rangle - 2\bigg(\sum_{j=1}^n\langle\nabla_0\psi,\nabla_j\psi\rangle^2\bigg)^{\frac12}\bigg)\\
&\geq \frac{Y^0}{2}\bigg(|\nabla_0\psi|^2 + \big\langle\nabla^j\psi,\nabla_j\psi\big\rangle - 2|\nabla_0\psi|\bigg(\sum_{j=1}^n\big\langle\nabla^j\psi,\nabla_j\psi\big\rangle^2\bigg)^{\frac12}\bigg)\\
&= \frac{Y^0}{2}\bigg(|\nabla_0\psi|^2 - \bigg(\sum_{j=1}^n\big\langle\nabla^j\psi,\nabla_j\psi\big\rangle^2\bigg)^{\frac12}\bigg)^2\\
&\geq 0\,.
\end{aligned}$$

□

Let us note for instance that in the case where $\nu\in\Gamma^\infty(TM)$ is the future-pointing unit-normal to a foliation $M\cong\mathbb{R}_t\times\Sigma$ by Cauchy surfaces $\Sigma_t=\{t\}\times\Sigma$, we have

$$T_{\langle\cdot,\cdot\rangle}[\psi](\nu,\nu) = \frac{1}{2}\big(\langle\nabla_\nu\psi,\nabla_\nu\psi\rangle + \big\langle\nabla^j\psi,\nabla_j\psi\big\rangle + \langle\psi,\mathcal{B}\psi\rangle\big)\,.$$

This is seen analogously as for the Klein-Gordon equation in Example 2.3.3. If the last term happens to satisfy a "positivity bound" of the form $\langle\psi,\mathcal{B}\psi\rangle \geq C\langle\psi,\psi\rangle$ with $C>0$, then the corresponding "energy" $\int_{\Sigma_t}T_{\langle\cdot,\cdot\rangle}[\psi](\nu,\nu)\,\mathrm{d}\mu_{\Sigma_t}$ basically controls the homogeneous first Sobolev norm of ψ over Σ_t.

While this seems very favorable, the price we have to pay when working with the tensor field (2.3.34) is that because of using a different inner product, the tensor field (2.3.34) will in general no longer be *divergence-free* (even if ψ solves the Dirac equation). The simple concrete reason is that $\langle\cdot,\cdot\rangle$ cannot generally be expected to be compatible with the connection ∇, a fact used crucially before in the computation of divergences. In general, this will make the tensor field (2.3.34) quite useless because although its associated "energies" now control some useful quantities, the time evolution of these quantities is no longer under control.

Nevertheless, there exists (at least) one way of choosing the positive definite inner product $\langle\cdot,\cdot\rangle$ which seems not quite so hopeless. Namely, we can use the fact that for any future-pointing timelike vector field $Z \in \Gamma^\infty(TM)$ the inner product

$$\langle\cdot,\cdot\rangle_Z := \prec\cdot,\gamma(Z)\cdot\succ_{\mathbb{S}M} \tag{2.3.35}$$

is positive definite (cf. Lemma 1.1.23). For this particular choice, the divergence of (2.3.34) will "only" involve additional derivatives of Z (as compared to Lemma 2.3.8):

Lemma 2.3.10. *Let $Z \in \Gamma^\infty(\mathbb{S}M)$ be a future-pointing timelike vector field. Let $\psi \in \Gamma^\infty(\mathbb{S}M)$, and denote by $T_Z[\psi]$ the tensor field given by (2.3.34) for $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_Z$. Then*

$$\begin{aligned}\nabla^\mu T_Z[\psi]_{\mu\nu} = {}& \mathsf{Re}\langle(\nabla^\mu\nabla_\mu - \mathcal{B})\psi,\psi\rangle_Z + \mathsf{Re}\langle R^\mu_\nu\psi,\nabla_\mu\psi\rangle_Z + \frac{1}{2}\langle\psi,(\nabla_\nu\mathcal{B})\psi\rangle_Z \\ & + \prec\nabla_\mu\psi,\gamma(\nabla^\mu Z)\nabla_\nu\psi\succ_{\mathbb{S}M} \\ & - \frac{1}{2}\Big[\prec\nabla_\alpha\psi,\gamma(\nabla^\mu Z)\nabla^\alpha\psi\succ_{\mathbb{S}M} + \prec\psi,\gamma(\nabla^\mu Z)\mathcal{B}\psi\succ_{\mathbb{S}M}\Big]\end{aligned} \tag{2.3.36}$$

PROOF. This follows easily using Lemma 2.3.8. □

One class of examples for which this energy-momentum tensor might indeed be very useful is given by *ultrastatic spacetimes*. These are simply products $M = \mathbb{R}_t \times \Sigma$ with product metrics $g = -\,\mathrm{d}t^2 + h$, where h is some fixed Riemannian metric on Σ (i.e. independent of t). On any such spacetime, the vector field $Z = \partial_t$ is a *parallel* timelike Killing field. Therefore in (2.3.36) all the terms containing ∇Z vanish.

In general situations, however, it seems that the energy-momentum tensor is not very helpful for energy estimates. Since one can always use the conserved current for energy estimates, this might not be such a big problem (see also the next section).

Remark 2.3.11. If one wants to couple the Dirac equation to the Einstein equations, then one has no choice but to use the original energy-momentum tensor (2.3.28).

2.3.4. Higher order energy estimates from lower order ones. In many applications one has an equation for which one can control certain norms of its solution, like the spatial L^2-norm for the Dirac equation, but one additionally needs to control higher order derivatives of solutions for some purpose. There is one basic approach to this, which manifests differently in various situations: one *differentiates the equation.*

Let us illustrate this briefly for the Dirac equation. Suppose to this end that $M = \mathbb{R}_t \times \Sigma$ is a foliation of the underlying spacetime into Cauchy surfaces. Then for any solution $\psi \in \Gamma^\infty(\mathbb{S}M)$ of the Dirac equation $(\mathcal{D}-m)\psi = \varphi$, current conservation leads to the Grönwall estimate

$$\|\psi(t,\cdot)\|_{L^2(\Sigma_t)} \leq \mathrm{e}^{Ct}\bigg(\|\psi|_{\Sigma_0}\|_{L^2(\Sigma_0)} + C\int_0^t \|\varphi(s,\cdot)\|_{L^2(\Sigma_s)}\,\mathrm{d}s\bigg).$$

If one can control the inhomogeneity φ, this at least provides an exponential bound on the growth of the L^2-norm of $\psi(t,\cdot)$. In the following we outline how one might "upgrade" this to also control L^2-norms of some spatial derivatives of ψ.

To this end, suppose that $A \in \mathrm{DiffOp}^k(\mathbb{S}M)$ is a differential operator of order k, which differentiates only in the spatial directions (i.e., those tangent to Σ). Applying A to the whole Dirac equation, we obtain

$$\begin{aligned} & A\varphi = A(\mathcal{D}-m)\psi = (\mathcal{D}-m)(A\psi) - [A,\mathcal{D}]\psi \\ \implies \quad & (\mathcal{D}-m)A\psi = A\varphi + [A,\mathcal{D}]\psi\,, \end{aligned}$$

where $[A,\mathcal{D}] = A\,\mathcal{D} - \mathcal{D}A$ denotes the commutator. Therefore it follows from the previous estimate that

$$\|A\psi(t,\cdot)\|_{L^2(\Sigma_t)} \leq \mathrm{e}^{Ct}\bigg(\|A\psi|_{\Sigma_0}\|_{L^2(\Sigma_0)} + C\int_0^t \|A\varphi(s,\cdot) + [A,\mathcal{D}]\psi(s,\cdot)\|_{L^2(\Sigma_s)}\,\mathrm{d}s \bigg)\,.$$

To see how this can be used, suppose first that we can find an operator A such that $[A,\mathcal{D}] = 0$. Then as long as $A\varphi$ can be controlled (certainly if $\varphi = 0$), we obtain an exponential bound on the growth of $\|A\psi(t,\cdot)\|_{L^2(\Sigma_t)}$. If one can even find an operator A which commutes with the Dirac operator and is *elliptic* on each Σ_t (in a suitable sense), then this yields an exponential bound of some first order Sobolev norm of $\psi(t,\cdot)$ on Σ_t.

In generic situations one will probably not find operators which commute with $\mathcal{D}$. But also then the previous estimate can be useful. Namely, the key observation in this case is that $[A,\mathcal{D}]$ is a differential operator of order zero, i.e. $[A,\mathcal{D}] \in \Gamma^\infty(\mathsf{End}(\mathbb{S}M))$ is just given by some "matrix-multiplication". Therefore one can try to estimate $\|[A,\mathcal{D}]\psi(s,\cdot)\|_{L^2(\Sigma_s)}$ in terms of the already controlled quantity $\|\psi(s,\cdot)\|_{L^2(\Sigma_s)}$. If this succeeds, then one will also obtain an exponential bound for $\|A\psi(t,\cdot)\|_{L^2(\Sigma_t)}$ by another Grönwall argument.

If one succeeds with the above for some first order operator A, one can try to repeat the previous idea with a second-order operator $B \in \mathrm{DiffOp}^2(\mathbb{S}M)$. Then $[B,\mathcal{D}] \in \mathrm{DiffOp}^1(\mathbb{S}M)$ is of first order, so if first derivatives of ψ are already controlled one can maybe obtain control of derivatives of order two of ψ. In this fashion, one can try to work oneself up to higher and higher derivatives.

Remark 2.3.12. The same idea of course also works for any other equation. The basic reason why this works is that the additional error terms one produces by differentiating the equation are always of lower order compared to the highest order derivatives of the differentiated equation. For a linear equation $Pu = w$ they are always of the form of a commutator $[A,P]$ if A is the operator one chooses to differentiate the equation.

CHAPTER 3

A Class of Static, Asymptotically Flat Spacetimes

In this chapter we introduce a special class of static, asymptotically flat spacetimes on which we will analyze the massive Dirac equation in Chapter 4. This class of spacetimes on the one hand contains certain interesting examples such as the Schwarzschild spacetime, and on the other hand is sufficiently simple to allow a pen and paper analysis of the Dirac equation.

3.1. Definition and Causal Properties

We start by defining the spacetimes we shall consider, studying their causal properties, and listing some examples.

3.1.1. The basic form of the spacetimes. We consider Lorentzian manifolds (M, g) which have the form

$$\begin{cases} M = \mathbb{R}_t \times (r_0, r_1)_r \times N \\ \quad g = \mathrm{e}^{2a(r)}[-\mathrm{d}t^2 + \mathrm{d}r^2] + R(r)^2 g_N \end{cases}, \tag{3.1.1}$$

where $(r_0, r_1) \subset \mathbb{R}$ is some interval, (N, g_N) is a connected, compact Riemannian manifold of dimension $\dim N = n - 1$ (without boundary), and $a, R \in C^\infty(r_0, r_1)$ are smooth functions of r with $R > 0$. We will also commonly write the first factor in the metric as

$$\mathrm{e}^{a(r)} =: 1 + A(r)\,, \tag{3.1.2}$$

depending on what is more convenient. In the following, we make some comments and observations about these spacetimes.

Any spacetime (M, g) of the above form is obviously static with complete timelike Killing vector field $K = \partial_t$. It is foliated by the K-orthogonal spacelike hypersurfaces

$$\Sigma_t := \{t\} \times (r_0, r_1) \times N \subset M \tag{3.1.3}$$

However, one should be aware that these spacelike hypersurfaces need not be Cauchy surfaces since (M, g) need not even be globally hyperbolic (cf. Proposition 3.1.2). Moreover, one should note that the integral curves of K are no geodesics in general. Indeed, they are all geodesics if and only if K is parallel, which is the case if and only if the function $a(r)$ is constant. Nevertheless, it might of course happen that some integral curves of K are geodesics even for non-constant $a(r)$.

Next, it is important to notice that (M, g) is a *warped product* (cf. [**O'N83**, Ch.7] for notation)

$$M = Q \times_R N\,, \tag{3.1.4}$$

where (Q, g_Q) is the $1 + 1$ dimensional spacetime given by

$$Q = \mathbb{R}_t \times (r_0, r_1)_r\,, \quad g_Q = \mathrm{e}^{2a(r)}[-\mathrm{d}t^2 + \mathrm{d}r^2]\,. \tag{3.1.5}$$

It might be of some interest to note that (Q, g_Q) as above is in fact the most general $1 + 1$ dimensional spacetime which admits a complete timelike Killing vector field and is homeomorphic to $\mathbb{R}^2$, cf. Theorem A.1.9 in Appendix A.1. Thus the spacetimes we consider are precisely all warped products $M = Q \times_R N$ which have a $1+1$ dimensional, simply connected base Q admitting a complete timelike Killing vector field K, a compact fiber N, and a warping function R satisfying $\mathscr{L}_K R = 0$.

What makes these spacetimes particularly convenient for the analysis of the Dirac equation (and other geometric wave equations) is of course the fact that in a certain sense all the essential functional dependence lies in the $1+1$ dimensional part Q and the warping function $R(r)$. This will become more clear in the actual analysis performed in Chapter 4. Notice moreover that (Q, g_Q) is conformally equivalent to an open subset of flat Minkowski spacetime $\mathbb{R}^{1,1}$, which will also play an important simplifying role.

Remark 3.1.1. From a classical general relativistic perspective, the case $N = S^2$ is maybe most natural. In this case, (M, g) is a usual spherically symmetric spacetime. From a more speculative physical perspective, also the cases $N = S^2 \times Y$ may be of interest, where Y then describes possible “extra dimensions”.

3.1.2. Causal properties. Being warped products, the causal properties of the spacetimes (3.1.1) are easily described. For background information on the notions of causality theory, we refer to the textbooks [**O'N83**, Ch. 14] and [**BEE96**, Ch. 3], or to the review article [**MS08**].

Generally, up to stable causality (i.e. excluding global hyperbolicity), the causal properties of a Lorentzian warped product are as good as those of its Lorentzian factor (cf. [**BEE96**, Sec. 3.6]). In our situation, since (Q, g_Q) given by (3.1.5) is conformally equivalent to an open subset of Minkowski spacetime and causal properties are conformally invariant, it follows that any spacetime (M, g) of the form (3.1.1) is stably causal.[1]

Concerning global hyperbolicity, a warped product is globally hyperbolic if and only if its Lorentzian factor is globally hyperbolic and its Riemannian factor is complete (cf. [**BEE96**, Thm. 3.68]). Since in our situation the Riemannian manifold (N, g_N) is assumed to be compact, it is automatically complete by Hopf-Rinow (cf. [**O'N83**, Ch. 5]). Therefore the only question is whether the Lorentzian factor (Q, g_Q) is globally hyperbolic. Since (Q, g_Q) is conformally equivalent to an open strip of Minkowski spacetime, and since global hyperbolicity is a conformally invariant property (as are all causal properties), it immediately follows that (Q, g_Q), and thus (M, g), is globally hyperbolic if and only if $r_0 = -\infty$ and $r_1 = +\infty$.

Let us summarize this discussion in a proposition. Notice in particular that all causal properties of (M, g) are completely independent of the functions $a(r)$ and $R(r)$.

Proposition 3.1.2. *Any spacetime (M, g) of the form (3.1.1) is stably causal. Moreover, it is globally hyperbolic if and only if $r_0 = -\infty$ and $r_1 = +\infty$.*

The examples in the following paragraph illustrate why we also want to keep the non-globally hyperbolic case in our class.

3.1.3. Graphical representation by Penrose diagrams. It is often very useful for matters of illustration that a spacetime of the form (3.1.1) can be graphically represented by a *Penrose diagram*. The reason why this is possible is that the 1+1 dimensional part (Q, g_Q) is conformally equivalent to an open subset of 1+1 dimensional Minkowski spacetime, and therefore one can simply use the Penrose diagram of Minkowski spacetime.

[1] Actually, any $1+1$ dimensional Lorentzian manifold homeomorphic to $\mathbb{R}^2$ is stably causal, no matter if it is conformally equivalent to a subset of Minkowski spacetime or not ([**BEE96**, Thm. 3.43]).

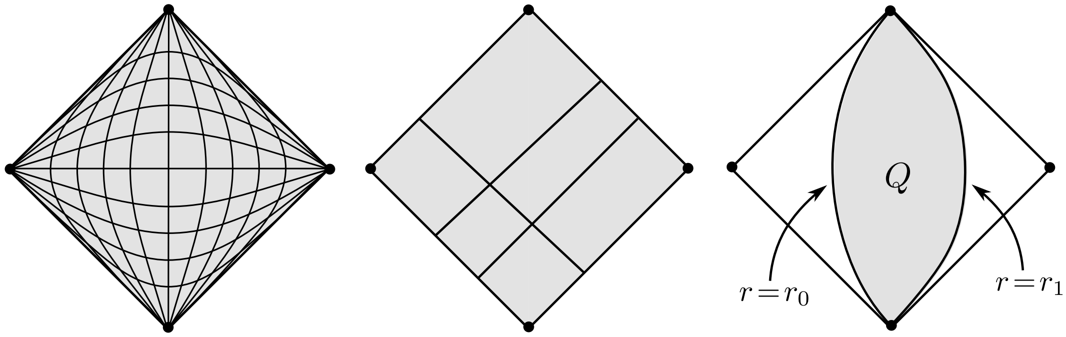

FIGURE 3.1. The left and the middle image show the Penrose diagram of $\mathbb{R}^{1,1}$ with some lines of constant t and r on the left, and some null geodesics in the middle. The right image is the Penrose diagram of Q.

To recall, the whole $1+1$ dimensional Minkowski spacetime $\mathbb{R}^{1,1}_{t,r}$ can be mapped conformally onto a bounded diamond-shaped region of $\mathbb{R}^{1,1}_{\tau,\rho}$ again. To distinguish between the two copies of $\mathbb{R}^{1,1}$ which appear here, we denote their points by (t,r) and (τ,ρ) respectively. To visualize the embedding, in the left image of figure 3.1 we have sketched its image together with some lines of constant t and r. Analytical details can be found in Appendix A.2. A crucial feature is that the embedding is *conformal*, i.e. the push-forward of the Minkowski metric to the interior region by the embedding is conformally equivalent to the Minkowski metric of the surrounding $\mathbb{R}^{1,1}_{\tau,\rho}$. A particularly useful consequence of this is that the causal structure of the embedded $\mathbb{R}^{1,1}_{t,r}$ coincides precisely with the one given by the surrounding $\mathbb{R}_{\tau,\rho}$. For instance, lightlike geodesics (of both) are given by straight lines at $45°$ inclination angle, as illustrated in the image in the middle of figure 3.1.

Let us now return to a spacetime $M = Q \times_R N$ of the form (3.1.1). Since the $1+1$ dimensional part $Q = \mathbb{R}_t \times (r_0, r_1)_r$ with metric $g_Q = \mathrm{e}^{2a}[-\mathrm{d}t^2 + \mathrm{d}r^2]$ is conformally equivalent to a strip of Minkowski spacetime $\mathbb{R}_{t,r}$, the image of Q in the Penrose diagram of Minkowski spacetime also gives a conformal representation of Q, which we call the *Penrose diagram* of Q. It is sketched in the right image of figure 3.1. We will also think of it as providing a graphical representation of the whole spacetime $M = Q \times_R N$ by taking every point of this diagram to represent a copy of N.

3.1.4. Some concrete examples. We illustrate the class of spacetimes of the form (3.3.1) by some concrete examples.

Example 3.1.3. Let us start with the simplest possible example, *Minkowski spacetime* $\mathbb{R}^{1,n} = \mathbb{R}_t \times \mathbb{R}^n$ with metric $g_{\mathbb{R}^{1,n}} = -\mathrm{d}t^2 + g_{\mathbb{R}^n}$. Here $g_{\mathbb{R}^n}$ is the flat Euclidean metric on $\mathbb{R}^n$. After removing the spatial origin $r = 0$, we can express the spatial part in polar coordinates, i.e. we use the diffeomorphism

$$(0,\infty)_r \times S^{n-1}_\omega \ni (r,\omega) \longmapsto r\omega \in \mathbb{R}^n \setminus \{0\}\,.$$

As is well-known, the Euclidean metric then takes the form

$$g_{\mathbb{R}^n} = \mathrm{d}r^2 + r^2 g_{S^{n-1}}\,.$$

so that Minkowski spacetime (with the line $\{r=0\}$ removed) is given by

$$\begin{cases} \mathbb{R}^{1,n}\setminus\{r=0\} = \mathbb{R}_t \times (0,\infty)_r \times S^{n-1} \\ \qquad\qquad g_{\mathbb{R}^{1,n}} = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2 g_{S^{n-1}} \end{cases}. \tag{3.1.6}$$

Obviously this is of the form (3.1.1), and it is *not* globally hyperbolic as $r_0 = 0 \neq -\infty$.

Of course, the whole Minkowski spacetime is globally hyperbolic, and we have only destroyed this property by the removal of the line $\{r = 0\}$. Still, if one wants to use the polar coordinate form 3.1.3, as is convenient in many situations, and wants to study for instance certain partial differential equations on Minkowski spacetime, one has to impose additional boundary conditions at $r = 0$ which reflect the fact that only the presentation in polar coordinates breaks down at $r = 0$ (cf. for instance [**Mas11b**]).

Example 3.1.4. Let $\mu > 0$. Then the *exterior Schwarzschild spacetime* of mass μ is, in polar coordinates, given by

$$\begin{cases} M = \mathbb{R}_t \times (2\mu, \infty)_r \times S^2 \\ g = -\left(1 - \frac{2\mu}{r}\right) \mathrm{d}t^2 + \left(1 - \frac{2\mu}{r}\right)^{-1} \mathrm{d}r^2 + r^2 g_{S^2} \end{cases} .$$

This is not yet of the form (3.1.1), but one can make a coordinate transformation $r \to r_*$ to achieve this. Concretely, introduce the *tortoise coordinate* $r_* = r_*(r)$ by

$$r_*(r) = r + 2\mu \log\left(\tfrac{r}{2\mu} - 1\right) .$$

Then a simple computation reveals that expressed in the coordinates (t, r_*, ω) the exterior Schwarzschild spacetime is given by

$$\begin{cases} M = \mathbb{R}_t \times (-\infty, \infty)_{r_*} \times S^2 \\ g = \left(1 - \frac{2\mu}{r(r_*)}\right)[-\,\mathrm{d}t^2 + \mathrm{d}r^2] + r_*(r)^2 g_{S^2} \end{cases} , \tag{3.1.7}$$

This now has the form (3.1.1), and we see that the exterior Schwarzschild spacetime is globally hyperbolic since the coordinate r_* ranges over all of $(-\infty, \infty)$.

These two examples have been placed into a more general framework in [**Mas12**] (see also [**Mas11b**], [**Mas11a**]), motivated by the structure of certain spherically symmetric solutions of the Einstein-Yang Mills equation constructed in [**SW93**] and [**SWY93**]. We briefly explain these spacetimes in the following two examples.

Example 3.1.5. In [**Mas12**, Ch. 4], a spacetime of the form

$$\begin{cases} M = \mathbb{R}_t \times (0, \infty)_r \times S^2 \\ g = -\frac{1}{T^2(r)} \mathrm{d}t^2 + K(r)^2 \mathrm{d}r^2 + r^2 g_{S^2} \end{cases}$$

is called a *spherically symmetric particle-like geometry* if the functions $T, K \in C^\infty[0, \infty)$ are smooth up to $r = 0$ and satisfy

$$K(0) = 1 , \quad T'(0) = K'(0) = 0 ,$$

as well as

$$T(r), K(r) \sim 1 + \mathcal{O}(r^{-1}) \quad \text{and} \quad \frac{T'(r)}{T(r)} + \frac{K'(r)}{K(r)} \sim 1 + \mathcal{O}(r^{-1}) \text{ as } r \to \infty .$$

Introducing the analogue of the tortoise coordinate in this context, i.e. setting (cf. the proof of Theorem A.1.9)

$$r_*(r) = \int_0^r T(s)K(s)\,\mathrm{d}s ,$$

one has $g = T^{-2}[-\,\mathrm{d}t^2 + \mathrm{d}r_*^2] + r^2 g_{S^2}$. Here $r = r(r_*)$ is now of course to be viewed as function of r_*. This has the form (3.1.1) with r_* taking values in the interval $(0, \infty)$.[2] Therefore (M, g) is not globally hyperbolic. However, the conditions for T and K at

[2]That r_* takes these values follows easily from the asymptotic behaviour of the functions T and K.

the origin and the fact that the factor $r^2 g_{S^2}$ tends to zero at the origin give a hint that the spacetime (M, g) might perhaps be extended to a globally hyperbolic spacetime by including the "point" $r = 0$ (similar to Minkowski spacetime). While this is not discussed concretely in [**Mas12**, Ch. 4], a possible indication is that, as shown there, the Cauchy problem for the wave equation is well-posed for suitable boundary conditions at $r = 0$.

Example 3.1.6. As a generalization of the exterior Schwarzschild spacetime, in [**Mas12**, Ch. 3] a spacetime of the form

$$\begin{cases} M = \mathbb{R}_t \times (r_0, \infty)_r \times S^2 \\ \;g = -\frac{1}{T^2(r)}\, \mathrm{d}t^2 + K(r)^2\, \mathrm{d}r^2 + r^2 g_{S^2} \end{cases},$$

is called a *spherically symmetric black hole* if the functions $T, K \in C^\infty(r_0, \infty)$ satisfy

$$T(r), K(r) \sim (r - r_0)^{-\frac{1}{2}} + \mathcal{O}(1) \text{ as } r \to r_0\,,$$

as well as the same condition at $r = \infty$ as in Example 3.1.5. Again one can introduce a tortoise-like coordinate via

$$r_*(r) = \int_a^r T(s)K(s)\, \mathrm{d}s$$

for some arbitrarily fixed $a \in (r_0, \infty)$, so that the metric takes the form (3.1.1). Moreover, this time the behaviour of T and K for $r \to r_0$ implies that r_* takes values in all of $\mathbb{R}$. Thus, any spherically symmetric black hole spacetime (as defined here) is globally hyperbolic.

Further concrete examples can be obtained by taking a product $(M, g) \oplus (Y, g_Y)$, where (M, g) is any of the previous examples (or a general spacetime of the form (3.1.1)), and (Y, g_Y) is any compact Riemannian manifold.

3.2. Asymptotic Flatness Conditions at $r = \infty$

In the analysis of solutions of the Dirac equation in the following chapter, we will always require that the underlying spacetimes (M, g) are of the form (3.1.1) with $r_1 = \infty$. Moreover, we will require that they satisfy the following asymptotic conditions at $r = \infty$:

Definition 3.2.1. (Decay and boundedness conditions at $r = \infty$)
Let (M, g) be of the form (3.1.1) with $r_1 = \infty$, and let $A(r) = \mathrm{e}^{a(r)} - 1$.

i.) We say that (M, g) is C^k*-bounded at* $r = \infty$ if for some $r_m > r_0$ we have

$$\|A\|_{C^k(r_m,\infty)}\,, \|R^{-1}\|_{C^k(r_m,\infty)} < \infty\,. \tag{3.2.1}$$

ii.) We say that (M, g) is C^k*-asymptotically flat at rate* $\alpha > 0$ *at* $r = \infty$ if for some $r_m > r_0$ there exists a constant $C > 0$ such that for all $j = 0, \ldots, k$ we have

$$\left|A^{(j)}(r)\right|, \left|\left(R(r)^{-1}\right)^{(j)}\right| \leq \frac{C}{(1+r)^\alpha} \qquad \forall r > r_m\,. \tag{3.2.2}$$

We simply say that (M, g) is C^k*-asymptotically flat at* $r = \infty$ if it is C^k-asymptotically flat of some rate $\alpha > 0$.

Clearly C^k-asymptotic flatness at $r = \infty$ implies C^k-boundedness at $r = \infty$. The reason for introducing both notions is in order to include spacetimes which are C^k-bounded for some $k \in \mathbb{N}$ but only C^ℓ-asymptotically flat for some $\ell < k$. For instance, the results in the next chapter only require that the underlying spacetime is C^1-asymptotically flat, but C^k-bounded for some $k \geq 1$.

The most important explicit example which satisfies the conditions of Definition 3.2.1 is once more the Schwarzschild spacetime.

Example 3.2.2. Consider the exterior Schwarzschild spacetime (M, g) of mass $\mu > 0$ as presented in Example 3.1.4. Using the tortoise coordinate $r_* \in (-\infty, \infty)$, the Schwarzschild metric has the form (3.1.1) for

$$a(r_*) = \tfrac{1}{2}\log\left(1 - \tfrac{2\mu}{r(r_*)}\right), \text{ hence } A(r_*) = \left(1 - \tfrac{2\mu}{r(r_*)}\right)^{\frac{1}{2}} - 1, \quad \text{and } R(r_*) = r(r_*),$$

where $r(r_*)$ is implicitly defined by the identity

$$r_*(r) = r + 2\mu \log\left(\tfrac{r}{2\mu} - 1\right).$$

Using that $r_*(r) \sim r$ for large values of r, it is not difficult to see that (M, g) is C^k-asymptotically flat of power $\alpha = 1$ at $r_* = \infty$ for any $k \in \mathbb{N}$. In particular, as remarked before, this implies that it is C^k-bounded at $r_* = \infty$ for any $k \in \mathbb{N}$.

Besides the Schwarzschild spacetime, also the spherically symmetric particle-like and spherically symmetric black hole geometries, which were described in Example 3.1.5 and Example 3.1.6 and include the Einstein-Yang Mills spacetimes constructed in [**SW93**] and [**SWY93**], are (at least) C^1-asymptotically flat of power $\alpha = 1$ at $r = \infty$.

Remark 3.2.3. One might wonder why we call (3.2.2) an asymptotic flatness condition. One reason is that the $1+1$ dimensional part (Q, g_Q) becomes asymptotically close to $\mathbb{R}^{1,1}$ for $r \to \infty$ since $\mathrm{e}^{2a(r)} = (1 + A(r))^2 \to 1$. However, also the remaining factor $(N, R^{-2} g_N)$ becomes flat at least in an intuitive sense due to the rescaling by $R(r)^{-2}$ which tends to infinity as $r \to \infty$. For $N = S^{n-1}$ see also the following remark.

Remark 3.2.4. In the existing literature, asymptotic flatness of a Riemannian manifold (M, g) is often defined in terms of certain decay properties of the metric and its derivative formulated in weighted Sobolev spaces (see for instance [**Bar86**, Def. 2.1]). To be more precise, first it is assumed that the complement of some compact subset of M is diffeomorphic to the complement of a ball in $\mathbb{R}^n$. Afterwards decay conditions are specified in the corresponding coordinates "at infinity", or in weighted Sobolev spaces defined using these coordinates. Our spacetimes can clearly only fit this setup for $N = S^{n-1}$. In this case it should not be difficult to show that our asymptotic flatness condition (3.2.2) imply some of those in the existing literature (on spatial slices $t = \textit{const.}$).

3.3. The Factorization of the Dirac Equation on these Spacetimes

Now we turn to the Dirac equation on spacetimes of the form (3.1.1). The goal in the following is to find suitable expressions for spinorial objects, such as the Dirac operator, which are adapted to the warped product structure (3.1.4) of the spacetime.[3] To this end we will make use of special properties of the Clifford algebras and spin representations in the involved dimensions and signatures. Although the content of this section is rather abstract, one may imagine it to be nothing but a convenient "choice of frame" on the spinor bundle with respect to which many of the spin geometric quantities take a more simple form. The most important outcome in the end is the "factorization formula" (3.3.32) for the Dirac operator, which will be used in the following chapter to separate variables in the Dirac equation.

[3] The special form of the Dirac operator on warped products has of course been studied before, and the content of this section is certainly influenced by the existing literature. See for instance [**KT03**] and further references therein.

To recall the setup, we consider warped product spacetimes of the form

$$M^{1,n} = Q^{1,1} \times N^{n-1}\,, \quad g = g_Q + R(q)^2 g_N\,. \tag{3.3.1}$$

Here (Q, g_Q) is a $1+1$ dimensional Lorentzian manifold, $R : Q \to (0, \infty)$ is a smooth positive function on Q, and (N, g_N) is an $n-1$ dimensional Riemannian manifold. Moreover, we assume that Q is oriented and time-oriented, N is oriented, and that both are equipped with spin structures $(\mathsf{Spin}^+(Q), \Theta_Q)$, $(\mathsf{Spin}(N), \Theta_N)$. On M we then choose the induced orientation and time-orientation and a "product type" spin structure as constructed in Section 3.3.5 below. Concerning concepts and notation for Clifford algebras, spinors etc. we refer to Chapter 1 and the literature cited therein.

3.3.1. Reduction of the orthonormal frame bundle. One basic consequence of the warped product structure is that we can further reduce the orthonormal frame bundle to the subgroup $\mathsf{SO}^+(1,1) \times \mathsf{SO}(n-1) \subset \mathsf{SO}^+(1,n)$ via the homomorphism

$$f : \mathsf{SO}^+(1,1) \times \mathsf{SO}(n-1) \to \mathsf{SO}^+(1,n)\,, \quad (A, B) \mapsto \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}. \tag{3.3.2}$$

In the following, whenever $E \to Q$ is a bundle over Q, we denote by $\widetilde{E} \to M$ the pullback bundle with respect to the projection $M = Q \times N \to Q$, and similar for bundles over N. We then have the following reduction of the frame bundle, the proof of which is clear.

Lemma 3.3.1. *The product* $\mathsf{SO}^+(M; Q \times_R N) := \widetilde{\mathsf{SO}^+(Q)} \times \widetilde{\mathsf{SO}(N)}$ *is an* $\mathsf{SO}^+(1,1) \times \mathsf{SO}(n-1)$*-principal bundle over* M*, and the map*

$$\mathsf{SO}^+(M; Q \times_R N) \ni (E, F) \mapsto (E, R^{-1}F) \in \mathsf{SO}^+(M) \tag{3.3.3}$$

is a reduction with respect to the homomorphism (3.3.2)*.*

The basic idea in the following is that we want to "lift" this reduction to spinors. To this end, we first need some algebraic preparations.

3.3.2. A certain realization of $\mathrm{Cl}(1,n)$**.** As a special case of (1.1.5), there exists an isomorphism

$$\mathrm{Cl}(1,n) \cong \mathrm{Cl}(1,1) \otimes \mathrm{Cl}(n-1)\,. \tag{3.3.4}$$

This isomorphism is very useful for describing the spinor bundles over warped products of the form (3.3.1). Therefore we now spell it out explicitly.

On the one hand, we always regard $\mathbb{R}^{1,1}$ and $\mathbb{R}^{n-1}$ as subspaces of $\mathbb{R}^{1,n}$ via $x \mapsto (x, 0)$ and $y \mapsto (0, y)$, respectively. On the other hand, we also regard them as subspaces of the respective Clifford algebras $\mathrm{Cl}(1,1)$, $\mathrm{Cl}(n-1)$, and $\mathrm{Cl}(1,n)$. It should always be clear from the context to which point of view we are referring momentarily.

Lemma 3.3.2. *Let* $e_0, e_1 \in \mathbb{R}^{1,1}$ *be an oriented and time-oriented orthonormal basis. Then the algebra homomorphism* $\mathrm{Cl}(1,1) \otimes \mathrm{Cl}(n-1) \to \mathrm{Cl}(1,n)$ *with*

$$v \otimes 1 \mapsto v \quad \forall v \in \mathbb{R}^{1,1} \qquad \text{and} \qquad 1 \otimes w \mapsto e_1 e_0 w \quad \forall w \in \mathbb{R}^{n-1} \tag{3.3.5}$$

is an isomorphism of algebras. Here the right-hand side denotes Clifford multiplication in $\mathrm{Cl}(1,n)$*. This isomorphism is independent of the particular choice of* e_0, e_1*.*

Proof. It is straight-forward to check that (3.3.5) really defines a homomorphism. Moreover, it is also easy to see that this homomorphism is injective, and so it must already be an isomorphism for dimensional reasons.

Concerning the claim about independence of the choice of basis, take another oriented and time-oriented orthonormal basis e_0', e_1'. Then there exists $A \in \mathsf{SO}^+(1,1)$ such that $e_0' = A^\mu_0 e_\mu$ and $e_1' = A^\nu_1 e_\nu$. Using the Clifford relations it follows that in $\mathrm{Cl}(1,n)$ we have

$$e_1'e_0' = (A^\mu_1 e_\mu)(A^\nu_0 e_\nu) = \underbrace{A^0_1 A^0_0 - A^1_1 A^1_0}_{=0\,(A\in\mathsf{O}(1,1))} + \underbrace{(A^1_1 A^0_0 - A^0_1 A^1_0)}_{=\det A = 1} e_1 e_0 = e_1 e_0\,.$$

This shows the independence. □

In the following, writing $\mathrm{Cl}(1,n) \cong \mathrm{Cl}(1,1) \otimes \mathrm{Cl}(n-1)$ always refers to the particular isomorphism in Lemma 3.3.2. Complexifying everything, we also obtain an isomorphism $\mathbb{C}\mathrm{l}(1,n) \cong \mathbb{C}\mathrm{l}(1,1) \otimes \mathbb{C}\mathrm{l}(n-1)$, the tensor product now being taken over $\mathbb{C}$ of course.

3.3.3. Consequences for spin groups. Recall that $\mathsf{Spin}^+(p,q) \subset \mathrm{Cl}(p,q)$ consists of all elements of the form $v_1 \cdots v_{2k}$ where $v_1, \dots, v_{2k} \in \mathbb{R}^{p,q}$ have positive or negative unit-length and the total number of either type, i.e. having either sign, is even.

Lemma 3.3.3. *The isomorphism* $\mathrm{Cl}(1,n) \cong \mathrm{Cl}(1,1) \otimes \mathrm{Cl}(n-1)$ *induces an embedding* $\varphi : \mathsf{Spin}^+(1,1) \times \mathsf{Spin}(n-1) \to \mathsf{Spin}^+(1,n)$ *by mapping a pair* $(v_1 \cdots v_{2k}, w_1 \cdots w_{2\ell})$ *to the image of* $(v_1 \cdots v_{2k}) \otimes (w_1 \cdots w_{2\ell})$*. Moreover, this embedding is compatible with the embedding* (3.3.2) *of orthogonal groups in the sense that the following diagram commutes*

$$\begin{array}{ccc} \mathsf{Spin}^+(1,1) \times \mathsf{Spin}(n-1) & \xrightarrow{\ \varphi\ } & \mathsf{Spin}^+(1,n) \\ \Big\downarrow{\scriptstyle \vartheta_{1,1} \times \vartheta_{n-1}} & & \Big\downarrow{\scriptstyle \vartheta_{1,n}} \\ \mathsf{SO}^+(1,1) \times \mathsf{SO}(n-1) & \xrightarrow{\ f\ } & \mathsf{SO}^+(1,n) \end{array}\quad, \tag{3.3.6}$$

where $\vartheta_{1,1}$*,* ϑ_{n-1}*, and* $\vartheta_{1,n}$ *denote the respective spin coverings.*

Proof. Let $v_1 \cdots v_{2k} \in \mathsf{Spin}^+(1,1)$ and $w_1 \cdots w_{2\ell} \in \mathsf{Spin}(n-1)$. Then the image of $(v_1 \cdots v_{2k}) \otimes (w_1 \cdots w_{2\ell})$ in $\mathrm{Cl}(1,n)$ under the homomorphism (3.3.4) is just

$$\begin{aligned} v_1 \cdots v_{2k}(e_1e_0w_1)\cdots(e_1e_0w_{2\ell}) &= v_1 \cdots v_{2k}\underbrace{(e_1e_0 \cdots e_1e_0)}_{2\ell \text{ times}} w_1 \cdots w_{2\ell} \\ &= v_1 \cdots v_{2k}\underbrace{(e_1e_0e_1e_0)}_{=1} \cdots (e_1e_0e_1e_0) w_1 \cdots w_{2\ell} \\ &= v_1 \cdots v_{2k} w_1 \cdots w_{2\ell}\,. \end{aligned}$$

Clearly this lies in $\mathsf{Spin}^+(1,n)$ since it is the product of an even number of timelike and an even number of spacelike unit vectors. It is also clear that this map is an embedding.

Concerning the second claim about the compatibility with the embedding of orthonormal groups, notice first that

$$(v_1 \cdots v_{2k} w_1 \cdots w_{2\ell})^{-1} = w_{2\ell} \cdots w_1 v_{2k} \cdots v_1\,,$$

as is easily verified. Therefore, for any $x \in \mathbb{R}^{1,1} \subset \mathbb{R}^{1,n}$ we have by the definition of the spin covering (cf. Sec. 1.1.4)

$$\begin{aligned} \vartheta_{1,n}(v_1 \cdots v_{2k} w_1 \cdots w_{2\ell})x &= v_1 \cdots v_{2k} w_1 \cdots w_{2\ell} \cdot x \cdot w_{2\ell} \cdots w_1 v_{2k} \cdots v_1 \\ &= v_1 \cdots v_{2k} w_1 \cdots w_{2\ell} w_{2\ell} \cdots w_1 x v_{2k} \cdots v_1 \\ &= v_1 \cdots v_{2k} x v_{2k} \cdots v_1 \\ &= \vartheta_{1,1}(v_1 \cdots v_{2k})x\,. \end{aligned}$$

Here we used that $x \in \mathbb{R}^{1,1} \subset \mathbb{R}^{1,n}$ of course commutes in $\mathrm{Cl}(1,n)$ with the product of any even number of vectors in $\mathbb{R}^{n-1} \subset \mathbb{R}^{1,n}$ since these are all orthogonal to x. Similarly, if $y \in \mathbb{R}^{n-1}$ then one finds that

$$\vartheta_{1,n}(v_1 \cdots v_{2k} w_1 \cdots w_{2\ell})y = \vartheta_{n-1}(w_1 \cdots w_{2\ell})y\,.$$

This now shows that with respect to the splitting $\mathbb{R}^{1,n} = \mathbb{R}^{1,1} \oplus \mathbb{R}^{n-1}$ we have

$$\vartheta_{1,n}(v_1 \cdots v_{2k} w_1 \cdots w_{2\ell}) = \begin{pmatrix} \vartheta_{1,1}(v_1 \cdots v_{2k}) & 0 \\ 0 & \vartheta_{n-1}(w_1 \cdots w_{2\ell}) \end{pmatrix}$$

as desired. □

3.3.4. Consequences for spin spaces. Let $\rho_{1,1} : \mathbb{C}\mathrm{l}(1,1) \to \mathsf{End}(\mathbb{S}_{1,1})$ and $\rho_{n-1} : \mathbb{C}\mathrm{l}(n-1) \to \mathsf{End}(\mathbb{S}_{n-1})$ be the spin representations, cf. Section 1.1.7.[4] Since $\rho_{1,1}$ and ρ_{n-1} are irreducible and finite-dimensional, it follows that the tensor product

$$\rho_{1,1} \otimes \rho_{n-1} : \mathbb{C}\mathrm{l}(1,n) \cong \mathbb{C}\mathrm{l}(1,1) \otimes \mathbb{C}\mathrm{l}(n-1) \to \mathsf{End}(\mathbb{S}_{1,1} \otimes \mathbb{S}_{n-1}) \tag{3.3.7}$$

is again an irreducible representation(cf. [**E**$^+$**11**, Thm. 3.10.2]).[5] But then it follows that (3.3.7) must be the complex spin representation in signature $(1,n)$, or one of the two in case that $\mathbb{C}\mathrm{l}(1,n)$ has two inequivalent irreducible representations. So we see that

$$\mathbb{S}_{1,n} \cong \mathbb{S}_{1,1} \otimes \mathbb{S}_{n-1}\,. \tag{3.3.8}$$

Next, let $\prec \cdot,\cdot \succ_{\mathbb{S}_{1,1}}$ and $\prec \cdot,\cdot \succ_{\mathbb{S}_{n-1}}$ be the inner products on spin spaces that are invariant under the respective spin groups, cf. Section 1.1.7. The first inner product has split signature $(1,1)$, whereas the second one is positive definite. Moreover, Clifford multiplication by vectors on $\mathbb{S}_{n-1}$ is skew-symmetric with respect to $\prec \cdot,\cdot \succ_{\mathbb{S}_{n-1}}$ and Clifford multiplication by vectors on $\mathbb{S}_{1,1}$ is symmetric with respect to $\prec \cdot,\cdot \succ_{\mathbb{S}_{1,1}}$.

From these two inner products we can of course form an indefinite inner product on $\mathbb{S}_{1,n} \cong \mathbb{S}_{1,1} \otimes \mathbb{S}_{n-1}$ by setting

$$\prec \phi \otimes \psi, \phi' \otimes \psi' \succ_{\mathbb{S}_{1,n}} = \prec \phi, \phi' \succ_{\mathbb{S}_{1,1}} \prec \psi, \psi' \succ_{\mathbb{S}_{n-1}} \tag{3.3.9}$$

on simple tensors and extending it sesquilinearly to all of $\mathbb{S}_{1,n}$. Note that this inner product has split signature $(\dim \mathbb{S}_{n-1}, \dim \mathbb{S}_{n-1})$.

Lemma 3.3.4. *Clifford multiplication on $\mathbb{S}_{1,n} \cong \mathbb{S}_{1,1} \otimes \mathbb{S}_{n-1}$ is symmetric with respect to the inner product* (3.3.9). *As a consequence, this inner product is* $\mathsf{Spin}^+(1,n)$*-invariant.*

PROOF. We show that Clifford multiplication by any vector is symmetric. To this end, note that it suffices to show this separately for vectors in $\mathbb{R}^{1,1}$ and vectors in $\mathbb{R}^{n-1}$. For vectors in $\mathbb{R}^{1,1}$ this follows from the fact that they act on $\mathbb{S}_{1,n} \cong \mathbb{S}_{1,1} \otimes \mathbb{S}_{n-1}$ simply by acting on the first factor and since this action is symmetric with respect to $\prec \cdot,\cdot \succ_{\mathbb{S}_{1,1}}$. Now let $v \in \mathbb{R}^{n-1}$. Denoting Clifford multiplication on $\mathbb{S}_{1,n}, \mathbb{S}_{1,1}, \mathbb{S}_{n-1}$ by $\gamma, \gamma^{(1,1)}, \gamma^{(n-1)}$, respectively, due to (3.3.5) we have

$$\begin{aligned}
\prec \gamma(v)(\phi \otimes \psi), \phi' \otimes \psi' \succ_{\mathbb{S}_{1,n}} &= \prec (\gamma^{(1,1)}(e_1 e_0)\phi) \otimes (\gamma^{(n-1)}(v)\psi, \phi' \otimes \psi' \succ_{\mathbb{S}_{1,n}} \\
&= \prec \gamma^{(1,1)}(e_1 e_0)\phi, \phi' \succ_{\mathbb{S}_{1,1}} \prec \gamma^{(n-1)}(v)\psi, \psi' \succ_{\mathbb{S}_{n-1}} \\
&= \prec \phi, \gamma^{(1,1)}(e_0 e_1)\phi' \succ_{\mathbb{S}_{1,1}} \cdot (- \prec \psi, \gamma^{(n-1)}(v)\psi' \succ_{\mathbb{S}_{n-1}}) \\
&= \prec \phi, \gamma^{(1,1)}(e_1 e_0)\phi' \succ_{\mathbb{S}_{1,1}} \cdot \prec \psi, \gamma^{(n-1)}(v)\psi' \succ_{\mathbb{S}_{n-1}}
\end{aligned}$$

[4]In the case where $\mathbb{C}\mathrm{l}(n-1)$ has two irreducible representations, simply choose one of them.

[5]The reason is that finite-dimensional representations of algebras are irreducible if they are surjective.

$$= \prec \phi \otimes \psi, \gamma(v)(\phi' \otimes \psi') \succ_{\mathbb{S}_{1,n}}$$

This shows symmetry of Clifford multiplication with respect to $\prec \cdot, \cdot \succ_{1,n}$. The statement about the $\mathsf{Spin}^+(1,n)$-invariance now follows from the fact that any element of $\mathsf{Spin}^+(1,n)$ can be written as an even product of unit vectors with the number of timelike and spacelike vectors in the product each being even. □

3.3.5. Matching spin structures on M, Q, and N. In order to decompose spinorial quantities on $M = Q \times_R N$ into spinorial quantities on Q and N, it is of course important that there is a connection between the spinor bundles on all three spaces in the first place, and thus even before that also between the underlying spin structures. In the following we will explain how one can construct a spin structure on the warped product $M = Q \times_R N$ from spin structures on Q and N in a natural way.

Remark 3.3.5. Let us make one clarifying remark in order to avoid possible confusion. Namely, one might wonder why we do not start instead from a spin structure on M as might be more natural since we are interested in M. To this end, recall from Theorem 1.2.2 that the set of spin structures on a manifold is in bijection to its first cohomology group with $\mathbb{Z}_2$-coefficients. Since $M = Q \times N$, by the *Künneth formula* (cf. [**Spa66**, Thm. 5.11]) we have

$$\begin{aligned} H^1(M;\mathbb{Z}_2) &\cong H^0(Q;\mathbb{Z}_2) \otimes H^1(N;\mathbb{Z}_2) \oplus H^1(Q;\mathbb{Z}_2) \otimes H^0(N;\mathbb{Z}_2) \\ &\overset{(*)}{=} H^1(N;\mathbb{Z}_2) \oplus H^1(Q;\mathbb{Z}_2) \\ &\overset{(**)}{=} H^1(N;\mathbb{Z}_2)\,, \end{aligned}$$

where the second equality $(*)$ holds if both Q and N are connected, and the third equality $(**)$ holds if Q is simply connected (as is the case for our actual spacetimes of interest (3.1.1)). This shows that spin structures on M and pairs of spin structures on Q and N are in 1:1 correspondence, so it does not matter after all from which side we start. The reason for starting from Q and N in the following is that this seemed to be more straight-forward to explain.

First, from the given spin structures $(\mathsf{Spin}^+(Q), \Theta_Q)$ on Q and $(\mathsf{Spin}(N), \Theta_N)$ on N we construct on M the $\mathsf{Spin}^+(1,1) \times \mathsf{Spin}(n-1)$-principal bundle

$$\mathsf{Spin}^+(M; Q \times_R N) := \widetilde{\mathsf{Spin}^+(Q)} \times \widetilde{\mathsf{Spin}(N)}\,. \tag{3.3.10}$$

Moreover, we define a map $\Theta_M : \mathsf{Spin}^+(M; Q \times_R N) \to \mathsf{SO}^+(M; Q \times_R N)$ by

$$\Theta_M(s_Q, s_N) := \big(\Theta_Q(s_Q), \tfrac{1}{R}\Theta_N(s_N)\big)\,. \tag{3.3.11}$$

Clearly then $(\mathsf{Spin}^+(M; Q \times_R N), \Theta_M)$ is a reduction of $\mathsf{SO}^+(M; Q \times_R N)$ with respect to $\vartheta_{1,1} \times \vartheta_{n-1}$, i.e. the following diagram commutes:

$$\begin{tikzcd} \mathsf{Spin}^+(M; Q \times_R N) \times (\mathsf{Spin}^+(1,1) \times \mathsf{Spin}(n-1)) \arrow[r, "\cdot"] \arrow[d, "\Theta_M \times (\vartheta_{1,1} \times \vartheta_{n-1})"] & \mathsf{Spin}^+(M; Q \times_R N) \arrow[d, "\Theta_M"] \arrow[dr] & \\ & & M \\ \mathsf{SO}^+(M, Q \times_R N) \times (\mathsf{SO}^+(1,1) \times \mathsf{SO}(n-1)) \arrow[r, "\cdot"] & \mathsf{SO}^+(M; Q \times_R N) \arrow[ur] & \end{tikzcd} \tag{3.3.12}$$

While this is not quite a spin structure on M, it is the reduction of a spin structure with respect to the homomorphism $\varphi : \mathsf{Spin}^+(1,1) \times \mathsf{Spin}(n-1) \to \mathsf{Spin}^+(1,n)$ described in Lemma 3.3.3. Namely, we can construct a $\mathsf{Spin}^+(1,n)$-principal bundle on M by replacing the fiber of $\mathsf{Spin}^+(M; Q\times_R N)$ by φ, i.e. we set [6]

$$\mathsf{Spin}^+(M) := \mathsf{Spin}^+(M; Q\times_R N) \times_\varphi \mathsf{Spin}^+(1,n)\,. \tag{3.3.13}$$

The natural right-action of $\mathsf{Spin}^+(1,n)$ on itself turns this into a $\mathsf{Spin}^+(1,n)$-principal bundle, and the map from $\mathsf{Spin}^+(M; Q\times_R N)$ to $\mathsf{Spin}^+(M)$ given by $s\mapsto [s,1]$ is a reduction with respect to φ. The proof of these statements is rather straight-forward and can be found, e.g., in [**Bau09**, Satz 2.18]. As a matter of fact, since φ is injective, we may regard $\mathsf{Spin}^+(M; Q\times_R N)\subset \mathsf{Spin}^+(M)$ as a submanifold.

We claim that $\mathsf{Spin}^+(M)$ is actually a spin structure on M in a natural sense. To this end, we need to extend the map Θ_M from $\mathsf{Spin}^+(M; Q\times_R N)$ to all of $\mathsf{Spin}^+(M)$. This is done as follows: For $p\in M$, take some element $s\in \mathsf{Spin}^+(M; Q\times_R N)|_p \subset \mathsf{Spin}^+(M)|_p$, and set $e=\theta_M(s)\in \mathsf{SO}^+(M; Q\times_R N)|_p\subset \mathsf{SO}^+(M)|_p$. This induces isomorphisms

$$\begin{aligned} \mathsf{Spin}^+(M; Q\times_R N)|_p &\cong \mathsf{Spin}^+(1,1)\times \mathsf{Spin}(n-1)\,,\\ \mathsf{Spin}^+(M)|_p &\cong \mathsf{Spin}^+(1,n)\\ \mathsf{SO}^+(M; Q\times_R N)|_p &\cong \mathsf{SO}^+(1,1)\times \mathsf{SO}(n-1)\,,\\ \mathsf{SO}^+(M)|_p &\cong \mathsf{SO}^+(1,n)\,. \end{aligned}$$

Using these isomorphisms, we simply define a map

$$\mathsf{Spin}^+(M)|_p \cong \mathsf{Spin}^+(1,n) \xrightarrow{\vartheta_{1,n}} \mathsf{SO}^+(1,n)\cong \mathsf{SO}^+(M)|_p\,,$$

and since everything here is equivariant, this does not depend on the particular choice of the element $s\in \mathsf{Spin}^+(M; Q\times_R N)|_p$. Clearly in this way we obtain an extension of Θ_M to a map from $\mathsf{Spin}^+(M)$ to $\mathsf{SO}^+(M)$. Moreover, choosing a local section of $\mathsf{Spin}^+(M; Q\times_R N)$ instead of an element at a fixed point $p\in M$ and repeating the previous arguments, one sees that Θ_M is smooth. Finally, it is clear by construction that $(\mathsf{Spin}^+(M),\Theta_M)$ is a spin structure.

Definition 3.3.6. Given spin structures $(\mathsf{Spin}^+(Q),\Theta_Q)$ and $(\mathsf{Spin}(N),\Theta_N)$, we call the spin structure $(\mathsf{Spin}^+(M),\Theta_M)$ constructed above the *associated product spin structure.*

Before moving on, let us note that

$$\mathsf{Spin}^+(M; Q\times_R N) = \{s\in \mathsf{Spin}^+(M)\mid \Theta_M(s)\in \mathsf{SO}^+(M; Q\times_R N)\}\,. \tag{3.3.14}$$

Just as $\mathsf{SO}^+(M; Q\times_R N)$ consists of orthonormal tangent frames where the first frame vectors are tangent to Q and the last ones to N, one may think of $\mathsf{Spin}^+(M; Q\times_R N)$ as consisting of spin frames for which the first frame vectors are “tangent” to Q and the last ones to N. This is what was meant in the introduction to this section under the slogan “convenient choice of spin frame”.

3.3.6. Consequences for the spinor bundle. Suppose now that $M = Q\times_R N$ is equipped with a product spin structure as described in the previous section. In this case the associated spinor bundle on M and its geometric structures can be expressed in terms of the corresponding quantities on Q and N.

Proposition 3.3.7. *The spinor bundle $\mathbb{S}M$ is naturally isomorphic to $\widetilde{\mathbb{S}Q}\otimes\widetilde{\mathbb{S}N}$. Moreover, with respect to this isomorphism the following holds:*

[6] Here we use the standard language for associated bundles, see for instance [**Bau09**, Sec. 2.3].

i.) Clifford multiplication by vectors $v \in T_q Q$ *and* $w \in T_\omega N$ *is given by*

$$\gamma(v) = \gamma^Q(v) \otimes 1\,, \qquad \gamma(w) = R(q) \cdot \gamma(e_1)\gamma(e_0)(1 \otimes \gamma^N(w))\,. \tag{3.3.15}$$

Here $e_0, e_1 \in T_q Q$ *is an arbitrary positively oriented orthonormal basis, and* R *is the warping function.*

ii.) For any $\psi = \psi_1 \otimes \psi_2, \phi = \phi_1 \otimes \phi_2 \in \mathbb{S}_p M \cong \widetilde{\mathbb{S}Q}|_p \times \widetilde{\mathbb{S}N}|_p$ *we have*

$$\prec \psi, \phi \succ_{\mathbb{S}M} = \prec \psi_1, \phi_1 \succ_{\mathbb{S}Q} \prec \psi_2, \phi_2 \succ_{\mathbb{S}N}\ . \tag{3.3.16}$$

Proof. By [**Bau09**, Satz 2.17], or as is easily checked by hand, the following holds: Suppose P is G-principal bundle, Q is reduction of P with respect to a homomorphism $\lambda : H \to Q$, and $\rho : G \to \mathsf{GL}(V)$ is a representation. Then the vector bundles $P \times_\rho V$ and $Q \times_{\rho\circ\lambda} V$ are isomorphic.

In our context, this means that

$$\begin{aligned}
\mathbb{S}M &= \mathsf{Spin}^+(M) \times_{\rho_{1,n}} \mathbb{S}_{1,n} \\
&\cong \mathsf{Spin}^+(M; Q \times_R N) \times_{\rho_{1,n}\circ\varphi} \mathbb{S}_{1,n} \\
&\cong \mathsf{Spin}^+(M; Q \times_R N) \times_{\rho_{1,1}\otimes\rho_{n-1}} (\mathbb{S}_{1,1} \times \mathbb{S}_{n-1}) \\
&\cong (\widetilde{\mathsf{Spin}^+(Q)} \times_{\rho_{1,1}} \mathbb{S}_{1,1}) \otimes (\widetilde{\mathsf{Spin}(N)} \times_{\rho_{n-1}} \mathbb{S}_{n-1}) \\
&\cong \widetilde{\mathbb{S}Q} \otimes \widetilde{\mathbb{S}N}\,.
\end{aligned}$$

The claim about Clifford multiplication follows immediately from (3.3.5) and the one about the inner product from (3.3.9). The appearance of the warping function is due to the fact that we always have to rescale tangents vectors of N in order to normalize correctly. □

In computations it quickly happens that one overlooks the rescaling by R^2 between the metrics g and g_N on $\widetilde{TN}$, which for instance manifests itself in formula (3.3.15) for Clifford multiplication. Therefore let us record one immediate consequence of (3.3.15).

Corollary 3.3.8. *Let* $p = (q, n) \in M$ *and let* $(E_0, E_1, E_2, \dots, E_n) \in \mathsf{SO}^+(M; Q \times_R N)|_p$ *be an adapted* g*-orthonormal basis. Let* $(e_2 = R(p)E_2, \dots, e_n = R(p)E_n) \in \mathsf{SO}(N)|_n$ *be the corresponding* g_N*-orthonormal basis. Denoting Clifford multiplication by* E_μ *on* $\mathbb{S}M$ *by* γ_μ*, and Clifford multiplication by* e_j *on* $\mathbb{S}N$ *by* γ_j^N*, it holds that*

$$\gamma_j = \gamma_1\gamma_0\gamma_j^N\,. \tag{3.3.17}$$

Next, we also want to find expressions for the spin connection and the Dirac operator which are adapted to the isomorphism $\mathbb{S}M \cong \widetilde{\mathbb{S}Q} \otimes \widetilde{\mathbb{S}N}$. To this end, we first need to make an observation concerning differentiation of sections of $\mathbb{S}M$.

In view of the isomorphism $\mathbb{S}M \cong \widetilde{\mathbb{S}Q} \otimes \widetilde{\mathbb{S}N}$, any $\psi \in \Gamma^\infty(\mathbb{S}M)$ can expressed as

$$\psi = \sum_{j=1}^{k} \psi_j^Q \otimes \psi_j^N \quad \text{with} \quad \psi_j^Q \in \Gamma^\infty(\widetilde{\mathbb{S}Q})\,, \psi_j^N \in \Gamma^\infty(\widetilde{\mathbb{S}N})\,. \tag{3.3.18}$$

Furthermore, sections of $\widetilde{\mathbb{S}Q}$ are nothing but sections of $\mathbb{S}Q$ which depend additionally on the "parameters" $n \in N$. Therefore, smooth sections of $\widetilde{\mathbb{S}Q}$ can simply be differentiated (fiberwise) in the "N-directions". Concretely, if $\phi \in \Gamma^\infty(\widetilde{\mathbb{S}Q})$ then for any fixed $q \in Q$, we can regard the map

$$N \ni n \longmapsto \phi(q, n) \in \mathbb{S}_q Q\,.$$

as an $\mathbb{S}_q Q$-valued smooth function on N. As such it may be differentiated just like any smooth function on N with values in a vector space. For $V \in \Gamma^\infty(TN)$, we will use the notation $\partial_V^N \phi \in \Gamma^\infty(\widetilde{\mathbb{S}Q})$ for this derivative, i.e.

$$(\partial_V^N \phi)(q,n) := (\mathrm{d}\phi(q,\cdot))(V|_n) \qquad \forall (q,n) \in M = Q \times N\,. \tag{3.3.19}$$

Similarly one can differentiate sections of $\widetilde{\mathbb{S}N}$ in Q-directions, and for $\phi \in \Gamma^\infty(\widetilde{\mathbb{S}N})$ and $X \in \Gamma^\infty(TN)$ we use the notation $\partial_X^Q \phi \in \Gamma^\infty(\widetilde{\mathbb{S}N})$ for the section defined analogously to (3.3.19).

With this observation in mind, the spin connection of $\mathbb{S}M$ can be expressed as follows.

Lemma 3.3.9. *In view of the identification $\mathbb{S}M \cong \widetilde{\mathbb{S}Q} \otimes \widetilde{\mathbb{S}N}$, the spin connection on M satisfies the following properties:*

i.) If $X \in \Gamma^\infty(TQ)$, then

$$\nabla_X^{\mathbb{S}M} = \nabla_X^{\mathbb{S}Q} \otimes 1 + 1 \otimes \partial_X^N\,. \tag{3.3.20}$$

ii.) If $Y \in \Gamma^\infty(TN)$, then

$$\nabla_Y^{\mathbb{S}M} = 1 \otimes \nabla_Y^{\mathbb{S}N} + \partial_Y^Q \otimes 1 + \frac{1}{2R}\gamma(\operatorname{grad} R)\gamma(Y)\,. \tag{3.3.21}$$

Proof. We will use the local formula (1.2.17) for the spin connection. To this end, let $E_0, E_1 \in \Gamma^\infty(TQ|_{U_Q})$ and $e_2, \dots e_n \in \Gamma^\infty(TN|_{U_N})$ be local orthonormal frames. Then $E_0, E_1, E_2 = R^{-1}e_2, \dots, E_n = R^{-1}e_n \in \Gamma^\infty(TM|_U)$ is a local orthonormal frame on $U = U_N \times U_Q \subset M$. Necessarily shrinking U_Q and U_N we may assume that U is simply connected. Then the local section $E = (E_0, \dots, E_n) : U \to \mathsf{SO}^+(M; Q \times_R N)$ lifts to a local section $s : U \to \mathsf{Spin}^+(M; Q \times_R N)$. By the local formula (1.2.17) for the spin connection we now have for any $X \in \Gamma^\infty(TM)$

$$\nabla_X^{\mathbb{S}M} \psi|_U = (\mathrm{d}\psi)(X) + \frac{1}{4}\sum_{\mu,\nu=0}^{n} \epsilon_\mu \epsilon_\nu \left\langle \nabla_X^M E_\mu, E_\nu \right\rangle_M \gamma_\mu \gamma_\nu \psi\,. \tag{$*$}$$

Here the first term has to be understood componentwise for the components of $\psi|_U$ with respect to the spin frame s. In the second term we have abbreviated $\gamma(E_\mu)$ by γ_μ.

Now let first $X \in \Gamma^\infty(TQ)$. Then it holds that (cf. [**O'N83**, Prop. 7.35])

$$\nabla_X^M Y = \nabla_X^Q Y \quad \forall Y \in \Gamma^\infty(TQ) \qquad \text{and} \qquad \nabla_X^M V = \frac{\mathrm{d}R(X)}{R} V \quad \forall V \in \Gamma^\infty(TN)\,.$$

We need to compute $\nabla_X^M E_j$ for $j = 2, \dots, n$. Here one has to be careful that the e_j are sections of TN, whereas the $E_j = R^{-1}e_j$ are *not* since R is a function on Q. Therefore we first need to apply the Leibniz rule in the following computation, which yields

$$\nabla_X^M E_j = \nabla_X^M \left(\frac{1}{R} e_j\right) = -\frac{\mathrm{d}R(X)}{R^2} e_j + \frac{1}{R}\nabla_X^M e_j = -\frac{\mathrm{d}R(X)}{R^2} e_j + \frac{1}{R}\frac{\mathrm{d}R(X)}{R} e_j = 0\,.$$

It follows that

$$\frac{1}{4}\sum_{\mu,\nu=0}^{n} \epsilon_\mu \epsilon_\nu \left\langle \nabla_X^M E_\mu, E_\nu \right\rangle_M \gamma_\mu \gamma_\nu = \frac{1}{4}\sum_{\alpha,\beta=0}^{1} \epsilon_\alpha \epsilon_\beta \langle \nabla_X^Q E_\alpha, E_\beta \rangle_Q \gamma_\alpha \gamma_\beta\,.$$

Formula (3.3.20) now follows from this, $(*)$, and noting that $\gamma_\alpha = \gamma_\alpha^Q \otimes 1$ for $\alpha, \beta = 0, 1$.

Next, let $Y \in \Gamma^\infty(TN)$. Then it holds that (cf. [**O'N83**, Prop. 7.35])

$$\nabla_Y^M X = \frac{\mathrm{d}R(X)}{R} Y \quad \forall X \in \Gamma^\infty(TQ)\,,$$

whereas

$$\tan_Q(\nabla^M_Y V) = \nabla^N_Y V\,, \quad \mathrm{nor}_Q(\nabla^M_Y V) = -\frac{\langle Y, V\rangle}{R}\,\mathrm{grad}\,R \quad \forall V \in \Gamma^\infty(TN)\,.$$

Therefore it follows that

$$\begin{aligned}
\frac{1}{4}\sum_{\mu,\nu=0}^{n} \epsilon_\mu\epsilon_\nu \left\langle \nabla^M_Y E_\mu, E_\nu\right\rangle_M \gamma_\mu\gamma_\nu &= -\frac{1}{2}\sum_{j=2}^{n}\left\langle \nabla^M_Y E_0, E_j\right\rangle_M \gamma_0\gamma_j \\
&\quad + \frac{1}{2}\sum_{j=2}^{n}\left\langle \nabla^M_Y E_1, E_j\right\rangle_M \gamma_1\gamma_j \\
&\quad + \frac{1}{4}\sum_{j,k=2}^{n}\left\langle \nabla^M_Y E_j, E_k\right\rangle_M \gamma_j\gamma_k \\
&= -\frac{\mathrm{d}R(E_0)}{R}\gamma_0\sum_{j=2}^{n}\left\langle Y, E_j\right\rangle_M \gamma_j \\
&\quad + \frac{\mathrm{d}R(E_1)}{R}\gamma_1\sum_{j=2}^{n}\left\langle Y, E_j\right\rangle_M \gamma_j \\
&\quad + \frac{1}{4}\sum_{j,k=2}^{n}\left\langle \nabla^N_Y e_j, e_k\right\rangle_N \gamma^N_j\gamma^N_k \\
&= \frac{1}{2R}\gamma(\mathrm{grad}\,R)\gamma(Y) + \frac{1}{4}\sum_{j,k=2}^{n}\left\langle \nabla^N_Y e_j, e_k\right\rangle_N \gamma^N_j\gamma^N_k\,.
\end{aligned}$$

Here, in the second step, we used the two additional facts that

$$\left\langle \nabla^M_Y E_j, E_k\right\rangle_M = \langle \frac{1}{R}\nabla^M_Y e_j, \frac{1}{R}e_k\rangle_M = \frac{1}{R^2}\left\langle \nabla^N_Y e_j, e_k\right\rangle_M = \left\langle \nabla^N_Y e_j, e_k\right\rangle_N\,,$$

and that due to Corollary 3.3.8 we have

$$\gamma_j\gamma_k = \gamma_1\gamma_0\gamma^N_j\gamma_1\gamma_0\gamma^N_k = \gamma_1\gamma_0\gamma_1\gamma_0\gamma^N_j\gamma^N_k = -\gamma_1^2\gamma_0^2\gamma^N_j\gamma^N_k = \gamma^N_j\gamma^N_k\,.$$

Formula (3.3.21) now follows together with $(*)$. □

Before we state the following consequent formula for the Dirac operator we introduce another notation. As one easily verifies, for any $\psi \in \Gamma^\infty(\mathbb{S}M)$ the map

$$(\gamma\otimes\partial^Q)\psi : \Gamma^\infty(T^*M\otimes TM) \longrightarrow \Gamma^\infty(\mathbb{S}M)\,, \quad (\omega, X) \longmapsto (\gamma(\omega)\otimes\partial^Q_X)\psi$$

is actually tensorial, i.e. $C^\infty(Q)$-linear. Therefore it makes sense to form the contraction

$$(\gamma^\alpha\otimes\partial^Q_\alpha)\psi|_p := (\gamma(\omega_0)\otimes\partial^Q_{X_0})\psi|_p + (\gamma(\omega_1)\otimes\partial^Q_{X_1})\psi|_p\,,$$

where $X_0, X_1 \in T_pM$ is any basis with dual basis $\omega_0, \omega_1 \in T^*_pM$. Letting p vary, this yields a first-order differential operator acting on $\Gamma^\infty(\mathbb{S}M)$, which we denote by $\gamma^\alpha\otimes\partial^Q_\alpha$, keeping in mind that α sums from 0 to 1 (as the reference to Q reminds us of). Similarly one defines $\partial^N_j\otimes\gamma^{N,j}$, where j sums from 2 to n. Note that this is similar to the definition of the Dirac operator 1.2.15 only that now Clifford multiplication and differentiation act on different "factors" of the spinor with respect to the factorization $\mathbb{S}M \cong \widetilde{\mathbb{S}Q}\otimes\widetilde{\mathbb{S}N}$.

Corollary 3.3.10. *In view of the isomorphism $\mathbb{S}M \cong \widetilde{\mathbb{S}Q} \otimes \widetilde{\mathbb{S}N}$, the Dirac operator on M satisfies*

$$\mathcal{D}_M = \mathcal{D}_Q \otimes 1 + \mathrm{i}\gamma^\alpha \otimes \partial_\alpha^Q + \frac{\mathrm{i}}{R}\gamma_1\gamma_0 \otimes \mathcal{D}_N + \frac{\mathrm{i}}{R}\gamma_1\gamma_0\partial_j^N \otimes \gamma^{N,j} + \mathrm{i}\frac{n-1}{2R}\gamma(\operatorname{grad} R)\,. \tag{3.3.22}$$

Proof. As in the previous proof, we work locally in an adapted orthonormal frame $(E_0, E_1, E_2, \dots, E_n) : U \to \mathsf{SO}^+(M; Q \times_R N)$, where $E_2 = R^{-1}e_1, \dots, E_n = R^{-1}e_n$ in terms of an orthonormal frame $e_2, \dots, e_n$ of TN. Denoting Clifford multiplication with E_μ by γ_μ and Clifford multiplication on N with e_j by γ_j^N, we simply compute by using (3.3.20) and (3.3.21) that

$$\begin{aligned}
\mathcal{D}_M &= \mathrm{i}\gamma^\mu \nabla_\mu^{\mathbb{S}M} \\
&= \mathrm{i}\gamma^0 \nabla_{E_0}^{\mathbb{S}M} + \mathrm{i}\gamma^1 \nabla_{E_1}^{\mathbb{S}M} + \frac{\mathrm{i}}{R}\gamma^j \nabla_{e_j}^{\mathbb{S}M} \\
&= \mathrm{i}\gamma^{Q,0}\nabla_{E_0}^{\mathbb{S}Q} \otimes 1 + \mathrm{i}\gamma^{Q,1}\nabla_{E_1}^{\mathbb{S}Q} \otimes 1 + \mathrm{i}\gamma^0 \otimes \partial_0^Q + \mathrm{i}\gamma^1 \otimes \partial_1^Q \\
&\quad + \frac{\mathrm{i}}{R}\gamma_1\gamma_0 \otimes \gamma^{N,j}\nabla_j^{\mathbb{S}N} + \frac{\mathrm{i}}{R}\gamma_1\gamma_0\partial_j^N \otimes \gamma^{N,j} + \frac{\mathrm{i}}{R}\gamma^j \frac{1}{2R}\gamma(\operatorname{grad} R)\gamma(e_j) \\
&= \mathcal{D}_Q \otimes 1 + \frac{\mathrm{i}}{R}\gamma_1\gamma_0 \otimes \mathcal{D}_N + \mathrm{i}\gamma^\alpha \otimes \partial_\alpha^Q + \mathrm{i}\gamma_1\gamma_0\partial_j^N \otimes \gamma^{N,j} - \frac{\mathrm{i}}{2R}\gamma(\operatorname{grad} R)\gamma^j\gamma_j\,.
\end{aligned}$$

To complete the proof one only has to note that $\gamma^j\gamma_j = -(n-1)$. ☐

3.3.7. The special case of a trivial spin structure on Q. In this last section, we work out additional formulas for the case that the spin structure on Q is trivial, i.e. $\mathsf{Spin}^+(Q) \cong Q \times \mathsf{Spin}^+(1,1)$. Notice that this is always the case for our actual manifolds of interest (3.1.1).

The trivialization of $\mathsf{Spin}^+(Q)$ induces trivializations $\mathsf{SO}^+(Q) \cong Q \times \mathsf{SO}^+(1,1)$ and $\mathbb{S}Q \cong Q \times \mathbb{S}_{1,1}$. In particular, we can fix a global oriented and time-oriented orthonormal tangent frame $(E_0, E_1) : Q \to \mathsf{SO}^+(Q)$ and a basis $\Xi_1, \Xi_2 \in \mathbb{S}_{1,1}$ such that $\mathbb{S}Q \cong Q \times \mathbb{C}^2$ and such that with respect to this identification we have[7]

$$\gamma^Q(E_0) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \gamma^Q(E_1) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{3.3.23}$$

and

$$\prec\cdot,\cdot\succ_{\mathbb{S}Q} = \langle \cdot, \gamma^Q(E_0)\cdot \rangle_{\mathbb{C}^2}\,. \tag{3.3.24}$$

Passing to M, we also have $\widetilde{\mathbb{S}Q} \cong M \times \mathbb{C}^2$. Moreover, we can use this identification to write

$$\mathbb{S}M \cong \widetilde{\mathbb{S}Q} \otimes \widetilde{\mathbb{S}N} \cong \widetilde{\mathbb{S}N} \oplus \widetilde{\mathbb{S}N}\,, \tag{3.3.25}$$

where the second isomorphism is just the map

$$\Xi_1^Q \otimes \psi^N \mapsto \begin{pmatrix} \psi^N \\ 0 \end{pmatrix} \quad \text{and} \quad \Xi_2^Q \otimes \psi^N \mapsto \begin{pmatrix} 0 \\ \psi^N \end{pmatrix} \qquad \forall \psi^N \in \widetilde{\mathbb{S}N}\,. \tag{3.3.26}$$

With respect to this new representation of $\mathbb{S}M$, Clifford multiplication etc. are of course given by slightly different formulas. The following Proposition gathers the results, the calculations are all straight-forward applications of (3.3.26) to the previously obtained formulas.

[7] Cf. Example 1.1.5 about concrete realizations of $\mathrm{Cl}(1,1)$.

Proposition 3.3.11. *Suppose that* $\mathsf{Spin}^+(Q) \cong Q \times \mathsf{Spin}^+(1,1)$*, and let* $(E_0, E_1) : Q \to \mathsf{SO}^+(Q)$ *be a corresponding global orthonormal frame. Then with respect to the isomorphism* $\mathbb{S}M \cong \widetilde{\mathbb{S}N} \oplus \widetilde{\mathbb{S}N}$ *from above, the following hold.*

i.) Clifford multiplication by E_0 *and* E_1 *on* $\mathbb{S}M$ *are given by*

$$\gamma(E_0) = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \quad \gamma(E_1) \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}. \tag{3.3.27}$$

Furthermore, Clifford multiplication by a vector $v \in TN$ *on* $\mathbb{S}M$ *is given by*

$$\gamma(v) = R \begin{pmatrix} 0 & \gamma^N(v) \\ \gamma^N(v) & 0 \end{pmatrix}. \tag{3.3.28}$$

ii.) The inner product $\prec\cdot,\cdot\succ_{\mathbb{S}M}$ *takes the form*

$$\prec\Psi, \Phi\succ_{\mathbb{S}M} = -\prec\psi_1, \phi_1\succ_{\mathbb{S}N} + \prec\psi_2, \phi_2\succ_{\mathbb{S}N} \tag{3.3.29}$$

for all $\Psi = (\psi_1, \psi_2), \Phi = (\phi_1, \phi_2) \in \mathbb{S}M \cong \widetilde{\mathbb{S}N} \oplus \widetilde{\mathbb{S}N}$.

iii.) For any $X \in \Gamma^\infty(TQ)$ *we have*

$$\nabla^{\mathbb{S}M}_X = \partial^Q_X + \frac{1}{2}\left\langle \nabla^Q_X E_0, E_1 \right\rangle \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \tag{3.3.30}$$

For any $V \in \Gamma^\infty(TN)$ *we have*

$$\nabla^{\mathbb{S}M}_V = \nabla^{\mathbb{S}N}_V + \frac{\mathrm{d}R(E_0)}{R} \begin{pmatrix} 0 & \gamma^N(V) \\ -\gamma^N(V) & 0 \end{pmatrix} + \frac{\mathrm{d}R(E_1)}{R} \begin{pmatrix} \gamma^N(V) & 0 \\ 0 & -\gamma^N(V) \end{pmatrix}. \tag{3.3.31}$$

iv.) The Dirac operator has the form

$$\begin{aligned} \mathcal{D}_M = \;& \mathrm{i}\begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \left(\partial^Q_{E_0} + \frac{1}{2}\left\langle \nabla^Q_{E_1} E_0, E_1 \right\rangle + \mathrm{i}\frac{n-1}{2}\frac{\mathrm{d}R(E_0)}{R} \right) \\ & + \mathrm{i}\begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \left(\partial^Q_{E_1} + \frac{1}{2}\left\langle \nabla^Q_{E_0} E_0, E_1 \right\rangle + \mathrm{i}\frac{n-1}{2}\frac{\mathrm{d}R(E_1)}{R} \right) \\ & + \frac{\mathrm{i}}{R}\begin{pmatrix} 0 & \mathcal{D}_N \\ \mathcal{D}_N & 0 \end{pmatrix}. \end{aligned} \tag{3.3.32}$$

We conclude this section by noting that on $\mathbb{S}M$ we can also introduce the *positive definite* inner product

$$\langle\cdot,\cdot\rangle_{E_0} := \prec\cdot, \gamma(E_0)\cdot\succ_{\mathbb{S}M}. \tag{3.3.33}$$

Due to (3.3.29) and (3.3.27), for any $\Psi = (\psi_1, \psi_2), \Phi = (\phi_1, \phi_2) \in \mathbb{S}M \cong \widetilde{\mathbb{S}N} \oplus \widetilde{\mathbb{S}N}$ we explicitly have

$$\langle\Psi, \Phi\rangle_{E_0} = \prec\psi_1, \phi_1\succ_{\mathbb{S}N} + \prec\psi_2, \phi_2\succ_{\mathbb{S}N}. \tag{3.3.34}$$

This inner product of course depends on the choice of E_0. Nevertheless it is important since it is positive definite.

3.3.8. The Dirac operator for our actual class of spacetimes. Finally, let us concretely take a spacetime (M, g) of the form (3.1.1). We always work with the global orthonormal frame

$$E_0 := \mathrm{e}^{-a(r)}\partial_t, \quad E_1 := \mathrm{e}^{-a(r)}\partial_r \tag{3.3.35}$$

and use it to trivialize $\mathsf{SO}^+(Q)$ and $\mathsf{Spin}^+(Q)$. In order to evaluate formula (3.3.32) for the Dirac operator completely, we have to compute the zero-order terms in the first two lines of (3.3.32). To this end, we first have

$$\mathrm{d}R(E_0) = \mathrm{e}^{-a}\partial_t R = 0, \quad \mathrm{d}R(E_1) = \mathrm{e}^{-a}\partial_r R = \mathrm{e}^{-a} R',$$

where we have written R' for the derivative of R. For the terms involving the Levi-Civita connection of Q, we use the well-known Koszul formula

$$2\langle\nabla^Q_X Y, Z\rangle = X\,\langle Y,Z\rangle + Y\,\langle X,Z\rangle - Z\,\langle X,Y\rangle + \langle[X,Y],Z\rangle + \langle[Z,X],Y\rangle + \langle[Z,Y],X\rangle\,,$$

which holds for all $X,Y,Z \in \Gamma^\infty(TQ)$. To evaluate this formula we first compute

$$[E_1,E_0] = [\mathrm{e}^{-a}\partial_r, \mathrm{e}^{-a}\partial_t] = \mathrm{e}^{-a}(\partial_r \mathrm{e}^{-a})\partial_t = (\partial_r \mathrm{e}^{-a})E_0\,,$$

from which it now follows that

$$\langle\nabla^Q_{E_1}E_0, E_1\rangle = \langle[E_1,E_0],E_1\rangle = 0\,, \quad \langle\nabla^Q_{E_0}E_0, E_1\rangle = \langle[E_1,E_0],E_0\rangle = a'\mathrm{e}^{-a}\,. \tag{3.3.36}$$

Plugging this into (3.3.32), we end up with

$$\boxed{\mathcal{D}_M = \mathrm{i}\mathrm{e}^{-a}\begin{pmatrix}\mathbb{1} & 0\\ 0 & -\mathbb{1}\end{pmatrix}\partial_t + \mathrm{i}\mathrm{e}^{-a}\begin{pmatrix}0 & \mathbb{1}\\ -\mathbb{1} & 0\end{pmatrix}\left(\partial_r + \frac{a'}{2} + \frac{n-1}{2}\frac{R'}{R}\right) + \frac{\mathrm{i}}{R}\begin{pmatrix}0 & \mathcal{D}_N\\ \mathcal{D}_N & 0\end{pmatrix}} \tag{3.3.37}$$

Remark 3.3.12. In concrete cases, i.e. with N given concretely, one can of course also explicitly compute the Dirac operator (in a local frame). For instance, if $N = S^{n-1}$ this is certainly possible, see, e.g., [**Nic95**, Sec. 2]. However, while these concrete computations are certainly more elementary than the abstract approach employed here, they can get quite messy. In particular, we did not have to worry about how to concretely pick a frame and coordinates on N. Moreover, the appearance of the Dirac operator of N in (3.3.37) is much clearer in the abstract approach.

3.4. Further Computations with the Dirac Equation on these Spacetimes

In the following we first explain how one can further simplify the Dirac equation on a spacetime (M,g) of the form (3.1.1) using some simple conformal transformations. Afterwards we define certain inner products on spinors on M related to the conserved current, which will be used in the following chapter.

Throughout, we assume that (M,g) has the form (3.1.1) and we assume that N is spin. Then also M is spin and we choose the spin structures on M and N in a compatible way as explained in Section 3.3.5, see also Remark 3.3.5.

3.4.1. Conformal rescaling of spinor fields. It is a useful observation that one can further simplify the expression (3.3.37) of the Dirac operator by a suitable rescaling of spinor fields. As the title of this section suggests, this rescaling is related to a conformal transformation, cf. Remark 3.4.2.

As for the actual computation, for $\psi \in \Gamma^\infty(\mathbb{S}M)$ we write

$$\psi = \mathrm{e}^{-\mu a}R^\nu\phi \tag{3.4.1}$$

for certain $\mu,\nu \in \mathbb{R}$ and some other section $\phi \in \Gamma^\infty(\mathbb{S}M)$. Then we have

$$\begin{aligned}\partial_r\psi &= -\mu a'\mathrm{e}^{-\mu a}R^\nu\phi + \mathrm{e}^{-\mu a}\nu R'R^{\nu-1}\phi + \mathrm{e}^{-\mu a}R^\nu\partial_r\phi\\ &= -\mu a'\psi + \nu\frac{R'}{R}\psi + \mathrm{e}^{-\mu a}R^\nu\partial_r\phi\,,\end{aligned}$$

which can be reformulated as

$$\left(\partial_r + \mu a' - \nu\frac{R'}{R}\right)\psi = \mathrm{e}^{-\mu a}R^\nu\partial_r\phi\,. \tag{3.4.2}$$

Comparing this to the expression (3.3.37) of the Dirac operator shows that if we put $\mu = \frac{1}{2}$ and $\nu = -\frac{n-1}{2}$, so that

$$\psi = \mathrm{e}^{-\frac{a}{2}} R^{-\frac{n-1}{2}} \phi \,, \tag{3.4.3}$$

then

$$\mathcal{D}\psi = \mathrm{e}^{-\frac{a}{2}} R^{-\frac{n-1}{2}} \left[\mathrm{i}\mathrm{e}^{-a} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \partial_t + \mathrm{i}\mathrm{e}^{-a} \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \partial_r + \frac{\mathrm{i}}{R} \begin{pmatrix} 0 & \mathcal{D}_N \\ \mathcal{D}_N & 0 \end{pmatrix} \right] \phi \,. \tag{3.4.4}$$

We summarize this finding in the following statement.

Lemma 3.4.1. *Let $\psi, \phi \in \Gamma^\infty(\mathbb{S}M)$ be related by* (3.4.3)*. Then ψ satisfies $(\mathcal{D}-m)\psi = 0$ if and only if ϕ satisfies*

$$\boxed{\left[\mathrm{i} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \partial_t + \mathrm{i} \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \partial_r + \mathrm{i}\frac{\mathrm{e}^a}{R} \begin{pmatrix} 0 & \mathcal{D}_N \\ \mathcal{D}_N & 0 \end{pmatrix} - \mathrm{e}^a m \right] \phi = 0} \,. \tag{3.4.5}$$

Proof. This follows immediately from the previous computations. □

Remark 3.4.2. Let $(M^{1,n}, g)$ be a Lorentzian spin manifold, and $\widetilde{g} = \mathrm{e}^{2f} g$ a conformally related metric. Then the Dirac operators for g and $\widetilde{g}$ are related by the identity

$$\mathcal{D} = \mathrm{e}^{\frac{n}{2} f} \circ \widetilde{\mathcal{D}} \circ \mathrm{e}^{-\frac{n}{2} f} \,, \tag{3.4.6}$$

where $\mathrm{e}^{\frac{n}{2} f}$ and $\mathrm{e}^{-\frac{n}{2} f}$ stand for the corresponding multiplication operators with these functions (cf. [**LM89**, Thm. 5.24] for the Riemannian case). This is precisely what happens in our situation: Writing $g = R^2 \big(R^{-2} \mathrm{e}^{2a} [-\,\mathrm{d}t^2 + \mathrm{d}r^2] + g_N \big)$, we see that g is conformally related to $R^{-2}\mathrm{e}^{2a}[-\,\mathrm{d}t^2 + \mathrm{d}r^2] + g_N$, and moreover the first part of this metric is conformal to the Minkowski metric. So in total we make two conformal transformations. If one carefully examines the correct powers of the identity (3.4.6) in these cases, one obtains precisely the rescaling (3.4.3).

3.4.2. Inner products on the spinor bundle. As we have seen in Section 3.3, in view of the isomorphism $\mathbb{S}M \cong \widetilde{\mathbb{S}N} \oplus \widetilde{\mathbb{S}N}$ the natural Hermitian inner product $\prec \cdot, \cdot \succ_{\mathbb{S}M}$ on the spinor bundle is given by

$$\prec \Psi, \Phi \succ_{\mathbb{S}M} = - \prec \psi_1, \phi_1 \succ_{\mathbb{S}N} + \prec \psi_2, \phi_2 \succ_{\mathbb{S}N} \tag{3.4.7}$$

for all $\Psi = (\psi_1, \psi_2), \Phi = (\phi_1, \phi_2) \in \mathbb{S}M$. Here $\prec \cdot, \cdot \succ_{\mathbb{S}N}$ is the positive definite inner product on $\mathbb{S}N$. Note again that since $\prec \cdot, \cdot \succ_{\mathbb{S}N}$ is positive definite (N is Riemannian), it follows that that $\prec \cdot, \cdot \succ_{\mathbb{S}M}$ has split signature. Therefore it is not well suited to measure the "size" of a spinor, for instance in decay estimates.

As was also already remarked before, however, for any future-directed timelike vector field $T \in \Gamma^\infty(TM)$ the inner product

$$\langle \cdot, \cdot \rangle_T := \prec \cdot, \gamma(T) \cdot \succ_{\mathbb{S}M} \tag{3.4.8}$$

is positive definite (cf. Lemma 1.1.23). Of course this construction depends on the choice of T, but in our specific setup there exists at least a natural choice: We will always take the normalized static vector field

$$Z = E_0 = \mathrm{e}^{-a} \partial_t \,. \tag{3.4.9}$$

We have already seen in Section 3.3 that then we have

$$\langle \Psi, \Phi \rangle_Z = \prec \psi_1, \phi_1 \succ_{\mathbb{S}N} + \prec \psi_2, \phi_2 \succ_{\mathbb{S}N} \tag{3.4.10}$$

for all $\Psi = (\psi_1, \psi_2), \Phi = (\phi_1, \phi_2) \in \mathbb{S}M \cong \widetilde{\mathbb{S}N} \oplus \widetilde{\mathbb{S}N}$.

This inner product in fact appears naturally as contraction of the current vector field $J[\psi]^\mu = \prec \psi, \gamma^\mu \psi \succ_{\mathbb{S}M}$ of a spinor field $\psi \in \Gamma^\infty(\mathbb{S}M)$ with the future-pointing timelike unit normal to any of the spacelike hypersurfaces $\Sigma_t = \{t\} \times (r_0, \infty)_r \times N \subset M$. Indeed, this unit normal is precisely given by $E_0 = \mathrm{e}^{-a}\partial_t$, so that

$$\langle J[\psi], E_0 \rangle = \prec \psi, \gamma(E_0)\psi \succ_{\mathbb{S}M} \ .$$

Moreover, motivated by current conservation (cf. Corollary 2.1.4), on the spacelike hypersurface $\Sigma = (r_0, \infty)_r \times N_\omega$ endowed with Riemannian metric $g_\Sigma = \mathrm{e}^{2a}\, \mathrm{d}r^2 + R(r)^2 g_{S^{n-1}}$ as induced by any constant t embedding into M, we introduce the L^2-inner product

$$\begin{aligned}
(\psi, \phi)_{L^2(\Sigma)} &:= \int_\Sigma \langle \psi, \phi \rangle_{E_0} \, \mathrm{d}\mu_\Sigma \\
&= \int_{r_0}^\infty \int_N \langle \psi(r,\omega), \phi(r,\omega) \rangle_{E_0} \, \mathrm{e}^{a(r)} R(r)^{n-1} \, \mathrm{d}\mu_N(\omega) \, \mathrm{d}r \\
&= \int_{r_0}^\infty (\psi(r,\cdot), \phi(r,\cdot))_{L^2(\mathbb{S}N \oplus \mathbb{S}N)} \mathrm{e}^{a(r)} R(r)^{n-1} \, \mathrm{d}r
\end{aligned} \tag{3.4.11}$$

for any $\psi, \phi \in \Gamma^\infty(\mathbb{S}M|_\Sigma)$.

It is useful to note that this L^2-inner product behaves nicely with respect to the conformal rescaling of spinors introduced in the previous section.

Lemma 3.4.3. *Let $\psi, \phi \in \Gamma^\infty(\mathbb{S}M|_\Sigma)$ be related by $\psi = \mathrm{e}^{-\frac{a}{2}} R^{-\frac{n-1}{2}} \phi$. Then we have*

$$\|\psi\|^2_{L^2(\Sigma)} = \|\phi\|^2_{L^2((r_0,\infty)\times N)} := \int_{r_0}^\infty \int_N \langle \phi(r,\omega), \phi(r,\omega) \rangle_{E_0} \, \mathrm{d}r \, \mathrm{d}\mu_N(\omega) \, . \tag{3.4.12}$$

Proof. This follows immediately from (3.4.11). $\square$

Here we introduced the following convention for integral norms on Σ, which we shall also use later: Whenever we write $L^2(\Sigma)$ we integrate with respect to the volume measure $\mathrm{d}\mu_\Sigma$ induced by any constant t embedding into M (the "physical" volume measure). Whenever we write $L^2((r_0, \infty) \times N)$ we integrate with respect to the product measure $\mathrm{d}r \, \mathrm{d}\mu_N$ (the "unphysical" volume measure).

3.4.3. Current conservation.

Recall from Section 2.1 that if our spacetime (M, g) is globally hyperbolic, then for any solution $\psi \in \Gamma^\infty_{sc}(\mathbb{S}M)$ of $(\mathcal{D} - m)\psi = 0$ the so-called *current* $\|\psi(t,\cdot)\|_{L^2(\Sigma)}$ is actually independent of t. Therefore it is also called the *conserved current* of ψ. Further, we know from Section 3.1.2 that (M, g) is globally hyperbolic if and only if $r_0 = -\infty$ and $r_1 = \infty$. Let us assume that $r_1 = \infty$. In case that $r_0 \neq -\infty$, the current will in general not be conserved but can "flow through the inner boundary at $r = r_0$". In the following Lemma, we compute this "defect" explicitly.

Lemma 3.4.4. *Let (M, g) be a spacetime of the form (3.1.1) with $r_1 = \infty$, and suppose that the functions $a(r)$ and $R(r)$ extend smoothly to $r = r_0$. Let $\psi \in \Gamma^\infty(\mathbb{S}M)$ be a solution of $(\mathcal{D}_M - m)\psi = 0$, and assume that also ψ extends to $r = r_0$. Moreover, assume that $\operatorname{supp}\psi|_{t=0} \subset (r_0, \widehat{r}_1) \times N$ for some $r_0 < \widehat{r}_1 < \infty$, so that $\|\psi(t,\cdot)\|_{L^2(\Sigma)} < \infty$ for all $t \in \mathbb{R}$ (by finite propagation speed). Then we have*

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\psi(t,\cdot)\|^2_{L^2(\Sigma)} = -\mathrm{e}^{a(r_0)} R(r_0)^{n-1} \left(\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \psi(t, r_0, \cdot) \middle| \psi(t, r_0, \cdot) \right)_{L^2(\mathbb{S}N \oplus \mathbb{S}N)} . \tag{3.4.13}$$

Proof. To simplify the computation, we use the conformal rescaling trick discussed before: If ψ satisfies the massive Dirac equation, then $\phi = \mathrm{e}^{\frac{a}{2}} R^{\frac{n-1}{2}} \psi$ satisfies the somewhat simpler rescaled Dirac equation (3.4.5). By "squaring away γ^0", this can be further rewritten as

$$\partial_t \phi = -\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \partial_r \phi - \frac{\mathrm{e}^a}{R} \begin{pmatrix} 0 & \mathcal{D}_N \\ -\mathcal{D}_N & 0 \end{pmatrix} \phi - \mathrm{i}\mathrm{e}^a m \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -1 \end{pmatrix} \phi \,. \tag{$*$}$$

Using how the L^2-inner product of spinors behaves under this rescaling (cf. Lemma 3.4.3), we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \|\psi(t,\cdot)\|^2_{L^2(\Sigma)} &= \frac{\mathrm{d}}{\mathrm{d}t} \int_{r_0}^{\infty} \int_N \langle \phi(t,\cdot), \phi(t,\cdot) \rangle_{E_0} \, \mathrm{d}r \, \mathrm{d}\mu_N \\
&= 2\,\mathsf{Re} \int_{r_0}^{\infty} \int_N \langle \partial_t \phi(t,\cdot), \phi(t,\cdot) \rangle_{E_0} \, \mathrm{d}r \, \mathrm{d}\mu_N \\
&= -2\,\mathsf{Re} \int_{r_0}^{\infty} \int_N \left\langle \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \partial_r \phi, \phi \right\rangle_{E_0} \mathrm{d}r \, \mathrm{d}\mu_N \\
&\quad - 2\,\mathsf{Re} \int_{r_0}^{\infty} \int_N \frac{\mathrm{e}^a}{R} \left\langle \begin{pmatrix} 0 & \mathcal{D}_N \\ -\mathcal{D}_N & 0 \end{pmatrix} \phi, \phi \right\rangle_{E_0} \mathrm{d}r \, \mathrm{d}\mu_N \\
&\quad - 2\,\mathsf{Re} \int_{r_0}^{\infty} \int_N m\mathrm{e}^a \left\langle \mathrm{i} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \phi, \phi \right\rangle_{E_0} \mathrm{d}r \, \mathrm{d}\mu_N \,.
\end{aligned}$$

In the last two terms the integrals are completely imaginary, so taking the real part kills these term. Indeed, for the third term this is obvious, and for the second term this follows since $\mathcal{D}_N$ is symmetric with respect to $(\cdot|\cdot)_{L^2(\mathbb{S}N)}$. In the remaining first term, we have

$$2\,\mathsf{Re} \left\langle \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \partial_r \phi, \phi \right\rangle_{E_0} = \partial_r \left\langle \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \phi, \phi \right\rangle_{E_0} .$$

Therefore, by the fundamental theorem of calculus it follows that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \|\psi(t,\cdot)\|^2_{L^2(\Sigma)} &= -\left(\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \phi(t, r_0, \cdot) \middle| \phi(t, r_0, \cdot) \right)_{L^2(\mathbb{S}N \oplus \mathbb{S}N)} \\
&= -\mathrm{e}^{a(r_0)} R(r_0)^{n-1} \left(\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \psi(t, r_0, \cdot) \middle| \psi(t, r_0, \cdot) \right)_{L^2(\mathbb{S}N \oplus \mathbb{S}N)} .
\end{aligned}$$

Note here that due to the assumptions on ψ (and a, R) there is no additional boundary term at $r = \infty$, and the boundary term at $r = r_0$ is well-defined. This concludes the proof. $\square$

3.4.4. Sobolev spaces on spacelike hypersurfaces. Besides the L^2-norm over spacelike hypersurface, we will also make use of Sobolev norms which we now describe. Let (M, g) again be of the form (3.1.1).

We again fix some time $t_1 \in \mathbb{R}$ and embed $\Sigma = (r_0, \infty) \times N$ into M via $x \mapsto (t_1, x)$. Since the metric g is t-static, any of these embeddings induces on Σ the same Riemannian metric $h_\Sigma = \mathrm{e}^{2a(r)} \, \mathrm{d}r^2 + R(r)^{-2} g_N$. Next, the isomorphism $\mathbb{S}M \cong \widetilde{\mathbb{S}N} \oplus \widetilde{\mathbb{S}N}$ induces an isomorphism $\mathbb{S}M|_\Sigma \cong \widetilde{\mathbb{S}N}|_\Sigma \oplus \widetilde{\mathbb{S}N}|_\Sigma$. By this isomorphism, sections of $\mathbb{S}M|_\Sigma$ can be identified with pairs of sections of $\mathbb{S}N$ which additionally depend on the parameter $r \in (r_0, r_1)$. Concerning differentiation of sections of $\mathbb{S}M|_\Sigma$ we simply use the induced covariant derivative, i.e. for $\psi \in \Gamma^\infty(\mathbb{S}M|_\Sigma)$ and $X \in \Gamma^\infty(T\Sigma)$ we set

$$\nabla^{\mathbb{S}\Sigma}_X \psi := \nabla^{\mathbb{S}M}_X \psi \,.$$

To be precise, on the right-hand side one should (locally) take a smooth extension of ψ and X to M and then restrict back to Σ afterwards. Since X is tangent to Σ the result does not depend on the specific choice of the extension. Furthermore, since M is t-static, everything is also independent of the choice of t_1. This can be seen explicitly from the form of the covariant derivatives in Section 3.3.7. Now we set

$$\|\psi\|^2_{H^k(\Sigma)} := \sum_{j=0}^{k} \int_\Sigma |(\nabla^{\mathbb{S}\Sigma})^k \psi|^2_{T,h} \, \mathrm{d}\mu_\Sigma \qquad \forall \psi \in \Gamma^\infty_c(\mathbb{S}M|_\Sigma)\,. \tag{3.4.14}$$

Here $(\nabla^{\mathbb{S}\Sigma})^k \psi \in \Gamma^\infty((T^*\Sigma)^{\otimes k} \otimes \mathbb{S}M|_\Sigma)$ is an iterated covariant derivative, defined inductively through

$$(\nabla^{\mathbb{S}\Sigma})^{k+1}_{X_0, X_1, \ldots, X_{k+1}} \psi := \nabla^{\mathbb{S}\Sigma}_{X_0}\big((\nabla^{\mathbb{S}\Sigma})^k_{X_1,\ldots,X_k}\psi\big) - \sum_{j=1}^{k} (\nabla^\Sigma)^k_{X_1,\ldots,\nabla_{X_0}X_j,\ldots,X_k}\psi\,, \tag{3.4.15}$$

and $|(\nabla^{\mathbb{S}\Sigma})^k\psi|_{T,h}$ is its norm defined by

$$\begin{aligned} |(\nabla^{\mathbb{S}\Sigma})^k\psi|^2_{T,h} &:= h^{i_1 j_1}_\Sigma \cdots h^{i_k j_k}_\Sigma \big\langle (\nabla^{\mathbb{S}\Sigma})^k_{i_1\ldots i_k}\psi, (\nabla^{\mathbb{S}\Sigma})^k_{j_1\ldots j_k}\psi \big\rangle_T \\ &= \sum_{i_1,\ldots,i_k=1}^{n} \big\langle (\nabla^{\mathbb{S}\Sigma})^k_{E_{i_1}\ldots E_{i_k}}\psi, (\nabla^{\mathbb{S}\Sigma})^k_{E_{i_1}\ldots E_{i_k}}\psi \big\rangle_T \,. \end{aligned} \tag{3.4.16}$$

Here the first line is to be understood in abstract index notation, and the second line is with respect to a (local) h-orthonormal frame $E_1, \ldots, E_n$.

Clearly $\|\cdot\|_{H^k(\Sigma)}$ is a norm on $\Gamma^\infty_c(\mathbb{S}M|_\Sigma)$, and the completion of $\Gamma^\infty_c(\mathbb{S}M|_\Sigma)$ is the space of Sobolev sections of $\mathbb{S}M|_\Sigma$ that vanish at the boundary/infinity of Σ. We will have no explicit need for these spaces since we will only use the norm (3.4.14) as a sort of "book-keeping device" for derivatives of smooth spinor fields. Therefore we do not discuss anything else relating to these Sobolev spaces here (but see Remark 3.4.5 for references).

The only further observation which we will make use of is the following: Suppose that $A : \Gamma^\infty(\mathbb{S}M|_\Sigma) \to \Gamma^\infty(\mathbb{S}M|_\Sigma)$ is a differential operator of order $m \in \mathbb{N}$ (with smooth coefficients). Then for any $k \in \mathbb{N}$ and any $\psi \in \Gamma^\infty_c(\mathbb{S}M|_\Sigma)$ it holds that

$$\|A\psi\|_{H^k(\Sigma)} \leq C(\operatorname{supp}\psi)\, \|\psi\|_{H^{k+m}(\Sigma)} \;, \tag{3.4.17}$$

where $C(\operatorname{supp}\psi) > 0$ is a constant which depends on the C^k-norm of the coefficients of A on $\operatorname{supp}\psi$. Clearly this follows immediately from the definition (3.4.14) of the Sobolev norms. Furthermore, if the derivatives up to order k of the coefficients of A are (globally) bounded, then clearly one can choose the constant in (3.4.17) independent of $\operatorname{supp}\psi$. Consequently, such operators are bounded operators between any pair of Sobolev spaces whose order differs by the order m of A.

Remark 3.4.5. There are a variety of ways to introduce Sobolev norms and Sobolev spaces on (Riemannian) manifolds. For compact manifolds they are basically all equivalent. In the noncompact setting this is no longer true, and also weighted Sobolev spaces are an important tool. We refer to [**Heb99**] for a textbook treatment, or to [**LM89**, Ch. 3] for the compact case. Further information about Sobolev spaces in the noncompact setting can also be found in [**Bar86**], [**CB09**, App. I], as well as in the references therein.

CHAPTER 4

Decay in Outgoing Null Directions

In this chapter we are going to analyze solutions of the massive Dirac equation in asymptotically flat spacetimes of the type introduced in Chapter 3. More precisely, the aim is to determine how smooth (spatially compactly supported) solutions decay as one moves out to infinity along outgoing null directions. Since the present chapter is the main part of the thesis, and also the longest chapter, we start with an overview.

4.1. Outline and Summary of this Chapter

The general goal is to study the behaviour of solutions of the massive Dirac equation $(\mathcal{D}{-}m)\psi = 0$ in a spacetime of the form $M = \mathbb{R}_t \times (r_0, \infty)_r \times N$, equipped with a metric of the form

$$g = \mathrm{e}^{2a(r)}[-\,\mathrm{d}t^2 + \mathrm{d}r^2] + R(r)^2 g_N \,,$$

as one moves out to infinity along outgoing radial null directions, i.e. as the outgoing radial null coordinate $v = t{+}r$ tends to infinity. More precisely, we are going to investigate how ψ decays for $v \to \infty$. In the following, we describe in more detail the particular estimate we will obtain. The geometric situation is sketched in figure 4.1 where, similar as in Penrose diagrams, every point in this diagram actually represents a copy of N.

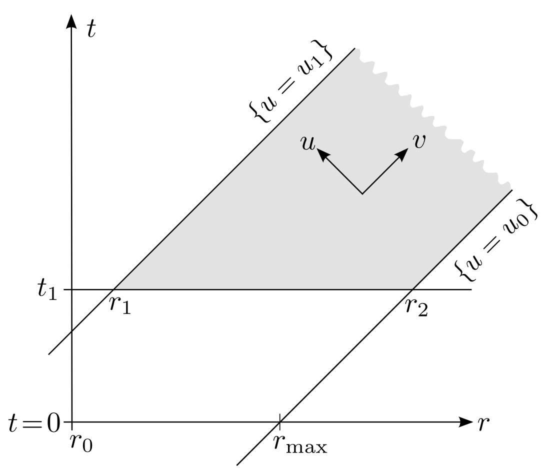

FIGURE 4.1. The geometric setup for the decay estimate in null directions.

First of all, if we want to obtain decay results to the future, we need to impose some decay properties of the initial values $\psi_0 := \psi|_{t=0}$.[1] Here we stick to the strictest possible decay condition, namely we impose that the support of ψ_0 does not extend to infinity. More precisely, we fix some $r_{\max} > r_0$ and restrict our considerations to solutions with

[1]In the massless case ($m = 0$) the necessity of such a condition is illustrated by constant solutions in flat Minkowski spacetime, which do not decay at all. In the massive case matters are more involved.

$\operatorname{supp}\psi_0 \subset (r_0, r_{\max}) \times N$. Next, set $u_0 = r_{\max}$ and fix some $u_1 > u_0$. Further, fix some $t_1 > 0$ such that $r_1 := t_1 - u_1 > r_0$, and set $r_2 := t_1 - u_0$. Figure 4.1 illustrates the various choices just described.

The estimate we are going to obtain at the end of this chapter is a decay estimate for ψ as the outgoing radial null coordinate $v = t + r$ tends to infinity, uniformly in the ingoing radial null coordinate $u = t - r$ as long as $u_0 \leq u \leq u_1$, and uniform in the N-coordinates. Or put more geometrically, we are going to show that ψ decays (in some specified way) as one moves out to infinity inside the grey-shaded strip in figure 4.1. It will turn that the decay (of suitable norms of ψ) is *superpolynomially* in v, i.e. as fast as any inverse power of v. More precisely, we are going to obtain a family of decay estimates (one for each inverse power of v). These depend on ψ through $\psi|_{t=t_1}$ and its derivatives. Notice that if M is globally hyperbolic, i.e. if $r_0 = -\infty$, one can always choose $t_1 = 0$.

Next, let us give a brief outline of the individual steps of the analysis. We start with a smooth solution $\psi \in \Gamma^\infty(\mathbb{S}M)$ of $(\mathcal{D}-m)\psi = 0$, which has the previously described support properties. Next, using the (warped) product structure (3.1.4) of the underlying spacetime, in Section 4.2 we show how by a separation of variables argument one can study ψ by studying solutions of the massive Dirac equation with a potential in a $1+1$ dimensional spacetime. Since we are interested in asymptotic behaviour in an outgoing null strip as in figure 4.1, we transform to the null coordinates v and u introduced above. In these coordinates the Dirac equation takes a particular form, see eq. (4.2.10), which we refer to as the *Dirac null system*. The analysis of this system makes up the central part of this chapter.

The general idea in studying this system is to decompose it into a "free" or "flat" part, and a part containing "curvature terms" (in a lose sense) and the influence of the N-part of the spacetime through the separation constant. The combination of these latter parts will be referred to as the "perturbation". The free part is chosen such that it has the nice feature of being explicitly solvable in terms of an integral representation (Sec. 4.4), from which one can rather easily determine decay properties of solutions of the free part alone, i.e. with the perturbation terms set to zero. The harder part then consists in showing that adding the perturbation does not change the decay properties of the free part. Here we will rely on the so-called *Lippmann-Schwinger equation* (or *Duhamel formula*) as explained in Section 4.6.1 and 4.6.2. To use this formula, we need suitable estimates of the perturbation and a priori estimates of the full solution. Such estimates will be derived in Section 4.5, basically by energy estimates in the grey region of figure 4.1, and at the end of Section 4.6 we use these estimates to establish decay of solutions of the Dirac null system as $v \to \infty$.

To finish, in Section 4.7 we show how to pass from decay properties of solutions of the Dirac null system to decay properties of the solution ψ of $(\mathcal{D}-m)\psi = 0$ we have started with. This is done by summing over the separation constant, using Plancharel type arguments.

4.2. From the Dirac Equation to the Dirac Null System

In the following, we assume that (M, g) is a spacetime of the form (3.1.1) with $r_1 = \infty$. To recall, this means that (M, g) is a warped product $M^{1,n} = Q^{1,1} \times_R N$, where

$$Q = \mathbb{R}_t \times (r_0, \infty)_r\,, \quad g_Q = \mathrm{e}^{2a(r)}[-\,\mathrm{d}t^2 + \mathrm{d}r^2]$$

with $a \in C^\infty(r_0, \infty)$, and where (N, g_N) is a connected, compact Riemannian spin manifold. The warping function R is only allowed to depend on r, i.e. $R \in C^\infty(r_0, \infty)$, and is of course required to be positive.

Choosing spin structures on M and Q in a compatible way as explained in Sec. 3.3.5), see also Remark 3.3.5, the following relations hold: First there exists an isomorphism $\mathbb{S}M \cong \widetilde{\mathbb{S}N} \oplus \widetilde{\mathbb{S}N}$, where $\widetilde{\mathbb{S}N}$ is the pullback of the spinor bundle $\mathbb{S}N$ of N by the projection $M = Q \times N \to N$. Moreover, with respect to this splitting the Dirac operator of M has the block form

$$\mathcal{D} = \mathrm{i}e^{-a} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \partial_t + \mathrm{i}e^{-a} \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \left(\partial_r + \frac{a'}{2} + \frac{n-1}{2} \frac{R'}{R} \right) + \frac{\mathrm{i}}{R} \begin{pmatrix} 0 & \mathcal{D}_N \\ \mathcal{D}_N & 0 \end{pmatrix}. \tag{4.2.1}$$

Furthermore, we recall from Section 3.4.1 that if $\psi \in \Gamma^\infty(\mathbb{S}M)$ satisfies $(\mathcal{D} - m)\psi = 0$, then the conformally rescaled spinor field $\phi = e^{\frac{a}{2}} R^{\frac{n-1}{2}} \psi \in \Gamma^\infty(\mathbb{S}M)$ satisfies

$$\left[\mathrm{i} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \partial_t + \mathrm{i} \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \partial_r + \mathrm{i} \frac{e^a}{R} \begin{pmatrix} 0 & \mathcal{D}_N \\ \mathcal{D}_N & 0 \end{pmatrix} - e^a m \right] \phi = 0\,. \tag{4.2.2}$$

This equation will be the starting point for the following analysis.

4.2.1. Decomposition into angular modes. The first step in the analysis of the Dirac equation is to get rid of the explicit N-dependence by a sort of separation of variables argument. This is possible due to the block form of (4.2.2), and the fact that the Dirac operator on a compact Riemannian manifold is a nice operator with respect to L^2-spaces. The following statement summarizes all of its properties we will need. All of this is well-known and we refer to [**LM89**, Ch. III] for a proof and background about differential operators, in particular elliptic operators, on compact Riemannian manifolds.

Theorem 4.2.1. *Let (N, g_N) be a compact Riemannian spin manifold with spinor bundle $\mathbb{S}N$ and Dirac operator $\mathcal{D}_N$. Then $\mathcal{D}_N$ is an elliptic differential operator acting on $\Gamma^\infty(\mathbb{S}N)$, and an essentially self-adjoint unbounded operator on $\Gamma^\infty(\mathbb{S}N) \subset \Gamma_{L^2}(\mathbb{S}N)$. As a consequence, the following hold:*

i.) The L^2-spectrum of $\mathcal{D}_N$ is real, discrete, and consists only of eigenvalues.
ii.) Each eigenspace is finite-dimensional, and all eigenspinors are smooth sections.
iii.) There exists an orthonormal (Hilbert) basis for $\Gamma_{L^2}(\mathbb{S}N)$ of eigenspinors of $\mathcal{D}_N$.

Remark 4.2.2. In some particular cases, typically situations with many symmetries, one can verify the properties in Theorem 4.2.1 by hand, and even explicitly compute the spectrum. For $N = S^{n-1}$ (and other symmetric spaces) this can be found, e.g., in [**Bär96**]. For the sphere one has

$$\sigma(\mathcal{D}_{S^{n-1}}) = \left\{ \pm \left(\frac{n-1}{2} + k \right) \,\middle|\, k \in \mathbb{N} \right\}, \quad \dim E_k^\pm = 2^{\lfloor \frac{n-1}{2} \rfloor} \binom{k+n-2}{k}, \tag{4.2.3}$$

where $E_k^\pm$ denotes the eigenspace corresponding to the eigenvalue $\pm(\frac{n-1}{2} + k)$.

Let us now return to our actual problem, and let $\phi \in \Gamma^\infty(\mathbb{S}M)$ be a solution of the rescaled Dirac equation (4.2.2). Due to the isomorphism $\mathbb{S}M \cong \widetilde{\mathbb{S}N} \oplus \widetilde{\mathbb{S}N}$ and Theorem 4.2.1 *iii.)*, for each fixed $(t, r) \in Q$ we can decompose $\phi(t, r, \cdot)$ as

$$\phi(t, r, \omega) = \sum_{\lambda \in \sigma(\mathcal{D}_N)} \begin{pmatrix} \phi_{\lambda,1}(t, r) \Xi_\lambda(\omega) \\ \phi_{\lambda,2}(t, r) \Xi_\lambda(\omega) \end{pmatrix}. \tag{4.2.4}$$

Here $\{\Xi_\lambda \mid \lambda \in \sigma(\mathcal{D}_N)\}$ is an L^2-orthonormal basis of eigenspinors of $\mathcal{D}_N$, and the convergence of (4.2.4) holds at least in $L^2(N)$.[2] In the following we refer to the separation constant λ as the "angular momentum", although this terminology is only really sensible if $N = S^{n-1}$. Furthermore, we call $(\phi_{\lambda,1}, \phi_{\lambda,2})$ the corresponding "angular momentum modes".

Next we project equation (4.2.2) onto the eigenspinors of $\mathcal{D}_N$. To this end, we first compute for each $\lambda \in \sigma(\mathcal{D}_N)$ separately that

$$\begin{aligned}
&\left[\mathrm{i}\begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1}\end{pmatrix}\partial_t + \mathrm{i}\begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0\end{pmatrix}\partial_r + \mathrm{i}\frac{\mathrm{e}^a}{R}\begin{pmatrix} 0 & \mathcal{D}_N \\ \mathcal{D}_N & 0\end{pmatrix} - \mathrm{e}^a m\right]\begin{pmatrix}\phi_{\lambda,1}\Xi_\lambda \\ \phi_{\lambda,2}\Xi_\lambda\end{pmatrix} \\
&\quad = \mathrm{i}\begin{pmatrix}\dot\phi_{\lambda,1}\Xi_\lambda \\ -\dot\phi_{\lambda,2}\Xi_\lambda\end{pmatrix} + \mathrm{i}\begin{pmatrix}\phi'_{\lambda,2}\Xi_\lambda \\ -\phi'_{\lambda,1}\Xi_\lambda\end{pmatrix} + \mathrm{i}\frac{\mathrm{e}^a}{R}\begin{pmatrix}\lambda\phi_{\lambda,2}\Xi_\lambda \\ \lambda\phi_{\lambda,1}\Xi_\lambda\end{pmatrix} - m\mathrm{e}^a\begin{pmatrix}\phi_{\lambda,1}\Xi_\lambda \\ \phi_{\lambda,2}\Xi_\lambda\end{pmatrix} \\
&\quad = \begin{pmatrix}\big(+ \mathrm{i}\dot\phi_{\lambda,1} + \mathrm{i}\phi'_{\lambda,2} + \mathrm{i}\frac{\mathrm{e}^a}{R}\lambda\phi_{\lambda,2} - m\mathrm{e}^a\phi_{\lambda,1}\big)\Xi_\lambda \\ \big(- \mathrm{i}\dot\phi_{\lambda,2} - \mathrm{i}\phi'_{\lambda,1} + \mathrm{i}\frac{\mathrm{e}^a}{R}\lambda\phi_{\lambda,1} - m\mathrm{e}^a\phi_{\lambda,2}\big)\Xi_\lambda\end{pmatrix}.
\end{aligned}$$

Here a dot denotes a t-derivative and a prime an r-derivative. By orthonormality of the eigenspinors Ξ_λ this implies that for every $\lambda \in \sigma(\mathcal{D}_N)$ we have

$$0 = \left\langle (4.2.2), \begin{pmatrix}\Xi_\lambda \\ 0\end{pmatrix}\right\rangle_{L^2(N)} = \mathrm{i}\dot\phi_{\lambda,1}(t,r) + \mathrm{i}\phi'_{\lambda,2}(t,r) + \mathrm{i}\frac{\mathrm{e}^a}{R}\lambda\phi_{\lambda,2}(t,r) - m\mathrm{e}^a\phi_{\lambda,1}(t,r)\,,$$

and also

$$0 = \left\langle (4.2.2), \begin{pmatrix}0 \\ \Xi_\lambda\end{pmatrix}\right\rangle_{L^2(N)} = -\mathrm{i}\dot\phi_{\lambda,2}(t,r) - \mathrm{i}\phi'_{\lambda,1}(t,r) + \mathrm{i}\frac{\mathrm{e}^a}{R}\lambda\phi_{\lambda,1}(t,r) - m\mathrm{e}^a\phi_{\lambda,2}(t,r)\,.$$

We can recombine these two equations into the 2×2 system

$$\boxed{\left[\mathrm{i}\begin{pmatrix} 1 & 0 \\ 0 & -1\end{pmatrix}\partial_t + \mathrm{i}\begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix}\partial_r + \mathrm{i}\lambda\frac{\mathrm{e}^a}{R}\begin{pmatrix} 0 & 1 \\ 1 & 0\end{pmatrix} - m\mathrm{e}^a\right]\begin{pmatrix}\phi_{\lambda,1}(t,r) \\ \phi_{\lambda,2}(t,r)\end{pmatrix} = 0}\,. \tag{4.2.5}$$

Notice that this has precisely the structure of the Dirac equation in $(1+1)$-dimensional Minkowski space with a "potential" given by the angular momentum and the mass term.

In the following the aim to study solutions of this reduced equation (4.2.5), with the focus on their behaviour along outgoing null geodesics. In the very end, we will return to the "full solution" $\phi \in \Gamma^\infty(\mathbb{S}M)$ in Section 4.7.

4.2.2. Null coordinates and the Dirac null system. The next step is to rewrite the system (4.2.5) in terms of the null coordinates

$$v = t + r\,, \qquad u = t - r\,, \tag{4.2.6}$$

since these are well-suited for what we want to analyze. The coordinates v and u are sketched in figure 4.2.

Suppose that $\phi = (\phi_1, \phi_2)$ is a solution of (4.2.5). To simplify notation, we will drop the explicit reference to the λ-dependence from now on. We want to compute how the equation (4.2.5) reads in the coordinates v and u. To this end, we first observe that

$$\partial_v = \frac{\partial_t + \partial_r}{2}\,, \qquad \partial_u = \frac{\partial_t - \partial_r}{2}\,. \tag{4.2.7}$$

[2]Since ϕ is assumed to be smooth, the convergence of (4.2.4) is actually better than just in L^2. But for us L^2-convergence suffices.

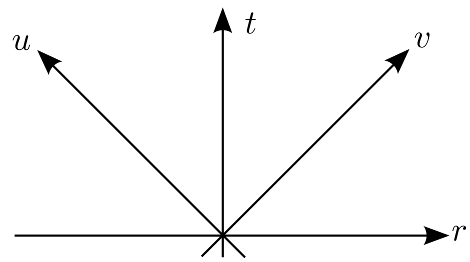

FIGURE 4.2. The ingoing and outgoing null coordinates u and v.

Next, we write out the system (4.2.5) explicitly as two equations again:

$$\begin{aligned}(1): \qquad & \mathrm{i}\partial_t\phi_1 + \mathrm{i}\partial_r\phi_2 - \mathrm{i}\lambda\frac{\mathrm{e}^a}{R}\phi_2 - \mathrm{e}^a m\phi_1 = 0\,,\\ (2): \qquad & -\mathrm{i}\partial_t\phi_2 - \mathrm{i}\partial_r\phi_1 - \mathrm{i}\lambda\frac{\mathrm{e}^a}{R}\phi_1 - \mathrm{e}^a m\phi_2 = 0\,.\end{aligned}$$

Taking the difference and the sum of these two equations, we obtain

$$\begin{aligned}(1'): \qquad & \mathrm{i}(\partial_t+\partial_r)(\phi_1+\phi_2) + \mathrm{i}\lambda\frac{\mathrm{e}^a}{R}(\phi_1-\phi_2) - \mathrm{e}^a m(\phi_1-\phi_2) = 0\,,\\ (2'): \qquad & \mathrm{i}(\partial_t-\partial_r)(\phi_1-\phi_2) - \mathrm{i}\lambda\frac{\mathrm{e}^a}{R}(\phi_1+\phi_2) - \mathrm{e}^a m(\phi_1+\phi_2) = 0\,.\end{aligned}$$

Note the appearance of ∂_v in (1') and of ∂_u in (2'). Setting

$$f := 2(\phi_1+\phi_2)\,, \qquad g := 2(\phi_1-\phi_2)\,, \tag{4.2.8}$$

we can therefore rewrite (1') and (2') as

$$\begin{aligned}(1'): \qquad & \mathrm{i}\partial_v f + \mathrm{i}\frac{\mathrm{e}^a}{2}\left(\frac{\lambda}{R} + \mathrm{i}m\right) g = 0\,,\\ (2'): \qquad & \mathrm{i}\partial_u g - \mathrm{i}\frac{\mathrm{e}^a}{2}\left(\frac{\lambda}{R} - \mathrm{i}m\right) f = 0\,.\end{aligned}$$

Introducing the "potential"

$$V_\lambda := -\frac{\mathrm{e}^a}{2}\left(\frac{\lambda}{R} + \mathrm{i}m\right)\,, \tag{4.2.9}$$

we can thus rewrite our system in the compact form

$$\boxed{\begin{cases}\partial_v f = V_\lambda g\\ \partial_u g = -\overline{V_\lambda} f\end{cases}} \tag{4.2.10}$$

Here $\overline{V_\lambda}$ is the complex conjugate of V_λ. This system will from now on be referred to as the *Dirac null system*.

Remark 4.2.3. The linear combination of the spinor components made in (4.2.8) corresponds precisely to the transformation from the *Dirac representation*, which we started with, to the *Weyl representation* (in $1+1$ dimensions). The simple form of the Dirac null system reflects the well-known fact that the Dirac equation in $(1+1)$ dimensions is a pair of transport equations along in- and outgoing null geodesics for the two Weyl-components of the spinor. The presence of the mass m and the angular momentum λ couples these two transport equations.

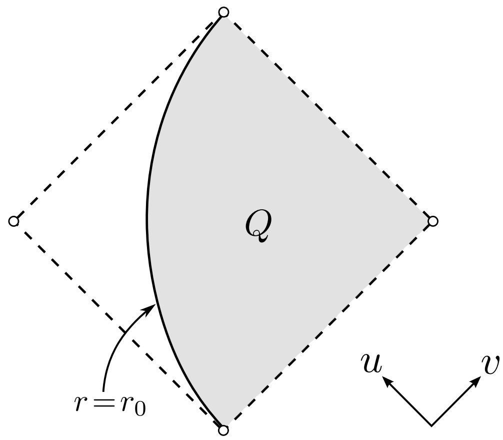

FIGURE 4.3. The Penrose diagram of Minkowski spacetime and $Q \subset \mathbb{R}^{1,1}$.

4.3. Decomposition into Free Part and Perturbation

This section contains the final rewritings of the Dirac equation before we can begin with the analytic part. Starting from the Dirac null system (4.2.10), we first derive a single, scalar integro-differential equation for the function f alone. Afterwards we decompose this scalar equation into a "free" part and a "perturbation".

4.3.1. From the Dirac null system to an integro-differential equation. As before, for $-\infty \leq r_0 < \infty$ fixed let

$$Q := \mathbb{R}_t \times (r_0, \infty)_r \subset \mathbb{R}^{1,1} \,. \tag{4.3.1}$$

Here t and r denote the standard coordinates in Minkowski spacetime, i.e. the ones in which the metric reads $\eta = -\,\mathrm{d}t^2 + \mathrm{d}r^2$. Let v and u denote the null coordinates introduced in (4.2.6). For the following argument it is useful to represent $\mathbb{R}^{1,1}$ and Q be its Penrose diagram as shown in figure 4.3, and as was explained in Section 3.1.3. Concerning these diagrams it is important to note that while the coordinate lines of the Cartesian coordinates t and r are distorted in this graphical representation, the coordinate lines of the null coordinates v and u are still straight lines and only the scale along them is distorted. This is a consequence of the fact that the representation of Q by its Penrose diagram is a *conformal* representation (see also figure 3.1 and App. A.2).

Suppose now that $f, g \in C^\infty(Q)$ satisfy the Dirac null system

$$\begin{cases} \partial_v f = V g \\ \partial_u g = -\overline{V} f \end{cases}$$

for some $V \in C^\infty(Q)$. Integrating the second equation, we write

$$g(v,u) = g(v,u_0) - \int_{u_0}^{u} \overline{V(v,x)} f(v,x)\, \mathrm{d}x \,. \tag{4.3.2}$$

At this point we make a crucial assumption about the support of f and g. Namely, we demand that

$$\boxed{\operatorname{supp} f, \operatorname{supp} g \subset J\big(\{0\}_t \times (r_0, r_1)_r\big)} \tag{4.3.3}$$

for some $r_1 > r_0$, where for $A \subset Q$ the set $J(A) = J^+(A) \cup J^-(A)$ denotes the causal future and past in Q. Figure 4.4 on the next page illustrates this condition. Equivalently,

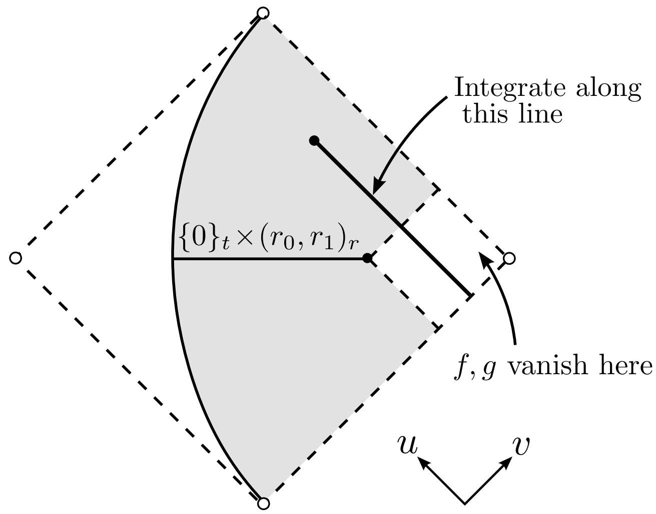

FIGURE 4.4. Illustration of the support assumption (4.3.3) and its consequence for the integral (4.3.2).

since f, g are assumed to satisfy the Dirac null system, we could demand that

$$\operatorname{supp}(f|_{t=0}), \operatorname{supp}(g|_{t=0}) \subset (r_0, r_1)_r\,. \tag{4.3.4}$$

The reason for the equivalence of these two conditions is that the Dirac null system is nothing but the Dirac equation in $\mathbb{R}^{1,1}$ (with a potential), which satisfies the usual finite propagation speed property. To see this, simply retrace the steps in the derivation of the Dirac null system in Section 4.2.2.

The relevant consequence of (4.3.3) for our purpose is that it implies that for any $v \in \mathbb{R}$ sufficiently large ($v > r_1$) we have $f(v, u) = g(v, u) = 0$ for all u sufficiently small ($u < -r_1$), see figure 4.4 for an illustration. Combining this with (4.3.2), it follows that

$$g(v, u) = -\int_{-\infty}^{u} \overline{V(v, x)} f(v, x)\, \mathrm{d}x \qquad v > r_1\,. \tag{4.3.5}$$

Note that because of (4.3.3) the integrand in fact vanishes for $x < -r_1$, so that the integral is really only over the finite interval $(-r_1, u)$.

Substituting (4.3.5) back into the first equation of the Dirac null system, we obtain the integro-differential equation

$$\partial_v f(v, u) = -V(v, u) \int_{-\infty}^{u} \overline{V(v, x)} f(v, x)\, \mathrm{d}x\,. \tag{4.3.6}$$

The important point is that this is an equation for f alone, i.e. which does not involve g. We will use it to study the behaviour of f as v tends to infinity.

Let us summarize our findings in the following Lemma.

Lemma 4.3.1. *Let $f, g \in C^\infty(Q)$ be a solution of the Dirac null system* (4.2.10) *for some $V \in C^\infty(Q)$. Assume that f and g satisfy the support assumption* (4.3.4) *for some $r_1 > r_0$. Then f satisfies the integro-differential equation*

$$\partial_v f(v, u) = -V(v, u) \int_{u_0}^{u} \overline{V(v, x)} f(v, x)\, \mathrm{d}x \qquad \forall v \geq v_0, u \geq u_0\,, \tag{4.3.7}$$

where $v_0 = r_1$ and $u_0 = -r_1$.

For later purposes we remark that g can be recovered from f by the integral (4.3.5).

Remark 4.3.2. Intuitively speaking, the support assumption (4.3.3) prevents that "unlimited energy comes in from spatial infinity". As one may observe while going through the estimates in the following, it seems reasonable that instead of the condition (4.3.3) it should be sufficient to demand that $f|_{t=0}, g|_{t=0}$ decay sufficiently fast for $r \to \infty$. More precisely, the final decay result (4.6.17) suggests that finite Sobolev norms of some order should be sufficient.

4.3.2. Splitting into free part and perturbation. In the second step we split the integro-differential equation (4.3.7) for f into a "free part", which has the merit of being exactly solvable, and a "perturbation". To this end, we insert into (4.3.7) the special form of the potential

$$V_\lambda(v,u) = -\frac{\mathrm{e}^{a(v,u)}}{2}\left(\frac{\lambda}{R(v,u)} + \mathrm{i}m\right) = -\frac{1-A(v,u)}{2}\left(\frac{\lambda}{R(v,u)} + \mathrm{i}m\right),$$

and multiply out. Here we set $\mathrm{e}^a = 1 - A$ since this is more convenient in the following. Dropping the v-dependence from the notation in the following computation, this yields

$$\begin{aligned}
\partial_v f(u) &\overset{(4.3.7)}{=} -\frac{1}{4}\int_{u_0}^u (1-A(u))\left(\frac{\lambda}{R(u)} + \mathrm{i}m\right)(1-A(x))\left(\frac{\lambda}{R(x)} - \mathrm{i}m\right) f(x)\,\mathrm{d}x \\
&= -\frac{1}{4}\int_{u_0}^u \left(\mathrm{i}m + \frac{\lambda}{R(u)} - \mathrm{i}mA(u) - \lambda\frac{A(u)}{R(u)}\right) \\
&\qquad \times \left(-\mathrm{i}m + \frac{\lambda}{R(x)} + \mathrm{i}mA(x) - \lambda\frac{A(x)}{R(x)}\right) f(x)\,\mathrm{d}x \\
&= -\frac{1}{4}\int_{u_0}^u \Bigg\{ m^2 + \mathrm{i}m\frac{\lambda}{R(x)} - m^2A(x) - \mathrm{i}m\lambda\frac{A(x)}{R(x)} \\
&\qquad\qquad - \mathrm{i}m\frac{\lambda}{R(u)} + \frac{\lambda^2}{R(x)R(u)} + \mathrm{i}m\lambda\frac{A(x)}{R(u)} - \lambda^2\frac{A(x)}{R(x)R(u)} \\
&\qquad\qquad - m^2A(u) - \mathrm{i}m\lambda\frac{A(u)}{R(x)} + m^2A(u)A(x) + \mathrm{i}m\lambda\frac{A(u)A(x)}{R(x)} \\
&\qquad\qquad + \mathrm{i}m\lambda\frac{A(u)}{R(u)} - \lambda^2\frac{A(u)}{R(u)R(x)} - \mathrm{i}m\lambda\frac{A(u)A(x)}{R(u)} + \lambda^2\frac{A(u)A(x)}{R(u)R(x)}\Bigg\} \\
&\qquad \times f(x)\,\mathrm{d}x \\
&= -\frac{m^2}{4}\int_{u_0}^u f(x)\,\mathrm{d}x \\
&\quad + \int_{u_0}^u \frac{1}{4}\Bigg\{\mathrm{i}m\lambda\left(\frac{1}{R(u)} - \frac{1}{R(x)}\right) + m^2(A(x)+A(u)) \\
&\qquad + \mathrm{i}m\lambda\left(\frac{A(x)}{R(x)} - \frac{A(u)}{R(u)} + \frac{A(u)}{R(x)} - \frac{A(x)}{R(u)}\right) - \frac{\lambda^2}{R(x)R(u)} - m^2A(x)A(u) \\
&\qquad + \mathrm{i}m\lambda\left(\frac{A(u)A(x)}{R(u)} - \frac{A(u)A(x)}{R(x)}\right) + \lambda^2\left(\frac{A(x)}{R(x)R(u)} + \frac{A(u)}{R(x)R(u)}\right) \\
&\qquad - \lambda^2\frac{A(u)A(x)}{R(u)R(x)}\Bigg\} f(x)\,\mathrm{d}x\,.
\end{aligned}$$

We abbreviate this lengthy expression as

$$\boxed{(\partial_v f)(v,u) = -\frac{m^2}{4}\int_{u_0}^{u} f(v,x)\,\mathrm{d}x + \int_{u_0}^{u} K_B(v,u,x) f(v,x)\,\mathrm{d}x\,,} \tag{4.3.8}$$

where the kernel $K_B(v,u,x)$ is explicitly given by

$$\begin{aligned}
&4K_B(v,u,x)\\
&= \mathrm{i}m\lambda\left(\frac{1}{R(v,u)} - \frac{1}{R(v,x)}\right) + m^2(A(v,x)+A(v,u))\\
&+ \mathrm{i}m\lambda\left(\frac{A(v,u)}{R(v,x)} - \frac{A(v,x)}{R(v,u)} + \frac{A(v,x)}{R(v,x)} - \frac{A(v,u)}{R(v,u)}\right) - \frac{\lambda^2}{R(v,x)R(v,u)} - m^2 A(v,x)A(v,u)\\
&+ \mathrm{i}m\lambda\left(\frac{A(v,u)A(v,x)}{R(v,u)} - \frac{A(v,u)A(v,x)}{R(v,x)}\right) + \lambda^2\left(\frac{A(v,x)}{R(v,x)R(v,u)} + \frac{A(v,u)}{R(v,x)R(v,u)}\right)\\
&- \lambda^2\frac{A(v,u)A(v,x)}{R(v,u)R(v,x)}
\end{aligned} \tag{4.3.9}$$

The grouping of the terms into the different lines of the right-hand side is done according to the total number of factors of both A and $\frac{1}{R}$ in each term.

In the first term in (4.3.8), which we call the "free part", we kept only the contribution of V_λ which does not depend on v and u. All other terms are put into K_B. The first indication that this splitting is useful is that dropping the "perturbation" we have

$$\partial_v f = -\frac{m^2}{4}\int_{u_0}^{u} f(v,x)\,\mathrm{d}x \quad\Longrightarrow\quad \Box f = -4\partial_u\partial_v f = m^2 f\,.$$

So we see that solutions of the free part alone satisfy the Klein-Gordon equation in $1+1$ dimensions. This will make it possible to explicitly solve the free part.

Remark 4.3.3. In the end it will of course also be crucial that the perturbation is really "small" compared to the free part in a suitable sense. The basic reasons why this is the case is that the terms of the perturbation decay as $v\to\infty$ if the metric satisfies the asymptotic flatness conditions of Definition 3.2.1.

4.4. Treatment of the Free Part

Up to now we have basically just rewritten the Dirac equation, ending up with the integro-differential equation (4.3.8). Now we start with the actual analysis. In this section we focus on the free part of (4.3.8), i.e. the equation $\partial_v f = -\frac{m^2}{4}\int f(v,x)\,\mathrm{d}x$. The aim is to analyze the behaviour of its solutions as $v\to\infty$. To this end, we are going to derive an integral representation of $f(v,u)$ in terms of $f(v_0,u)$ for some fixed v_0. This amounts to solving a *Goursat problem*, since the hypersurface $\{v=v_0\}$ is *characteristic* for the equation.[3] From this integral representation, we will then be able to "read off" the behaviour as $v\to\infty$.

In the following treatment it is absolutely crucial that the mass is nonzero:

$$\boxed{\textbf{Nonzero mass assumption:}\quad \mathbf{m\neq 0}}\,. \tag{4.4.1}$$

For $m=0$ one can derive analogous representation formulas, but the decay properties are fundamentally different.

[3] More precisely, it is characteristic for the related Klein-Gordon equation.

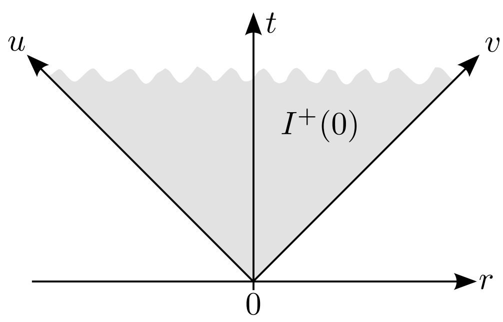

FIGURE 4.5. Sketch of the domain $I^+(0) \subset \mathbb{R}^{1,1}$ in which we consider the Goursat problem (4.4.2).

4.4.1. The Goursat problem for the Klein-Gordon equation. As we have already seen, if $\partial_v f = -\frac{m^2}{4}\int f(v,x)\,\mathrm{d}x$, then f satisfies the Klein-Gordon equation

$$(\Box - m^2)f = 0\,.$$

To see this, just differentiate the integro-differential equation in u and use that the wave operator is given by $\Box = -4\partial_v\partial_u$ when expressed in the null coordinates v and u. Therefore in the following we study the Goursat problem

$$\begin{cases} -4\partial_v\partial_u\phi(v,u) = m^2\phi & \text{in } I^+(0) \subset \mathbb{R}^{1,1} \\ \phi(0,u) = \phi_0(u) & \forall u \geq 0 \\ \phi(v,0) = \phi_1(v) & \forall v \geq 0 \end{cases} \tag{4.4.2}$$

Here the domain $I^+(0)$ is the future timecone of the origin in $\mathbb{R}^{1,1}$, i.e.

$$I^+(0) = \{p \in \mathbb{R}^{1,1} \mid v(p), u(p) > 0\}\,, \tag{4.4.3}$$

where $v = t + r$ and $u = t - r$ are the usual null coordinates, see figure 4.5. In the following, we will exclusively work in these coordinates. To this end, let us note that the Minkowski metric takes the form $\eta = -\,\mathrm{d}v\,\mathrm{d}u$ in these coordinates.

Our aim is to derive an integral representation for ϕ in terms of ϕ_0 and ϕ_1. We will actually demonstrate existence and uniqueness of the problem (4.4.2) as well. Let us note that since the equation is translation-invariant in $\mathbb{R}^{1,1}$, once we have accomplished this for the domain $I^+(0)$ we have also obtained it for any other domain $I^+(p)$. Also all formulas can be adjusted by a simple translation in the variables.

Remark 4.4.1. At first sight it might seem odd that we consider an initial value problem for the second order equation $(\Box - m^2)\phi = 0$ where we only prescribe ϕ on the initial hypersurface and not also its ”time” derivative (or normal derivative). However, that this is the correct initial data is a general key feature of the Goursat problem. The reason for this becomes particularly clear in our example if we write the equation as $-4\partial_u\partial_v\phi = m^2\phi$. Namely, notice that one of the derivatives on the left-hand side is *tangential* to our initial hypersurface. Therefore the order of the equation is effectively reduced from two to one, giving an intuitive argument why we should only prescribe ϕ as initial data. Another heuristic argument is that for a characteristic surface, such as $\partial I^+(0)$, the normal vector is actually tangential. Therefore one cannot independently specify characteristic initial values both for ϕ and its normal derivative, since the latter can already be obtained from the characteristic initial values.

4.4.2. Uniqueness of solutions of the Goursat problem. First we show uniqueness for the Goursat problem (4.4.2) by a simple energy estimate argument using the *energy-momentum tensor formalism* or *vector field method* (see Section 2.3.3 for a general explanation of this method). In the following, $\Omega \subset \mathbb{R}^{1,1}$ will always denote an open subset of $\mathbb{R}^{1,1}$.

Definition 4.4.2. For $\phi \in \mathcal{C}^\infty(\Omega)$, the *Klein-Gordon energy-momentum tensor* $T^\phi \in \Gamma^\infty(T^*\Omega \otimes T^*\Omega)$ of mass $m \geq 0$ is the symmetric 2-tensor defined by

$$T^\phi_{\mu\nu} := \mathsf{Re}\Big[(\partial_\mu \phi)\overline{(\partial_\nu \phi)}\Big] - \frac{1}{2}\Big[(\partial_\alpha \phi)\overline{(\partial^\alpha \phi)} + m^2|\phi|^2\Big]\,\eta_{\mu\nu}\,. \tag{4.4.4}$$

This is to be understood in abstract index notation, and $\eta_{\mu\nu}$ denotes the Minkowski metric.

The relevance of this tensor field for the Klein-Gordon equation comes about when computing its divergence.

Lemma 4.4.3. (Divergence of energy-momentum tensor) *Let $\phi \in \mathcal{C}^\infty(\Omega)$ and $X \in \Gamma^\infty(T\Omega)$. Then*

$$\nabla^\mu(T^\phi_{\mu\nu}X^\nu) = \mathsf{Re}\Big[(\Box\,\phi - m^2\phi)\overline{\mathrm{d}\phi(X)}\Big] + \frac{1}{2}T^\phi_{\mu\nu}(\mathscr{L}_X\eta)^{\mu\nu}\,. \tag{4.4.5}$$

In particular, if $\Box\,\phi = m^2\phi$ and X is an η-Killing vector field (i.e. $\mathscr{L}_X\eta = 0$), then the one-form $\mathbb{J}^\phi[X]_\mu := T^\phi_{\mu\nu}X^\nu$ is divergence-free.[4]

Proof. Using the Leibniz rule, we find

$$\begin{aligned}
\partial^\mu(T^\phi_{\mu\nu}X^\nu) &= \mathsf{Re}\big[(\partial^\mu\partial_\mu\phi)(X\overline{\phi})\big] + \mathsf{Re}\big[(\partial_\mu\phi)\partial^\mu(X\overline{\phi})\big] \\
&\quad - \underbrace{\frac{1}{2}\eta_{\mu\nu}X^\nu\partial^\mu\big[(\partial_\alpha\phi)\overline{(\partial^\alpha\phi)} + m^2|\phi|^2\big]}_{=\mathsf{Re}[(\partial_\alpha\phi)(X\partial^\alpha\overline{\phi})]+\mathsf{Re}[m^2\phi(X\overline{\phi})]} - \frac{1}{2}(\eta_{\mu\nu}\partial^\mu X^\nu)\big[(\partial_\alpha\phi)\overline{(\partial^\alpha\phi)} + m^2|\phi|^2\big] \\
&= \mathsf{Re}\big[(\Box\,\phi - m^2\phi)(X\overline{\phi})\big] + \mathsf{Re}\big[(\partial_\mu\phi)(\partial^\mu X^\nu)(\partial_\nu\overline{\phi})\big] \\
&\quad - \frac{1}{2}(\eta_{\mu\nu}\partial^\mu X^\nu)\big[(\partial_\alpha\phi)\overline{(\partial^\alpha\phi)} + m^2|\phi|^2\big] \\
&= \mathsf{Re}\big[(\Box\,\phi - m^2\phi)(X\overline{\phi})\big] + T^\phi_{\mu\nu}\partial^\mu X^\nu\,.
\end{aligned}$$

To finish the proof we note that since $T^\phi_{\mu\nu}$ is symmetric, we have

$$T^\phi_{\mu\nu}\partial^\mu X^\nu = \frac{1}{2}T^\phi_{\mu\nu}\partial^{(\mu}X^{\nu)} = \frac{1}{2}T^\phi_{\mu\nu}(\mathscr{L}_X\,\eta)^{\mu\nu}\,.$$

The last equality is easy to verify, for instance using the Koszul formula for the Lie derivative (cf. [**Lee03**, eq. (18.12)]). ∎

For a given solution ϕ of the Klein-Gordon equation, one can thus construct *conservation laws* from Killing vector fields X by integrating the divergence of $\mathbb{J}^\phi[X]$ over a bounded domain and converting the integral into boundary integrals using Gauss' theorem.

[4] In index-free notation, equation (4.4.5) reads $\mathrm{div}(T^\phi(X,\cdot)) = \mathsf{Re}[(\Box\,\phi - m^2\phi)\,\mathrm{d}\overline{\phi}(X)] + \frac{1}{2}T^\phi \bullet \mathscr{L}_X\eta$, where $\bullet$ denotes a complete (metric) contraction of two 2-tensors.

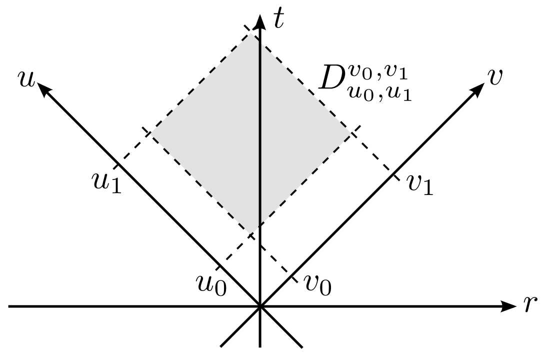

FIGURE 4.6. Sketch of the diamond-shaped domain $D^{v_0,v_1}_{u_0,u_1}$.

In our situation, we will consider domains whose boundaries are null hypersurfaces. Since the metric degenerates on such hypersurfaces, Gauß' theorem takes a slightly different form, as we now explain. More precisely, we consider diamond-shaped domains

$$D^{v_0,v_1}_{u_0,u_1} := \{p \in \mathbb{R}^{1,1} \mid v_0 \leq v(p) \leq v_1 \, , \, u_0 \leq u(p) \leq u_1\} \subset \mathbb{R}^{1,1} \tag{4.4.6}$$

as sketched in figure 4.14. Here $v = t + r$ and $u = t - r$ again denote the usual null coordinates. Concerning integration, note that since in null coordinates the Minkowski metric reads $\eta = -\,\mathrm{d}v\,\mathrm{d}u$, and since (∂_u, ∂_v) is positively oriented with respect to the usual orientation, the Minkowskian volume element is given in null coordinates by

$$\mathrm{d}\mu_\eta = \mathrm{d}u \wedge \mathrm{d}v \, . \tag{4.4.7}$$

Now we come to Gauß' law for null boundaries.

Lemma 4.4.4. (Gauss' for null domains) *Let $D = D^{v_0,v_1}_{u_0,u_1}$, and let $X \in \Gamma^1(TD) \cap \Gamma^0(T\overline{D})$. Assume that $\int_D |\operatorname{div} X| \,\mathrm{d}u\,\mathrm{d}v < \infty$. Then it holds that*

$$\begin{aligned} -\int_D \operatorname{div} X \,\mathrm{d}u \wedge \mathrm{d}v = &\int_{u_0}^{u_1} \eta(X, \partial_u)|_{v_1} \,\mathrm{d}u + \int_{v_0}^{v_1} \eta(X, \partial_v)|_{u_1} \,\mathrm{d}v \\ &- \int_{u_0}^{u_1} \eta(X, \partial_u)|_{v_0} \,\mathrm{d}u - \int_{v_0}^{v_1} \eta(X, \partial_v)|_{u_0} \,\mathrm{d}v \, . \end{aligned} \tag{4.4.8}$$

PROOF. First of all, we have

$$\eta = -\,\mathrm{d}t^2 + \mathrm{d}x^2 = -(\mathrm{d}t + \mathrm{d}x)(\mathrm{d}t - \mathrm{d}x) = -\,\mathrm{d}v\,\mathrm{d}u \, . \tag{$*$}$$

Since the ordered basis (∂_u, ∂_v) is positively oriented, the volume form of η is given by $\Omega_\eta = \mathrm{d}u \wedge \mathrm{d}v$. Using this together with the identity $\mathscr{L}_X \Omega_\eta = \operatorname{div} X \cdot \eta$ and Cartan's magic formula for the Lie derivative, it follows that

$$\operatorname{div} X \cdot \mathrm{d}u \wedge \mathrm{d}v = \operatorname{div} X \cdot \Omega_\eta = \mathscr{L}_X \Omega_\eta = (\iota_X \circ \mathrm{d} + \mathrm{d} \circ \iota_X)\Omega_\eta = \mathrm{d}(\iota_X \Omega_\eta) \, .$$

Here we used that $\mathrm{d}\Omega_\eta = 0$ since Ω_η already has top degree. Using the assumptions on X we can now apply Stokes' theorem as stated for instance in [**Sau06**, Ch.I §4, Thm.1, p.38], which yields

$$\int_D \operatorname{div} X \,\mathrm{d}u \wedge \mathrm{d}v = \int_D \mathrm{d}(\iota_X \Omega_\eta) = \int_{\partial D} \iota_X \Omega_\eta \, .$$

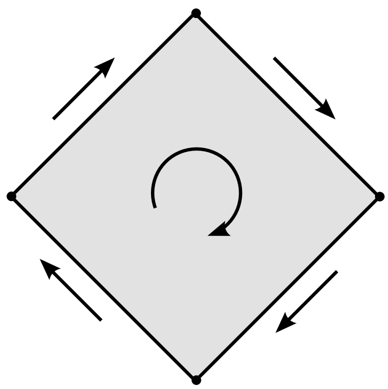

FIGURE 4.7. The correct orientation for Stokes' theorem in Lemma 4.4.4.

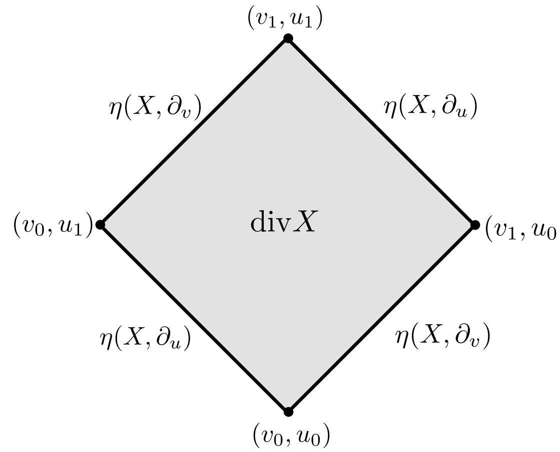

FIGURE 4.8. Stokes theorem for diamond-shaped regions.

In order to arrive at (4.4.8), we only have to express $\iota_X \Omega_\eta$ more explicitly. To this end, notice that by $(*)$ we can decompose X as

$$X = -\eta(X, \partial_u)\partial_v - \eta(X, \partial_v)\partial_u \,,$$

and thus

$$\begin{aligned} \iota_X \Omega_g &= -\eta(X, \partial_v)\iota_{\partial_u}\Omega_g - \eta(X, \partial_u)\iota_{\partial_v}\Omega_g \\ &= -\eta(X, \partial_v)\, \mathrm{d}v + \eta(X, \partial_u)\, \mathrm{d}u \,. \end{aligned}$$

Taking care of the correct boundary orientation for the application of Stokes theorem as sketched in figure 4.7, formula (4.4.8) follows. □

Figure 4.8 illustrates the Gauß' law for a diamond. Over each boundary part we have written the term one has to integrate over that part.

Now we apply this to the case where the vector field X in (4.4.8) is the contraction of the energy momentum tensor T^ϕ with either of the Killing fields ∂_u or ∂_v of the Minkowski metric, and where $\phi \in C^\infty(\overline{D^{v_0,v_1}_{u_0,u_1}})$ is a solution of the Klein-Gordon equation. Then by (4.4.5) the divergence of X vanishes, so that we obtain a pure relation between the boundary terms. Concerning these, we have

$$T^\phi(\partial_u, \partial_u) = |\partial_u \phi|^2 \tag{4.4.9}$$

$$T^\phi(\partial_v, \partial_v) = |\partial_v \phi|^2 \tag{4.4.10}$$

$$T^\phi(\partial_u, \partial_v) = T^\phi(\partial_v, \partial_u) = \frac{m^2}{2}|\phi|^2 \,. \tag{4.4.11}$$

This together with Lemma 4.4.3 and Lemma 4.4.4 now immediately yields the following identities.

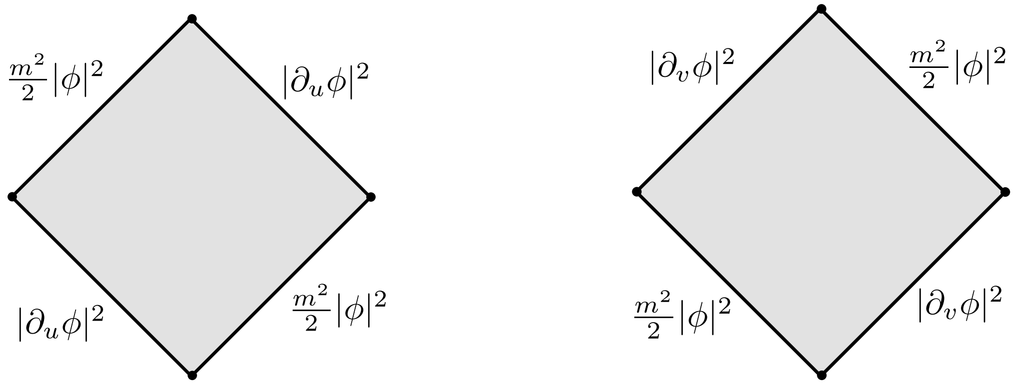

FIGURE 4.9. Null conservation laws for the wave equation.

Corollary 4.4.5. *Let $\phi \in C^2(D^{v_0,v_1}_{u_0,u_1}) \cap C^1(\overline{D^{v_0,v_1}_{u_0,u_1}})$ be a solution of the Klein-Gordon equation. Then it holds that*

$$\begin{aligned}&\int_{u_0}^{u_1} |\partial_u\phi(v_1,u)|^2 \,\mathrm{d}u^2 + \int_{v_0}^{v_1} \frac{m^2}{2}|\phi(v,u_1)|^2 \,\mathrm{d}v^2 \\ =&\int_{u_0}^{u_1} |\partial_u\phi(v_0,u)|^2 \,\mathrm{d}u^2 + \int_{v_0}^{v_1} \frac{m^2}{2}|\phi(v,u_0)|^2 \,\mathrm{d}v^2 \,,\end{aligned} \tag{4.4.12}$$

and

$$\begin{aligned}&\int_{u_0}^{u_1} \frac{m^2}{2}|\phi(v_1,u)|^2 \,\mathrm{d}u^2 + \int_{v_0}^{v_1} |\partial_v\phi(v,u_1)|^2 \,\mathrm{d}v^2 \\ =&\int_{u_0}^{u_1} \frac{m^2}{2}|\phi(v_0,u)|^2 \,\mathrm{d}u^2 + \int_{v_0}^{v_1} |\partial_v\phi(v,u_0)|^2 \,\mathrm{d}v^2 \,.\end{aligned} \tag{4.4.13}$$

PROOF. The regularity assumptions on ϕ imply that $T^\phi(\partial_u,\cdot), T^\phi(\partial_v,\cdot) \in \Gamma^1(TD) \cap \Gamma^0(T\overline{D})$. Further, since ϕ satisfies the Klein-Gordon equation, these vector fields are both divergence-free. Therefore the assumptions of Lemma 4.4.4 are satisfied, and together with (4.4.9) – (4.4.11) this immediately yields the asserted identities. □

Figure 4.9 gives a graphical representation of these conservation laws, where as before we have written over each boundary part the term one has to integrate over this part.

From these identities we can immediately deduce uniqueness for smooth solutions of the Goursat problem (4.4.2).

Corollary 4.4.6. (Uniqueness of solutions of the Goursat problem) *Suppose that $\phi \in C^2(D^{0,v_1}_{0,u_1}) \cap C^1(\overline{D^{0,v_1}_{0,u_1}})$ solves the Goursat problem (4.4.2) with $\phi_0 = \phi_1 = 0$. Then it follows that $\phi = 0$. In particular, this implies that for any given smooth initial conditions ϕ_0, ϕ_1 there exists at most one smooth solution of (4.4.2).*

PROOF. Let $p \in I^+(0)$ be given. Choose a diamond $D^{0,v_1}_{0,u_1} \subset I^+(0)$ such that $p \in \partial D^{0,v_1}_{0,u_1}$. From the conservation laws (4.4.12) and (4.4.13) and continuity of ϕ it follows that ϕ vanishes on the boundary. So $\phi(p) = 0$ as desired. Concerning uniqueness one only has to apply this to the difference of two solutions with the same initial values. □

Now that we know that there is at most one smooth solution for the Goursat problem (4.4.2), we can start looking for a representation formula of a given solution in terms of its

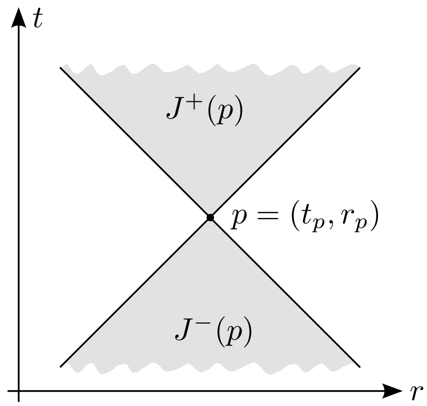

FIGURE 4.10. The causal past and past of a point $p \in \mathbb{R}^{1,1}$.

characteristic initial data. Note once more that since a solution of the Goursat problem (4.4.2) is uniquely determined by its values on the characteristic hypersurface alone, we should look for a representation formula which only involves the initial values.

4.4.3. The causal fundamental solution for the Klein-Gordon equation. The representation formula for solutions of the Goursat problem (4.4.2) we are going to derive will be a convolution integral of the (derivative of) the characteristic initial data with the so-called *causal fundamental solution* of the Klein-Gordon equation. In this section we therefore present an explicit formula for the causal fundamental solution. This is similar in spirit to the content explained in Section 2.3.1, only that here we solve a characteristic initial value problem and not a Cauchy problem.

First, recall that a *Green's function*[5] of the Klein-Gordon equation at a fixed point $p \in \mathbb{R}^{1,1}$ is a tempered distribution $S_p \in \mathcal{S}'(\mathbb{R}^{1,1})$ satisfying

$$(\Box - m^2) S_p = \delta_p \,. \tag{4.4.14}$$

Here $\delta_p \in \mathcal{S}'(\mathbb{R}^{1,1})$ is the Dirac delta distribution at p. This equation does not have a unique solution since one may always add solutions of the homogeneous Klein-Gordon equation. However, one can single out unique solutions by demanding additional properties. As is well-known, for every point $p \in \mathbb{R}^{1,1}$ there exist two unique Green's functions $S_p^\pm \in \mathcal{S}'(\mathbb{R}^{1,1})$ with the additional property that

$$\operatorname{supp} S_p^\pm \subset J^\pm(p) \,. \tag{4.4.15}$$

Here for $p = (t_p, r_p) \in \mathbb{R}^{1,1}$, the sets $J^\pm(p) \subset \mathbb{R}^{1,1}$ are the *causal future* and *causal past* of p, which are defined by

$$J^\pm(p) = \big\{(t,x) \in \mathbb{R}^{1,1} \mid \pm(t - t_p) > 0 \,, \, -(t-t_p)^2 + (x - x_p)^2 \leq 0\big\} \subset \mathbb{R}^{1,1} \,, \tag{4.4.16}$$

see figure 4.10 for an illustration. These two Green's functions are called *advanced Green's function* and *retarded Green's function*. Their existence and uniqueness is for instance shown in [**BGP07**, Thm. 3.3.1] in a much more general setting than considered here.

[5]Be careful that the expressions *Green's function* and *fundamental solution* are not used consistently throughout the literature. Sometimes what we call Green's function is called fundamental solution. Here we follow the same convention which is used in [**Fin06**].

Proposition 4.4.7. (Advanced and retarded Green's function) *Let $p = (t_p, r_p) \in \mathbb{R}^{1,1}$. The advanced and retarded Green's functions $S_p^{\pm} \in \mathcal{S}'(\mathbb{R}^{1,1})$ are explicitly given by the continuous functions*

$$S_p^{\pm}(t,r) = \frac{1}{2} J_0(m|p-q|_\eta) \mathbb{1}_{J^\pm(p)}(q)\,. \tag{4.4.17}$$

Here $J_0 \in C^\infty(\mathbb{R})$ is the zeroth Bessel function of the first kind (cf. [**BW10**, Ch. 7]*), $|p-q|_\eta = \sqrt{(t_p-t_q)^2 - (r_p-r_q)^2}$ is the Minkowskian distance, and $\mathbb{1}_{J^\pm(p)}$ is the characteristic function of $J^\pm(p)$.*

Proof. We focus on the advanced Green's function, the retarded one works analogously. Moreover, due to translation-invariance of the whole problem it suffices to consider the case $p = 0$. To ease notation, we will drop the reference to the point $p = 0$ from the notation.

Instead of verifying directly that S^+ is a Green's function (clearly it has the correct support properties), we proceed a little different. Namely, let us first find a solution $\widehat{S}^+ \in \mathcal{S}'(\mathbb{R}_t \times \mathbb{R}_k)$ of

$$(-\partial_t^2 - k^2 - m^2)\widehat{S}^+ = \delta(t)\,. \tag{$*$}$$

Note that this is just the Fourier transform in r of the Green's function equation (4.4.14). The idea is to first construct the spatial Fourier transform of S^+ of S^+, and then to transform back. For every $t \neq 0$ the right-hand side of ($*$) vanishes, and we can explicitly solve the equation by

$$\widehat{S}^+(t,k) = \begin{cases} a(k)\cos(\omega_k t) + b(k)\sin(\omega_k t) & t > 0 \\ c(k)\cos(\omega_k t) + d(k)\sin(\omega_k t) & t < 0 \end{cases},$$

where we have set $\omega_k := \sqrt{|k|^2 + m^2}$. The idea to produce a delta function in ($*$) is to choose the coefficients $a(k), b(k), c(k), d(k)$ in such a way that $\widehat{S}^+(t,k)$ is continuous in t, whereas $\partial_t \widehat{S}^+(t,k)$ jumps by one across $t = 0$. This will hold if we set

$$c(k) = a(k) \quad \text{and} \quad \big(b(k) - d(k)\big) = \frac{1}{\omega_k}\,. \tag{$**$}$$

This leaves us with the choice of two of the four coefficients, which once more illustrates that there is no unique Green's function. We will choose these coefficients such that the support condition of the advanced Green's function is satisfied. To this end, it suffices to notice that $\widehat{S}^+(t,k)$ should clearly vanish for $t < 0$, which will only hold if we set $c(k) = d(k) = 0$. By ($**$) this implies that $a(k) = 0$ and $b(k) = \frac{1}{\omega_k}$, so that

$$\widehat{S}^+(t,k) = \begin{cases} \frac{\sin(\omega_k t)}{\omega_k} & t > 0 \\ 0 & t < 0 \end{cases}.$$

At this point, notice that $\widehat{S}^+$ is actually a continuous function and that, moreover, $\widehat{S}^+(t,\cdot)$ is bounded (actually decays) as a function of k for fixed t. Therefore $\widehat{S}^+(t,\cdot) \in \mathcal{S}'(\mathbb{R}_k)$, so that we can indeed take the inverse Fourier transform in k (in the distributional sense). This yields a distribution $S^+ \in \mathcal{S}'(\mathbb{R}_t \times \mathbb{R}_r)$, and we claim that S^+ is the advanced Green's function. To this end, notice that the difference of S^+ and the advanced Green's function is an (a priori distributional) solution of the *homogeneous* Klein-Gordon equation which vanishes identically for $t < 0$. By uniqueness of the Cauchy problem of the Klein-Gordon equation (and regularity of solutions with smooth data), it follows that this difference must vanish everywhere. Hence S^+ coincides with the advanced Green's function.

Finally, it remains to show that S^+ coincides with the function (4.4.17). To this end, we need to explicitly compute the inverse Fourier transform

$$S^+(t,r) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin(\sqrt{k^2+m^2}\,t)}{\sqrt{k^2+m^2}} e^{ikr}\,dk\,.$$

In this integral, let us substitute k for a new integration variable x, defined implicitly through the equation $k =: m\sinh x$. Notice that then we have

$$\sqrt{k^2+m^2} = m\sqrt{\sinh^2 x + 1} = m\cosh x = \frac{dk}{dx}\,,$$

so that

$$S^+(t,r) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \sin(mt\cosh x) e^{imr\sinh x}\,dx\,.$$

To evaluate this integral, we use the so-called *Mehler-Sonine integral* for the Bessel functions J_ν of the first kind, which states that (cf. [**O$^+$10**, p. 224, (10.9.8.)])

$$J_\nu(z) = \frac{2}{\pi}\int_0^{\infty} \sin(z\cosh x - \tfrac{1}{2}\nu\pi)\cosh(\nu x)\,dx\,.$$

By this identity we have

$$S^+(t,0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \sin(mt\cosh x)\,dx = \frac{1}{\pi}\int_0^{\infty} \sin(mt\cosh x)\,dx = \frac{1}{2}J_0(mt)\,,$$

which coincides with the right-hand side of (4.4.17) for $r = 0$.

To finish the proof, notice that the advanced Green's function S^+ is invariant under proper Lorentz-transformations. To see this, notice first that any Lorentz transformation $\mathcal{L}(S^+)$ of S^+ is again a Green's function since the wave operator commutes with Lorentz transformations. Secondly, notice that a proper Lorentz transformation also leaves the set $J^+(0)$ invariant, so that $\operatorname{supp}\mathcal{L}(S^+) = J^+(0)$. Therefore it follows from uniqueness of the advanced Green's function that $\mathcal{L}(S^+) = S^+$. We use this invariance property as follows. On the one hand, any point $(t,r) \in I^+(0)$ can be mapped to $(\sqrt{t^2-r^2}, 0)$ by a Lorentz transformation. Therefore

$$S^+(t,r) = S^+(\sqrt{t^2-r^2}, 0) = \frac{1}{2}J_0\big(m\sqrt{t^2-r^2}\,\big) \qquad \forall (t,r) \in I^+(0)\,.$$

On the other hand, any point outside of $J^+(0)$ can be related to a point in the lower half-plane $\{t < 0\}$ by a proper Lorentz transformation. Since S^+ vanishes in the lower half-plane $\{t < 0\}$, it thus follows that it vanishes completely outside of $J^+(0)$. This completes the proof. $\square$

Now we come to the causal fundamental solution. Let again $p \in \mathbb{R}^{1,1}$ be fixed, then the *causal fundamental solution at* p (of the Klein-Gordon equation) is the distribution $K_p \in \mathcal{S}'(\mathbb{R}^{1,1})$ defined by

$$K_p := S_p^+ - S_p^-\,. \tag{4.4.18}$$

Notice that K_p is a solution of the *homogeneous* Klein-Gordon equation

$$(\Box - m^2)K_p = 0\,, \tag{4.4.19}$$

and $\operatorname{supp} K_p \subset J(p)$, where $J(p) = J^-(p) \cup J^+(p)$ is the causal cone at p. From the explicit expressions (4.4.17) for $S_p^\pm$, we immediately obtain an explicit expression for K_p.

Proposition 4.4.8. *Let $p \in \mathbb{R}^{1,1}$. The causal fundamental solution at p of the Klein-Gordon equation is explicitly given by the continuous function*

$$K_p(q) = \frac{1}{2} J_0(m|p-q|_\eta) \mathbb{1}_{J(p)}(q) \operatorname{sign}(t_p - t_q)\,, \tag{4.4.20}$$

where sign *is the sign function, i.e.* $\operatorname{sign}(s) = +1$ *for* $s > 0$ *and* $\operatorname{sign}(s) = -1$ *for* $s < 0$.

Notice that K_p is also continuous in p, so that we can view it as a function on $\mathbb{R}^{1,1} \times \mathbb{R}^{1,1}$. We will do so in the following and write $K(p,q) := K_p(q)$. Notice moreover that $K(p,q) = -K(q,p)$, which implies that K satisfies the Klein-Gordon equation in *both* the p- and the q-variable.

Finally, we note that in the null coordinates v and u, the causal fundamental solution is given by

$$K(v,u) = \frac{1}{2} J_0(m\sqrt{vu})\theta(vu) \operatorname{sign}(v+u)\,, \tag{4.4.21}$$

where $\theta = \mathbb{1}_{[0,\infty)}$ is the Heaviside function, i.e. the characteristic function of $[0,\infty) \subset \mathbb{R}$. Here we commit the slight crime of using the same symbol for K both as a function of t and r and as a function of v and u. This should not cause any confusion in the following, since only the null coordinates v and u are used. We will make use of this formula later.

4.4.4. Solving the Goursat problem with the causal fundamental solution. With the just collected information about the causal fundamental solution we will now derive the announced representation formula for solutions of the Goursat problem (4.4.2). The idea is that since the causal fundamental solution $K(p,q)$ is a solution of the homogeneous Klein-Gordon equation (in either variable), so will be any convolution integral $\int K(p,q) f(q)\,\mathrm{d}q$. Therefore the plan is to choose the function f in such a way that this integral also has the correct initial values of the Goursat problem (4.4.2).

Proposition 4.4.9. *Let $\phi_0, \phi_1 \in C^\infty[0,\infty)$ with $\phi_0(0) = \phi_1(0)$. Then there exists a (unique) smooth solution $\phi \in C^\infty(J^+(0))$ of the Goursat problem* (4.4.2)*. Expressed in null coordinates, it is explicitly given by*

$$\begin{aligned}\phi(v,u) &= 2\int_0^\infty K(v,u-x)\phi_0'(x)\,\mathrm{d}x + 2\int_0^\infty K(v-y,u)\phi_1'(y)\,\mathrm{d}y + 2\phi_0(0)K(v,u)\\ &= \int_0^u J_0(m\sqrt{v(u-x)})\phi_0'(x)\mathrm{d}x + \int_0^v J_0(m\sqrt{(v-y)u})\phi_1'(y)\mathrm{d}y + \phi_0(0)J_0(m\sqrt{vu})\,.\end{aligned} \tag{4.4.22}$$

Proof. Since we already know from Corollary 4.4.6 that there can be at most one smooth solution with characteristic initial values ϕ_0 and ϕ_1, it suffices to show that the right-hand side of (4.4.22) is in fact a smooth solution with these initial values.

First of all, let us address smoothness of the right-hand side of (4.4.22). At first sight, one might be skeptical about this due to the appearance of the square root. To clarify this, we use the (convergent) Taylor series of J_0 (cf. [**BW10**, Ch.7, eq.(7.1.2)]), which reads

$$J_0(z) \sum_{k=0}^\infty \frac{(-1)^k}{k!k!} \left(\frac{z}{2}\right)^{2k}\,.$$

Since here only *even* powers of z appear, the square root in (4.4.22) does not cause any problem, and hence the right-hand side of (4.4.22) is smooth on all of $J^+(0)$.

Next, since K satisfies the Klein-Gordon equation in both variables, so does the right-hand side of (4.4.22). Just to be sure, let us verify this explicitly for the first integral. Denoting this term by $I_1(v,u)$, we have

$$\partial_u I_1 = \phi_0'(u) + \int_0^u J_0'(m\sqrt{v(u-x)})\,\frac{m\sqrt{v}}{2\sqrt{u-x}}\,\phi_0'(x)\,\mathrm{d}x\,,$$

and

$$\begin{aligned}\partial_v\partial_u I_1 &= \int_0^u J_0''(m\sqrt{v(u-x)})\,\frac{m\sqrt{u-x}}{2\sqrt{v}}\,\frac{m\sqrt{v}}{2\sqrt{u-x}}\,\phi_0'(x)\,\mathrm{d}x \\ &\quad + \int_0^u J_0'(m\sqrt{v(u-x)})\,\frac{m}{4\sqrt{v(u-x)}}\,\phi_0'(x)\,\mathrm{d}x\,.\end{aligned}$$

Using that $\Box = -4\partial_v\partial_u$ it follows that

$$\begin{aligned}&(\Box - m^2) I_1(v,u) \\ &= -\int_0^u \left[m^2 J_0''(m\sqrt{v(u-x)}) + \frac{m J_0'(m\sqrt{v(u-x)})}{\sqrt{v(u-x)}} + m^2 J_0(m\sqrt{v(u-x)})\right]\phi_0'(x)\,\mathrm{d}x \\ &= -\int_0^u \left[J_0''(m\sqrt{v(u-x)}) + \frac{J_0'(m\sqrt{v(u-x)})}{m\sqrt{v(u-x)}} + J_0(m\sqrt{v(u-x)})\right] m^2\phi_0'(x)\,\mathrm{d}x\,.\end{aligned}$$

Now we note that the terms in brackets cancel due to the Bessel equation (cf. [**BW10**, (7.0.1)])

$$J_0''(z) + \frac{J_0'(z)}{z} + J_0(z) = 0\,,$$

which shows that $(\Box - m^2)I_1 = 0$. Similarly one sees that the second term also satisfies the Klein-Gordon equation, and for the last term it follows even simpler.

It remains to verify that the right-hand side of (4.4.22) has the correct initial values at $v=0$ and $u=0$. Using that $J_0(0)=1$, it follows that

$$\phi(0,u) = \int_0^u \phi_0'(x)\,\mathrm{d}x + \phi_0(0) = \phi_0(u)\,.$$

Since $\phi_0(0) = \phi_1(0)$, we also have

$$\phi(v,0) = \int_0^v \phi_1'(y)\,\mathrm{d}y + \phi_1(0) = \phi_1(v)\,.$$

This concludes the proof. $\Box$

Remark 4.4.10. One can also understand the integral formula (4.4.22) from the perhaps more familiar integral formula for solutions of the Cauchy problem, which reads

$$\phi(p) = \int_\Sigma (\nu_\Sigma K(p,q))\phi(q)\,\mathrm{d}\mu_\Sigma(q) + \int_\Sigma K(p,q)(\nu_\Sigma\phi)(q)\,\mathrm{d}\mu_\Sigma(q)\,. \tag{4.4.23}$$

Here $\Sigma \subset \mathbb{R}^{1,1}$ is any Cauchy hypersurface with future-directed unit-normal ν_Σ, and induced volume element $\mathrm{d}\mu_\Sigma$. In this formula, the normal ν_Σ acts as directional derivative on K (in the p-variable) and ϕ. To make the connection to formula (4.4.22) for the Goursat problem, choose a family of Cauchy hypersurfaces Σ_ϵ which approach the cone $\mathcal{C}$ (where initial data for the Goursat problem are posed) as sketched in figure 4.11. The important point to note is that in the limit as the Cauchy hypersurfaces Σ_ϵ approach the null hypersurface $\mathcal{C}$, the normal ν_{Σ_ϵ} becomes *tangent* to $\mathcal{C}$. Therefore, in the limit, we

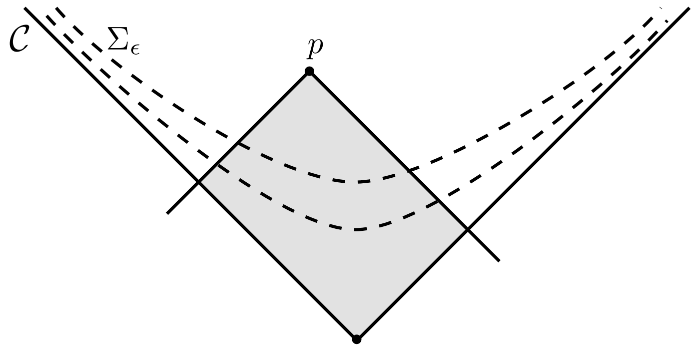

FIGURE 4.11. Approximating the Goursat problem by regular Cauchy problems.

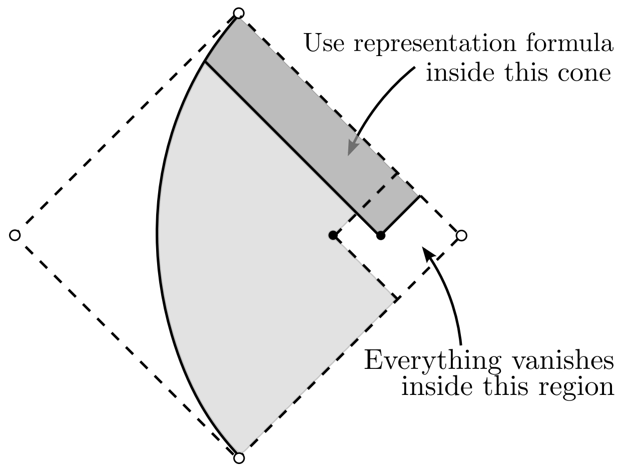

FIGURE 4.12. Illustration for why $\phi_1 \equiv 0$ in the situation where we are going to apply formula (4.4.22).

can partially integrate in the first integral to move the derivative from K to ϕ. Carrying this out, one can recover formula (4.4.22).

4.4.5. The basic decay estimate for solutions of the simple propagation. Using the representation formula (4.4.22), we now show that solutions of the Goursat problem (4.4.2), and thus also solutions of the free part of equation (4.3.8), decay for $v \to \infty$. From (4.4.22) one sees that a general solution ϕ of the Goursat problem (4.4.2) need not decay at all since the second integral (the one over v) does not necessarily decay. However, due to the *support assumption* (4.3.3) which we made in our problem, we only need to consider the case where $\phi_1 \equiv 0$. To see this more clearly, figure 4.12 illustrates the precise cone in which we will apply formula (4.4.22).

The previous argument shows that for our purposes it suffices to study the decay in v of functions of the form

$$\phi(v,u) = \int_0^u J_0(m\sqrt{v(u-x)})\phi_0'(x)\,\mathrm{d}x \tag{4.4.24}$$

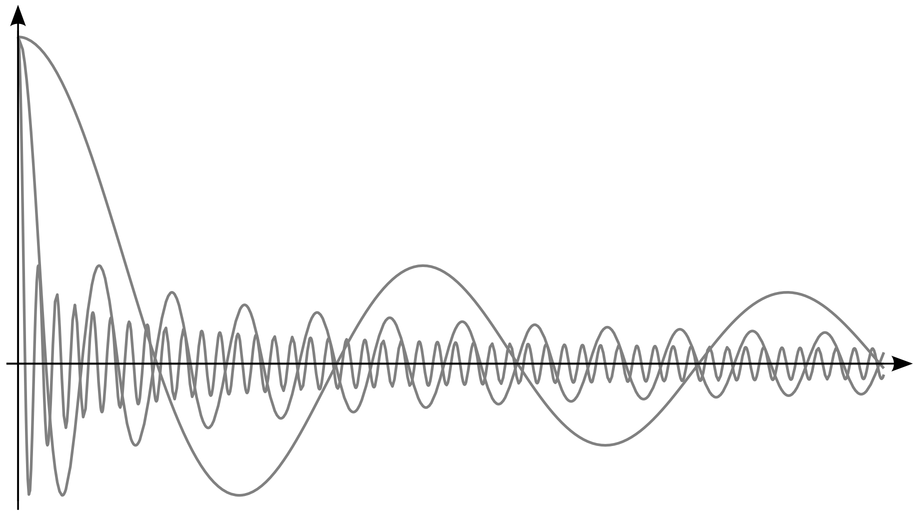

FIGURE 4.13. The plot shows the functions $J_0(x), J_0(5x), J_0(20x)$ which oscillate ever faster as the "frequency" increases from 1 over 5 to 20.

for "initial values"

$$\phi_0 \in \underline{C}^\infty(0,\infty) := \{\varphi \in C^\infty(0,\infty) \mid \exists \epsilon > 0 : \operatorname{supp}\varphi \subset [\epsilon,\infty)\}\,.^6 \tag{4.4.25}$$

The reason why the integral (4.4.24) decays (fast) for $v \to \infty$ is that the Bessel function $J_0(m\sqrt{v(u-x)})$ oscillates ever faster (as function of x) as the "frequency" v increases, see figure 4.13. This produces cancellations in the integral (4.4.24) ("destructive interferences"), similar as in the well-known method of (non-) stationary phase (cf. [**Zwo12**, Ch.3]). In practice these oscillations are exploited by an integration by parts, and the starting point is the following identity for Bessel functions.

Lemma 4.4.11. *For any $k \in \mathbb{N}$ and all $u, v \geq 0$, we have*

$$(vu)^{\frac{k}{2}} J_k(\sqrt{vu}) = \frac{2}{v}\frac{\mathrm{d}}{\mathrm{d}u}\Big((vu)^{\frac{k+1}{2}} J_{k+1}(\sqrt{vu})\Big)\,. \tag{4.4.26}$$

PROOF. The assertion is an immediate consequence of the well-known identity (cf. [**BW10**, Ch.7, eq.(7.1.4)])

$$\left(\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\right)^n \Big[z^k J_k(z)\Big] = z^{k-n} J_{k-n}(z)\,. \tag{$*$}$$

To see this, for fixed $v > 0$ we introduce the new variable $U = \sqrt{vu}$. Then we have

$$\frac{\mathrm{d}}{\mathrm{d}u} = \frac{\mathrm{d}U}{\mathrm{d}u}\frac{\mathrm{d}}{\mathrm{d}u} = \frac{v}{2U}\frac{\mathrm{d}}{\mathrm{d}u}\,,$$

so that

$$\begin{aligned}\frac{2}{v}\frac{\mathrm{d}}{\mathrm{d}u}\Big((vu)^{\frac{k+1}{2}} J_{k+1}(\sqrt{vu})\Big) &= \frac{1}{U}\frac{\mathrm{d}}{\mathrm{d}U}\Big(U^{k+1}J_{k+1}(U)\Big)\\ &\overset{(*)}{=} U^k J_k(U)\\ &= (vu)^{\frac{k}{2}} J_k(\sqrt{vu})\,.\end{aligned}$$

[6]The underline in the notation should emphasize that the functions in this space are supposed to vanish at the lower boundary of the interval $(0,\infty)$.

This is just what we wanted to show. $\square$

We now use this identity to successively integrate by parts in the integrals (4.4.24), producing decaying factors in this way.

Lemma 4.4.12. *Let $\phi_0 \in \underline{C}^\infty(0,\infty)$, and let $\phi(v,u)$ be given by the integral (4.4.24). Then for all $k \geq 0$ we have*

$$\begin{aligned} \phi(v,u) &= \left(\frac{2}{v}\right)^k \frac{1}{m^{2(k+1)}} \int_0^{m^2 u} (vz)^{\frac{k}{2}} J_k(\sqrt{vz}) \phi_0^{(k+1)}\left(u - \frac{z}{m^2}\right) \mathrm{d}z \\ &= \left(\frac{2}{m^2 v}\right)^k \int_0^u \left(m\sqrt{v(u-x)}\right)^k J_k(m\sqrt{v(u-x)}) \phi_0^{(k+1)}(x)\, \mathrm{d}x\,. \end{aligned} \tag{4.4.27}$$

Proof. Substituting x for $z = m^2(u-x)$ in the integral (4.4.24) and integrating by parts using (4.4.26) gives

$$\begin{aligned} \phi(v,u) &= \int_0^u J_0\left(m\sqrt{v(u-x)}\right) \phi_0'(x)\, \mathrm{d}x \\ &= \frac{1}{m^2} \int_0^{m^2 u} J_0(\sqrt{vz}) \phi_0'\left(u - \frac{z}{m^2}\right) \mathrm{d}z \\ &\overset{(4.4.26)}{=} \frac{1}{m^2} \int_0^{m^2 u} \frac{2}{v} \frac{\mathrm{d}}{\mathrm{d}z} \underbrace{\left[(vz)^{\frac{1}{2}} J_1(\sqrt{vz})\right]}_{\text{vanishes at } z=0} \underbrace{\phi_0'\left(u - \frac{z}{m^2}\right)}_{\text{vanishes at } z=m^2 u} \mathrm{d}z \\ &= -\frac{1}{m^2} \frac{2}{v} \int_0^{m^2 u} (vz)^{\frac{1}{2}} J_1(\sqrt{vz}) \frac{\mathrm{d}}{\mathrm{d}z}\left[\phi_0'\left(u - \frac{z}{m^2}\right)\right] \mathrm{d}z \\ &= \frac{1}{m^4} \frac{2}{v} \int_0^{m^2 u} (vz)^{\frac{1}{2}} J_1(\sqrt{vz}) \phi_0''\left(u - \frac{z}{m^2}\right) \mathrm{d}z \\ &= \ldots \text{(repeat the previous steps another } (k-1) \text{ times)} \\ &= \left(\frac{2}{v}\right)^k \frac{1}{m^{2(k+1)}} \int_0^{m^2 u} (vz)^{\frac{k}{2}} J_k(\sqrt{vz}) \phi_0^{(k+1)}\left(u - \frac{z}{m^2}\right) \mathrm{d}z\,. \end{aligned}$$

Notice that we needed the assumption that $\phi_0 \in \underline{C}^\infty(0,\infty)$ in order to avoid boundary terms at 0 when integrating by parts. This shows the first identity in (4.4.27), the second one follows by resubstituting z for x. $\square$

The relevant point to note about (4.4.27) is of course the decaying factor v^{-k} in front of the integral. In order to deduce decay in v, we also need to take care of the appearance of v inside the integral. Here we can use that all Bessel functions are bounded as $x \to 0$, and decay as $x^{-\frac{1}{2}}$ for $x \to \infty$ (cf. [**BW10**, Ch.7, eq.(7.4.8)]). Therefore for every $k \in \mathbb{N}$ there exists a constant $C_k > 0$ such that

$$|J_k(x)| \leq C_k |x|^{-\frac{1}{2}} \qquad \forall x > 0\,. \tag{4.4.28}$$

Using this we can now prove decay.

In the following, we use the standard L^2-Sobolev norms

$$\|\varphi\|^2_{H^k(a,b)} = \sum_{j=0}^k \|\varphi^{(j)}\|^2_{L^2(a,b)}\,, \tag{4.4.29}$$

where $\varphi^{(j)}$ denotes the j-th derivative of φ.

Proposition 4.4.13. (Decay of free propagation 1) *For every $k \in \mathbb{N}$ there exists a constant $C_k > 0$ such that the following holds: For every $\phi_0 \in \underline{C}^\infty(0,\infty)$, the function $\phi(v,u)$ defined by* (4.4.24) *satisfies*

$$|\phi(v,u)| \leq \frac{C_k}{(1+v)^{\frac{2k+1}{4}}}(1+u)^{\frac{k+1}{2}} \|\phi_0\|_{H^{k+1}(0,u)} \qquad \forall v,u \geq 0\,. \tag{4.4.30}$$

Proof. On the one hand, we have the global bound

$$\begin{aligned}
|\phi(v,u)| &\leq \int_0^u \underbrace{\left|J_0(m\sqrt{v(u-x)})\right|}_{\leq 1} \left|\phi_0'(x)\right| \mathrm{d}x \\
&\leq \int_0^u \left|\phi_0'(x)\right| \mathrm{d}x \\
&\leq u^{\frac{1}{2}} \|\phi_0\|_{H^1(0,u)} \\
&\leq u^{\frac{1}{2}} \|\phi_0\|_{H^{k+1}(0,u)}\,.
\end{aligned} \tag{$*$}$$

Here we used Hölder's inequality in the second to last step. Notice that $\phi_0 \in \underline{C}^\infty(0,\infty)$ implies that $\|\phi_0\|_{H^k(0,u)} < \infty$ for all $u \in (0,\infty)$. On the other hand, from (4.4.27) and (4.4.28) it follows that, for a suitable constant $C_k > 0$ changing from line to line,

$$\begin{aligned}
|\phi(v,u)| &\leq \frac{C_k}{v^k} \int_0^u \left(m\sqrt{v(u-x)}\right)^k \underbrace{\left|J_k(m\sqrt{v(u-x)}\right|}_{\leq C_k\left(m\sqrt{v(u-x)}\right)^{-\frac{1}{2}}} |\phi_0^{(k+1)}(x)| \,\mathrm{d}x \\
&\leq \frac{C_k}{v^{\frac{k}{2}+\frac{1}{4}}} \int_0^u (u-x)^{\frac{k}{2}-\frac{1}{4}} |\phi_0^{(k+1)}(x)| \,\mathrm{d}x \\
&\leq \frac{C_k}{v^{\frac{k}{2}+\frac{1}{4}}} u^{\frac{k}{2}+\frac{1}{4}} \|\phi_0\|_{H^{k+1}(0,u)}\,.
\end{aligned} \tag{$**$}$$

The last step was again an application of Hölder's inequality. Combining $(*)$ and $(**)$, we obtain

$$\begin{aligned}
|\phi(v,u)| &\leq \min\left\{u^{\frac{1}{2}} \|\phi_0\|_{H^{k+1}(0,u)}\,,\, \frac{C_k}{v^{\frac{k}{2}+\frac{1}{4}}} u^{\frac{k}{2}+\frac{1}{4}} \|\phi_0\|_{H^{k+1}(0,u)}\right\} \\
&\leq \frac{C_k}{(1+v)^{\frac{k}{2}+\frac{1}{4}}} \left(u^{\frac{1}{2}} + u^{\frac{k}{2}+\frac{1}{4}}\right) \|\phi_0\|_{H^{k+1}(0,u)} \\
&\leq \frac{C_k}{(1+v)^{\frac{2k+1}{4}}} (1+u)^{\frac{k+1}{2}} \|\phi_0\|_{H^{k+1}(0,u)}\,.
\end{aligned}$$

∎

Instead of a Sobolev norm, we can of course also use pointwise norms.

Proposition 4.4.14. (Decay of free propagation 2) *For every $k \in \mathbb{N}$ there exists a constant $C_k > 0$ such that the following holds: For every $\phi_0 \in \underline{C}^\infty(0,\infty)$, the function $\phi(v,u)$ defined by* (4.4.24) *satisfies*

$$|\phi(v,u)| \leq \frac{C_k}{(1+v)^{\frac{2k+1}{4}}}(1+u)^{\frac{k+3}{2}} \|\phi_0\|_{C^{k+1}(0,u)} \qquad \forall v,u \geq 0\,. \tag{4.4.31}$$

Proof. One can proceed similar as in the proof of Proposition 4.4.13, only that in both estimates instead of using Hölder's inequality in the end one rather directly estimates the ϕ_0-derivatives by their supremum and evaluates the remaining integral explicitly. Notice that $\phi_0 \in \underline{C}^\infty(0,\infty)$ implies that $\|\phi_0\|_{C^{k+1}(0,u)} < \infty$ for all $u \in (0,\infty)$. $\square$

4.4.6. Some function space estimates related to the free propagation. So far we have obtained a direct *pointwise* estimate for ϕ defined by (4.4.24) in terms of Sobolev norms of the initial data ϕ_0. We are now going to derive some related estimates in various functions spaces. To achieve a more clearly structured presentation, we introduce the '“free propagator in v-direction”.

Definition 4.4.15. For every $v \in (0,\infty)$, we define $U_0(v) : \underline{C}^\infty((0,\infty)_u) \to \underline{C}^\infty((0,\infty)_u)$ by

$$(U_0(v)\phi_0)(u) := \phi(v,u)\,, \tag{4.4.32}$$

where $\phi \in \mathcal{C}^\infty((0,\infty)_v \times (0,\infty)_u)$ is the unique solution of the Goursat problem (4.4.2) with $\phi_1 = 0$. We call this family of operators the *free propagation in v-direction.*

From the representation formula (4.4.23) for solutions of the Goursat problem (4.4.2) we already know that $U_0(v)$ is an “integro-differential operator” Moreover, by Lemma 4.4.12 we know that we can actually represent it in different ways. To write these in compact form, we introduce the following “integral kernels”.

Definition 4.4.16. For every $k \in \mathbb{N}$ we define $K_0^{(k)} : (0,\infty)_v \times (0,\infty)_u \times (0,\infty)_x \to \mathbb{R}$ by

$$K_0^{(k)}(v,u;x) := \left(\frac{2}{m^2 v}\right)^k \left(m\sqrt{v(u-x)}\right)^k J_k(m\sqrt{v(u-x)})\theta(u-x)\,. \tag{4.4.33}$$

With these kernels, for any $k \in \mathbb{N}$ we have

$$(U_0(v)\phi_0)(u) = \int_0^\infty K_0^{(k)}(v,u;x)\phi_0^{(k+1)}(x)\,\mathrm{d}x \qquad \forall \phi_0 \in \underline{C}^\infty(0,\infty)\,. \tag{4.4.34}$$

If we want to stress which particular of these representations we use, we will write $U_0^{(k)}(v)$ instead of $U_0(v)$ for clarity.

In the remainder of this section we will express the decay estimates obtained in the previous section in terms of operator norm estimates of the operators $U_0(v)$. To this end, the following Sobolev spaces will be used.

Definition 4.4.17. For each $u_1 \in (0,\infty)$ we define the space

$$\underline{H}^k(0,u_1) := \overline{\underline{C}^\infty(0,u_1)}^{\|\cdot\|_{H^k(0,u_1)}}\,, \tag{4.4.35}$$

where the closure is taken in the usual L^2-Sobolev norm (4.4.29). Here similar to (4.4.25) for any interval $(a,b) \subset \mathbb{R}$ we set

$$\underline{C}^\infty(a,b) := \{\varphi \in C^\infty(a,b) \mid \exists \epsilon > 0 : \operatorname{supp}\varphi \subset [a+\epsilon, b)\}\,. \tag{4.4.36}$$

Now we estimate the free propagator $U_0(v)$ in various function spaces.

Lemma 4.4.18. **(H^{k+1} to C^0 estimate of the free propagation)** *For each $k \in \mathbb{N}$, let $C_k > 0$ be the same constant as in Proposition 4.4.13. Then for any $\phi_0 \in \underline{C}^\infty(0,\infty)$ the function $\phi \in C^\infty((0,\infty)_v \times (0,\infty)_u)$ given by (4.4.24) satisfies*

$$\|\phi(v,\cdot)\|_{C^0(0,u_1)} \le \frac{C_k}{(1+v)^{\frac{2k+1}{4}}}(1+u_1)^{\frac{k+1}{2}}\,\|\phi_0\|_{H^{k+1}(0,u_1)} \qquad \forall u_1 > 0\,. \tag{4.4.37}$$

Consequently, for any $u_1 > 0$ the free propagation $U_0^{(k)}(v)$ extends to a family of bounded operators

$$U_0^{H^{k+1},C^0}(v) : \underline{H}^{k+1}(0,u_1) \longrightarrow C^0(0,u_1)\,. \tag{4.4.38}$$

The action of $U_0^{H^{k+1},C^0}(v)$ is explicitly given by (4.4.34).

Proof. This is obvious from (4.4.30). $\square$

Lemma 4.4.19. (H^{k+1} to L^p estimate of the free propagation) *For each $k \in \mathbb{N}$, let $C_k > 0$ be the same constant as in Proposition 4.4.13. Then for any $\phi_0 \in \underline{C}_0^\infty(0,\infty)$ the function $\phi \in C^\infty((0,\infty)_v \times (0,\infty)_u)$ given by (4.4.24) satisfies*

$$\|\phi(v,\cdot)\|_{L^p(0,u_1)} \le \frac{C_k}{(1+v)^{\frac{2k+1}{4}}}(1+u_1)^{\frac{k+1}{2}+\frac{1}{p}}\,\|\phi_0\|_{H^{k+1}(0,u_1)} \qquad \forall u_1 > 0\,. \tag{4.4.39}$$

Consequently, for any $u_1 > 0$ the free propagation $U_0^{(k)}(v)$ extends to a family of bounded operators

$$U_0^{H^{k+1},L^p}(v) : \underline{H}^{k+1}(0,u_1) \longrightarrow L^p(0,u_1)\,, \tag{4.4.40}$$

The action of $U_0^{H^{k+1},L^p}(v)$ is explicitly given by (4.4.34).

Proof. We simply take the L^p-norm of the estimate (4.4.30) and make the crude estimate

$$\begin{aligned}\left\|u^{\frac{1}{2}} + u^{\frac{k}{2}+\frac{1}{4}}\right\|_{L^p(0,u_1)} &= \left(\int_0^{u_1} \left(u^{\frac{1}{2}} + u^{\frac{k}{2}+\frac{1}{4}}\right)^p \mathrm{d}u\right)^{\frac{1}{p}} \\ &\le \left(u_1^{\frac{1}{2}} + u_1^{\frac{k}{2}+\frac{1}{4}}\right)\cdot u_1^{\frac{1}{p}} \\ &= u_1^{\frac{1}{2}+\frac{1}{p}} + u_1^{\frac{k}{2}+\frac{1}{4}+\frac{1}{p}} \\ &\le (1+u_1)^{\frac{k+1}{2}+\frac{1}{p}}\,.\end{aligned}$$

$\square$

Lemma 4.4.20. (C^{k+1} to C^0 estimate of the free propagation) *For each $k \in \mathbb{N}$, let $C_k > 0$ be the same constant as in Proposition 4.4.13. Then for any $\phi_0 \in \underline{C}^\infty(0,\infty)$,the function $\phi \in C^\infty((0,\infty)_v \times (0,\infty)_u)$ given by (4.4.24) satisfies*

$$\|\phi(v,\cdot)\|_{C^0(0,u_1)} \le \frac{C_k}{(1+v)^{\frac{2k+1}{4}}}(1+u_1)^{\frac{k+3}{2}}\,\|\phi_0\|_{C^{k+1}(0,u_1)} \qquad \forall u_1 > 0\,. \tag{4.4.41}$$

Consequently, for any $u_1 > 0$ the free propagation $U_0^{(k)}(v)$ extends to a family of bounded operators

$$U_0^{C^{k+1},C^0}(v) : C^{k+1}(0,u_1) \longrightarrow C^0(0,u_1)\,. \tag{4.4.42}$$

The action of $U_0^{C^{k+1},C^0}(v)$ is explicitly given by (4.4.34).

Proof. This is obvious from (4.4.31). $\square$

Lemma 4.4.21. (C^{k+1} to L^p estimate of the free propagation) *For each $k \in \mathbb{N}$, let $C_k > 0$ be the same constant as in Proposition 4.4.13. Then for any $\phi_0 \in \underline{C}^\infty(0,\infty)$ the function $\phi \in C^\infty((0,\infty)_v \times (0,\infty)_u)$ given by (4.4.24) satisfies*

$$\|\phi(v,\cdot)\|_{L^p(0,u_1)} \le \frac{C_k}{(1+v)^{\frac{2k+1}{4}}}(1+u_1)^{\frac{k+3}{2}+\frac{1}{p}}\,\|\phi_0\|_{C^{k+1}(0,u_1)} \qquad \forall u_1 > 0\,. \tag{4.4.43}$$

Consequently, for any $u_1 > 0$ the free propagation $U_0^{(k)}(v)$ extends to a family of bounded operators

$$U_0^{C^{k+1},L^p}(v) : C^{k+1}(0,u_1) \longrightarrow L^p(0,u_1)\,. \tag{4.4.44}$$

The action of $U_0^{C_0^{k+1}L^p}(v)$ is explicitly given by (4.4.34).

Proof. We simply take the L^p-norm of the estimate (4.4.31) and make the crude estimate

$$\begin{aligned}\left\|u^{\frac{3}{2}} + u^{\frac{k}{2}+\frac{3}{4}}\right\|_{L^p(0,u_1)} &= \left(\int_0^{u_1} \left(u^{\frac{3}{2}} + u^{\frac{k}{2}+\frac{3}{4}}\right)^p \mathrm{d}u\right)^{\frac{1}{p}} \\ &\leq \left(u_1^{\frac{3}{2}} + u_1^{\frac{k}{2}+\frac{3}{4}}\right) \cdot u_1^{\frac{1}{p}} \\ &= u_1^{\frac{3}{2}+\frac{1}{p}} + u_1^{\frac{k}{2}+\frac{3}{4}+\frac{1}{p}} \\ &\leq (1+u_1)^{\frac{k+3}{2}+\frac{1}{p}}\,.\end{aligned}$$

$\square$

4.5. Treatment of the Perturbation

Having derived decay properties of the free part alone, the next step is to perturbatively take into account the remainder. As in most perturbative approaches this requires three different inputs: Firstly, one of course needs a general strategy for how to deal with the perturbation at all. Here we will use a simple identity which is known as *Lippmann-Schwinger equation* or *Duhamal's formula* (see Section 4.6.1). Secondly, one needs suitable *estimates of the remainder terms* to ensure that they may really be treated as a "small perturbation" of the simple part. This is where we will make use of the asymptotic structure of the metrics as introduced in Section 3.2 (see Section 4.5.3 and Section 4.5.4). And finally, when closing the estimates (and typically also in the remainder estimates already), one needs suitable *a priori estimates* for general solutions of the full equation. Here one needs to use the structure of the full equation, e.g. by exploiting symmetry properties or conserved quantities. We start with these a priori estimates.

4.5.1. A priori estimates for solutions of the Dirac null system. In the following, we always consider solutions of the Dirac null system (4.2.10) which are defined on a diamond-shaped domain

$$D_{u_0,u_1}^{v_0,v_1} := \{p \in \mathbb{R}^{1,1} \mid v_0 \leq v(p) \leq v_1\,,\, u_0 \leq u(p) \leq u_1\} \subset \mathbb{R}^{1,1} \tag{4.5.1}$$

as sketched in figure 4.14, where v and u are the null coordinates defined in (4.2.6).[7] More precisely, we assume that $f, g \in C^\infty(D_{u_0,u_1}^{v_0,v_1})$ satisfy

$$\begin{cases} \partial_v f = V g\,, \\ \partial_u g = -\overline{V} g\,. \end{cases} \tag{4.5.2}$$

for some given (complex-valued) function $V \in C^\infty(D_{u_0,u_1}^{v_0,v_1})$.

The goal in the following is to derive "energy estimate" for f, i.e. L^2-estimates in the u-variable of f at fixed v-coordinate. Since the Dirac null system is just the Dirac

[7] Be aware that the splitting $\mathbb{R}^{1,1} = \mathbb{R}_t \times \mathbb{R}_r$ always refers to the usual coordinates t and r and not to the null coordinates v and u. This is why we do not simply write $D_{u_0,u_1}^{v_0,v_1} = [v_0,v_1]_v \times [u_0,u_1]_u$, since it may lead to confusion.

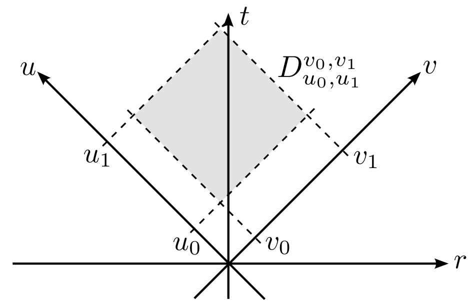

FIGURE 4.14. Sketch of the diamond-shaped domain $D^{v_0,v_1}_{u_0,u_1}$.

equation expressed in null coordinates, the basic structure we can make use of is the *conserved current* (cf. Section 2.1).

Proposition 4.5.1. (Current conservation) *Suppose that $f, g \in \mathcal{C}^\infty(D^{v_0,v_1}_{u_0,u_1})$ satisfy the Dirac null system (4.5.2) for some $V \in C^\infty(D^{v_0,v_1}_{u_0,u_1})$. Suppose moreover that $\partial_t V = 0$. Then for all $n \in \mathbb{N}$, we have*

$$\begin{aligned} &\int_{u_0}^{u} |\partial_t^n f(v_1,u)|^2 \, \mathrm{d}u + \int_{v_0}^{v} |\partial_t^n g(v,u_1)|^2 \, \mathrm{d}v \\ &= \int_{u_0}^{u} |\partial_t^n f(v_1,u)|^2 \, \mathrm{d}u + \int_{v_0}^{v} |\partial_t^n g(v,u_0)|^2 \, \mathrm{d}v \,. \end{aligned} \tag{4.5.3}$$

PROOF. We start with the case $n = 0$, and consider the vector field

$$j = |f|^2 \partial_v + |g|^2 \partial_u \,.$$

Using that f and g satisfy the Dirac null system (4.5.2), we find that

$$\begin{aligned} \operatorname{div} j &= \partial_v |f|^2 + \partial_u |g|^2 \\ &= 2\,\mathsf{Re}(\overline{f} \partial_v f) + 2\,\mathsf{Re}(g \partial_u \overline{g}) \\ &= 2\,\mathsf{Re}(\overline{f} V g) - 2\,\mathsf{Re}(g V \overline{f}) \\ &= 0 \,. \end{aligned}$$

From this and Gauss' theorem for diamond-shaped null domains (cf. Lemma 4.4.4), the assertion follows for the case $n = 0$.

Now let $n \in \mathbb{N}$ be arbitrary. Since $\partial_t V = 0$, differentiating the whole Dirac null system by t one finds that $\partial_t^n f, \partial_t^n g$ also satisfy the Dirac null system. Therefore one can simply repeat the previous steps to see that (4.5.3) holds for general $n \in \mathbb{N}$. ☐

We will also need similar estimates involving u-derivatives of f. Our starting point for deriving such estimates is the following identity, which expresses u-derivatives of f in terms of v-derivatives of f to which we can apply the Dirac null system, and t-derivatives of f which are already under control by Proposition 4.5.1.

Lemma 4.5.2. *Suppose that $f, g \in \mathcal{C}^\infty(D^{v_0,v_1}_{u_0,u_1})$ satisfy the Dirac null system* (4.5.2) *for some $V \in C^\infty(D^{v_0,v_1}_{u_0,u_1})$. Then for any $n \in \mathbb{N}$, we have*

$$\begin{aligned} \partial_u^n f \;&= \partial_u^{n-1} \partial_t f - \big(\partial_u^{n-1} V\big) g \\ &\quad + \sum_{k=1}^{n-1} \sum_{\ell=0}^{k-1} \binom{n-1}{k} \binom{k-1}{\ell} \big(\partial_u^{n-1-k} V\big) \big(\partial_u^{k-1-\ell} \overline{V}\big) \big(\partial_u^{\ell} f\big) \,. \end{aligned} \tag{4.5.4}$$

Proof. Using the definition of the null coordinates u and v and the Dirac null system (4.5.2), we have

$$\partial_u^n f = \partial_u^{n-1}(\partial_u f) = \partial_u^{n-1}(\partial_t f - \partial_v f) = \partial_u^{n-1}\partial_t f - \partial_u^{n-1}(Vg)\,. \tag{$*$}$$

By the Leibniz rule and another application of the Dirac null system (4.5.2), the last term can be rewritten as

$$\begin{aligned}
\partial_u^{n-1}(Vg) &= \sum_{k=0}^{n-1}\binom{n-1}{k}\big(\partial_u^{n-1-k}V\big)\big(\partial_u^k g\big)\\
&= \big(\partial_u^{n-1}V\big)g + \sum_{k=1}^{n-1}\binom{n-1}{k}\big(\partial_u^{n-1-k}V\big)\big(\partial_u^{k-1}\underbrace{\partial_u g}_{=-\overline{V}f}\big)\\
&= \big(\partial_u^{n-1}V\big)g - \sum_{k=1}^{n-1}\sum_{\ell=0}^{k-1}\binom{n-1}{k}\binom{k-1}{\ell}\big(\partial_u^{n-1-k}V\big)\big(\partial_u^{k-1-\ell}\overline{V}\big)\big(\partial_u^\ell f\big)\,.
\end{aligned}$$

Going back into $(*)$ with this, the assertion follows. □

If we want to use (4.5.4) to estimate $\partial_u^n f$, the basic problem is that the function g reappears on the right-hand side since we have used the equation for f. Therefore we also need an estimate for g in the u-variable, and this is not directly available from energy identities. Perhaps the simplest way to obtain such an estimate is to integrate the second equation of the Dirac null system (4.5.2).

Lemma 4.5.3. *Suppose that $f, g \in C^\infty(D_{u_0,u_1}^{v_0,v_1})$ satisfy the Dirac null system* (4.5.2) *for some $V \in C^\infty(D_{u_0,u_1}^{v_0,v_1})$. Then for any $v \in [v_0, v_1]$, we have*

$$\begin{aligned}
\|g(v,\cdot)\|_{L^2(u_0,u_1)} \;\le\; & |g(v,u_0)|\cdot|u_1-u_0|^{\frac{1}{2}}\\
&+ \|V(v,\cdot)\|_{L^\infty(u_0,u_1)}\,\|f(v,\cdot)\|_{L^2(u_0,u_1)}\cdot|u_1-u_0|\,.
\end{aligned} \tag{4.5.5}$$

Proof. Integrating the equation for g in the Dirac null system (4.5.2), we get

$$g(v,u) = g(v,u_0) + \int_{u_0}^{u}\overline{V}(v,x)f(v,x)\,\mathrm{d}x\,.$$

Now we simply take the L^2-norm in the u-variable of this and obtain

$$\|g(v,\cdot)\|_{L^2(u_0,u_1)} \le |g(v,u_0)|\cdot|u_1-u_0|^{\frac{1}{2}} + \left(\int_{u_0}^{u_1}\left|\int_{u_0}^{u}\overline{V}(v,x)f(v,x)\,\mathrm{d}x\right|^2\mathrm{d}u\right)^{\frac{1}{2}}.$$

Using Hölder's inequality we further estimate the second term as

$$\begin{aligned}
\int_{u_0}^{u_1}\left|\int_{u_0}^{u}\overline{V}(v,x)f(v,x)\,\mathrm{d}x\right|^2\mathrm{d}u &\le \int_{u_0}^{u_1}\left(\int_{u_0}^{u}|V(v,x)|^2\,\mathrm{d}x\right)\left(\int_{u_0}^{u}|f(v,x)|^2\,\mathrm{d}x\right)\mathrm{d}u\\
&\le \|V(v,\cdot)\|^2_{L^\infty(u_0,u_1)}\,\|f(v,\cdot)\|^2_{L^2(u_0,u_1)}\,|u_1-u_0|^2\,.
\end{aligned}$$

Putting this back into the previous estimate, we end up with (4.5.5). □

While (4.5.5) does still depend on g, it only does so by the values of g on the boundary $\{u = u_0\}$ of $D_{u_0,u_1}^{v_0,v_1}$. In our application, however, the *support assumption* (4.3.3) ensures that g vanishes on this boundary (see figure 4.12 and Section 4.3.1). Therefore, as far as we are concerned, the right-hand side of (4.5.5) is independent of g. More precisely, we obtain the following recursive L^2-estimate for the u-derivatives of f at fixed v.

Lemma 4.5.4. *Suppose that $f, g \in C^\infty(D_{u_0,u_1}^{v_0,v_1})$ satisfy the Dirac null system (4.5.2) for some $V \in C^\infty(D_{u_0,u_1}^{v_0,v_1})$, and suppose moreover that $g(\cdot, u_0) \equiv 0$. Then for any $n \in \mathbb{N}$, we have*

$$\begin{aligned}\|\partial_u^n f(v,\cdot)\|_{L^2(u_0,u_1)} \leq{}& \left\|\partial_u^{n-1}\partial_t f(v,\cdot)\right\|_{L^2(u_0,u_1)} \\ &+ \|V(v,\cdot)\|^2_{C_u^{n-1}(u_0,u_1)} \sum_{k=1}^{n-1}\sum_{\ell=0}^{k-1} \binom{n-1}{k}\binom{k-1}{\ell} \left\|\partial_u^\ell f(v,\cdot)\right\|_{L^2(u_0,u_1)} \\ &+ \|f(v_0,\cdot)\|_{L^2(u_0,u_1)} \|V(v,\cdot)\|^2_{C_u^{n-1}(u_0,u_1)} |u_1 - u_0| .\end{aligned} \tag{4.5.6}$$

Proof. This follows immediately by taking the L^2-norm of the recursion formula (4.5.4) for $\partial_u^n f$ and then using the L^2-estimate (4.5.5) for the term containing g together with the vanishing assumption of $g(\cdot, u_0)$. □

We can now use this recursive estimate to inductively establish higher order a priori estimates for $f(v,\cdot)$ in terms of V and the initial data $f(v_0,\cdot)$ alone, i.e. not containing g anymore. Be aware in the following that all functions are to be understood as functions of v and u, and the symbol ∂_t stands for the differential operator $\partial_t = \partial_v + \partial_u$ in these coordinates.

Proposition 4.5.5. (Higher order a priori estimates) *For every $n \in \mathbb{N}$ there exists a constant $C_n > 0$ such that the following holds: Let $D_{u_0,u_1}^{v_0,v_1} \subset \mathbb{R}^{1,1}$ be any diamond-shaped null domain, and let $f, g, \in C^\infty(D_{u_0,u_1}^{v_0,v_1})$ be a solution of the Dirac null system (4.5.2) for some $V \in C^\infty(D_{u_0,u_1}^{v_0,v_1})$ with $\partial_t V = 0$. Assume moreover that $g(\cdot, u_0) \equiv 0$. Then it holds that*

$$\begin{aligned}\|f(v,\cdot)\|_{H_u^n(u_0,u_1)} \leq{}& C_n \cdot \left(1 + \|V(v,\cdot)\|^{2n}_{C_u^{n-1}(u_0,u_1)}\right) \cdot \left(1 + |u_1 - u_0|\right) \\ &\times \left\{\|\partial_t^n f(v_0,\cdot)\|_{L_u^2(u_0,u_1)} + \ldots + \|f(v_0,\cdot)\|_{L_u^2(u_0,u_1)}\right\}\end{aligned} \tag{4.5.7}$$

for all $v \in [v_0, v_1]$.

Proof. The proof proceeds by induction on n. For brevity, we will drop the explicit reference to the interval (u_0, u_1) from the notation in the following.

For $n = 0$, current conservation (4.5.3) and $g(\cdot, u_0) = 0$ yield

$$\|f(v,\cdot)\|_{L_u^2} \leq \|f(v_0,\cdot)\|_{L_u^2} \leq \left(1 + |u_1 - u_0|\right) \cdot \|f(v_0,\cdot)\|_{L_u^2} .$$

This shows (4.5.7) for $n = 0$ with $C_0 = 1$.

Suppose now that (4.5.7) holds for all the natural numbers $0, 1, \ldots, n-1$. Notice once more that due to $\partial_t V = 0$ also $\partial_t^k f, \partial_t^k g$ satisfy the Dirac null system for any $k \in \mathbb{N}$. Therefore the estimate (4.5.7) also holds for u-derivatives of order at most $n-1$ of $\partial_t^k f$. With this in mind and (4.5.6), we have[8]

$$\begin{aligned}\|\partial_u^n f(v,\cdot)\|_{L_u^2} \leq{}& \underbrace{\|\partial_t f(v,\cdot)\|_{H_u^{n-1}}}_{\text{A}} + \underbrace{C_n \|V(v,\cdot)\|^2_{C_u^{n-1}} \|f(v,\cdot)\|_{H_u^{n-1}}}_{\text{B}} \\ &+ \underbrace{\|V(v,\cdot)\|^2_{C_u^{n-1}} |u_1 - u_0| \, \|f(v_0,\cdot)\|_{L_u^2}}_{\text{C}}\end{aligned}$$

[8]In the following, the constant C_n varies from line to line, but always only depends on n.

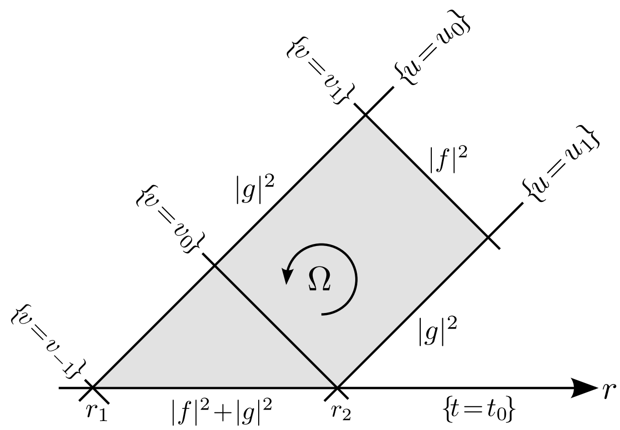

FIGURE 4.15. Estimating Goursat data in terms of Cauchy data.

$$
\begin{aligned}
\leq\ & \left.\begin{aligned} & C_{n-1}\big(1+\|V(v,\cdot)\|^{2(n-1)}_{C^{n-2}_u}\big)\big(1+|u_1-u_0|\big) \\ & \times\Big\{\big\|\partial_t^{n-1}(\partial_t f)(v_0,\cdot)\big\|_{L^2_u}+\ldots+\|(\partial_t f)(v_0,\cdot)\|_{L^2_u}\Big\}\end{aligned}\right\}\text{A} \\
& \left.\begin{aligned} + \ & C_n\,\|V(v,\cdot)\|^2_{C^{n-1}_u}\,C_{n-1}\big(1+\|V(v,\cdot)\|^{2(n-1)}_{C^{n-2}_u}\big)\big(1+|u_1-u_0|\big) \\ & \times\Big\{\big\|\partial_t^{n-1}f(v_0,\cdot)\big\|_{L^2_u}+\ldots+\|f(v_0,\cdot)\|_{L^2_u}\Big\}\end{aligned}\right\}\text{B} \\
& \quad + \|V(v,\cdot)\|^2_{C^{n-1}_u}\,|u_1-u_0|\,\|f(0,\cdot)\|_{L^2_u}\Big\}\ \text{C} \\
\leq\ & C_n\cdot\big(1+\|V(v,\cdot)\|^{2n}_{C^{n-1}_u}\big)\big(1+|u_1-u_0|\big) \\
& \times\Big\{\|\partial_t^n f(v_0,\cdot)\|_{L^2_u}+\ldots+\|f(v_0,\cdot)\|_{L^2_u}\Big\}\,.
\end{aligned}
$$

This is precisely what we wanted to show. □

4.5.2. Control of Goursat data by Cauchy data. So far we have obtained estimates for the Sobolev norms of $f(v,\cdot)$ in the u-variable in terms of norms of $f(v_0,\cdot)$. Ultimately, however, it would be nice to estimate $f(v,\cdot)$ against Cauchy data, i.e. against $f|_{t=t_0}$. To this end we have to relate the functions $f|_{v=v_0}$ and $f|_{t=t_0}$, which can be accomplished using current conservation once more, integrating over a domain $\Omega\subset\mathbb{R}^{1,1}$ as sketched in figure 4.15. This figure also illustrates the notation used in the following.

Lemma 4.5.6. *Let $u_0<u_1$, $v_1\in\mathbb{R}$, and $t_0\in\mathbb{R}$ be given. Set*

$$\Omega:=\Omega^{t_0,v_1}_{u_0,u_1}:=\{p\in\mathbb{R}^{1,1}\mid t(p)\geq t_0\,,v(p)\leq v_1\,,u_0\leq u(p)\leq u_1\}\,, \tag{4.5.8}$$

and suppose that $f,g\in C^\infty(\Omega)$ satisfy the Dirac null system (4.5.2) *for some $V\in C^\infty(\Omega)$ with $\partial_t V=0$. Then for any $n\geq 0$, we have*

$$
\begin{aligned}
&\int_{u_0}^{u_1}|\partial_t^n f(v_1,u)|^2\,\mathrm{d}u+\int_{v_{-1}}^{v_1}|\partial_t^n g(v,u_1)|^2\,\mathrm{d}v \\
&=\int_{r_1}^{r_2}\big\{|(\partial_t^n f|_{t=t_0})(r)|^2+|(\partial_t^n g|_{t=t_0})(r)|^2\big\}\,\mathrm{d}r+\int_{v_0}^{v_1}|\partial_t^n g(v,u_0)|^2\,\mathrm{d}v\,,
\end{aligned}
\tag{4.5.9}
$$

where $r_1=t_0-u_1$, $r_2=t_0-u_0$, $v_0=t_0+r_2$, and $v_{-1}=t_0+r_1$.

PROOF. As already said, (4.5.9) follows similar as Proposition 4.5.1 by integrating the divergence-free current vector field over Ω, and converting this integral into boundary terms using Gauss' theorem. □

Combining this with Proposition 4.5.5 and making once more the assumption that $g(\cdot, u_0) \equiv 0$, we obtain the following estimate between "Goursat data" and Cauchy data. The notation is as before, and as illustrated in figure 4.15.

Proposition 4.5.7. *For every $n \in \mathbb{N}$ there exists a constant $C_n > 0$ such that the following holds: Let $\Omega^{t_0,v_1}_{u_0,u_1}$ be as in (4.5.8), and let $f, g \in C^\infty(\Omega^{t_0,v_1}_{u_0,u_1})$ be a solution of the Dirac null system (4.5.2) for some $V \in C^\infty(\Omega^{t_0,v_1}_{u_0,u_1})$ with $\partial_t V = 0$. Assume moreover that $g(\cdot, u_0) \equiv 0$. Then it holds that*

$$\begin{aligned} \|f(v,\cdot)\|_{H^n_u(u_0,u_1)} \leq C_n \cdot \big(1 + \|V(v,\cdot)\|^{2n}_{C^{n-1}_u(u_0,u_1)}\big) \cdot \big(1 + |u_1 - u_0|\big) \\ \times \Big\{ \|\partial_t^n (f,g)|_{t=t_0}\|_{L^2_r(r_1,r_2)} + \ldots + \|(f,g)|_{t=t_0}\|_{L^2_r(r_1,r_2)} \Big\} \end{aligned} \tag{4.5.10}$$

for all $v \in [v_0, v_1]$, where $r_1 = t_0 - u_1$ and $r_2 = t_0 - u_0$. Here we use the short-hand notation $\|(f,g)|_{t=t_0}\|^2_{L^2_r(r_1,r_2)} = \|f|_{t=0}\|^2_{L^2(r_1,r_2)} + \|g|_{t=0}\|^2_{L^2(r_1,r_2)}$.

PROOF. First one estimates from v to v_0 using (4.5.7) and the vanishing assumption $g(\cdot, u_0) \equiv 0$. Afterwards one estimates from v_0 to t_0 using (4.5.9). □

4.5.3. General estimates of some special integral operators. Our next goal is to derive suitable error estimates for the perturbation terms. To this end, recall that our aim is to study solutions of the integro-differential equation

$$(\partial_v f)(v,u) = -\frac{m^2}{4} \int_{u_0}^u f(v,x)\,\mathrm{d}x + \int_{u_0}^u K_B(v,u,x) f(v,x)\,\mathrm{d}x\,,$$

where the "perturbation kernel" K_B is explicitly given by (4.3.9). Notice that the perturbation part, i.e. the second term, can be rewritten as

$$(B_v \phi)(u) = \int_{u_0}^u K_B(v,u,x)\phi(x)\,\mathrm{d}x = \int_{u_0}^{u_1} \underbrace{K_B(v,u,x)\theta(u-x)}_{=:k_B(v,u,x)} \phi(x)\,\mathrm{d}x\,,$$

where $\theta = 1\!\!1_{[0,\infty)}$ is again the Heaviside function. This is an integral operator with a kernel $k_B(v,u,x)$ of a rather special form. In this section we gather some general properties of such integral operators which will be useful for us.

More precisely, we will study integral operators B, acting on functions defined on a fixed interval $(u_0, u_1)_u$, which have kernels $k_B(u,x)$ of the type

$$k_B(u,x) := K_B(u,x)\theta(u-x) \tag{4.5.11}$$

for some smooth function $K_B \in C^\infty((0,u_1)^2)$. The corresponding integral operator B is then given by

$$(B\phi)(u) := \int_{u_0}^{u_1} k_B(u,x)\phi(x)\,\mathrm{d}x = \int_{u_0}^u K_B(u,x)\phi(x)\,\mathrm{d}x\,. \tag{4.5.12}$$

The right-hand side is certainly well-defined if $\phi \in C^\infty_c(u_0,u_1)$ has compact support. If the kernel k_B enjoys further regularity properties, one can extend the action of B to other spaces by continuity. This and the related operator norm estimates will be studied in the following.

First we establish a sort of "Leibniz rule" for derivatives of $B\phi$ for $\phi \in C_c^\infty(u_0, u_1)$. To this end, the following notation will be convenient: For a function $K : X \times X \to \mathbb{C}$, we denote its restriction to the diagonal $\Delta \subset X \times X$ by

$$K^\Delta(x) := k(x,x)\,. \tag{4.5.13}$$

Be aware in the following that Δ does of course *not* commute with taking derivatives.

Lemma 4.5.8. *Let $K_B \in C^\infty((u_0, u_1)^2)$. Then for any $\phi \in C_c^\infty(u_0, u_1)$ we have $B\phi \in C^\infty(u_0, u_1)$, where $B\phi$ is defined by* (4.5.12)*. Moreover, for any $k \in \mathbb{N}$ we have*

$$\begin{aligned}\partial_u^{k+1}(B\phi)(u) &= \sum_{n=0}^{k}\sum_{j=0}^{n}\binom{n}{j}\left[\partial_u^{n-j}\left((\partial_u^{k-n}K_B)^\Delta\right)(u)\right]\cdot\left[(\partial_u^j\phi)(u)\right] \\ &\quad + \int_{u_0}^{u}(\partial_u^{k+1}K_B)(u,x)\phi(x)\,\mathrm{d}x\,.\end{aligned} \tag{4.5.14}$$

Proof. Firstly, $B\phi$ is smooth since in the right-hand side of (4.5.12) both the integrand and the upper boundary of the integral depend smoothly on u.

To show (4.5.14), we proceed by induction on k. For $k=0$, we have

$$\begin{aligned}\partial_u(B\phi)(u) &= \partial_u\int_{u_0}^{u}K_B(u,x)f(v,x)\,\mathrm{d}x \\ &= K_B(u,u)f(v,u) + \int_{u_0}^{u}(\partial_u K_B)(u,x)f(v,x)\,\mathrm{d}x\,,\end{aligned}$$

which agrees with (4.5.14) in this case.

Now suppose that the assertion is true for some $k \in \mathbb{N}$. Then we have

$$\begin{aligned}\partial_u^{k+1}(B\phi) &= \partial_u\sum_{n=0}^{k-1}\sum_{j=0}^{n}\binom{n}{j}\left[\partial_j^{n-j}(\partial_u^{k-1-n}K_B)^\Delta\right]\cdot\left[(\partial_u^j\phi)(u)\right] \\ &\quad + \partial_u\int_{u_0}^{u}(\partial_u^k K_B)(u,x)\phi(x)\,\mathrm{d}x \\ &= \partial_u(\partial_u^k B\phi) \\ &= \sum_{n=0}^{k-1}\sum_{j=0}^{n}\binom{n}{j}\left[\partial_j^{n-j+1}(\partial_u^{k-1-n}K_B)^\Delta\right]\cdot\left[(\partial_u^j\phi)(u)\right] \\ &\quad + \sum_{n=0}^{k-1}\sum_{j=0}^{n}\binom{n}{j}\left[\partial_j^{n-j}(\partial_u^{k-1-n}K_B)^\Delta\right]\cdot\left[(\partial_u^{j+1}\phi)(u)\right] \\ &\quad + (\partial_u^k K_B)^\Delta\phi + \int_{u_0}^{u}(\partial_u^{k+1}K_B)(u,x)\phi(x)\,\mathrm{d}x \\ &\quad \text{[Set } n = m-1 \text{ in first and second line, and } \ell = j-1 \text{ in second line]} \\ &= \sum_{m=1}^{k}\sum_{j=0}^{m-1}\binom{m-1}{j}\left[\partial_j^{m-j}(\partial_u^{k-m}K_B)^\Delta\right]\cdot\left[(\partial_u^j\phi)(u)\right] \\ &\quad + \sum_{m=1}^{k}\sum_{\ell=1}^{m}\binom{m-1}{\ell-1}\left[\partial_j^{m-\ell}(\partial_u^{k-m}K_B)^\Delta\right]\cdot\left[(\partial_u^\ell\phi)\right]\end{aligned}$$

$$+ (\partial_u^k K_B)^\Delta \phi + \int_{u_0}^u (\partial_u^{k+1} K_B)(u,x)\phi(x)\,\mathrm{d}x$$

[Replace ℓ by j again in second line and regroup the sums]

$$= \sum_{m=1}^k \underbrace{\binom{m-1}{0}}_{=\binom{m}{0}} \left[\partial_u^m \left((\partial_u^{k-m} K_B)^\Delta\right)\right][\phi]$$

$$+ \sum_{m=1}^k \sum_{j=1}^{m-1} \left[\binom{m-1}{j} + \binom{m-1}{j-1}\right] \left[\partial_u^{m-j}\left((\partial_u^{k-m} K_B)^\Delta\right)\right][\partial_u^j \phi]$$

$$+ \sum_{m=1}^k \underbrace{\binom{m-1}{m-1}}_{=\binom{m}{m}} \left[(\partial_u^{k-m} K_B)^\Delta\right][\partial_u^m \phi]$$

$$+ \underbrace{\sum_{j=0}^0 \binom{0}{0} \left[\partial_u^{0-j}\left((\partial_u^k K_B)^\Delta\right)\right][\partial_u^j \phi]}_{=(\partial_u^k K_B)^\Delta}$$

$$+ \int_{u_0}^u (\partial_u^{k+1} K_B)(u,x)\phi(x)\,\mathrm{d}x$$

$$= \sum_{m=0}^k \sum_{j=0}^m \binom{m}{j} \left[\partial_u^{m-j}\left((\partial_u^{k-m} K_B)^\Delta\right)(u)\right] \cdot \left[(\partial_u^j \phi)(u)\right]$$

$$+ \int_{u_0}^u (\partial_u^{k+1} K_B)(u,x)\phi(x)\,\mathrm{d}x\,.$$

This shows that the assertion is also true for $k+1$, and thus completes the induction. □

Now we extend the operator B from $C_c^\infty(u_0,u_1)$ to other spaces.

Lemma 4.5.9. (L^2 to L^2 estimate) *Let $K_B \in C^\infty((u_0,u_1)^2)$, and suppose that K_B is bounded, i.e. $\|K_B\|_{L^\infty((u_0,u_1)^2)} < \infty$. Then the operator B defined by (4.5.12) extends by continuity to a bounded operator on $L^2(u_0,u_1)$ with*

$$\|B\|_{L^2 \to L^2} \leq \|K_B(u,x)\theta(u-x)\|_{L^2((u_0,u_1)^2)} \leq \frac{|u_1-u_0|^2}{2}\,\|K_B\|_{L^\infty((u_0,u_1)^2)}\,. \tag{4.5.15}$$

The first estimate still holds if only $K_B(u,x)\theta(u-x) \in L^2((u_0,u_1)^2)$.

Proof. The first estimate follows by Hölder. The second estimate follows by a crude estimate of $\|K_B(u,x)\theta(u-x)\|_{L^2((u_0,u_1)^2)}$. □

Lemma 4.5.10. (H^k to H^{k+1} estimate) *Let $K_B \in C^\infty((u_0,u_1)^2)$ and suppose that the first $k+1$ derivatives of K_B are bounded, i.e. $\|K_B\|_{C^{k+1}((u_0,u_1)^2)} < \infty$. Then there exists a numerical constant $C_k > 0$ such that for all $\phi \in C_c^\infty(u_0,u_1)$ we have*

$$\|B\phi\|_{H^{k+1}(u_0,u_1)} \leq C_k(1+|u_1-u_0|)\,\|K_B\|_{C^{k+1}((u_0,u_1)^2}\,\|\phi\|_{H^k(u_0,u_1)}\,. \tag{4.5.16}$$

Consequently, B extends to a bounded operator from $H^k(u_0,u_1)$ to $H^{k+1}(u_0,u_1)$.

Proof. For any $j \leq k+1$, the derivative $\partial^j(B\phi)$ can be expressed by formula (4.5.14). We now estimate the L^2-norms of the two parts in (4.5.14) separately. Since $j \leq k+1$, the first part (i.e. the sum) can be estimated using Hölder's inequality as

$$\begin{aligned}\Big\|\sum\sum\cdots\Big\|_{L^2(u_0,u_1)} &\leq C_j \left\|K_B^\Delta\right\|_{H^{j-1}(u_0,u_1)} \|\phi\|_{H^{j-1}(u_0,u_1)} \\ &\leq C_j \sqrt{|u_1-u_0|}\, \|K_B\|_{C^k((u_0,u_1)^2)} \|\phi\|_{H^k(u_0,u_1)} \,.\end{aligned}$$

For the L^2-norm squared of the second part, by Hölder's inequality we have

$$\begin{aligned}\int_{u_0}^{u_1} \left|\int_{u_0}^{u} (\partial_u^j K_B)(u,x)\phi(x)\,\mathrm{d}x\right| \mathrm{d}u &\leq \int_{u_0}^{u_1} \left(\int_{u_0}^{u} |(\partial_u^j K_B)(u,x)|^2\,\mathrm{d}x\right)\left(\int_{u_0}^{u} |\phi(x)|^2\,\mathrm{d}x\right)\mathrm{d}u \\ &\leq \|K_B\|^2_{H^{k+1}((u_0,u_1)^2)} \|\phi\|^2_{L^2(u_0,u_1)} \\ &\leq |u_1-u_0|^2 \|K_B\|^2_{C^{k+1}((u_0,u_1)^2)} \|\phi\|^2_{H^k(u_0,u_1)} \,.\end{aligned}$$

Together these estimates now imply that

$$\begin{aligned}\left\|\partial^j(B\phi)\right\|_{L^2(u_0,u_1)} &\leq C_j\big(\sqrt{|u_1-u_0|}+|u_1-u_0|\big) \|K_B\|_{C^{k+1}((u_0,u_1)^2)} \|\phi\|_{H^k(u_0,u_1)} \\ &\leq C_j(1+|u_1-u_0|) \|K_B\|_{C^{k+1}((u_0,u_1)^2)} \|\phi\|_{H^k(u_0,u_1)} \,.\end{aligned}$$

Since this works for all $j \leq k+1$, the claim about $\|B\phi\|_{H^{k+1}(0,u_1)}$ follows (for a suitable numerical constant $C_k = 0$). ◻

Instead of using L^2-Sobolev spaces, one can also use C^k-spaces.

Lemma 4.5.11. **(C^0 to C^0 estimate)** *Let $K_B \in C^\infty((u_0,u_1)^2)$ and suppose that K_B is bounded, i.e. $\|K_B\|_{L^\infty((u_0,u_1)^2)} < \infty$. Then for all $\phi \in C^\infty_c(u_0,u_1)$ we have*

$$\|B\phi\|_{C^0(u_0,u_1)} \leq |u_1-u_0|\, \|K_B\|_{L^\infty((u_0,u_1)^2)} \|\phi\|_{C^0(u_0,u_1)} \,. \tag{4.5.17}$$

Consequently, B extends to a bounded operator on the space $C^0_b(u_0,u_1)$ of bounded continuous functions.[9]

Proof. This is obvious from the definition of B in (4.5.12). ◻

Lemma 4.5.12. **(C^k to C^{k+1} estimate)** *Let $K_B \in C^\infty((u_0,u_1)^2)$ and suppose that $\|K_B\|_{C^{k+1}((u_0,u_1)^2)} < \infty$. Then there exists a constant $C_k > 0$ such that for all $\phi \in C^\infty_c(u_0,u_1)$ we have*

$$\|B\phi\|_{C^{k+1}(u_0,u_1)} \leq C_k(1+|u_1-u_0|)\, \|K_B\|_{C^{k+1}((u_0,u_1)^2)} \|\phi\|_{C^k(u_0,u_1)} \,. \tag{4.5.18}$$

Consequently, B extends as bounded operator to the space $C^k_b(u_0,u_1)$ of functions with bounded derivatives up to order k.

Proof. This follows immediately from the Leibniz rule (4.5.14). ◻

[9] Actually, it even extends to $L^\infty(u_0,u_1)$ by the same argument.

4.5.4. Decay and boundedness estimates of the remainder terms. In order to apply the general estimates derived in the previous section to our concrete situation, we need to estimate the specific integral kernel K_B^λ in our case. To recall, for fixed $r_0 \in \mathbb{R}$ we are given two smooth functions

$$A : (r_0, \infty)_r \to (0,1) \quad \text{and} \quad R : (r_0, \infty)_r \to (0, \infty)\,.^{10}$$

Next, instead of using the radial coordinate r and the usual time coordinate t of our problem, the whole analysis is performed in the null coordinates

$$v = t + r \quad \text{and} \quad u = t - r\,.$$

The condition $r > r_0$ translates into $v - u > 2r_0$. In the following, we always view A and R as functions of the coordinates v and u, and we simply write $A(v,u)$ and $R(v,u)$ although it would be more accurate to write $A(\frac{v-u}{2})$ and $R(\frac{v-u}{2})$.

In terms of these functions and the additional parameters $m \in \mathbb{R}$ (the mass) and $\lambda \in \sigma(D_N) \subset \mathbb{R}$ (the "angular momentum"), the kernel $K_B = K_B^\lambda$ of our integral operator is the function $K_B^\lambda : (v_0, \infty)_v \times (u_0, u_1)_u \times (u_0, u_1)_x \longrightarrow \mathbb{C}$ given by

$$\begin{aligned}
4K_B^\lambda(v,u,x) =& \mathrm{i}m\lambda \left(\frac{1}{R(v,u)} - \frac{1}{R(v,x)} \right) + m^2 (A(v,x) + A(v,u)) \\
&+ \mathrm{i}m\lambda \left(\frac{A(v,u)}{R(v,x)} - \frac{A(v,x)}{R(v,u)} + \frac{A(v,x)}{R(v,x)} - \frac{A(v,u)}{R(v,u)} \right) \\
&- \frac{\lambda^2}{R(v,x)R(v,u)} - m^2 A(v,x)A(v,u) \\
&+ \mathrm{i}m\lambda \left(\frac{A(v,u)A(v,x)}{R(v,u)} - \frac{A(v,u)A(v,x)}{R(v,x)} \right) \\
&+ \lambda^2 \left(\frac{A(v,x)}{R(v,x)R(v,u)} + \frac{A(v,u)}{R(v,x)R(v,u)} \right) \\
&- \lambda^2 \frac{A(v,u)A(v,x)}{R(v,u)R(v,x)}
\end{aligned} \tag{4.5.19}$$

Let us remark that for given $u_0 < u_1$, the number v_0 has to be chosen in such a way that $v - u > 2r_0$ for all $v \in (v_0, \infty)$ and $u \in (u_0, u_1)$. The particular choice is not so important here since we will only be interested in large values of v.

In order to apply the various estimates derived in the previous section, we need bounds of derivatives of K_B^λ. Such bounds can be derived assuming C^k-bounds of A and R^{-1} of the type assumed in our asymptotic flatness conditions in Def. 3.2.1. Depending on whether we assume boundedness or decay of A and R^{-1}, we obtain different estimates of K_B^λ. Concerning constants in the estimates, it should be pointed out that it will be important later on how the estimate depends on angular momentum λ.

Lemma 4.5.13. (Boundedness estimate of the perturbation kernel) *Let $r_1 > r_0$, and assume that A and R^{-1} satisfy the C^k-boundedness assumption* (3.2.1), *i.e. assume that there exists $C_k^g > 0$ such that*

$$\|A\|_{C^k(r_1,\infty)}\,, \|R^{-1}\|_{C^k(r_1,\infty)} \leq C_k^g\,. \tag{4.5.20}$$

[10]These functions appeared at the very beginning as coefficients of the Lorentzian metrics of our spacetime (cf. (3.1.1)), and they reappear in the Dirac equation in some way.

Then for any multiindex $\beta \in \mathbb{N}^3$ with $|\beta| \leq k$, and all $v, u, x \in \mathbb{R}$ with $\frac{v-u}{2}, \frac{v-x}{2} > r_1$, the perturbation kernel (4.5.19) *satisfies*

$$|\partial^\beta K_B^\lambda(v,u,x)| \leq C_k(1+C_k^g)^4(1+\lambda^2)\,, \tag{4.5.21}$$

where $C_k > 0$ is a purely combinatorical constant depending on the structure of K_B^λ.

Proof. It follows from formula (4.5.19) for the perturbation kernel K_B^λ that $\partial^\beta K_B^\lambda$ is a sum of terms with each term being a product of derivatives of the four functions $A(\frac{v-u}{2}), A(\frac{v-x}{2})$ and $R(\frac{v-u}{2})^{-1}, R(\frac{v-x}{2})^{-1}$ of order at most $|\beta| \leq k$ (in either of the variables v, u, x). Moreover, each term in the sum is a product of at least one and at most four of such factors. Each factor can be estimated by (4.5.20), and noting that λ appears at most quadratically in K_B^λ one obtains (4.5.21) with C_k depending on the total number of terms and on m. $\square$

Lemma 4.5.14. (Decay estimate of the perturbation kernel) *Let $r_1 > r_0$, and assume that A and R^{-1} satisfy the C^k-asymptotic flatness assumption* (3.2.2), *i.e. assume there exist $\alpha > 0$ and $C_k^{g,dec} > 0$ such that*

$$|\partial_r^j A(r)|, \left|\partial_r^j R(r)^{-1}\right| \leq \frac{C_k^{g,dec}}{(1+r-r_1)^\alpha} \qquad \forall r > r_1\,, j \leq k\,.^{11} \tag{4.5.22}$$

Then for any multiindex $\beta \in \mathbb{N}^3$ with $|\beta| \leq k$, and any $u_0 < u_1$, the perturbation kernel (4.5.19) *satisfies*

$$|\partial^\beta K_B^\lambda(v,u,x)| \leq \frac{C_k(1+C_k^{g,dec})^4(1+\lambda^2)}{(1+v)^\alpha} \tag{4.5.23}$$

for all $v \in \mathbb{R}$ and $u, x \in [u_0, u_1]$ with $\frac{v-u}{2}, \frac{v-x}{2} > r_1$, where $C_k > 0$ depends on the structure of K_B^λ, and on r_1, u_0, u_1.

Proof. As in the previous proof, it follows from formula (4.5.19) for the perturbation kernel K_B^λ that $\partial^\beta K_B^\lambda$ is a sum of terms with each term being a product of derivatives of the four functions $A(\frac{v-u}{2}), A(\frac{v-x}{2})$ and $R(\frac{v-u}{2})^{-1}, R(\frac{v-x}{2})^{-1}$ of order at most $|\beta| \leq k$ (in either of the variables v, u, x). Moreover, each term in the sum is a product of at least one and at most four of such factors.

Estimating each factor by (4.5.22), it follows that any product of at most four of such factors can be estimated by

$$\left(\frac{C_k^{g,\mathrm{dec}}}{(1+\frac{v-u}{2}-r_1)^\alpha}\right)^{k_1}\left(\frac{C_k^{g,\mathrm{dec}}}{(1+\frac{v-x}{2}-r_1)^\alpha}\right)^{k_2} \leq \frac{(1+C_k^{g,\mathrm{dec}})^4}{(1+\frac{v-u}{2}-r_1)^{k_1\alpha}(1+\frac{v-x}{2}-r_1)^{k_2\alpha}} \tag{$*$}$$

for some $k_1, k_2 \in \mathbb{N}$ with $1 \leq k_1 + k_2 \leq 4$. To finish the proof, we write

$$\frac{1}{1+\frac{v-u}{2}-r_1} = \frac{1+v}{1+\frac{v-u}{2}-r_1}\cdot\frac{1}{1+v}\,,$$

and note that the first factor on the right-hand side is a bounded function of $v \in \mathbb{R}$ and $u \in [u_0, u_1]$ in the allowed range of these variables, i.e. $\frac{v-u}{2} > r_1$. Therefore it can be absorbed in the constant $C_k > 0$. The same holds with x in place of u. Finally, concerning the λ-dependence, one uses again λ appears at most quadratically in K_B^λ. $\square$

[11] Since r_1 may be negative, the subtraction of r_1 in the deniominator makes sure that the demoniator is always larger than one, hence positive and nonzero in particular.

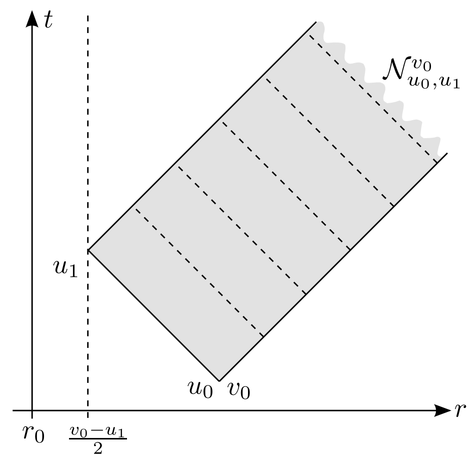

FIGURE 4.16. An outgoing null strip $\mathcal{N}^{v_0}_{u_0,u_1} \subset \mathbb{R}^{1,1}$.

Since it is similar in spirit to the previous two estimates, let us include at this point an estimate of the potential V_λ in the Dirac null system which we will need later.

Lemma 4.5.15. (Boundedness estimate of the Dirac null system potential) *Let $r_1 > r_0$, and assume that A and R^{-1} satisfy the C^k-boundedness assumption* (3.2.1)*, i.e. assume that there exists $C^g_k > 0$ such that* (4.5.20) *holds. Then the function V_λ, defined by* (4.2.9)*, satisfies*

$$|\partial_r^j V_\lambda(r)| \leq C_k(1 + C^g_k)^2(1 + |\lambda|) \qquad \forall r > r_1 \, , j \leq k \, , \tag{4.5.24}$$

where $C_k > 0$ is a combinatorical constant.

PROOF. One reasons similarly as in the proof of Lemma 4.5.13, only this time considering the structure of V_λ as given by (4.2.9) instead of that of K^λ_B. $\square$

4.6. Decay in Null Directions of Individual Angular Momentum Modes

Now we put together the previously obtained estimates to obtain a decay result in null directions for f and g. We begin by explaining the general method which will be used to take care of the perturbations.

4.6.1. The Lippmann-Schwinger equation in an outgoing null strip. The starting point is the Dirac null system (4.2.10). We are going to study the behaviour of solutions f, g of this system as one moves out to infinity in an *outgoing null strip* in the t, r-plane as sketched in figure 4.16. More precisely, if $v = t + r$ and $u = t - r$ again denote the null coordinates, then for a fixed interval $(u_0, u_1) \subset \mathbb{R}$ and some fixed $v_0 \in \mathbb{R}$ with $v_0 - u_1 > 2r_0$, we study the asymptotic behaviour of f and g in the *outgoing null strip*

$$\mathcal{N}^{v_0}_{u_0,u_1} := \{p \in \mathbb{R}^{1,1} \mid v_0 \leq v(p) \leq \infty \, , \, u_0 < u(p) < u_1\} \subset \mathbb{R}^{1,1} \, , \tag{4.6.1}$$

which is sketched in figure 4.16, as v tends to infinity.

To be clear, the solutions f and g will always be assumed to be defined on the *whole* spacetime $\mathbb{R}_t \times (r_0, \infty)_r$ so that they are perfectly regular at the boundaries of the null strip. Moreover, we will assume that f and g satisfy the support condition (4.3.3). Therefore, by choosing u_0 sufficiently small (possibly very negative) and v_0 sufficiently large, we can achieve that $f(v, u) = g(v, u) = 0$ for all $v \in (v_0, \infty)$ and all $u \leq u_0'$ with some suitably fixed $u_0' \in (u_0, u_1)$, see figure 4.12. As we have seen in Section 4.3.2, the function f then satisfies the integro-differential equation

$$\partial_v f(v,u) = -V(v,u) \int_{u_0}^{u} \overline{V(v,x)} f(v,x)\, \mathrm{d}x \qquad \forall (v,u) \in \mathcal{N}^{v_0}_{u_0,u_1}\,,$$

where V is given by (4.2.9). In Section 4.3.2 we expanded the right-hand side and rewrote this integro-differential equation for f as

$$(\partial_v f)(v,u) = -\frac{m^2}{4} \int_{u_0}^{u} f(v,x)\, \mathrm{d}x + \int_{u_0}^{u} K_B^{\lambda}(v,u,x) f(v,x)\, \mathrm{d}x\,, \tag{$*$}$$

where the "perturbation kernel" K_B^λ is explicitly given by (4.3.9).

In order to explain how to proceed from here, we write $f(v,u) = f_v(u)$. Then for fixed v one can view f_v as a function in the space $\underline{C}^\infty(u_0,u_1)$, defined by (4.4.36), which additionally depends smoothly on the parameter $v \in (v_0, \infty)$. Using this notation, $(*)$ has the structure

$$\partial_v f_v = A f_v + B_v f_v\,, \tag{4.6.2}$$

where A and B_v are linear operators on $\underline{C}^\infty(u_0,u_1)$, acting by

$$(Af_v)(u) := -\frac{m^2}{4} \int_{u_0}^{u} f_v(x)\, \mathrm{d}x \quad \text{and} \quad (B_v f_v)(u) := \int_{u_0}^{u} K_B^\lambda(v,u,x) f_v(x)\, \mathrm{d}x\,. \tag{4.6.3}$$

As we have seen in Section 4.4, for any given v_0 and $f_0 \in \underline{C}^\infty(u_0,u_1)$ the free equation $\partial_v f_v = A f_v$ has a unique solution $f_v \in \underline{C}^\infty(u_0,u_1)$, defined for all $v \in (v_0,\infty)$, with initial conditions $f_{v_0} = f_0$. Since the equation is linear, f_v depends linearly on f_0, so that

$$f_v = U_A^{v_0,v} f_0 \tag{4.6.4}$$

for some linear operator $U_A^{v_0,v}$ acting on $\underline{C}^\infty(u_0,u_1)$. It also follows from uniqueness of solutions that this family of operators satisfies the usual semi-group properties $U_A^{v_0,v} U_A^{v,w} = U_A^{v_0,w}$ and $U_A^{v_0,v_0} = \mathsf{id}$. Moreover, since the equation is autonomous, i.e. does not explicitly depend on v, the operator $U_A^{v_0,v}$ only depends on the difference $v - v_0$, so we may write $U_A^{v_0,v} = U_A(v-v_0)$.

Returning to the full equation (4.6.2), we now show that solutions of (4.6.2) satisfy a convenient identity, known as *Lippmann-Schwinger equation*, or *Duhamel's formula.*

Lemma 4.6.1. *Suppose that* $f \in C^\infty((v_0,v_1)_v \times (u_0,u_1)_u)$ *satisfies* $f_v = f(v,\cdot) \in \underline{C}^\infty(u_0,u_1)$ *for all* $v \in (v_0,v_1)$. *Suppose moreover that* f *satisfies the equation* (4.6.2). *Then* f *satisfies the identity*

$$f_v = U_A(v-v_0) f_{v_0} + \int_{v_0}^{v} U_A(v-s) B_s f_s\, \mathrm{d}s\,. \tag{4.6.5}$$

Proof. One easily sees that the right-hand side of (4.6.5) defines a smooth function of v and u, which for fixed v is a smooth function in $\underline{C}^\infty(u_0,u_1)$. Moreover, a straightforward computation shows that it satisfies equation (4.6.2) and coincides with f_{v_0} at $v = v_0$. Consequently, if we can show that there can be at most one such solution, the right-hand side must coincide with f.

To see uniqueness of such solutions, we simply rewrite (4.6.2) back in the form

$$\partial_v f(v,u) = -V(v,u) \int_{u_0}^{u} \overline{V(v,x)} f(v,x)\, \mathrm{d}x,$$

and introduce the function $g(v,u) = \int_{u_0}^{u} \overline{V(v,x)} f(v,x)\, \mathrm{d}x$. Then f, g satisfy the Dirac null system (4.5.2), and $f(\cdot,u_0) = g(\cdot,u_0) \equiv 0$ since $f_v \in \underline{C}^\infty(u_0,u_1)$. A simple application of

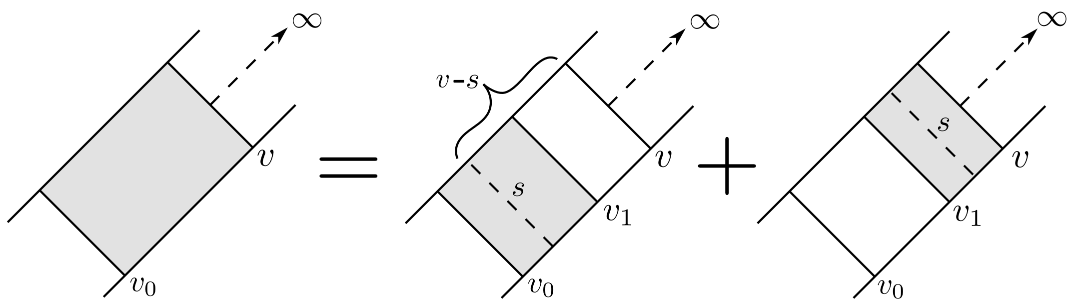

FIGURE 4.17. Illustration of the division of the integral (4.6.6) appearing in the Lippmann-Schwinger equation.

current conservation (4.5.3) for the Dirac null system now yields the desired uniqueness. □

Equation (4.6.5) is known under the name *Lippmann-Schwinger equation* in the physics literature. In the mathematics literature it is also known as *Duhamel's formula*. We will use this simple identity to estimate the solution f_v in the way explained now.

4.6.2. Using the Lippmann-Schwinger equation to derive estimates. The Lippmann-Schwinger equation (4.6.5) can be used to estimate f_v in situations where the second term in equation (4.6.2), i.e. the one given by the operator B_v, may be regarded as a "small perturbation" to A in a suitable sense. This is how we shall use the Lippmann-Schwinger equation, so let us explain how this can be done.

We want to show that f_v decays as v tends to infinite. To this end, one first needs to show that the free propagation $U_A(v - v_0)f_{v_0}$ decays as $v \to \infty$. In our case this was accomplished in Section 4.4.5. Next, one splits the integral term in the Lippmann-Schwinger equation into two parts as

$$\int_{v_0}^{v} U_A(v-s)B_s f_s \,\mathrm{d}s = \int_{v_0}^{v_1} U_A(v-s)B_s f_s \,\mathrm{d}s + \int_{v_1}^{v} U_A(v-s)B_s f_s \,\mathrm{d}s\,, \tag{4.6.6}$$

where $v_1 = v_1(v)$ is some suitably chosen intermediate point between v_0 and v. This splitting is illustrated in figure 4.17. The key idea is now to choose the intermediate point $v_1 = v_1(v)$ in such a way that in the first integral, $v - s$ is always of order v and therefore $U_A(v - s)$ gives decay as v tends to infinity. In the second integral, $v - s$ can become arbitrarily small so that $U_A(v - s)$ cannot be used to get decay. However, in the second integral s is always of order v, so *if* the perturbation B_s decays for large s, then this may cause the second term to decay as well as v tends to infinity. The particular choice of intermediate time for which this works (if it works at all) depends on the specific properties of B_s.

This splitting thus allows to exploit the decay of the free propagation and the perturbation in two separate ways. Of course, this is only a rough sketch of the basic idea. In the following we will carry out the corresponding actual estimates, using what we have already accomplished.

4.6.3. Specification of the precise geometric setup. Before we start putting together estimates, we restate clearly the precise geometric setup we are working in.

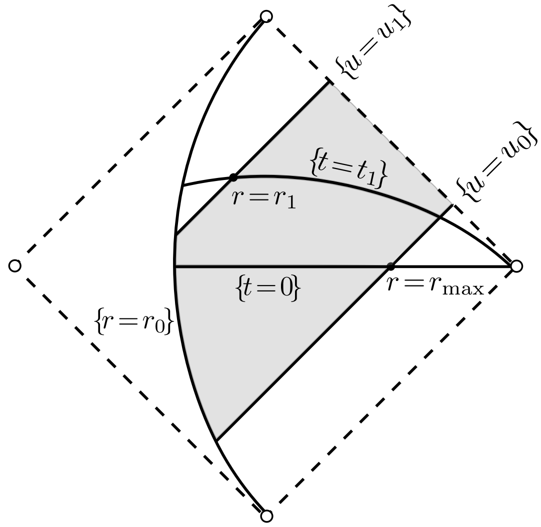

FIGURE 4.18. The precise geometric setup.

Figure 4.18 serves to illustrate the following description in a Penrose diagram. As before, the underlying spacetime is $Q = \mathbb{R}_t \times (r_0, \infty)_r$.

First we fix $r_{\max} \in (r_0, \infty)$, and we consider only solutions whose initial data at time $t = 0$ lies inside $(r_0, r_{\max})$. Next, we fix an outgoing null strip with lower boundary $u_0 = -r_{\max}$ and upper boundary $u_1 > u_0$ arbitrary. Finally, we choose $t_1 > 0$ sufficiently large such that the hypersurface $\{t = t_1\}$ intersects the upper boundary $\{u = u_1\}$ of our chosen null strip. We denote the radial coordinate of the intersection point by r_1.

All our estimates will then be in the part of the null strip which lies in the future of the hypersurface $\{t = t_1\}$. We will derive decay estimates for $v \to \infty$, depending on data at time $t = t_1$. Concerning the coefficients $A(r)$ and $R(r)$, this has the effect that only their restrictions to (r_1, ∞) play a role.

Remark 4.6.2. In the case where the underlying spacetime Q is globally hyperbolic, i.e. $r_0 = -\infty$, we can of course simply choose $t_1 = 0$. We come back to this simpler case at the very end in Section 4.7.3. Also in the non-globally hyperbolic case, one can try to relate initial data at $t = t_1$ to initial data at $t = 0$ by using the computation for the change in time of the Dirac current of Section 3.4.3. We will not do this.

4.6.4. Estimating the integral term. In the following we estimate the integral term in the Lippmann-Schwinger equation (4.6.5). As explained, the idea is to split the integral at a suitable intermediate point between v_0 and v and estimate each part separately, using different ideas. The particular choice of the intermediate point which works for our purposes is as follows. We fix $0 < \gamma < 1$ and set

$$v_\gamma := v - (v - v_0)^\gamma \,. \tag{4.6.7}$$

For $v - v_0 \geq 1$, we indeed have $v_0 < v_\gamma < v$, and since we are anyways interested in the behaviour for large v in the end, we may as well assume that $v - v_0 \geq 1$. From now on v_γ always refers to (4.6.7).

We begin with the first part of the integral, where we can exploit the decay of the free propagation. For this we need to assume only boundedness of the coefficients A and

R^{-1} that appear in the Dirac null system. The geometric quantities $r_{\max}, r_1, u_0, u_1, t_1$ are from now on fixed as explained in the previous section.

Lemma 4.6.3. (Decay from decaying free evolution) *Assume that the A and R^{-1} satisfy the C^k-boundedness condition (3.2.1) for some $k > 1$, i.e. assume there exists $C_k^g > 0$ such that $\|A\|_{C^{k+1}(r_1,\infty)}, \left\|R^{-1}\right\|_{C^{k+1}(r_1,\infty)} < C_k^g$. Let $f, g \in C^\infty(Q)$ be a smooth solution of the Dirac null system (4.2.10) satisfying the support condition (4.3.3), i.e. assume that $\mathrm{supp} f|_{t=0}, \mathrm{supp}\, g|_{t=0} \subset (r_0, r_{\max})$. Then there exists $C_k > 0$ such that*

$$\begin{aligned}\left|\int_{v_0}^{v_\gamma} (U_0(v-s)B_s f_s)(u)\,\mathrm{d}s\right| \leq \frac{C_k(1+C_k^g)^{4k}}{(1+(v-v_0))^{\gamma\frac{2k-3}{4}}} \big(1+|u_1-u_0|\big)^{\frac{k+5}{2}} \big(1+|\lambda|\big)^{2k} \\ \times \sum_{j=0}^{k} \|\partial_t^j (f,g)|_{t=t_1}\|_{L^2(r_1,\infty)}\end{aligned} \tag{4.6.8}$$

for all $v \in (v_0+1, \infty)$.

PROOF. We start by estimating

$$\left|\int_{v_0}^{v_\gamma} (U_0(v-s)(B_s f_s))(u)\,\mathrm{d}s\right| \leq \int_{v_0}^{v_\gamma} \|U_0(v-s)\|_{H^{k+1}\to C^0} \|B_s\|_{H^k\to H^{k+1}} \|f_s\|_{H^k}\,\mathrm{d}s\,. \quad (*)$$

For the free propagator U_0, we use Lemma 4.4.18 which yields the estimate

$$\|U_0(v-s)\|_{H^{k+1}\to C^0} \leq \frac{C_k}{(1+(v-s))^{\frac{k+2}{4}}} (1+|u_1-u_0|)^{\frac{k+1}{2}}\,.$$

For the perturbation B_s we have by Lemma 4.5.10 and Lemma 4.5.13

$$\|B_s\|_{H^k\to H^{k+1}} \leq C_k(1+|u_1-u_0|)\,\|K_{B_s}\|_{C^{k+1}} \leq C_k(1+|u_1-u_0|)(1+C_k^g)^4(1+|\lambda|)^2\,.$$

Here we used the boundedness assumption of the coefficients A and R^{-1}. For f_s we use the a priori bounds from Proposition 4.5.7 together with Lemma 4.5.15 to obtain

$$\begin{aligned}\|f_s\|_{H^k} &\leq C_k\big(1+|u_1-u_0|\big)\big(1+\|V\|_{C^{k-1}}^{2k}\big) \sum_{j=0}^{k} \|\partial_t^j (f,f)|_{t=t_1}\|_{L^2(r_1,\infty)} \\ &\leq C_k\big(1+C_k^g\big)^{4k}\big(1+|u_1-u_0|\big)\big(1+|\lambda|\big)^{2k} \sum_{j=0}^{k} \|\partial_t^j f|_{t=t_1}\|_{L^2(r_1,\infty)}\,.\end{aligned}$$

Putting these three estimates back into $(*)$, we obtain

$$\begin{aligned}\left|\int_{v_0}^{v_\gamma} (U_0(v-s)(B_s f_s))(u)\,\mathrm{d}s\right| &\leq C_k\big(1+C_k^g\big)^{4k}\big(1+|u_1-u_0|\big)^{\frac{k+5}{2}}\big(1+|\lambda|\big)^{2k} \\ &\quad\times \sum_{j=0}^{k} \|\partial_t^j f|_{t=0}\|_{L_r^2} \int_{v_0}^{v_\gamma} \frac{1}{(1+(v-s))^{\frac{k}{2}+\frac{1}{4}}}\,\mathrm{d}s\end{aligned}$$

The remaining integral can be estimated directly as

$$\begin{aligned}\int_{v_0}^{v_\gamma} \frac{1}{(1+(v-s))^{\frac{2k+1}{2}}}\,\mathrm{d}s &= \frac{4}{2k-3}\left[\frac{1}{(1+(v-s))^{\frac{2k-3}{4}}}\right]_{s=v_0}^{s=v_\gamma} \\ &\leq \frac{4}{2k-3}\frac{1}{(1+(v-v_\gamma))^{\frac{2k-3}{4}}}\end{aligned}$$

$$
\begin{aligned}
&= \frac{4}{2k-3} \frac{1}{(1+(v-v_0)^\gamma))^{\frac{2k-3}{4}}} \\
&\leq \frac{4}{2k-3} \frac{1}{(1+(v-v_0))^{\gamma \frac{2k-3}{4}}} \,.
\end{aligned}
$$

Here the last estimate holds since $0 < \gamma < 1$. This concludes the proof. $\square$

Next we treat the second part of the integral, where we want to use the decay of the functions A and R^{-1}, i.e. the asymptotic flatness assumptions on our spacetime.

Lemma 4.6.4. (Decay from decaying potentials) *Assume that A and R^{-1} satisfy the C^1-asymptotic flatness condition* (3.2.2) *of some rate $\alpha > 0$, i.e. assume there exists $C_1^{g,dec} > 0$ such that*

$$
|A(r)|, |A'(r)|, \left|\frac{1}{R(r)}\right|, \left|\left(\frac{1}{R(r)}\right)'\right| \leq \frac{C_k^{g,dec}}{(1+r-r_1)^\alpha} \qquad \forall r > r_1 \,.
$$

Let $f, g \in \mathcal{C}^\infty(Q)$ be a smooth solution of the Dirac null system (4.2.10) *satisfying the support condition* (4.3.3)*, i.e. assume that* $\operatorname{supp} f|_{t=0}, \operatorname{supp} g|_{t=0} \subset (r_0, r_{\max})$*. Fix $0 < \gamma < \min\{1, \alpha\}$ and let v_γ be as in* (4.6.7)*. Then there exist a constant $C(v_0, \gamma, \alpha) > 0$ and some $v_1 = v_1(v_0, \gamma) \geq v_0$ such that we have*

$$
\begin{aligned}
\left|\int_{v_\gamma}^v (U_0(v-s) B_s f_s)(u) \, \mathrm{d}s\right| \leq \;& \frac{C(v_0, \gamma, \alpha)(1 + C_1^{g,dec})^4 (1 + |u_1 - u_0|)^{\frac{3}{2}} (1 + |\lambda|)^2}{(1 + (v - v_0))^{\alpha - \gamma}} \\
& \times \|(f, g)|_{t=t_1}\|_{L^2(r_1, \infty)}
\end{aligned}
\tag{4.6.9}
$$

for all $v \geq v_1$.

Proof. We start with the estimate

$$
\left|\int_{v_\gamma}^v (U_0(v-s) B_s f_s)(u) \, \mathrm{d}s\right| \leq \int_{v_\gamma}^v \|U_0(v-s)\|_{H^1 \to C^0} \|B_s\|_{L^2 \to H^1} \|f_s\|_{L^2} \, \mathrm{d}s \tag{$*$}
$$

Since we plan to extract the desired decay from B_s, the free propagation U_0 and f_s need merely be bounded. For the free propagation this is accomplished using Lemma 4.4.18, which gives the estimate

$$
\|U_0(v-s)\|_{H^1 \to C^0} \leq C_0 (1 + |u_1 - u_0|)^{\frac{1}{2}} \,.
$$

To estimate f_s, we use Proposition 4.5.7 to obtain

$$
\|f_s\|_{L^2_u} \leq \|(f, g)|_{t=t_1}\|_{L^2(r_1, \infty)} \,.
$$

To extract the decay from B_s, we use Lemma 4.5.10 together with Lemma 4.5.14, which yields

$$
\|B_s\|_{L^2 \to H^1} \leq C_0 (1 + |u_1 - u_0|) \|K_{B_s}\|_{C^1_{u,x}((u_0, u_1)^2)} \leq \frac{C_1 (1 + C_g^{1,\mathrm{dec}})^4 (1 + \lambda)^2}{(1 + s)^\alpha} \,.
$$

Here we used the decaying assumptions for A and R^{-1}. Putting everything back into $(*)$, the remaining integral over s can be estimated by

$$
\int_{v_\gamma}^v \frac{1}{(1+s)^\alpha} \, \mathrm{d}s \leq \frac{v - v_\gamma}{(1 + v_\gamma)^\alpha} = \frac{(v - v_0)^\gamma}{(1 + v - (v - v_0)^\gamma)^\alpha} \sim \frac{v^\gamma}{v^\alpha} \qquad (\text{as } v \to \infty) \,.
$$

Here we used in the end that $\gamma < \alpha$ and $\gamma < 1$. Since moreover the left-hand-side is bounded for $v \to v_0$, it follows that there exists a constant $C = C(v_0, \gamma, \alpha)$ such that

$$\int_{v_\gamma}^{v} \frac{1}{(1+s)^\alpha}\,\mathrm{d}s \le \frac{C(v_0,\gamma,\alpha)}{(1+(v-v_0))^{\alpha-\gamma}} \qquad \forall v \ge v_0\,.$$

This concludes the proof. ∎

Compared to (4.6.8), the decay in (4.6.9) is not as good. The key observation, however, is that we can improve the decay of the second integral if we already know that f_v decays in v. This will eventually allow us to bootstrap us up to the same decay as the one of the free propagation. Before we do this bootstrapping, we work out precisely how the decay of the second integral can actually be improved.

Lemma 4.6.5. (Improving the decay coming from the potentials) *Assume that A and R^{-1} satisfy the C^1-asymptotic flatness condition* (3.2.2) *of some rate $\alpha > 0$, i.e. assume there exists $C_1^{g,dec} > 0$ such that*

$$|A(r)|, |A'(r)|, \left|\frac{1}{R(r)}\right|, \left|\left(\frac{1}{R(r)}\right)'\right| \le \frac{C_k^{g,dec}}{(1+r-r_1)^\alpha} \qquad \forall r > r_1\,.$$

Let $f \in C^\infty(Q)$ be any smooth function which satisfies

$$|f(v,u)| \le \frac{C(u_0,u_1,v_0,f|_{t=t_1})}{(1+(v-v_0))^\mu}$$

for some $\mu > 0$, for all $v \ge v_0$, $u \in [u_0,u_1]$, and for some $C(u_0,u_1,v_0,f|_{t=t_1}) > 0$. Fix $0 < \gamma < \min\{1,\alpha\}$ and set $v_\gamma = v-(v-v_0)^\gamma$ for any $v \ge v_0$. Then there exists a constant $C(v_0,\gamma,\alpha) > 0$ and some $v_1 \ge v_0$ such that

$$\left|\int_{v-v^\gamma}^{v} (U_0(v-s)B_s f_s)(u)\,\mathrm{d}s\right| \le \frac{C(v_0,\gamma,\alpha)(1+C_1^{g,dec})^4(1+|u_1-u_0|)^{\frac{5}{2}}(1+|\lambda|)^2}{(1+(v-v_0))^{\alpha-\gamma+\mu}} \times C(u_0,u_1,v_0,f|_{t=t_1}) \tag{4.6.10}$$

for all $v \ge v_1$.

Proof. We start with the estimate

$$\left|\int_{v_\gamma}^{v} (U_0(v-s)B_s f_s)(u)\,\mathrm{d}s\right| \le \int_{v_\gamma}^{v} \|U_0(v-s)\|_{C^1\to C^0}\,\|B_s\|_{C^0\to C^1}\,\|f_s\|_{C^0}\,\mathrm{d}s\,.$$

The free propagator U_0 can be estimated by Lemma 4.4.20 by

$$\|U_0(v-s)\|_{C^1\to C^0} \le C(1+|u_1-u_0|)^{\frac{3}{2}}\,.$$

For the perturbation, we use Lemma 4.5.12 and Lemma 4.5.14 to estimate

$$\|B_s\|_{C^0\to C^1} \le C(1+|u_1-u_0|)\,\|K_{B_s}\|_{C^1} \le \frac{C(1+C_1^4)}{(1+s)^\alpha}\frac{(1+v_0)(1+\lambda^2)}{(v_0-u_1)}\,.$$

Together with the additionally assumed decay of f_s, this yields

$$\left|\int_{v-v^\gamma}^{v} (U_0(v-s)B_s f_s)(u)\,\mathrm{d}s\right| \le C(1+C_1^4)C_0(u_0,u_1,v_0,\lambda,f_0)(1+|u_1-u_0|)^{\frac{5}{2}} \times \frac{(1+v_0)(1+\lambda^2)}{(v_0-u_1)}\int_{v_\gamma}^{v}\frac{1}{(1+s)^{\alpha+\mu}}\,\mathrm{d}s\,.$$

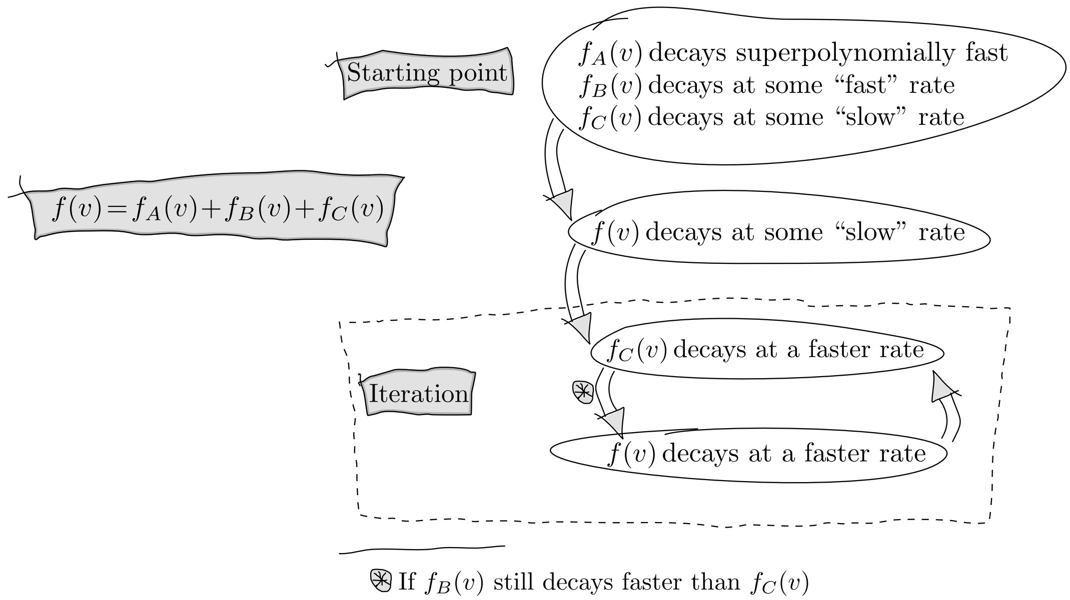

FIGURE 4.19. Sketch of the iterative scheme to prove decay of f in v.

The remaining integral can now be estimated strictly analogously as in the proof of Lemma 4.6.4 to give the desired decay in v. ☐

Finally, we come to the bootstrapping argument which will result in a decay result with speed of decay about as fast as for the free propagation. The logic behind the bootstrapping scheme is illustrated in figure 4.19.

Proposition 4.6.6. *Assume that A and R^{-1} satisfy the C^k-boundedness condition* (3.2.1) *for some $k \geq 2$ and the C^1-asymptotic flatness condition* (3.2.2) *at some rate $\alpha > 0$, i.e. assume there exist constants $C_k^g > 0$ and $C_1^{g,dec} > 0$ such that*

$$\|A\|_{C^{k+1}(r_1,\infty)}\,, \left\|R^{-1}\right\|_{C^{k+1}(r_1,\infty)} < C_k^g\,,$$

and

$$|A(r)|, |A'(r)|, \left|R(r)^{-1}\right|, \left|(R(r)^{-1})'\right| \leq \frac{C_1^{g,dec}}{(1+r-r_1)^\alpha} \qquad \forall r > r_1\,.$$

Fix $0 < \gamma < \min\{1, \frac{\alpha}{4}\}$. Then any solution $f, g \in C^\infty(Q)$ of the Dirac null system (4.2.10) *which satisfies the support condition* (4.3.3) *satisfies*

$$|f(v,u)| \leq \frac{C_k}{(1+(v-v_0))^{\gamma\frac{2k-3}{4}}} \sum_{j=0}^{k} \|\partial_t^j (f,g)|_{t=t_1}\|_{L^2(r_1,\infty)} \tag{4.6.11}$$

for all $v \geq v_1$ and all $u \in [u_0, u_1]$. The constant $C_k > 0$ is explicitly given by

$$C_k = C_{k,0}(1+C_k^g)^{4k}(1+C_1^{g,dec})^{\frac{2k-3}{3}}(1+|u_1-u_0|)^{\frac{k+5}{2}}(1+|\lambda|)^{2k}\,, \tag{4.6.12}$$

where $C_{k,0} = C(k, v_1, \alpha, \gamma) > 0$ is independent of $f, g, u_0, u_1, \lambda, C_k^g, C_1^{g,dec}$.

PROOF. Before going into the details, let us take a look at the general structure of the argument, which is also illustrated in figure 4.19. By the Lippmann-Schwinger equation we can write

$$f(v) = \underbrace{U_0(v-v_0)f_{v_0}}_{=:f_A(v)} + \underbrace{\int_{v_0}^{v_\gamma} U_0(v-s)B_s f_s \,\mathrm{d}s}_{=:f_B(v)} + \underbrace{\int_{v_\gamma}^{v} U_0(v-s)B_s f_s \,\mathrm{d}s}_{=:f_C(v)} .$$

Here we suppressed the explicit reference to the u-dependence from the notation. In the whole proof, u is restricted to the interval $[u_0, u_1]$. By what we have shown so far, we know the following:

(A) $f_A(v)$ decays superpolynomially fast by Lemma 4.4.18. Concretely, for any $m \in \mathbb{N}$, control of $\|f_{v_0}\|_{H^m(u_0,u_1)}$ gives decay in v at power $-\frac{2m-1}{4}$.

(B) For A and R^{-1} bounded as assumed, by Lemma 4.6.3 L^2-control of sufficiently high time derivatives of f, g at $t = t_1$ gives decay of $f_B(v)$ in v at power $-\gamma\frac{2k-3}{4}$.

(C) A priori, if A and $R-1$ decay as assumed, then by Lemma 4.6.4, $f_C(v)$ only decays in v at power $-(\alpha-\gamma)$, at the cost of an L^2-norm of f, g at $t = t_1$. However, together with (A) and (B) this shows that $f(v)$ itself decays in v at power $-(\alpha-\gamma)$ (given that k is sufficiently large). Therefore, by Lemma 4.6.5 it follows that $f_C(v)$ actually decays like $-2(\alpha-\gamma)$.

Now we can repeat the procedure described in (C) until we obtain that $f_C(v)$ decays faster than $f_B(v)$, i.e. at a power higher than $-\gamma\frac{2k-3}{4}$. Then we have to stop since the decay of the term f_B cannot be improved.

In the following, we carry out the details of this scheme. The most effort is to work out the behaviour of the constants in the estimates under this iterative procedure. Recall here that we need to keep track of the angular momentum separation constant λ. We split the analysis into several steps.

Step 1: Getting the decay started

We start by spelling out explicitly how the three terms f_A, f_B, f_C decay. Since in the beginning it is f_C which decays the slowest, and hence dictates the speed of decay, we start with this term. By Lemma 4.6.4 and the assumed decay of A and R^{-1}, we have

$$|f_C(v)| \leq \frac{C_C(v_0,\gamma,\alpha)(1+C_1^{g,\mathrm{dec}})^4(1+|u_1-u_0|)^{\frac{3}{2}}(1+|\lambda|)^2}{(1+(v-v_0))^{\alpha-\gamma}} \|(f,g)|_{t=t_1}\|_{L^2(r_1,\infty)} \quad \text{(C1)}$$

for all $v \geq v_1$. Here we put an index C to the constant as a reminder that it comes from the estimate for f_C.

Coming to f_A and f_B, keep in mind that for f_A we actually have a decay estimate at every polynomial rate of decay, and that for f_B we also have estimates at various polynomial rates, up to some maximal rate (dictated by the assumed estimates on A and R^{-1}). However, and this is an important point, in the estimates with higher decay rates the constants of the estimates are larger (such as the dependence on λ). Therefore in order to keep these constants small, at every step of the proof it is favourable to use decay estimates for f_A and f_B which only just about match the rate of decay of f_C that we have at the respective stage. To this end, it is useful to keep in mind that in terms of integers $k^A, k^B > 0$, the rate of decay for f_A is counted as $-\frac{2k^A-1}{4}$ (cf. Lemma4.4.18), and the rate of f_B is counted as $-\gamma\frac{2k^B-3}{4}$ (cf. Lemma 4.6.3). So the aim is always to choose k^A and k^B as small as possible.

Since at the present stage the rate of decay of f_C is $-(\alpha-\gamma)$, we set

$$k_1^A := \left\lceil 2(\alpha-\gamma)+\frac{1}{2}\right\rceil \quad \text{and} \quad k_1^B := \left\lceil 2\frac{\alpha-\gamma}{\gamma}+\frac{3}{2}\right\rceil . \tag{$k_1^{A/B}$}$$

where for $x\in\mathbb{R}^+$ we denote by $\lceil x\rceil\in\mathbb{N}$ the next natural number above or equal to x. Note that $k_1^B\geq 2$ and $k_1^B>k_1^A$ since $0<\gamma<\min\{1,\frac{\alpha}{4}\}$. In terms of these integers, we then have the following decay of f_A and f_B: For f_A, combining Lemma 4.4.18, Proposition 4.5.5, and Proposition 4.5.7, one gets

$$|f_A(v)| \leq \frac{C_{k_1^A}(1+C^g_{k_1^A})^{4k_1^A}(1+|u_1-u_0|)^{\frac{k_1^A+2}{2}}(1+|\lambda|)^{2k_1^A}}{(1+v)^{\frac{2k_1^A-1}{4}}}\, \|(f,g)|_{t=t_1}\|_{L^2(r_1,\infty)} \, . \tag{A1}$$

For f_B it follows directly from Lemma 4.6.3 that

$$\begin{aligned} |f_B(v)| \leq{}& \frac{C_{k_1^B}(1+C^g_{k_1^B})^{4k_1^B}(1+|u_1-u_0|)^{\frac{k_1^B+5}{2}}(1+|\lambda|)^{2k_1^B}}{(1+v_0)^{\gamma\frac{2k_1^B-3}{4}}} \\ &\times \sum_{j=0}^{k_1^B}\|\partial_t^j(f,g)|_{t=t_1}\|_{L^2(r_1,\infty)} \, . \end{aligned} \tag{B1}$$

Combining the estimates (A1), (B1), (C1) and keeping in mind the choice ($k_1^{A/B}$) of k_1^A and k_1^B, it follows that[12]

$$\begin{aligned} |f(v)| &\leq |f_A(v)|+|f_B(v)|+|f_C(v)| \\ &\leq \frac{C_1(1+C^g_{k_1^B})^{4k_1^B}(1+C_1^{g,\mathrm{dec}})^4(1+|u_1-u_0|)^{\frac{k_1^B+5}{2}}(1+|\lambda|)^{2k_1^B}}{(1+(v-v_0))^{\alpha-\gamma}} \\ &\quad\times\sum_{j=0}^{k_1^B}\|\partial_t^j(f,g)|_{t=t_1}\|_{L^2(r_1,\infty)} \, . \end{aligned} \tag{E1}$$

Here $C_1=\max\{C_C,C_{k_1^A},C_{k_1^B}\}$. Note that all other dependencies in the estimate, for instance on $|\lambda|$ are dictated by the term f_B.

STEP 2: IMPROVING THE DECAY

Starting from the decay estimate (E1) for f which we just established, we now use Lemma 4.6.5 to get a better decay estimate for the term f_C. Afterwards we combine this improved decay of f_C with stronger decay estimates of f_A and f_B (which follow as before) to obtain an improvement over (E1).

To begin with, it follows from (E1) and Lemma 4.6.5 that

$$\begin{aligned} |f_C(v)| \leq{}& \frac{C_1C(v_0,\gamma,\alpha)(1+C^g_{k_1^B})^{4k_1^B}(1+C_1^{g,\mathrm{dec}})^8(1+|u_1-u_0|)^{\frac{k_1^B+10}{2}}(1+|\lambda|)^{2(k_1^B+1)}}{(1+(v-v_0))^{2(\alpha-\gamma)}} \\ &\times\sum_{j=0}^{k_1^B}\|\partial_t^j(f,g)|_{t=t_1}\|_{L^2(r_1,\infty)} \, . \end{aligned} \tag{C2}$$

[12]The "E" in the tag (E1) stands for "estimate".

Next, we adjust the decay estimates (A1) for f_A and (B1) for f_B to the rate of decay in (C2). To this end, similarly as before we set

$$k_2^A := \left\lceil 4(\alpha-\gamma)+\frac{1}{2}\right\rceil \quad \text{and} \quad k_2^B := \left\lceil 4\frac{\alpha-\gamma}{\gamma}+\frac{3}{2}\right\rceil . \tag{$k_2^{A/B}$}$$

It then follows as before that we have the estimates

$$|f_A(v)| \leq \frac{C_{k_2^A}(1+C^g_{k_2^A})^{4k_2^A}(1+|u_1-u_0|)^{\frac{k_2^A+2}{2}}(1+|\lambda|)^{2k_2^A}}{(1+v)^{\frac{2k_2^A-1}{4}}}\|(f,g)|_{t=t_1}\|_{L^2(r_1,\infty)} \tag{A2}$$

and

$$\begin{aligned}|f_B(v)| \leq \; & \frac{C_{k_2^B}(1+C^g_{k_2^B})^{4k_2^B}(1+|u_1-u_0|)^{\frac{k_2^B+5}{2}}(1+|\lambda|)^{2k_2^B}}{(1+v_0)^{\gamma\frac{2k_2^B-3}{4}}} \\ & \times \sum_{j=0}^{k_2^B}\|\partial_t^j(f,g)|_{t=t_1}\|_{L^2(r_1,\infty)} .\end{aligned} \tag{B2}$$

Next, we again combine the estimates (A2), (B2), and (C2) in order to obtain an improved estimate for f. Here the only thing which needs some attention is which constants in the three estimates (A2), (B2), (C2) dominate which others. To this end, notice that in (A2) and (B2) we have simply replaced k_1^A and k_1^B by k_2^A and k_2^B, whereas in (C2), k_1^B is replaced by k_1^B+5 and k_1^B+1 in two respective places. To determine which of these two changes results in a larger growth, we use Lemma 4.6.7 which is postponed to after this proof (here we need the condition $\gamma \leq \frac{\alpha}{4}$). From this result one basically infers that k_2^B always dominates, and thus we obtain the estimate

$$\begin{aligned}|f(v)| &\leq |f_A(v)|+|f_B(v)|+|f_C(v)| \\ &\leq \frac{C_2(1+C^g_{k_2^B})^{4k_2^B}(1+C_1^{g,\mathrm{dec}})^8(1+|u_1-u_0|)^{\frac{k_2^B+5}{2}}(1+|\lambda|)^{2k_2^B}}{(1+(v-v_0))^{2(\alpha-\gamma)}} \\ &\quad \times \sum_{j=0}^{k_2^B}\|\partial_t^j(f,g)|_{t=t_1}\|_{L^2(r_1,\infty)} .\end{aligned} \tag{E2}$$

Step 3: How often can we repeat this procedure?

Restarting the previous step from the estimate (E2) instead of (E1), we can further improve the decay of f. As noted in the general outline at the beginning of the proof, we can keep on repeating Step 2 until we reach the point at which f_C decays faster than f_B. To determine when this happens, notice that from the arguments in Step 2 it is clear that after repeating Step 2 a total number of m times, we end up with the estimate

$$\begin{aligned}|f(v)| &\leq |f_A(v)|+|f_B(v)|+|f_C(v)| \\ &\leq \frac{C_m(1+C^g_{k_m^B})^{4k_m^B}(1+C_1^{g,\mathrm{dec}})^{4m}(1+|u_1-u_0|)^{\frac{k_m^B+5}{2}}(1+|\lambda|)^{2k_m^B}}{(1+(v-v_0))^{m(\alpha-\gamma)}} \\ &\quad \times \sum_{j=0}^{k_2^m}\|\partial_t^j(f,g)|_{t=t_1}\|_{L^2(r_1,\infty)} .\end{aligned} \tag{Em}$$

Therefore, to find out at which point our scheme can no longer be applied, we have to find the maximal $m \in \mathbb{N}$ such that

$$m(\alpha-\gamma) \leq \gamma \frac{2k-3}{4}\,,$$

where $k \in \mathbb{N}$ given in the assumptions of the current proposition. It follows that the maximal such m is given by

$$m_{\max} = \left\lfloor \frac{\gamma}{\alpha-\gamma} \frac{2k-3}{4} \right\rfloor\,,$$

where for $x \in \mathbb{R}^+$ we denote by $\lfloor x \rfloor$ the next smaller or equal natural number to x.

Step 4: Getting all the constants right

It remains to verify the asserted constants in (4.6.15). First, concerning the rate of decay, notice that this will be just the same rate of decay as is allowed by the term f_B, which is

$$\frac{1}{(1+(v-v_0))^{\gamma\frac{2k-3}{4}}}\,.$$

Next, if we want to control the constants, we have to estimate $k^B_{m_{\max}}$. To this end, we first have the trivial estimate

$$m_{\max} = \left\lfloor \frac{\gamma}{\alpha-\gamma} \frac{2k-3}{4} \right\rfloor \leq \frac{\gamma}{\alpha-\gamma} \frac{2k-3}{4}\,,$$

which leads to

$$k^B_{m_{\max}} = \left\lceil 2m_{\max}\frac{\alpha-\gamma}{\gamma} + \frac{3}{2} \right\rceil \leq \left\lceil 2\left(\frac{\gamma}{\alpha-\gamma}\frac{2k-3}{4}\right)\frac{\alpha-\gamma}{\gamma} + \frac{3}{2} \right\rceil = \left\lceil \frac{2k-3}{2} + \frac{3}{2} \right\rceil = k\,.$$

This shows that the final estimate we are going to obtain before we can no longer apply Step 1 is

$$|f(v)| \leq \frac{C(1+C^g_k)^{4k}(1+C^{g,\mathrm{dec}}_1)^{\frac{2k-3}{3}}(1+|u_1-u_0|)^{\frac{k+5}{2}}(1+|\lambda|)^{2k}}{(1+(v-v_0))^{\gamma\frac{2k-3}{4}}} \times \sum_{j=0}^{k} \|\partial_t^j (f,g)|_{t=t_1}\|_{L^2(r_1,\infty)}\,.$$

Here we also used that

$$4m_{\max} \leq \frac{\gamma}{\alpha-\gamma}(2k-3) \leq \frac{2k-3}{3}\,,$$

which follows due to $\gamma < \frac{\alpha}{4}$. This now concludes the proof. □

We still need to catch up on the estimates of k^B_j that we used in the previous proof.

Lemma 4.6.7. *Let $\alpha > 0$ and $0 < \gamma < \min\{1, \frac{\alpha}{4}\}$. For each $j \in \mathbb{N}$, set*

$$k_j := \left\lceil 2j\frac{\alpha-\gamma}{\gamma} + \frac{3}{2} \right\rceil\,.$$

Then it holds that

$$k_{j+1} \geq k_j + 5 \qquad \forall j \in \mathbb{N}\,. \tag{4.6.13}$$

PROOF. Set $\widetilde{k}_j = 2j\frac{\alpha-\gamma}{\gamma} + \frac{3}{2}$. Then we can estimate

$$k_{j+1} - k_j = \lceil \widetilde{k}_{j+1} \rceil - \lceil \widetilde{k}_j \rceil \geq \widetilde{k}_{j+1} - (\widetilde{k}_j + 1) = 2\frac{\alpha-\gamma}{\gamma} - 1\,.$$

Continuing with the right-hand side, we have

$$2\frac{\alpha-\gamma}{\gamma} - 1 \geq 5 \quad\Longleftrightarrow\quad \frac{\alpha-\gamma}{\gamma} \geq 3 \quad\Longleftrightarrow\quad \alpha \geq 4\gamma \quad\Longleftrightarrow\quad \frac{\alpha}{4} \geq \gamma\,.$$

This concludes the proof. □

4.6.5. From decay of f to decay of g. Having obtained decay of $f(v,u)$ as v tends to infinity, one obtains decay of $g(v,u)$ using that (cf. Section 4.3.1)

$$g(v,u) = -\int_{u_0}^{u} \overline{V}(v,x) f(v,x)\,\mathrm{d}x\,. \tag{4.6.14}$$

Proposition 4.6.8. *Assume that A and R^{-1} satisfy the C^k-boundedness condition* (3.2.1) *for some $k \geq 2$ and the C^1-asymptotic flatness condition* (3.2.2) *at some power $\alpha > 0$, i.e. assume there exist constants $C_k^g > 0$ and $C_1^{g,dec} > 0$ such that*

$$\|A\|_{C^{k+1}(r_1,\infty)}\,, \left\|R^{-1}\right\|_{C^{k+1}(r_1,\infty)} < C_k^g\,,$$

and

$$|A(r)|, |A'(r)|, \left|R(r)^{-1}\right|, \left|(R(r)^{-1})'\right| \leq \frac{C_1^{g,dec}}{(1+r-r_1)^{\alpha}} \qquad \forall r > r_1\,.$$

Fix $0 < \gamma < \min\{1, \frac{\alpha}{4}\}$. Let $f, g \in C^\infty(Q)$ be any solution of the Dirac null system (4.2.10) *which satisfies the support condition* (4.3.3)*. Then we have*

$$\boxed{|f(v,u)|, |g(v,u)| \leq \frac{C_k'}{(1+(v-v_0))^{\gamma\frac{2k-3}{4}}} \sum_{j=0}^{k} \|\partial_t^j (f,g)|_{t=t_1}\|_{L^2(r_1,\infty)}} \tag{4.6.15}$$

for all $v \geq v_1$ and all $u \in [u_0, u_1]$. The constant $C_k' > 0$ is explicitly given by

$$C_k' = C_{k,0}'(1+C_k^g)^{4k+2}(1+C_1^{g,dec})^{\frac{2k-3}{3}}(1+|u_1-u_0|)^{\frac{k+7}{2}}(1+|\lambda|)^{2k+1}\,, \tag{4.6.16}$$

where $C_{k,0}' = C(k, v_1, \alpha, \gamma) > 0$ is independent of $f, g, u_0, u_1, \lambda, C_k^g, C_1^{g,dec}$.

PROOF. Using (4.6.14), this is an immediate consequence of the estimate (4.6.15) of f and the estimate (4.5.24) for the potential V. Concerning the constant, starting from the constant C_k in the estimate of f given by (4.6.12), the estimate for V raises the power of the term $(1+C_k^g)$ by two, and the power of the term $(1+|\lambda|)$ by one. The integral raises the power of the term $(1+|u_1-u_0|)$ by one. □

4.6.6. Superpolynomial decay for C^∞-bounded coefficients. In Proposition 4.6.8 we have seen that for coefficients A and R^{-1} which are C^1-asymptotically flat and whose C^{k+1}-norm is bounded, solutions f, g of the Dirac null system (4.2.10) which satisfy the support condition (4.3.3) decay like $v^{-\frac{2k-3}{4}}$ for $v \to \infty$. Consequently, if we assume that *all* derivatives of A and R^{-1} are bounded, such solutions will decay *superpolynomially fast*, i.e. as fast as any inverse power of v. Note that we do not need to assume stronger decay conditions of the coefficients.

In the following statement we prefer to write the decay rate as v^{-k} instead of $v^{-\gamma\frac{2k-3}{4}}$ to point out more clearly that solutions decay as fast as at any inverse power of v. Therefore some of the other constants change slightly in their dependence on k as compared to previous statements.

Theorem 4.6.9. (Superpolynomial Decay in Null Directions) *Suppose that*

i.) A and R^{-1} satisfy the C^k-boundedness condition (3.2.1) *for every $k \in \mathbb{N}$, i.e. assume there exist constants $C_k^g > 0$ such that*

$$\|A\|_{C^{k+1}(r_1,\infty)}, \left\|R^{-1}\right\|_{C^{k+1}(r_1,\infty)} < C_k^g \qquad \forall k \in \mathbb{N},$$

ii.) A and R^{-1} satisfy the C^1-asymptotic flatness condition (3.2.2) *at some power $\alpha >$, i.e. assume there exists $C_1^{g,dec} > 0$ such that*

$$|A(r)|, |A'(r)|, |R(r)^{-1}|, |(R(r)^{-1})'| \leq \frac{C_1^{g,dec}}{(1+r-r_1)^\alpha} \qquad \forall r \geq r_1 .$$

Fix $0 < \gamma < \min\{1, \frac{\alpha}{4}\}$, and let $f, g \in C^\infty(Q)$ be any solution of the Dirac null system (4.2.10) *which satisfies the support condition* (4.3.3)*. Then for every $k \in \mathbb{N}$ we have*

$$\boxed{|f(v,u)|, |g(v,u)| \leq \frac{C_k''}{(1+(v-v_0))^k} \sum_{j=0}^{m(k)} \|\partial_t^j (f,g)|_{t=t_1}\|_{L^2(r_1,\infty)}} \tag{4.6.17}$$

for all $v \geq v_1$, $u \in [u_0, u_1]$. Here $m(k) \in \mathbb{N}$ is determined implicitly by $k = \lfloor \gamma \frac{2m(k)-3}{4} \rfloor$, and the constant $C_k'' > 0$ is given by

$$C_k'' = C_{k,0}''(1+C_k^g)^{4m(k)+2}(1+C_1^{g,dec})^{\frac{2m(k)-3}{3}}(1+|u_1-u_0|)^{\frac{m(k)+7}{2}}(1+|\lambda|)^{2m(k)+1}, \tag{4.6.18}$$

where $C_{k,0}'' = C(k, v_1, \alpha, \gamma) > 0$ is independent of $f, g, u_0, u_1, \lambda, C_k^g, C_1^{g,dec}$.

Proof. Due to the assumption that *all* derivatives of A and R^{-1} are bounded, the previous Proposition 4.6.8 applies for any $k \in \mathbb{N}$. This shows (4.6.17) and (4.6.18) up to the claims about $m(k)$. First a quick look back at the decay rate in (4.6.15) shows that the implicit equation $k = \lfloor \gamma \frac{2m(k)-3}{4} \rfloor$ indeed picks the correct $m(k) \in \mathbb{N}$ in order for (4.6.17) to hold. ☐

Remark 4.6.10. One can of course determine some bounds for the number $m(k)$ in the estimate (4.6.17). It is straight-forward to see that one always has $m(k) > 2k + \frac{3}{2}$. Concerning an upper bound, it holds that $m(k) < 2(k+1) + \frac{3}{2}$ in case the decay rate of the metric satisfies $\alpha > 1$. If $\alpha < 1$, then at least one can show that $m(k) < \frac{8(k+1)}{\alpha} + \frac{3}{2}$.

4.7. Decay in Null Directions of General Solutions

Now we finally return to the actual solution $\psi \in \Gamma^\infty(\mathbb{S}M)$ of the massive Dirac equation $(\mathcal{D}-m)\psi = 0$ on M, which we are actually interested in. Using its decomposition into angular modes and the mode estimates derived in the previous sections, we now obtain various estimates for ψ. To express these estimates, we will use certain norms on spinors which are described in Section 3.4.

For simplicity, all the statements in the following are for the case that the spacetime is C^∞-bounded at infinity. One can show analogous results in the case of C^k-boundedness for finite k in the same manner, using the corresponding mode decay results of before.

4.7.1. Decay of $L^2(N)$-averages. In Theorem 4.6.9 we have obtained decay of the angular momentum modes f_λ and g_λ. Since the dependence on the λ in these decay estimates is explicitly known, we can turn them into estimates of the $L^2(N)$-average of ϕ (and ψ) by summing over λ and using Parseval's identity. More precisely, we will estimate the quantity

$$\|\psi(v,u,\cdot)\|^2_{L^2_\omega(N)} := \int_N |\psi(v,u,\omega)|^2_T \,\mathrm{d}\mu_N(\omega)\,, \tag{4.7.1}$$

where v and u are the usual null coordinates (4.2.6).[13]

Since we will need it, let us also recall that on $\Sigma = (r_0,\infty)_r \times N$ we use two L^2-inner products which differ by the volume measure which is used (cf. Section 3.4.2): On the one hand we use the volume measure $\mathrm{d}\mu_\Sigma$ which is induced from any constant t embedding of Σ into M, and the corresponding L^2-space is denoted by $L^2(\Sigma)$. On the other hand, we also use the measure $\mathrm{d}r\,\mathrm{d}\mu_N$ and the corresponding L^2-space is denoted by $L^2((r_0,\infty)\times N)$. The use of these two different measures is related to the conformal rescaling of spinor fields which we did at the beginning.

Theorem 4.7.1. (Decay of $L^2(N)$-average) *Assume that (M,g) is a spacetime of the form (3.1.1), which is C^∞-bounded and C^1-asymptotically flat at $r=\infty$ in the sense of Definition 3.2.1. Fix $r_{\max} > r_0$ and $u_1 > r_{\max}$, and let u_0, t_1, r_1, v_0 be as described in Section 4.6.3. Let $\psi \in \Gamma^\infty(\mathbb{S}M)$ be a solution of $(\mathcal{D}-m)\psi = 0$ with* $\operatorname{supp}\psi|_{t=0} \subset (r_0, r_{\max})\times N$*. Then for every $k\in\mathbb{N}$ there exist $C_k > 0$ and $m(k)\in\mathbb{N}$ such that*

$$\|\psi(v,u,\cdot)\|_{L^2(N)} \le C_k \frac{R(\frac{v-u}{2})^{-\frac{n-1}{2}}}{(1+(v-v_0))^k}\,\|\psi|_{t=t_1}\|_{H^{3m(k)+1}(\Sigma\cap\{r>r_1\})} \tag{4.7.2}$$

for all $v\ge v_0$ and all $u\in[u_0,u_1]$. The constant C_k is explicitly given by

$$C_k = C_{k,0} C_N (1+C^g_k)^{6m(k)+2}(1+C^{g,dec}_1)^{\frac{2m(k)-3}{3}}(1+|u_1-u_0|)^{\frac{m(k)+7}{2}}\,, \tag{4.7.3}$$

where $C_{k,0}$ is purely combinatorical and C_N depends on $\mathcal{D}_N$. Further, $m(k)\in\mathbb{N}$ is as in Theorem 4.6.9, and depends on the rate of decay at infinity of the metric.

Proof. Let $\alpha > 0$ be the rate of asymptotic flatness, and fix $0<\gamma<\min\{1,\frac{\alpha}{4}\}$. Next, let $\phi = e^{\frac{a}{2}} R^{\frac{n-1}{2}}$ be the rescaled spinor field which satisfies the rescaled Dirac equation (3.4.5), and let

$$\phi(t,r,\omega) = \sum_\lambda \begin{pmatrix} \phi_1(t,r)\Xi_\lambda \\ \phi_2(t,r)\Xi_\lambda \end{pmatrix} \tag{$*$}$$

be its $L^2(N)$-orthonormal decomposition into the eigenbasis $\{\Xi_\lambda\}_\lambda$ of $\mathcal{D}_N$. Further, for each $\lambda\in\sigma(\mathcal{D}_N)$, set

$$f_\lambda = 2(\phi_{\lambda,1}+\phi_{\lambda,2})\,, \qquad g_\lambda = 2(\phi_{\lambda,1}-\phi_{\lambda,2})\,.$$

Then, as seen in Section 4.2.2, f_λ, g_λ, as functions of the null coordinates v and u, satisfy the Dirac null system (4.2.10) with potential V_λ given by (4.2.9). Moreover, due to the support assumptions on $\psi|_{t=0}$ we clearly have $\operatorname{supp} f_\lambda|_{t=0}, \operatorname{supp} g_\lambda|_{t=0} \subset (r_0, r_{\max})$, so

[13] For the definition of (4.7.1) it does not matter of course if we use the coordinates v,u, or the coordinates t,r, or any other coordinates.

that the support condition (4.3.3) is satisfied. Therefore we can apply Theorem 4.6.9, from which it follows that

$$|f_\lambda(v,u)|, |g_\lambda(v,u)| \leq \frac{C_k}{(1+(v-v_0))^k}(1+|\lambda|)^{2m(k)+1} \sum_{j=0}^{m(k)} \|\partial_t^j (f_\lambda, g_\lambda)|_{t=t_1}\|_{L^2(r_1,\infty)_r} \,,$$

where $m(k) \in \mathbb{N}$ is implicitly determined by $k = \lfloor \gamma \frac{2m(k)-3}{4} \rfloor$, and where the constant $C_k > 0$ coincides with (4.6.18) up to the factor involving λ, which is pulled out in the estimate used here.

Clearly the same estimate also holds with f_λ, g_λ replaced by $\phi_{\lambda,1}, \phi_{\lambda,2}$. Therefore, using Parseval's identity for the orthonormal basis $\{\Xi_\lambda\}_\lambda$ it follows that

$$\begin{aligned}
\|\phi(v,u,\cdot)\|^2_{L^2(N)} &= \sum_\lambda \left(|\phi_{\lambda,1}(v,u)|^2 + |\phi_{\lambda,2}(v,u)|^2\right) \|\Xi_\lambda\|^2_{L^2(N)} \\
&\leq \sum_\lambda \frac{C_k^2}{(1+(v-v_0))^{2k}}(1+|\lambda|)^{2(2m(k)+1)} \\
&\quad \times \left(\sum_{j=0}^{m(k)} \|\partial_t^j(\phi_{\lambda,1}, \phi_{\lambda,2})|_{t=t_1}\|_{L^2(r_1,\infty)_r} \right)^2 \|\Xi_\lambda\|^2_{L^2(N)} \,.
\end{aligned}$$

By the Hölder inequality (for sums), we have

$$\left(\sum_{j=0}^{m(k)} \|\partial_t^j(\phi_{\lambda,1}, \phi_{\lambda,2})|_{t=t_1}\|_{L^2(r_1,\infty)_r} \right)^2 \leq (m(k)+1) \sum_{j=0}^{m(k)} \|\partial_t^j(\phi_{\lambda,1}, \phi_{\lambda,2})|_{t=t_1}\|^2_{L^2(r_1,\infty)_r} \,,$$

so that our previous estimate continues with [14]

$$\begin{aligned}
\|\phi(v,u,\cdot)\|^2_{L^2(N)} &\leq \frac{C_k^2}{(1+(v-v_0))^{2k}} \\
&\quad \times \sum_{j=0}^{m(k)} \sum_\lambda \|\partial_t^j(\phi_{\lambda,1}, \phi_{\lambda,2})|_{t=t_1}\|^2_{L^2(r_1,\infty)} (1+|\lambda|)^{2(2m(k)+1)} \|\Xi_\lambda\|^2_{L^2(N)} \\
&= \frac{C_k^2}{(1+(v-v_0))^{2k}} \\
&\quad \times \sum_{j=0}^{m(k)} \sum_\lambda \|\partial_t^j(\phi_{\lambda,1}, \phi_{\lambda,2})|_{t=t_1}\|^2_{L^2(r_1,\infty)} \|(1+|\mathcal{D}_N|)^{2m(k)+1}\Xi_\lambda\|^2_{L^2(N)} \\
&= \frac{C_k^2}{(1+(v-v_0))^{2k}} \sum_{j=0}^{m(k)} \|\partial_t^j (1+|\mathcal{D}_N|)^{2m(k)+1}\phi|_{t=t_1}\|^2_{L^2((r_1,\infty)\times N)} \,.
\end{aligned}$$

Scaling back to ψ, this yields the estimate

$$\begin{aligned}
\|\psi(v,u,\cdot)\|^2_{L^2(N)} &= \mathrm{e}^{-a(v,u)} R(v,u)^{-(n-1)} \|\phi(v,u,\cdot)\|^2_{L^2(N)} \\
&\leq \mathrm{e}^{-a(v,u)} R(v,u)^{-(n-1)} \frac{C_k^2}{(1+(v-v_0))^{2k}}
\end{aligned}$$

[14] We absorb the factor $m(k)+1$ into the constant C_k. From now on, such purely combinatorical changes of the constant will no longer be explicitly mentioned.

$$\times \sum_{j=0}^{m(k)} \|\partial_t^j (1+|\mathcal{D}_N|)^{2m(k)+1}\phi|_{t=t_1}\|^2_{L^2((r_1,\infty)\times N)}$$
$$= \mathrm{e}^{-a(v,u)} R(v,u)^{-(n-1)} \frac{C_k^2}{(1+(v-v_0))^{2k}}$$
$$\times \sum_{j=0}^{m(k)} \|(1+|\mathcal{D}_N|)^{2m(k)+1}\partial_t^j \psi|_{t=t_1}\|^2_{L^2(\Sigma\cap\{r>r_1\})} \,. \qquad (*)$$

In the last step we used that the $L^2((r_0,\infty)\times N)$-norm of ϕ is the same as the $L^2(\Sigma)$-norm of ψ (cf. Lemma 3.4.3), and that this of course also applies to any non-radial derivatives of ϕ and ψ. Note also that we interchanged ∂_t^j with $(1+|\mathcal{D}_N|)^{2m(k)+1}$, which is possible since these operators commute. Concerning the two factors e^{-a} and $R^{-(n-1)}$, notice that the first one simply goes to one at infinity, so we absorb it into the constant. The second factor decay since $R\to\infty$, so we keep it in the estimate.

Next we use that ψ satisfies the Dirac equation, which allows to convert t-derivatives into spatial derivatives and zeroth order terms. Concretely, one can rewrite the equation $(\mathcal{D}-m)\psi=0$ with $\mathcal{D}$ explicitly given by (4.2.1) as

$$\partial_t\psi = -\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \left(\partial_r + \tfrac{a'}{2} + \tfrac{n-1}{2}\tfrac{R'}{R}\right)\psi - \frac{\mathrm{e}^a}{R}\begin{pmatrix} 0 & \mathcal{D}_N \\ -\mathcal{D}_N & 0 \end{pmatrix}\psi - \mathrm{i}m\mathrm{e}^a\psi\,.$$

Rearranging terms, we abbreviate this expression in the form

$$\partial_t\psi =: A(r)\partial_r\psi + B(r)\,\mathcal{D}_N\psi + C(r)\psi$$

for certain "block-matrices" $A(r), B(r), C(r)$ which depend on r through the functions $a(r), R(r)$ and their first derivatives. Iterating this identity, one easily sees that

$$\partial_t^j\psi = \sum_{i=0}^{j}\sum_{\ell=0}^{i} A_{i\ell}^{(j)}(r)\partial_r^\ell\,\mathcal{D}_N^{i-\ell}\psi\,,$$

where $A_{i\ell}^{(j)}(r)$ are again block-matrices which depend on r through the functions $a(r)$ and $R(r)$ and their first j derivatives. Moreover, the functional dependence on e^a and R^{-1} (and their derivatives) is polynomial of order at most 2^j. Therefore it follows from (3.4.17) that

$$\left\|(1+|\mathcal{D}_N|)^{2m(k)+1}\partial_t^j\psi|_{t=t_1}\right\|_{L^2(\Sigma\cap\{r>r_1\})} \le C_N(1+C_j^g)^{2^j}\,\|\psi|_{t=t_1}\|_{H^{2m(k)+j+1}(\Sigma\cap\{r>r_1\})}\,,$$

where $C_N>0$ denotes the operator norm of $(1+|\mathcal{D}_N|)^{2m(k)+1}$, viewed as bounded operator from $\Gamma_{H^{2m(k)+1}}(\mathbb{S}N)$ to $\Gamma_{L^2}(\mathbb{S}N)$.

Together with $(*)$ this now yields the desired estimate. □

4.7.2. Pointwise decay estimates. Since the Dirac operator $\mathcal{D}$, given by (4.2.1), commutes with $\mathcal{D}_N$, it is not difficult to pass from the decay of the $L^2(N)$-average of ψ to pointwise decay using that $\mathcal{D}_N$ is elliptic together with some standard Sobolev embedding results. Recall for the following that $n-1=\dim N$.

Theorem 4.7.2. (Pointwise decay) *Let $s>\frac{n-1}{2}$, and assume the same assumptions as in Theorem 4.7.1 hold. Then for every $k\in\mathbb{N}$ there exist $C_{k,s}>0$ and $m(k)\in\mathbb{N}$ such that*

$$|\psi(v,u,\omega)|_T \le C_{k,s}\frac{R(\frac{v-u}{2})^{-\frac{n-1}{2}}}{(1+(v-v_0))^k}\,\|\psi|_{t=t_1}\|_{H^{3m(k)+1+s}(\Sigma\cap\{r>r_1\})} \qquad (4.7.4)$$

for all $v \geq v_0$, $u \in [u_0, u_1]$, $\omega \in N$. Here the constant $C_{k,s} > 0$ is explicitly given by

$$C_k = C_{k,0} C_{N,s} (1 + C_k^g)^{6m(k)+2} (1 + C_1^{g,dec})^{\frac{2m(k)-3}{3}} (1 + |u_1 - u_0|)^{\frac{m(k)+7}{2}} , \tag{4.7.5}$$

where $C_{k,0} > 0$ is purely numerical, and $C_{N,s} > 0$ depends on the optimal constants in the Sobolev embedding $\Gamma_{H^s}(\mathbb{S}N) \hookrightarrow \Gamma^0(\mathbb{S}N)$ and in elliptic estimates for $\mathcal{D}_N$. The number $m(k) \in \mathbb{N}$ is as in Theorem 4.7.1.

Proof. By the embedding $\Gamma_{H^s}(\mathbb{S}N) \hookrightarrow \Gamma^0(\mathbb{S}N)$ (cf. [**LM89**, Ch.III, Thm.2.15]) and the standard elliptic estimate for $\mathcal{D}_N$ (cf. [**LM89**, Ch.III, Thm.5.2]) we have (the constant $C_{N,S}$ again changes from line to line)

$$\begin{aligned}
|\psi(v,u,\omega)|_T &\leq C_{N,s} \left\| \psi(v,u,\cdot) \right\|_{H^s(N)} \\
&\leq C_{N,S} \big(\left\| \mathcal{D}_N^s \psi(v,u,\cdot) \right\|_{L^2(N)} + \left\| \psi(v,u,\cdot) \right\|_{L^2(N)} \big) \\
&\leq C_k C_{N,s} \frac{\big(1 - A(\frac{v-u}{2})\big)^{-\frac{1}{2}} R(\frac{v-u}{2})^{-\frac{n-1}{2}}}{(1 + (v - v_0)^k} \\
&\quad \times \big(\left\| \mathcal{D}_N^s \psi|_{t=t_1} \right\|_{H^{3m(k)+1}(\Sigma)} + \left\| \psi|_{t=t_1} \right\|_{H^{3m(k)+1}(\Sigma)} \big) \\
&\leq C_{k,s} \frac{\big(1 - A(\frac{v-u}{2})\big)^{-\frac{1}{2}} R(\frac{v-u}{2})^{-\frac{n-1}{2}}}{(1 + (v - v_0)^k} \left\| \psi|_{t=t_1} \right\|_{H^{3m(k)+1+s}(\Sigma)} .
\end{aligned}$$

Here the third estimate is of course an application of Theorem 4.7.1. Notice that we use here that both ψ and $\mathcal{D}_N^s \psi$ are solutions of the Dirac equation, so that the decay estimate of Theorem 4.7.1 applies to both. $\square$

4.7.3. The globally hyperbolic case. In this last section, we restrict to the globally hyperbolic case, which apples for instance to the exterior Schwarzschild spacetime (cf. Example 3.1.4). Recall that a spacetime (M, g) of the form (3.1.1) is globally hyperbolic if and only if $r_0 = -\infty$ (and $r_1 = \infty$, which was always satisfied so far anyways), cf. Sec. 3.1.2.

A special feature in the globally hyperbolic case is that any outgoing null strip completely intersects the Cauchy hypersurface $\{t = 0\}$. Therefore, in the notation used before, we can always set $t_1 = 0$ which makes the results a bit easier to state. Figure 4.20 illustrates this in a Penrose diagram.

The following two statements are simply restatements of the previous decay estimates in the globally hyperbolic situation. We start with the decay of $L^2(N)$-averages.

Theorem 4.7.3. (Decay of $L^2(N)$-average) *Let (M, g) be a globally hyperbolic spacetime of the form (3.1.1), and assume that it is C^∞-bounded and C^1-asymptotically flat at $r = \infty$ in the sense of Definition 3.2.1. Let $\psi \in \Gamma_{sc}^\infty(\mathbb{S}M)$ be a spatially compactly supported solution of $(\mathcal{D} - m)\psi = 0$. Fix $u_0 < u_1$ with $\operatorname{supp} \psi|_{t=0} \subset (-u_1, -u_0) \times N$. Then for every $k \in \mathbb{N}$ there exist $C_k > 0$ and $m(k) \in \mathbb{N}$ such that*

$$\left\| \psi(v,u,\cdot) \right\|_{L^2(N)} \leq C_k \frac{R(\frac{v-u}{2})^{-\frac{n-1}{2}}}{(1 + (v - v_0))^k} \left\| \psi|_{t=0} \right\|_{H^{3m(k)+1}(\Sigma)} \tag{4.7.6}$$

for all $v > 0$ and all $u \in [u_0, u_1]$. The constant C_k is explicitly given by

$$C_k = C_{k,0} C_N (1 + C_k^g)^{6m(k)+2} (1 + C_1^{g,dec})^{\frac{2m(k)-3}{3}} (1 + |u_1 - u_0|)^{\frac{m(k)+7}{2}} , \tag{4.7.7}$$

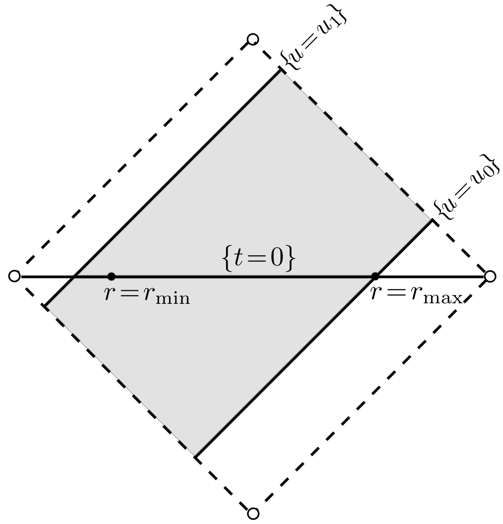

FIGURE 4.20. The situation in case that (M, g) is globally hyperbolic, illustrated by the respective Penrose diagram.

where C_k^g and $C_1^{g,dec}$ are the constants in the boundedness respective decay estimates for the metric coefficients $A(r), R(r)^{-1}$ on the interval $[-u_1, \infty)$, $C_{k,0}$ is purely combinatorical, and C_N depends on $\mathcal{D}_N$. The number $m(k) \in \mathbb{N}$ is as in Theorem 4.7.1.

Also the pointwise decay estimates can of course be restated. Recall once more that $\dim N = n - 1$.

Theorem 4.7.4. (Pointwise decay) *Let $s > \frac{n-1}{2}$, and assume that the same assumptions as in Theorem 4.7.3 hold. Then for every $k \in \mathbb{N}$ there exist $C_{k,s} > 0$ and $m(k) \in \mathbb{N}$ such that*

$$|\psi(v, u, \omega)|_T \leq C_{k,s} \frac{R(\frac{v-u}{2})^{-\frac{n-1}{2}}}{(1 + (v - v_0))^k} \, \|\psi|_{t=0}\|_{H^{3m(k)+1+s}(\Sigma)} \tag{4.7.8}$$

for all $v > 0$, $u \in [u_0, u_1]$, $\omega \in N$. Here the constant $C_{k,s} > 0$ is explicitly given by

$$C_k = C_{k,0} C_{N,s} (1 + C_k^g)^{6m(k)+2} (1 + C_1^{g,dec})^{\frac{2m(k)-3}{3}} (1 + |u_1 - u_0|)^{\frac{m(k)+7}{2}}, \tag{4.7.9}$$

where C_k^g and $C_1^{g,dec}$ are the constants in the boundedness respective decay estimates for the metric coefficients $A(r), R(r)^{-1}$ on the interval $[-u_1, \infty)$, $C_{k,0} > 0$ is purely numerical, and $C_{N,s} > 0$ depends on the constant in the Sobolev embedding $\Gamma_{H^s}(\mathbb{S}N) \hookrightarrow \Gamma^0(\mathbb{S}N)$ and on constants in elliptic estimates for $\mathcal{D}_N$. The number $m(k) \in \mathbb{N}$ is as in Theorem 4.7.1.

We conclude this chapter by formulating an immediate corollary of the previous decay result, which makes a less precise statement but is simpler to read and still contains the main point.

Corollary 4.7.5. *Let (M, g) be a globally hyperbolic spacetime of the form (3.1.1), and assume that it is C^∞-bounded and C^1-asymptotically flat at $r = \infty$ in the sense of Definition 3.2.1. Let $\psi \in \Gamma_{sc}^\infty(\mathbb{S}M)$ be a solution of the massive Dirac equation $(\mathcal{D} - m)\psi = 0$. Then for each $k \in \mathbb{N}$ and each finite interval $[u_0, u_1]$ with $\operatorname{supp} \psi|_{t=0} \subset (-u_1, -u_0) \times N$*

there exist $C_k > 0$ and $s(k) \in \mathbb{N}$ such that

$$|\psi(v,u,\omega)|_T \leq \frac{C_k}{(1+v)^k} \, \|\psi|_{t=0}\|_{H^{s(k)}(\Sigma)} \qquad \forall v > 0 \tag{4.7.10}$$

for all $u \in [u_0, u_1], \omega \in N$. The constants $C_k > 0$ and $s(k)$ are independent of ψ.

In words, spatially compactly supported solutions of the massive Dirac equation decay superpolynomially fast in outgoing null directions.

CHAPTER 5

Outlook and some Open Questions

To close this thesis, let me point to some possible improvements of the results obtained in this thesis, and give an outlook on remaining open problems and possible future work. At least on some of these questions I have the intention to continue working myself in the future.

Possible improvements of the results. I begin with a list of a few possible improvements of the results obtained in Section 4.7, which I believe to be rather straight-forward to implement.

The non-static case: One of the obvious questions is whether the results remain the same in case the underlying spacetime is no longer static. I believe this should still be the case at least if the spacetime still has the form $M = \mathbb{R}_t \times (r_0, \infty)_r \times N$ with metric

$$g = \mathrm{e}^{2a(t,r)}[-\,\mathrm{d}t^2 + \mathrm{d}r^2] + R(t,r)^2 g_N \,, \tag{5.0.11}$$

and if one assumes that the function $a(t,r)$ and $R(t,r)$ have the same asymptotic behaviour as specified in Definition 3.2.1, only that this time we also demand boundedness and decay conditions for t-derivatives (over each hypersurfaces $\{t\} \times (r_0, \infty)$ with constants uniform in t). The reason why I believe the results to remain the same is that *current conservation*, which is the crucial a priori estimate used in order to keep solutions bounded, still holds in the non-static case. Of course, the explicit estimates become more involved, in particular when differentiating the equation in t (as done in Section 4.5.1), but as long as time-derivatives of the metric coefficients remain bounded all the a priori estimates should work out. Therefore, if a and R^{-1} decay sufficiently at infinity, one should be able to use the same perturbative approach based on the Klein-Gordon equation in flat 1+1 dimensional Minkowski spacetime.

Removing the spatially compact support assumption: We always assumed that the solution of the massive Dirac equation is spatially compactly supported. One should be able to replace this assumption by a sort of weaker "finite energy condition" instead. More precisely, notice that on first glance the estimates in the final decay results in Section 4.7 still make sense if the initial data has finite Sobolev norm (of any order for super-polynomial decay). In detail, one has to be more cautious since also the constant in the estimate depends on the size of the support of the initial data. Therefore a simple density argument does not directly help. Nevertheless, I believe that one should still be able to generalize to some sort of finite energy solutions by slightly adjusting the estimates which were used. For instance, going back to Section 4.3.1 where the support condition was first introduced, one might simply try to keep the first term $g(v, u_0)$ in (4.3.2), which basically refers to $g|_{t=0}$, and see if suitable conditions on the initial data can be deduced in this way.

Non-compact N: Next, one can of course ask what happens for a non-compact Riemannian manifold N in the decomposition $M = Q \times_R N$ of the underlying spacetime. To answer this question, we recall that it were basically two properties of N which entered into the analysis: Firstly, we used an L^2-decomposition of the solution of the Dirac equation into eigenspinors of $\mathcal{D}_N$ (cf. Section 4.2.1). And secondly, we used a Sobolev embedding and elliptic estimates for $\mathcal{D}_N$ in order to deduce pointwise estimates (cf. Section 4.7.2). Concerning the first property, it should be sufficient that $\mathcal{D}_N$ is essentially self-adjoint on $\Gamma_c^\infty(\mathbb{S}N) \subset \Gamma_{L^2}(\mathbb{S}N)$ and possesses an orthonormal basis of "generalized eigenspinors", i.e. smooth solutions (not necessarily in L^2) of $\mathcal{D}_N \Xi_\lambda = \lambda \Xi_\lambda$ for every $\lambda \in \sigma(\mathcal{D}_N)$ which form a complete orthonormal basis in an appropriate sense (see [**Shu92**, App. 2.2] for this notion). In this case one should be able to decompose any solution of the Dirac equation on M into these generalized eigenspinors and proceed analogously to Chapter 4. This, and also the validity of the Sobolev embedding and elliptic estimate for $\mathcal{D}_N$ are satisfied if (M, g) has *bounded geometry* (see [**Shu92**]). Therefore I believe that one should be able to extent the results of this thesis to this case in a rather straight-forward way.

Spacetimes with more "degrees of freedom": The spacetimes considered in this thesis had the particular feature of being warped products $M = Q \times_R N$ over a 1+1 dimensional base Q. This had the effect that the coefficients in the Dirac equation essentially only depended on the coordinates t and r on Q (actually even only on r, but see the first point of this list).

One can of course ask what happens if one makes the spacetimes more complicated, for instance by taking a 1+2 dimensional base Q (or, more generally, a $1+k$ dimensional base). As long as this base manifold is conformally equivalent to piece of Minkowski spacetime in a similar way as was true in the situation of this thesis, one can probably proceed similarly as in Chapter 4. Apart from the fact that all estimates become more complicated, the most important point one needs to take care of is to obtain a suitable representation formula for solutions of the corresponding free equation in flat Minkowski spacetime, similar to how it was done in Section 4.4. This representation formula was the catalyst of all the decay estimates in the end. It should be possible in higher dimensions to do this using explicit expressions for the Green's function of the massive Dirac or Klein-Gordon equation in higher dimensional Minkowski spacetime, which will of course be more complicated however. One should also be aware that the general complexity is increased by the fact that the number of components of spinors grows exponentially fast with respect to the dimension of the spacetime. Apart from these complications, I would expect the general approach to work rather similarly.

Spacetimes which are no warped products: As remarked in the previous points, the spacetimes in this thesis are warped products $M = Q \times_R N$, and one can of course ask how crucial this is for the analysis and if it can be relaxed. Different than the previous points, this is not so clear since the approach taken in this thesis relied heavily on separating off the N-dependence and working in the lower dimensional spacetime Q. One possible way how one could try to generalize is relaxing the condition that M is a warped product to the condition that it is only a *bundle* over Q with typical fiber N.

Open problems and future directions. Here I want to indicate a few open problems which remain, as well as possible further working directions.

Exponential decay? One can ask if the decay results obtained in this thesis are actually optimal, or if one can perhaps even show exponential decay. The reason for bringing up this question here is the "potential wall"-picture made on page xi of the introduction. A vague idea to approach this could be to try to take some ideas from proofs of exponential decay of eigenfunctions of Schrödinger operators (see for instance [**HS96**, 3]). This might be a little far fetched since Schrödinger operators are elliptic, whereas the Dirac equation on a Lorentzian manifold is hyperbolic, but one could certainly try if some ideas can be used.

Behaviour at ι^+? The decay estimates which were obtained in this thesis are always inside an outgoing null strip of finite size (compare to figure 4.18). One glaring question this leaves open is the behaviour of ι^+.

Firstly, one may note that the constants in the estimates diverge if one lets the size of the strip tend to infinity (cf. (4.7.7)). This indicates that the superpolynomial decay inside a null strip does not hold in timelike directions, as should not be surprising in view of the results in [**FKSY03**]. Nevertheless, one can wonder if it is perhaps possible to increase the size of the strip simultaneously as the outgoing null coordinate v increases in precisely such a way that the decay in v in the estimate wins against the growth of the constant. If one really wants to proceed in this way, one might also wish to tighten some of the estimates made in Chapter 4 to optimize the dependence of the constant on the size of the strip. In any case such an approach seems rather delicate.

Another idea to study the behaviour in timelike directions could be to pick up again the idea of "compactifying" the underlying spacetime somehow and trying to extend the equation or its solution to the boundary attached in this way. As indicated in the introduction, this did not work out for the lightlike boundary obtained using the Penrose conformal compactification of Minkowski spacetime (at least not in the way the author tried), but perhaps it can be made to work in timelike directions. For instance, perhaps one can "blow up" the point ι^+ in the Penrose diagram (see for instance [**ACN12**, Sec. 2] for the notion of blowup) in such a way that the equation has a nice form in the blown-up spacetime as one tends to infinity in timelike directions.

Another idea could be to try to use completely different compactifications, for instance the one for asymptotically Minkowski spacetimes used in [**BVW12**]. In this context, maybe instead of really continuing the equation to the boundary, one can at least try to set up a functional framework to study asymptotics similar to how it is done in [**BVW12**].

Model problems? In order to understand the behaviour around ι^+, one might try to construct some simple model problems which one can perhaps either solve explicitly or at least study more easily than actual problems. One idea would be to take a simple transport or wave equation in a diamond-shaped region of Minkowski spacetime (resembling the Penrose conformal compactification), and adding a simple explicit potential which diverges on the boundary.

More generally, it might be interesting to study characteristic initial value problems with a "singularity" on the initial hypersurface to see if a solution has any chance of "reaching" the initial hypersurface or if it will just propagate parallel to it or do something else.

Spacetimes with a different asymptotic structure? Extending the ideas of conformal scattering to (non-conformally invariant) massive equations in asymptotically simple spacetimes, as outlined in the introduction, seems to be difficult. As was also outlined in the introduction, one basic problem seems to be the fact that the conformally attached boundary of an asymptotically simple spacetime is lightlike. Therefore one could try to repeat the attempt of doing conformal scattering for massive equations in other spacetimes whose conformal boundary is spacelike. Such spacetimes are known as *asymptotically de Sitter spacetimes*, and trying to do conformal scattering for massive equations on such spacetimes is currently studied by J.-P. Nicolas and the author of this thesis.

APPENDIX A

Miscellaneous Results in Lorentzian Geometry

A.1. Static Spacetimes

In the following we collect some results about spacetimes admitting timelike Killing vector fields which are common folklore and which can for instance be found in the articles [**CMP03**], [**Ś05**], [**SS07**], and [**JS08**].

A.1.1. Definitions and some examples. First, let us recall that a vector field $K \in \Gamma^\infty(TM)$ is a *Killing field* if and only if $\mathscr{L}_K\, g = 0$, or, equivalently, if the (local) flow generated by K consists of isometries. Explicitly this means the following.

Lemma A.1.1. *A vector field $K \in \Gamma^\infty(TM)$ is a Killing vector field if and only if*

$$\langle \nabla_X K, Y\rangle = -\langle X, \nabla_Y K\rangle \qquad \forall X, Y \in \Gamma^\infty(TM)\,. \tag{A.1.1}$$

Proof. Using the Leibniz rule for the Lie derivative, we have

$$\begin{aligned}
0 &= (\mathscr{L}_K\, g)(X,Y)\\
&= Kg(X,Y) - g([K,X],Y) - g(X,[K,Y])\\
&= g(\nabla_K X, Y) + g(X, \nabla_K Y) - g(\nabla_K X - \nabla_X K, Y) - g(X, \nabla_K Y - \nabla_Y K)\\
&= \langle \nabla_X K, Y\rangle + \langle X, \nabla_Y K\rangle\ .
\end{aligned}$$

□

Coming to the central theme of this section, (M,g) is said to be *stationary* there exists a timelike Killing vector field $K \in \Gamma^\infty(TM)$. It is said to be *static* if there exists a timelike Killing vector field $K \in \Gamma^\infty(TM)$ whose orthogonal distribution $K^\perp \subset TM$ is completely integrable.[1] Explicitly, this condition on K means the following.

Lemma A.1.2. *Let $K \in \Gamma^\infty(TM)$ be a timelike Killing vector field. Then $K^\perp \subset TM$ is completely integrable if and only if the* curl *of the normalized vector field $\widetilde{K} := -\langle K,K\rangle^{-1} K$, i.e. the 2-form*

$$\mathrm{curl}\widetilde{K} = \mathrm{d}\widetilde{K}^\flat \tag{A.1.2}$$

vanishes. In this case one calls K (or rather $\widetilde{K}$) irrotational.[2]

Proof. Notice first that for any vector field K, then clearly $\mathrm{d}\widetilde{K}^\flat(X,Y) = 0$ whenever X and Y are collinear to K. Moreover, if K is Killing then for any vector field $X \perp K$ we have by the Koszul formula for the exterior derivative

$$\mathrm{d}\widetilde{K}^\flat(X,K) = X\,\underbrace{\widetilde{K}^\flat(K)}_{=1} - K\,\underbrace{\widetilde{K}^\flat(X)}_{=0} - \widetilde{K}^\flat([X,K])$$

[1] We refer [**Lee03**, Ch. 19] for the notion of (integrable) distributions of the tangent bundle.

[2] Notice that of course we have $K^\perp = \widetilde{K}^\perp$.

$$\begin{aligned}
&= \frac{1}{\langle K,K\rangle}\big(\underbrace{\langle K,\nabla_X K\rangle}_{=-\langle\nabla_K K,X\rangle} - \langle K,\nabla_K X\rangle\big)\\
&= -\frac{1}{\langle K,K\rangle}\big(\langle\nabla_K K,X\rangle + \langle K,\nabla_K X\rangle\big)\\
&= -\frac{1}{\langle K,K\rangle}K\underbrace{\langle K,X\rangle}_{=0}\\
&= 0\,.
\end{aligned}$$

Therefore the only obstacle for the vanishing of $\mathrm{d}\widetilde{K}^\flat$ is $\mathrm{d}\widetilde{K}^\flat(X,Y)$ for both $X,Y\perp K$. Using the Koszul formula again, in this case we have

$$\mathrm{d}\widetilde{K}^\flat(X,Y) = X\underbrace{\widetilde{K}^\flat(Y)}_{=0} - Y\underbrace{\widetilde{K}^\flat(X)}_{=0} - \widetilde{K}^\flat([X,Y]) = -\langle K,K\rangle^{-1}\langle K,[X,Y]\rangle\,.$$

This shows that $\mathrm{d}\widetilde{K}^\flat = 0$ if and only if $[X,Y]\perp K$ for any vector fields $X,Y\perp K$. But this is precisely the criterion for the complete integrability of $K^\perp$, cf. [**Lee03**, Thm. 19.10]. □

It is important to keep in mind that a Lorentzian manifold can have various different timelike Killing vector fields (or none at all). Let us illustrate this by some examples.

Example A.1.3. Consider $1+1$ dimensional Minkowski spacetime $\mathbb{R}^{1,1} = \mathbb{R}_t\times\mathbb{R}_x$ with metric $g = -\mathrm{d}t^2 + \mathrm{d}x^2$. Then $K = \partial_t$ is an irrotational Killing vector field. The leaves of its orthogonal distribution are the spacelike hypersurfaces $\Sigma_t = \{t\}\times\mathbb{R}\subset\mathbb{R}^{1,1}$. But there exist many more irrotational Killing fields: For any $\beta\in\mathbb{R}$, consider the Lorentz transformation

$$A_\beta = \begin{pmatrix}\cosh\beta & \sinh\beta\\ \sinh\beta & \cosh\beta\end{pmatrix}\,.$$

Then also the vector field

$$K_\beta := A_\beta K = \cosh\beta\,\partial_t + \sinh\beta\,\partial_x$$

is an irrotational timelike Killing vector field.

Example A.1.4. Consider the *Rindler wedge*

$$M = \{(t,x)\in\mathbb{R}^{1,1}\mid x>0, x^2-t^2>0\}\,,$$

equipped with the flat Minkowski metric $g = -\mathrm{d}t^2 + \mathrm{d}x^2$. Then of course ∂_t is an irrotational timelike Killing field of (M,g). More interestingly, also the vector field

$$K := x\partial_t + t\partial_x$$

is an irrotational timelike Killing field. To see this, one can make a coordinate transformation from (t,x) to $(\tau,\sigma)\in\mathbb{R}\times(0,\infty)$ defined via

$$t = \sigma\sinh\tau\,,\qquad x = \sigma\cosh\tau\,.$$

As one easily verifies, in these coordinates we have

$$g = -\sigma^2\,\mathrm{d}\tau^2 + \mathrm{d}\sigma^2\,,\qquad K = \partial_\tau\,.$$

This shows that K is an irrotational timelike Killing vector field.

A.1.2. Standard static spacetimes. A static spacetime (M, g) with irrotational timelike Killing vector field K is said to be *standard static* if there exists a diffeomorphism such that

$$\begin{cases} M \cong \mathbb{R}_t \times \Sigma_x \\ \ g \cong -\beta(x)\,\mathrm{d}t^2 + g_\Sigma \\ K \cong \partial_t \end{cases}, \tag{A.1.3}$$

where (Σ, g_Σ) is a Riemannian manifold, and $\beta \in C^\infty(\Sigma)^+$ a smooth, positive function.

It is natural to ask for conditions when a given static spacetime is standard static. Note that this is a question on both the spacetime and the Killing field. For instance, the Rindler wedge is not standard static with respect to the Killing field ∂_t, but is standard static with respect to the Killing field $K = x\partial_t + t\partial_x$.

Now if (M, g, K) is standard static, then the Killing field K is necessarily *complete*, i.e. its flow is defined for all times. As we now show, if M is simply connected then this condition is also sufficient.

Theorem A.1.5. *Let (M, g, K) be static, and assume that M is simply connected and that K is complete. Then (M, g, K) is standard static.*

Proof. Since (M, g, K) is static, by Lemma A.1.2 we have $\mathrm{d}\widetilde{K}^\flat = 0$, where $\widetilde{K} = -\langle K, K\rangle^{-1} K$. Since M is simply connected, it follows that $\widetilde{K}^\flat = \mathrm{d}T$ for some $T \in C^\infty(M)$. Without loss we may assume that $0 \in \mathrm{im}(T)$. Since $\widetilde{K}$ is timelike, we have $\mathrm{d}T \neq 0$ everywhere. Therefore $\Sigma := T^{-1}(\{0\}) \subset M$ is an embedded spacelike hypersurface. We claim that the flow Φ of K provides the desired diffeomorphism $\Phi : \mathbb{R} \times \Sigma \to M$ such that (A.1.3) holds.

To see this, observe first that

$$\frac{\mathrm{d}}{\mathrm{d}t} T(\phi(t,p)) = \mathrm{d}T(K|_{\phi(t,p)}) = \widetilde{K}^\flat(K|_{\phi(t,p)}) = 1\,.$$

From this it follows that any integral curve of K intersects $\Sigma = T^{-1}(\{0\})$ exactly once, and so Φ is bijective. Since any flow is a local diffeomorphism, it follows that Φ is indeed a diffeomorphism. Moreover, by construction of Φ it is clear that $K \cong \partial_t$.

Finally, that the metric takes the desired form follows immediately from the fact that K is a Killing field. The function β is explicitly given by $\beta = -\langle K, K\rangle|_\Sigma$, and g_Σ is the Riemannian metric induced on $\Sigma \subset M$ by the Lorentzian metric g of M. $\square$

Let us comment on one possible error in reasoning. Namely, one should be aware that even in a standard static spacetime the integral curves of the Killing field K need *not* be geodesics. Indeed, the integral curves of K will be all geodesics if and only if the Killing field K is additionally *parallel*, which is the case if and only if the function β is constant. Of course, it might always happen that some of the integral curves of K are geodesics. For instance, in the exterior Schwarzschild spacetime the circular planetary orbits are integral curves of the standard static Killing field (cf. [**Wal84**, Sec. 6.3] or [**Str04**, Sec. 3.2]).

Remark A.1.6. Let (M, g, K) be standard static and assume that Σ in (A.1.3) is compact. Then any other standard static Killing vector field K' for (M, g) is necessarily a constant multiple of K (cf. [**SS07**]). However, there may still be other static Killing vector fields (then *not* standard static), see Example A.1.10. If Σ is simply connected, then this is excluded by Theorem A.1.5.

Remark A.1.7. Let (M, g, K) be standard static. Then the following are equivalent:

(1) (M, g) is globally hyperbolic.
(2) $\beta^{-1} g_\Sigma$ is a complete Riemannian metric.
(3) $\Sigma_t = \{t\} \times \Sigma \subset M$ is a Cauchy hypersurface for every $t \in \mathbb{R}$.

Indeed this follows easily from the fact that global hyperbolicity is a conformally invariant property and the fact that a product spacetime $(\mathbb{R}, -\,\mathrm{d}t^2) \oplus (N, g_N)$ is globally hyperbolic if and only if (N, g_N) is a complete Riemannian manifold (cf. [**BEE96**, Ch. 3]).

Conversely, even if (M, g) is globally hyperbolic, then not every static Killing vector field K is standard static. This is illustrated by Example A.1.10 below.

A.1.3. Spacetimes of dimension $1+1$. Suppose now that (M, g) is $1+1$ dimensional. Then matters are more simple, as illustrated by the following Lemma and its consequences.

Lemma A.1.8. *In $1+1$ dimensions, $K^\perp$ is completely integrable for any timelike vector field $K \in \Gamma^\infty(M)$.*

Proof. This just follows since any 1 dimensional distribution is trivially completely integrable. To see this, suppose that X, Y are vector fields orthogonal to K. We need to show that $[X, Y]$ is again orthogonal to K. This is a local issue, and locally $K^\perp$ can be spanned by a single vector field $Z \perp K$. Writing $X = fZ$ and $Y = gZ$ for some functions f, g (locally), we have $[X, Y] = (Xg)Y - (Yf)X$. This is clearly orthogonal to K. $\square$

Theorem A.1.9. *Let (M, g) be a $1+1$ dimensional Lorentzian manifold. Assume that*

i.) M is homeomorphic to $\mathbb{R}^2$,
ii.) $K \in \Gamma^\infty(TM)$ is a complete timelike Killing vector field.

Then there exists a diffeomorphism such that

$$\begin{cases} M \cong \mathbb{R}_t \times (a, b)_x \\ \ \ g \cong \beta(x)\big[-\mathrm{d}t^2 + \mathrm{d}x^2\big] \end{cases} \qquad \textit{and} \qquad K \cong \partial_t\,. \tag{A.1.4}$$

Proof. By Lemma A.1.8 it follows that (M, g, K) is static. Since K is assumed to be complete and M is simply connected, it follows further from Theorem A.1.5 that there exists a diffeomorphism such that (A.1.3) is satisfied. Now Σ must either be an open interval or S^1, and by simple connectedness of M the S^1-case is ruled out. Therefore, dropping the diffeomorphism from the notation, we have $M = \mathbb{R}_t \times (c, d)_y$, $K = \partial_t$, and the metric is given by $g = -\beta(y)\,\mathrm{d}t^2 + f(y)\,\mathrm{d}y^2$, where $\beta = -\langle K, K\rangle|_\Sigma$ and $f \in C^\infty(c, d)^+$ is some smooth positive function. To go from this to (A.1.4), fix some $y_0 \in (a, b)$ and define

$$x(y) = \int_{y_0}^{y} \left(\frac{f(s)}{\beta(s)}\right)^{\frac{1}{2}} \mathrm{d}s\,.$$

Since the functions β and f are smooth and strictly positive, this defines a diffeomorphism between the interval $(c, d)_y$ and another interval $(a, b)_x$, where

$$a = -\int_a^{y_0} \left(\frac{f(s)}{\beta(s)}\right)^{\frac{1}{2}} \mathrm{d}s \in [-\infty, 0)\,, \qquad b = \int_{y_0}^{b} \left(\frac{f(s)}{\beta(s)}\right)^{\frac{1}{2}} \mathrm{d}s \in (0, \infty]\,.$$

Notice that then we have

$$\mathrm{d}x = \frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}y = \left(\frac{f(y)}{\beta(y)}\right)^{\frac{1}{2}} \mathrm{d}y\,,$$

from which it follows that, expressed in t and x, we have

$$g = \beta(x)\big[-\mathrm{d}t^2 + \mathrm{d}x^2\big]$$

as desired. □

Example A.1.10. ([**SS07**, Ex. 1]) Take the spacetime $M = \mathbb{R}_t \times S^1_\theta$ with product metric $g = -\mathrm{d}t^2 + \mathrm{d}\theta^2$. Then $K = 2\partial_t + \partial_\theta$ is a timelike Killing vector field which is necessarily irrotational since we are in $1+1$ dimensions. The orthogonal distribution $K^\perp$ is spanned by the vector field $Z = \partial_t + 2\partial_\theta$. As is not difficult to see, any leave of the integral manifold of $K^\perp$, i.e. any integral curve of Z, is diffeomorphic to $\mathbb{R}$. Therefore it is *not* possible to write (M, g, K) as in (A.1.3). Locally this is of course still possible.

A.2. Penrose Diagrams

A *Penrose diagram* of a spacetime (M, g) is a schematic (conformal) representation of (M, g) by a (typically bounded) subset of $1+1$ dimensional Minkowski spacetime. Since bounded subsets of $1+1$ dimensional Minkowski spacetime can be drawn on paper, this yields a simple graphical representation of (M, g) which is often useful to illustrate certain arguments. In the following we illustrate this for the example of Minkowski spacetime.

A.2.1. The Penrose diagram of $\mathbb{R}^{1,1}$ itself. We start with $\mathbb{R}^{1,1}$ itself. Denoting the standard Cartesian coordinates on $\mathbb{R}^{1,1}$ by t, r the Minkowski metric takes the form $\eta = -\mathrm{d}t^2 + \mathrm{d}r^2$.

As first step, we transform to the null coordinates

$$v := t + r\,, \quad u := t - r\,. \tag{A.2.1}$$

Expressed in these coordinates, the Minkowski metric reads

$$\eta = -\frac{1}{2}(\mathrm{d}v \otimes \mathrm{d}u + \mathrm{d}u \otimes \mathrm{d}v) = -\frac{1}{2}(\mathrm{d}v\,\mathrm{d}u + \mathrm{d}u\,\mathrm{d}v)\,. \tag{A.2.2}$$

Now comes the decisive step. We further transform coordinates from u, v to new coordinates U, V which are defined by

$$v =: \tan V\,, \quad u =: \tan U\,. \tag{A.2.3}$$

The range of these new coordinates is the finite interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ for each. The interesting thing is that the Minkowski metric, expressed in the coordinates V, U, now reads

$$\eta = \frac{1}{\cos^2 U \cos^2 V}\left[-\frac{1}{2}(\mathrm{d}V\,\mathrm{d}U + \mathrm{d}U\,\mathrm{d}V)\right]. \tag{A.2.4}$$

We observe that the term inside the brackets has the same form as (A.2.2) again. Therefore we have actually constructed an explicit conformal embedding of $\mathbb{R}^{1,1}$ into a bounded subset of $\mathbb{R}^{1,1}$ itself. To see this even more clearly, we introduce the coordinates

$$T := \frac{V + U}{2}\,, \quad R := \frac{V - U}{2}\,, \tag{A.2.5}$$

in which the Minkowski metric now takes the form

$$\eta = \frac{1}{\cos^2(T+R)\cos^2(T-R)}\big[-\mathrm{d}T^2 + \mathrm{d}R^2\big]\,. \tag{A.2.6}$$

Therefore the map

$$\wp_{1,1} : \mathbb{R}^{1,1}_{t,r} \ni (t, r) \longmapsto (T, R) \in \mathbb{R}^{1,1}_{T,R}$$

is the mentioned conformal embedding.

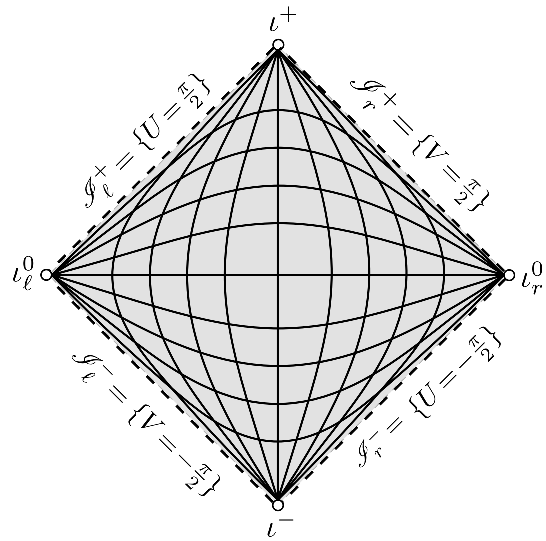

FIGURE A.1. The Penrose diagram of $\mathbb{R}^{1,1}$.

The image of this embedding (with its causal structure) is what is called the *Penrose diagram of* $\mathbb{R}^{1,1}$. It is sketched in figure A.1. The fact that the embedding is conformal implies in particular that the causal structure induced by the original metric $-\mathrm{d}t^2 + \mathrm{d}r^2$ is the same as the causal structure induced by the Minkowski metric $-\mathrm{d}T^2 + \mathrm{d}R^2$, which can be simply "read off" the illustration.

That the whole causal structure becomes so immediately visible in a bounded (i.e. completely drawable) domain is perhaps the key benefit of using Penrose diagrams.

Remark A.2.1. One can further embed $\mathbb{R}^{1,1}$ conformally into the *Einstein cylinder* $\mathbb{R}_t \times S^1_\theta$ with metric $-\mathrm{d}t^2 + \mathrm{d}\theta^2$ by taking the quotient of $\mathbb{R}^{1,1}_{T,R}$ under the equivalence relation $(T, R) \sim (T, R + 2\pi)$. This then identifies the points ι^0_ℓ and ι^0_r.

A.2.2. The Penrose diagram of $\mathbb{R}^{1,n}$. The first step is to introduce polar coordinates on the spatial part of $\mathbb{R}^{1,n}$, i.e. employ the usual diffeomorphism $\mathbb{R}^{1,n} \setminus \{r = 0\} \cong \mathbb{R}_t \times (0,\infty)_r \times S^{n-1}_\omega$ given by

$$\mathbb{R}_t \times (0,\infty)_r \times S^{n-1}_\omega \ni (t, r, \omega) \longmapsto (t, r\omega) \in \mathbb{R}_t \times \mathbb{R}^n_x \,.$$

As is well-known, the Minkowski metric then takes the form

$$\eta_{1,n} = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2 g_{S^{n-1}} \,.$$

Next one introduces the *radial null coordinates*

$$v = t + r\,, \qquad u = t - r\,.$$

As one easily verifies, Minkowski spacetime (modulo $r = 0$) then corresponds to the manifold $\{(v, u) \in \mathbb{R}^2 \mid u < v\} \times S^{n-1}_\omega$. Moreover, one easily checks that the Minkowski metric is now given as

$$\eta_{1,n} = -\frac{1}{2}\big(\mathrm{d}v\,\mathrm{d}u + \mathrm{d}u\,\mathrm{d}v\big) + \frac{1}{4}(v - u)^2 g_{S^{n-1}} \,.$$

Now, similar as in the previous section, we compactify the range of v and u by introducing yet another set of variables V, U defined by

$$v := \tan V\,, \qquad u := \tan U\,.$$

Expressed in these coordinates, Minkowski spacetime (module $r = 0$) corresponds to the manifold $\{(V, U) \in (-\frac{\pi}{2}, \frac{\pi}{2})^2 \mid U < V\} \times S^{n-1}_\omega$. Moreover, using some simple trigonometric identities, one also checks that the Minkowski metric is given by

$$\eta_{1,n} = \frac{1}{4\cos^2 V \cos^2 U}\Big[-2(\mathrm{d}V\,\mathrm{d}U + \mathrm{d}U\,\mathrm{d}V) + \sin^2(V-U) g_{S^{n-1}}\Big]\,.$$

In the final step we introduce the coordinates

$$T := V + U\,, \qquad \chi := V - U\,.$$

Another simple verification shows that Minkowski spacetime (module $r = 0$) then corresponds to the manifold $\{(T, \chi) \in (-\pi, \pi) \times (0, \pi) \mid \chi \leq \pi - |T|\} \times S^{n-1}_\omega$ and the Minkowski metric is given by

$$\eta_{1,n} = \frac{1}{\Omega^2(T,\chi)}\Big[-\mathrm{d}T^2 + \mathrm{d}\chi^2 + \sin^2\chi\, g_{S^{n-1}}\Big]\,, \tag{A.2.7}$$

where the conformal factor is given by $\Omega(T,\chi) = 2\cos\left(\frac{T+\chi}{2}\right)\cos\left(\frac{T-\chi}{2}\right)$.

To obtain the global picture of what is going on, one has to observe that the expression $\mathrm{d}\chi^2 + \sin^2\chi\, g_{S^{n-1}}$ is precisely the usual round metric on S^n expressed via the polar angle $\chi \in (0, \pi)$, i.e. where one writes $S^n \setminus \{\mathrm{NP}, \mathrm{SP}\} \cong (0, \pi) \times S^{n-1}$ via

$$(0, \pi) \times S^{n-1} \ni (\chi, \omega) \longmapsto (\sin\chi \cdot \omega, \cos\chi) \in S^n\,.$$

Therefore, by the previous calculations we have actually explicitly constructed an open conformal embedding

$$\wp_{1,n} : \mathbb{R}^{1,n} \longrightarrow \mathbb{E}^{1,n}\,, \tag{A.2.8}$$

where $\mathbb{E}^{1,n} = \mathbb{R}_T \times S^n$ denotes the *Einstein cylinder*, equipped with the Lorentzian metric $g_{\mathbb{E}^{1,n}} = -\,\mathrm{d}T^2 + g_{S^n}$.

Remark A.2.2. It is interesting to note that besides flat Minkowski spacetime also the other "Lorentzian model spacetimes" *de Sitter spacetime* and *(universal) anti de Sitter spacetime* possess open conformal embeddings into the Einstein cylinder, see for instance [**HE73**, Sec. 5.2]. Just as for Minkowski spacetime, these diagrams are a great way to visualize the conformal and causal structure of these spacetimes.

Bibliography

[ACN12] B. Ammann, C. Carvalho, and V. Nistor, *Regularity for eigenfunctions of Schrödinger operators*, Letters in Mathematical Physics **101** (2012), 49–84.

[Ali10] S. Alinhac, *Geometric analysis of hyperbolic differential equations*, London Mathematial Society Lecture Note Series, vol. 374, Cambridge University Press, Cambridge, 2010.

[ALN04] B. Ammann, R. Lauter, and V. Nistor, *On the geometry of Riemannian manifolds with a Lie structure at infinity*, International Journal of Mathematics and Mathematical Sciences **1–4** (2004), 161–193.

[ALN07] ———, *Pseudodifferential operators on manifolds with a Lie structure at infinity*, Annals of Mathematics **165** (2007), 717–747.

[Ati13] M. Atiyah, *What is a spinor?*, IHÈS, Conference en l'honneur de Jean-Pierre Bourguignon, video recording of conference talk, Available online at `https://www.youtube.com/watch?v=SBdW978Ii_E` (as of 29.10.2014), 2013.

[Bar86] R. Bartnik, *The mass of an asymptotically flat manifold*, Communications in Pure and Applied Mathematics **XXXIX** (1986), 661–693.

[Bär96] C. Bär, *The Dirac operator on space forms of positive curvature*, Journal of the Mathematical Society of Japan **48** (1996), no. 1, 69–83.

[Bär14] ———, *Green-hyperbolic operators on globally hyperbolic spacetimes*, Communications in Mathematical Physics (2014), 1–31.

[Bau81] H. Baum, *Spin-Strukturen und Dirac-Operatoren über pseudoriemannschen Mannigfaltigkeiten*, Texte zur Mathematik, vol. 41, Teubner, Leipzig, 1981.

[Bau09] ———, *Eichfeldtheorie*, Springer, Berlin, 2009.

[BEE96] J. K. Beem, P. E. Ehrlich, and K. L. Easley, *Global Lorentzian geometry*, 2 ed., Pure and Applied Mathematics, vol. 202, Marcel Dekker, New York, Basel, Hong Kong, 1996.

[BGM05] C. Bär, P. Gauduchon, and A. Moroianu, *Generalized cylinders in semi-Riemannian and spin geometry*, Mathematische Zeitschrift **249** (2005), 545–580.

[BGP07] C. Bär, N. Ginoux, and F. Pfäffle, *Wave equations on Lorentzian manifolds and quantization*, European Mathematical Society, Zürich, 2007.

[Ble81] D. Bleecker, *Gauge theory and variational principles*, Dover, Mineola, 1981.

[BS03] A. N. Bernal and M. Sánchez, *On smooth Cauchy hypersurfaces and Geroch's splitting theorem*, Communications in Mathematical Physics **243** (2003), 461–470.

[BS05] Antonio N. Bernal and M. Sánchez, *Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes*, Communications in Mathematical Physics **257** (2005), 43–50.

[BSZ90] J. C. Baez, I. E. Segal, and Z.-F. Zhou, *The global Goursat problem and scattering for nonlinear waves*, Journal of Functional Analysis **93** (1990), 239–269.

[BTW14] C. Bär and R. Tagne Wafo, *Initial value problems for wave equations on manifolds*, arXiv:1408.4995 [math.AP] **preprint** (2014), 1–25.

[BVW12] D. Baskin, A. Vasy, and J. Wunsch, *Asymptotics of radiation fields in asymptotically Minkowski space*, arXiv:1212.5141 [math.AP] **preprint** (2012), 1–67.

[BW10] R. Beals and R. Wong, *Special functions*, Cambridge University Press, Cambridge, 2010.

[Car13] E. Cartan, *Les groupes projectives qui ne laissent invariante aucun multiplicite plane*, Bulletin de la Societe Mathematique de France **41** (1913), 53–96.

[Car66] ———, *The theory of spinors*, Hermann, Paris, 1966.

[CB09] Y. Choquet-Bruhat, *General relativity and Einstein's equations*, Oxford Mathematical Monographs, Oxford University Press, Oxford, 2009.

[CD02] P. T. Chruściel and E. Delay, *Existence of non trivial, asymptotically vacuum, asymptotically simple space-times*, Classical and Quantum Gravity **19** (2002), L71–L79, erratum Classical and Quantum Gravity **19** (2002), 3389.

[CGP10] P. T. Chruściel, G. J. Galloway, and D. Pollack, *Mathematical general relativity: A sampler*, Bulletin of the American Mathematical Society **47** (2010), 567–638.

[CK93] D. Christodoulou and S. Klainerman, *The global nonlinear stability of Minkowski space*, Princeton University Press, Princeton, 1993.

[CMP03] E. Caponio, A. Masiello, and P. Piccione, *Some global properties of static spacetimes*, Mathematische Zeitschrift **244** (2003), no. 3, 457–468.

[DF99] P. Deligne and D. Freed, *Classical field theory*, Quantum Fields and Strings: A Course for Mathematicians (Providence) (Pierre Deligne et al., eds.), American Mathematical Society, 1999.

[DG97] J. Dereziński and C. Gérard, *Scattering theory of classical and quantum N-particle systems*, Springer, Heidelberg, 1997.

[Dir28] P. A. M. Dirac, *The quantum theory of the electron*, Proceedings of the Royal Society of London A **117** (1928), no. 778, 610–624.

[DMP06] C. Dappiaggi, V. Moretti, and N. Pinamonti, *Rigorous steps towards holography in asymptotically flat spacetimes*, Reviews in Mathematical Physics **18** (2006), 349–416.

[E$^+$11] P. Etingof et al., *Introduction to representation theory*, American Mathematical Society, Providence, 2011.

[Eva10] L. C. Evans, *Partial differential equations*, 2 ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, 2010.

[Fin98] F. Finster, *Local* $U(2,2)$ *symmetry in relativistic quantum mechanics*, Journal of Mathematical Physics **39** (1998), 6276–6290.

[Fin02] ———, *Partielle Differentialgleichungen aus der Geometrie und Physik II - Hyperbolische Probleme*, Lecture Notes, Max-Planck-Institut für Mathematik in den Naturwissenschaften, available online at `http://www.mis.mpg.de/de/publications/andere-reihen/ln/lecturenote-1502.html`, 2002.

[Fin06] ———, *The principle of the fermionic projector*, American Mathematical Society/IP, Providence, 2006.

[FJ98] G. Friedlander and M. Joshi, *Introduction to the theory of distributions*, 2 ed., Cambridge University Press, Cambridge, 1998.

[FKSY00] F. Finster, N. Kamran, J. Smoller, and S.-T. Yau, *Nonexistence of time-periodic solutions of the Dirac equation in an axisymmetric black hole geometry*, Communications on Pure and Applied Mathematics **53** (2000), 902–929.

[FKSY02] ———, *Decay rates and probability estimates for massive Dirac particles in the Kerr-Newman black hole geometry*, Communications in Mathematical Physics **230** (2002), 201–244.

[FKSY03] ———, *The long-time dynamics of Dirac particles in the Kerr-Newman black hole geometry*, Advances in Theoretical and Mathematical Physics **7** (2003), 25–52.

[FKSY09] ———, *Linear waves in the kerr geometry: A mathematical voyage to black hole physics*, Bulletin of the American Mathematical Society **46** (2009), 635–659.

[FR04] M. Forger and H. Römer, *Current and the energy-momentum tensor in classical field theory: A fresh look at an old problem*, Annals of Physics **309** (2004), 306–389.

[Fri00] T. Friedrich, *Dirac Operators in Riemannian Geometry*, Graduate Studies in Mathematics, vol. 25, American Mathematical Society, Providence, 2000.

[GMS97] G. Giachetta, L. Mangiarotti, and G. Sardanashvily, *New Lagrangian and Hamiltonian methods in field theory*, World Scientific, Singapore, 1997.

[Gre78] W. Greub, *Multilinear algebra*, 2 ed., Springer-Verlag, New York, Berlin, Heidelberg, 1978.

[GW14] A. R. Gover and A. Waldron, *Boundary calculus for conformally compact manifolds*, Indiana University Mathematics Journal **63** (2014), 119–163.

[Har90] F. Reese Harvey, *Spinors and calibrations*, Perspectives in Mathematics, vol. 9, Academic Press Inc., San Diego, 1990.

[HE73] S.W. Hawking and G.F.R Ellis, *The large scale structure of space-time*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 1973.

[Heb99] E. Hebey, *Nonlinear analysis on manifolds*, Courant Lecture Notes, vol. 5, American Mathematical Society, Providence, 1999.

[Hin13] P. Hintz, *Global well-posedness of quasilinear wave equations on asymptotically de Sitter spaces*, arXiv:1311.6859 [math.AP] **preprint** (2013), 1–87.

[Hör97] L. Hörmander, *Lectures on nonlinear hyperbolic equations*, Mathematics and Aplications, vol. 26, Springer, New York, 1997.

[HS96] P. D. Hislop and I. M. Sigal, *Introduction to spectral theory: with applications to Schrödinger operators*, Applied Mathematical Sciences, vol. 113, Springer, New York, 1996.

[HT94] S. A. Huggett and K. P. Tod, *An introduction to twistor theory*, 2 ed., London Mathematical Society Student Texts, vol. 4, Cambridge University Press, Cambridge, 1994.

[Hus94] D. Husemoller, *Fibre bundles*, 3 ed., Graduate Texts in Mathematics, vol. 20, Springer, New York, 1994.

[HV13] P. Hintz and A. Vasy, *Semilinear wave equations on asymptotically de Sitter, Kerr-de Sitter and Minkowski spacetimes*, arXiv:1306.4705 [math.AP] **preprint** (2013), 1–78.

[HV14] ———, *Global analysis of quasilinear wave equations on asymptotically Kerr-de Sitter spaces*, arXiv:1404.1348 [math.AP] **preprint** (2014), 1–45.

[Joh82] F. John, *Partial differential equations*, 4 ed., Applied Mathematical Sciences, vol. 1, Springer, New York, 1982.

[Jou10] J. Joudioux, *Problème de Cauchy caractéristique et scattering conforme en relativité générale*, Ph.D. thesis, Université de Bretagne Occidentale, Brest, 2010, Available online at `https://tel.archives-ouvertes.fr/tel-00517339`.

[Jou12] ———, *Conformal scattering for a nonlinear wave equation*, Journal of Hyperbolic Differential Equations **9** (2012), no. 1, 1–65.

[JS08] M. A. Javaloyes and M. Sánchez, *A note on the existence of standard splittings for conformally stationary spacetimes*, Classical and Quantum Gravity **25** (2008), no. 16, 168001 (7pp).

[Kic96] S. Kichenassamy, *Nonlinear wave equations*, Monographs and textbooks in pure and applied mathematics, vol. 194, Marcel Decker, New York, 1996.

[KT03] M. Kraus and C. Tretter, *A new method for eigenvalue estimates for Dirac operators on certain manifolds with S^k-symmetry*, Differential Geometry and its Applications **19** (2003), 1–14.

[Lax06] P. D. Lax, *Hyperbolic partial differential equations*, American Mathematical Society, Providence, 2006.

[Lee03] J. M. Lee, *Introduction to smooth manifolds*, Graduate Texts in Mathematics, vol. 218, Springer-Verlag, Berlin, Heidelberg, New York, 2003.

[LM89] H. B. Lawson and M. L. Michelsohn, *Spin geometry*, Princeton University Press, Princeton, 1989.

[Mas11a] M. P. Masarik, *The wave equation in a general spherically symmetric black hole geometry*, Advances in Theoretical and Mathematical Physics **15** (2011), no. 6, 1789–1815.

[Mas11b] ———, *The wave equation in a general spherically symmetric particlelike geometry*, arXiv:1106.4466 [gr-qc] **preprint** (2011), 1–20.

[Mas12] ———, *Decay of solutions to the wave equation in static spherically symmetric spacetimes*, Ph.D. thesis, University of Michigan, Ann Arbor, 2012, Available online at `http://deepblue.lib.umich.edu/handle/2027.42/91416`.

[Mel95] R. B. Melrose, *Geometric scattering theory*, Cambridge University Press, Cambridge, 1995.

[MN04] L. J. Mason and J.-P. Nicolas, *Conformal scattering and the Goursat problem*, Journal of Hyperbolic Differential Equations **1** (2004), no. 2, 197–233.

[MN09] ———, *Regularity at space-like and null infinity*, Journal of the Institute of Mathematics of Jussieu **8** (2009), no. 1, 179–208.

[MN12] ———, *Peeling of Maxwell and Dirac fields on a Schwarzschild background*, Journal of Geometry and Physics **62** (2012), no. 4, 876–889.

[MS74] J. Milnor and J. Stasheff, *Characteristic classes*, Princeton University Press, Princeton, 1974.

[MS08] E. Minguzzi and M. Sanchez, *The causal hierarchy of spacetimes*, Recent Developments in Pseudo-Riemannian Geometry (Zürich) (H. Baum and D. Alekseevsk, eds.), European Mathematical Society Publishing House, 2008, pp. 299–358.

[MS13] O. Müller and M. Sánchez, *An invitation to Lorentzian geometry*, Jahresberichte der Deutschen Mathematiker-Vereinigung **115** (2013), no. 3–4, 153–183.

[MSBV13] R. Melrose, A. Sá Barreto, and A. Vasy, *Asymptotics of solutions of the wave equation on de Sitter-Schwarzschild space*, Communications in Partial Differential Equations **39** (2013), 512–529.

[Mül14] O. Müller, *Geometry of globally hyperbolic manifolds*, Habilitation, Universität Regensburg, 2014.

[MVW08] R. Melrose, A. Vasy, and J. Wunsch, *Propagation of singularities for the wave equation on edge manifolds*, Duke Mathematical Journal **144** (2008), no. 1, 109–193.

[Nic95] J.-P. Nicolas, *Scattering of linear Dirac fields by a spherically symmetric black hole*, Ann. Inst. Henri Poincaré - Physique Théoretique **61** (1995), no. 2, 145–179.

[Nic02] ———, *Dirac fields on asymptotically flat space-times*, Dissertationes Math. **408** (2002), 1–85.

[Nic13] ———, *Conformal scattering on the Schwarzschild metric*, arXiv:1312.1386 [gr-qc] **preprint** (2013), 1–33.

[O$^+$10] F. W. Olver et al. (eds.), *NIST handbook of mathematical functions*, Cambridge, Cambridge University Press, 2010.

[O'N83] B. O'Neill, *Semi-Riemannian geometry*, Pure and Applied Mathematics, vol. 103, Academic Press, New York, 1983.

[Pen65] R. Penrose, *Zero rest-mass fields including gravitation: asymptotic behaviour*, Proceedings of the Royal Society of London A **284** (1965), no. 1387, 159–203.

[Pet06] P. Petersen, *Riemannian geometry*, 2 ed., Graduate Texts in Mathematics, vol. 171, Springer, New York, 2006.

[PR86] R. Penrose and W. Rindler, *Spinors and space-time 1&2*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 1986.

[Ren08] A. D. Rendall, *Partial differential equations in general relativity*, Oxford Graduate Texts in Mathematics, vol. 16, Oxford University Press, Oxford, 2008.

[Rin09] H. Ringström, *The Cauchy problem in general relativity*, EMS Publishing House, Zürich, 2009.

[RS79] M. Reed and B. Simon, *Methods of modern mathematical physics 3: Scattering theory*, Academic Press, San Diego, 1979.

[Ś05] M. Sánchez, *On the geometry of static spacetimes*, Nonlinear Analysis: Theory, Methods & Applications **63** (2005), no. 5, e455–e463.

[Sak67] J. J. Sakurai, *Advanced quantum mechanics*, Addison-Wesley Publishing Company, Reading, 1967.

[Sak94] ———, *Modern quantum mechanics*, revised ed., Addison-Wesley Publishing Company, Reading, 1994.

[Sau06] F. Sauvigny, *Partial differential equations 1*, Springer, Berlin, 2006.

[Sch95] M. Schottenloher, *Geometrie und Symmetrie in der Physik*, Vieweg, Braunschweig, 1995.

[Shu92] M. A. Shubin, *Spectral theory of elliptic operators on non-compact manifolds*, Astérisque **207** (1992), 35–108.

[Sog08] C. D. Sogge, *Lectures on non-linear wave equations*, International Press, Boston, 2008.

[Spa66] E. H. Spanier, *Algebraic topology*, Springer, New York, 1966.

[SS07] M. Sánchez and José M. M. Senovilla, *A note on the uniqueness of global static decompositions*, Classical and Quantum Gravity **24** (2007), no. 23, 6121–6126.

[Str04] N. Straumann, *General Relativity*, Springer, Berlin, 2004.

[SW93] J. A. Smoller and A. G. Wasserman, *Existence of infinitely smooth, static, global solutions of the Einstein/Yang-Mills equations*, Communications in Mathematical Physics **151** (1993), no. 2, 303–325.

[SWY93] J. A. Smoller, A. G. Wasserman, and S.-T. Yau, *Existence of black hole solutions of the Einstein-Yang/Mills equations*, Communications in Mathematical Physics **154** (1993), no. 2, 377–401.

[Tao06] T. Tao, *Nonlinear dispersive equations*, Regional Conference Series in Mathematics, vol. 106, American Mathematical Society, Providence, 2006.

[Tay11] M. Taylor, *Partial differential equations III*, 2 ed., Springer, New York, 2011.

[Tha92] B. Thaller, *The dirac equation*, Texts and Monographs in Physics, Springer, New York, 1992.

[Vas08] A. Vasy, *Propagation of singularities for the wave equation on manifolds with corners*, Annals of Mathematics **168** (2008), 749–812.

[Vas10] ———, *The wave equation on asymptotically de Sitter-like spaces*, Advances in Mathematics **223** (2010), 49–97.

[Vas12] ———, *The wave equation on asymptotically anti de Sitter spaces*, Analysis and PDE **5** (2012), 81–144.

[Vas13] ———, *Microlocal analysis of asymptotically hyperbolic and Kerr-de Sitter spaces (with an appendix by Semyon Dyatlov)*, Inventiones Mathematicae **194** (2013), 381–513.

[Wal84] R. M. Wald, *General Relativity*, University of Chicago Press, Chicago, 1984.

[Wal12] S. Waldmann, *Geometric wave equations*, Lecture Notes, Department Mathematik, Friedrich-Alexander Universität Erlangen-Nürnberg, Available at `http://arxiv.org/abs/1208.4706`, 2012.

[Wei05] S. Weinberg, *The quantum theory of fields*, Cambridge University Press, Cambridge, 2005.

[Yaf92] D. R. Yafaev, *Mathematical scattering theory: General theory*, American Mathematical Society, Providence, 1992.

[Yaf10] ———, *Mathematical scattering theory: Analytic theory*, American Mathematical Society, Providence, 2010.

[Zwo12] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics, vol. 138, American Mathematical Society, Providence, 2012.

Jan-Hendrik Treude
Basic Research Community