Conformal Scattering and Analysis of Asymptotic Properties of Gauge Theories in General Relativity

Grigalius Taujanskas
Keble College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2019

Acknowledgements

First and foremost, I would like to thank my supervisor Lionel Mason, whose guidance and knowledge throughout the last three years have been invaluable and from whom I have learnt a lot about mathematics, and mathematicians too. His ideas—as well as his generous attitude towards my intellectual independence—have much to do with the shape of this thesis. I would also like to thank Jean-Philippe Nicolas, who has been a great mentor, collaborator, friend, and essentially an unofficial second supervisor. Where I have struggled to understand, he has taken precious time to explain. I would also like to thank all of my colleagues at Oxford—too many to name—but particularly Qian Wang, Paul Tod and Jan Sbierski, for listening to my rambling seminars and coming back with constructive feedback. I also thank them for sharing with me their view of the land of mathematical research, which has been instrumental in shaping me as a scientist. Outside of Oxford, I thank Mihalis Dafermos for his time, patience, and hikes in remote parts of Japan.

I would also like to express my gratitude to the Engineering and Physical Sciences Research Council of the United Kingdom for financial support[1]. Additionally, I thank the Mittag-Leffler institute in Sweden, where the final stages of this thesis were completed.

Lastly, I thank my parents Jūratė and Astijus, who have since childhood allowed and encouraged me to pursue the arcane art of mathematics, only questioning why anyone would want to do so every so often. And Amelia Hassoun, for always sharing a nectarine with me, even though they are clearly the highest form of currency in this strange world of humans.

[1]Grant number [EP/L05811/1].

Abstract

The study of scattering is a fundamental aspect of mathematical physics. In the second half of the 20th century, mathematical tools were developed by Penrose, Friedlander, Lax, Phillips, and others that have revealed an intriguing underlying geometry of scattering theories.

In the first part of this thesis we study the conformal scattering of Maxwell potentials on a class of asymptotically flat asymptotically simple spacetimes. We construct scattering operators as isomorphisms between Hilbert spaces on past and future null infinity, and (in the flat case) explore the structure of these Hilbert spaces with respect to the symmetries of the spacetime.

In the second part we study the Maxwell-scalar field system on de Sitter space. We prove small data peeling estimates at all orders, and construct bounded and invertible, but nonlinear, scattering operators. We discover that sufficiently regular solutions decay exponentially in time and disperse as linear waves, and find a curious asymptotic decoupling of the scalar field.

In the third part of this thesis we study the Yang–Mills–Higgs equations on the Einstein cylinder. By localizing Eardley & Moncrief's famous L^∞ estimates, we extend them to the Einstein cylinder, and remove a small data restriction in a classical theorem of Choquet-Bruhat and Christodoulou from 1981. By using conformal transformations, we deduce large data decay rates for Yang–Mills–Higgs fields on Minkowski and de Sitter spacetimes.

Declaration

This thesis is based on work done while a graduate student at the Mathematical Institute, Oxford University, and the Mittag-Leffler Institute, Sweden, in the period between September 2016 and October 2019.

All of the work presented in this thesis is the sole work of the author. At the time of submission, chapter 3 is unpublished in any form. Chapter 4 is based on the paper

> G. Taujanskas, *Conformal scattering of the Maxwell-scalar field system on de Sitter space*, J. Hyperbolic Differ. Equ. **16** (04), 743-791 (2019).

Chapter 5 is based on the paper

> G. Taujanskas, *Large data decay of Yang–Mills–Higgs fields on Minkowski and de Sitter spacetimes*, J. Math. Phys. **60**, 121504 (2019).

Contents

Nathan Rosen and I arrived at the interesting result that gravitational waves do not exist.

— Albert Einstein, [63]

1 Introduction

1.1 General Relativity and Asymptotic Structure of Spacetime

In 1915 Albert Einstein wrote down the equations of general relativity (GR) and proposed three tests of his theory: the precession of the perihelion of Mercury, gravitational lensing of light by the Sun, and gravitational redshift. The amount the orbit of Mercury precesses had been known to be approximately 43 arcseconds per century more than predicted by Newtonian theory, with many ad-hoc fixes failing to consistently account for the discrepancy. Einstein's calculations showed that in the framework of his new theory, the precession of the perihelion of Mercury should be corrected by 42.98" per century due to the effects of spacetime curvature[1]. Gravitational lensing effects in line with the predictions of GR were confirmed in 1919, whereas sensitive gravitational redshift effects were measured by 1954. This started a programme of increasingly stringent tests of various predictions of GR, including Shapiro's measurement of relativistic time delay, the Eötvös experiment, and tests of local Lorentz invariance, all of them confirming the predictions of GR to a stunning degree of accuracy [124].

This programme culminated on 14 September 2015, almost exactly a hundred years after the birth of GR, in a remarkable observation[2] by the gravitational wave detector LIGO in Livingston, WA. Efforts to detect gravitational waves—dubbed the last remaining test of general relativity—had been ongoing for more than half a century, their effects being so minuscule that Einstein himself had doubted that they could ever be observed. A cataclysmic event approximately a light-year away, lasting a mere 0.2 seconds, of two black holes inspiralling and merging had been observed by LIGO. With masses approximately 36 and 29 times the mass of the Sun, the two black holes accelerated from about 30% the speed of light to about 60% the speed of light across the 0.2-second

[1] Einstein's theory also correctly accounted for a smaller discrepancy of 8.6" per century in the precession of the perihelion of Venus.

[2] Codename GW150914.

duration of the detectable signal, and merged into a Kerr[3] black hole of 62 solar masses. The enormous impact of the merger released the 3 missing solar masses of energy, which was radiated away in the form of gravitational waves. During the final 20 milliseconds of the event, the power of the radiated gravitational waves peaked at about 3.6×10^{49} watts—approximately 50 times greater than the combined power of the light radiated by all the stars in the observable universe [1, 2]. Yet—despite the magnitude of the astrophysical event leading to GW150914—the resulting ripple in spacetime reaching the Earth changed the length of a 4-kilometre LIGO arm by just a thousandth of the width of a proton. Still, the instruments of LIGO observed a characteristic waveform of the merger which was satisfyingly consistent with large-scale numerical relativity simulations, thereby laying general relativity's last test to rest [1].

The discovery of gravitational waves was heralded as the beginning of a new era of gravitational wave astronomy [120, 121]. And although gravitational waves were predicted by the theory of general relativity back in 1916, there is much that is not yet well-understood mathematically about GR, as Einstein's equations

$$\mathrm{R}_{ab} - \frac{1}{2}\mathrm{R} g_{ab} + \lambda g_{ab} = -8\pi\gamma \mathbf{T}_{ab} \tag{1.1.1}$$

have proven extraordinarily difficult to study. Only a handful of exact solutions of Einstein's equations are known to date, most others being constructed from small perturbations of known initial data sets [24, 26, 27, 29, 30]. The reason understanding GR is difficult is that, unlike other physical theories, GR does not have a well-defined dynamical variable evolving on a fixed background. In Newtonian theory, for example, the background is Euclidean space and the dynamical variable describing the physics is the Newtonian potential. In GR, on the other hand, the spacetime metric is both; the evolution of the metric *and* the spacetime on which the metric lives must therefore be constructed at the same time. Indeed, in appropriate coordinates[4] the equations (1.1.1) in vacuum become

$$\Box_g g_{ab} = Q_{ab}(g)(\partial g, \partial g). \tag{1.1.2}$$

Without a non-dynamical reference point, extracting physical meaning from dynamical fields is much more difficult [39]. What is more, the notion of an isolated system (a system that at some rate approaches an 'empty' solution at infinity), fairly trivial in other theories, becomes much more blurred in GR without a non-dynamical reference point.

The key idea underpinning most of the work in this thesis, that in one stroke takes care of both of these difficulties, is due to Penrose [93, 94]. Since 'infinity is metrically far away', consider the replacement of the physical metric g_{ab} with a new, *rescaled* metric $\hat{g}_{ab} = \Omega^2 g_{ab}$. Such a transformation is called a *conformal transformation*, since it preserves angles[5] between intersecting curves. The purpose of the function Ω—the conformal factor—is to 'bring infinity to a finite region' by approaching zero asymptotically at an appropriate rate. For a class of spacetimes called *asymptotically simple*, this procedure

[3] The post-merger object was found to be a rotating sub-extremal Kerr black hole with the angular momentum parameter $|a| \approx 2M/3$, for M the mass of the resulting black hole [119].

[4] The wave coordinates x^a satisfying $\Box_g x^a = 0$. These are the Lorentzian analogue of harmonic coordinates in Euclidean space. The nonlinearity Q_{ab} appearing on the right-hand side of (1.1.2) is quadratic in ∂g.

[5] As a result, a conformal transformation preserves causal structure: a timelike (null, spacelike) curve with respect to g_{ab} remains timelike (null, spacelike) with respect to $\hat{g}_{ab}$.

turns out to be possible and well-behaved. In these cases one can attach to the original spacetime $\mathscr{M}$ a boundary $\mathscr{I} := \{\Omega = 0\}$—called *null infinity*—consisting of the set of points which are infinitely distant with respect to the physical metric g_{ab}, and define a new spacetime $\hat{\mathscr{M}} := \mathscr{M} \cup \mathscr{I}$. With respect to the rescaled metric $\hat{g}_{ab}$, one requires that the attached points $\mathscr{I}$ are appropriately smooth, which results in a definition which at the same time avoids any consideration of separate components of the metric or its curvature tensors, and replaces asymptotic considerations in $\mathscr{M}$ with local differential geometry near $\mathscr{I}$ in the rescaled spacetime $\hat{\mathscr{M}}$.

Physical processes occurring in $\mathscr{M}$ imbue the boundary $\mathscr{I}$ with various forms of radiation, at which point one can detach $\mathscr{I}$ and make it into an abstract manifold, stripping away everything non-asymptotic. Various fields in $\mathscr{M}$ are found to yield corresponding fields on this abstract manifold $\mathscr{I}$, which provides a sort of summary of all the available asymptotic information in $\mathscr{M}$. One finds that the fields so defined on $\mathscr{I}$ split, on the grounds of being mathematically distinct, into two classes. The first class is that of universal (or geometrical) fields, which are at least locally the same no matter what the original physical spacetime $\mathscr{M}$. The second class is that of physical fields, which are not. That as soon as such a distinction becomes available one adopts it is unsurprising, as many of the difficulties in GR arise from the metric being both a geometrical and a physical object. On $\mathscr{I}$—in the asymptotic limit—then, GR begins to look like other physical theories, and one is able to separate the geometry from the physics. In the case of zero cosmological constant λ, if the metric g_{ab} is asymptotically flat, the universal fields give rise to the BMS group $\mathfrak{B}$, the asymptotic symmetry group of the spacetime, which is the group of all diffeomorphisms of $\mathscr{I}$ which leave the universal fields invariant. As one might expect, $\mathfrak{B}$ is similar to the Poincaré group, the group of symmetries of Minkowski space consisting of the Lorentz group $\mathcal{L}$ and spacetime translations, but turns out to not quite be identical to the Poincaré group. The group of translations becomes enlarged in the asymptotic limit, and is replaced by an infinite dimensional group of so-called *supertranslations* $\mathcal{S}$ of $\mathscr{I}$. The BMS group then takes the form $\mathfrak{B} = \mathcal{S} \rtimes \mathcal{L}$ [5, 39]. In the case of positive cosmological constant λ, for asymptotically de Sitter spacetimes, it turns out that the largest possible symmetry group of $\mathscr{I}$ is simply the de Sitter group $\mathrm{SO}(4,1)$ [3]. In this case, however, $\mathscr{I}$ may be curved and have no symmetry group at all.

By adopting Penrose's conformal framework, in this thesis we study the asymptotic behaviour of fields in both the asymptotically flat and asymptotically de Sitter regimes.

1.2 Scattering

1.2.1 Scattering in Quantum Field Theory

Relativistic quantum field theory (QFT) is a tremendously successful theoretical framework of the mechanics of elementary particles. Boasting impressive agreement with experiments, its crowning achievement is the Standard Model of Particle Physics, fundamentally an $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$ gauge theory, which includes leptons, the photon, the $W^{\pm}$ and Z^0 force bosons, quarks and gluons, and the most recently discovered Higgs boson. The Standard Model for short, it was discovered through scattering experiments: the practice of throwing building blocks of nature at one another and observing how they break apart. The observables which encode the likelihood of a given scattering process taking place are the so-called *scattering amplitudes*. To a quantum field theorist, these scattering amplitudes come as terms in a perturbation expansion of a functional called an *S-matrix* [38, 113].

Despite the immense experimental success of the Standard Model, however, the perturbation expansion of a quantum S-matrix famously does not exist due to both infrared (low energy) and ultraviolet (high energy) divergences. It has been known since Dyson's argument back in 1952 that large oscillations are expected to arise in terms of high order (typically beyond order $n = \alpha^{-1} \approx 137$)[6] and overwhelm the successful lowest order terms on which most conventional wisdom of quantum field theory rests [35]. In general, therefore, the perturbation expansion is mathematically unreliable, although it should be stressed that the theoretical predictions of perturbative quantum electrodynamics (QED) are in extraordinary[7] agreement with experiments. The curious conflict between the predictive success of relativistic quantum field theories and the divergence of the (renormalized) perturbation expansions of their S-matrices require a non-perturbative approach to QFT. Unfortunately, to date no non-trivial non-perturbative quantum field theory is known in four spacetime dimensions [113].

Nonetheless, one may try to take a classical approach. The terms in the perturbation expansions of S-matrices are computable using Feynman diagrams, which can be divided into *tree-level* and *loop-level* diagrams. At the tree-level, an expansion in terms of Feynman diagrams corresponds to classical scattering of the underlying fields. On top of this one has loop-level diagrams, which encode quantum corrections to classical scattering and become important at higher energies, but which may reasonably be ignored at lower energies. Non-perturbative classical scattering therefore corresponds to tree-level scattering in QFT, and can shed light on non-perturbative aspects of the quantum S-matrix.

1.2.2 Infrared Divergences and the Memory Effect

At the classical level, the infamous infrared divergences [72, 86] in interacting quantum field theories correspond to the so-called memory effect [17, 105]. In the case of gravity, the memory effect is the permanent relative displacement of an array of test particles that results from the passage of a burst of gravitational radiation (gravitational waves). To make this precise, one integrates the geodesic deviation equation, and defines the *memory tensor* as the map taking the initial deviation of the test particles (at $t = -\infty$) to their final deviation (at $t = +\infty$). It turns out that the memory tensor decomposes into two parts, called ordinary memory and null memory. Further, it turns out that null memory comes from a supertranslation of null infinity—an asymptotic symmetry of the spacetime. The memory tensor then enters the charge-flux conservation law from past null infinity to future null infinity associated with this supertranslation. On the quantum side, what this means is that, in a scattering theory from a Hilbert space of in-states $\mathcal{H}_{\text{in}}$ to a Hilbert space of out-states $\mathcal{H}_{\text{out}}$, the memory tensor introduces a ω^{-1} term at low frequencies ω in the Fourier transform of the field as it evolves from past null infinity to future null infinity, which of course diverges at $\omega = 0$. The effect is

[6] Here α is the fine structure constant, $\alpha = (4\pi\varepsilon_0\hbar c)^{-1}e^2$. It is directly related to the coupling constant determining the strength of the interaction between electrons and photons.

[7] Up to the 11th significant figure, to be precise, based on the current state-of-the-art calculations of QED diagrams up to four loops [49, 50]. Other quantum field theories, like QCD (quantum chromodynamics), have large coupling constants, meaning that more terms in the perturbative expansion need to computed to achieve a comparable level of precision.

also present in electromagnetism: one can replace the metric tensor with the Maxwell potential and study a corresponding *electromagnetic* memory effect [17]. In this case, the memory effect is not a displacement of test particles, but rather a momentum kick (a residual velocity) to charges in a detector after an electromagnetic wave has passed through it. If Maxwell's equations have a non-zero right-hand side, a similar divergent low frequency term shows up in the electromagnetic case.

A complete understanding of how these infrared divergences are related to classical aspects of field theory calls for a thorough study of classical scattering theories from past null infinity to future null infinity. This is in large part the motivation for the work presented in this thesis, however these issues remain to be understood.

1.2.3 Classical Scattering

Perhaps the most famous treatment of classical scattering on flat space was developed by Lax and Phillips in the 1960s [73, 74]. Much like one imagines quantum field theoretic scattering comparing the incoming particles before the interaction with the outgoing particles afterwards, the scattering theory of Lax and Phillips was developed to compare the asymptotic behaviour of an evolving system in the far past with its asymptotic behaviour in the far future, after having been scattered by a disturbance.

The essentials of the Lax–Phillips scattering theory are most easily understood by considering the free scalar wave equation on flat space [91],

$$\partial_t^2\phi - \Delta\phi = 0, \tag{1.2.1}$$

for which one expects no scattering to occur as there is no disturbance. By rewriting it in Hamiltonian form as $\partial_t\Phi = -iH\Phi$, where

$$\Phi = \begin{pmatrix} \phi \\ \partial_t\phi \end{pmatrix} \quad \text{and} \quad H = -i\begin{pmatrix} 0 & 1 \\ \Delta & 0 \end{pmatrix},$$

one associates to the equation (1.2.1) the operator H, which is self-adjoint on the space $\dot{H}^1(\mathbb{R}^3)\oplus L^2(\mathbb{R}^3)$. The spectrum of H is the whole real line $\mathbb{R}$, and for each eigenvalue $\sigma \in \mathbb{R}$ the operator H has a 2-sphere[8] of plane wave eigenfunctions indexed by $(\sigma, \omega) \in \mathbb{R} \times \mathbb{S}^2$. By taking the $\dot{H}^1 \oplus L^2$ inner product with these eigenfunctions, one obtains a map

$$\Phi(t, x) \longmapsto \tilde{\Phi}(\sigma, \omega)$$

that is an isometry from $\dot{H}^1(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)$ to $L^2(\mathbb{R} \times \mathbb{S}^2)$. This map provides a spectral representation of H and its propagator, $H\Phi \mapsto \sigma\tilde{\Phi}$ and $\mathrm{e}^{itH}\Phi \mapsto \mathrm{e}^{it\sigma}\tilde{\Phi}$. Taking the Fourier transform of $\tilde{\Phi}$ in σ, one obtains a new representation $\mathfrak{T}^+\Phi$ of Φ, which satisfies

$$\mathfrak{T}^+(\mathrm{e}^{itH}\Phi)(r, \omega) = (\mathfrak{T}^+\Phi)(r - t, \omega).$$

This is the famous Lax–Phillips *translation representation*, which is an isometry from $\dot{H}^1(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)$ to $L^2(\mathbb{R} \times \mathbb{S}^2)$. The existence of a translation representation for the equation (1.2.1) can in a sense be interpreted as a no scattering condition: the evolution of the scalar field is described simply as a translation of the initial data.

[8] Except for $\sigma = 0$, in which case the 2-sphere collapses to a point and there is only one eigenfunction $(1, 0)$. In general they have the form $(\mathrm{e}^{-i\sigma x\cdot\omega}, i\sigma\mathrm{e}^{-i\sigma x\cdot\omega})$.

The Lax–Phillips scattering theory goes further, expressing $\mathfrak{T}^+\Phi$ in terms of Radon transforms of the initial data for Φ, and, using the inverse Radon transform, constructing the inverse of $\mathfrak{T}^+$. Performing a similar construction for $-H$ instead of H, Lax and Phillips construct the past translation representation $\mathfrak{T}^-$ and its inverse ($\mathfrak{T}^+$ being the *future* translation representation), and define the *scattering operator* as

$$\mathscr{S} := \mathfrak{T}^+ \circ (\mathfrak{T}^-)^{-1}.$$

A key element of this construction is the asymptotic profile property,

$$(\mathfrak{T}^+\Phi)(s,\omega) = -\lim_{r\to+\infty} r(\partial_t\phi)(r,(r+s)\omega), \tag{1.2.2}$$

which, as we shall see shortly, essentially reinterprets the translation representation as Friedlander's radiation field [45]. Moreover, the explicit inversion of $\mathfrak{T}^+$ provides a formula for the solution ϕ in terms of its translation representation via

$$\phi(t,x) = \frac{1}{2\pi}\int_{\mathbb{S}^2} (\mathfrak{T}^+\Phi)(x\cdot\omega + t,\omega)\,\mathrm{dv}_{\mathfrak{s}_2}, \tag{1.2.3}$$

which, as it turns out, is exactly Whittaker's formula for solutions to (1.2.1) [123]. The theory of Lax–Phillips constructs a complete scattering theory in the sense that the solution to (1.2.1) is completely characterized by its asymptotic profiles via (1.2.3). In other words, the scattering operator $\mathscr{S}$ is invertible. The reliance on spectral techniques, however, requires that the background geometry of a Lax–Phillips scattering theory be flat and static.

In 1980 Friedlander explained the underpinning geometry of the Lax–Phillips approach in terms of Penrose's null infinity [45, 94], and constructed the first genuinely *conformal* scattering theory for a free (conformal) wave equation on a static asymptotically flat spacetime with sufficiently fast decay at infinity—fast enough to allow a smooth conformal compactification. Friedlander's construction relied on his previous work on the wave equation in curved space, and crucially on his result that the solution could be recovered from the *radiation fields* [42, 43, 44]

$$\begin{aligned}\hat\phi^+(u,\omega) &:= \sqrt{1+u^2}\lim_{r\to+\infty} r\phi(r+u,r,\omega),\\ \hat\phi^-(v,\omega) &:= \sqrt{1+v^2}\lim_{r\to+\infty} r\phi(-r+v,r,\omega).\end{aligned} \tag{1.2.4}$$

The radiation fields (1.2.4) are equivalent to the Lax–Phillips asymptotic profiles (cf. (1.2.2)), and are simply the restrictions of the rescaled solution $\hat\phi$ to $\mathscr{I}^\pm$. Friedlander rephrased the Lax–Phillips scattering theory as the well-posedness of a characteristic Cauchy problem from $\mathscr{I}$, so that knowledge of the radiation fields $\hat\phi^\pm$ determined the solution in the interior of the spacetime[9]. It seems, however, that Friedlander's aim was to recover the full richness, in particular the translation representation, of the Lax–Phillips theory in curved space. This is perhaps what motivated his choice of a static background, since the existence of a translation representation requires a timelike Killing vector field that extends as a null generator of null infinity [91]. This requirement is purely technical, however. The relevant information for the development of a scattering theory—particularly the invertibility of the scattering operator $\mathscr{S}$—is the asymptotics of the physical fields, and it is the method of extraction of these asymptotics from the field equations by spectral techniques that is incompatible with generic time dependence. On the other hand, the techniques of conformal scattering are geometrical, and therefore insensitive to time dependence.

[9]Friedlander's construction was subsequently adapted, on flat space, for a wave equation with a cubic nonlinearity [10, 11, 12, 13, 14] by Baez, Segal and Zhou.

1.2.4 Conformal Scattering

Short of the re-interpretation of the asymptotic profiles as the translation representation, Friedlander's conformal scattering construction can be performed on a large class of non-stationary backgrounds on which the Lax–Phillips theory is unavailable. It is therefore suitable for handling scattering in time-dependent settings. In 1990 Hörmander [58] presented a general method for resolving the characteristic Cauchy problem for linear wave equations on generic spatially compact backgrounds, which eventually allowed Mason and Nicolas to construct scattering theories[10] for Maxwell, wave, and Dirac equations [78, 79] on a large class of asymptotically flat non-stationary spacetimes constructed by Corvino, Schoen, Chruściel, Delay, Klainerman, Nicolò, and others [26, 27, 29, 30, 67, 68, 69]. Since then Joudioux has extended their results to include a wave equation with a cubic nonlinearity [60, 61]. At the same time, there has been much interest in conformal scattering of linear theories on black hole spacetimes [56, 62, 82, 89, 122][11].

On the other hand, even though the cosmological constant is experimentally known to be positive [98, 97, 99, 106], there has been less work on conformal scattering theories on asymptotically de Sitter spacetimes. In part, this is because for linear equations the requisite energy estimates are well-known, and consist only of the resolution of a regular Cauchy problem. Mostly, however, the experimental verification of the positivity of the cosmological constant has only come in the last few decades, and has only sparked a resurgence of interest in the quantum field theory and gravity with $\lambda > 0$ recently. For nonlinear systems such as the Maxwell-scalar field system

$$\begin{aligned} &\Box\phi + L_1(A)\phi = 0, \\ &\Box A_a + L_1'(\phi)A_a = Q(\phi, \partial_a\phi), \end{aligned} \tag{1.2.5}$$

much more work is required to establish even the most basic of estimates that allows the construction of a conformal scattering theory. The system (1.2.5), of physical interest in its own right, is also a background-decoupled semi-linear model for Einstein's equations (1.1.2), as it has a similar nonlinearity structure. It is therefore of considerable interest—as outlined in section 1.1 and earlier in section 1.2—to study conformal scattering, especially of the equations (1.2.5) and related systems, on both asymptotically flat and asymptotically de Sitter spacetimes. Indeed, perhaps unsurprisingly, the two avenues are closely related mathematically.

To introduce the details of the subject, we present conformal scattering constructions for the linear wave equation on Minkowski and de Sitter spacetimes in section 2.4.

[10] Mason and Nicolas also proved that conformal scattering is equivalent to conventional scattering—in the sense of Lax–Phillips—in the cases addressed in [78].

[11] It should be mentioned here that there have been plenty of works studying relativistic scattering theory without employing the conformal method, notably by Dimock and Kay in the 1980s [33, 34] and later by Bachelot [7, 8] and collaborators, a programme that eventually led to rigorous proofs of the Hawking effect [9, 80]. More recently, a definitive scattering theory for linear waves on the exterior of a Kerr black hole was constructed by Dafermos, Rodnianski and Shlapentokh-Rothman [32].

1.3 Outline of Thesis

In this thesis we study the conformal scattering and asymptotic properties of several incarnations of the Yang–Mills–Higgs equations. In chapter 2 we introduce our conventions, outline the complete conformal compactifications of Minkowski and de Sitter spacetimes, describe in detail the various fields—linear and nonlinear—studied in the following chapters, and present an introductory conformal scattering construction for the linear wave equation.

In chapter 3 we study the conformal scattering of Maxwell *potentials*, first on Minkowski space and subsequently on curved spacetimes. In the case of Minkowski space, we construct an explicit conformal factor which is equal to r^{-1} near spatial infinity i^0, so that spatial infinity remains a singular point of the spacetime. We find a complete gauge fixing condition which allows us to convert the natural energies for Maxwell's equations into norms on the potential on both the initial surface and null infinity, and construct the trace operators

$$\begin{aligned} \mathfrak{T}^{\pm} : \dot{H}^1 \oplus L^2 &\longrightarrow \mathcal{H}^1(\mathscr{I}^{\pm}), \\ (\mathbf{A}, \dot{\mathbf{A}}) &\longmapsto \hat{A}_2^{\pm}, \end{aligned}$$

where $\hat{A}_2^{\pm}$ are the $\mathbb{S}^2$ components of the Maxwell potential on $\mathscr{I}^{\pm}$. We therefore find that characteristic data for Maxwell potentials consists of one complex function on null infinity. To solve the characteristic Cauchy problem, we find a reduction on $\mathscr{I}$ of the complete gauge fixing condition, and use it to reconstruct a missing component of the data on null infinity. We then solve the characteristic Cauchy problem using the techniques outlined in appendix A.2.2 to show that the trace operators $\mathfrak{T}^{\pm}$ are invertible.

In the case of curved space, the complete gauge fixing condition used in Minkowski space does not extend, and we have to use a slight modification of the argument and the scattering theory of Mason and Nicolas [78] for Maxwell fields. A key feature here is reconstructing the data for the potential from the data for the field, both on the initial surface and on null infinity. The main theorem of chapter 3 is the following.

Theorem. *On a large class of curved non-stationary spacetimes there exist linear isomorphisms*

$$\begin{aligned} \mathfrak{T}^{\pm} : \dot{H}^1 \oplus L^2 &\longrightarrow \mathcal{H}^1(\mathscr{I}^{\pm}), \\ (\mathbf{A}, \dot{\mathbf{A}}) &\longmapsto \hat{A}_2^{\pm}, \end{aligned}$$

called the future (past) Maxwell trace operators which map finite energy initial data for the Maxwell potential to finite energy characteristic data for the Maxwell potential. The associated conformal scattering operator

$$\begin{aligned} \mathscr{S} : \mathcal{H}^1(\mathscr{I}^{-}) &\longrightarrow \mathcal{H}^1(\mathscr{I}^{+}), \\ \mathscr{S} &= \mathfrak{T}^{+} \circ (\mathfrak{T}^{-})^{-1}, \end{aligned}$$

is then a linear isomorphism of Hilbert spaces mapping past characteristic data for the Maxwell potential to its future characteristic data.

Furthermore, in the flat case we briefly study the effects of different choices of timelike vector field on the domains and co-domains of the operators $\mathfrak{T}^{\pm}$ and $\mathscr{S}$, and spell out the implications for the decay of the initial and characteristic data. Our construction extends the results of [78], and touches for the first time on the different flavours of conformal scattering obtained by different choices of timelike vector field.

In chapter 4 we consider the nonlinear Maxwell-scalar field system on de Sitter space and prove small data energy estimates of the peeling type[12]. These estimates rely crucially on the subcritical nature of the nonlinearity of the Maxwell-scalar field system in four dimensions. We find that, using a careful choice of gauge, it is possible to control all components of the Maxwell potential and the scalar field, and close the estimates using a nonlinear Grönwall inequality. The results of chapter 4 are the following.

Theorem. *For any $m \in \mathbb{N}$, the $H^m \oplus H^{m-1}$ norm on null infinity of the rescaled solution of the Maxwell-scalar field system is equivalent to the $H^m \oplus H^{m-1}$ norm of the initial data, provided the initial data is sufficiently small.*

The estimates also allow us to define a sequence of bounded and invertible trace and scattering operators; however, these operators are nonlinear, and we do not prove them continuous.

Corollary. *For any $m \in \mathbb{N}$, $m \geqslant 2$, there exist bounded, invertible, but nonlinear, trace operators $\mathfrak{T}^{\pm}_m$ mapping small $H^m \oplus H^{m-1}$ Maxwell-scalar field initial data to small $H^m \oplus H^{m-1}$ Maxwell-scalar field data on $\mathscr{I}^{\pm}$. Furthermore, for each $m \geqslant 2$ there exists a bounded invertible nonlinear scattering operator given by*

$$\mathscr{S}_m = \mathfrak{T}^+_m \circ (\mathfrak{T}^-_m)^{-1}.$$

For data possessing sufficiently many derivatives, we deduce that the physical fields decay exponentially along timelike geodesics approaching null infinity, extending the results of Melrose, Sá Barreto and Vasy [81]. In addition to their interpretation in terms of peeling and conformal scattering, our results may also be seen as a fixed background stability result in the spirit of Friedrich, Svedberg and Ringström [47, 48, 100, 114].

Theorem. *Small data solutions to the Maxwell-scalar field system on de Sitter space decay exponentially in proper time along timelike curves approaching $\mathscr{I}$.*

Finally, at the end of chapter 4 we examine the implications of the initial data having a specified number of derivatives on the asymptotic form of the solution. In particular, we show that the slowest-decaying component of the scalar field asymptotically decouples from the Maxwell potential. This chapter forms the basis of the material published in [117].

In chapter 5 we study large data solutions to the Yang–Mills–Higgs equations. Using a conformal patching construction, we extend Eardley and Moncrief's L^∞ estimates for the Yang–Mills–Higgs equations [36] to the Einstein cylinder, and use them to extend Choquet-Bruhat and Christodoulou's small data well-posedness result on the Einstein cylinder [23] to large data.

[12] See section 2.4.2 for a description of peeling in de Sitter space.

Theorem. *For $H^2 \times H^1$ initial data on the Einstein cylinder there exists a unique global solution to the Yang–Mills–Higgs equations which is uniformly bounded on any compact subset of the Einstein cylinder.*

By conformally embedding Minkowski and de Sitter spaces in the Einstein cylinder, we then deduce large data decay rates for solutions to the Yang–Mills–Higgs equations on both spacetimes. This settles the question of whether large data solutions to the Yang–Mills–Higgs equations can form singularities or disperse as linear waves, and improves on our earlier result for small data in chapter 4.

Theorem. *For $H^2 \times H^1$ initial data on Minkowski space (de Sitter space) there exists a global unique solution which decays polynomially (exponentially) along timelike and null (timelike) curves approaching $\mathscr{I}$.*

Chapter 5 contains the material published in [118].

2

Background Material

2.1 Conventions and Notation

Our conventions will for the most part be consistent with Penrose & Rindler [95, 96]. In particular, we will be using their abstract index notation, whereby indices on tensor fields will not refer to any particular chart but merely serve as markers for the *type* of tensor field being considered. A 4-dimensional Lorentzian manifold $\mathscr{M}$ will be called a *spacetime* if it is connected, orientable, time-orientable, admits a Lorentzian metric g_{ab}, contains no closed timelike curves, and satisfies the Einstein equations (1.1.1) for some $\mathbf{T}_{ab}$. We will work exclusively on 4-dimensional spacetimes and our metric signature shall be $(+,-,-,-)$. We will denote the Riemann and Ricci curvature tensors by R_{abcd} and R_{ab}, and the scalar curvature by R. The Riemann curvature tensor will be defined in Penrose conventions: $2\nabla_{[a}\nabla_{b]}X^c = -\mathrm{R}^c{}_{dab}X^d$. We will denote by λ the cosmological constant, and by Λ the multiple $\frac{1}{24}\mathrm{R}$ of the scalar curvature. The Weyl tensor shall be denoted by C_{abcd}, and Φ_{ab} will denote the trace-free part of the Ricci tensor. We will take the connection ∇ to be the Levi–Civita connection of the spacetime metric g_{ab}, and will denote by $\Box = \nabla^a\nabla_a$ the corresponding covariant wave operator. For Riemannian metrics h_{ab} (such as the metric induced by g_{ab} on a Cauchy surface Σ of $\mathscr{M}$), if Σ is non-compact we will denote the Levi–Civita connection in bold, $\boldsymbol{\nabla}$, and the corresponding Laplace–Beltrami operator by $\boldsymbol{\Delta}$. If Σ is compact, we will instead denote the connection by $\not{\nabla}$, and the Laplace–Beltrami operator by $\not{\Delta}$. Symmetrizations and anti-symmetrizations of indices will be weighted by the number of permutations of the indices, for example $T_{[ab]} = \frac{1}{2}(T_{ab} - T_{ba})$.

The letters θ and ϕ will denote the standard coordinates on the unit 2-sphere $\mathbb{S}^2$, while the standard metric on $\mathbb{S}^2$ will be denoted by $\mathfrak{s}_2$. The Levi–Civita connection on $\mathbb{S}^2$ will be denoted by $\nabla_{\mathbb{S}^2}$, with the corresponding Laplace–Beltrami operator $\Delta_{\mathbb{S}^2}$. More generally, $\mathbb{S}^n$ will denote the standard unit n-sphere with metric $\mathfrak{s}_n$. The metric on Minkowski space will be denoted by η_{ab}. The symbol $\mathfrak{E}$ will denote the Einstein cylinder $\mathbb{R}\times\mathbb{S}^3$ with its natural metric $\mathfrak{e} = 1 \oplus (-\mathfrak{s}_3)$.

Much of this thesis involves the usage of conformal transformations $\hat{g}_{ab} = \Omega^2 g_{ab}$ of the spacetime metric. Denoting by $\hat{g}_{ab}$ the 'rescaled' metric, we will also denote by $\hat{\nabla}$

the Levi–Civita connection of $\hat{g}_{ab}$, and by $\widehat{\Box}$ the wave operator of $\hat{g}_{ab}$. More generally, unhatted quantities will be physical, whereas hatted quantities will refer to the rescaled spacetime. We will depart slightly from this convention in chapter 4, where, for ease of notation, we will denote physical quantities with a tilde, and the rescaled quantities plainly. The tensor Υ_a will denote the quantity $\partial_a \log \Omega$, which will appear frequently in conformal transformation laws of curvature tensors. Null infinity will be denoted by $\mathscr{I}$, and equality *on* $\mathscr{I}$ will be denoted by $\approx$.

Finally, we will denote by $\lesssim$ inequality up to a constant, so that $f \lesssim g$ will mean $f \leqslant Cg$. If $f \lesssim g$ and $g \lesssim f$, we will write $f \simeq g$. For a coordinate x, we will frequently use the shorthand ∂_x for the vector field $\partial/\partial x$. The standard Sobolev and Hölder spaces will be denoted by $W^{m,p}$ and $\mathcal{C}^{m,\delta}$ and are defined in appendix A.3.

2.2 Physical Fields

In this section we work on a generic spacetime $(\mathscr{M}, g)$ and introduce the various physical fields that we will be studying throughout this thesis. All of the following field equations will come from the principle of least action and a choice of Lagrangian density $\mathcal{L}$, and be given by the Euler–Lagrange equations

$$\frac{\delta S}{\delta \psi} = \frac{\partial \mathcal{L}}{\partial \psi} - \nabla_a \left(\frac{\partial \mathcal{L}}{\partial(\nabla_a \psi)} \right) = 0.$$

Each theory will have a canonical stress-energy tensor

$$\mathbf{T}_{ab} := 2\frac{\delta \mathcal{L}}{\delta g^{ab}} - g_{ab}\mathcal{L},$$

however we will sometimes modify it to fit our purposes.

2.2.1 Maxwell's Equations

The Lagrangian for Maxwell's equations is

$$\mathcal{L} = -\frac{1}{4} F_{ab} F^{ab}, \tag{2.2.1}$$

where $F = F_{ab}\,\mathrm{d}x^a \wedge \mathrm{d}x^b$ is a real 2-form called the Maxwell field. The Euler–Lagrange equations for (2.2.1) are

$$\nabla^a F_{ab} = 0 \iff \mathrm{d} * F = 0, \tag{2.2.2}$$

where $*$ is the Hodge-star operator associated to the metric g_{ab}, along with the Bianchi identity

$$\nabla_{[a} F_{bc]} = 0 \iff \mathrm{d}F = 0. \tag{2.2.3}$$

The equation (2.2.3) states that the 2-form F is closed, so by the Poincaré lemma F is locally exact: there exists a real 1-form $A = A_a \mathrm{d}x^a$ such that

$$F = \mathrm{d}A.$$

The 1-form A is called the Maxwell potential. Since $\mathrm{d}^2 = 0$, the Maxwell potential is only determined by the Maxwell field F up to exact 1-forms $\mathrm{d}\chi$, so that the potentials A

and $A + \mathrm{d}\chi$ give rise to the same Maxwell field. The freedom of adding any such $\mathrm{d}\chi$ to A is called the gauge freedom in the Maxwell potential, and the addition of $\mathrm{d}\chi$ to A is called a gauge transformation of A. In indices, gauge transformations are

$$A_a \rightsquigarrow A_a + \nabla_a \chi, \tag{2.2.4}$$

and the Maxwell field is given in terms of the Maxwell potential by

$$F_{ab} = 2\partial_{[a} A_{b]} = \nabla_a A_b - \nabla_b A_a.$$

The Bianchi identities (2.2.3) reduce to a triviality when written in terms of the potential, while the equation of motion (2.2.2) becomes

$$\Box A_b - \nabla_b(\nabla_a A^a) + \mathrm{R}_{ab} A^a = 0. \tag{2.2.5}$$

The stress-energy tensor for (2.2.1) is given by

$$\mathbf{T}_{ab} = -F_a{}^c F_{bc} + \frac{1}{4} g_{ab} F_{cd} F^{cd}, \tag{2.2.6}$$

and satisfies

$$\nabla^a \mathbf{T}_{ab} = (\nabla^a F_{ac}) F^c_{\ b} + \frac{3}{2} F^{cd} \nabla_{[b} F_{cd]}. \tag{2.2.7}$$

If the field equations (2.2.2) and (2.2.3) are satisfied, then of course the stress-energy tensor is conserved,

$$\nabla^a \mathbf{T}_{ab} = 0.$$

Furthermore, the stress-energy tensor $\mathbf{T}_{ab}$ is manifestly trace-free,

$$\mathbf{T}_a{}^a = 0.$$

Conformal invariance

Consider a conformal transformation $\hat{g}_{ab} = \Omega^2 g_{ab}$ of the spacetime $(\mathscr{M}, g)$ for some smooth strictly positive function Ω. The covariant derivative of a 1-form α_b on $\mathscr{M}$ then transforms as

$$\hat{\nabla}_a \alpha_b = \nabla_a \alpha_b - \Upsilon_a \alpha_b - \Upsilon_b \alpha_a + g_{ab} \Upsilon_c \alpha^c,$$

where $\Upsilon_a = \partial_a \log \Omega$. From this we observe that

$$\hat{\nabla}_a \alpha_b - \hat{\nabla}_b \alpha_a = \nabla_a \alpha_b - \nabla_b \alpha_a,$$

so that if we set

$$\hat{A}_a := A_a, \tag{2.2.8}$$

then

$$\hat{F}_{ab} = F_{ab}. \tag{2.2.9}$$

That is, if the Maxwell potential is taken to have conformal weight zero ($\hat{A}_a = \Omega^0 A_a$), then the Maxwell field is also conformally covariant with conformal weight zero. This choice of conformal weight has the feature that the field equations remain conformally invariant. Indeed, the Lagrangian (2.2.1) then transforms as

$$\mathcal{L} = -\frac{1}{4}F_{ab}F^{ab} = -\frac{1}{4}\Omega^4\hat{F}_{ab}\hat{F}^{ab} = \Omega^4\hat{\mathcal{L}},$$

where $\hat{\mathcal{L}} = -\frac{1}{4}\hat{F}_{ab}\hat{F}^{ab}$ (the indices of $\hat{F}_{ab}$ being raised and lowered with $\hat{g}_{ab}$), so the action S transforms as

$$S = \int_{\mathscr{M}} \mathcal{L}\,\mathrm{dv} = \int_{\mathscr{M}} \hat{\mathcal{L}}\,\Omega^4\,\Omega^{-4}\,\widehat{\mathrm{dv}} = \hat{S}, \tag{2.2.10}$$

where $\hat{S} = \int_{\mathscr{M}} \hat{\mathcal{L}}\,\widehat{\mathrm{dv}}$, and where we have used the fact that $\mathrm{dv} = \Omega^{-4}\,\widehat{\mathrm{dv}}$. The Euler–Lagrange equations of S are thus equivalent to the Euler–Lagrange equations of $\hat{S}$,

$$\Big(\nabla^a F_{ab} = 0, \quad \nabla_{[a}F_{bc]} = 0\Big) \iff \Big(\hat{\nabla}^a \hat{F}_{ab} = 0, \quad \hat{\nabla}_{[a}\hat{F}_{bc]} = 0\Big).$$

Moreover, it is simple to check that the stress-energy tensor (2.2.6) has conformal weight -2,

$$\hat{\mathbf{T}}_{ab} = \Omega^{-2}\mathbf{T}_{ab}, \tag{2.2.11}$$

where $\hat{\mathbf{T}}_{ab} = -\hat{F}_a{}^c\hat{F}_{bc} + \frac{1}{4}\hat{g}_{ab}\hat{F}_{cd}\hat{F}^{cd}$. The rescaled stress-energy tensor $\hat{\mathbf{T}}_{ab}$ satisfies an identity of the form (2.2.7) with all the relevant quantities hatted, and is of course conserved as a consequence of the rescaled field equations $\hat{\nabla}^a\hat{F}_{ab} = 0$ and $\hat{\nabla}_{[a}\hat{F}_{bc]} = 0$.

Popular gauge choices

The gauge freedom (2.2.4) allows one to impose various gauge conditions on the Maxwell potential A_a. The appropriate choice of gauge is often dictated by the particular problem at hand, and we shall make use of various gauge conditions throughout this thesis. To mention the most popular ones, we suppose for a moment that $\mathscr{M}$ is Minkowski space and that t is the standard time coordinate on $\mathscr{M}$. The first important gauge condition is the *Lorenz*[1] *gauge*, in which

$$\nabla_a A^a = 0. \tag{2.2.12}$$

On Minkowski space this may be expanded as

$$\partial_t \mathfrak{a} - \boldsymbol{\nabla}\cdot\mathbf{A} = 0,$$

where $\mathfrak{a}$ is the component of A_a in the direction of ∂_t, and $\mathbf{A}$ are the remaining (spatial) components of A_a. Given any (sufficiently smooth) instance of the potential A'_a, one imposes the Lorenz gauge on $A_a = A'_a + \nabla_a\chi$ by solving

$$\Box\chi = -\nabla_a A'^a$$

[1] Although the Lorenz gauge has the welcome feature of being Lorentz invariant, it is named after the 19th century Danish mathematician Ludvig Lorenz, *not* the Dutch 1902 Nobel Prize physicist Hendrik Lorentz. The scientific legacies of the two seem to have blended more than once, as they also claim the eponymous Lorentz–Lorenz equation relating the refractive index of a substance to its polarizability, discovered by Lorenz in 1869, and independently by Lorentz in 1878.

on $\mathscr{M}$. The gauge is then fixed up to solutions $\chi_{\text{res.}}$ of the free wave equation $\Box\chi_{\text{res.}} = 0$. The gauge condition (2.2.12) generalizes easily to curved spacetimes.

Another important gauge condition is the *temporal gauge*

$$\mathfrak{a} = 0, \tag{2.2.13}$$

which is imposed on $A_a = A'_a + \nabla_a\chi$ by choosing any function χ such that $\partial_t\chi = -A'_t$. The residual gauge freedom $\chi_{\text{res.}}$ in the temporal gauge is rather substantial ($\partial_t\chi_{\text{res.}} = 0$, but the spatial dependence of $\chi_{\text{res.}}$ is allowed to be arbitrary), so the temporal gauge is frequently termed an incomplete gauge fixing condition.

A third important gauge condition is the *Coulomb gauge*

$$\boldsymbol{\nabla} \cdot \mathbf{A} = 0. \tag{2.2.14}$$

The Coulomb gauge is imposed by solving $\Delta\chi = -\boldsymbol{\nabla} \cdot \mathbf{A}'$, and has the residual gauge freedom of $\Delta\chi_{\text{res.}} = 0$. Both the temporal and the Coulomb gauges require a choice of a timelike vector field or a foliation by spacelike hypersurfaces to generalize to curved space.

2.2.2 The Conformal Wave Equation

As we will be studying primarily conformally invariant theories, we will be interested in the wave equation with a particular zeroth order perturbation. Consider the Lagrangian

$$\mathcal{L} = \frac{1}{2}\nabla_a\phi\nabla^a\phi - \frac{1}{12}\mathrm{R}\phi^2, \tag{2.2.15}$$

where ϕ is a real scalar field on $\mathscr{M}$. The Euler–Lagrange equation for (2.2.15) is

$$\Box\phi + \frac{1}{6}\mathrm{R}\phi = 0, \tag{2.2.16}$$

and the canonical stress-energy tensor for (2.2.15) is

$$\boldsymbol{\Theta}_{ab} = \nabla_a\phi\nabla_b\phi - \frac{1}{2}g_{ab}\nabla_c\phi\nabla^c\phi - \frac{1}{6}\mathrm{G}_{ab}\phi^2, \tag{2.2.17}$$

which (for metrics with an appropriately signed Einstein tensor G_{ab}) is positive-definite and symmetric. The tensor $\boldsymbol{\Theta}_{ab}$ satisfies the conservation law

$$\nabla^a\boldsymbol{\Theta}_{ab} = \left(\Box\phi + \frac{1}{6}\mathrm{R}\phi\right)\nabla_b\phi - \frac{1}{3}\phi\mathrm{R}_{ab}\nabla^a\phi,$$

so is conserved whenever (2.2.16) is satisfied and the background $\mathscr{M}$ is vacuum with zero cosmological constant. To have a conserved stress-energy tensor $\mathbf{T}_{ab}$ on more general backgrounds, one may choose

$$\mathbf{T}_{ab} = \nabla_a\phi\nabla_b\phi - \frac{1}{2}g_{ab}\nabla_c\phi\nabla^c\phi + \frac{1}{12}g_{ab}\mathrm{R}\phi^2, \tag{2.2.18}$$

which is also positive-definite and symmetric, and which satisfies the conservation law

$$\nabla^a\mathbf{T}_{ab} = \left(\Box\phi + \frac{1}{6}\mathrm{R}\phi\right)\nabla_b\phi + \frac{1}{12}\phi^2\nabla_b\mathrm{R}.$$

For (2.2.18) to be conserved, the background spacetime need only satisfy $\nabla_b\mathrm{R} = 0$ instead of $\mathrm{R}_{ab} = 0$. Neither $\mathbf{T}_{ab}$ nor $\boldsymbol{\Theta}_{ab}$ is trace-free, however, with

$$\mathbf{T}_a{}^a = -\nabla_a\phi\nabla^a\phi + \frac{1}{3}\mathrm{R}\phi^2 \quad \text{and} \quad \boldsymbol{\Theta}_a{}^a = -\nabla_a\phi\nabla^a\phi + \frac{1}{6}\mathrm{R}\phi^2.$$

Conformal invariance

If under the conformal transformation $\hat{g}_{ab} = \Omega^2 g_{ab}$ the scalar field ϕ is chosen to have weight -1,

$$\hat{\phi} := \Omega^{-1}\phi,$$

then, as a result of the identity (B.1.3), the Lagrangian (2.2.15) transforms as

$$\mathcal{L} = \Omega^4\hat{\mathcal{L}} + \frac{1}{2}\Omega^4\hat{\nabla}^a(\hat{\phi}^2\Upsilon_a),$$

where $\hat{\mathcal{L}} = \frac{1}{2}\hat{\nabla}_a\hat{\phi}\hat{\nabla}^a\hat{\phi} - \frac{1}{12}\hat{\mathrm{R}}\hat{\phi}^2$. The action then transforms as

$$S = \hat{S} + \frac{1}{2}\int_{\mathscr{M}} \hat{\nabla}^a(\hat{\phi}^2\Upsilon_a)\,\widehat{\mathrm{dv}}. \tag{2.2.19}$$

For compactly supported scalar fields $\phi \in \mathcal{C}_c^\infty(\mathscr{M})$, one may integrate the second term in (2.2.19) by parts to see that it vanishes. Since the Euler–Lagrange equation (2.2.16) arises from a local variation of the action, it follows that (2.2.16) is conformally invariant[2],

$$\left(\Box\phi + \frac{1}{6}\mathrm{R}\phi = 0\right) \iff \left(\widehat{\Box}\hat{\phi} + \frac{1}{6}\hat{\mathrm{R}}\hat{\phi} = 0\right).$$

The term R/6 is sometimes called the *conformal mass*, in reference to the massive Klein–Gordon equation

$$\Box\phi + m^2\phi = 0.$$

Remark 2.2.1. Unlike in the case of Maxwell's equations, neither the canonical stress-energy tensor (2.2.17) nor the alternative stress-energy tensor (2.2.18) for ϕ is conformally covariant. There does exist a 'new-and-improved' conserved stress-energy tensor which *is* conformally covariant,

$$\mathbf{T}'_{ab} = \nabla_a\phi\nabla_b\phi - \frac{1}{4}g_{ab}\nabla_c\phi\nabla^c\phi - \frac{1}{2}\phi\nabla_a\nabla_b\phi + \Lambda g_{ab}\phi^2 - \frac{1}{4}\mathrm{R}_{ab}\phi^2, \tag{2.2.20}$$

which is moreover symmetric and trace-free when the field equation (2.2.16) is satisfied [96]. However, $\mathbf{T}'_{ab}$ is not positive-definite and is also second order in derivatives of ϕ, making it unhelpful for performing low regularity energy estimates. In flat space $\mathbf{T}'_{ab}$ differs from $\mathbf{T}_{ab}$ by the spacetime divergence $\nabla^c(\phi g_{a[b}\nabla_{c]}\phi)$. The problem of finding a symmetric, first order, positive-definite stress-energy tensor for the scalar field (2.2.16) which is also conserved, trace-free and conformally covariant seems to be related to the slightly garbled conformal transformation rule (2.2.19) of its action, and consequently the asymptotics of ϕ. We do not wish to clutter the presentation with unnecessary discussion, but include the following perhaps curious observation.

Klainerman and Machedon's [66] null form $Q_0(\phi,\phi)$ for the wave equation $\Box\phi = 0$, at least on flat space, is given by

$$Q_0(\phi,\phi) = \dot{\phi}^2 - |\boldsymbol{\nabla}\phi|^2 = \nabla_a\phi\nabla^a\phi = \frac{1}{2}\Box(\phi^2).$$

[2] One can in fact show that $\Box\phi + \frac{1}{6}\mathrm{R}\phi = \Omega^3\left(\widehat{\Box}\hat{\phi} + \frac{1}{6}\hat{\mathrm{R}}\hat{\phi}\right)$.

One notices that, up to factors of 2, $Q_0(\phi,\phi)$ is in fact equal to the Lagrangian $\mathcal{L}$ on flat space, which is moreover equal to the trace of the canonical stress-energy tensor,

$$\boldsymbol{\Theta}_a{}^a = -\nabla_a\phi\nabla^a\phi = -\frac{1}{2}\Box(\phi^2) = -2\mathcal{L}.$$

This observation extends to the conformal wave equation $\Box\phi + (\mathrm{R}/6)\phi = 0$ on any spacetime, so long as we *redefine* the null form $Q_0(\phi,\phi)$ to be

$$Q_0(\phi,\phi) := \frac{1}{2}\Box(\phi^2) = \nabla_a\phi\nabla^a\phi - \frac{1}{6}\mathrm{R}\phi^2 = -\boldsymbol{\Theta}_a{}^a = -2\mathcal{L}.$$

2.2.3 The Maxwell-Scalar Field System

One may couple the conformal wave equation (2.2.16) to Maxwell's equations (2.2.2) by writing down the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{ab}F^{ab} + \frac{1}{2}\mathrm{D}_a\phi\overline{\mathrm{D}^a\phi} - \frac{1}{12}\mathrm{R}|\phi|^2, \tag{2.2.21}$$

where ϕ is a complex scalar field, and the Maxwell field is, as before, the real 2-form $F_{ab} = 2\nabla_{[a}A_{b]}$. The coupling is effected via the differential operator

$$\mathrm{D}_a := \nabla_a + iA_a$$

called the *gauge covariant derivative*. That is, the gauge covariant derivative of the scalar field ϕ is $\mathrm{D}_a\phi = \nabla_a\phi + iA_a\phi$, with the term $iA_a\phi$ interpreted as the coupling of a scalar field ϕ with unit electric charge with the electromagnetic potential A_a. The Euler–Lagrange equations for (2.2.21) are

$$\nabla^b F_{ab} = \mathrm{Im}(\bar{\phi}\mathrm{D}_a\phi) \quad \text{and} \quad \mathrm{D}^a\mathrm{D}_a\phi + \frac{1}{6}\mathrm{R}\phi = 0, \tag{2.2.22}$$

along with the Bianchi identity

$$\nabla_{[a}F_{bc]} = 0. \tag{2.2.3}$$

The Maxwell-scalar field system (2.2.22) is the simplest classical field theory exhibiting a non-trivial gauge dependence. Indeed, as in the case of pure Maxwell's equations, the Maxwell field F_{ab} is invariant under gauge transformations of A_a of the form

$$A_a \rightsquigarrow A_a + \nabla_a\chi.$$

These gauge transformations then effect

$$\mathrm{D}_a\phi = \nabla_a\phi + iA_a\phi \rightsquigarrow \nabla_a\phi + i(A_a + \nabla_a\chi)\phi = \mathrm{e}^{-i\chi}\mathrm{D}_a(\mathrm{e}^{i\chi}\phi),$$

so that if one makes the corresponding transformation

$$\phi \rightsquigarrow \mathrm{e}^{-i\chi}\phi,$$

the Lagrangian (2.2.21), and consequently the field equations (2.2.22), remain unchanged under gauge transformations. Note that the gauge covariant derivative D_a is a connection on a principal bundle $P \to \mathscr{M}$ with fibre $\mathrm{U}(1)$. This connection is represented by the real

1-form A_a on $\mathscr{M}$ in any trivialisation of P, where the factor of i in D_a comes from the Lie algebra of U(1) being $\mathfrak{u}(1) = i\mathbb{R}$. The scalar field ϕ is a section of a complex line bundle over $\mathscr{M}$ associated to P by the representation $\mathrm{e}^{i\chi}$ of U(1).

A stress-energy tensor for (2.2.22) is

$$\mathbf{T}_{ab}[\phi, A] = -F_{ac}F_b{}^c + \frac{1}{4}g_{ab}F_{cd}F^{cd} + \overline{\mathrm{D}_{(a}\phi}\mathrm{D}_{b)}\phi - \frac{1}{2}g_{ab}\overline{\mathrm{D}_c\phi}\mathrm{D}^c\phi + \frac{1}{12}g_{ab}\mathrm{R}|\phi|^2, \quad (2.2.23)$$

which is written down by adding the stress-energy tensors (2.2.6) and (2.2.18), and employing the change $\nabla_a \to \mathrm{D}_a$. The tensor $\mathbf{T}_{ab}$ satisfies the conservation law

$$\nabla^a\mathbf{T}_{ab} = F_b{}^a\left(\nabla^c F_{ac} - \mathrm{Im}\left(\bar{\phi}\mathrm{D}_a\phi\right)\right) + \mathrm{Re}\left(\left(\mathrm{D}^a\mathrm{D}_a\phi + \frac{1}{6}\mathrm{R}\phi\right)\overline{\mathrm{D}_b\phi}\right) + \frac{3}{2}F^{ac}\nabla_{[b}F_{ac]} + \frac{1}{12}|\phi|^2\nabla_b\mathrm{R}.$$

If the field equations (2.2.22) and the Bianchi identity (2.2.3) are satisfied, then

$$\nabla^a\mathbf{T}_{ab} = \frac{1}{12}|\phi|^2\nabla_b\mathrm{R}, \quad (2.2.24)$$

and $\mathbf{T}_{ab}$ is conserved if the background spacetime has constant scalar curvature.

Conformal invariance and gauge choices

The conformal transformation rule for the Lagrangian (2.2.21) is the same as for the conformal wave equation. That is, if the fields (A_a, ϕ) are taken to have conformal weights $(0, -1)$,

$$\hat{A}_a := A_a, \qquad \hat{\phi} := \Omega^{-1}\phi,$$

then the action satisfies (2.2.19). This may be seen by computing the conformal transformation of the Lagrangian (2.2.21) directly, by using the conformal transformation rules (B.1.6) and (B.1.3) for the Christoffel symbols Γ^a_{bc} and the scalar curvature R. As before, the field equations (2.2.22) remain conformally invariant.

As in the case of the free Maxwell's equations (section 2.2.1), various gauge choices are possible for the system (2.2.22). In the case of flat space (as well as most other spacetimes) all three of the most popular gauge choices (Lorenz, Coulomb, and temporal) are allowed individually. However, unlike in the case of the free Maxwell's equations on flat space, no two of these are allowed simultaneously. In chapter 4 we shall define and use a certain *strong Coulomb gauge* on the Einstein cylinder; this is a complete gauge fixing condition that resembles as much as possible the simultaneous imposition of the Coulomb and temporal gauges in flat space, and will allow us to perform very precise estimates on the potential.

2.2.4 The Yang–Mills–Higgs Equations

Finally, we introduce the full Yang–Mills–Higgs equations, a non-commutative version of Maxwell's equations coupled to a nonlinear scalar field. For an appropriate choice of the gauge group (G = U(1)) the Yang–Mills equations reduce to Maxwell's equations, whereas the Yang–Mills–Higgs equations reduce to the Maxwell-scalar field system.

Let G be a connected matrix Lie group with a compact semi-simple Lie algebra $\mathfrak{g}$. In particular, we assume that $\mathfrak{g}$ is represented by a subalgebra of the algebra of real matrices equipped with the usual matrix commutator, and admits a positive-definite Ad-invariant scalar product $\langle\cdot,\cdot\rangle$ given by

$$\langle X, Y\rangle = -\operatorname{Tr}(XY) \quad \forall X, Y \in \mathfrak{g}.$$

Let $\{\theta_\alpha\}$ be the generators of $\mathfrak{g}$ in such a representation and let $f_{\alpha\beta}{}^{\gamma}$ be the structure constants of $\mathfrak{g}$, so that

$$[\theta_\alpha, \theta_\beta] = f_{\alpha\beta}{}^{\gamma}\theta_\gamma.$$

Since $\mathfrak{g}$ is semi-simple, $f_{\alpha\beta\gamma}$ can be chosen to be totally antisymmetric in its indices, and the generators can be chosen to be real antisymmetric matrices satisfying

$$\langle\theta_\alpha, \theta_\beta\rangle = \delta_{\alpha\beta}.$$

Let $P \to \mathscr{M}$ be a principal G-bundle over $\mathscr{M}$. The Yang–Mills potential A is defined to be a connection on P, and in any trivialization of P over a coordinate patch $\mathscr{U}$ of $\mathscr{M}$ is given by a $\mathfrak{g}$-valued 1-form on $\mathscr{U}$,

$$A = A_a(x)\,\mathrm{d}x^a, \qquad A_a(x) = A_a^\alpha(x)\theta_\alpha \in \mathfrak{g}$$

for some real-valued functions A_a^α on $\mathscr{U}$. The curvature of A (or the Yang–Mills field) is then the $\mathfrak{g}$-valued 2-form

$$F = F_{ab}(x)\,\mathrm{d}x^a \wedge \mathrm{d}x^b = (F_{ab}^\alpha(x)\theta_\alpha)\,\mathrm{d}x^a \wedge \mathrm{d}x^b$$

given by

$$F_{ab} = \nabla_a A_b - \nabla_b A_a + [A_a, A_b]$$

in $\mathscr{U}$, where ∇_a is the Levi–Civita connection on $\mathscr{M}$. We define the Higgs field ϕ to be a section of the real[3] vector bundle associated to the representation $\{\theta_\alpha\}$. We denote the inner product of such sections by $\phi\cdot\psi = \phi_\alpha\psi_\alpha$, and write, for example, $|\phi|^2 = \phi\cdot\phi = \phi_\alpha\phi_\alpha$. The gauge-covariant derivative D_a of ϕ is defined to be

$$\mathrm{D}_a\phi = \nabla_a\phi + A_a\phi. \tag{2.2.25}$$

The Yang–Mills field is then the commutator of two gauge-covariant derivatives:

$$F_{ab}\phi = [\mathrm{D}_a, \mathrm{D}_b]\phi.$$

Under a gauge transformation

$$\begin{aligned} A_a &\rightsquigarrow UA_aU^{-1} + U\partial_a U^{-1}, \\ F_{ab} &\rightsquigarrow UF_{ab}U^{-1}, \\ \phi &\rightsquigarrow U\phi, \text{ and} \\ \mathrm{D}_a\phi &\rightsquigarrow U\mathrm{D}_a\phi \end{aligned}$$

[3]Note that ϕ being a section of a real vector bundle does not preclude it being a complex scalar field, as in the case of section 2.2.3, since $\mathbb{C}$ is a (two-dimensional) vector space over $\mathbb{R}$.

for any smooth G-valued function U on $\mathscr{M}$. The gauge transformation U here is related to the gauge transformation χ in the abelian case by exponentiation. Notice that in contrast to the abelian case of the Maxwell-scalar field system of section 2.2.3, the fields F_{ab} and $\mathrm{D}_a\phi$ are not gauge-invariant. Furthermore, the gauge-covariant derivative $\mathrm{D}_a\phi$ here seemingly differs from the gauge-covariant derivative we defined in section 2.2.3 by a factor of i; this is only because there we made the i (as in $\mathfrak{u}(1) = i\mathbb{R}$) explicit to make the Maxwell potential A_a a real function. Note also that D_a acts on ϕ by (2.2.25) since ϕ belongs to the fundamental representation of G ($\phi \rightsquigarrow U\phi$). Since F_{ab} belongs to the adjoint representation ($F_{ab} \rightsquigarrow UF_{ab}U^{-1}$), the operator D_a acts on F_{ab} by

$$\mathrm{D}_a F_{bc} = \nabla_a F_{bc} + [A_a, F_{bc}].$$

The Lagrangian for the conformally invariant Yang–Mills–Higgs equations is

$$\mathcal{L} = -\frac{1}{4}\langle F_{ab}, F^{ab}\rangle + \frac{1}{2}(\mathrm{D}_a\phi)\cdot(\mathrm{D}^a\phi) - \frac{1}{12}\mathrm{R}|\phi|^2 - \frac{1}{4}\lambda_0|\phi|^4, \tag{2.2.26}$$

where $\lambda_0 \geqslant 0$ is a constant. The Euler–Lagrange equations associated to (2.2.26) are

$$\mathrm{D}^b F_{ab} = -((\mathrm{D}_a\phi)\cdot\theta_\alpha\phi)\theta_\alpha \quad \text{and} \quad \mathrm{D}^a\mathrm{D}_a\phi + \frac{1}{6}\mathrm{R}\phi + \lambda_0|\phi|^2\phi = 0, \tag{2.2.27}$$

along with the Bianchi identity

$$\mathrm{D}_{[a}F_{bc]} = 0. \tag{2.2.28}$$

If $\lambda_0 \neq 0$, there is a new cubic interaction term in the equation for ϕ that is absent in the Maxwell-scalar field system (2.2.22). From a PDE point of view, however, this cubic nonlinearity is of the same type as the forcing term $\mathrm{Im}(\bar{\phi}\mathrm{D}_a\phi)$ in the equation for the Maxwell field. Note moreover that since $\lambda_0 \geqslant 0$, the equation for ϕ is a *defocussing* nonlinear wave equation [110].

The canonical stress-energy tensor for (2.2.26) is

$$\begin{aligned}\mathbf{\Theta}_{ab} = &-\langle F_{ac}, F_b{}^c\rangle + \frac{1}{4}g_{ab}\langle F_{cd}, F^{cd}\rangle \\ &+ (\mathrm{D}_a\phi)\cdot(\mathrm{D}_b\phi) - \frac{1}{2}g_{ab}(\mathrm{D}_c\phi)\cdot(\mathrm{D}^c\phi) - \frac{1}{6}\mathrm{G}_{ab}|\phi|^2 + \frac{1}{4}\lambda_0 g_{ab}|\phi|^4,\end{aligned} \tag{2.2.29}$$

with an alternative choice being

$$\begin{aligned}\mathbf{T}_{ab} = &-\langle F_{ac}, F_b{}^c\rangle + \frac{1}{4}g_{ab}\langle F_{cd}, F^{cd}\rangle \\ &+ (\mathrm{D}_a\phi)\cdot(\mathrm{D}_b\phi) - \frac{1}{2}g_{ab}(\mathrm{D}_c\phi)\cdot(\mathrm{D}^c\phi) + \frac{1}{12}g_{ab}\mathrm{R}|\phi|^2 + \frac{1}{4}\lambda_0 g_{ab}|\phi|^4.\end{aligned} \tag{2.2.30}$$

As a consequence of the field equations (2.2.27), $\mathbf{\Theta}_{ab}$ and $\mathbf{T}_{ab}$ satisfy the conservation laws

$$\nabla^a\mathbf{\Theta}_{ab} = -\frac{1}{3}\mathrm{R}_{ab}\phi\cdot\nabla^a\phi \quad \text{and} \quad \nabla^a\mathbf{T}_{ab} = \frac{1}{12}|\phi|^2\nabla_b\mathrm{R}.$$

As in the case of the Maxwell-scalar field system, it can be checked that the equations (2.2.27) are conformally invariant under $g_{ab} \rightsquigarrow \hat{g}_{ab} = \Omega^2 g_{ab}$. That is, the conformally rescaled fields

$$\hat{A}_a = A_a, \qquad \hat{F}_{ab} = F_{ab}, \qquad \hat{\phi} = \Omega^{-1}\phi.$$

satisfy the rescaled field equations

$$\hat{\mathrm{D}}^b \hat{F}_{ab} = -((\hat{\mathrm{D}}_a \hat{\phi}) \cdot \theta_\alpha \hat{\phi})\theta_\alpha, \qquad \hat{\mathrm{D}}^a \hat{\mathrm{D}}_a \hat{\phi} + \frac{1}{6}\mathrm{R}\hat{\phi} + \lambda_0 |\hat{\phi}|^2 \hat{\phi} = 0$$

if and only if the physical fields (A_a, F_{ab}, ϕ) satisfy the equations (2.2.27). As in the case of the Maxwell-scalar field system, neither of the stress-energy tensors $\mathbf{\Theta}_{ab}$ or $\mathbf{T}_{ab}$ are a conformally covariant quantity.

2.3 Two Conformal Compactifications

As we shall be making extensive use of these, in this section we introduce the well-known natural conformal compactifications of 4-dimensional Minkowski and de Sitter spacetimes in detail.

2.3.1 Minkowski Space

Four dimensional Minkowski space $(\mathscr{M} = \mathbb{R}^4, \eta)$ is the maximally symmetric solution to the vacuum Einstein equations (1.1.1) with zero cosmological constant. The metric η in the usual spherical coordinates is given by

$$\eta = \mathrm{d}t^2 - \mathrm{d}r^2 - r^2 \mathfrak{s}_2,$$

where $\mathfrak{s}_2$ is the standard metric on the 2-sphere $\mathbb{S}^2_{\theta,\phi}$. We consider a conformal transformation of the metric $\hat{\eta} := \Omega^2 \eta$, where

$$\Omega = \frac{2}{\sqrt{1+(t-r)^2}\sqrt{1+(t+r)^2}}. \tag{2.3.1}$$

Defining

$$\begin{aligned} \tau &:= \arctan(t+r) + \arctan(t-r) \in (-\pi, \pi) \quad \text{and} \\ \zeta &:= \arctan(t+r) - \arctan(t-r) \in [0, \pi), \end{aligned}$$

the conformal factor becomes

$$\Omega = 2\cos\left(\frac{\tau-\zeta}{2}\right)\cos\left(\frac{\tau+\zeta}{2}\right),$$

and the rescaled Minkowski metric $\hat{\eta}$ reads

$$\hat{\eta} = \mathrm{d}\tau^2 - \mathrm{d}\zeta^2 - (\sin^2\zeta)\,\mathfrak{s}_2. \tag{2.3.2}$$

In the new coordinates (τ, ζ) physical Minkowski space is given by the diamond

$$\mathscr{M} = \{(\tau, \zeta) \,:\, |\tau| + \zeta < \pi,\, \zeta \geqslant 0\} \times \mathbb{S}^2, \tag{2.3.3}$$

as shown in fig. 2.1. Quotienting the metric (2.3.2) by the symmetry group SO(3) of the $\mathfrak{s}_2$ factors also gives the Penrose diagram of Minkowski space, fig. 2.2.

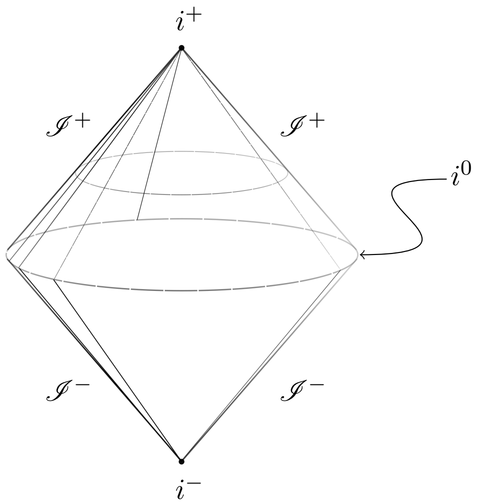

Figure 2.1: Compactified Minkowski space.

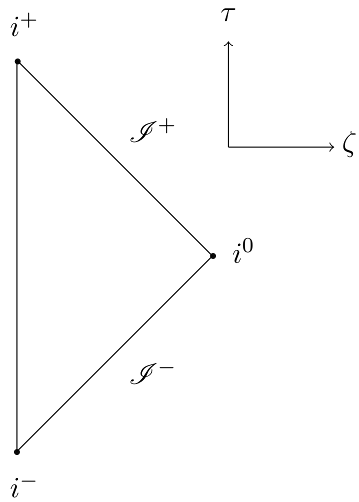

Figure 2.2: The Penrose diagram for Minkowski space.

Identifying $0 \sim \pi$ for the variable ζ, $\zeta \in [0, \pi]/\sim$, one sees that the spacelike part of the metric $\hat{\eta}$ coincides with the standard metric $\mathfrak{s}_3$ on the 3-sphere $\mathbb{S}^3$, $\mathrm{d}\zeta^2 + (\sin^2 \zeta)\, \mathfrak{s}_2 = \mathfrak{s}_3$. It is then clear that the metric

$$\hat{\eta} = \mathrm{d}\tau^2 - \mathfrak{s}_3 \tag{2.3.2}$$

extends smoothly for all $\tau \in \mathbb{R}$ and $\mathbb{S}^3_{\zeta,\theta,\phi}$, and in fact represents the *Einstein cylinder* $(\mathfrak{E}, \mathfrak{e}) = (\mathbb{R} \times \mathbb{S}^3, \hat{\eta})$. The manifold (2.3.3) has a boundary (as a submanifold of $(\mathfrak{E}, \mathfrak{e})$) given by

$$\partial\mathscr{M} = \{\Omega = 0\} = \{(\tau, \zeta) \,:\, |\tau| + \zeta = \pi,\, \zeta \geqslant 0\} \times \mathbb{S}^2 =: \mathscr{I}, \tag{2.3.4}$$

known as *null infinity*. The hypersurface $\mathscr{I}$ is the union of the past and future sections

$$\mathscr{I}^\pm = \{(\tau, \zeta) \,:\, \pm\tau + \zeta = \pi,\, \zeta \in (0, \pi)\} \times \mathbb{S}^2 \tag{2.3.5}$$

and the three points

$$i^\pm = (\pm\pi, 0) \tag{2.3.6}$$

and

$$i^0 = (0, \pi). \tag{2.3.7}$$

The surfaces $\mathscr{I}^\pm$ are called *future (past) null infinities*, the points $i^\pm$ are called *future (past) timelike infinities*, and i^0 is called *spatial* (or *spacelike*) *infinity*. The motivation for this nomenclature may be seen by examining the behaviour of geodesics in physical Minkowski space: if an inextendible geodesic is timelike, it intersects $\mathscr{I}$ at i^- and i^+ only; if it is null, it intersects $\mathscr{I}$ at a point on $\mathscr{I}^-$ and a point on $\mathscr{I}^+$, and if it is spacelike, it is a closed curve through i^0 [94]. Coincidentally[4], the hypersurfaces $\mathscr{I}^\pm$ are null with respect to the metric $\hat{\eta}$, with null generators $\partial_\tau \pm \partial_\zeta$.

[4]This is a consequence of Minkowski space being a solution to Einstein's equations with a vanishing cosmological constant. Null infinity $\mathscr{I}$ is only null in the case of $\lambda = 0$.

Furthermore, the timelike infinities $i^{\pm}$ and spacelike infinity i^0 are genuine points, not 2-spheres, since the 2-sphere factors have radius $\sin\zeta$. They are the endpoints of $\mathscr{I}^{\pm}$, $\partial\mathscr{I}^{\pm} = i^{\pm} \cup i^0$. One sees that $i^{\pm}$, i^0 and the hypersurfaces $\mathscr{I}^{\pm}$ are all smooth with respect to the metric (2.3.2), and that $\hat{\mathscr{M}} := \mathscr{M} \cup \mathscr{I}$ is a genuine smooth compact manifold with boundary called *compactified Minkowski space*. The embedding of $(\hat{\mathscr{M}}, \hat{\eta})$ into the Einstein cylinder $(\mathfrak{E}, \mathfrak{e}) = (\mathbb{R} \times \mathbb{S}^3, \hat{\eta})$ may be visualised as in fig. 2.3.

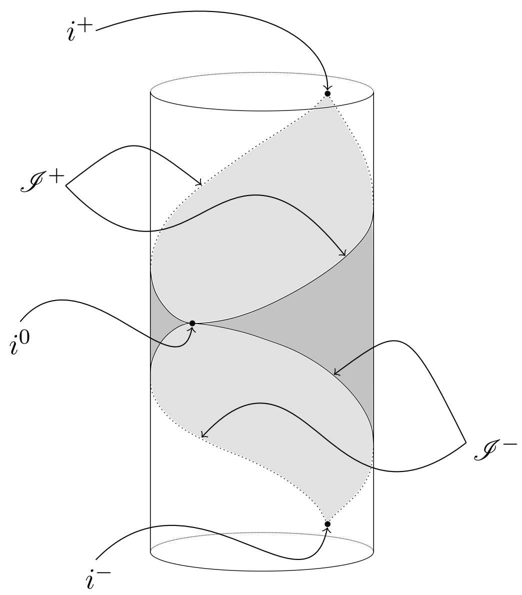

Figure 2.3: The conformal embedding of Minkowski space in the Einstein cylinder may be visualised as a diamond wrapped around a cylinder.

Figure 2.3 shows that the future null cone $\mathscr{I}^-$ of the point i^- is refocused at the point i^0, which is situated antipodally to i^- on the cylinder. Similarly, the future null cone $\mathscr{I}^+$ of i^0 is refocused at i^+. In fact one may regard the union $\mathscr{I}^+ \cup \mathscr{I}^-$ as the null cone of a single point, i^0, with $\mathscr{I}^+$ representing the future cone and $\mathscr{I}^-$ representing the past cone.

Remark 2.3.1. Such a complete compactification which produces a completely non-singular $\mathscr{I}$ (including $i^{\pm}$ and i^0) is unique to flat space (in the case of $\lambda = 0$). The situation is less forgiving in curved space: it turns out, for example, that the points $i^{\pm}$ and i^0 of the Schwarzschild solution are all necessarily singular, as the eigenvalues of the Weyl tensor of the compactified spacetime blow up at these points [94]. Indeed, they are proportional to

$$\frac{m}{r^3\Omega^2},$$

which is unbounded in the neighbourhood of each of $i^{\pm}$ and i^0 for any plausible choice of Ω that brings the relevant point to a finite distance (for example $\Omega = r^{-2}$ for i^0). This will be relevant in chapter 3, where we shall work on curved spacetimes and use an incomplete conformal compactification which will leave i^0 a singular point of the rescaled metric.

2.3.2 De Sitter Space

Four dimensional de Sitter space dS_4 is the maximally symmetric solution to the vacuum Einstein equations (1.1.1) with positive cosmological constant $\lambda > 0$, and the positive scalar curvature analogue[5] of Minkowski space. It may be defined as the hyperboloid[6]

$$|x|^2 - x_0^2 = \frac{1}{H^2}$$

in $(4+1)$-dimensional Minkowski space

$$\eta_5 = \mathrm{d}x_0^2 - \mathrm{d}|x|^2 - |x|^2 \mathfrak{s}_3,$$

where $|x| = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}$. If we set

$$x_0 = \frac{1}{H}\sinh(H\alpha), \qquad |x| = \frac{1}{H}\cosh(H\alpha)$$

so that α is a coordinate on dS_4, the metric η_5 descends to the metric $\tilde{g}$ on dS_4,

$$\tilde{g} = \mathrm{d}\alpha^2 - \frac{1}{H^2}\cosh^2(H\alpha)\, \mathfrak{s}_3. \tag{2.3.8}$$

This provides a global coordinate system on dS_4 and is known as the closed slicing of de Sitter space. These coordinates make manifest the $\mathbb{R} \times \mathbb{S}^3$ topology of dS_4: the metric (5.5.1) can be visualized as a compact spacelike slice expanding in time α, as depicted in fig. 2.4.

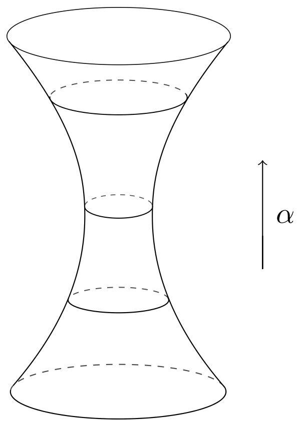

Figure 2.4: The closed slicing of dS_4.

To conformally compactify dS_4, we need a further change of coordinates

$$\tan\left(\frac{\tau}{2}\right) = \tanh\left(\frac{H\alpha}{2}\right).$$

[5]The maximally symmetric solution to the vacuum Einstein equations with *negative* cosmological constant is called anti-de Sitter space AdS_4, which we do not study in this thesis. Anti-de Sitter space has a timelike $\mathscr{I}$, and therefore fails to be globally hyperbolic. It is also expected to be a dynamically unstable solution of Einstein's equations [18, 31, 84, 85].

[6]Note that the parameter H corresponds to the Hubble constant in vacuum.

In terms of τ the metric $\tilde{g}$ becomes

$$\tilde{g} = \frac{1}{H^2 \cos^2 \tau}(\mathrm{d}\tau^2 - \mathfrak{s}_3), \tag{2.3.9}$$

where $\tau \in (-\pi/2, \pi/2)$. The conformal factor

$$\Omega = H \cos \tau$$

therefore conformally embeds dS_4 in the Einstein cylinder $(\mathfrak{E}, \mathfrak{e})$,

$$\Omega^2 \tilde{g} = \mathrm{d}\tau^2 - \mathfrak{s}_3 = \mathfrak{e}.$$

In this conformal scale the hypersurfaces $\{\tau = \pm\pi/2\}$ are regular, in contrast to the physical metric (2.3.9). We thus identify compactified de Sitter space $\widehat{\mathrm{dS}}_4$ with the subset $[-\pi/2, \pi/2] \times \mathbb{S}^3$ of the Einstein cylinder $\mathfrak{E}$ by attaching to $((-\pi/2, \pi/2) \times \mathbb{S}^3, \mathfrak{e})$ the boundary $\mathscr{I} := \{\Omega = 0\} = \{|\tau| = \pi/2\}$. This boundary is the union of two disjoint smooth surfaces

$$\mathscr{I}^+ = \left\{\tau = \frac{\pi}{2}\right\} \quad \text{and} \quad \mathscr{I}^- = \left\{\tau = -\frac{\pi}{2}\right\},$$

which we call future null infinity and past null infinity respectively, analogously to the case of compactified Minkowski space. All inextendible null geodesics in dS_4 acquire two endpoints in $\widehat{\mathrm{dS}}_4$, one on $\mathscr{I}^-$ and one on $\mathscr{I}^+$. Note, however, that here $\mathscr{I}^\pm$ are *spacelike* hypersurfaces of $(\mathfrak{E}, \mathfrak{e})$.

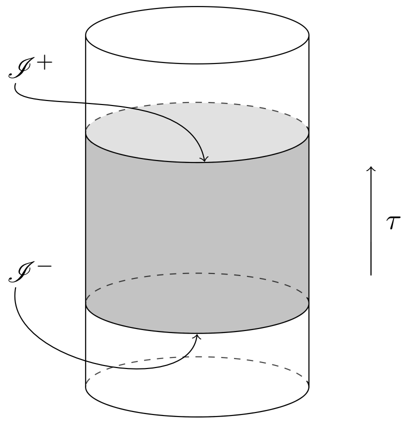

Figure 2.5: Compactified de Sitter space $\widehat{\mathrm{dS}}_4$ embedded in the Einstein cylinder $\mathfrak{E}$.

Remark 2.3.2. The fact that $\mathscr{I}$ is spacelike is a consequence of dS_4 being a solution to the vacuum Einstein equations ((1.1.1) with $\mathbf{T}_{ab} = 0$) with a positive cosmological constant λ,

$$\widetilde{\mathrm{R}}_{ab} = \lambda \tilde{g}_{ab}.$$

Indeed, in general the norm squared *on* $\mathscr{I}$ of the normal to $\mathscr{I}$ is

$$(\nabla_a \Omega)(\nabla^a \Omega) \approx \frac{1}{3}\lambda.$$

In the case of dS_4, $\lambda = 3H^2$, so that $\nabla_a \Omega \nabla^a \Omega \approx H^2 > 0$.

Writing the 3-sphere metric as $\mathfrak{s}_3 = \mathrm{d}\zeta^2 + (\sin^2\zeta)\,\mathfrak{s}_2$ for $\zeta \in [0,\pi]/\sim$ and quotienting the metric $\mathfrak{e}$ by the SO(3) symmetry group of $\mathfrak{s}_2$, we obtain the Penrose diagram for dS_4, fig. 2.6.

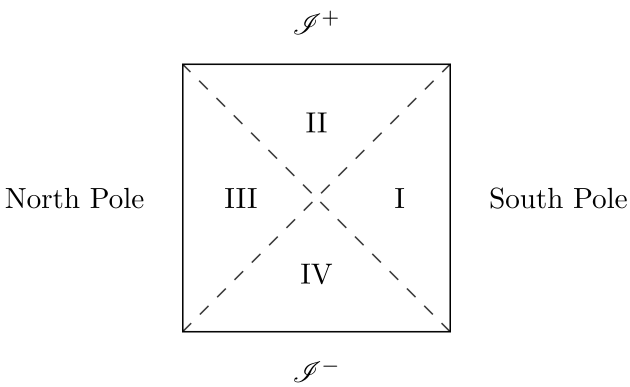

Figure 2.6: The Penrose diagram for dS_4.

The coordinate ζ varies from 0 to π going from left to right, with the vertical lines $\{\zeta = 0\}$ and $\{\zeta = \pi\}$ representing the North Pole and the South Pole of the 3-spheres respectively. The coordinate τ varies from $-\pi/2$ to $\pi/2$ going up, with the horizontal lines $\{\tau = -\pi/2\}$ and $\{\tau = \pi/2\}$ representing past and future null infinities $\mathscr{I}^{\pm}$, as remarked earlier. The dashed lines are the past and future cosmological horizons for an observer at the South Pole: a classical observer sitting at $\{\zeta = \pi\}$ can never observe the region $\mathrm{II} \cup \mathrm{III}$, and can never send a signal to the region $\mathrm{III} \cup \mathrm{IV}$. Thus region I is the region of communications for an observer at the South Pole, while region III is completely inaccessible.

2.4 Two Conformal Scattering Constructions

To introduce the details of conformal scattering, we sketch the construction for the linear wave equation on compactified Minkowski and de Sitter spacetimes.

2.4.1 The Wave Equation on Minkowski Space

Consider the free wave equation

$$\Box\phi = 0 \tag{2.4.1}$$

on Minkowski space $(\mathscr{M} = \mathbb{R}^4, \eta)$, and consider its complete conformal compactification $\hat{\eta} = \Omega^2\eta = \mathfrak{e}$ as outlined in section 2.3.1. Since for η the scalar curvature vanishes and for $\mathfrak{e}$ it is equal to 6, it follows that (2.4.1) is satisfied if and only if the *rescaled* solution

$$\hat{\phi} := \Omega^{-1}\phi \tag{2.4.2}$$

satisfies

$$\widehat{\Box}\hat{\phi} + \hat{\phi} = 0 \tag{2.4.3}$$

(see section 2.2.2 for a description of the conformal wave equation). Consider initial data $(\phi, \partial_t\phi)|_\Sigma = (\phi_0, \phi_1)$ on a Cauchy surface $\Sigma = \{t = 0\} \times \mathbb{R}^3$ in Minkowski space. The conformal transformation (2.4.2) induces initial data

$$\hat{\phi}_0 = \frac{1}{2}(1 + r^2)\phi_0, \qquad \hat{\phi}_1 = \frac{1}{4}(1 + r^2)^2\phi_1$$

on the Cauchy surface $\hat{\Sigma} = \{\tau = 0\} \times \mathbb{S}^3 = \Sigma \cup i^0$, which we assume extends[7] as a pair of $H^1 \oplus L^2$ functions on $\mathbb{S}^3$. The stress-energy tensor for (2.4.3) on $\mathfrak{E}$ is

$$\hat{\mathbf{T}}_{ab} = \hat{\nabla}_a \hat{\phi} \hat{\nabla}_b \hat{\phi} - \frac{1}{2} \mathfrak{e}_{ab} \hat{\nabla}_c \hat{\phi} \hat{\nabla}^c \hat{\phi} + \frac{1}{2} \mathfrak{e}_{ab} \hat{\phi}^2, \tag{2.4.4}$$

which is symmetric and satisfies the conservation law

$$\hat{\nabla}^a \hat{\mathbf{T}}_{ab} = (\widehat{\Box} \hat{\phi} + \hat{\phi}) \hat{\nabla}_b \hat{\phi}.$$

As a result of (2.4.3), $\hat{\mathbf{T}}_{ab}$ is conserved, and gives rise to the conserved current $\hat{J}_b = \hat{T}^a \hat{\mathbf{T}}_{ab}$, where $\hat{T}^a = \partial_\tau$ is a timelike Killing vector field with respect to $\mathfrak{e}$. By integrating $\hat{\nabla}^b \hat{J}_b = 0$ over the compact region bounded by $\hat{\Sigma}$ and $\mathscr{I}^+$ and using the divergence theorem, one obtains an energy estimate

$$\|\hat{\phi}_0\|^2_{H^1(\hat{\Sigma})} + \|\hat{\phi}_1\|^2_{L^2(\hat{\Sigma})} \simeq \|\hat{\phi}^+\|^2_{H^1(\mathscr{I}^+)}, \tag{2.4.5}$$

where the norm $H^1(\mathscr{I}^+)$ is the natural norm on $\mathscr{I}^+$ involving only derivatives which are tangent to $\mathscr{I}^+$, and $\hat{\phi}^+$ is the restriction to $\mathscr{I}^+$ of the solution $\hat{\phi}$ to (2.4.3).

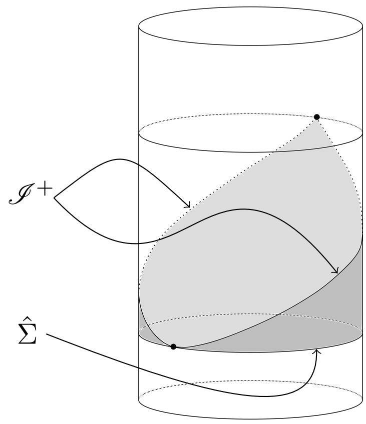

Figure 2.7: We obtain the basic energy estimate (2.4.5) by integrating $\hat{\nabla}_b \hat{J}^b = 0$ in the compact region bounded by $\hat{\Sigma}$ and $\mathscr{I}^+$.

It is classical that for initial data $(\hat{\phi}_0, \hat{\phi}_1) \in H^1(\hat{\Sigma}) \oplus L^2(\hat{\Sigma})$ there exists a finite energy solution $\hat{\phi}$ to (2.4.3) which has a well-defined restriction $\hat{\phi}^+ \in H^1(\mathscr{I}^+)$ on $\mathscr{I}^+$; this defines a bounded linear operator

$$\begin{aligned} \mathfrak{T}^+ : H^1(\hat{\Sigma}) \oplus L^2(\hat{\Sigma}) &\longrightarrow H^1(\mathscr{I}^+), \\ (\hat{\phi}_0, \hat{\phi}_1) &\longmapsto \hat{\phi}^+ \end{aligned} \tag{2.4.6}$$

called the *future trace operator*. This operator is injective by the estimate (2.4.5), and in fact surjective as a consequence of the fact that one can resolve the characteristic Cauchy problem from the data $\hat{\phi}^+ \in H^1(\mathscr{I}^+)$ on $\mathscr{I}^+$ to recover the full finite energy solution $\hat{\phi}$—this is Hörmander's construction for general weakly spacelike Cauchy problems [58], the details of which are given in appendix A.2.2. The operator $\mathfrak{T}^+$ is therefore an isomorphism

[7]This involves an assumption on the rate of decay of the initial data (ϕ_0, ϕ_1), but we do not wish to explore this point at this time. The details of these decay assumptions may be found in chapter 5.

between $H^1(\hat{\Sigma}) \oplus L^2(\hat{\Sigma})$ and $H^1(\mathscr{I}^+)$. Performing the same construction to the past gives the past trace operator $\mathfrak{T}^-$, and one defines the *conformal scattering operator* $\mathscr{S}$ by

$$\begin{gathered}\mathscr{S} : H^1(\mathscr{I}^-) \longrightarrow H^1(\mathscr{I}^+),\\ \mathscr{S} = \mathfrak{T}^+ \circ (\mathfrak{T}^-)^{-1}.\end{gathered}$$

If the rescaled solution $\hat{\phi}$ happens to be sufficiently smooth, the scaling $\hat{\phi} = \Omega^{-1}\phi$ by the conformal factor (2.3.1) allows one to deduce decay rates of the physical solution ϕ along null and timelike directions. Indeed, if $\hat{\phi}$ is at least $\mathcal{C}^0(\hat{\mathscr{M}})$, then it is bounded on $\mathscr{I}$, and the physical solution ϕ satisfies

$$|\phi| = \Omega|\hat{\phi}|.$$

For $u = t - r$ or $v = t + r$ fixed, this leads to decay along null directions in the form

$$\begin{aligned}\lim_{r\to\infty} r\phi(r+u, r, \omega) &= \frac{1}{\sqrt{1+u^2}}\hat{\phi}^+(u,\omega),\\ \lim_{r\to\infty} r\phi(-r+v, r, \omega) &= \frac{1}{\sqrt{1+v^2}}\hat{\phi}^-(v,\omega)\end{aligned}$$

for some continuous functions $\hat{\phi}^\pm \in \mathcal{C}^0(\mathbb{R}\times\mathbb{S}^2)$—exactly the characterstic data on $\mathscr{I}^\pm$ (cf. the radiation fields (1.2.4)). Along timelike directions, keeping r fixed leads to

$$\lim_{t\to\pm\infty} \frac{1}{2}t^2\phi(t,r,\omega) = \hat{\phi}(i^\pm),$$

where $\hat{\phi}(i^\pm)$ are the values, finite as $\hat{\phi} \in \mathcal{C}^0(\hat{\mathscr{M}})$, of the rescaled solution at timelike infinities $i^\pm$.

If the initial data is more regular, one can commute the Killing field $\hat{T}^a = \partial_\tau$ into the equation (2.4.3) and derive higher order estimates for the solution $\hat{\phi}$,

$$\|\hat{\phi}_0\|^2_{H^{m+1}(\hat{\Sigma})} + \|\hat{\phi}_1\|^2_{H^m(\hat{\Sigma})} \simeq \sum_{l=0}^{m} \|\hat{T}^l\hat{\phi}\|^2_{H^1(\mathscr{I}^+)}. \tag{2.4.7}$$

These estimates are in fact a reincarnation of the *peeling* property, originally termed the peeling-off of principal null directions, discovered by Sachs for zero rest mass spin-1 and spin-2 fields [103, 104]. More generally, a zero rest mass field of spin s may be represented as a symmetric rank $2s$ spinor, which possesses $2s$ principal null directions (see proposition 3.5.18 of [95]). Consider the expansion of such a spin-s field in powers of $1/r$ along an outgoing null geodesic; the peeling property is the statement that the part of the field decaying like r^{-m} $(1 \leqslant m \leqslant 2s)$ has $2s - m$ of its principal null directions aligned with the null geodesic along which the original expansion was performed [91]. The peeling behaviour of zero rest mass fields was further studied by Newman and Penrose [87] using what is now known as the Newman–Penrose formalism, as well as the conformal method [93, 94]. It has been shown that the peeling property is generic [77, 78, 88], and in fact equivalent simply to regularity at null infinity in the conformal picture [91], as per the estimates (2.4.7). This equivalence may be seen by expanding $\hat{\phi}$ in a Taylor series around $\mathscr{I}^+$: the order to which one can perform the expansion depends on the number of transverse-to-$\mathscr{I}^+$ derivatives that $\hat{\phi}$ admits, which is precisely what the norm $\|\hat{T}^l\hat{\phi}\|^2_{H^1(\mathscr{I}^+)}$ measures. One makes the following definition.

Definition 2.4.1. The physical solution ϕ to (2.4.1) on Minkowski space $\mathscr{M}$ is said to peel at order $m \in \mathbb{N}$ towards $\mathscr{I}^+$ if

$$\sum_{l=0}^{m} \|\hat{T}^l \hat{\phi}\|^2_{H^1(\mathscr{I}^+)} < \infty,$$

where $\hat{\phi}$ is the rescaled solution on the Einstein cylinder (2.4.2).

The estimate (2.4.7) then shows that the optimal function space of initial data on $\hat{\Sigma}$ giving rise to peeling of order m is precisely the pair of Sobolev spaces $H^{m+1}(\hat{\Sigma}) \oplus H^m(\hat{\Sigma})$.

2.4.2 The Wave Equation on de Sitter Space

A similar construction may be carried out on de Sitter space, where one considers the conformal wave equation on dS_4

$$\tilde{\Box}\tilde{\phi} + 2H^2\tilde{\phi} = 0. \tag{2.4.8}$$

As we have seen, the conformal factor $\Omega = H\cos\tau$ embeds $(\mathrm{dS}_4, \tilde{g})$ into the Einstein cylinder $(\mathfrak{E}, \mathfrak{e})$,

$$\mathfrak{e} = \Omega^2 \tilde{g},$$

and with the associated rescaling of the solution to (2.4.8),

$$\hat{\phi} = \Omega^{-1}\tilde{\phi}, \tag{2.4.9}$$

the equation (2.4.8) becomes the conformal wave equation (2.4.3) on $\mathfrak{E}$, as in the case of compactified Minkowski space. Natural initial data $(\tilde{\phi}_0, \tilde{\phi}_1)$ on $\tilde{\Sigma} \simeq \{\alpha = 0\} \times \mathbb{S}^3$ for (2.4.8) in fact defines initial data for (2.4.3) on $\hat{\Sigma} = \{\tau = 0\} \times \mathbb{S}^3$ without any modification, on account of the fact that the Cauchy surface $\tilde{\Sigma}$ in dS_4 is already a 3-sphere, $\{\alpha = 0\} = \{\tau = 0\}$, and the conformal factor Ω is initially just a constant,

$$\Omega\big|_{\alpha=0} = H.$$

We again use the stress-energy tensor (2.4.4) on the Einstein cylinder, and define the same conserved current $\hat{J}_b = \hat{T}^a \hat{\mathbf{T}}_{ab}$. By integrating $\hat{\nabla}_b \hat{J}^b = 0$ over the region of compactified de Sitter space bounded by $\hat{\Sigma}$ and $\mathscr{I}^+$, we obtain the energy estimate

$$\|\hat{\phi}_0\|^2_{H^1(\hat{\Sigma})} + \|\hat{\phi}_1\|^2_{L^2(\hat{\Sigma})} \simeq \|\hat{\phi}^+\|^2_{H^1(\mathscr{I}^+)} + \|(\hat{T}\hat{\phi})^+\|^2_{L^2(\mathscr{I}^+)}, \tag{2.4.10}$$

where $\hat{\phi}^+$ and $(\hat{T}\hat{\phi})^+$ are the restrictions of $\hat{\phi}$ and $\hat{T}\hat{\phi} = \partial_\tau \hat{\phi}$ to $\mathscr{I}^+$.

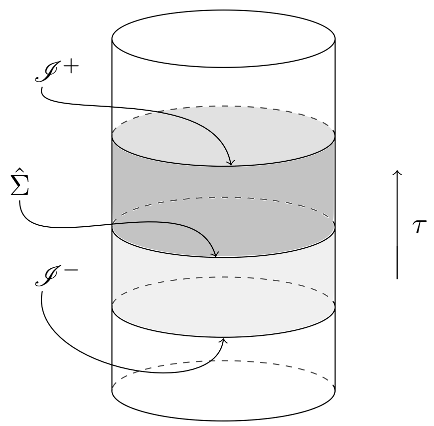

Figure 2.8: We obtain the basic energy estimate (2.4.10) by integrating $\hat{\nabla}_b \hat{J}^b = 0$ in the compact region bounded by $\hat{\Sigma}$ and $\mathscr{I}^+$.

In contrast to the case of compactified Minkowski space where $\mathscr{I}^+$ is null, here the estimate (2.4.10) picks up a time derivative term in the energy on $\mathscr{I}^+$. This is simply a consequence of the fact that here $\mathscr{I}^+$ is spacelike, and the correct set of data on $\mathscr{I}^+$ includes both $\hat{\phi}$ and $\partial_\tau \hat{\phi}$. The reason for why $\partial_\tau \hat{\phi}$ is absent in the null case can be seen in the construction in appendix A.2.2, where one finds that the measure carrying $\partial_\tau \hat{\phi}$ degenerates as a spacelike hypersurface is taken to approach a null hypersurface.

Resolving the Cauchy problem from both $\hat{\Sigma}$ *and* $\mathscr{I}^+$ is trivial in this case, and one constructs the future trace operator

$$\begin{aligned} \mathfrak{T}^+ : H^1(\hat{\Sigma}) \oplus L^2(\hat{\Sigma}) &\longrightarrow H^1(\mathscr{I}^+) \oplus L^2(\mathscr{I}^+), \\ (\hat{\phi}_0, \hat{\phi}_1) &\longmapsto (\hat{\phi}^+, (\hat{T}\hat{\phi})^+) \end{aligned} \tag{2.4.11}$$

as before. The operator $\mathfrak{T}^+$ is then clearly an isomorphism of Hilbert spaces, its invertibility requiring only the resolution of a regular Cauchy problem from $\mathscr{I}^+$ backwards in time. Constructing $\mathfrak{T}^-$ similarly, we define

$$\begin{aligned} \mathscr{S} : H^1(\mathscr{I}^-) \oplus L^2(\mathscr{I}^-) &\longrightarrow H^1(\mathscr{I}^-) \oplus L^2(\mathscr{I}^-), \\ \mathscr{S} &= \mathfrak{T}^+ \circ (\mathfrak{T}^-)^{-1}. \end{aligned}$$

If $\hat{\phi}$ is $\mathcal{C}^0(\widehat{\mathrm{dS}_4})$, then the conformal scaling $\tilde{\phi} = \Omega\hat{\phi}$ gives rise to decay rates of ϕ towards $\mathscr{I}$ as before. Since $\mathscr{I}$ is spacelike here, it is natural to consider decay along timelike directions approaching $\mathscr{I}$. One finds that

$$\lim_{\alpha \to \pm\infty} \mathrm{e}^{H\alpha} \tilde{\phi}(\alpha, \omega) = 2H\hat{\phi}^\pm(\omega)$$

for some functions[8] $\phi^\pm \in \mathcal{C}^0(\mathbb{S}^3)$, which are exactly the restrictions of the rescaled solution $\hat{\phi}$ to $\mathscr{I}^\pm$.

As before, for sufficiently regular data one can also commute the Killing field $\hat{T}$ into the equation (2.4.3) to derive peeling estimates similar to (2.4.7); here, however, because $\mathscr{I}^+$ is spacelike, derivatives transverse to $\mathscr{I}^+$ can be re-expressed in terms of derivatives tangential to $\mathscr{I}^+$ using the evolution equation

$$\partial_\tau^2 \hat{\phi} - \Delta\!\!\!/\, \hat{\phi} + \hat{\phi} = 0.$$

The peeling estimates then become

$$\|\hat{\phi}_0\|^2_{H^{m+1}(\hat{\Sigma})} + \|\hat{\phi}_1\|^2_{H^m(\hat{\Sigma})} \simeq \|\hat{\phi}\|^2_{H^{m+1}(\mathscr{I}^+)} + \|\hat{T}\hat{\phi}\|^2_{H^m(\mathscr{I}^+)}, \tag{2.4.12}$$

and, analogously to the case of Minkowski space, we may make the following definition.

Definition 2.4.2. The physical solution $\tilde{\phi}$ to (2.4.8) on de Sitter space dS_4 is said to peel at order $m \in \mathbb{N}$ if

$$\|\hat{\phi}\|^2_{H^{m+1}(\mathscr{I}^+)} + \|\hat{T}\hat{\phi}\|^2_{H^m(\mathscr{I}^+)} < \infty,$$

where $\hat{\phi}$ is the rescaled solution on the Einstein cylinder (2.4.9).

The estimate (2.4.12) then shows that the optimal function space of initial data for which the solution peels at order m is exactly $H^{m+1}(\hat{\Sigma}) \oplus H^m(\hat{\Sigma})$.

[8] Here ω denotes the coordinates on $\mathbb{S}^3$.

3

Conformal Scattering of Maxwell Potentials

3.1 Introduction

For nonlinear gauge theories the field F_{ab} is insufficient to describe the full dynamics of the system, and one must work with the potential A_a. With the view of approaching scattering questions for more general gauge theories, in this chapter we tackle the problem of conformal scattering of free Maxwell potentials.

As mentioned in remark 2.3.1, it turns out that the natural framework in the Schwarzschild case—or anything that looks like the Schwarzschild solution near i^0—is to forget about a complete compactification of the spacetime, and instead consider a *rescaling* of the spacetime by a conformal factor that is asymptotically equivalent to $\Omega = R \coloneqq r^{-1}$. The solution to the wave equation is then rescaled according to $\hat{\phi} = \Omega^{-1}\phi \approx r\phi$, and one recovers exactly the scaling required to obtain Friedlander's radiation field (1.2.4) on $\mathscr{I}$. Indeed, one can see a hint as to why this should be the correct rescaling by writing out the wave operator in spherical coordinates,

$$\Box\phi = r^{-1}\left((\partial_t + \partial_r)(\partial_t - \partial_r)(r\phi) - r^{-2}\Delta_{\mathbb{S}^2}(r\phi)\right),$$

where the quantity $r\phi$ appears naturally. The conformal factor $\Omega = R$ then brings *null* infinity to a finite region, but leaves i^0 at infinity. One reason why one might wish to construct a scattering theory using such a conformal factor even on flat space, where a complete compactification *is* available, is to be able to compare the results obtained there to results on more general spacetimes, for which a complete compactification will be unavailable. For example, the conformal factor (2.3.1) is *not* uniformly equivalent to r^{-1} near i^0, which makes comparing asymptotics obtained using the two frameworks problematic. A conformal scattering theory for the wave equation (as well as Maxwell and Dirac fields) with $\Omega = R$ near i^0 has been constructed by Mason & Nicolas [78, 79]. Here, we study Maxwell's equations in terms of the potential.

This chapter is divided into two parts. In the first part we construct a complete conformal scattering theory for Maxwell potentials, using a conformal factor such as the one described above, on flat space. Spoiled for choice of Killing and conformally Killing vector fields, here we also investigate the effects of different choices of multiplier vector fields on spaces of scattering data. We discover, for example, that using the Morawetz vector field K_0 [83] yields a scattering theory which is in a certain sense a strict subset of the scattering theory with ∂_t. In the second part we extend the scattering theory of Mason & Nicolas [78] to Maxwell potentials on a large class of curved spacetimes containing matter. Here we must use a strategy similar to the one we employ for the Morawetz vector field on flat space in the first part, where we must solve an elliptic PDE and an ODE to reconstruct the potential from the field.

In this chapter we make heavy use of the Newman–Penrose formalism, which we describe briefly in appendix C.1.

3.2 Geometric and Analytic Framework

3.2.1 Asymptotically Simple and Corvino–Schoen–Chruściel–Delay Spacetimes

In (the second part of) this chapter we work on spacetimes constructed by Chruściel–Delay [26, 27], Corvino [29], and Corvino–Schoen [30]. These are asymptotically flat, asymptotically simple spacetimes with null and timelike infinities of specifiable regularity, which are in addition diffeomorphic to the Schwarzschild or Kerr solution in a neighbourhood of spacelike infinity. These spacetimes are generically non-stationary and may contain matter, providing an ideal testing ground for conformal scattering. As a consequence of their structure near spatial infinity, their conformal compactifications are necessarily singular at i^0. The construction of Chruściel–Delay permits spacetimes with a $\mathcal{C}^k$ conformal compactification for any finite k (however fails to produce spacetimes with a $\mathcal{C}^\infty$ conformal compactification), and in what follows we shall simply assume that a sufficiently large order of differentiability k has been chosen. We will refer to such $\mathcal{C}^k$ differentiability as *smooth*. Smooth asymptotically simple spacetimes are then defined as follows [4, 5, 94, 95, 96].

Definition 3.2.1 (Asymptotically simple spacetimes)**.** Let $(\mathscr{M}, g)$ be a smooth globally hyperbolic spacetime. We say $(\mathscr{M}, g)$ is *asymptotically simple* if there exists another globally hyperbolic spacetime $(\hat{\mathscr{M}}, \hat{g})$ such that

1. the spacetime $\hat{\mathscr{M}}$ is a manifold with boundary $\partial\hat{\mathscr{M}} = \mathscr{I}$, and $\hat{\mathscr{M}} \setminus \mathscr{I}$ is diffeomorphic to $\mathscr{M}$,

2. there exists a smooth function Ω on $\hat{\mathscr{M}}$ such that $\hat{g}_{ab} = \Omega^2 g_{ab}$ and $\Omega > 0$ in $\mathscr{M}$, $\Omega = 0$ on $\mathscr{I}$, and $\mathrm{d}\Omega \neq 0$ on $\mathscr{I}$, and

3. every inextendible null geodesic in $\mathscr{M}$ acquires two distinct endpoints on $\mathscr{I}$.

The condition $\mathrm{d}\Omega \neq 0$ on $\mathscr{I}$ ensures that Ω can be used as a coordinate on $\hat{\mathscr{M}}$, in particular to perform Taylor expansions to capture the decay of fields near $\mathscr{I}$. The above definition extends the notion of conformal compactification used in sections 2.3.1 and 2.3.2 to a much larger class of *curved* spacetimes. As already alluded to in section 2.3.1, the last point in the above definition warrants the following terminology.

Definition 3.2.2. The boundary $\mathscr{I}$ of the compactified spacetime $\hat{\mathscr{M}}$ is called *null infinity*.

If $\mathscr{M}$ happens to be vacuum (in fact it is enough that the trace of the matter stress-energy tensor vanishes asymptotically near $\mathscr{I}$), then as a hypersurface of the unphysical spacetime $\hat{\mathscr{M}}$, $\mathscr{I}$ is spacelike, null, or timelike, depending on whether the cosmological constant λ is positive, zero, or negative respectively. We focus in this chapter on the case $\lambda = 0$. The spacetimes of Corvino–Schoen–Chruściel–Delay are asymptotically simple, and in addition to the conditions of definition 3.2.1 satisfy the following.

4. The physical spacetime $(\mathscr{M}, g_{ab})$ satisfies Einstein's equations with $\lambda = 0$,

$$\mathrm{R}_{ab} - \frac{1}{2}\mathrm{R}g_{ab} = -8\pi\gamma\mathbf{T}_{ab},$$

   where $\Omega^{-2}\mathbf{T}_{ab}$ has a smooth limit on $\mathscr{I}$,

5. the boundary $\mathscr{I}$ of $\hat{\mathscr{M}}$ is the union of three points i^0, i^-, and i^+, and two smooth null hypersurfaces $\mathscr{I}^+$ and $\mathscr{I}^-$; the hypersurface $\mathscr{I}^+$ is the past lightcone of i^+, whereas $\mathscr{I}^-$ is the future lightcone of i^-,

6. the metric $\hat{g}_{ab}$ is smooth at $i^\pm$ and $\mathscr{I}^\pm$, and

7. the physical spacetime $\mathscr{M}$ is diffeomorphic to the Schwarzschild or Kerr solution outside the domain of influence of a given compact subset of a Cauchy surface Σ.

The points $i^\pm$ are called future and past timelike infinities, the point i^0 is called spacelike infinity, and the null hypersurfaces $\mathscr{I}^\pm$ are called future and past null infinities respectively. Spacelike infinity i^0 will generically be a singular point of the conformally rescaled metric. Condition 4 ensures that the matter fields in the physical spacetime $\mathscr{M}$ decay sufficiently fast at infinity. We call spacetimes satisfying the above decay condition for the matter fields *asymptotically vacuum*.

3.2.2 Orthogonal Decompositions

In order to pose the Cauchy problem for Maxwell's equations, we perform an orthogonal $3+1$ decomposition of the physical and rescaled metrics g_{ab} and $\hat{g}_{ab}$. Since $\mathscr{M}$ is globally hyperbolic, there exists a smooth time function $t : \mathscr{M} \to \mathbb{R}$ such that $\nabla^a t$ is uniformly timelike on $\mathscr{M}$, where ∇ is the Levi–Civita connection of g. The level sets $\{\Sigma_t\}_t$ of t define a uniformly spacelike foliation of $\mathscr{M}$. If $\mathscr{M}$ is diffeomorphic to $\mathbb{R}^4$, then each Σ_t is diffeomorphic to some $\Sigma \simeq \mathbb{R}^3$, and the flow of the vector field $\nabla^a t$ effects the identification $\mathscr{M} \simeq \mathbb{R}_t \times \Sigma$. The metric g_{ab} then decomposes as

$$\mathrm{d}s^2 = g_{ab}\,\mathrm{d}x^a\,\mathrm{d}x^b = N^2\mathrm{d}t^2 - h,$$

where h is a smooth Riemannian metric on Σ_t for each t, and N is a smooth non-vanishing *lapse* function. The unit normal to the hypersurfaces Σ_t is

$$T^a = \frac{1}{N}\frac{\partial}{\partial t}, \quad \text{i.e.} \quad T_a\mathrm{d}x^a = N\mathrm{d}t,$$

so the metric can be written as

$$g_{ab} = T_aT_b - h_{ab}.$$

The Levi–Civita connection of g_{ab} decomposes as

$$\nabla_a = T_a \nabla_T + \nabla_a^{\perp},$$

where $\nabla_a^{\perp} = -h_a^b \nabla_b$ is the part of ∇_a orthogonal to T^a, $T^a \nabla_a^{\perp} = 0$. It is the 4-dimensional covariant derivative ∇ projected onto Σ_t, and differs from the Levi–Civita connection $\boldsymbol{\nabla}$ of $(\Sigma_t, h(t))$ by the extrinsic curvature κ_{ab} of Σ_t. Indeed,

$$\nabla_a^{\perp} T_b = -h_a^c \nabla_c T_b = \kappa_{ab} = \kappa_{(ab)},$$

so that for any X_a such that $T^a X_a = 0$

$$\boldsymbol{\nabla}_a X_b - \nabla_a^{\perp} X_b = \kappa_a{}^c T_b X_c.$$

We define the trace of the extrinsic curvature by

$$\operatorname{Tr} \kappa = \kappa^a{}_a = -h^{ab} \nabla_a T_b.$$

A similar decomposition may be performed for the conformally rescaled metric $\hat{g}_{ab} = \Omega^2 g_{ab}$. Here we choose a smooth time function τ such that $\hat{\nabla}^a \tau$ is uniformly timelike and such that $\tau(i^{\pm}) = \pm\tau_{\max}$, $0 < \tau_{\max} < \infty$, where $\hat{\nabla}$ is the Levi–Civita connection of $\hat{g}$. The level sets $\{\hat{\Sigma}_\tau\}_\tau$ of τ define a uniformly spacelike foliation of $\hat{\mathscr{M}}$ such that the leaves $\hat{\Sigma}_\tau$ are transverse to $\mathscr{I}$, and, as $\tau \to \pm\tau_{\max}$, shrink to the points $i^{\pm}$. With respect to this foliation the rescaled metric decomposes as

$$\hat{g}_{ab} = \hat{T}_a \hat{T}_b - \hat{h}_{ab}, \qquad \hat{g}_{ab}\, \mathrm{d}x^a\, \mathrm{d}x^b = \hat{N}^2 \mathrm{d}\tau^2 - \hat{h},$$

where $\hat{T}^a$ is the unit normal to $\hat{\Sigma}_\tau$ with respect to $\hat{g}_{ab}$, and $\hat{h}_{ab}$ is a smooth Riemannian metric on $\hat{\Sigma}_\tau$ for each τ. As before, the Levi–Civita connection $\hat{\nabla}$ of $\hat{g}_{ab}$ decomposes as

$$\hat{\nabla}_a = \hat{T}_a \hat{\nabla}_{\hat{T}} + \hat{\nabla}_a^{\perp}.$$

We assume that the functions t and τ are such that the initial leaf of the rescaled foliation $\{\hat{\Sigma}_\tau\}_\tau$ agrees with the initial leaf of the physical foliation $\{\Sigma_t\}_t$, $\hat{\Sigma}_0 = \Sigma_0$. The vector fields $\hat{T}^a$ and T^a are therefore parallel on $\hat{\Sigma}_0 = \Sigma_0 =: \Sigma$, and the above decomposition of the metric gives the relation

$$\hat{T}^a|_\Sigma = \Omega^{-1} T^a|_\Sigma.$$

We also assume that the time derivative of the conformal factor vanishes on Σ,

$$\partial_t \Omega|_\Sigma \propto \partial_\tau \Omega|_\Sigma = 0.$$

The uniformly spacelike foliation $\{\Sigma_t\}_t$ of the physical spacetime extends to an asymptotically null foliation of the rescaled spacetime. Indeed, the unit normal T^a with respect to g_{ab} has norm Ω^2 with respect to $\hat{g}_{ab}$, which tends to zero as $\Omega \to 0$. Conversely, the uniformly spacelike foliation $\{\hat{\Sigma}_\tau\}_\tau$ of the rescaled spacetime corresponds to a foliation of the physical spacetime by hyperboloids which are asymptotically null.

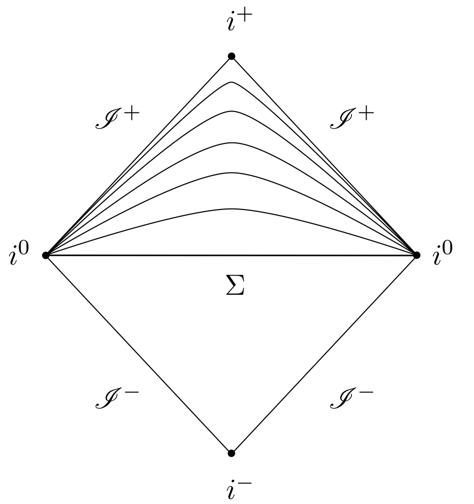

Figure 3.1: The asymptotically null foliation $\{\Sigma_t\}_t$ of $\hat{\mathscr{M}}$.

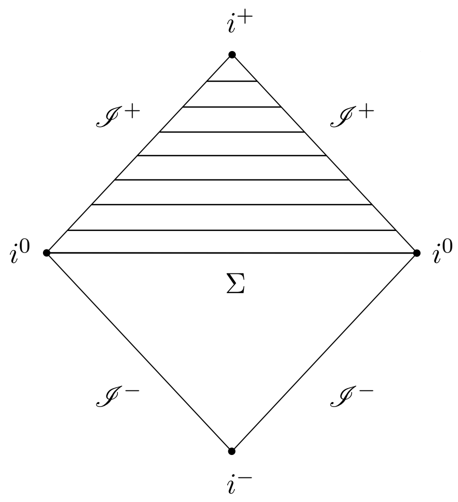

Figure 3.2: The foliation $\{\hat{\Sigma}_\tau\}_\tau$ of $\hat{\mathscr{M}}$ whose leaves are transverse to $\mathscr{I}^+$.

The foliation $\{\hat{\Sigma}_\tau\}_\tau$ will be used to pose and solve the Cauchy problem in the rescaled spacetime $\hat{\mathscr{M}}$.

We define the projection onto (Σ_t, h_{ab}) of a 1-form A_a on $\mathscr{M}$ by

$$\mathbf{A}_\alpha := -h^a_\alpha A_a.$$

The Σ_t-covariant derivative $\boldsymbol{\nabla}_\alpha$ applied to $\mathbf{A}_\beta$ is then given by

$$\boldsymbol{\nabla}_\alpha \mathbf{A}_\beta = h^b_\beta h^a_\alpha \nabla_a \mathbf{A}_b = -h^b_\beta h^a_\alpha \nabla_a (h^c_b A_c),$$

and more generally the Σ_t-covariant derivative of a tensor field $T^{a_1\dots a_n}{}_{b_1\dots b_m}$ is

$$\boldsymbol{\nabla}_\gamma T^{\alpha_1\dots\alpha_n}{}_{\beta_1\dots\beta_m} = (-1)^{n+m+1} h^c_\gamma h^{\alpha_1}_{a_1} \dots h^{\alpha_n}_{a_n} h^{b_1}_{\beta_1} \dots h^{b_m}_{\beta_m} \nabla_c T^{a_1\dots a_n}{}_{b_1\dots b_m}.$$

The factor of $(-1)^{n+m+1}$ is included to compensate for the successive changes of signature each time the projector h^a_b is applied: note that $h^{ac}h_{cb} = -h^a_b = \delta^a_b - T^a T_b$. For tensors on Σ_t, $-h^a_b$ of course acts as δ^a_b since all tensors on Σ_t are orthogonal to T^a with respect to the full metric g_{ab}.

3.2.3 The Schwarzschildean Neighbourhood of Spacelike Infinity

The spacetimes of Corvino–Schoen–Chruściel–Delay are diffeomorphic to the Schwarzschild (or Kerr) spacetime in a neighbourhood of i^0. For simplicity, we consider the Schwarzschild case. The metric near i^0 is then given by

$$g_{ab}\,\mathrm{d}x^a\,\mathrm{d}x^b = F(r)\,\mathrm{d}t^2 - F(r)^{-1}\mathrm{d}r^2 - r^2\mathfrak{s}_2, \tag{3.2.1}$$

where $F(r) = 1 - 2mr^{-1}$, with inverse metric

$$g^{ab}\partial_a \odot \partial_b = F(r)^{-1}\partial_t^2 - F(r)\,\partial_r^2 - r^{-2}\mathfrak{s}_2^{-1}.$$

We define the Eddington–Finkelstein coordinates

$$u = t - r_*, \qquad r_* = r + 2m\log\left(\frac{r}{2m} - 1\right),$$

and the inverted radial coordinate

$$R = \frac{1}{r}.$$

The metric (3.2.1) in the coordinates (u, r, θ, ϕ) becomes

$$g_{ab}\,\mathrm{d}x^a\,\mathrm{d}x^b = F(r)\,\mathrm{d}u^2 + 2\,\mathrm{d}u\,\mathrm{d}r - r^2\mathfrak{s}_2,$$

with the inverse metric

$$g^{ab}\partial_a \odot \partial_b = 2\,\partial_u \odot \partial_r - F(r)\,\partial_r^2 - r^{-2}\mathfrak{s}_2^{-1}.$$

For the conformally rescaled metric $\hat{g}_{ab} = \Omega^2 g_{ab}$, we choose the conformal factor Ω so that it is equal or equivalent to R near i^0; with $\Omega = R$ the rescaled metric becomes

$$\hat{g}_{ab}\,\mathrm{d}x^a\,\mathrm{d}x^b = R^2F(R)\,\mathrm{d}u^2 - 2\,\mathrm{d}u\,\mathrm{d}R - \mathfrak{s}_2, \tag{3.2.2}$$

where $F(R) = 1 - 2mR = F(r)$. The rescaled inverse metric is then

$$\hat{g}^{ab}\partial_a \odot \partial_b = -2\,\partial_u \odot \partial_R - R^2F(R)\,\partial_R^2 - \mathfrak{s}_2^{-1}.$$

The rescaled volume form is given by $\widehat{\mathrm{dv}} = \mathrm{d}u \wedge \mathrm{d}R \wedge \mathrm{dv}_{\mathfrak{s}_2}$.

3.2.4 Newman–Penrose Tetrads

We will work with two related Newman–Penrose (NP) tetrads, one associated to the physical spacetime $\mathscr{M}$, and another to the rescaled spacetime $\hat{\mathscr{M}}$[1]. On the physical spacetime $\mathscr{M}$ we define an NP tetrad $(l, n, m, \bar{m})$ by aligning l^a and n^a with incoming and outgoing null congruences respectively such that wherever the metric g_{ab} agrees with the Schwarzschild metric, the tetrad $(l, n, m, \bar{m})$ takes the concrete form

$$l^a = \partial_u - \frac{1}{2}F(r)\partial_r, \qquad l_a = \frac{1}{2}F(r)\,\mathrm{d}u + \mathrm{d}r, \tag{3.2.3}$$

$$n^a = \partial_r, \qquad n_a = \mathrm{d}u, \tag{3.2.4}$$

$$m^a = \frac{1}{\sqrt{2}r}\left(\partial_\theta + \frac{i}{\sin\theta}\partial_\phi\right), \qquad m_a = -\frac{r}{\sqrt{2}}(\mathrm{d}\theta + i\sin\theta\,\mathrm{d}\phi), \tag{3.2.5}$$

$$\bar{m}^a = \frac{1}{\sqrt{2}r}\left(\partial_\theta - \frac{i}{\sin\theta}\partial_\phi\right), \qquad \bar{m}_a = -\frac{r}{\sqrt{2}}(\mathrm{d}\theta - i\sin\theta\,\mathrm{d}\phi), \tag{3.2.6}$$

and extending m^a as a $\mathcal{C}^k$-smooth complex null vector everywhere orthogonal to l^a and n^a. We assume that the vector fields l^a and n^a are $\mathcal{C}^k$-smooth and real. We obtain the rescaled NP tetrad on $\hat{\mathscr{M}}$ by the rescaling

$$\hat{l}^a = l^a, \qquad \hat{l}_a = \Omega^2 l_a, \tag{3.2.7}$$

$$\hat{n}^a = \Omega^{-2}n^a, \qquad \hat{n}_a = n_a, \tag{3.2.8}$$

$$\hat{m}^a = \Omega^{-1}m^a, \qquad \hat{m}_a = \Omega m_a, \tag{3.2.9}$$

$$\hat{\bar{m}}^a = \Omega^{-1}\bar{m}^a, \qquad \hat{\bar{m}}_a = \Omega\bar{m}_a. \tag{3.2.10}$$

[1]We secretly assume that the spacetimes $\mathscr{M}$ and $\hat{\mathscr{M}}$ admit spinor structures; that is, they admit spin structures and are space- and time-orientable. This is a mild topological restriction that we do not concern ourselves with.

We assume that $\hat{l}^a$ restricted to $\mathscr{I}^+$ is a generator of $\mathscr{I}^+$ and that $\hat{\nabla}^a\Omega$ is proportional to $\hat{l}^a$ on $\mathscr{I}^+$. We also assume that the vector fields (and their corresponding 1-forms) $(\hat{l}, \hat{n}, \hat{m}, \bar{\hat{m}})$ are all $\mathcal{C}^k$-smooth throughout $\hat{\mathscr{M}}$. Furthermore, we suppose that $\hat{n}_a \neq 0$ on $\mathscr{I}^+$, and use it to define a measure on $\mathscr{I}^+$:

$$\widehat{\mathrm{dv}}_{\mathscr{I}^+} := \hat{n}^\flat \wedge (i\hat{m}^\flat \wedge \bar{\hat{m}}^\flat).$$

In the Schwarzschild sector this will of course be given by $\widehat{\mathrm{dv}}_{\mathscr{I}^+} = \mathrm{d}u \wedge \mathrm{dv}_{\mathfrak{S}_2}$. Finally, we assume that the vector field T^a is given by

$$T^a = al^a + \frac{1}{2a}n^a \tag{3.2.11}$$

for some non-negative function a on $\mathscr{M}$. Note that (3.2.11) is simply the assumption that T^a is independent of the angular vector fields m^a and $\bar{m}^a$, as the coefficient in front of n^a is fixed by normalization. In terms of the rescaled tetrad T^a is then given by

$$T^a = a\hat{l}^a + \frac{\Omega^2}{2a}\hat{n}^a,$$

and becomes proportional to the generator $\hat{l}^a$ of $\mathscr{I}^+$ on $\mathscr{I}^+$. In the Schwarzschild sector we explicitly have that the unit normal to surfaces Σ_t of constant t is $T^a = F(r)^{-1/2}\partial_t$, which in terms of the physical NP tetrad is given by

$$T^a = F(r)^{-1/2}l^a + \frac{1}{2}F(r)^{1/2}n^a.$$

In the rescaled Schwarzschild coordinates (u, R, θ, ϕ) the rescaled tetrad $(\hat{l}, \hat{n}, \hat{m}, \bar{\hat{m}})$ takes the form

$$\hat{l}^a = \partial_u + \frac{1}{2}R^2F(R)\partial_R, \qquad \hat{l}_a = \frac{1}{2}R^2F(R)\,\mathrm{d}u - \mathrm{d}R, \tag{3.2.12}$$

$$\hat{n}^a = -\partial_R, \qquad \hat{n}_a = \mathrm{d}u, \tag{3.2.13}$$

$$\hat{m}^a = \frac{1}{\sqrt{2}}\left(\partial_\theta + \frac{i}{\sin\theta}\partial_\phi\right), \qquad \hat{m}_a = -\frac{1}{\sqrt{2}}(\mathrm{d}\theta + i\sin\theta\,\mathrm{d}\phi), \tag{3.2.14}$$

$$\bar{\hat{m}}^a = \frac{1}{\sqrt{2}}\left(\partial_\theta - \frac{i}{\sin\theta}\partial_\phi\right), \qquad \bar{\hat{m}}_a = -\frac{1}{\sqrt{2}}(\mathrm{d}\theta - i\sin\theta\,\mathrm{d}\phi). \tag{3.2.15}$$

Of course, by setting $m = 0 \iff F(r) \equiv 1$ in (3.2.3)–(3.2.6), one recovers an NP tetrad for physical Minkowski space. Similarly, setting $F(R) \equiv 1$ in (3.2.12)–(3.2.15) gives an NP tetrad for rescaled Minkowski space wherever $\Omega = R$.

Maxwell Components

We define the components of the physical Maxwell potential A_a and the physical Maxwell field F_{ab} with respect to the physical NP tetrad $(l, n, m, \bar{m})$ by

$$A_0 := l^aA_a, \qquad A_1 := n^aA_a, \qquad A_2 := m^aA_a$$

and

$$F_0 := F_{ab}l^am^b, \qquad F_1 := \frac{1}{2}F_{ab}(l^an^b + \bar{m}^am^b), \qquad F_2 := F_{ab}\bar{m}^an^b.$$

We denote

$$\mathfrak{a} := T^a A_a,$$

and define the electric and magnetic fields with respect to the foliation (Σ_t, h_{ab}) by

$$E_a := T^b F_{ba} = -h_a^\alpha E_\alpha =: \mathbf{E}_a$$

and

$$B_a := \frac{1}{2}\varepsilon_a{}^{bc} F_{bc} = -\frac{1}{2}\varepsilon_{\alpha\beta\gamma} h_a^\alpha h_b^\beta h_c^\gamma F_{bc} = -h_a^\alpha B_\alpha =: \mathbf{B}_a,$$

where ε_{abc} is the volume form of h_{ab}. The components of the rescaled Maxwell potential $\hat{A}_a$ and rescaled Maxwell field $\hat{F}_{ab}$ with respect to $(\hat{l}, \hat{n}, \hat{m}, \bar{\hat{m}})$, as well as $\hat{\mathbf{E}}_a$ and $\hat{\mathbf{B}}_a$, are defined in the same way.

One can check that the components of the stress-energy tensor $\mathbf{T}_{ab}$ with respect to the tetrad $(l, n, m, \bar{m})$ are given by

$$\mathbf{T}_{ab} l^a l^b = 2|F_0|^2, \qquad \mathbf{T}_{ab} l^a n^b = 2|F_1|^2, \qquad \mathbf{T}_{ab} n^a n^b = 2|F_2|^2,$$

and similarly for the rescaled stress-energy tensor $\hat{\mathbf{T}}_{ab}$ with respect to the rescaled tetrad $(\hat{l}, \hat{n}, \hat{m}, \bar{\hat{m}})$. The components of the Maxwell field F_{ab} are given in terms of the components of the Maxwell potential A_a by

$$F_0 = (\text{þ} - \bar{\rho})A_2 + \kappa A_1 - (\eth + \bar{\pi})A_0 - \sigma \bar{A}_2, \tag{3.2.16}$$

$$F_1 = \frac{1}{2}\big(- (\text{þ}' - \mu + \bar{\mu})A_0 + (\text{þ} + \rho - \bar{\rho})A_1 + (\bar{\eth} - \bar{\tau} - \pi)A_2 - (\eth + \tau + \bar{\pi})\bar{A}_2\big), \tag{3.2.17}$$

$$F_2 = -(\text{þ}' + \bar{\mu})\bar{A}_2 - \nu A_0 + (\bar{\eth} + \bar{\tau})A_1 - \lambda A_2. \tag{3.2.18}$$

3.3 Partially Compactified Minkowski Space

In this simplest case we work on Minkowski space $(\mathscr{M} = \mathbb{R}^4, \eta)$,

$$\eta = \mathrm{d}t^2 - \mathrm{d}r^2 - r^2 \mathfrak{s}_2, \tag{3.3.1}$$

where the coordinates t and r range over $\mathbb{R}$ and $[0, \infty)$ respectively. We define the retarded[2] Bondi coordinate

$$u := t - r.$$

In terms of the new coordinates (u, r, θ, ϕ) the metric η becomes

$$\eta = \mathrm{d}u^2 + 2\,\mathrm{d}u\,\mathrm{d}r - r^2 \mathfrak{s}_2. \tag{3.3.2}$$

We consider the conformally rescaled unphysical metric

$$\hat{\eta}_{ab} := \Omega^2 \eta_{ab},$$

[2]The coordinate u is adapted for constructing future null infinity $\mathscr{I}^+$, but a similar construction can be performed with the advanced coordinate $v = t + r$ to construct $\mathscr{I}^-$.

where Ω is a smooth radial conformal factor such that $\Omega = R$ near $\mathscr{I}^+$ and $\Omega = R\chi(u)$ near i^+ for some smooth rapidly decaying function $\chi(u)$.

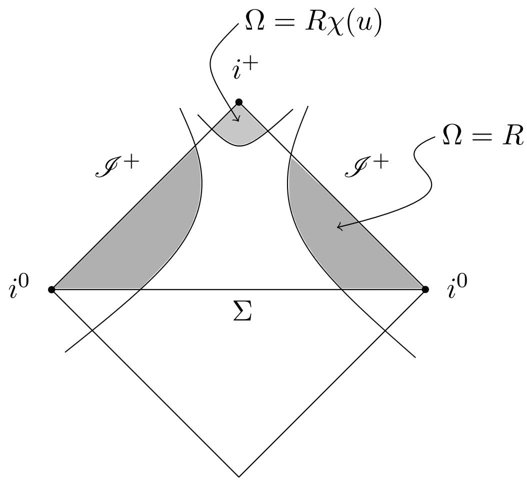

Figure 3.3: We choose a smooth conformal factor Ω which is equal to R near $\mathscr{I}^+$ and brings i^+ to a finite distance.

It should be understood that in the white region near $\mathscr{I}^+$ in fig. 3.3 the conformal factor Ω interpolates smoothly between R and $R\chi(u)$. Near $\mathscr{I}^+$ the rescaled metric then takes the form

$$\hat{\eta} = \chi^2 R^2 \mathrm{d}u^2 - 2\chi^2 \,\mathrm{d}u\,\mathrm{d}R - \chi^2 \mathfrak{s}_2, \tag{3.3.3}$$

and is now regular (although degenerate) at $R = 0$, unlike the physical metric η at $r = \infty$. Clearly the set $\mathscr{I} = \{R = 0\}$ defines a boundary of the spacetime $(\mathscr{M}, \hat{\eta})$, and we define $\hat{\mathscr{M}} := \mathscr{M} \cup \mathscr{I}$. By construction, $\hat{\mathscr{M}} \setminus \mathscr{I}$ is diffeomorphic to $\mathscr{M}$. Future null infinity $\mathscr{I}^+$ is given by

$$\mathscr{I}^+ = \mathbb{R}_u \times \{R = 0\} \times \mathbb{S}^2,$$

which is topologically the open cylinder $\mathbb{R} \times \mathbb{S}^2$. The metric $\hat{\eta}$ is degenerate on $\mathscr{I}^+$ (it has signature $(0, -, -)$, as can be seen from (3.3.3)), but one may integrate over $\mathscr{I}^+$ with respect to the measure $\widehat{\mathrm{dv}}_{\mathscr{I}^+} = \chi^2 \,\mathrm{d}u \wedge \mathrm{dv}_{\mathfrak{s}_2}$. Since ∂_u is tangent to $\mathscr{I}^+$ and $\hat{\eta}(\partial_u, \partial_R) = -\chi^2$, the vector field ∂_R is transverse to $\mathscr{I}^+$. In the (u, R)-plane spacelike infinity is the point given by $i^0 = (-\infty, 0)$, and future timelike infinity is given by $i^+ = (\infty, 0)$. The inverse metric to (3.3.3) is

$$\hat{\eta}^{-1} = -2\chi^{-2}\partial_u \odot \partial_R - \chi^{-2}R^2\partial_R \otimes \partial_R - \chi^{-2}\mathfrak{s}_2^{-1}.$$

Remark 3.3.1. The purpose of choosing the conformal factor to be

$$\Omega = R\chi(u) \tag{3.3.4}$$

near i^+ is to bring i^+ to a finite distance with respect to $\hat{\eta}$ from the interior of $\hat{\mathscr{M}}$. For this to be true the function $\chi(u)$ needs to decay sufficiently fast as $u \to \infty$, however it is clear that plenty such functions exist. With i^+ a finite point in $\hat{\mathscr{M}}$, $\mathscr{I}^+$ is now the backwards lightcone of i^+ in $\hat{\mathscr{M}}$. This will be crucial for the solution of the Goursat problem.

A Newman–Penrose tetrad $(l, n, m, \bar{m})$ for the metric (3.3.2) is given by setting $F(r) = 1 \iff m = 0$ in (3.2.3)–(3.2.6). In this case the normal $T^a = \partial_t = \partial_u$ to the surfaces Σ_t of constant t becomes simply

$$T^a = l^a + \frac{1}{2}n^a.$$

We obtain a Newman–Penrose tetrad $(\hat{l}, \hat{n}, \hat{m}, \bar{\hat{m}})$ on $\hat{\mathscr{M}}$ by the rescaling (3.2.7)–(3.2.10), so that near $\mathscr{I}^+$

$$\hat{l}^a = \partial_u + \frac{1}{2}R^2\partial_R, \qquad \hat{l}_a = \frac{1}{2}\chi^2R^2\mathrm{d}u - \chi^2\mathrm{d}R,$$
$$\hat{n}^a = -\chi^{-2}\partial_R, \qquad \hat{n}_a = \mathrm{d}u,$$
$$\hat{m}^a = \frac{1}{\sqrt{2}\chi}\left(\partial_\theta + \frac{i}{\sin\theta}\partial_\phi\right), \qquad \hat{m}_a = -\frac{\chi}{\sqrt{2}}(\mathrm{d}\theta + i\sin\theta\,\mathrm{d}\phi),$$
$$\bar{\hat{m}}^a = \frac{1}{\sqrt{2}\chi}\left(\partial_\theta - \frac{i}{\sin\theta}\partial_\phi\right), \qquad \bar{\hat{m}}_a = -\frac{\chi}{\sqrt{2}}(\mathrm{d}\theta - i\sin\theta\,\mathrm{d}\phi).$$

3.3.1 A Priori Energy Estimates

The volume form on the rescaled spacetime $\hat{\mathscr{M}}$ is given by

$$\widehat{\mathrm{dv}} = \hat{l}^\flat \wedge \hat{n}^\flat \wedge (i\hat{m}^\flat \wedge \bar{\hat{m}}^\flat),$$

and, where $\Omega = R\chi$, is explicitly given by $\widehat{\mathrm{dv}} = \chi^4\,\mathrm{d}u\wedge\mathrm{d}R\wedge\mathrm{dv}_{\mathfrak{s}_2}$. For the multiplier vector field

$$K^a = \partial_t = \partial_u = l^a + \frac{1}{2}n^a = \hat{l}^a + \frac{1}{2}\Omega^2\hat{n}^a$$

we compute the energy density 3-form

$$\begin{aligned} K^a\hat{\mathbf{T}}_a{}^b\partial_b \lrcorner \widehat{\mathrm{dv}} &= \left(\hat{l}^a + \frac{1}{2}\Omega^2\hat{n}^a\right)\hat{\mathbf{T}}_{ac}(\hat{l}^b\hat{n}^c + \hat{l}^c\hat{n}^b - \hat{m}^c\bar{\hat{m}}^b - \bar{\hat{m}}^c\hat{m}^b)\partial_b \lrcorner \widehat{\mathrm{dv}} \\ &= -\left(2|\hat{F}_1|^2 + \Omega^2|\hat{F}_2|^2\right)\hat{l}^\flat \wedge (i\hat{m}^\flat \wedge \bar{\hat{m}}^\flat) + \left(2|\hat{F}_0|^2 + \Omega^2|\hat{F}_1|^2\right)\hat{n}^\flat \wedge (i\hat{m}^\flat \wedge \bar{\hat{m}}^\flat) \\ &\quad + \dots, \end{aligned}$$

where the ellipsis represents contractions of $\widehat{\mathrm{dv}}$ with either $\hat{m}$ or $\bar{\hat{m}}$. One immediately reads off the energy on $\mathscr{I}^+$,

$$\mathcal{E}_{\mathscr{I}^+}[\hat{F}] \simeq \int_{\mathscr{I}^+} |\hat{F}_0|^2\,\widehat{\mathrm{dv}}_{\mathscr{I}^+} = \|\hat{F}_0\|^2_{L^2(\mathscr{I}^+)}. \tag{3.3.5}$$

Similarly, on the initial surface $\Sigma = \{t = 0\}$ whose unit normal with respect to $\hat{\eta}$ is

$$\hat{T}^a\Big|_\Sigma = \left(\Omega^{-1}\hat{l}^a + \frac{1}{2}\Omega\hat{n}^a\right)\Big|_\Sigma,$$

the energy density 3-form is

$$\begin{aligned} K^a\hat{\mathbf{T}}_a{}^b\partial_b \lrcorner \widehat{\mathrm{dv}}|_\Sigma &= K^a\hat{T}^b\hat{\mathbf{T}}_{ab}(\hat{T} \lrcorner \widehat{\mathrm{dv}})|_\Sigma \\ &= \left(\hat{l}^a + \frac{1}{2}\Omega^2\hat{n}^a\right)\left(\frac{1}{\Omega}\hat{l}^b + \frac{1}{2}\Omega\hat{n}^b\right)\hat{\mathbf{T}}_{ab}\,\widehat{\mathrm{dv}}_\Sigma \\ &= \left(\frac{2}{\Omega}|\hat{F}_0|^2 + 2\Omega|\hat{F}_1|^2 + \frac{\Omega^3}{2}|\hat{F}_2|^2\right)\widehat{\mathrm{dv}}_\Sigma, \end{aligned}$$

so that

$$\mathcal{E}_\Sigma[\hat{F}] \simeq \int_\Sigma \left(\Omega^{-1}|\hat{F}_0|^2 + \Omega|\hat{F}_1|^2 + \Omega^3|\hat{F}_2|^2\right) \widehat{\mathrm{dv}}_\Sigma, \tag{3.3.6}$$

where $\widehat{\mathrm{dv}}_\Sigma = (\hat{T} \lrcorner \widehat{\mathrm{dv}})|_\Sigma$.

Theorem 3.3.2. *For smooth compactly supported Maxwell initial data on Σ one has the energy estimate*

$$\mathcal{E}_{\mathscr{I}^+} = \mathcal{E}_\Sigma. \tag{3.3.7}$$

Proof. For smooth compactly supported initial data on Σ the solution $\hat{F}_{ab} = F_{ab}$ is smooth on $\hat{\mathscr{M}}$ and is supported away from i^0 by finite speed of propagation, as depicted in fig. 3.4.

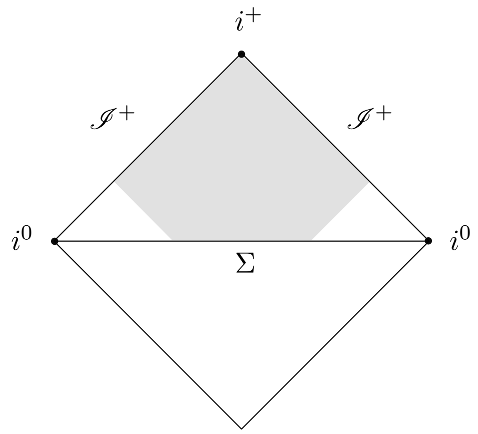

Figure 3.4: For compactly supported data the support of the solution remains supported away from spacelike infinity i^0.

We integrate the divergence

$$\hat{\nabla}^b(K^a\hat{\mathbf{T}}_{ab}) = \hat{\nabla}^{(b}K^{a)}\hat{\mathbf{T}}_{ab} + K^a\hat{\nabla}^b\hat{\mathbf{T}}_{ab}$$

over the region bounded by Σ and $\mathscr{I}^+$. Since the stress-energy tensor $\hat{\mathbf{T}}_{ab}$ is conserved and $K^a = \partial_t$ is conformally[3] Killing on the rescaled spacetime $(\hat{\mathscr{M}}, \hat{\eta})$, one sees that the current $\hat{J}_b = K^a\hat{\mathbf{T}}_{ab}$ is exactly conserved,

$$\hat{\nabla}^b\hat{J}_b = 0,$$

as $\hat{\mathbf{T}}_a{}^a = 0$. Although the region is not compact in the i^0 direction, the solution $\hat{F}_{ab}$ is compactly supported, so we have

$$0 = -\mathcal{E}_\Sigma[\hat{F}] + \mathcal{E}_{\mathscr{I}^+}[\hat{F}]. \tag{3.3.8}$$

To be completely rigorous, one may see the above by introducing, say, a null hypersurface emanating from a point on Σ outside of the support of the initial data towards $\mathscr{I}^+$ to make the region of integration compact. The restriction of the energy density on this hypersurface is identically zero, and sending it off towards i^0 returns the identity (3.3.8). □

[3]This follows from the fact that $K^a = \partial_t$ is exactly Killing on the physical spacetime $\mathscr{M}$, and the identity $\pounds_K\hat{\eta}_{ab} = \pounds_K(\Omega^2\eta_{ab}) = (\Omega^{-2}\pounds_K\Omega^2)\hat{\eta}_{ab} = 2(\Omega^{-1}\partial_u\Omega)\hat{\eta}_{ab}$, where $\pounds_K$ denotes the Lie derivative along K. In fact, K^a is also exactly Killing with respect to $\hat{\eta}$ where $\Omega = R$, since $\partial_u\Omega = 0$ there. However, this is only true near $\mathscr{I}^+$ and where $\chi = 1$.

Conformal Invariance of Energies

On any hypersurface $\mathcal{H}$ of $\hat{\mathscr{M}} \supset \mathscr{M}$ the energies induced by the rescaled stress-energy tensor $\hat{\mathbf{T}}_{ab}$ and the physical stress-energy tensor $\mathbf{T}_{ab}$ are equal, on account of the conformal covariance (2.2.11) of $\mathbf{T}_{ab}$. Indeed, for any multiplier vector field K^a

$$\begin{aligned}\int_{\mathcal{H}} K^a \mathbf{T}_a{}^b \partial_b \lrcorner \,\mathrm{dv} &= \int_{\mathcal{H}} K^a \mathbf{T}_{ac} g^{bc} \partial_b \lrcorner \,\mathrm{dv} \\ &= \int_{\mathcal{H}} K^a \Omega^2 \hat{\mathbf{T}}_{ac} \Omega^2 \hat{g}^{bc} \partial_b \lrcorner \,\Omega^{-4} \widehat{\mathrm{dv}} = \int_{\mathcal{H}} K^a \hat{\mathbf{T}}_a{}^b \partial_b \lrcorner \,\widehat{\mathrm{dv}}.\end{aligned} \tag{3.3.9}$$

In particular, the initial energy (3.3.6) is

$$\begin{aligned}\mathcal{E}_\Sigma[\hat{F}] &= \int_\Sigma K^a \hat{\mathbf{T}}_{ab} \hat{T}^b (\hat{T} \lrcorner \,\widehat{\mathrm{dv}}) \\ &= \int_\Sigma K^a \mathbf{T}_{ab} T^b (T \lrcorner \,\mathrm{dv}) \\ &= \frac{1}{2} \int_\Sigma (|\mathbf{E}|^2 + |\mathbf{B}|^2) \,\mathrm{dv}_\Sigma =: \mathcal{E}_\Sigma[F],\end{aligned}$$

where $\mathrm{dv}_\Sigma = (T \lrcorner \,\mathrm{dv})\big|_\Sigma$. One also notes that $\widehat{\mathrm{dv}}_\Sigma = \Omega^3 \,\mathrm{dv}$ and

$$F_0 = \Omega \hat{F}_0, \qquad F_1 = \Omega^2 \hat{F}_1, \qquad F_2 = \Omega^3 \hat{F}_2, \tag{3.3.10}$$

so the expression (3.3.6) can be rewritten as

$$\mathcal{E}_\Sigma[\hat{F}] \simeq \int_\Sigma \left(|F_0|^2 + |F_1|^2 + |F_2|^2\right) \mathrm{dv}_\Sigma = \|F_0\|^2_{L^2(\Sigma)} + \|F_1\|^2_{L^2(\Sigma)} + \|F_2\|^2_{L^2(\Sigma)} \simeq \mathcal{E}_\Sigma[F].$$

3.3.2 Equations of Motion and Gauge Fixing

The equations (2.2.5) on $\mathscr{M}$ read

$$\Box A_b - \nabla_b(\nabla_a A^a) = 0, \tag{3.3.11}$$

and, by the conformal invariance (2.2.10) of Maxwell's equations, are equivalent to

$$\hat{\Box} \hat{A}_b - \hat{\nabla}_b(\hat{\nabla}_a \hat{A}^a) + \hat{\mathrm{R}}_{ab} \hat{A}^a = 0 \tag{3.3.12}$$

on $\hat{\mathscr{M}}$. The energies defined by (2.2.6), when written out in terms of the potential A_a, contain antisymmetrized derivatives of A_a and do not define norms on the potential without a choice of gauge. To construct the trace and scattering operators of section 1.2.4 as maps between Hilbert spaces, one thus aims to fix the gauge in such a way that the natural energies on the initial surface Σ and $\mathscr{I}^+$ become norms on the Maxwell potential. We will impose a gauge on the physical field A_a, and show that it leads to an admissible gauge fixing condition on $\hat{A}_a$ throughout $\hat{\mathscr{M}}$, all the way up to and on $\mathscr{I}^+$. To reduce the natural energy in the physical spacetime to a norm in A_a on $\Sigma \simeq \mathbb{R}^3$, the obvious choice of gauge is the Coulomb gauge (4.3.3), since then for a smooth compactly supported potential on Σ

$$\begin{aligned}\mathcal{E}_\Sigma[F] &= \frac{1}{2} \int_\Sigma \left(|\mathbf{E}|^2 + |\mathbf{B}|^2\right) \mathrm{dv}_\Sigma \\ &= \frac{1}{2} \int_\Sigma \left(|\dot{\mathbf{A}}|^2 - 2\dot{\mathbf{A}} \cdot \boldsymbol{\nabla} \mathfrak{a} + |\boldsymbol{\nabla} \mathfrak{a}|^2 + |\boldsymbol{\nabla} \mathbf{A}|^2 - \boldsymbol{\nabla}_j \mathbf{A}_i \boldsymbol{\nabla}^i \mathbf{A}^j\right) \mathrm{dv}_\Sigma \\ &= \frac{1}{2} \int_\Sigma \left(|\dot{\mathbf{A}}|^2 + |\boldsymbol{\nabla} \mathfrak{a}|^2 + |\boldsymbol{\nabla} \mathbf{A}|^2\right) \mathrm{dv}_\Sigma,\end{aligned}$$

where $\boldsymbol{\nabla}$ is the Levi–Civita connection on Σ, $\mathbf{A}$ denotes the projection of A_a onto Σ, $\dot{\mathbf{A}} = \partial_t \mathbf{A}$, and in the last line we have integrated by parts and used the Coulomb gauge $\boldsymbol{\nabla} \cdot \mathbf{A} = 0 = \boldsymbol{\nabla} \cdot \dot{\mathbf{A}}$. Now if one contracts (3.3.11) with $T^a = \partial_t$, one ends up with the elliptic equation

$$\boldsymbol{\Delta} \mathfrak{a} = 0 \quad \text{on } \Sigma_t \tag{3.3.13}$$

for each $t \in \mathbb{R}$. We have the following result.

Proposition 3.3.3. *On Minkowski space* $(\mathscr{M} = \mathbb{R}^4, \eta)$ *one may impose the gauges*

$$\boldsymbol{\nabla} \cdot \mathbf{A} = 0, \qquad \mathfrak{a} = 0, \qquad \textit{and} \qquad \nabla_a A^a = 0 \tag{3.3.14}$$

simultaneously. We call the gauge (3.3.14) *the* temporal-Coulomb gauge*.*

Proof. Let $A_a = (\mathfrak{a}, \mathbf{A})$ be any smooth solution to Maxwell's equations on $\mathscr{M}$. We impose the Coulomb gauge $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$, which has the residual gauge freedom $\chi_{\text{res.}}$, where $\boldsymbol{\Delta} \chi_{\text{res.}} = 0$ on Σ_t for all t. The solutions to $\boldsymbol{\Delta} \chi_{\text{res.}} = 0$ are constants on $\Sigma_t \simeq \mathbb{R}^3$, so a residual gauge transformation effects

$$\mathfrak{a} \rightsquigarrow \mathfrak{a} + \partial_t \chi_{\text{res.}},$$

where $\chi_{\text{res.}}$ is a function only of t. From (3.3.13), $\mathfrak{a}$ is also a function only of t, so we may choose $\chi_{\text{res.}}$ so that the gauge transformed component $\mathfrak{a}$ is identically zero (by choosing $\chi_{\text{res.}}$ to be the negative of the antiderivative of $\mathfrak{a}$). Then $\nabla_a A^a = \partial_t \mathfrak{a} - \boldsymbol{\nabla} \cdot \mathbf{A} = 0$ follows automatically. □

To study $\hat{A}_a$ on the rescaled spacetime $\hat{\mathscr{M}}$ (and in particular to ensure that we can solve for $\hat{A}_a$ up to $\mathscr{I}$), we must convert the gauge condition (3.3.14) into a gauge condition on the rescaled Maxwell potential $\hat{A}_a$ and solve the system (3.3.12). Under a conformal transformation $g_{ab} \rightsquigarrow \hat{g}_{ab} = \Omega^2 g_{ab}$ the spacetime divergence $\nabla_a A^a$ transforms as

$$\nabla_a A^a = \Omega^2 (\hat{\nabla}_a \hat{A}^a - 2\Upsilon_a \hat{A}^a),$$

so Lorenz gauge in the physical spacetime $\mathscr{M}$ is equivalent to the gauge condition

$$\hat{\nabla}_a \hat{A}^a = 2\Upsilon_a \hat{A}^a = 2\Omega^{-1} (\hat{\nabla}^a \Omega) \hat{A}_a \tag{3.3.15}$$

on $\hat{\mathscr{M}}$. The equation (3.3.12) then reads

$$\hat{\Box} \hat{A}_b - 2\hat{\nabla}_b (\Omega^{-1} (\hat{\nabla}^a \Omega) \hat{A}_a) + \hat{\mathrm{R}}_{ab} \hat{A}^a = 0. \tag{3.3.16}$$

Here the appearance of the Ω^{-1} factor in (3.3.16) is problematic; as it stands, solutions to (3.3.16) may develop singularities on $\mathscr{I} = \{\Omega = 0\}$. The extra temporal gauge condition $\mathfrak{a} = 0$ on $\mathscr{M}$ ensures that this cannot happen: recalling that $T^a = l^a + \frac{1}{2} n^a = \hat{l}^a + \frac{1}{2} \Omega^2 \hat{n}^a$, the temporal gauge transported to $\hat{\mathscr{M}}$ reads

$$0 = \hat{A}_0 + \frac{1}{2} \Omega^2 \hat{A}_1. \tag{3.3.17}$$

Now for $\Omega = R\chi(u)$ one computes

$$\hat{\nabla}^a \Omega = -\chi^{-1} \hat{l}^a + R \left(\chi' + \frac{1}{2} R \chi \right) \hat{n}^a,$$

so using (3.3.17), one has

$$\Omega^{-1}(\hat{\nabla}^a\Omega)\hat{A}_a = (R+\chi'\chi^{-1})\hat{A}_1 =: \varsigma\hat{A}_1,$$

where $\varsigma := R+\chi'\chi^{-1}$, showing that the coefficients of the equation (3.3.16) are in fact non-singular up to $\mathscr{I}^+$. Equation (3.3.16) then becomes

$$\widehat{\Box}\hat{A}_b - 2\hat{\nabla}_b(\varsigma\hat{A}_1) + \hat{\mathrm{R}}_{ab}\hat{A}^a = 0. \tag{3.3.18}$$

This will ensure that the solution to (3.3.16) is in fact smooth throughout the partially compactified spacetime $\hat{\mathscr{M}}$, including on $\mathscr{I}^+$. Some further care is needed before the Cauchy problem for equation (3.3.16) can be solved, however, since the background spacetime is singular at i^0. To circumvent this, we proceed as follows.

For smooth compactly supported initial data $(\hat{A}_a, \hat{\nabla}_{\hat{T}}\hat{A}_a)|_\Sigma$ the putative solution will have support bounded away from i^0, as depicted below.

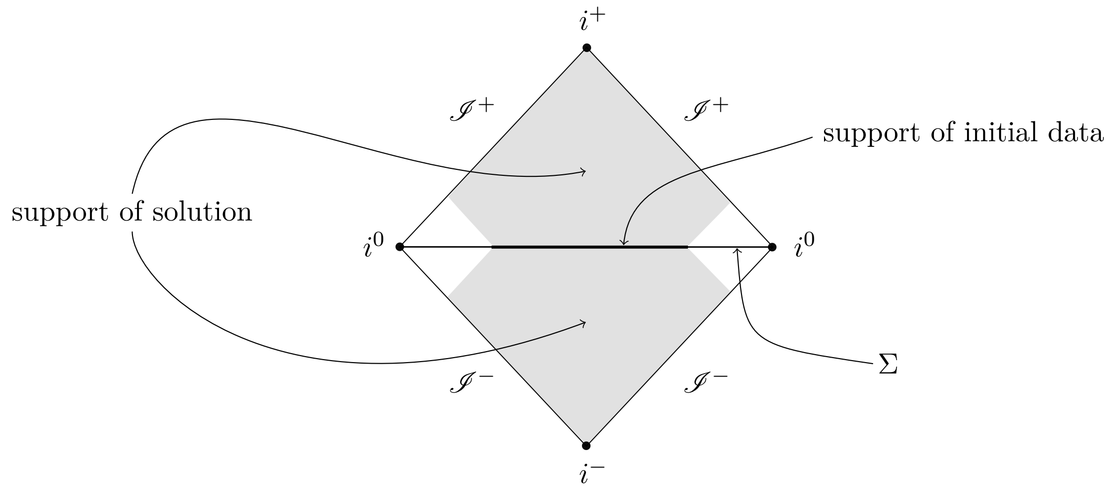

Figure 3.5: The shape of the support of a solution with compactly supported initial data.

With this in mind, we regularize the background spacetime without affecting the solution in the following way. Keeping it uniformly spacelike, we deform the initial surface Σ away from the support of the initial data in such a way that the new deformed initial surface $\check{\Sigma}$ intersects $\mathscr{I}^+$ in the future of i^0, and the support of the solution in the future of Σ remains to the future of the deformed surface $\check{\Sigma}$. We then cut off and discard the past of the deformed surface $\check{\Sigma}$, as illustrated in fig. 3.6.

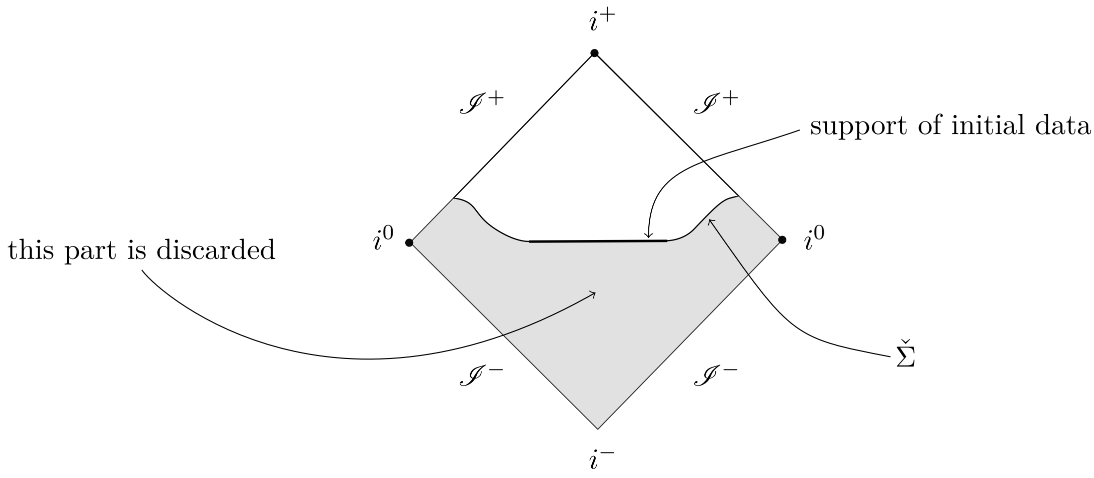

Figure 3.6: The deformation $\check{\Sigma}$ of the initial surface Σ avoids i^0 without affecting the solution.

The part of $\hat{\mathscr{M}}$ lying in the future of $\check{\Sigma}$ is now a completely regular compact globally hyperbolic Lorentzian manifold with boundary, and from the point of view of the solution to (3.3.18) in $J^+(\Sigma)$ is indistinguishable from $\hat{\mathscr{M}}$. It may be extended to a smooth globally hyperbolic Lorentzian manifold without boundary, say $(\hat{\mathscr{M}}^e = \mathbb{R} \times \mathbb{S}^3, \hat{h})$, where $\hat{h}$ agrees with $\hat{\eta}$ on $J^+(\check{\Sigma}) \cap \hat{\mathscr{M}}$. By theorem A.2.2, the original smooth compactly supported data $(\hat{A}_a, \hat{\nabla}_{\hat{T}} \hat{A}_a)|_\Sigma$ gives rise to a unique smooth solution $\hat{A}_a$ on $\hat{\mathscr{M}}^e$, which solves (3.3.18) in $J^+(\Sigma) \cap \hat{\mathscr{M}}$ and whose support remains away from i^0. In particular, the components $\hat{A}_0$ and $\hat{A}_1$ of this solution are smooth up to $\mathscr{I}^+$, and so the temporal gauge condition (3.3.17) can be extended smoothly *onto* $\mathscr{I}^+$, where it becomes

$$\hat{A}_0 \approx 0. \tag{3.3.19}$$

With the gauge condition (3.3.17) now satisfied throughout $\hat{\mathscr{M}}$, the equation (3.3.18) in fact consists of three, not four independent equations, since the component $\hat{A}_0$ can be determined from $\hat{A}_1$. We have thus proven the following.

Theorem 3.3.4. *For smooth compactly supported initial data* $(A_a, \nabla_T A_a)|_\Sigma$ *for Maxwell's equations in the temporal-Coulomb gauge* $\mathfrak{a}|_\Sigma = \boldsymbol{\nabla} \cdot \mathbf{A}|_\Sigma = \boldsymbol{\nabla} \cdot \dot{\mathbf{A}}|_\Sigma = 0$ *there exists a unique smooth* rescaled *solution* $\hat{A}_a$ *on* $\hat{\mathscr{M}}$. *The support of* $\hat{A}_a$ *remains away from* i^0, *and* $\hat{A}_a$ *satisfies the gauge conditions* (3.3.15) *and* (3.3.17) *throughout* $\hat{\mathscr{M}}$. *In particular,* $\hat{A}_0 \approx 0$ *on* $\mathscr{I}^+$.

Proof. First, it is clear that for smooth compactly supported initial data *for the field* there exists a unique smooth solution F_{ab} on $\mathscr{M}$, for example by theorem A.2.2, and that the initial gauge constraints

$$\mathfrak{a}|_\Sigma = \boldsymbol{\nabla} \cdot \mathbf{A}|_\Sigma = \boldsymbol{\nabla} \cdot \dot{\mathbf{A}}|_\Sigma = 0$$

are propagated throughout $\mathscr{M}$. By proposition 3.3.3, we may impose the temporal-Coulomb gauge on this solution throughout $\mathscr{M}$. Once rescaled initial data $(\hat{A}_a, \hat{\nabla}_{\hat{T}} \hat{A}_a)|_\Sigma$ is obtained from the physical initial data $(A_a, \nabla_T A_a)|_\Sigma$, the above construction goes through to extend the solution $\hat{A}_a = A_a$ to $\mathscr{I}^+$, and ensure that the relevant gauge conditions are satisfied. The rescaled initial data is easily constructed from the physical initial data; one has

$$\hat{A}_a = A_a$$

and

$$\hat{\nabla}_{\hat{T}} \hat{A}_a = \hat{T}^b \hat{\nabla}_b \hat{A}_a = \Omega^{-1} T^b (\nabla_b A_a - \Upsilon_b A_a - \Upsilon_a A_b + \eta_{ab} \eta^{cd} \Upsilon_c A_d).$$

The first of these immediately gives the data for $\hat{A}_a$ in terms of the data for A_a, while for the time derivative restriction to Σ gives

$$\hat{\nabla}_{\hat{T}} \hat{A}_a|_\Sigma = \Omega^{-1} (\nabla_T A_a + T_a \eta^{cd} \Upsilon_c A_d)|_\Sigma, \tag{3.3.20}$$

since $T^b \Upsilon_b|_\Sigma = 0$ and $\mathfrak{a}|_\Sigma = 0$. Smoothness and compact support of $(A_a, \nabla_T A_a)|_\Sigma$ then imply the smoothness and compact support of $(\hat{A}_a, \hat{\nabla}_{\hat{T}} \hat{A}_a)|_\Sigma$. □

Gauge Reduction on $\mathscr{I}^+$

In addition to the gauge $\hat{A}_0 \approx 0$ on $\mathscr{I}^+$, the triple gauge fixing condition (3.3.14) in fact gives rise to a kind of second-order gauge reduction on $\mathscr{I}^+$, as we shall see now. The Coulomb gauge $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$ on surfaces of constant t in the coordinates (t, r, θ, ϕ) reads

$$\boldsymbol{\nabla} \cdot \mathbf{A} = \partial_r A_r + \frac{2}{r} A_r + \frac{1}{r^2} \nabla_{\mathbb{S}^2} \cdot A_{\mathbb{S}^2} = 0,$$

where $\nabla_{\mathbb{S}^2} \cdot A_{\mathbb{S}^2} = \mathfrak{s}_2^{ab} \nabla_a^{\mathbb{S}^2} A_b = \nabla_{\mathbb{S}^2} \cdot \hat{A}_{\mathbb{S}^2}$ is the divergence of A_a on the 2-spheres. Changing coordinates

$$(t, r, \theta, \phi) \rightsquigarrow (u, R, \theta, \phi),$$

we note that

$$\partial_r = -l^a + \frac{1}{2} n^a = -\hat{l}^a + \frac{1}{2} \Omega^2 \hat{n}^a. \tag{3.3.21}$$

It should be noted that the ∂_r on the left hand side of (3.3.21) denotes the partial derivative with respect to r *while keeping* t *constant*, in contrast to the ∂_r of (3.2.4), where the variable u is kept constant (hence the apparent disagreement between (3.3.21) and (3.2.4) due to the slight abuse of notation). The gauge condition (3.3.17) then shows that

$$A_r = -\hat{A}_0 + \frac{1}{2} \Omega^2 \hat{A}_1 = \Omega^2 \hat{A}_1.$$

Recalling that near $\mathscr{I}^+$ the conformal factor is $\Omega = R\chi(u)$, the temporal-Coulomb gauge condition becomes

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{A} &= \left(-\hat{D} + \frac{1}{2} \Omega^2 \hat{\Delta} \right) (\Omega^2 \hat{A}_1) + 2R\Omega^2 \hat{A}_1 + R^2 \nabla_{\mathbb{S}^2} \cdot \hat{A}_{\mathbb{S}^2} \\ &= -\hat{D}(R^2 \chi^2 \hat{A}_1) + R^2 \nabla_{\mathbb{S}^2} \cdot \hat{A}_{\mathbb{S}^2} + \mathcal{O}(R^3) \\ &= -R^2 \partial_u (\chi^2 \hat{A}_1) + R^2 \nabla_{\mathbb{S}^2} \cdot \hat{A}_{\mathbb{S}^2} + \mathcal{O}(R^3) = 0. \end{aligned}$$

Dividing by R^2 and taking the limit $R \to 0$, the temporal-Coulomb gauge reduces to the condition

$$\partial_u (\chi^2 \hat{A}_1) \approx \nabla_{\mathbb{S}^2} \cdot \hat{A}_{\mathbb{S}^2} \tag{3.3.22}$$

on $\mathscr{I}^+$.

Remark 3.3.5. One may express the 2-sphere divergence $\nabla_{\mathbb{S}^2} \cdot \hat{A}_{\mathbb{S}^2}$ in terms of the NP components of $\hat{A}_a$ as

$$\nabla_{\mathbb{S}^2} \cdot \hat{A}_{\mathbb{S}^2} = \bar{\hat{\eth}} \hat{A}_2 + \hat{\eth} \bar{\hat{A}}_2 = 2 \operatorname{Re} \hat{\eth} \bar{\hat{A}}_2.$$

3.3.3 Energies and the Trace Operators

Since $\hat{A}_0 \approx 0$ and the angular derivatives are tangential to $\mathscr{I}^+$, we have $\eth \hat{A}_0 \approx 0$. Noting that in rescaled Minkowski space $\hat{\kappa} = 0 = \hat{\sigma}$ and $\hat{\rho} = \bar{\hat{\rho}}$ (see (C.3.2) and (C.1.1)), the expansion (3.2.16) for $\hat{F}_0$ reduces to

$$\hat{F}_0 \approx \hat{\text{þ}} \hat{A}_2 - \hat{\rho} \hat{A}_2.$$

Squaring this we have

$$|\hat{F}_0|^2 \approx |\hat{\text{þ}} \hat{A}_2|^2 - \hat{\rho} \hat{\text{þ}}(|\hat{A}_2|^2) + |\hat{\rho}|^2 |\hat{A}_2|^2.$$

Since $\hat{\text{þ}}$ is a tangential derivative to $\mathscr{I}^+$, we may integrate the second term in the above by parts. Since $\hat{A}_2$ is a $(1,-1)$-scalar, its complex conjugate $\bar{\hat{A}}_2$ is a $(-1,1)$-scalar, and so $|\hat{A}_2|^2$ is a $(0,0)$-scalar. For $\hat{A}_2$ compactly supported on $\mathscr{I}^+$ (that is, supported away from i^0 *and* i^+) we then have

$$\begin{aligned} \int_{\mathscr{I}^+} \hat{\rho} \hat{\text{þ}}(|\hat{A}_2|^2) \, \widehat{\mathrm{dv}}_{\mathscr{I}^+} &= \int_{\mathscr{I}^+} \hat{\rho} \hat{D}(|\hat{A}_2|^2) \, \mathrm{d}u \wedge \mathrm{dv}_{\mathfrak{S}_2} \\ &= - \int_{\mathscr{I}^+} (\hat{D}\rho) |\hat{A}_2|^2 \, \mathrm{d}u \wedge \mathrm{dv}_{\mathfrak{S}_2} \\ &= - \int_{\mathscr{I}^+} (\hat{\text{þ}} \hat{\rho} + \hat{\varepsilon} \rho + \bar{\hat{\varepsilon}} \hat{\rho}) |\hat{A}_2|^2 \, \widehat{\mathrm{dv}}_{\mathscr{I}^+}, \end{aligned}$$

where $\hat{D} \approx \partial_u$. Due to the vanishing of the shear of $\mathscr{I}^+$ ($\hat{\sigma} \approx 0$) and the fact that it is geodetic ($\hat{\kappa} \approx 0$), the spin coefficient curvature equation for $\hat{\text{þ}} \hat{\rho}$ on $\mathscr{I}^+$ reduces to[4]

$$\hat{\text{þ}} \hat{\rho} \approx \hat{\rho}^2 + \hat{\Phi}_{00}. \tag{3.3.23}$$

Thus

$$\int_{\mathscr{I}^+} -\hat{\rho} \hat{\text{þ}}(|\hat{A}_2|^2) \, \widehat{\mathrm{dv}}_{\mathscr{I}^+} = \int_{\mathscr{I}^+} (\hat{\rho}^2 + \hat{\Phi}_{00} + \hat{\varepsilon} \hat{\rho} + \bar{\hat{\varepsilon}} \hat{\rho}) |\hat{A}_2|^2 \, \widehat{\mathrm{dv}}_{\mathscr{I}^+}.$$

Proposition 3.3.6. *On rescaled Minkowski space $(\hat{\mathscr{M}}, \hat{\eta})$ the quantity*

$$\hat{\rho}^2 + \hat{\Phi}_{00} + \hat{\varepsilon} \hat{\rho} + \bar{\hat{\varepsilon}} \hat{\rho}$$

is non-negative definite on $\mathscr{I}^+$ for an appropriate choice of the function $\chi(u)$.

Proof. We first note that

$$\Upsilon_a l^a = \left(\partial_u - \frac{1}{2} \partial_r \right) \log(r^{-1} \chi(u)) = \frac{\chi'(u)}{\chi(u)} + \frac{1}{2r} \approx \frac{\chi'}{\chi},$$

and so by the transformation rules (C.1.1),

$$\hat{\varepsilon} = \varepsilon + \Upsilon_a l^a \approx \frac{\chi'}{\chi} \quad \text{and} \quad \hat{\rho} = \rho - \Upsilon_a l^a \approx -\frac{\chi'}{\chi}.$$

[4] Note that the equation (3.3.23) is one of the Sachs equations with $\hat{\sigma} \approx 0$, [103, 104].

Using the formula (B.1.4) for the conformal change of the trace-free part of the Ricci tensor, the fact that $\mathscr{M}$ is Ricci-flat, and the identity $Dl^a = 0$ valid on $\mathscr{M}$, we have

$$\begin{aligned}\hat{\Phi}_{00} &= \Phi_{00} - l^a D\Upsilon_a + (\Upsilon_a l^a)^2 = -D(\Upsilon_a l^a) + \Upsilon_a Dl^a + (\Upsilon_a l^a)^2 \\ &= -\left(\partial_u - \frac{1}{2}\partial_r\right)\left(\frac{\chi'}{\chi} + \frac{1}{2r}\right) + \left(\frac{\chi'}{\chi} + \frac{1}{2r}\right)^2 \approx \frac{2(\chi')^2 - \chi\chi''}{\chi^2}.\end{aligned}$$

Altogether we therefore have

$$\hat{\rho}^2 + \hat{\Phi}_{00} + \hat{\varepsilon}\hat{\rho} + \bar{\hat{\varepsilon}}\hat{\rho} \approx \frac{1}{\chi^2}\left((\chi')^2 - \chi\chi''\right).$$

From here it is clear that there exist many possible choices for $\chi(u)$. One explicit choice near i^+ is

$$\chi(u) = \mathrm{e}^{-u^2},$$

which gives

$$\frac{1}{\chi^2}\left((\chi')^2 - \chi\chi''\right) = 2.$$

□

Therefore for $\hat{A}_2 \in \mathcal{C}_c^\infty(\mathscr{I}^+)$

$$\mathcal{E}_{\mathscr{I}^+}[\hat{A}] \simeq \int_{\mathscr{I}^+} |\hat{F}_0|^2 \,\widehat{\mathrm{dv}}_{\mathscr{I}^+} \simeq \int_{\mathscr{I}^+} \left(|\hat{\text{þ}}\hat{A}_2|^2 + \varpi_\chi^2 |\hat{A}_2|^2\right) \widehat{\mathrm{dv}}_{\mathscr{I}^+},$$

where ϖ_χ^2 is a smooth non-negative function which is equal to $\hat{\rho}^2 + 2 = 4u^2 + 2$ in a neighbourhood of i^+, and where $\mathcal{C}_c^\infty(\mathscr{I}^+)$ is the space of smooth compactly supported functions on $\mathscr{I}^+$ (which are therefore supported away from both i^0 and i^+). Note that it is always possible to affinely reparametrize $\mathscr{I}^+$ so that $\hat{\varepsilon} + \bar{\hat{\varepsilon}} \approx 0$, and choose the vector $\hat{l}^a = \hat{o}^A \hat{o}^{A'}$ so that the spinor $\hat{o}^A$ has parallely propagated flag planes, $\hat{\varepsilon} - \bar{\hat{\varepsilon}} \approx 0$. Then $\hat{\varepsilon} \approx 0$, and $\hat{\text{þ}} = \hat{D}$. This would, of course, entail a modification of the conformal factor.

Definition 3.3.7. For the component $\hat{A}_2$ of the Maxwell potential we define the function space $\mathcal{H}^1(\mathscr{I}^+)$ on $\mathscr{I}^+$ by completion of $\mathcal{C}_c^\infty(\mathscr{I}^+)$ in the norm

$$\|\hat{A}_2\|^2_{\mathcal{H}^1(\mathscr{I}^+)} := \int_{\mathscr{I}^+} \left(|\hat{\text{þ}}\hat{A}_2|^2 + \varpi_\chi^2 |\hat{A}_2|^2\right) \widehat{\mathrm{dv}}_{\mathscr{I}^+}.$$

The space $\mathcal{H}^1(\mathscr{I}^+)$ will be the space of characteristic data on $\mathscr{I}^+$ for the free Maxwell equations. Note that $\mathcal{H}^1(\mathscr{I}^+)$ is not quite the usual H^1 Sobolev space on $\mathbb{R} \times \mathbb{S}^2$ as it does not involve any derivatives in the $\mathbb{S}^2$ directions.

Remark 3.3.8. Note that because in our setting the value of ϖ_χ^2 near i^+ is $2 + 4u^2$, the finiteness of the norm $\|\hat{A}_2\|^2_{\mathcal{H}^1(\mathscr{I}^+)}$ imposes a corresponding decay rate on $\hat{A}_2$ towards i^+. This is simply a feature of the compactification and may be relaxed by unfolding the i^+ end of $\mathscr{I}^+$ a posteriori.

We have already seen that in the temporal-Coulomb gauge the initial energy $\mathcal{E}_\Sigma$ as given in (3.3.6) is neatly expressed in terms of the physical potential A_a as

$$\begin{aligned}\mathcal{E}_\Sigma[F] &= \frac{1}{2}\int_\Sigma \left(|\mathbf{E}|^2 + |\mathbf{B}|^2\right) \mathrm{dv}_\Sigma \\ &= \frac{1}{2}\int_\Sigma \left(|\dot{\mathbf{A}}|^2 + |\boldsymbol{\nabla}\mathbf{A}|^2\right) \mathrm{dv}_\Sigma =: \mathcal{E}_\Sigma[\mathbf{A}].\end{aligned} \tag{3.3.24}$$

Definition 3.3.9. For the initial data $(\dot{\mathbf{A}},\mathbf{A})|_\Sigma$ for the free Maxwell equations we define the function space $\dot{H}^1_C(\Sigma)\oplus L^2_C(\Sigma)$ of Coulomb gauge initial data by completion of smooth compactly supported Coulomb gauge initial data in the norm

$$\|(\dot{\mathbf{A}},\mathbf{A})\|^2_{\dot{H}^1\oplus L^2} = \int_\Sigma \left(|\boldsymbol{\nabla}\mathbf{A}|^2 + |\dot{\mathbf{A}}|^2\right)\mathrm{d}v_\Sigma\,.$$

The space $\dot{H}^1(\Sigma)\oplus L^2(\Sigma)$ will be the space of initial data for the free Maxwell equations on Minkowski space.

Construction of Trace Operators

Let

$$\mathcal{D}^\infty_c(\Sigma) := \{(\mathbf{A},\dot{\mathbf{A}}) \in \mathcal{C}^\infty_c(\Sigma)\oplus\mathcal{C}^\infty_c(\Sigma)\ :\ \boldsymbol{\nabla}\cdot\mathbf{A} = 0 = \boldsymbol{\nabla}\cdot\dot{\mathbf{A}}\}$$

be the space of smooth compactly supported Coulomb gauge initial data for the physical Maxwell equations in the temporal-Coulomb gauge (3.3.14). An element $\mathscr{a} = (\mathbf{A},\dot{\mathbf{A}})$ of $\mathcal{D}^\infty_c(\Sigma)$ defines smooth compactly supported initial data for the rescaled Maxwell equations (3.3.16) in the temporal-Coulomb gauge as follows. First,

$$\hat{\mathfrak{a}}\big|_\Sigma = 0, \qquad\qquad \hat{\mathbf{A}}\big|_\Sigma = \mathbf{A}\big|_\Sigma.$$

For the time derivative part of the initial data, one computes the inverse relation to (3.3.20),

$$\nabla_T A_b\big|_\Sigma = \Omega\left(\hat{\nabla}_{\hat{T}}\hat{A}_b + \hat{T}^a\Upsilon_a\hat{A}_b + \Upsilon_b\hat{\mathfrak{a}} - \hat{T}_b\Upsilon_a\hat{A}^a\right)\Big|_\Sigma,$$

so since $\partial_t\Omega|_\Sigma = 0$,

$$-\hat{h}_a^b\hat{\nabla}_{\hat{T}}\hat{A}_b\big|_\Sigma = \Omega^{-1}\dot{\mathbf{A}}_a\big|_\Sigma.$$

Note that $\dot{\mathbf{A}}$ is supported away from i^0, so that Ω^{-1} is smooth on its support. Therefore we can solve (3.3.16) as described in section 3.3.2 (theorem 3.3.4) to get a unique smooth solution $\hat{A}_a$ satisfying the gauge conditions (3.3.15) and (3.3.17). Using the smoothness of $\hat{A}_a$, one may take the trace of this solution on $\mathscr{I}^+$ to get a smooth restriction $\hat{A}_a|_{\mathscr{I}^+}$. Theorem A.2.2 ensures that $\hat{A}_a$ depends linearly on the initial data, so one has the linear map

$$\begin{aligned}\mathfrak{T}^+ : \mathcal{D}^\infty_c(\Sigma) &\longrightarrow \mathcal{C}^\infty(\mathscr{I}^+),\\ (\mathbf{A},\dot{\mathbf{A}}) &\longmapsto \hat{A}_2^+,\end{aligned} \tag{3.3.25}$$

where $\hat{A}_2^+ = \hat{A}_2|_{\mathscr{I}^+}$ is supported away from i^0 and decays towards i^+ at the rate described in remark 3.3.8. The energy estimate (3.3.7) implies that there exists a constant $C > 0$ such that for all $\mathscr{a} \in \mathcal{D}^\infty_c(\Sigma)$

$$\|\mathfrak{T}^+\mathscr{a}\|_{\mathcal{H}^1(\mathscr{I}^+)} \leqslant C\|\mathscr{a}\|_{\dot{H}^1\oplus L^2}, \tag{3.3.26}$$

and

$$\|\mathscr{a}\|_{\dot{H}^1\oplus L^2} \leqslant C\|\mathfrak{T}^+\mathscr{a}\|_{\mathcal{H}^1(\mathscr{I}^+)}. \tag{3.3.27}$$

By (3.3.26) and the density of $\mathcal{D}^\infty_c(\Sigma)$ in $\dot{H}^1_C(\Sigma)\oplus L^2_C(\Sigma)$, the bounded linear operator $\mathfrak{T}^+$ extends uniquely to a bounded linear operator on $\dot{H}^1_C(\Sigma)\oplus L^2_C(\Sigma)$. Moreover, the reverse estimate (3.3.27) ensures that $\mathfrak{T}^+$ is a bijection from $\dot{H}^1_C(\Sigma)\oplus L^2_C(\Sigma)$ to its image, and that the image is a closed subspace of $\mathcal{H}^1(\mathscr{I}^+)$.

Definition 3.3.10. The bounded linear operator

$$\mathfrak{T}^+ : \dot{H}^1_C(\Sigma) \oplus L^2_C(\Sigma) \longrightarrow \mathcal{H}^1(\mathscr{I}^+)$$

that takes the initial data for (3.3.16) on Σ to the characteristic data on $\mathscr{I}^+$ is called the *future trace operator* for the free Maxwell equations in the gauge (3.3.14).

To show that $\mathfrak{T}^+$ is surjective (and hence an isomorphism between $\dot{H}^1_C(\Sigma) \oplus L^2_C(\Sigma)$ and $\mathcal{H}^1(\mathscr{I}^+)$), it is enough to show that for every $\mathscr{b} \in \mathcal{C}^\infty_c(\mathscr{I}^+)$ there exists a unique $\mathscr{a} \in \dot{H}^1_C(\Sigma) \oplus L^2_C(\Sigma)$ such that $\mathfrak{T}^+\mathscr{a} = \mathscr{b}$. Indeed, then the inverse trace operator can be extended to $\mathcal{H}^1(\mathscr{I}^+)$ as follows. For any $\mathscr{b} \in \mathcal{H}^1(\mathscr{I}^+)$ we can find a sequence $\{\mathscr{b}_n\}_n \subset \mathcal{C}^\infty_c(\mathscr{I}^+)$ such that $\mathscr{b}_n \to \mathscr{b}$ in $\mathcal{H}^1(\mathscr{I}^+)$. Then for each n there exists a unique $\mathscr{a}_n \in \dot{H}^1_C(\Sigma) \oplus L^2_C(\Sigma)$ such that $\mathscr{b}_n = \mathfrak{T}^+\mathscr{a}_n$, and

$$\|\mathfrak{T}^+\mathscr{a}_n - \mathscr{b}\|_{\mathcal{H}^1(\mathscr{I}^+)} \longrightarrow 0. \tag{3.3.28}$$

The above estimates easily imply that the sequence $\{\mathscr{a}_n\}_n$ is Cauchy, since

$$\|\mathscr{a}_n - \mathscr{a}_m\|_{\dot{H}^1 \oplus L^2} \lesssim \|\mathfrak{T}^+\mathscr{a}_n - \mathfrak{T}^+\mathscr{a}_m\|_{\mathcal{H}^1(\mathscr{I}^+)} \leqslant \|\mathscr{b}_n - \mathscr{b}_m\|_{\mathcal{H}^1(\mathscr{I}^+)}.$$

Therefore there exists $\mathscr{a} \in \dot{H}^1_C(\Sigma) \oplus L^2_C(\Sigma)$ such that $\mathscr{a}_n \to \mathscr{a}$ in $\dot{H}^1 \oplus L^2$, and by (3.3.28) $\mathfrak{T}^+\mathscr{a} = \mathscr{b}$. Proving that for every $\mathscr{b} \in C^\infty_c(\mathscr{I}^+)$ there exists a unique $\mathscr{a} \in \dot{H}^1_C(\Sigma) \oplus L^2_C(\Sigma)$ such that $\mathfrak{T}^+\mathscr{a} = \mathscr{b}$ amounts to solving the Goursat problem from $\mathscr{I}^+$.

The Goursat Problem

There are two key observations that allow us to solve the Goursat problem for (3.3.18). The first is that the equation for the component $\hat{A}_1$ decouples from the other equations near $\mathscr{I}^+$. The second is that our second order gauge reduction on $\mathscr{I}^+$ provides the characteristic data for $\hat{A}_1$ given the characteristic data for $\hat{A}_2$. These are intimately tied to the fact that the physical background is flat, and the form of the conformal factor near $\mathscr{I}^+$.

Proposition 3.3.11. *Near $\mathscr{I}^+$ the equation for the component $\hat{A}_1$ in the system (3.3.18) decouples from the other components and becomes*

$$\widehat{\Box}\hat{A}_1 + 2\chi'\chi^{-1}\hat{\Delta}\hat{A}_1 = 0, \tag{3.3.29}$$

where $\hat{\Delta} = \hat{n}^a\hat{\nabla}_a$.

Proof. We contract (3.3.18) with $\hat{n}^b$,

$$\hat{n}^b\widehat{\Box}\hat{A}_b - 2\hat{n}^b\hat{\nabla}_b(\varsigma\hat{A}_1) + \hat{n}^b\hat{\mathrm{R}}_{ab}\hat{A}^a = 0,$$

and compute the terms arising in this contraction explicitly in terms of the components $\hat{A}_{0,1,2}$. First, one expands

$$\hat{n}^b\hat{\mathrm{R}}_{ab}\hat{A}^a = \hat{A}_1\hat{\mathrm{R}}_{ab}\hat{l}^a\hat{n}^b + \hat{A}_0\hat{\mathrm{R}}_{ab}\hat{n}^a\hat{n}^b - \hat{A}_2\hat{\mathrm{R}}_{ab}\bar{\hat{m}}^a\hat{n}^b - \bar{\hat{A}}_2\hat{\mathrm{R}}_{ab}\hat{m}^a\hat{n}^b. \tag{3.3.30}$$

Using the conformal transformation rule (B.1.2) and recalling that Ω is radial and $\mathscr{M}$ Ricci-flat, one sees that the last two terms in (3.3.30) vanish,

$$\hat{n}^b\hat{\mathrm{R}}_{ab}\hat{A}^a = \hat{A}_1\hat{\mathrm{R}}_{ab}\hat{l}^a\hat{n}^b + \hat{A}_0\hat{\mathrm{R}}_{ab}\hat{n}^a\hat{n}^b = \hat{A}_1(\Omega^2\hat{\Phi}_{22} - 2\hat{\Phi}_{11} + 6\hat{\Lambda}).$$

Now the conformal transformation rules (C.1.1) for the spin coefficients and the values of the spin coefficients on physical Minkowski space (C.3.1) give the following values of the relevant rescaled spin coefficients near $\mathscr{I}^+$,

$$\hat{\gamma} = 0, \qquad \hat{\nu} = 0, \qquad \hat{\pi} = 0, \qquad \hat{\lambda} = 0, \qquad \hat{\varepsilon} = \bar{\hat{\varepsilon}} = \frac{\chi'}{\chi} + \frac{R}{2},$$

and moreover $\hat{\alpha} + \bar{\hat{\beta}} = 0$ and $\hat{\mu} = 0$. Thus the only non-vanishing directional derivative of $\hat{n}^a$ is $\hat{D}\hat{n}^a = -2\hat{\varepsilon}\hat{n}^a$. To compute $\hat{n}^b \widehat{\Box} \hat{A}_b = \widehat{\Box}\hat{A}_1 - 2\hat{\nabla}_a \hat{A}_b \hat{\nabla}^a \hat{n}^b - \hat{A}_b \widehat{\Box}\hat{n}^b$, we calculate

$$\hat{\nabla}_a \hat{A}_b \hat{\nabla}^a \hat{n}^b = \hat{D}\hat{A}_b \hat{\Delta}\hat{n}^b + \hat{\Delta}\hat{A}_b \hat{D}\hat{n}^b - \bar{\hat{\delta}}\hat{A}_b \hat{\delta}\hat{n}^b - \hat{\delta}\hat{A}_b \bar{\hat{\delta}}\hat{n}^b = -2\hat{\varepsilon}\hat{\Delta}\hat{A}_1$$

and

$$\widehat{\Box}\hat{n}^b = \hat{\nabla}_c(\hat{n}^c \hat{D}\hat{n}^b + \hat{l}^c \hat{\Delta}\hat{n}^b - \hat{m}^c \bar{\hat{\delta}}\hat{n}^b - \bar{\hat{m}}^c \hat{\delta}\hat{n}^b) = -2\hat{n}^b(\hat{\Lambda} - \hat{\Phi}_{11}).$$

Altogether we thus have

$$\widehat{\Box}\hat{A}_1 + \hat{L}_1 \hat{A}_1 = 0,$$

where

$$\hat{L}_1 = 2\chi'\chi^{-1}\hat{\Delta} + (R^2\chi^2\hat{\Phi}_{22} - 2\hat{\Delta}\varsigma + 8\hat{\Lambda} - 4\hat{\Phi}_{11}).$$

We may compute the NP Ricci curvature scalars near $\mathscr{I}^+$, which turn out to be

$$\hat{\Phi}_{01,02,12,22} = 0, \qquad \hat{\Lambda} = 0, \qquad \hat{\Phi}_{00} = R\chi'\chi^{-1} - \chi''\chi^{-1} + 2(\chi'\chi^{-1})^2, \qquad \hat{\Phi}_{11} = \frac{1}{2}\chi^{-2}.$$

Noting that $\hat{n}^a = -\chi^{-2}\partial_R$, $\hat{L}_1$ then manifestly simplifies to just

$$\hat{L}_1 = 2\chi'\chi^{-1}\hat{\Delta}.$$

□

Since we wish to solve the system (3.3.18) from the characteristic data in $\mathcal{H}^1(\mathscr{I}^+)$ (cf. (3.3.25)), we must find a way to recover the data for the component $\hat{A}_1$ on $\mathscr{I}^+$ from the component $\hat{A}_2$ on $\mathscr{I}^+$. This is exactly the purpose of the identity (3.3.22); given characteristic data $\hat{A}_2^+ \in \mathcal{C}_c^\infty(\mathscr{I}^+)$, we solve

$$\partial_u(\chi^2 \hat{A}_1^+) \approx 2 \operatorname{Re} \hat{\eth} \bar{\hat{A}}_2^+ \tag{3.3.31}$$

on $\mathscr{I}^+$ for $\hat{A}_1^+ \in \mathcal{C}^\infty(\mathscr{I}^+) \cap \dot{H}^1(\mathscr{I}^+)$[5]. Then for this $\hat{A}_1^+$ we choose a spacelike hypersurface $\mathcal{S}$ which intersects $\mathscr{I}^+$ in the past of the support of $\hat{A}_2^+$ and in the future of i^0, and solve the Goursat problem for (3.3.29) in the future of $\mathcal{S}$, as depicted in fig. 3.7.

[5]Note that generically the solution $\hat{A}_1$ to (3.3.31) with $\hat{A}_2 \in \mathcal{C}_c^\infty(\mathscr{I}^+)$ will be smooth, but can only be compactly supported in one direction (near i^0 *or* near i^+). If we choose $\hat{A}_1^+$ to vanish near i^+, then it follows that $\hat{A}_1^+ \in \dot{H}^1(\mathscr{I}^+)$.

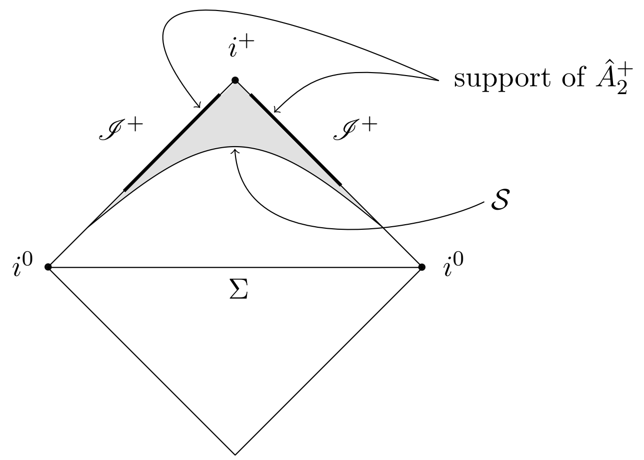

Figure 3.7: We first solve the Goursat problem for $\hat{A}_1$ in the shaded region $J^+(\mathcal{S})$.

As $\hat{A}_1^+ \in H^1(\mathscr{I}^+ \cap J^+(\mathcal{S}))$, using Hörmander's result (theorem A.2.5), we obtain a unique finite energy solution $\hat{A}_1$ to (3.3.29) in $J^+(\mathcal{S})$, with data on $\mathscr{I}^+$ given by the solution to (3.3.31). For a sufficiently regular foliation $\{\mathcal{S}_\tau\}_\tau$ of $J^+(\mathcal{S})$, where, say, $\mathcal{S}_0 = \mathcal{S}$ and $\mathcal{S}_1 = \mathscr{I}^+ \cap J^+(\mathcal{S})$, this solution has the regularity

$$\hat{A}_1 \in \mathcal{C}^0([0,1]_\tau; H^1(\mathcal{S}_\tau)) \cap \mathcal{C}^1([0,1]_\tau; L^2(\mathcal{S}_\tau, \mathrm{dv}^0_{\mathcal{S}_\tau})).$$

We next solve the equation for the component $\hat{A}_2$ in $J^+(\mathcal{S})$; a calculation shows that it is

$$\widehat{\Box}\hat{A}_2 + 4\hat{\alpha}\left(\hat{\eth}\hat{A}_2 - \bar{\hat{\eth}}\hat{A}_2\right) + \left(\chi^{-2} - 8\hat{\alpha}^2\right)\hat{A}_2 = 2R\hat{\delta}\hat{A}_1 \tag{3.3.32}$$

near $\mathscr{I}^+$. This is a forced linear wave equation, where the forcing term is

$$2R\hat{\delta}\hat{A}_1 \in \mathcal{C}^0([0,1]_\tau; L^2(\mathcal{S}_\tau)) \subset L^1_{\mathrm{loc}}([0,1]_\tau; L^2(\mathcal{S}_\tau)),$$

so we may apply Hörmander's theorem once more to get a unique finite energy solution

$$\hat{A}_2 \in \mathcal{C}^0([0,1]_\tau; H^1(\mathcal{S}_\tau)) \cap \mathcal{C}^1([0,1]_\tau; L^2(\mathcal{S}_\tau, \mathrm{dv}^0_{\mathcal{S}_\tau})).$$

Having solved for $\hat{A}_1$ and $\hat{A}_2$, we recover the remaining component using the temporal gauge

$$\hat{A}_0 = -\frac{1}{2}\Omega^2\hat{A}_1. \tag{3.3.17}$$

Clearly these components then define $H^1 \oplus L^2$ initial data on $\mathcal{S}$, and so the solution $\hat{A}_a$ can be extended to the entire spacetime using standard existence theorems and energy estimates (see appendix A.2.1). It remains to check that, with the definition (3.3.17) of the remaining component $\hat{A}_0$, the Coulomb gauge condition $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$ is propagated from $\mathscr{I}^+$ into the interior of the spacetime. We set

$$\omega := \boldsymbol{\nabla} \cdot \mathbf{A} \quad \text{and} \quad \psi = R^{-2}\omega,$$

where we recall that near $\mathscr{I}^+$

$$\begin{aligned}\psi &= -\hat{D}(\chi^2\hat{A}_1) + \frac{1}{2}R^2\chi^4\hat{\Delta}\hat{A}_1 + \nabla_{\mathbb{S}^2} \cdot \hat{A}_{\mathbb{S}^2} \\ &= -2\chi'\chi\hat{A}_1 - \chi^2\partial_u\hat{A}_1 - R^2\chi^2\partial_R\hat{A}_1 + \nabla_{\mathbb{S}^2} \cdot \hat{A}_{\mathbb{S}^2}.\end{aligned}$$

Acting on scalars the wave operator near $\mathscr{I}^+$ is

$$\widehat{\Box} = \chi^{-2}\left(-2\partial_u\partial_R - 2\left(R + \chi'\chi^{-1}\right)\partial_R - R^2\partial_R^2 - \Delta_{\mathbb{S}^2}\right),$$

and, using the commutators

$$\begin{aligned}
[\widehat{\Box}, \partial_u] &= 2\chi'\chi^{-1}\widehat{\Box} + 2\chi^{-2}(\chi''\chi^{-1} - (\chi')^2\chi^{-2})\partial_R, \\
[\widehat{\Box}, \partial_R] &= 2\chi^{-2}(R\partial_R^2 + \partial_R), \\
[\widehat{\Box}, \nabla_{\mathbb{S}^2}\cdot] &= \chi^{-2}\nabla_{\mathbb{S}^2}\cdot,
\end{aligned}$$

a lengthy direct calculation using (3.3.32) shows that

$$\widehat{\Box}\psi - 2\chi'\chi^{-1}\hat{\Delta}\psi = -2R\chi^2\left(\widehat{\Box}\hat{A}_1 + 2\chi'\chi^{-1}\hat{\Delta}\hat{A}_1\right).$$

As a consequence of (3.3.29), the function ψ therefore satisfies the homogeneous wave equation

$$\widehat{\Box}\psi - 2\chi'\chi^{-1}\hat{\Delta}\psi = 0.$$

The function ψ has zero characteristic data on $\mathscr{I}^+$ by virtue of our definition (3.3.31) of the data for $\hat{A}_1$, and hence by the uniqueness of Hörmander's construction we have

$$\psi \equiv 0 \quad \text{in } J^+(\mathcal{S}).$$

Along with the temporal gauge, the fact that $\omega = R^2\psi = 0$ then implies the usual Coulomb gauge condition $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$ in the physical spacetime, which is propagated to the rest of the spacetime by standard results.

The Scattering Operator

By completing the function space $\mathcal{C}_c^\infty(\mathscr{I}^+)$ in the norm $\|\cdot\|_{\mathcal{H}^1(\mathscr{I}^+)}$ on $\mathscr{I}^+$ as outlined in section 3.3.3, we have therefore constructed a linear isomorphism

$$\mathfrak{T}^+ : \dot{H}^1_C(\Sigma) \oplus L^2_C(\Sigma) \longrightarrow \mathcal{H}^1(\mathscr{I}^+)$$

with inverse

$$(\mathfrak{T}^+)^{-1} : \mathcal{H}^1(\mathscr{I}^+) \longrightarrow \dot{H}^1_C(\Sigma) \oplus L^2_C(\Sigma).$$

An analogous construction can be performed to the past of the initial surface Σ to construct the *past trace operator*

$$\mathfrak{T}^- : \dot{H}^1_C(\Sigma) \oplus L^2_C(\Sigma) \longrightarrow \mathcal{H}^1(\mathscr{I}^-),$$

which is an isomorphism by the same token. We are now in a position to define the scattering operator $\mathscr{S}$.

Definition 3.3.12. The linear isomorphism of Hilbert spaces

$$\mathscr{S} := \mathfrak{T}^+ \circ (\mathfrak{T}^-)^{-1} : \mathcal{H}^1(\mathscr{I}^-) \longrightarrow \mathcal{H}^1(\mathscr{I}^+)$$

taking finite energy characteristic data for the Maxwell potential on $\mathscr{I}^-$ to finite energy characteristic data on $\mathscr{I}^+$ is called the *conformal scattering operator of Maxwell potentials in temporal-Coulomb gauge* on partially compactified Minkowski space.

Alternative Formulations

The preceding construction of the scattering operator $\mathscr{S}$ is predicated on the usage of the multiplier Killing vector field

$$K^a = \partial_u,$$

which, via natural energy estimates (cf. section 3.3.1), defines the norms on $\mathscr{I}^\pm$

$$\mathcal{E}^K_{\mathscr{I}^+} \simeq \int_{\mathscr{I}^+} |\hat{F}_0|^2 \, \widehat{\mathrm{dv}}_{\mathscr{I}^+} \simeq \int_{\mathscr{I}^+} \left(|\hat{\text{þ}} \hat{A}_2|^2 + \varpi_\chi^2 |\hat{A}_2|^2 \right) \widehat{\mathrm{dv}}_{\mathscr{I}^+},$$
$$\mathcal{E}^K_{\mathscr{I}^-} \simeq \int_{\mathscr{I}^-} |\hat{F}_2|^2 \, \widehat{\mathrm{dv}}_{\mathscr{I}^-} \simeq \int_{\mathscr{I}^-} \left(|\hat{\text{þ}}' \hat{A}_2|^2 + \varpi_\chi^2 |\hat{A}_2|^2 \right) \widehat{\mathrm{dv}}_{\mathscr{I}^-}.$$

However, one has many alternative choices for K^a on Minkowski space. Indeed, inspecting the proof of theorem 3.3.2, one sees that any uniformly timelike conformally[6] Killing vector field on $\mathscr{M}$ will do. One particular choice which is tied to the conformal structure of Minkowski space is the Morawetz vector field

$$K_0^a = u^2 \partial_u + 2r(u+r)\partial_r, \tag{3.3.33}$$

discovered by Cathleen Morawetz in 1961 in her study of the decay of solutions to the wave equation in the exterior of an obstacle [83]. The vector field K_0^a is conformally Killing on $(\mathscr{M}, \eta)$,

$$\pounds_{K_0} \eta_{ab} = 4(u+r)\eta_{ab},$$

and in fact exactly Killing with respect to $R^2 \eta_{ab}$,

$$\pounds_{K_0} (R^2 \eta_{ab}) = 0.$$

If one uses K_0^a instead of K^a in the energy estimates, one arrives at the following energies on Σ and $\mathscr{I}^\pm$,

$$\begin{aligned} \mathcal{E}^{K_0}_\Sigma &\simeq \int_\Sigma r^2 \left(|F_0|^2 + |F_1|^2 + |F_2|^2 \right) \mathrm{dv}_\Sigma \\ &\simeq \int_\Sigma r^2 \left(|\mathbf{E}|^2 + |\mathbf{B}|^2 \right) \mathrm{dv}_\Sigma \\ &= \int_\Sigma \left(r^2 |\dot{\mathbf{A}}|^2 + r^2 |\boldsymbol{\nabla} \mathbf{A}|^2 - 2|\mathbf{A}|^2 \right) \mathrm{dv}_\Sigma, \end{aligned} \tag{3.3.34}$$

and

$$\mathcal{E}^{K_0}_{\mathscr{I}^+} \simeq \int_{\mathscr{I}^+} \left(u^2 |\hat{F}_0|^2 + \chi^2 |\hat{F}_1|^2 \right) \widehat{\mathrm{dv}}_{\mathscr{I}^+},$$

which may be written[7], after a residual conformal rescaling of $\mathscr{I}^+$, as

$$\mathcal{E}^{K_0}_{\mathscr{I}^+} \simeq \int_{\mathscr{I}^+} \left(u^2 |\hat{\text{þ}} \hat{A}_2|^2 + |\hat{\eth} \bar{\hat{A}}_2|^2 \right) \widehat{\mathrm{dv}}_{\mathscr{I}^+}. \tag{3.3.35}$$

[6]The multiplier vector field is allowed to be merely *conformally* Killing on $\mathscr{M}$ because the Maxwell stress-energy tensor is traceless. This is not the case in a theory that involves a scalar field, see remark 2.2.1.

[7]There is a slight subtlety here in converting the energy on $\mathscr{I}^+$ into a norm on $\hat{A}_2$. In writing the field components $\hat{F}_{0,1}$ in terms of the potential, it turns out to be useful to conformally rescale $\mathscr{I}^+$ to set $\chi \equiv 1$ everywhere on $\mathscr{I}^+$, which in particular sets $\varpi_\chi^2 = 0$ and puts i^+ at infinity. The energy estimates work in the same way, but the residual rescaling allows us to use the reduced temporal-Coulomb gauge (3.3.22) to obtain the formula (3.3.35).

A similar expression exists on $\mathscr{I}^-$. While the energies $\mathcal{E}^{K_0}_{\mathscr{I}^\pm}$ on $\mathscr{I}^\pm$ define weighted Sobolev norms in terms of $\hat{A}_2$, the energy $\mathcal{E}^{K_0}_{\Sigma}$ on Σ no longer defines a (weighted) Sobolev norm in terms of $(\dot{\mathbf{A}}, \mathbf{A})$ due to the presence of the negative-definite term $-2|\mathbf{A}|^2$. This means that the space of data on Σ has to be defined slightly differently in this context. As before, we have the trace operators

$$\begin{aligned}\mathfrak{T}^\pm_{K_0} : \mathcal{D}^\infty_c(\Sigma) &\longrightarrow \mathcal{C}^\infty(\mathscr{I}^\pm),\\ (\mathbf{A}, \dot{\mathbf{A}}) &\longmapsto \hat{A}^\pm_2\end{aligned}$$

from smooth initial data to smooth characteristic data, but instead of completing $\mathcal{D}^\infty_c(\Sigma)$ in the norm $\dot{H}^1 \oplus L^2$, we shall show that the pairs $(\mathbf{A}, \dot{\mathbf{A}})$ are in 1-to-1 correspondence with finite energy Maxwell fields on Σ in the natural energy space (3.3.34). Indeed, in Coulomb gauge on Σ one has

$$\begin{aligned}\mathbf{E} &= \dot{\mathbf{A}} \in \mathcal{C}^\infty_c(\Sigma),\\ \mathbf{B} &= \frac{1}{2}\boldsymbol{\nabla} \times \mathbf{A} \in \mathcal{C}^\infty_c(\Sigma).\end{aligned}$$

The time derivative component is therefore recovered trivially, whereas to recover $\mathbf{A}$ from $\mathbf{B}$ on Σ we take the curl,

$$\boldsymbol{\nabla} \times \mathbf{B} = \frac{1}{2}\left(\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) - \boldsymbol{\Delta}\mathbf{A}\right) = -\frac{1}{2}\boldsymbol{\Delta}\mathbf{A}. \tag{3.3.36}$$

For $\mathbf{B} \in \mathcal{C}^\infty_c(\Sigma)$ there exists a unique solution $\mathbf{A} \in \mathcal{C}^\infty_c(\Sigma)$ to (3.3.36), which we write as $\mathbf{A} = \boldsymbol{\Delta}^{-1}(-2\boldsymbol{\nabla}\times\mathbf{B})$. If $r\mathbf{E} \in L^2(\Sigma)$, it is obvious that $r\dot{\mathbf{A}} \in L^2(\Sigma)$. Also, $r\mathbf{B} \in L^2(\Sigma) \implies \boldsymbol{\nabla}\times\mathbf{B} \in H^{-1}_{\text{loc}}(\Sigma)$, and so by standard elliptic regularity results (theorem A.1.4) $\mathbf{A} \in H^1_{\text{loc}}(\Sigma)$. We define

$$r^{-1}H^1_C(\Sigma)^{\text{curl}} := \left\{\mathbf{A} \in H^1_{\text{loc}}(\Sigma) \,:\, \boldsymbol{\nabla}\cdot\mathbf{A} = 0,\ r(\boldsymbol{\nabla}\times\mathbf{A}) \in L^2(\Sigma)\right\}$$

and

$$r^{-1}L^2_C(\Sigma) := \left\{\dot{\mathbf{A}} \in L^2_{\text{loc}}(\Sigma) \,:\, \boldsymbol{\nabla}\cdot\dot{\mathbf{A}} = 0,\ r\dot{\mathbf{A}} \in L^2(\Sigma)\right\}.$$

Then the operator $\mathfrak{T}^+_{K_0}$ extends as an isomorphism

$$\mathfrak{T}^+_{K_0} : r^{-1}H^1_C(\Sigma)^{\text{curl}} \oplus r^{-1}L^2_C(\Sigma) \longrightarrow u^{-1}\mathcal{H}^1(\mathscr{I}^+),$$

where $u^{-1}\mathcal{H}^1(\mathscr{I}^+)$ is the completion of $\mathcal{C}^\infty_c(\mathscr{I}^+)$ in the norm $(\mathcal{E}^{K_0}_{\mathscr{I}^+})^{1/2}$, and similarly for $\mathfrak{T}^-_{K_0}$. We then define the scattering operator associated to K_0 by

$$\mathscr{S}_{K_0} := \mathfrak{T}^+_{K_0} \circ (\mathfrak{T}^-_{K_0})^{-1} : v^{-1}\mathcal{H}^1(\mathscr{I}^-) \longrightarrow u^{-1}\mathcal{H}^1(\mathscr{I}^+),$$

which is again an isomorphism of Hilbert spaces.

Remark 3.3.13. Notice that the space $r^{-1}H^1_C(\Sigma)^{\text{curl}} \oplus r^{-1}L^2_C(\Sigma)$ is a subspace of $H^1_C(\Sigma) \oplus L^2_C(\Sigma)$, and that $u^{-1}\mathcal{H}^1(\mathscr{I}^+)$ is a subspace of $\mathcal{H}^1(\mathscr{I}^+)$. In other words, the vector field K^a defines a weaker—more general—scattering theory between $\mathscr{I}^-$ and $\mathscr{I}^+$ than the vector field K^a_0. The construction for K^a_0 shows that the faster-decaying characteristic data on $\mathscr{I}^-$ scatters to the correspondingly faster-decaying characteristic data on $\mathscr{I}^+$.

Indeed, the scattering operator $\mathscr{S}$ maps data that is $\hat{F}_2^- = \mathcal{O}(v^{-1})$ on $\mathscr{I}^-$, through data that is $F_{0,1,2} = \mathcal{O}(r^{-2})$ on Σ, to data that is $\hat{F}_0^+ = \mathcal{O}(u^{-1})$ on $\mathscr{I}^+$,

$$\mathscr{S} \,:\, \hat{F}_2^- = \mathcal{O}(v^{-1}) \overset{(\mathfrak{T}^-)^{-1}}{\rightsquigarrow} F_{0,1,2} = \mathcal{O}(r^{-2}) \overset{\mathfrak{T}^+}{\rightsquigarrow} \hat{F}_0^+ = \mathcal{O}(u^{-1}).$$

Equivalently, in terms of the potential

$$\mathscr{S} \,:\, \hat{A}_2^- = \mathcal{O}(\log v) \overset{(\mathfrak{T}^-)^{-1}}{\rightsquigarrow} \mathbf{A} = \mathcal{O}(r^{-1}),\ \dot{\mathbf{A}} = \mathcal{O}(r^{-2}) \overset{\mathfrak{T}^+}{\rightsquigarrow} \hat{A}_2^+ = \mathcal{O}(\log u).$$

On the other hand,

$$\mathscr{S}_{K_0} \,:\, \hat{F}_2^- = \mathcal{O}(v^{-2}) \overset{\left(\mathfrak{T}^-_{K_0}\right)^{-1}}{\rightsquigarrow} F_{0,1,2} = \mathcal{O}(r^{-3}) \overset{\mathfrak{T}^+_{K_0}}{\rightsquigarrow} \hat{F}_0^+ = \mathcal{O}(u^{-2}),$$
$$\mathscr{S}_{K_0} \,:\, \hat{A}_2^- = \mathcal{O}(v^{-1}) \overset{\left(\mathfrak{T}^-_{K_0}\right)^{-1}}{\rightsquigarrow} \mathbf{A} = \mathcal{O}(r^{-2}),\ \dot{\mathbf{A}} = \mathcal{O}(r^{-3}) \overset{\mathfrak{T}^+_{K_0}}{\rightsquigarrow} \hat{A}_2^+ = \mathcal{O}(u^{-1}).$$

3.4 Curved Spacetimes

We now turn to the case of Corvino–Schoen–Chruściel–Delay spacetimes $(\mathcal{M}, g)$ as described in section 3.2. By design, much of the work we have done in the case of partially compactified Minkowski space will assist us here in extending the scattering theory of Mason & Nicolas [78] to Maxwell potentials. As before, we shall choose the conformal factor Ω so that $\hat{\mathcal{M}}$ is partially compactified: i^+ is at a finite distance and $\mathscr{I}^+$ is the backwards lightcone of i^+, and $\Omega = R$ in a Schwarzschildean neighbourhood of i^0.

3.4.1 Structure of $\mathscr{I}$

The topological structure of null infinity of all asymptotically flat asymptotically simple spacetimes is the same, and essentially identical to the topology of null infinity of Minkowski space [94, 96]. Indeed, one has the following theorem.

Theorem 3.4.1 (Penrose, 1965)**.** *In any asymptotically simple spacetime $\mathcal{M}$ for which $\mathscr{I}$ is everywhere null, the topology of each of $\mathscr{I}^\pm$ is given by*

$$\mathscr{I}^+ \simeq \mathscr{I}^- \simeq \mathbb{R} \times \mathbb{S}^2,$$

and the rays generating $\mathscr{I}^\pm$ can be taken to be the $\mathbb{R}$ factors.

For these spacetimes future (or past) null infinity $\mathscr{I}^+$ is therefore a null 3-dimensional manifold ruled by the integral curves of $\hat{l}^a \propto \hat{\nabla}^a\Omega$. The geometry of the interior of the spacetime $\mathcal{M}$ provides the following structure on null infinity. The pullback $\hat{q}_{ab}$ to $\mathscr{I}^+$ of the metric $\hat{g}_{ab}$ gives a degenerate metric on $\mathscr{I}^+$ which has signature $(0, -, -)$. Furthermore, one has a residual conformal freedom: if the original metric g_{ab} was related to the rescaled metric by $\hat{g}_{ab} = \Omega^2 g_{ab}$, then for any ω which is smooth and nowhere vanishing on $\mathscr{I}^+$ the rescaling $\Omega \rightsquigarrow \omega\Omega$ is still permissible; we have already employed such a residual rescaling in section 3.3, to derive the identity (3.3.35). If one considers $\mathscr{I}^+$ as an abstract 3-manifold with topology $\mathbb{R} \times \mathbb{S}^2$, a pair of fields $(\hat{q}_{ab}, \hat{l}^a)$ satisfying a set of conditions forms the *universal structure* on $\mathscr{I}^+$ of asymptotically flat asymptotically simple spacetimes, alluded to in section 1.1 [5]. These conditions are that $\hat{l}^a$ be complete,

$\hat{q}_{ab}$ be of signature $(0,-,-)$, $\hat{q}_{ab}\hat{l}^b = 0$, $\pounds_{\hat{l}}\hat{q}_{ab} = 0$, and furthermore that any two such pairs $(\hat{q}_{ab}, \hat{l}^a)$ and $(\check{q}_{ab}, \check{l}^a)$ be related[8] by

$$\check{q}_{ab} = \omega^2 \hat{q}_{ab} \quad \text{and} \quad \check{l}^a = \omega^{-1}\hat{l}^a.$$

The BMS groups $\mathfrak{B}$ is then characterized intrinsically on $\mathscr{I}^+$ (as opposed to with reference to the asymptotic behaviour of the symmetries of $\mathscr{M}$) as the subgroup of diffeomorphisms of $\mathscr{I}^+$ which preserves this universal structure.

It turns out that further information concerning the structure of $\mathscr{I}^+$ is provided by the fact that the physical spacetime $\mathscr{M}$ is asymptotically vacuum. The trace-free part Φ_{ab} of the Ricci tensor

$$\Phi_{ab} := -\frac{1}{2}\left(\mathrm{R}_{ab} - \frac{1}{4}\mathrm{R}g_{ab}\right)$$

transforms under a conformal transformation as

$$\begin{aligned}\Phi_{ab} &= \hat{\Phi}_{ab} + \hat{\nabla}_a\Upsilon_b - \frac{1}{4}\hat{g}_{ab}\hat{\nabla}^c\Upsilon_c + \Upsilon_a\Upsilon_b - \frac{1}{4}\hat{g}_{ab}\hat{g}^{cd}\Upsilon_c\Upsilon_d \\ &= \hat{\Phi}_{ab} + \Omega^{-1}\hat{\nabla}_a\hat{\nabla}_b\Omega - \frac{1}{4}\Omega^{-1}\hat{g}_{ab}\hat{\nabla}^c\hat{\nabla}_c\Omega.\end{aligned}$$

One has, according to definition 3.2.1, that $\Omega^{-2}\mathrm{R}_{ab}$ has a continuous limit on $\mathscr{I}^+$, so multiplying the above by Ω and taking the limit $\Omega \to 0$ ensures that $\Omega\Phi_{ab} \approx 0$, and gives the *asymptotic Einstein condition*

$$\hat{\nabla}_a\hat{\nabla}_b\Omega \approx \frac{1}{4}\hat{g}_{ab}\hat{\nabla}^c\hat{\nabla}_c\Omega. \tag{3.4.1}$$

The normal $\hat{\nabla}^b\Omega$ to $\mathscr{I}^+$ is proportional to $\hat{l}^a$, $\hat{\nabla}^b\Omega \approx f\hat{l}^b$ for some non-vanishing scalar function f, so the condition (3.4.1) reads

$$f\hat{\nabla}_a\hat{l}_b + \hat{l}_b\hat{\nabla}_a f \approx \frac{1}{4}\hat{g}_{ab}(f\hat{\nabla}_c\hat{l}^c + \hat{D}f). \tag{3.4.2}$$

Multiplying by $\hat{l}_c$ and antisymmetrizing shows that (see [96], (7.1.58))

$$\hat{l}_{[a}\hat{\nabla}_b\hat{l}_{c]} \approx 0 \iff (\hat{\kappa} \approx 0, \quad \hat{\rho} \approx \bar{\hat{\rho}}), \tag{3.4.3}$$

where the conditions on the spin coefficients $\hat{\kappa}$ and $\hat{\rho}$ may be rapidly obtained from the hypersurface orthogonal condition by contracting with $\hat{l}^a\hat{m}^b$ and $\hat{m}^{[a}\bar{\hat{m}}^{b]}$ respectively. The vanishing of the spin coefficient $\hat{\kappa}$ on $\mathscr{I}^+$ tells us that $\mathscr{I}^+$ is generated by null geodesics, whereas the condition $\hat{\rho} \approx \bar{\hat{\rho}}$ says that the vectors $\hat{l}^a$ are *twist-free* on $\mathscr{I}^+$. Contracting (3.4.2) with $\hat{m}^a\hat{m}^b$, we also get

$$\hat{\sigma} \approx 0, \tag{3.4.4}$$

which is the statement that the vectors $\hat{l}^a$ are *shear-free* on $\mathscr{I}^+$. We say the hypersurface $\mathscr{I}^+$ is geodetic, twist-free and shear-free. Since the vectors $\hat{l}^a$ are geodetic on $\mathscr{I}^+$, they are parallely propagated, $\hat{D}\hat{l}^a = \hat{\nabla}_{\hat{l}}\hat{l}^a \approx s\hat{l}^a$ for some function s, which vanishes identically if

[8] Here we have secretly chosen a conformal frame in which $\hat{\nabla}_a\hat{l}_b = 0$, though we do not work in such a frame in general.

the geodesics are affinely parametrized. Contracting with $\hat{n}^a$, one sees that the function s is given by $s = \hat{n}^a \hat{D} \hat{l}_a$. This is in fact the real part of another spin coefficient, $\hat{\varepsilon} + \bar{\hat{\varepsilon}} = \hat{n}^a \hat{D} \hat{l}_a$, so the condition for the geodesics generated by $\hat{l}^a$ on $\mathscr{I}^+$ to be affinely parametrized is

$$\hat{\varepsilon} + \bar{\hat{\varepsilon}} \approx 0. \tag{3.4.5}$$

It is always possible to reparametrize a geodesic affinely, and here we will assume that the original parametrization has been made to that effect. The condition for the imaginary part of $\hat{\varepsilon}$ to vanish, $\hat{\varepsilon} - \bar{\hat{\varepsilon}} = 0$, can be translated as the statement that the spinor field $\hat{o}^A$ has parallelly propagated flag planes, where $\hat{l}^a = \hat{o}^A \hat{o}^{A'}$. If $\mathscr{I}^+$ is affinely parametrized with parallelly propagated flag planes, then of course $\hat{\varepsilon} \approx 0$.

Another consequence of being asymptotically Einstein is that the Weyl tensor vanishes[9] on null infinity, $\hat{\mathrm{C}}_{abcd} \approx 0$ ([96], (9.6.32)).

3.4.2 Scattering of Maxwell Fields

A conformal scattering theory for Maxwell fields has been constructed by Mason and Nicolas [78]; we do not re-prove their results here. We will, however, make use of their energy estimates and their resolution of the Goursat problem for $\hat{F}_{ab}$.

Energy Estimates

Theorem 3.4.2 (Appendix A of [78]). *For smooth compactly supported Maxwell data on* Σ *one has the energy estimate*

$$\mathcal{E}_{\mathscr{I}^+} \simeq \int_{\mathscr{I}^+} |\hat{F}_0|^2 \,\widehat{\mathrm{dv}}_{\mathscr{I}^+} \simeq \int_{\Sigma} \left(|F_0|^2 + |F_1|^2 + |F_2|^2 \right) \mathrm{dv}_{\Sigma} \simeq \mathcal{E}_{\Sigma}. \tag{3.4.6}$$

Sketch proof. Estimate (3.4.6) is proved by considering the regions near i^0 and i^+ separately. Near i^0 one uses the fact that the data is compactly supported, and employs the exact Killing field $K^a = \partial_t$ in the Schwarzschild spacetime; it is conformally Killing in the rescaled spacetime and the stress-energy tensor (2.2.6) is trace-free, which ensures that the estimate near i^0 has no error terms, $\hat{\nabla}^b (K^a \hat{\mathbf{T}}_{ab}) = 0$. Near i^+ one has to use special features of $\mathscr{I}$ and the multiplier vector field $K^a = \hat{\nabla}^a \Omega$. One chooses a conformal scale such that i^+ is a finite point and such that $\hat{\mathrm{R}}_{ab} = 0$ at i^+, and $\hat{l}^a \hat{\mathrm{R}}_{ab} = 0 = \hat{\mathrm{R}}$ on $\mathscr{I}^+$. One then works on a foliation of the level sets of Ω near $\mathscr{I}^+$ to show that the Killing form of $\hat{\nabla}^a \Omega$ decays sufficiently fast to close the estimate. □

Solving the Goursat Problem

Theorem 3.4.3 (Appendix C of [78]). *For characteristic data* $\hat{F}_0^+ \in L^2(\mathscr{I}^+)$ *the full solution*

$$F_0, F_1, F_2 \in \mathcal{C}^0(\mathbb{R}_t; L^2(\Sigma_t))$$

on the physical spacetime $\mathcal{M}$ *is uniquely recoverable and satisfies the Maxwell constraint equations on each spacelike hypersurface* Σ_t.

[9] A more general version of this statement is provided by the peeling theorem [67, 69], which states that for suitably regular spacetimes the components of the Weyl tensor Ψ_k, $0 \leqslant k \leqslant 4$, satisfy the decay conditions $\Psi_k = \mathcal{O}(r^{-5+k})$ as $r \to +\infty$ for an appropriately defined radial coordinate r.

Sketch proof. The proof is essentially a rephrasing of the ideas presented in appendix A.2.2, the main tool being the energy estimates. Some care is required to ensure that the solution satisfies the constraint equations, which is achieved by projecting a smooth extension of the Taylor series of the solution to Maxwell's equations onto a solution of the constraint equations. □

3.4.3 Extension to Potentials

In a generic curved spacetime of Corvino–Schoen–Chruściel–Delay type, if one imposes the Coulomb gauge $\boldsymbol{\nabla}^\alpha \mathbf{A}_\alpha = 0$ on the slices (Σ_t, h_{ab}), the component $\mathfrak{a} = T^a A_a$ no longer satisfies an unsourced elliptic equation as in the flat case. Instead, $\mathfrak{a}$ satisfies an equation of the form

$$\boldsymbol{\Delta}\mathfrak{a} = \kappa \cdot f + (\nabla\kappa) \cdot g$$

for sources f and g. The presence of the extrinsic curvature κ therefore generically prevents $\mathfrak{a}$ from being zero, making the Coulomb and temporal gauges incompatible. What we *can* do in generic curved spacetimes is to impose the temporal gauge throughout the physical spacetime $\mathscr{M}$, and supplement it with the Coulomb gauge only on the initial surface Σ. This is clearly possible, as the time derivative of the gauge transformation $\nabla_T \chi$ is fixed by imposing the temporal gauge, but one is free to choose $\chi|_\Sigma$ in such a way as to ensure $\boldsymbol{\nabla}_\alpha \mathbf{A}^\alpha|_\Sigma = 0$. The temporal gauge in the physical spacetime $\mathscr{M}$ is $\mathfrak{a} = 0$, and in terms of NP components reads

$$0 = T^a A_a = aA_0 + \frac{1}{2a} A_1 = a\hat{A}_0 + \frac{\Omega^2}{2a}\hat{A}_1$$

for some function a which is bounded with bounded reciprocal on $\mathscr{I}^+$. For smooth solutions $\hat{A}_a$ on $\hat{\mathscr{M}}$, the temporal gauge therefore extends to null infinity as

$$\hat{A}_0 \approx 0. \tag{3.4.7}$$

This condition is necessary for our treatment of the energy on $\mathscr{I}^+$.

Remark 3.4.4. It turns out that the physical Lorenz gauge $\nabla_a A^a = 0$ also reduces to $\hat{A}_0 \approx 0$ on $\mathscr{I}^+$, but it is prone to introducing conformal singularities. Indeed, under a conformal transformation

$$\nabla_a A^a = \Omega^2(\hat{\nabla}_a \hat{A}^a - 2\Upsilon_a \hat{A}^a),$$

where the singular quantity $\Upsilon_a \hat{A}^a \sim \Omega^{-1}(\partial_a \Omega)\hat{A}^a$ appears in the rescaled field equations (2.2.5). On the other hand, the temporal gauge is only tied to the foliation $\{\Sigma_t\}_t$, and is in that sense conformally invariant.

Energy on $\mathscr{I}^+$

As in section 3.3.3, the condition (3.4.7) also implies $\eth \hat{A}_0 \approx 0$, and by (3.4.3) and (3.4.4), in the temporal gauge the expression (3.2.16) for $\hat{F}_0$ reduces to $\hat{F}_0 \approx \hat{\text{þ}}\hat{A}_2 - \hat{\rho}\hat{A}_2$. Choosing an affine parametrization of $\mathscr{I}^+$ and a spin frame $(\hat{o}^A, \hat{\iota}^A)$ such that $\hat{l}^a = \hat{o}^A\hat{o}^{A'}$ and such that $\hat{o}^A$ has parallely propagated flag planes, one can ensure that $\hat{\varepsilon} \approx 0$, and

$$\hat{F}_0 \approx \hat{D}\hat{A}_2 - \hat{\rho}\hat{A}_2. \tag{3.4.8}$$

As before, for $\hat{A}_2 \in \mathcal{C}_c^\infty(\mathscr{I}^+)$, squaring (3.4.8), integrating by parts and using the Sachs equation (3.3.23) then gives

$$\int_{\mathscr{I}^+} |\hat{F}_0|^2 \, \widehat{\mathrm{dv}}_{\mathscr{I}^+} = \int_{\mathscr{I}^+} \left(|\hat{D}\hat{A}_2|^2 + (2|\hat{\rho}|^2 + \hat{\Phi}_{00})|\hat{A}_2|^2 \right) \widehat{\mathrm{dv}}_{\mathscr{I}^+}.$$

Proposition 3.4.5. *There exists a conformal scale on $\mathscr{I}^+$ in which $\hat{\Phi}_{00} \approx 0$.*

Proof. Suppose a conformal transformation bringing $\mathscr{I}^+$ and i^+ to a finite distance has been performed, giving the first rescaling of the metric $\check{g}_{ab} = \Omega^2 g_{ab}$. Consider a second conformal rescaling given by

$$\hat{g}_{ab} = \omega^2 \check{g}_{ab} = \omega^2\Omega^2 g_{ab}.$$

We use the conformal transformation rule of the trace-free part of the Ricci tensor,

$$\check{\Phi}_{ab} = \hat{\Phi}_{ab} + \omega^{-1}\hat{\nabla}_a\hat{\nabla}_b\omega - \frac{1}{4}\omega^{-1}\hat{g}_{ab}\hat{\Box}\omega.$$

Contracting this with $\hat{l}^a\hat{l}^b = \check{l}^a\check{l}^b$ and noting that $\hat{\varepsilon} \approx 0 \approx \hat{\kappa}$, one obtains the second order ODE on $\mathscr{I}^+$

$$\check{\Phi}_{00} \approx \hat{\Phi}_{00}\omega + \check{D}^2\omega,$$

which contains only derivatives which are tangential to $\mathscr{I}^+$. By solving

$$\check{D}^2\omega - \check{\Phi}_{00}\omega \approx 0$$

along $\mathscr{I}^+$ (for data prescribed at i^+, for example $\omega(i^+) = 1$ and $(\check{D}\omega)(i^+) = 0$), one obtains the values of a conformal factor ω on $\mathscr{I}^+$ that sets $\hat{\Phi}_{00} \approx 0$. There are various ways of extending ω into the interior of the spacetime. By considering the conformal transformation of the scalar curvature,

$$\hat{\mathrm{R}} = \omega^{-2}(\check{\mathrm{R}} + 6\omega^{-1}\check{\Box}\omega),$$

one observes that an extension is provided by solving the characteristic Cauchy problem from $\mathscr{I}^+$ for ω satisfying the conformally invariant wave equation. In addition, this sets the newly rescaled scalar curvature to zero, $\hat{\mathrm{R}} = 0$, at least where this extension is adopted. □

Using proposition 3.4.5, one obtains the expression

$$\mathcal{E}_{\mathscr{I}^+} \simeq \int_{\mathscr{I}^+} |\hat{F}_0|^2 \, \widehat{\mathrm{dv}}_{\mathscr{I}^+} = \int_{\mathscr{I}^+} \left(|\hat{D}\hat{A}_2|^2 + 2|\hat{\rho}|^2|\hat{A}_2|^2 \right) \widehat{\mathrm{dv}}_{\mathscr{I}^+} =: \|\hat{A}_2\|^2_{\mathcal{H}^1(\mathscr{I}^+)} \tag{3.4.9}$$

for the energy on $\mathscr{I}^+$ in terms of the potential.

Definition 3.4.6. For the component $\hat{A}_2$ of the Maxwell potential we define the function space $\mathcal{H}^1(\mathscr{I}^+)$ on $\mathscr{I}^+$ by requiring that the norm

$$\|\hat{A}_2\|^2_{\mathcal{H}^1(\mathscr{I}^+)} := \int_{\mathscr{I}^+} \left(|\hat{D}\hat{A}_2|^2 + 2|\hat{\rho}|^2|\hat{A}_2|^2 \right) \widehat{\mathrm{dv}}_{\mathscr{I}^+}$$

be finite. The space $\mathcal{H}^1(\mathscr{I}^+)$ will be the space of characteristic data on $\mathscr{I}^+$ for the free Maxwell's equations.

Energy on Σ and Reconstruction of Data

As before, we construct the space of data on Σ by working with the physical potential A_a. The energy estimates of theorem 3.4.2 are obtained by choosing the multiplier vector field K^a so that $K^a = kT^a$ on Σ for some smooth function k satisfying $c^{-1} \leqslant k \leqslant c$ for some $c > 0$. In the Schwarzschild sector this function will be equal to $k = \sqrt{F(r)}$, so that $K^a = \partial_t$. The initial energy is then

$$\mathcal{E}_\Sigma = \int_\Sigma kT^aT^b\mathbf{T}_{ab}\,\mathrm{dv}_\Sigma = \int_\Sigma \frac{k}{2}(|\mathbf{E}|^2 + |\mathbf{B}|^2)\,\mathrm{dv}_\Sigma \simeq \int_\Sigma \frac{1}{2}(|\mathbf{E}|^2 + |\mathbf{B}|^2)\,\mathrm{dv}_\Sigma\,.$$

In the temporal gauge $\mathfrak{a} = 0$ we can express the electric field $\mathbf{E}_a$ in terms of the potential as

$$E_a = T^b(\nabla_b A_a - \nabla_a A_b) = \nabla_T A_a - A_b(T_a\nabla_T T^b + \kappa^b{}_a),$$

which upon projection to Σ maintains the same form,

$$\mathbf{E}_a = \dot{\mathbf{A}}_a + \mathbf{A}_b T_a \nabla_T T^b + \kappa_a{}^b\mathbf{A}_b, \tag{3.4.10}$$

where $\mathbf{A}_a = -h_a^\alpha A_\alpha$ and $\dot{\mathbf{A}}_a = \nabla_T \mathbf{A}_a$. As for the magnetic part, we note that $h_b^\beta h_c^\gamma F_{\beta\gamma} = \mathbf{F}_{bc} = \boldsymbol{\nabla}_b\mathbf{A}_c - \boldsymbol{\nabla}_c\mathbf{A}_b$, and so

$$\mathbf{B}_a = \varepsilon_a{}^{bc}\boldsymbol{\nabla}_b\mathbf{A}_c, \tag{3.4.11}$$

where ε_{abc} is the volume form on Σ. It turns out that it is not possible to easily convert the energy $\mathcal{E}_\Sigma$ to a norm in the potential, even in the Coulomb gauge, due to curvature terms that arise when integrating by parts. Specifically, a term involving the Ricci curvature $-\mathbf{R}_{ab}$ of Σ appears, which is non-positive definite on Schwarzschild. Instead, we adopt the method used in section 3.3.3.

Contracting (3.4.11) with the volume form on Σ and differentiating, one obtains

$$\boldsymbol{\nabla}^c(\mathbf{B}_a\varepsilon^a{}_{bc}) = \boldsymbol{\nabla}_b(\boldsymbol{\nabla}\cdot\mathbf{A}) - \mathbf{R}_{ab}\mathbf{A}^a - \boldsymbol{\Delta}\mathbf{A}_b,$$

which, after imposing the Coulomb gauge on Σ, becomes the elliptic equation

$$\boldsymbol{\Delta}\mathbf{A}_b + \mathbf{R}_{ab}h^{ac}\mathbf{A}_c = -\boldsymbol{\nabla}^c(\mathbf{B}_a\varepsilon^a{}_{bc}). \tag{3.4.12}$$

By standard elliptic regularity results (theorem A.1.4), for $\mathbf{B} \in L^2(\Sigma)$ there exists a unique $\mathbf{A} \in H^1_{\mathrm{loc}}(\Sigma)$, which will define part of the potential initial data on Σ given $(\mathbf{E},\mathbf{B})|_\Sigma$. The remaining part of the initial data is obtained from (3.4.10),

$$\dot{\mathbf{A}}_a|_\Sigma = \left(\mathbf{E}_a - \mathbf{A}_b T_a\nabla_T T^b + \kappa_a{}^b\mathbf{A}_b\right)\Big|_\Sigma,$$

which is $L^2_{\mathrm{loc}}(\Sigma)$ if $\mathbf{E} \in L^2_{\mathrm{loc}}(\Sigma)$.

On $\mathscr{I}^+$, the characteristic data $\hat{A}_2^+$ is easily reconstructed from $\hat{F}_0^+$ using the ODE (3.4.8). By construction, for $\hat{F}_0^+ \in L^2(\mathscr{I}^+)$ the characteristic data for the potential is $\hat{A}_2^+ \in \mathcal{H}^1(\mathscr{I}^+)$, where the space $\mathcal{H}^1(\mathscr{I}^+)$ is as in definition 3.4.6.

The Trace and Scattering Operators

We define the space of initial data by

$$H^1_C(\Sigma)^{\text{curl}} \oplus L^2_C(\Sigma) := \Big\{(\mathbf{A}, \dot{\mathbf{A}}) \in H^1_{\text{loc}}(\Sigma) \oplus L^2_{\text{loc}}(\Sigma) \text{ s.t. } \boldsymbol{\nabla} \cdot \mathbf{A} = 0, \\ \text{and } \|\mathbf{E}\|_{L^2(\Sigma)} + \|\mathbf{B}\|_{L^2(\Sigma)} < \infty\Big\}, \tag{3.4.13}$$

where $\mathbf{E}$ and $\mathbf{B}$ are defined on Σ by (3.4.10) and (3.4.11). The space (3.4.13) is a normed vector space when endowed with the norm $\|\mathbf{E}\|_{L^2} + \|\mathbf{B}\|_{L^2}$. It is straightforward to check that whenever this vanishes, it follows that $(\mathbf{A}, \dot{\mathbf{A}}) = (0, 0)$. Also, since the relations (3.4.10) and (3.4.11) are linear in $\mathbf{A}$ and $\dot{\mathbf{A}}$, the above norm is subadditive and 1-homogeneous in $(\mathbf{A}, \dot{\mathbf{A}})$. By theorem 3.4.2 and theorem 3.4.3, and the constructions in [78], it then follows that there exist invertible bounded linear operators

$$\begin{aligned} \mathfrak{T}^{\pm} : H^1_C(\Sigma)^{\text{curl}} \oplus L^2_C(\Sigma) &\longrightarrow \mathcal{H}^1(\mathscr{I}^{\pm}), \\ (\mathbf{A}, \dot{\mathbf{A}}) &\longmapsto \hat{A}_2^{\pm}, \end{aligned}$$

and the associated scattering operator

$$\begin{aligned} \mathscr{S} : \mathcal{H}^1(\mathscr{I}^-) &\longrightarrow \mathcal{H}^1(\mathscr{I}^+), \\ \hat{A}_2^- &\longmapsto \hat{A}_2^+. \end{aligned}$$

As before, $\mathscr{S}$ is an isomorphism.

4

The Maxwell-Scalar Field System on de Sitter Space

4.1 Introduction

Although asymptotically flat spacetimes are well suited for studying scattering[1] around isolated systems, astronomical observations have by now shown that the cosmological constant λ in our universe is positive [4], though its magnitude is estimated to be only 10^{-122} [16, 98, 97, 99, 106]. However small, a positive cosmological constant nonetheless pins de Sitter space as the primary model for the shape of the empty universe. From the point of view of cosmology, and indeed quantum gravity, it is therefore natural to study scattering on asymptotically de Sitter spacetimes [112]. From an analytic point of view, it has been known since the work of Friedrich [47, 48] that de Sitter space is a stable solution of Einstein's equations with $\lambda > 0$. Moreover, a recent and much celebrated result of Hintz and Vasy has shown that Kerr–de Sitter black holes are stable [57]. The techniques of Hintz and Vasy are novel and involve much microlocal machinery, but also employ a conformal compactification (which is nonetheless of a different type to the one used here).

We have seen in section 2.3.2 that dS_4 is conformal to the Einstein cylinder $\mathfrak{E}$, the global time coordinate α being related to the conformal time τ by a fairly simple reparametrization. The vector field $T^a := \partial/\partial\tau$ is a uniformly timelike Killing field in $\mathfrak{E}$ (and therefore conformally Killing in dS_4), and provides a uniformly spacelike foliation of $\mathfrak{E}$ given by $\mathcal{F} = \{\mathbb{S}^3_\tau := \mathbb{S}^3 \times \{\tau\} : \tau \in \mathbb{R}\}$. In this chapter we prove explicit small data energy estimates for the Maxwell-scalar field system (2.2.22) on the Einstein cylinder with respect to the foliation $\mathcal{F}$. We begin with the natural conserved energy, which yields L^2 energy estimates for the Maxwell field and the gauge-covariant derivative of the scalar field. Using the smallness of the data and a carefully chosen gauge, we show that the natural energies may be controlled by Sobolev norms of the fields, and commute derivatives into the equations to derive higher order estimates. We discover an estimate

[1] For the Maxwell-scalar field system on Minkowski space, decay and stability results have been obtained by Lindblad–Sterbenz [76], Yang–Yu [125], and Candy–Kauffman–Lindblad [20].

algebra, which relies crucially on the nonlinearity in (2.2.22) being subcritical in four dimensions (the Maxwell-scalar field system (2.2.22) is critical in five dimensions, and *supercritical* in dimensions greater than five [53, 71]). We therefore show that small data solutions to the Maxwell-scalar field system exhibit the peeling property in the sense of definition 2.4.2, and that the optimal space of initial data for which the solution peels at order m is the space of $H^{m+1} \oplus H^m$ initial data. Furthermore, the estimates allow us to define, for each order of regularity m, bounded and invertible *small data* trace operators $\mathfrak{T}^{\pm}_m$ and a small data scattering operator $\mathscr{S}_m$. However, the trace and scattering operators are nonlinear, and we do not prove them continuous.

Conformally transforming the solutions to (2.2.22) back to physical de Sitter space, we deduce exponential decay rates along timelike curves approaching $\mathscr{I}^{\pm}$ (see fig. 2.6). We describe this behaviour both in terms of the global timelike coordinate α, and also in terms of the static coordinate t, defined in section 4.1.2. Using the static coordinate, we furthermore discover a certain asymptotic dispersion property whereby the slowest decaying mode—the part of the scalar field decaying like e^{-Ht}—uncouples from the Maxwell potential, and turns out to satisfy an uncharged wave equation.

This chapter is organized as follows. In the rest of this section we state our conventions and define a set of static coordinates on de Sitter space. In section 4.2 we state our results. In sections 4.3 and 4.4 we derive the required gauge fixing conditions, formulate the Cauchy problem, and state a classical existence theorem. Sections 4.5–4.7 contain the inductive energy estimates on which our results rest. Finally, in sections 4.8.1–4.8.3 we finish the proofs of the main results and derive the decay rates mentioned above.

4.1.1 Conventions

Our main estimates will be performed on the Einstein cylinder $(\mathfrak{E} = \mathbb{R} \times \mathbb{S}^3, \mathfrak{e})$. We will use the Roman indices $a, b, \dots$ to refer to tensors on $\mathfrak{E}$ and contractions with respect to the full spacetime metric $\mathfrak{e}$, and use the Greek indices $\mu, \nu, \dots$ to refer to tensors on $\mathbb{S}^3$ and contractions with respect to the metric $\mathfrak{s}_3$. At a certain point we will also use the indices i, j and k to refer to a basis of vector fields on $\mathbb{S}^3$, but this will be made explicit at the time. For ease of notation, we will depart slightly from the notation in the rest of this thesis and denote by ϕ and A_a (instead of $\hat{\phi}$ and $\hat{A}_a$) the scalar field and Maxwell potential on the Einstein cylinder, and by $\tilde{\phi}$ and $\tilde{A}_a$ the conformally related physical fields on de Sitter space,

$$\phi = \Omega^{-1}\tilde{\phi}, \qquad A_a = \tilde{A}_a. \tag{4.1.1}$$

More generally, unhatted quantities will refer to quantities on $\mathfrak{E}$, and tilded quantities will refer to dS_4. As $\mathfrak{e} = 1 \oplus (-\mathfrak{s}_3)$, we shall have $\nabla = \nabla^{\mathbb{R}} \oplus \nabla^{\mathfrak{s}_3} = \partial \oplus \not\nabla$. For a 1-form A on $\mathfrak{E}$ we will use $\mathbf{A}$ to denote the projection of A onto $\mathbb{S}^3$, A_0 to denote the component of A along ∂_τ, and dot (as in $\dot{\mathbf{A}}$) to denote differentiation with respect to conformal time τ. In this chapter L^p, $W^{m,p}$, H^m and $\mathcal{C}^{m,\delta}$ will denote the Lebesgue, Sobolev and Hölder spaces on $\mathbb{S}^3$, unless specifically stated otherwise.

4.1.2 Static Coordinates

The spacetime dS_4 is static, however this is not apparent from the coordinates (5.5.1). A more appropriate set of coordinates for an observer in de Sitter space is given by the following construction (a priori only valid in region I of fig. 2.6). We define new coordinates (t, r) by

$$r = \frac{\sin\zeta}{H\cos\tau} \in (0, \infty), \qquad \tanh(Ht) = \frac{\sin\tau}{\cos\zeta} \in (-\infty, \infty)$$

for $\tau \in (-\pi/2, \pi/2)$ and $\zeta \in (0, \pi)$. The metric on dS_4 in these coordinates then takes the form

$$\tilde{g} = F(r)\,\mathrm{d}t^2 - F(r)^{-1}\mathrm{d}r^2 - r^2\mathfrak{s}_2, \tag{4.1.2}$$

where $F(r) = (1 - H^2r^2)$. The cosmological horizons represented by the dashed lines in fig. 2.6 are now given by $\{r = 1/H\}$, $\mathscr{I}$ is given by $\{r = \infty\}$, the North and South Poles are at $\{r = 0\}$, and the four corners of the Penrose diagram are at $\{t = \pm\infty\}$. The vector field ∂_t is manifestly a timelike Killing vector in the region $\{r < 1/H\}$, but becomes null on the cosmological horizon $\{r = 1/H\}$. We may extend these coordinates to regions II, III and IV, however the vector field ∂_t starts to behave strangely in these regions. In particular, it is future-pointing in the region I, past-pointing in the region III, and spacelike in the regions II and IV. The arrows in fig. 4.1 represent the flow along ∂_t.

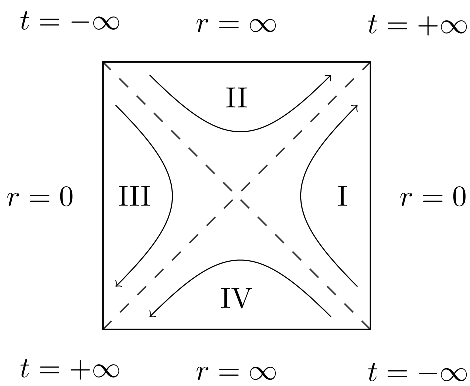

Figure 4.1: Static coordinates on dS_4.

Since all timelike geodesics in region I end up in the top right corner of fig. 4.1, that point plays the role of future timelike infinity i^+. In this sense the coordinates (4.1.2) single out a preferred point on $\mathscr{I}^+ \simeq \mathbb{S}^3$.

4.2 Main Theorems

4.2.1 Geometric and Approximate Energies

In this chapter we use the stress-energy tensor (2.2.23) for the system (2.2.22) on $\mathfrak{E}$. We will denote the Maxwell sector of the energy by

$$\mathbf{T}_{ab}[A] = -F_{ac}F_b{}^c + \frac{1}{4}\mathfrak{e}_{ab}F_{cd}F^{cd},$$

and the charged scalar field sector of the energy by

$$\mathbf{T}_{ab}[\phi] = \overline{\mathrm{D}_{(a}\phi}\mathrm{D}_{b)}\phi - \frac{1}{2}\mathfrak{e}_{ab}\overline{\mathrm{D}_c\phi}\mathrm{D}^c\phi + \frac{1}{2}\mathfrak{e}_{ab}|\phi|^2.$$

The full stress-energy tensor is then given by

$$\mathbf{T}_{ab}[\phi, A] = \mathbf{T}_{ab}[A] + \mathbf{T}_{ab}[\phi]. \tag{4.2.1}$$

By (2.2.24) and the fact that $\mathfrak{E}$ has constant scalar curvature, $\mathbf{T}_{ab}$ is conserved. It is therefore suitable for defining a conserved energy for the system (2.2.22),

$$\mathcal{E}_\tau[\phi, A] := \int_{\mathbb{S}^3_\tau} \mathbf{T}_{00}[\phi, A]\,\mathrm{dv}_{\mathfrak{s}_3} = \int_{\mathbb{S}^3_\tau} \mathbf{T}_{ab}[\phi, A]T^a T^b\,\mathrm{dv}_{\mathfrak{s}_3}\,. \tag{4.2.2}$$

Since $T^a = \partial_\tau$ is Killing on $\mathfrak{E}$, this satisfies

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{E}_\tau[\phi, A] = 0$$

if the field equations (2.2.22) are satisfied. We call (4.2.2) the *geometric* energy for the system (2.2.22). We also define the geometric energies for the individual sectors by

$$\mathcal{E}_\tau[\phi] := \int_{\mathbb{S}^3_\tau} \mathbf{T}_{00}[\phi]\,\mathrm{dv}_{\mathfrak{s}_3}, \qquad \mathcal{E}_\tau[A] := \int_{\mathbb{S}^3_\tau} \mathbf{T}_{00}[A]\,\mathrm{dv}_{\mathfrak{s}_3}\,.$$

The sectorial geometric energies $\mathcal{E}_\tau[\phi]$ and $\mathcal{E}_\tau[A]$ are not conserved individually and can exchange energy throughout the evolution, but of course the total geometric energy $\mathcal{E}_\tau[\phi, A] = \mathcal{E}_\tau[\phi] + \mathcal{E}_\tau[A]$ is. For $m \geqslant 1$ we also define the Sobolev-type approximate energies

$$\begin{aligned}
\mathrm{S}_m[\phi] &:= \|\dot{\phi}\|^2_{H^{m-1}} + \|\phi\|^2_{H^m}, & \mathrm{S}_m[A] &:= \mathrm{S}_m[\mathbf{A}] + \mathrm{S}_m[A_0],\\
\mathrm{S}_m[\mathbf{A}] &:= \|\dot{\mathbf{A}}\|^2_{H^{m-1}} + \|\mathbf{A}\|^2_{H^m}, & \mathrm{S}_m[\phi, \mathbf{A}] &:= \mathrm{S}_m[\phi] + \mathrm{S}_m[\mathbf{A}],\\
\mathrm{S}_m[A_0] &:= \|A_0\|^2_{H^m}, & \mathrm{S}_m[\phi, A] &:= \mathrm{S}_m[\phi, \mathbf{A}] + \mathrm{S}_m[A_0],
\end{aligned}$$

where $H^0 = L^2$. Furthermore, for brevity we will often simply write S_m to mean $\mathrm{S}_m[\phi, A]$.

4.2.2 Scaling of Initial Energies

By differentiating the relationship $\tan(\tau/2) = \tanh(H\alpha/2)$ we find

$$\mathrm{d}\tau = \frac{H}{\cosh(H\alpha)}\mathrm{d}\alpha,$$

so raising indices with $\mathfrak{e}^{-1} = \Omega^{-2}\tilde{g}^{-1}$ we find that ∂_τ and ∂_α are related by

$$\partial_\tau = \frac{\cosh(H\alpha)}{H}\partial_\alpha.$$

Furthermore, the conformal factor Ω in the global coordinates (5.5.1) is given by

$$\Omega = H\cos\tau = \frac{H}{\cosh(H\alpha)}.$$

Consider the rescaled approximate energies

$$\mathrm{S}_m[\phi, A](\tau) = \|\dot{\phi}\|^2_{H^{m-1}}(\tau) + \|\phi\|^2_{H^m}(\tau) + \|\dot{\mathbf{A}}\|^2_{H^{m-1}}(\tau) + \|\mathbf{A}\|^2_{H^m}(\tau) + \|A_0\|^2_{H^m}(\tau)$$

and consider smooth initial data for (2.2.22) on the Cauchy surface $\Sigma = \{\tau = 0\} \times \mathbb{S}^3 = \{\alpha = 0\} \times \mathbb{S}^3$. On Σ the conformal factor Ω is a constant and has vanishing τ derivative, $\partial_\tau \Omega|_\Sigma = 0$, so the rescaled scalar field ϕ is related to the physical scalar field $\tilde{\phi}$ by

$$\phi|_\Sigma = (\Omega^{-1}\tilde{\phi})|_\Sigma = \frac{1}{H}\tilde{\phi}|_\Sigma,$$

while their time derivatives are related by

$$\dot{\phi}|_\Sigma = (\Omega^{-1}\partial_\tau\tilde{\phi} - (\partial_\tau\Omega)\Omega^{-2}\tilde{\phi})|_\Sigma = \frac{1}{H^2}\partial_\alpha\tilde{\phi}|_\Sigma.$$

The norms of the rescaled and physical scalar field are therefore equivalent initially,

$$\|\dot{\phi}\|^2_{H^{m-1}}(\tau = 0) + \|\phi\|^2_{H^m}(\tau = 0) \simeq \|\partial_\alpha\tilde{\phi}\|^2_{H^{m-1}}(\alpha = 0) + \|\tilde{\phi}\|^2_{H^m}(\alpha = 0).$$

One similarly checks that

$$\|\dot{\mathbf{A}}\|^2_{H^{m-1}}(\tau = 0) + \|\mathbf{A}\|^2_{H^m}(\tau = 0) \simeq \|\partial_\alpha\tilde{\mathbf{A}}\|^2_{H^{m-1}}(\alpha = 0) + \|\tilde{\mathbf{A}}\|^2_{H^m}(\alpha = 0)$$

and

$$\|A_0\|^2_{H^m}(\tau = 0) \simeq \|\tilde{A}_\alpha\|^2_{H^m}(\alpha = 0),$$

where $\tilde{A}_\alpha = (\partial_\alpha)^a\tilde{A}_a$. Thus

$$\mathrm{S}_m[\phi, \mathbf{A}](\tau = 0) \simeq \mathrm{S}_m[\tilde{\phi}, \tilde{\mathbf{A}}](\alpha = 0), \tag{4.2.3}$$

and also $\mathrm{S}_m[A_0](\tau = 0) \simeq \mathrm{S}_m[\tilde{A}_\alpha](\alpha = 0)$.

Definition 4.2.1. Let $\tilde{\Sigma}$ be a Cauchy surface in dS_4 and Σ the corresponding Cauchy surface in $\widehat{\mathrm{dS}}_4$, and consider data for the Maxwell-scalar field system on Σ. We say the data

$$(\phi_0, \mathbf{A}_0, \phi_1, \mathbf{A}_1, a_0) = (\phi, \mathbf{A}, \dot{\phi}, \dot{\mathbf{A}}, A_0)|_\Sigma$$

is *admissible* if it satisfies the strong Coulomb gauge[2] and a_0 solves the elliptic equation

$$-\not{\Delta} a_0 + |\phi_0|^2 a_0 = -\operatorname{Im}(\bar{\phi}_0\phi_1)$$

on Σ.

Theorem 4.2.2 (Energy Estimates)**.** *Let $m \in \mathbb{N}$. For sufficiently $\mathrm{S}_m[\phi, \mathbf{A}]$-small admissible data on Σ for the Maxwell-scalar field system on $\widehat{\mathrm{dS}}_4 \simeq \mathbb{S}^3 \times [-\pi/2, \pi/2]$ in strong Coulomb gauge one has*

$$\mathrm{S}_m[\phi, A](0) \simeq \mathrm{S}_m[\phi, A](\tau)$$

for all $\tau \in [-\pi/2, \pi/2]$. In particular,

$$\mathrm{S}_m[\phi, A](\mathscr{I}^-) \simeq \mathrm{S}_m[\phi, A](\mathscr{I}^+),$$

where $\mathscr{I}^\pm = \{\tau = \pm\pi/2\}$ is the future (past) null infinity of de Sitter space dS_4.

[2] See section 4.3.1.

Theorem 4.2.3 (Scattering for Small Data)**.** *For $m \geqslant 2$ let S^0_m be the subset of $H^m(\Sigma)^2 \times H^{m-1}(\Sigma)^2 \times H^m(\Sigma)$ of distributions u_0 of admissible data on Σ and let $\mathrm{S}^\pm_m$ be the subset of $H^m(\mathscr{I}^\pm)^2 \times H^{m-1}(\mathscr{I}^\pm)^2 \times H^m(\mathscr{I}^\pm)$ of distributions $u^\pm$ of admissible data on $\mathscr{I}^\pm$, all equipped with the natural norm $\sqrt{\mathrm{S}_m}$. Denote by B_ε the open ball of radius ε in $(H^m)^2 \times (H^{m-1})^2 \times H^m$, and write $\mathrm{S}^0_{m,\varepsilon} = \mathrm{S}^0_m \cap B_\varepsilon$ and $\mathrm{S}^\pm_{m,\varepsilon} = \mathrm{S}^\pm_m \cap B_\varepsilon$. Then for every $m \geqslant 2$ there exist $\varepsilon_0, \varepsilon_1 > 0$, $0 < \delta \ll 1$ and sets $\mathscr{D}^\pm_{m,\varepsilon_1}$ with $\mathrm{S}^\pm_{m,\delta} \subset \mathscr{D}^\pm_{m,\varepsilon_1} \subset \mathrm{S}^\pm_{m,\varepsilon_1}$ such that*

- *there exist bounded invertible nonlinear future and past trace operators*

$$\mathfrak{T}^\pm_m : \mathrm{S}^0_{m,\varepsilon_0} \longrightarrow \mathscr{D}^\pm_{m,\varepsilon_1} \subset \mathrm{S}^\pm_{m,\varepsilon_1}$$

  such that $u^\pm = \mathfrak{T}^\pm_m(u_0)$ is the future (past) Maxwell-scalar field development of u_0 on $\widehat{\mathrm{dS}}_4$ restricted to $\mathscr{I}^\pm$, and

- *there exists a bounded invertible nonlinear scattering operator*

$$\mathscr{S}_m : \mathscr{D}^-_{m,\varepsilon_1} \longrightarrow \mathscr{D}^+_{m,\varepsilon_1},$$
$$\mathscr{S}_m = \mathfrak{T}^+_m \circ (\mathfrak{T}^-_m)^{-1}$$

  such that $u^+ = \mathscr{S}_m(u^-)$ is the Maxwell-scalar field development of u^- on $\widehat{\mathrm{dS}}_4$ restricted to $\mathscr{I}^+$.

Theorem 4.2.4 (Small Data Decay Rates)**.** *Let $\tilde{\phi} = \Omega\phi$ and $\tilde{A}_a = A_a$ be the physical fields related to the conformally rescaled fields ϕ and A_a by (5.3.1), and suppose $\mathrm{S}_2[\tilde{\phi}, \tilde{\mathbf{A}}]$ is small initially. Then the Maxwell-scalar field development $(\tilde{\phi}, \tilde{A})$ of this initial data satisfies the estimates*

$$|\tilde{\phi}| \lesssim \mathrm{e}^{-H|\alpha|}, \qquad |\tilde{A}_\alpha| \lesssim \mathrm{e}^{-H|\alpha|}, \qquad |\tilde{\mathbf{A}}|_{\mathfrak{s}_3} \lesssim 1$$

as $|\alpha| \to \infty$. Furthermore, in the static coordinates (4.1.2)

$$|\tilde{\phi}| \lesssim_r \mathrm{e}^{-H|t|}, \qquad |\tilde{A}_t| \lesssim_r \mathrm{e}^{-H|t|}, \qquad |\tilde{A}_r| \lesssim_r \mathrm{e}^{-H|t|}, \qquad \frac{1}{r}|\tilde{A}|_{\mathfrak{s}_2} \lesssim_r \mathrm{e}^{-H|t|}$$

as $|t| \to \infty$ and r is fixed. Moreover, if $\mathrm{S}_3[\tilde{\phi}, \tilde{\mathbf{A}}]$ is small initially then there exists a constant c such that

$$\tilde{\phi} \sim c\tilde{\Phi}_1 + \mathcal{O}\left(\mathrm{e}^{-2Ht}\right) \quad \text{as} \quad t \to +\infty,$$

where $\tilde{\Phi}_1 = F(r)^{-1/2}\mathrm{e}^{-Ht}$ is the e^{-Ht} eigenmode of the linear uncharged conformally invariant spherically symmetric wave operator on dS_4.

4.3 Field Equations and Gauge Fixing

The field equations (2.2.22) written in terms of the Maxwell potential A_a take the form

$$\begin{aligned} &\Box A_a - \nabla_a(\nabla^b A_b) + \mathrm{R}_{ab}A^b = -\operatorname{Im}(\bar{\phi}\mathrm{D}_a\phi), \\ &\Box\phi + 2iA_a\nabla^a\phi + \left(\frac{1}{6}\mathrm{R} - A_aA^a + i\nabla^a A_a\right)\phi = 0. \end{aligned} \tag{4.3.1}$$

We shall be commuting differential operators into these equations, so it will be convenient to introduce the operators representing their left-hand sides. For any 1-form ω and any scalar field ψ we set

$$\mathscr{F}[\omega]_a := \Box\omega_a - \nabla_a(\nabla_b\omega^b) + \mathrm{R}_{ab}\omega^b \quad \text{and} \quad \mathscr{W}[\psi] := \mathrm{D}^a\mathrm{D}_a\psi + \frac{1}{6}\mathrm{R}\psi.$$

The system (4.3.1) is then equivalent to

$$\mathscr{F}[A]_a = -\operatorname{Im}(\bar{\phi}\mathrm{D}_a\phi) \quad \text{and} \quad \mathscr{W}[\phi] = 0. \tag{4.3.2}$$

4.3.1 Strong Coulomb Gauge

We will work in the Coulomb gauge adapted to the foliation $\mathcal{F}$,

$$\not\nabla \cdot \mathbf{A} = 0, \tag{4.3.3}$$

but will also need to use the residual gauge freedom to fix the gauge fully. More precisely, given a solution (A, ϕ) to the Maxwell-scalar field system (4.3.1), a general gauge transformation sends $\phi \rightsquigarrow \mathrm{e}^{-i\chi}\phi$ and $A_a \rightsquigarrow A_a + \nabla_a\chi$, and (4.3.3) is imposed by solving the elliptic equation

$$\not\Delta\chi = -\not\nabla \cdot \mathbf{A}$$

on $\mathbb{S}^3_\tau$ for every fixed τ. This does not determine χ uniquely: there is still the residual gauge freedom of $\chi \rightsquigarrow \chi + \chi_{\mathrm{res.}}$, where $\chi_{\mathrm{res.}}$ solves

$$\not\Delta\chi_{\mathrm{res.}} = 0$$

on each $\mathbb{S}^3_\tau$. Because $\mathbb{S}^3$ is compact, the kernel of the Laplacian $\not\Delta$ is just the vector space of constant functions, i.e. those $\chi_{\mathrm{res.}}$ which satisfy $\not\nabla\chi_{\mathrm{res.}} = 0$, but the τ dependence in the $\chi_{\mathrm{res.}}$ is still arbitrary. Thus in the Coulomb gauge we have the residual gauge freedom

$$\phi \rightsquigarrow \mathrm{e}^{-i\chi_{\mathrm{res.}}(\tau)}\phi, \qquad A_0 \rightsquigarrow A_0 + \dot{\chi}_{\mathrm{res.}}(\tau), \qquad \mathbf{A} \rightsquigarrow \mathbf{A},$$

which allows one to choose

$$\dot{\chi}_{\mathrm{res.}}(\tau) = -\frac{1}{|\mathbb{S}^3|}\int_{\mathbb{S}^3} A_0(\tau)\,\mathrm{dv}_{\mathfrak{s}_3} =: -\bar{A}_0(\tau)$$

and so impose the additional gauge condition

$$\bar{A}_0(\tau) = 0.$$

This determines $\chi_{\text{res.}}$ up to the addition of a global constant, so there is very little remaining gauge freedom. Indeed, constants are irrelevant for the gauge transformation of A_a and only impart a constant phase change for ϕ, so we have now fixed the gauge as completely as possible. We call this stronger gauge fixing condition

$$\not{\nabla} \cdot \mathbf{A} = 0, \qquad \bar{A}_0 = 0 \tag{4.3.4}$$

strong Coulomb gauge. For us, the most useful feature of the strong Coulomb gauge will be the fact that in this gauge A_0 will obey the Poincaré inequality on each leaf $\mathbb{S}^3_\tau$ of $\mathcal{F}$,

$$\|A_0\|_{L^2}(\tau) \leqslant C\|\not{\nabla} A_0\|_{L^2}(\tau).$$

In strong Coulomb gauge the field equations (4.3.1) are equivalent to the system

$$\begin{aligned} &\Box\phi + 2iA_0\dot{\phi} - 2i\mathbf{A}\cdot\not{\nabla}\phi + (1 - A_0^2 + |\mathbf{A}|^2 + i\dot{A}_0)\phi = 0, \\ &\Box\mathbf{A} + (2 + |\phi|^2)\mathbf{A} = -\operatorname{Im}(\bar{\phi}\not{\nabla}\phi) + \not{\nabla}\dot{A}_0, \\ -&\not{\Delta} A_0 + |\phi|^2 A_0 = -\operatorname{Im}(\bar{\phi}\dot{\phi}), \\ &\not{\nabla}\cdot\mathbf{A} = 0, \\ &\bar{A}_0(\tau) = 0. \end{aligned} \tag{4.3.5}$$

We do not prescribe initial data for A_0 since it is non-dynamical: it is completely determined by ϕ and $\dot{\phi}$ via the elliptic equation on each slice of constant τ. It is convenient to incorporate the constraint $\not{\nabla} \cdot \mathbf{A} = 0$ into the equations by projecting the equation for $\mathbf{A}$ onto divergence free 1-forms on $\mathbb{S}^3$. Let $\mathcal{P}$ be this projection (see appendix A.1.3); then since

$$\not{\nabla} \cdot \Box\mathbf{A} = \Box(\not{\nabla}\cdot\mathbf{A}) - 2\not{\nabla}\cdot\mathbf{A} = 0$$

and

$$\operatorname{curl}\not{\nabla}\dot{A}_0 = 0,$$

applying $\mathcal{P}$ to the equation for $\mathbf{A}$ gives

$$\Box\mathbf{A} + 2\mathbf{A} + \mathcal{P}(|\phi|^2\mathbf{A}) = -\mathcal{P}\operatorname{Im}(\bar{\phi}\not{\nabla}\phi).$$

Thus the system (4.3.5) is equivalent to

$$\begin{aligned} &\Box\phi + 2iA_0\dot{\phi} - 2i\mathbf{A}\cdot\not{\nabla}\phi + (1 - A_0^2 + |\mathbf{A}|^2 + i\dot{A}_0)\phi = 0, \\ &\Box\mathbf{A} + 2\mathbf{A} + \mathcal{P}(|\phi|^2\mathbf{A}) = -\mathcal{P}\operatorname{Im}(\bar{\phi}\not{\nabla}\phi), \\ -&\not{\Delta} A_0 + |\phi|^2 A_0 = -\operatorname{Im}(\bar{\phi}\dot{\phi}), \\ &\bar{A}_0(\tau) = 0, \end{aligned} \tag{4.3.6}$$

provided one considers divergence-free initial data for $\mathbf{A}$ and $\dot{\mathbf{A}}$. Indeed, it is straightforward to show that $v = \not{\nabla} \cdot \mathbf{A}$ satisfies $\Box v = 0$, so $v \equiv 0$ whenever $v = 0$ and $\dot{v} = 0$ initially.

In addition to the restriction $\not{\nabla} \cdot \mathbf{A}_0 = 0 = \not{\nabla} \cdot \mathbf{A}_1$ on the initial data, the extra gauge condition $\bar{A}_0 = 0$ restricts the set of initial data further. Suppose we prescribe initial data $\phi|_\Sigma = \phi_0$ and $\dot{\phi}|_\Sigma = \phi_1$. We must then solve for $A_0|_\Sigma = a_0$ by solving

$$-\not{\Delta} a_0 + |\phi_0|^2 a_0 = -\operatorname{Im}(\bar{\phi}_0\phi_1), \tag{4.3.7}$$

so we must choose the initial data so that this solution has $\bar{a}_0 = 0$. Because A_0 is non-dynamical, it is not possible to write down an evolution equation for $\bar{A}_0$, but the gauge $\bar{A}_0 = 0$ is propagated nonetheless. While A_0 need not be part of the initial data (prescribing $(\phi, \mathbf{A}, \dot{\phi}, \dot{\mathbf{A}})|_\Sigma = (\phi_0, \mathbf{A}_0, \phi_1, \mathbf{A}_1)$ is enough), we may consider A_0 as part of the initial data if it is equal to the a_0 obtained by solving the elliptic equation (4.3.7) initially.

We call data satisfying the above conditions *admissible*.

Remark 4.3.1. The condition $\bar{a}_0 = 0$ is a condition on the initial data for ϕ and can be seen explicitly as follows. Consider the operator

$$L := -\not{\Delta} + |\phi_0|^2$$

on $\mathbb{S}^3$ and assume that ϕ_0 is not identically zero (if it is, then the equation becomes $\not{\Delta} a_0 = 0$ and we can trivially choose the zero solution). We can classify the kernel of L if the data (ϕ_0, ϕ_1) is sufficiently regular, say $(\phi_0, \phi_1) \in H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3)$. Multiplying the equation $Lu = 0$ by u and integrating we get

$$\int_{\mathbb{S}^3} |\not{\nabla} u|^2 \, \mathrm{dv}_{\mathfrak{s}_3} + \int_{\mathbb{S}^3} |\phi_0|^2 u^2 \, \mathrm{dv}_{\mathfrak{s}_3} = 0,$$

so that $\not{\nabla} u = 0$. If $u \in H^2(\mathbb{S}^3)$, then by Sobolev embedding $u \in \mathcal{C}^0(\mathbb{S}^3)$; then continuity of u and $\|\phi_0 u\|_{L^2} = 0$ imply that $u \equiv 0$. Thus as an operator from $H^2(\mathbb{S}^3)$ to $L^2(\mathbb{S}^3)$[3], L has trivial kernel. It follows from standard elliptic theory that the equation $Lu = \psi$ has a unique solution $u \in H^2(\mathbb{S}^3)$ for $\psi \in L^2(\mathbb{S}^3)$, which we write as $u = L^{-1}\psi$. Since $(\phi_0, \phi_1) \in H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3)$ ensures[4] that $\bar{\phi}_0 \phi_1 \in L^2(\mathbb{S}^3)$, we have

$$a_0 = -L^{-1} \operatorname{Im}(\bar{\phi}_0 \phi_1) = (\not{\Delta} - |\phi_0|^2)^{-1} \operatorname{Im}(\bar{\phi}_0 \phi_1).$$

The requirement $\bar{a}_0 = 0$ may thus be written as the condition

$$\int_{\mathbb{S}^3} (\not{\Delta} - |\phi_0|^2)^{-1} \operatorname{Im}(\bar{\phi}_0 \phi_1) \, \mathrm{dv}_{\mathfrak{s}_3} = 0 \tag{4.3.8}$$

on the initial data (ϕ_0, ϕ_1).

Remark 4.3.2. The system (4.3.6) in principle exhibits the null structure of Klainerman and Machedon [66], which has been exploited to prove the finite energy well-posedness of the Maxwell-scalar field system on flat space [65]. Although the requisite null form estimates have also been established on certain spacetimes with compact spacelike slices [51, 111], a corresponding finite energy well-posedness result has not been proven, even on the Einstein cylinder.

[3] For $\phi_0, u \in H^2(\mathbb{S}^3)$ it is straightforward to check that $|\phi_0|^2 u \in L^2(\mathbb{S}^3)$, so L does indeed map into $L^2(\mathbb{S}^3)$.

[4] In fact, $H^2(\mathbb{S}^3) \cdot H^1(\mathbb{S}^3) \subset H^1(\mathbb{S}^3)$, by Sobolev embedding.

4.4 Well-Posedness

To solve the Cauchy problem for (4.3.6) we make use of a classical theorem of Choquet-Bruhat. Let I be an interval in $\mathbb{R}$ and let

$$E_m(I \times \mathbb{S}^n) := \bigcap_{k=0}^{m} \mathcal{C}^k(I; H^{m-k}(\mathbb{S}^n))$$

be the standard finite m-energy space for hyperbolic systems. The following theorem elucidates why first order (that is, H^1) energy estimates are insufficient to construct a scattering theory for the Maxwell-scalar field system and why H^2 estimates are good enough $(2 > 3/2)$.

Theorem 4.4.1 (Y. Choquet-Bruhat, [22])**.** *Consider the system (2.2.22) on $\mathbb{R}_\tau \times \mathbb{S}^n$. Let $T^a = \partial_\tau$ be the timelike unit normal to $\mathbb{S}^n_\tau := \mathbb{S}^n \times \{\tau\}$, set $\mathbf{E}_b := T^a F_{ab}$, and suppose that we are given data ϕ_0, $a_b \in H^m(\mathbb{S}^n_0)$ and ϕ_1, $\mathbf{e}_b \in H^{m-1}(\mathbb{S}^n_0)$ satisfying the constraint*

$$\bar{\nabla} \cdot \mathbf{e} = -a_0 |\phi_0|^2 - \operatorname{Im}(\bar{\phi}_0 \phi_1), \qquad (\dagger)$$

where $\bar{\nabla}$ is the Levi–Civita connection on $\mathbb{S}^n_0$ and $a_0 = T^b a_b$. Then there exists an interval $I_\sigma = (-\sigma, \sigma) \subset \mathbb{R}$ and a solution (ϕ, A_a) in $E_m(I_\sigma \times \mathbb{S}^n)$ satisfying the system (2.2.22) and the Lorenz gauge condition $\nabla_a A^a = 0$ such that

$$A_b|_{\mathbb{S}^n_0} = a_b, \qquad \mathbf{E}_b|_{\mathbb{S}^n_0} = \mathbf{e}_b, \qquad \phi|_{\mathbb{S}^n_0} = \phi_0, \qquad \dot{\phi}|_{\mathbb{S}^n_0} = \phi_1$$

if $m > n/2$. The supremum of such numbers $\sigma > 0$ depends continuously on

$$M_1 = \|a\|_{H^m} + \|\phi_0\|_{H^m} + \|\phi_1\|_{H^{m-1}} + \|\mathbf{e}\|_{H^{m-1}}$$

and tends to infinity as M_1 tends to zero. The solution (ϕ, A_a) is unique in $E_m(I_\sigma \times \mathbb{S}^n)$ up to gauge transformations preserving the Lorenz gauge.

Corollary 4.4.2. *Consider the system (4.3.6) on $\mathfrak{E} = \mathbb{R} \times \mathbb{S}^3$ and suppose that for $m \geqslant 2$ we are given data ϕ_0, $\mathbf{A}_0 \in H^m(\mathbb{S}^3_0)$ and ϕ_1, $\mathbf{A}_1 \in H^{m-1}(\mathbb{S}^3_0)$ satisfying the strong Coulomb gauge initially. Then there exists an interval $I_\sigma = (-\sigma, \sigma) \subset \mathbb{R}$ and a solution $(\phi, A_0, \mathbf{A})$ in $E_m(I_\sigma \times \mathbb{S}^3)$ satisfying the system (4.3.6) and the strong Coulomb gauge conditions $\bar{A}_0 = 0$, $\bar{\nabla} \cdot \mathbf{A} = 0$ such that*

$$\mathbf{A}|_{\mathbb{S}^3_0} = \mathbf{A}_0, \qquad \dot{\mathbf{A}}|_{\mathbb{S}^3_0} = \mathbf{A}_1, \qquad \phi|_{\mathbb{S}^3_0} = \phi_0, \qquad \dot{\phi}|_{\mathbb{S}^3_0} = \phi_1.$$

The supremum of such numbers $\sigma > 0$ depends continuously on

$$M_2 = \|a_0\|_{H^m} + \|\mathbf{A}_0\|_{H^m} + \|\mathbf{A}_1\|_{H^{m-1}} + \|\phi_0\|_{H^{m-1}} + \|\phi_1\|_{H^{m-1}} \simeq \mathrm{S}_m[\phi, A](0)^{1/2}$$

and tends to infinity as M_2 tends to zero, where a_0 is determined by ϕ_0 and ϕ_1 via the elliptic equation (4.3.7) on $\mathbb{S}^3_0$. The solution $(\phi, A_0, \mathbf{A})$ is unique in $E_m(I_\sigma \times \mathbb{S}^3)$ up to gauge transformations preserving the strong Coulomb gauge[5]*.*

[5] Recall that the gauge transformations preserving the strong Coulomb gauge are just the trivial ones $\chi = \mathrm{e}^{i\theta}$ for global constants $\theta \in \mathbb{R}$.

Proof. Given admissible $\phi_0 \in H^m(\mathbb{S}^3_0)$ and $\phi_1 \in H^{m-1}(\mathbb{S}^3_0)$, the equation

$$-\not{\Delta} a_0 + |\phi_0|^2 a_0 = -\operatorname{Im}(\bar{\phi}_0 \phi_1)$$

on $\mathbb{S}^3_0$ has a unique solution a_0 in H^m which by (4.3.8) satisfies $\bar{a}_0 = 0$. We define $\mathbf{E} := \mathbf{A}_1 - \not{\nabla} a_0$, which by construction satisfies (†). We may thus apply theorem 4.4.1. Note that we do not prescribe $\dot{A}_0$, but instead construct it so that the Lorenz gauge condition is satisfied initially. The Lorenz gauge is then propagated by the equations (2.2.22) in Lorenz gauge (but note that, of course, the strong Coulomb gauge is not). We thus have a solution (ϕ, A_a) in $E_m(I_\sigma \times \mathbb{S}^3)$ of (2.2.22) satisfying $\nabla_a A^a = 0$ throughout $I_\sigma \times \mathbb{S}^3$. Now perform a gauge transformation as in section 4.3.1 to convert this solution to a solution $(\phi, A_0, \mathbf{A})$ in $E_m(I_\sigma \times \mathbb{S}^3)$ of (4.3.6) satisfying the strong Coulomb gauge. It is easy to see that this gauge transformation preserves E_m regularity, while uniqueness up to gauge transformations is also clear. As for the continuous dependence of σ on the data, we note that

$$\begin{aligned} M_1 &= \|a\|_{H^2} + \|\phi_0\|_{H^2} + \|\phi_1\|_{H^1} + \|\mathbf{E}\|_{H^1} \\ &\lesssim \|a_0\|_{H^2} + \|\mathbf{A}_0\|_{H^2} + \|\phi_0\|_{H^2} + \|\phi_1\|_{H^1} + \|\not{\nabla} a_0\|_{H^1} + \|\mathbf{A}_1\|_{H^1} \\ &\lesssim \|a_0\|_{H^2} + \|\mathbf{A}_0\|_{H^2} + \|\mathbf{A}_1\|_{H^1} + \|\phi_0\|_{H^2} + \|\phi_1\|_{H^1} = M_2, \end{aligned}$$

and similarly $M_2 \lesssim M_1$. Thus $M_1 \simeq M_2$ and we are done. □

4.5 Energies

4.5.1 The Maxwell Sector

We treat the Maxwell and the scalar field sectors of the energy-momentum tensor $\mathbf{T}_{ab}$ separately. The energy-momentum tensor for the Maxwell sector in terms of the Maxwell field F_{ab} is

$$\mathbf{T}_{ab}[F] = -F_a{}^c F_{bc} + \frac{1}{4}\mathfrak{e}_{ab} F_{cd} F^{cd},$$

and becomes

$$\begin{aligned} \mathbf{T}_{ab}[A] = &- \nabla_a A^c \nabla_b A_c + \nabla^c A_a \nabla_b A_c + \nabla_a A^c \nabla_c A_b - \nabla^c A_a \nabla_c A_b \\ &+ \frac{1}{2}\mathfrak{e}_{ab}\left(\nabla_c A_d \nabla^c A^d - \nabla_c A_d \nabla^d A^c\right) \end{aligned}$$

in terms of the potential A_a. The Maxwell sector energy density with respect to the foliation $\mathcal{F}$ is given by the component

$$\begin{aligned} \mathbf{T}_{00}[A] &= \mathbf{T}_{ab} T^a T^b \\ &= -\dot{A}^c \dot{A}_c + 2\dot{A}_c \nabla^c A_0 - \nabla^c A_0 \nabla_c A_0 + \frac{1}{2}\left(\nabla_c A_d \nabla^c A^d - \nabla_c A_d \nabla^d A^c\right), \end{aligned}$$

where in the above we have denoted by $A_0 := T^a A_a$ and $\dot{A}_a := T^b \nabla_b A_a$. Note that the metric $\mathfrak{e}$ splits as the direct sum $\mathfrak{e} = 1 \oplus (-\mathfrak{s}_3)$, so in particular the full connection ∇ also splits as $\nabla = \nabla^{\mathbb{R}} \oplus \nabla^{\mathfrak{s}_3} = \partial_\tau \oplus \not{\nabla}$ (see appendix C.3.3). Furthermore, there is no curvature in the τ direction, so ∂_τ commutes with the 3-sphere derivatives, $[\partial_\tau, \not\nabla] = 0$. Expanding the above, we therefore have

$$\mathbf{T}_{00}[A] = \frac{1}{2}|\dot{\mathbf{A}}|^2 + \frac{1}{2}|\not\nabla A_0|^2 + \frac{1}{2}|\not\nabla \mathbf{A}|^2 - \dot{\mathbf{A}} \cdot \not\nabla A_0 - \frac{1}{2}(\not\nabla_\mu \mathbf{A}_\nu)(\not\nabla^\nu \mathbf{A}^\mu). \tag{4.5.1}$$

In the Coulomb gauge $\not\nabla \cdot \mathbf{A} \equiv 0 \equiv \not\nabla \cdot \dot{\mathbf{A}}$ the last two terms in (4.5.1) become non-negative-definite upon integration by parts:

$$\int_{\mathbb{S}^3} -\dot{\mathbf{A}} \cdot \not\nabla A_0 \, \mathrm{dv}_{\mathfrak{s}_3} = \int_{\mathbb{S}^3} A_0 \not\nabla \cdot \dot{\mathbf{A}} \, \mathrm{dv}_{\mathfrak{s}_3} = 0,$$

and

$$\int_{\mathbb{S}^3} -\frac{1}{2}(\not\nabla_\mu \mathbf{A}_\nu)(\not\nabla^\nu \mathbf{A}^\mu) \, \mathrm{dv}_{\mathfrak{s}_3} = \int_{\mathbb{S}^3} \left(\frac{1}{2}\mathbf{A}^\mu \not\nabla_\mu \not\nabla_\nu \mathbf{A}^\nu - \frac{1}{2}\mathrm{R}_{\mu\nu}\mathbf{A}^\mu \mathbf{A}^\nu \right) \mathrm{dv}_{\mathfrak{s}_3} = \int_{\mathbb{S}^3} |\mathbf{A}|^2 \, \mathrm{dv}_{\mathfrak{s}_3} .$$

Thus the Maxwell sector energy on 3-spheres of constant τ is

$$\mathcal{E}_\tau[A] \simeq \|\mathbf{A}\|^2_{H^1}(\tau) + \|\dot{\mathbf{A}}\|^2_{L^2}(\tau) + \|\not\nabla A_0\|^2_{L^2}(\tau) = \mathrm{S}_1[\mathbf{A}](\tau) + \|\not\nabla A_0\|^2_{L^2}(\tau).$$

Using the additional gauge condition $\bar{A}_0(\tau) = 0$, one has that $\|A_0\|^2_{L^2(\mathbb{S}^3)} \lesssim \|\not\nabla A_0\|^2_{L^2(\mathbb{S}^3)}$, so

$$\mathcal{E}_\tau[A] \simeq \mathrm{S}_1[A](\tau) \tag{4.5.2}$$

for all $\tau \in \mathbb{R}$.

Higher Order Energies

More generally, for a 1-form α set

$$\begin{aligned} \mathbf{T}_{ab}[\alpha] \coloneqq & -\nabla_a \alpha^c \nabla_b \alpha_c + \nabla^c \alpha_a \nabla_b \alpha_c + \nabla_a \alpha^c \nabla_c \alpha_b - \nabla^c \alpha_a \nabla_c \alpha_b \\ & + \frac{1}{2}\mathfrak{e}_{ab} \left(\nabla_c \alpha_d \nabla^c \alpha^d - \nabla_c \alpha_d \nabla^d \alpha^c \right). \end{aligned}$$

When $\alpha_a = A_a$, this is, of course, just the Maxwell energy-momentum tensor written out in terms of the potential. As in (4.5.1), we have

$$\mathbf{T}_{00}[\alpha] = \frac{1}{2}|\dot{\boldsymbol{\alpha}}|^2 + \frac{1}{2}|\not\nabla \alpha_0|^2 + \frac{1}{2}|\not\nabla \boldsymbol{\alpha}|^2 - \dot{\boldsymbol{\alpha}}_\mu \not\nabla^\mu \alpha_0 - \frac{1}{2}(\not\nabla_\mu \boldsymbol{\alpha}_\nu)(\not\nabla^\nu \boldsymbol{\alpha}^\mu).$$

Integrating by parts as before we obtain

$$\begin{aligned} \mathcal{E}_\tau[\alpha] \coloneqq & \int_{\mathbb{S}^3} \mathbf{T}_{00}[\alpha] \, \mathrm{dv}_{\mathfrak{s}_3} \\ = & \frac{1}{2}\int_{\mathbb{S}^3} |\dot{\boldsymbol{\alpha}}|^2 \, \mathrm{dv}_{\mathfrak{s}_3} + \frac{1}{2}\int_{\mathbb{S}^3} |\not\nabla \alpha_0|^2 \, \mathrm{dv}_{\mathfrak{s}_3} + \frac{1}{2}\int_{\mathbb{S}^3} |\not\nabla \boldsymbol{\alpha}|^2 \, \mathrm{dv}_{\mathfrak{s}_3} \\ & + \int_{\mathbb{S}^3} \alpha_0 \not\nabla_\mu \dot{\boldsymbol{\alpha}}^\mu \, \mathrm{dv}_{\mathfrak{s}_3} - \frac{1}{2}\int_{\mathbb{S}^3} |\not\nabla \cdot \boldsymbol{\alpha}|^2 \, \mathrm{dv}_{\mathfrak{s}_3} + \int_{\mathbb{S}^3} |\boldsymbol{\alpha}|^2 \, \mathrm{dv}_{\mathfrak{s}_3} . \end{aligned}$$

For our second order estimates we will want to set $\alpha_a = X_i^\mu \not\nabla_\mu A_a \coloneqq \not\nabla_i A_a$ and sum over i for a basis of vector fields $\{X_i\}_i$ on $\mathbb{S}^3$ (e.g. a basis of left-invariant vector fields on $\mathbb{S}^3 \simeq \mathrm{SU}(2)$). The first term in the above is then clearly

$$\sum_i |\dot{\boldsymbol{\alpha}}|^2 = \sum_i \not\nabla_i \dot{\mathbf{A}}_\mu \not\nabla_i \dot{\mathbf{A}}^\mu = |\not\nabla \dot{\mathbf{A}}|^2,$$

the second term becomes

$$\sum_i |\not\nabla \alpha_0|^2 = \sum_i \not\nabla_\mu \not\nabla_i A_0 \not\nabla^\mu \not\nabla_i A_0 = |\not\nabla^2 A_0|^2 + \text{l.o.t.s},$$

the third term becomes

$$\sum_i |\not\nabla \boldsymbol{\alpha}|^2 = \sum_i \not\nabla_\mu \not\nabla_i \mathbf{A}_\nu \not\nabla^\mu \not\nabla_i \mathbf{A}^\nu = |\not\nabla^2 \mathbf{A}|^2 + \text{l.o.t.s},$$

the fourth term, after commuting derivatives to impose the Coulomb gauge $\not\nabla \cdot \mathbf{A} = 0$, is

$$\sum_i \alpha_0 \not\nabla_\mu \dot{\boldsymbol{\alpha}}^\mu = \sum_i \not\nabla_i A_0 \not\nabla_\mu \not\nabla_i \dot{\mathbf{A}}^\mu = \text{l.o.t.s},$$

and the fifth term similarly becomes

$$\sum_i |\not\nabla \cdot \boldsymbol{\alpha}|^2 = \sum_i \not\nabla_\mu \not\nabla_i \mathbf{A}^\mu \not\nabla_\nu \not\nabla_i \not{\mathbf{A}}^\nu = \text{l.o.t.s},$$

where in the above we have written $\not\nabla_j := X_j^\mu \not\nabla_\mu$, and the lower order terms are at most quadratic and of order zero and one in derivatives of A_a. The sixth and final term is

$$\sum_i |\boldsymbol{\alpha}|^2 = \sum_i |\not\nabla_i \mathbf{A}|^2 = \text{l.o.t.s.}$$

The lower order terms can be controlled by $\mathcal{E}_\tau[A] \simeq \mathrm{S}_1[A](\tau)$, so we can find a constant $C > 0$ large enough such that

$$\begin{aligned} C\mathcal{E}_\tau[A] + \sum_i \mathcal{E}_\tau[\not\nabla_i A] &\simeq \|\mathbf{A}\|^2_{H^2}(\tau) + \|\dot{\mathbf{A}}\|^2_{H^1}(\tau) + \|\not\nabla A_0\|^2_{H^1}(\tau) \\ &\simeq \mathrm{S}_2[\mathbf{A}](\tau) + \|\not\nabla A_0\|^2_{H^1}(\tau). \end{aligned}$$

As before, the strong Coulomb gauge implies $\|A_0\|_{L^2} \lesssim \|\not\nabla A_0\|_{L^2}$, and so

$$\mathcal{E}_\tau[A] + \sum_i \mathcal{E}_\tau[\not\nabla_i A] \simeq \mathrm{S}_2[A](\tau). \tag{4.5.3}$$

More generally, it is straightforward to see that in the strong Coulomb gauge one has

$$\sum_{k=0}^{m-1} \mathcal{E}_\tau[\not\nabla^k A] \simeq \mathrm{S}_m[A](\tau),$$

where $\mathcal{E}_\tau[\not\nabla^k A]$ denotes $\sum_{i_1,\ldots,i_k} \mathcal{E}_\tau[\not\nabla_{i_1} \ldots \not\nabla_{i_k} A]$.

4.5.2 The Scalar Field Sector

The energy-momentum tensor for the scalar field sector is

$$\mathbf{T}_{ab}[\phi] = \overline{\mathrm{D}_{(a}\phi}\mathrm{D}_{b)}\phi - \frac{1}{2}\mathfrak{e}_{ab}\overline{\mathrm{D}_c\phi}\mathrm{D}^c\phi + \frac{1}{2}\mathfrak{e}_{ab}|\phi|^2,$$

and we calculate

$$\begin{aligned} \mathbf{T}_{00}[\phi] &= |\mathrm{D}_0\phi|^2 - \frac{1}{2}\overline{\mathrm{D}_c\phi}\mathrm{D}^c\phi + \frac{1}{2}|\phi|^2 \\ &= \frac{1}{2}|\mathrm{D}_0\phi|^2 + \frac{1}{2}\overline{\not{\mathrm{D}}_\mu\phi}\not{\mathrm{D}}^\mu\phi + \frac{1}{2}|\phi|^2 \end{aligned}$$

and therefore

$$\mathcal{E}_\tau[\phi] = \frac{1}{2}\|\mathrm{D}_0\phi\|_{L^2}^2(\tau) + \frac{1}{2}\|\not{\mathrm{D}}\phi\|_{L^2}^2(\tau) + \frac{1}{2}\|\phi\|_{L^2}^2(\tau),$$

where $\mathrm{D}_0\phi = \dot{\phi} + iA_0\phi$ and $\not{\mathrm{D}}_\mu = \not{\nabla}_\mu + i\mathbf{A}_\mu$. More generally, we set

$$\mathbf{T}_{ab}[\psi] = \overline{\mathrm{D}_{(a}\psi}\mathrm{D}_{b)}\psi - \frac{1}{2}\mathfrak{e}_{ab}\overline{\mathrm{D}_c\psi}\mathrm{D}^c\psi + \frac{1}{2}\mathfrak{e}_{ab}|\psi|^2$$

and

$$\mathcal{E}_\tau[\psi] = \frac{1}{2}\|\mathrm{D}_0\psi\|_{L^2}^2(\tau) + \frac{1}{2}\|\not{\mathrm{D}}\psi\|_{L^2}^2(\tau) + \frac{1}{2}\|\psi\|_{L^2}^2(\tau)$$

for any complex scalar field ψ on $\mathfrak{E}$. As with the Maxwell sector, we will want to choose $\psi = \not{\nabla}_i\phi$ for our second order estimates.

Conversion Between Geometric and Approximate Energies

Proposition 4.5.1. *For any fixed $\tau \in \mathbb{R}$ and any sufficiently smooth complex scalar field ψ on $\mathfrak{E}$ there exists $\varepsilon > 0$ such that if $\mathrm{S}_1[\mathbf{A}](\tau) \leqslant \varepsilon$, then*

$$\|\not{\nabla}\psi\|_{L^2}^2(\tau) \lesssim \mathcal{E}_\tau[\psi].$$

Proof. We suppress the τ variable. Clearly

$$\|\not{\nabla}\psi\|_{L^2}^2 \lesssim \|\not{\mathrm{D}}\psi\|_{L^2}^2 + \|\mathbf{A}\psi\|_{L^2}^2 \lesssim \mathcal{E}[\psi] + \|\mathbf{A}\|_{L^6}^2\|\psi\|_{L^3}^2.$$

Now since $\mathbb{S}^3$ is compact, $\|\psi\|_{L^3} \lesssim \|\psi\|_{L^6}$, and by Sobolev embedding

$$\|\mathbf{A}\|_{L^6}^2 \lesssim \|\not{\nabla}\mathbf{A}\|_{L^2}^2 + \|\mathbf{A}\|_{L^2}^2 \lesssim \mathrm{S}_1[\mathbf{A}]$$

and

$$\|\psi\|_{L^6}^2 \lesssim \|\not{\nabla}\psi\|_{L^2}^2 + \|\psi\|_{L^2}^2 \lesssim \|\not{\nabla}\psi\|_{L^2}^2 + \mathcal{E}[\psi].$$

This gives

$$\|\not{\nabla}\psi\|_{L^2}^2 \leqslant C(1+\mathrm{S}_1[\mathbf{A}])\mathcal{E}[\psi] + C\mathrm{S}_1[\mathbf{A}]\|\not{\nabla}\psi\|_{L^2}^2 \leqslant C\varepsilon\|\not{\nabla}\psi\|_{L^2}^2 + C(1+\varepsilon)\mathcal{E}[\psi],$$

so

$$\|\not{\nabla}\psi\|_{L^2}^2 \leqslant C\left(\frac{1+\varepsilon}{1-\varepsilon C}\right)\mathcal{E}[\psi] \lesssim \mathcal{E}[\psi]$$

for $\varepsilon > 0$ small enough. □

Proposition 4.5.2. *For any fixed $\tau \in \mathbb{R}$ and any sufficiently smooth complex scalar field ψ on $\mathfrak{E}$ there exists $\varepsilon > 0$ such that if $\mathrm{S}_1[\mathbf{A}](\tau) \leqslant \varepsilon$, then*

$$\|\dot{\psi}\|_{L^2}^2(\tau) \lesssim (1+\mathrm{S}_1[A_0](\tau))\,\mathcal{E}_\tau[\psi].$$

Proof. Working similarly to the previous proposition,

$$\|\dot{\psi}\|_{L^2}^2 \lesssim \|\mathrm{D}_0\psi\|_{L^2}^2 + \|A_0\psi\|_{L^2}^2 \lesssim \mathcal{E}[\psi] + \|A_0\|_{L^6}^2\|\psi\|_{L^3}^2 \lesssim \mathcal{E}[\psi] + \|A_0\|_{H^1}^2\|\psi\|_{L^6}^2.$$

Also $\|\psi\|_{L^6}^2 \lesssim \|\not{\nabla}\psi\|_{L^2}^2 + \|\psi\|_{L^2}^2$, so

$$\|\dot{\psi}\|_{L^2}^2 \lesssim \mathcal{E}[\psi] + \mathrm{S}_1[A_0](\|\not{\nabla}\psi\|_{L^2}^2 + \|\psi\|_{L^2}^2) \lesssim (1+\mathrm{S}_1[A_0])\,\mathcal{E}[\psi] + \mathrm{S}_1[A_0]\|\not{\nabla}\psi\|_{L^2}^2.$$

Proposition 4.5.1 now gives the result for small $\mathrm{S}_1[\mathbf{A}]$. □

Proposition 4.5.3. *For any fixed $\tau \in \mathbb{R}$ and any sufficiently smooth complex scalar field ψ on $\mathfrak{E}$ one has*

$$\|\not{D}\psi\|^2_{L^2}(\tau) \lesssim \mathrm{S}_1[\psi](\tau)\,(1 + \mathrm{S}_1[\mathbf{A}](\tau)).$$

Proof. This is a simple consequence of the compactness of $\mathbb{S}^3$ and Sobolev embedding as above,

$$\|\not{D}\psi\|^2_{L^2} \lesssim \|\not{\nabla}\psi\|^2_{L^2} + \|\mathbf{A}\psi\|^2_{L^2} \lesssim \|\not{\nabla}\psi\|^2_{L^2} + \|\mathbf{A}\|^2_{L^6}\|\psi\|^2_{L^6} \lesssim \mathrm{S}_1[\psi] + \mathrm{S}_1[\mathbf{A}]\mathrm{S}_1[\psi].$$

□

Proposition 4.5.4. *For any fixed $\tau \in \mathbb{R}$ and any sufficiently smooth complex scalar field ψ on $\mathfrak{E}$ one has*

$$\|\mathrm{D}_0\psi\|^2_{L^2}(\tau) \lesssim (1 + \mathrm{S}_1[A_0](\tau))\,\mathrm{S}_1[\psi](\tau).$$

Proof. This follows from the same splitting and embedding as in the previous propositions,

$$\|\mathrm{D}_0\psi\|^2_{L^2} \lesssim \|\dot{\psi}\|^2_{L^2} + \|A_0\psi\|^2_{L^2} \lesssim \mathrm{S}_1[\psi](1 + \|A_0\|^2_{H^1} \lesssim (1 + \mathrm{S}_1[A_0])\,\mathrm{S}_1[\psi].$$

□

Theorem 4.5.5. *For any fixed $\tau \in \mathbb{R}$ and any sufficiently smooth complex scalar field ψ on $\mathfrak{E}$ there exists $\varepsilon > 0$ such that if $\mathrm{S}_1[A] \leqslant \varepsilon$, then*

$$\mathrm{S}_1[\psi](\tau) \simeq \mathcal{E}[\psi](\tau).$$

Proof. Suppose $\mathrm{S}_1[A]$ is small. Then in particular both $\mathrm{S}_1[\mathbf{A}]$ and $\mathrm{S}_1[A_0]$ are small, so by proposition 4.5.1 $\|\not{\nabla}\psi\|^2_{L^2} \lesssim \mathcal{E}[\psi]$. By proposition 4.5.2, $\|\dot{\psi}\|^2_{L^2} \lesssim \mathcal{E}[\psi]$, so

$$\mathrm{S}_1[\psi] \lesssim \mathcal{E}[\psi].$$

Conversely, by propositions 4.5.3 and 4.5.4, $\|\not{D}\psi\|^2_{L^2} \lesssim \mathrm{S}_1[\psi]$ and $\|\mathrm{D}_0\psi\|^2_{L^2} \lesssim \mathrm{S}_1[\psi]$, so

$$\mathcal{E}[\psi] \lesssim \mathrm{S}_1[\psi].$$

□

In particular, $\mathcal{E}[\phi] \simeq \mathrm{S}_1[\phi]$ and $\mathcal{E}[\not{\nabla}\phi] \simeq \mathrm{S}_1[\not{\nabla}\phi]$. Since $\mathrm{S}_1[\phi]+\mathrm{S}_1[\not{\nabla}\phi] \simeq \mathrm{S}_2[\phi]$, one then has

$$\mathcal{E}_\tau[\phi] + \mathcal{E}_\tau[\not{\nabla}\phi] \simeq \mathrm{S}_2[\phi](\tau) \tag{4.5.4}$$

if $\mathrm{S}_1[A](\tau)$ is sufficiently small. Similarly,

$$\sum_{k=0}^{m-1} \mathcal{E}_\tau[\not{\nabla}^k\phi] \simeq \mathrm{S}_m[\phi](\tau)$$

if $\mathrm{S}_1[A](\tau)$ is sufficiently small.

4.5.3 Elliptic Estimates

As we have already seen, one useful feature of the Coulomb gauge is that the field equation for A_0 becomes elliptic,

$$-\not\Delta A_0 + |\phi|^2 A_0 = -\operatorname{Im}(\bar{\phi}\dot{\phi}). \tag{4.5.5}$$

But even though the component A_0 is non-dynamical, it still carries energy. This energy is controlled by $\dot{\phi}$ as follows.

Proposition 4.5.6. *The non-dynamical component A_0 satisfies the a priori estimates*

$$\|\not\nabla A_0\|_{L^2}^2(\tau) + \|\phi A_0\|_{L^2}^2(\tau) + \|A_0\|_{L^2}^2(\tau) \lesssim \|\dot{\phi}\|_{L^2}^2(\tau)$$

for every fixed $\tau \in \mathbb{R}$.

Proof. Multiplying equation (4.5.5) by A_0 and integrating, we have

$$\|\not\nabla A_0\|_{L^2}^2 + \|\phi A_0\|_{L^2}^2 = -\int_{\mathbb{S}^3} \operatorname{Im}(\bar{\phi}\dot{\phi})A_0 \,\mathrm{dv}_{\mathfrak{s}_3} \leqslant \|\phi A_0\|_{L^2}\|\dot{\phi}\|_{L^2} \leqslant \frac{1}{2}\|\phi A_0\|_{L^2}^2 + \frac{1}{2}\|\dot{\phi}\|_{L^2}^2,$$

which gives the first two estimates. The third estimate follows from the Poincaré inequality for A_0. □

4.6 A Priori Energy Estimates

4.6.1 Conservation of Energy

For general sufficiently smooth α_a, ψ one computes

$$\begin{aligned} \nabla^a \mathbf{T}_{ab}[\alpha] &= \mathscr{F}[\alpha]^a \left(\nabla_a \alpha_b - \nabla_b \alpha_a\right), \\ \nabla^a \mathbf{T}_{ab}[\psi] &= \frac{1}{2}\overline{\mathscr{W}[\psi]}\mathrm{D}_b\psi + \frac{1}{2}\mathscr{W}[\psi]\overline{\mathrm{D}_b\psi} + F_{ab}\operatorname{Im}(\bar{\psi}\mathrm{D}^a\psi). \end{aligned} \tag{4.6.1}$$

When $\alpha_a = A_a$ and $\psi = \phi$, the field equations $\mathscr{F}[A]_a = -\operatorname{Im}(\bar{\phi}\mathrm{D}_a\phi)$ and $\mathscr{W}[\phi] = 0$ imply that

$$\begin{aligned} \nabla^a \mathbf{T}_{ab}[\phi, A] &= \nabla^a(\mathbf{T}_{ab}[A] + \mathbf{T}_{ab}[\phi]) \\ &= F_{ab}\left(\operatorname{Im}(\bar{\phi}\mathrm{D}^a\phi) - \operatorname{Im}(\bar{\phi}\mathrm{D}^a\phi)\right) = 0. \end{aligned}$$

We use the identities (4.6.1) to derive a priori energy estimates for ϕ and A_a at all orders.

4.6.2 H^1 Estimates

Consider admissible initial data for the system (4.3.6). We can make no a priori assumptions about the smallness of the non-dynamical component A_0, but we will of course be able to extract all the required information about A_0 using the elliptic equation (4.5.5).

Theorem 4.6.1. *There exists an $\varepsilon > 0$ such that if $\mathrm{S}_1[\phi, \mathbf{A}](0) \leqslant \varepsilon$, then*

$$\mathrm{S}_1[\phi, A](\tau) \simeq \mathrm{S}_1[\phi, A](0)$$

for all $\tau \in \mathbb{R}$.

Proof. Since $\nabla^a \mathbf{T}_{ab}[\phi, A] = 0$ and $T^a = \partial_\tau$ is Killing on $\mathfrak{E}$, integrating $\mathsf{e}_1 := \nabla^a(T^b \mathbf{T}_{ab}[\phi, A]) = 0$ over the region $\mathbb{S}^3 \times [0, \tau]$ for any $\tau > 0$ immediately gives

$$0 = \int_{\mathbb{S}^3 \times [0,\tau]} \mathsf{e}_1 \,\mathrm{dv} = \int_{\mathbb{S}^3_\tau} \mathbf{T}_{00}[\phi, A] \,\mathrm{dv}_{\mathfrak{s}_3} - \int_{\mathbb{S}^3_0} \mathbf{T}_{00}[\phi, A] \,\mathrm{dv}_{\mathfrak{s}_3},$$

i.e.

$$\mathcal{E}_\tau[\phi] + \mathcal{E}_\tau[A] = \mathcal{E}_\tau[\phi, A] = \mathcal{E}_0[\phi, A] = \mathcal{E}_0[\phi] + \mathcal{E}_0[A]. \tag{4.6.2}$$

Now the smallness assumption $\mathrm{S}_1[\phi, \mathbf{A}](0) \leqslant \varepsilon$ implies that $\mathrm{S}_1[\mathbf{A}](0) \leqslant \varepsilon$ and $\mathrm{S}_1[\phi](0) \leqslant \varepsilon$, so by proposition 4.5.6

$$\|\nabla\!\!\!\!/ A_0\|^2_{L^2}(0) \lesssim \mathrm{S}_1[\phi](0) \leqslant \varepsilon,$$

and so $\mathrm{S}_1[A](0) \lesssim \varepsilon$. Then by theorem 4.5.5, $\mathcal{E}_0[\phi] \simeq \mathrm{S}_1[\phi](0)$. Now equation (4.5.2) reads $\mathcal{E}_\tau[A] \simeq \mathrm{S}_1[A](\tau)$, which in particular holds at $\tau = 0$, so we have $\mathcal{E}_0[\phi] + \mathcal{E}_0[A] \simeq \mathrm{S}_1[\phi](0) + \mathrm{S}_1[A](0)$, and so by (4.6.2)

$$\mathcal{E}_\tau[\phi] + \mathcal{E}_\tau[A] \simeq \mathrm{S}_1[\phi](0) + \mathrm{S}_1[A](0).$$

This means that $\mathcal{E}_\tau[\phi] + \mathcal{E}_\tau[A]$ is small too, $\mathcal{E}_\tau[\phi, A] \lesssim \varepsilon$. In particular, $\mathcal{E}_\tau[A] \simeq \mathrm{S}_1[A](\tau)$ is small, so again by theorem 4.5.5, $\mathcal{E}_\tau[\phi] \simeq \mathrm{S}_1[\phi](\tau)$. We deduce that

$$\mathrm{S}_1[\phi](\tau) + \mathrm{S}_1[A](\tau) \simeq \mathrm{S}_1[\phi](0) + \mathrm{S}_1[A](0) \tag{4.6.3}$$

for all $\tau > 0$. The same argument works for $\tau < 0$.

□

4.6.3 H^2 Estimates

Commutators

Proposition 4.6.2. *One has the following bounds on the commutators of $\nabla\!\!\!\!/$ with the field equation operators $\mathscr{F}$ and $\mathscr{W}$:*

$$|[\nabla\!\!\!\!/, \mathscr{F}]A|_{\mathbb{S}^3} \lesssim |\nabla\!\!\!\!/^2 \mathbf{A}| + |\nabla\!\!\!\!/ \mathbf{A}| + |\nabla\!\!\!\!/ \dot{A}_0|,$$

and

$$\begin{aligned}|[\nabla\!\!\!\!/, \mathscr{W}]\phi| \lesssim{}& |\dot{\phi} \nabla\!\!\!\!/ A_0| + |\phi \nabla\!\!\!\!/ \dot{A}_0| + |\phi A_0 \nabla\!\!\!\!/ A_0| + |\nabla\!\!\!\!/^2 \phi| + |\phi \nabla\!\!\!\!/^2 \mathbf{A}| \\ &+ |\mathbf{A} \nabla\!\!\!\!/ \phi| + |\nabla\!\!\!\!/ \phi| + |\phi \nabla\!\!\!\!/ \mathbf{A}| + |\mathbf{A}\phi| + |\nabla\!\!\!\!/ \phi \nabla\!\!\!\!/ \mathbf{A}| + |\phi \mathbf{A} \nabla\!\!\!\!/ \mathbf{A}|.\end{aligned}$$

Proof. Note that in the following the index i always refers to a contraction with a basis vector field X_i. The operator $\mathscr{F}_\mu$ on A_a is given by $\mathscr{F}[A]_\mu = \Box \mathbf{A}_\mu - \nabla\!\!\!\!/_\mu \dot{A}_0 - 2\mathbf{A}_\mu$, so for any i

$$\begin{aligned}|[\nabla\!\!\!\!/_i, \mathscr{F}]A|_{\mathbb{S}^3} &= |\nabla\!\!\!\!/_i \mathscr{F}[A]_\mu - \mathscr{F}[\nabla\!\!\!\!/_i A]_\mu| \\ &= \left|\nabla\!\!\!\!/_i \left(\Box \mathbf{A}_\mu - \nabla\!\!\!\!/_\mu \dot{A}_0 - 2\mathbf{A}_\mu\right) - \Box(\nabla\!\!\!\!/_i \mathbf{A}_\mu) + \nabla\!\!\!\!/_\mu \nabla\!\!\!\!/_i \dot{A}_0 + 2\nabla\!\!\!\!/_i \mathbf{A}_\mu\right| \\ &= \left|\nabla\!\!\!\!/_i \nabla\!\!\!\!/^\nu \nabla\!\!\!\!/_\nu \mathbf{A}_\mu - \nabla\!\!\!\!/^\nu \nabla\!\!\!\!/_\nu (X_i^\lambda \nabla\!\!\!\!/_\lambda \mathbf{A}_\mu) + \nabla\!\!\!\!/_\mu X_i^\nu \nabla\!\!\!\!/_\nu \dot{A}_0\right| \\ &\leqslant C\left[|\nabla\!\!\!\!/^2 \mathbf{A}| + |\nabla\!\!\!\!/ \mathbf{A}| + |\nabla\!\!\!\!/ \dot{A}_0|\right],\end{aligned}$$

where the constant C depends only on the geometry of $\mathbb{S}^3$. To calculate the other commutator we need a couple of preliminary formulae. Let ψ be any sufficiently regular complex scalar field. Then

$$[\not\nabla_i,\, \mathrm{D}_0]\psi = \not\nabla_i(\dot{\psi} + iA_0\psi) - \mathrm{D}_0\not\nabla_i\psi = i\psi\not\nabla_i A_0,$$

and similarly

$$[\not\nabla_i,\, \not{\mathrm{D}}_\mu]\psi = -(\not\nabla_\mu X_i^\nu)\not\nabla_\nu\psi + i\psi\not\nabla_i\mathbf{A}_\mu,$$

so

$$\begin{aligned}[\not\nabla_i,\, \mathrm{D}_0\mathrm{D}_0]\phi &= \mathrm{D}_0[\not\nabla_i,\, \mathrm{D}_0]\phi + [\not\nabla_i,\, \mathrm{D}_0]\mathrm{D}_0\phi \\ &= \mathrm{D}_0(i\phi\not\nabla_i A_0) + i\mathrm{D}_0\phi\not\nabla_i A_0 \\ &= i\phi\not\nabla_i\dot{A}_0 + 2i\dot{\phi}\not\nabla_i A_0 - 2\phi A_0\not\nabla_i A_0.\end{aligned}$$

Further, for any vector field $\mathbf{V}$ on $\mathbb{S}^3$

$$\begin{aligned}[\not\nabla_i,\, \not{\mathrm{D}}_\mu]\mathbf{V}^\mu &= \not\nabla_i(\not\nabla_\mu\mathbf{V}^\mu + i\mathbf{A}_\mu\mathbf{V}^\mu) - (\not\nabla_\mu + i\mathbf{A}_\mu)(\not\nabla_i\mathbf{V}^\mu) \\ &= \not\nabla_i\not\nabla_\mu\mathbf{V}^\mu - \not\nabla_\mu\not\nabla_i\mathbf{V}^\mu + i(\not\nabla_i\mathbf{A}_\mu)\mathbf{V}^\mu \\ &\leqslant C\left[|\not\nabla\mathbf{V}| + |\mathbf{V}| + |\mathbf{V}\not\nabla\mathbf{A}|\right],\end{aligned}$$

where, as before, C depends only on the geometry of $\mathbb{S}^3$. Then

$$\begin{aligned}[\not\nabla_i,\, \not{\mathrm{D}}_\mu\not{\mathrm{D}}^\mu]\phi &= \not{\mathrm{D}}^\mu[\not\nabla_i,\, \not{\mathrm{D}}_\mu]\phi + [\not\nabla_i,\, \not{\mathrm{D}}_\mu]\not{\mathrm{D}}^\mu\phi \\ &\leqslant \not{\mathrm{D}}^\mu\left(-\not\nabla_\mu X_i^\nu\not\nabla_\nu\phi + i\phi\not\nabla_i\mathbf{A}_\mu\right) + C\left[|\not\nabla\not{\mathrm{D}}\phi| + |\not{\mathrm{D}}\phi| + |\not{\mathrm{D}}\phi\not\nabla\mathbf{A}|\right] \\ &\leqslant -\not\Delta X_i^\nu\not\nabla_\nu\phi - \not\nabla_\mu X_i^\nu\not\nabla^\mu\not\nabla_\nu\phi + i\not\nabla^\mu\phi\not\nabla_i\mathbf{A}_\mu + i\phi\not\nabla_\mu\not\nabla_i\mathbf{A}_\mu \\ &\quad - i\mathbf{A}^\mu\not\nabla_\mu X_i^\nu\not\nabla_\nu\phi - \phi\mathbf{A}^\mu\not\nabla_i\mathbf{A}_\mu \\ &\quad + C\left[|\not\nabla^2\phi| + |\not\nabla(\mathbf{A}\phi)| + |\not\nabla\phi| + |\mathbf{A}\phi| + |\not\nabla\phi\not\nabla\mathbf{A}| + |\mathbf{A}\phi\not\nabla\mathbf{A}|\right] \\ &\lesssim |\not\nabla\phi| + |\not\nabla^2\phi| + |\not\nabla\phi\not\nabla\mathbf{A}| + |\phi\not\nabla^2\mathbf{A}| \\ &\quad + |\phi\not\nabla\mathbf{A}| + |\mathbf{A}\not\nabla\phi| + |\phi\mathbf{A}\not\nabla\mathbf{A}| + |\mathbf{A}\phi|.\end{aligned}$$

Putting these together, we have

$$\begin{aligned}[\not\nabla_i,\, \mathscr{W}]\phi &= [\not\nabla_i,\, \mathrm{D}^a\mathrm{D}_a + 1]\phi \\ &= [\not\nabla_i,\, \mathrm{D}_0\mathrm{D}_0]\phi - [\not\nabla_i,\, \not{\mathrm{D}}^\mu\not{\mathrm{D}}_\mu]\phi \\ &\lesssim |\phi\not\nabla\dot{A}_0| + |\dot{\phi}\not\nabla A_0| + |\phi A_0\not\nabla A_0| + |\not\nabla\phi| + |\not\nabla^2\phi| + |\not\nabla\phi\not\nabla\mathbf{A}| \\ &\quad + |\phi\not\nabla^2\mathbf{A}| + |\phi\not\nabla\mathbf{A}| + |\mathbf{A}\not\nabla\phi| + |\phi\mathbf{A}\not\nabla\mathbf{A}| + |\mathbf{A}\phi|.\end{aligned}$$

□

Most of the terms in the above estimates we can control by the energy directly, with the exception of time derivatives of A_0. These terms we shall control using the elliptic equation for A_0 and the evolution equation for ϕ.

Proposition 4.6.3. *For any fixed $\tau \in \mathbb{R}$ there exists $\varepsilon > 0$ such that if $\mathrm{S}_1[\phi] < \varepsilon$ and A_a satisfies the strong Coulomb gauge, then*

$$\|\dot{A}_0\|^2_{H^1}(\tau) \lesssim \mathrm{S}_2[\phi](\tau)(1 + \mathrm{S}_1[A](\tau))^2.$$

Proof. First note that in the strong Coulomb gauge $\bar{A}_0(\tau) = 0$ for all τ, and so $\dot{\bar{A}}_0(\tau) = 0$ for all τ as well. Thus $\|\dot{A}_0\|_{L^2} \lesssim \|\nabla\!\!\!/ \dot{A}_0\|_{L^2}$, and we only need to estimate $\|\nabla\!\!\!/ \dot{A}_0\|_{L^2}$. Differentiating (4.5.5) in τ, we have

$$-\Delta\!\!\!/ \dot{A}_0 + |\phi|^2 \dot{A}_0 = -\operatorname{Im}(\bar{\phi}\ddot{\phi}) - \bar{\phi}\dot{\phi}A_0 - \dot{\bar{\phi}}\phi A_0.$$

Multiplying through by $\dot{A}_0$ and integrating we have

$$\|\nabla\!\!\!/ \dot{A}_0\|_{L^2}^2 + \|\phi\dot{A}_0\|_{L^2}^2 \leqslant \|\phi\dot{A}_0\|_{L^2}\|\ddot{\phi}\|_{L^2} + 2\|\phi\dot{A}_0\|_{L^2}\|\dot{\phi}A_0\|_{L^2},$$

which gives

$$\|\nabla\!\!\!/ \dot{A}_0\|_{L^2}^2 + \delta\|\phi\dot{A}_0\|_{L^2}^2 \lesssim \|\ddot{\phi}\|_{L^2}^2 + \|\dot{\phi}A_0\|_{L^2}^2 \tag{4.6.4}$$

for some $0 < \delta < 1$. We thus need to estimate $\|\ddot{\phi}\|_{L^2}$, for which we shall use the field equation for ϕ,

$$\Box\phi + 2iA_0\dot{\phi} - 2i\mathbf{A}\cdot\nabla\!\!\!/\phi + (1 - A_0^2 + |\mathbf{A}|^2 + i\dot{A}_0)\phi = 0.$$

We estimate

$$|\ddot{\phi}|^2 \lesssim |\Delta\!\!\!/\phi|^2 + |A_0\dot{\phi}|^2 + |\mathbf{A}\nabla\!\!\!/\phi|^2 + |\phi|^2 + |A_0^2\phi|^2 + |\mathbf{A}^2\phi|^2 + |\dot{A}_0\phi|^2. \tag{4.6.5}$$

With the exception of the term $|\dot{A}_0\phi|^2$, the right-hand side of (4.6.5) will be easily controlled as we will see shortly. To deal with the problematic term we will use smallness of the data. Integrating (4.6.5) over the 3-sphere we have

$$\begin{aligned}
\|\ddot{\phi}\|_{L^2}^2 &\lesssim \|\Delta\!\!\!/\phi\|_{L^2}^2 + \|A_0\dot{\phi}\|_{L^2}^2 + \|\mathbf{A}\nabla\!\!\!/\phi\|_{L^2}^2 + \|\phi\|_{L^2}^2 + \|A_0^2\phi\|_{L^2}^2 + \|\mathbf{A}^2\phi\|_{L^2}^2 + \|\dot{A}_0\phi\|_{L^2}^2 \\
&\lesssim \|\phi\|_{H^2}^2 + \|A_0\|_{L^3}^2\|\dot{\phi}\|_{L^6}^2 + \|\mathbf{A}\|_{L^3}^2\|\nabla\!\!\!/\phi\|_{L^6}^2 \\
&\quad + \|A_0\|_{L^6}^4\|\phi\|_{L^6}^2 + \|\mathbf{A}\|_{L^6}^4\|\phi\|_{L^6}^2 + \|\dot{A}_0\|_{L^3}^2\|\phi\|_{L^6}^2 \\
&\lesssim \|\phi\|_{H^2}^2 + \|A_0\|_{H^1}^2\|\dot{\phi}\|_{H^1}^2 + \|\mathbf{A}\|_{H^1}^2\|\phi\|_{H^2}^2 \\
&\quad + \|A_0\|_{H^1}^4\|\phi\|_{H^1}^2 + \|\mathbf{A}\|_{H^1}^4\|\phi\|_{H^1}^2 + \|\dot{A}_0\|_{H^1}^2\|\phi\|_{H^1}^2 \\
&\lesssim \mathrm{S}_2[\phi] + \mathrm{S}_1[A]\mathrm{S}_2[\phi] + \mathrm{S}_1[\mathbf{A}]\mathrm{S}_2[\phi] + \mathrm{S}_1[A]^2\mathrm{S}_1[\phi] + \mathrm{S}_1[\mathbf{A}]^2\mathrm{S}_1[\phi] + \|\dot{A}_0\|_{H^1}^2\mathrm{S}_1[\phi] \\
&\lesssim \mathrm{S}_2[\phi](1 + \mathrm{S}_1[A])^2 + \|\dot{A}_0\|_{H^1}^2\mathrm{S}_1[\phi].
\end{aligned}$$

Putting this into (4.6.4) gives

$$\|\nabla\!\!\!/ \dot{A}_0\|_{L^2}^2 \lesssim \mathrm{S}_2[\phi](1 + \mathrm{S}_1[A])^2 + \|\dot{A}_0\|_{H^1}^2\mathrm{S}_1[\phi],$$

so provided $\mathrm{S}_1[\phi]$ is small enough the Poincaré inequality gives

$$\|\nabla\!\!\!/ \dot{A}_0\|_{L^2}^2 \lesssim \mathrm{S}_2[\phi](1 + \mathrm{S}_1[A])^2.$$

□

Estimate Algebra

For ease of presentation we outline a schematic procedure to track how we bound the various terms arising in our H^2 estimates. The idea is simply to track the number of derivatives and their Sobolev exponents of the error terms and check that they do not exceed certain critical values. Let f denote either A_a or ϕ, let ∂ denote either the $\mathbb{S}^3$-derivatives $\not\nabla$ or the τ-derivative ∂_τ, and let ∂^2 denote either $\not\nabla^2$ or $\partial_\tau \not\nabla$ (that is, not ∂_τ^2). Then all the error terms that we encounter will in fact be of the form

$$\| |\partial^2 f|^m |\partial f|^k |f|^l \|_{L^1(\mathbb{S}^3)},$$

where m, k, and l are non-negative integers and in particular $m = 0$, 1, or 2.

If $m = 0$, we have

$$\| |\partial f|^k |f|^l \|_{L^1} \leqslant \|f\|^l_{L^\infty} \|\partial f\|^k_{L^k}.$$

Now since $\mathbb{S}^3$ is compact, the Lebesgue spaces $L^p(\mathbb{S}^3)$ form a decreasing sequence in p,

$$L^\infty(\mathbb{S}^3) \hookrightarrow \cdots \hookrightarrow L^p(\mathbb{S}^3) \hookrightarrow \cdots \hookrightarrow L^q(\mathbb{S}^3) \hookrightarrow \cdots \hookrightarrow L^1(\mathbb{S}^3),$$

$p > q$, where $\hookrightarrow$ denotes continuous inclusion. As $\mathbb{S}^3$ has dimension 3, by Sobolev embedding we also have

$$H^1(\mathbb{S}^3) \hookrightarrow L^6(\mathbb{S}^3) \quad \text{and} \quad H^2(\mathbb{S}^3) \hookrightarrow \mathcal{C}^{0,\frac{1}{2}}(\mathbb{S}^3) \hookrightarrow L^\infty(\mathbb{S}^3),$$

so provided $k \leqslant 6$ we have

$$\| |\partial f|^k |f|^l \|_{L^1} \lesssim \|f\|_2^l \|f\|_2^k = \|f\|_2^{k+l},$$

where

$$\|f\|_2 := \|f\|_{H^2(\mathbb{S}^3)} + \|\dot{f}\|_{H^1(\mathbb{S}^3)}$$

(notice that the norm $\|\cdot\|_2^2$ is the familiar Sobolev-type approximate energy S_2).

If $m = 1$, we perform the splitting

$$\| |\partial^2 f| |\partial f|^k |f|^l \|_{L^1} = \int |\partial^2 f| |\partial f|^k |f|^l \leqslant \int |\partial^2 f|^2 + \int |\partial f|^{2k} |f|^{2l} \leqslant \|f\|_2^2 + \| |\partial f|^{2k} |f|^{2l} \|_{L^1}.$$

Now provided $2k \leqslant 6$, the second term in the above may be dealt with as in the case $m = 0$, so we have

$$\| |\partial^2 f| |\partial f|^k |f|^l \|_{L^1} \lesssim \|f\|_2^2 + \|f\|_2^{2(k+l)}.$$

Finally, when $m = 2$ it will in fact turn out that k is necessarily zero, so we will have

$$\| |\partial^2 f|^2 |f|^l \|_{L^1} \leqslant \|f\|^l_{L^\infty} \|f\|_2^2 \lesssim \|f\|_2^{l+2}.$$

It will thus be sufficient to use the following prescription. For terms involving no $|\partial^2 f|$ (i.e. $m = 0$) we shall check whether $k \leqslant 6$, and if so, conclude that the term is bounded by $\|f\|_2^{k+l}$; for terms involving $|\partial^2 f|$ (i.e. $m = 1$), we shall check whether $k \leqslant 3$, and if so, conclude that the term is bounded by $\|f\|_2^2 + \|f\|_2^{2(k+l)}$; finally, for terms with $m = 2$

we shall check whether $k = 0$, and if so, conclude that these are bounded by $\|f\|_2^{l+2}$. In the estimates that follow we will write down a term to be estimated,

$$|\partial^2 f|^m |\partial f|^k |f|^l,$$

and underneath note down its 'signature' (m, k, l), as in

$$\underset{(m,k,l)}{|\partial^2 f|^m |\partial f|^k |f|^l}.$$

If the criteria outlined above are met (that is, $k \leqslant 6$ for $m = 0$, $k \leqslant 3$ for $m = 1$, and $k = 0$ for $m = 2$), we shall tick the triplet,

$$\underset{(m,k,l)\checkmark}{|\partial^2 f|^m |\partial f|^k |f|^l}.$$

Altogether this notation will thus mean that

$$\| |\partial^2 f|^m |\partial f|^k |f|^l \|_{L^1(\mathbb{S}^3)} \lesssim Q(\|f\|_2)$$

for some polynomial Q with positive coefficients.

H^2 Error Terms

We now take $\alpha_a = \bar\nabla_i A_a$ and $\psi = \bar\nabla_i \phi$ in (4.6.1) and estimate the second order error terms

$$\mathsf{e}_2 := \sum_i T^b \left(\nabla^a \mathbf{T}_{ab}[\bar\nabla_i A] + \nabla^a \mathbf{T}_{ab}[\bar\nabla_i \phi] \right).$$

Equation (4.6.1) gives

$$\begin{aligned}
\mathsf{e}_2 &= \sum_i -\mathscr{F}[\bar\nabla_i A]^\mu (\bar\nabla_\mu \bar\nabla_i A_0 - \bar\nabla_i \dot{\mathbf{A}}_\mu) \\
&+ \sum_i \left(\frac{1}{2} \overline{\mathscr{W}[\bar\nabla_i \phi]} \mathrm{D}_0(\bar\nabla_i \phi) + \frac{1}{2} \mathscr{W}[\bar\nabla_i \phi] \overline{\mathrm{D}_0(\bar\nabla_i \phi)} - (\bar\nabla_\mu A_0 - \dot{\mathbf{A}}_\mu) \operatorname{Im}(\bar\nabla_i \bar\phi \bar{\mathrm{D}}^\mu \bar\nabla_i \phi) \right) \\
&=: \mathsf{e}_2^1 + \mathsf{e}_2^2,
\end{aligned}$$

and we consider e_2^1 and e_2^2 separately. We have

$$\begin{aligned}
|\mathsf{e}_2^1| &= \left| \sum_i -\mathscr{F}[\bar\nabla_i A]^\mu (\bar\nabla_\mu \bar\nabla_i A_0 - \bar\nabla_i \dot{\mathbf{A}}_\mu) \right| \\
&\leqslant \sum_i \left| (\bar\nabla_i \mathscr{F}[A]^\mu - [\bar\nabla_i, \mathscr{F}][A]^\mu) \left(\bar\nabla_\mu \bar\nabla_i A_0 - \bar\nabla_i \dot{\mathbf{A}}_\mu \right) \right| \\
&\lesssim \left| \bar\nabla(\bar\phi \bar{\mathrm{D}} \phi) \right| \left[|\bar\nabla^2 A_0| + |\bar\nabla A_0| + |\bar\nabla \dot{\mathbf{A}}| \right] \\
&+ \left[|\bar\nabla^2 \mathbf{A}| + |\bar\nabla \mathbf{A}| + |\bar\nabla \dot{A}_0| \right] \left[|\bar\nabla^2 A_0| + |\bar\nabla A_0| + |\bar\nabla \dot{\mathbf{A}}| \right] \\
&\lesssim \left[|\bar\nabla \phi|^2 + |\bar\nabla \phi||\phi||\mathbf{A}| + |\bar\nabla^2 \phi||\phi| + |\bar\nabla \mathbf{A}||\phi|^2 + |\bar\nabla \phi||\phi||\mathbf{A}| \right] \left[|\bar\nabla^2 A_0| + |\bar\nabla A_0| + |\bar\nabla \dot{\mathbf{A}}| \right] \\
&+ \left[|\bar\nabla^2 \mathbf{A}| + |\bar\nabla \mathbf{A}| + |\bar\nabla \dot{A}_0| \right] \left[|\bar\nabla^2 A_0| + |\bar\nabla A_0| + |\bar\nabla \dot{\mathbf{A}}| \right] \\
&\lesssim \underset{(1,2,0)\checkmark}{|\bar\nabla^2 A_0||\bar\nabla \phi|^2} + \underset{(1,1,2)\checkmark}{|\bar\nabla^2 A_0||\bar\nabla \phi||\phi||\mathbf{A}|} + \underset{(2,0,1)\checkmark}{|\bar\nabla^2 A_0||\bar\nabla^2 \phi||\phi|} + \underset{(1,1,2)\checkmark}{|\bar\nabla^2 A_0||\bar\nabla \mathbf{A}||\phi|^2}
\end{aligned}$$

$$
\begin{aligned}
&+ \underset{(1,1,2)\checkmark}{|\slashed{\nabla}^2 A_0||\slashed{\nabla}\phi||\phi||\mathbf{A}|} + \underset{(0,3,0)\checkmark}{|\slashed{\nabla} A_0||\slashed{\nabla}\phi|^2} + \underset{(0,2,2)\checkmark}{|\slashed{\nabla} A_0||\slashed{\nabla}\phi||\phi||\mathbf{A}|} + \underset{(1,1,1)\checkmark}{|\slashed{\nabla}^2\phi||\slashed{\nabla} A_0||\phi|} \\
&+ \underset{(0,2,2)\checkmark}{|\slashed{\nabla} A_0||\slashed{\nabla}\mathbf{A}||\phi|^2} + \underset{(0,2,2)\checkmark}{|\slashed{\nabla} A_0||\slashed{\nabla}\phi||\phi||\mathbf{A}|} + \underset{(1,2,0)\checkmark}{|\slashed{\nabla}\dot{\mathbf{A}}||\slashed{\nabla}\phi|^2} + \underset{(1,1,2)\checkmark}{|\slashed{\nabla}\dot{\mathbf{A}}||\slashed{\nabla}\phi||\phi||\mathbf{A}|} \\
&+ \underset{(2,0,1)\checkmark}{|\slashed{\nabla}\dot{\mathbf{A}}||\slashed{\nabla}^2\phi||\phi|} + \underset{(1,1,2)\checkmark}{|\slashed{\nabla}\dot{\mathbf{A}}||\slashed{\nabla}\mathbf{A}||\phi|^2} + \underset{(1,1,2)\checkmark}{|\slashed{\nabla}\dot{\mathbf{A}}||\slashed{\nabla}\phi||\phi||\mathbf{A}|} + \underset{(2,0,0)\checkmark}{|\slashed{\nabla}^2 A_0||\slashed{\nabla}^2\mathbf{A}|} \\
&+ \underset{(1,1,0)\checkmark}{|\slashed{\nabla}^2 A_0||\slashed{\nabla}\mathbf{A}|} + \underset{(2,0,0)\checkmark}{|\slashed{\nabla}^2 A_0||\slashed{\nabla}\dot{A}_0|} + \underset{(1,1,0)\checkmark}{|\slashed{\nabla}^2\mathbf{A}||\slashed{\nabla} A_0|} + \underset{(0,2,0)\checkmark}{|\slashed{\nabla} A_0||\slashed{\nabla}\mathbf{A}|} + \underset{(1,1,0)\checkmark}{|\slashed{\nabla}\dot{A}_0||\slashed{\nabla} A_0|} \\
&+ \underset{(2,0,0)\checkmark}{|\slashed{\nabla}^2\mathbf{A}||\slashed{\nabla}\dot{\mathbf{A}}|} + \underset{(1,1,0)\checkmark}{|\slashed{\nabla}\dot{\mathbf{A}}||\slashed{\nabla}\mathbf{A}|} + \underset{(2,0,0)\checkmark}{|\slashed{\nabla}\dot{A}_0||\slashed{\nabla}\dot{\mathbf{A}}|}
\end{aligned}
$$

and

$$
\begin{aligned}
|\mathsf{e}_2^2| &= \left|\sum_i \left(\frac{1}{2}\overline{\mathscr{W}[\slashed{\nabla}_i\phi]}\mathrm{D}_0(\slashed{\nabla}_i\phi) + \frac{1}{2}\mathscr{W}[\slashed{\nabla}_i\phi]\overline{\mathrm{D}_0(\slashed{\nabla}_i\phi)} - (\slashed{\nabla}_\mu A_0 - \dot{\mathbf{A}}_\mu)\,\mathrm{Im}(\slashed{\nabla}_i\phi\slashed{D}^\mu\slashed{\nabla}_i\phi)\right)\right| \\
&\leqslant \sum_i \Big[|\mathscr{W}[\slashed{\nabla}_i\phi]||\mathrm{D}_0(\slashed{\nabla}_i\phi)| + |\slashed{\nabla} A_0 - \dot{\mathbf{A}}||\slashed{\nabla}_i\phi||\slashed{D}\slashed{\nabla}_i\phi|\Big] \\
&\lesssim \sum_i |[\slashed{\nabla}_i,\, \mathscr{W}]\phi|\Big[|\slashed{\nabla}_i\dot{\phi}| + |A_0\slashed{\nabla}_i\phi|\Big] + \Big[|\slashed{\nabla} A_0| + |\dot{\mathbf{A}}|\Big]|\slashed{\nabla}\phi|\Big[|\slashed{\nabla}^2\phi| + |\mathbf{A}\slashed{\nabla}\phi| + |\slashed{\nabla}\phi|\Big] \\
&\lesssim \Big[|\dot{\phi}\slashed{\nabla} A_0| + |\phi\slashed{\nabla}\dot{A}_0| + |\phi A_0\slashed{\nabla} A_0| + |\slashed{\nabla}^2\phi| + |\phi\slashed{\nabla}^2\mathbf{A}| + |\slashed{\nabla} A_0||\slashed{\nabla}\phi|^2 + |\dot{\mathbf{A}}||\slashed{\nabla}\phi|^2 \\
&\quad + |\mathbf{A}\slashed{\nabla}\phi| + |\slashed{\nabla}\phi| + |\phi\slashed{\nabla}\mathbf{A}| + |\mathbf{A}\phi| + |\slashed{\nabla}\phi\slashed{\nabla}\mathbf{A}| + |\phi\mathbf{A}\slashed{\nabla}\mathbf{A}|\Big]\Big[|\slashed{\nabla}\dot{\phi}| + |A_0\slashed{\nabla}\phi|\Big] \\
&\quad + |\slashed{\nabla}^2\phi||\slashed{\nabla} A_0||\slashed{\nabla}\phi| + |\slashed{\nabla} A_0||\slashed{\nabla}\phi|^2|\mathbf{A}| + |\slashed{\nabla}^2\phi||\slashed{\nabla}\phi||\dot{\mathbf{A}}| \\
&\quad + |\slashed{\nabla}\phi|^2|\dot{\mathbf{A}}||\mathbf{A}| + |\slashed{\nabla} A_0||\slashed{\nabla}\phi|^2 + |\dot{\mathbf{A}}||\slashed{\nabla}\phi|^2 \\
&\lesssim \underset{(1,2,0)\checkmark}{|\slashed{\nabla}\dot{\phi}||\slashed{\nabla} A_0||\dot{\phi}|} + \underset{(2,0,1)\checkmark}{|\slashed{\nabla}\dot{A}_0||\slashed{\nabla}\dot{\phi}||\phi|} + \underset{(1,1,2)\checkmark}{|\slashed{\nabla}\dot{\phi}||\slashed{\nabla} A_0||\phi||A_0|} + \underset{(2,0,0)\checkmark}{|\slashed{\nabla}^2\phi||\slashed{\nabla}\dot{\phi}|} + \underset{(2,0,1)\checkmark}{|\slashed{\nabla}^2\mathbf{A}||\slashed{\nabla}\dot{\phi}||\phi|} \\
&\quad + \underset{(1,1,1)\checkmark}{|\slashed{\nabla}\dot{\phi}||\slashed{\nabla}\phi||\mathbf{A}|} + \underset{(1,1,0)\checkmark}{|\slashed{\nabla}\dot{\phi}||\slashed{\nabla}\phi|} + \underset{(1,1,1)\checkmark}{|\slashed{\nabla}\dot{\phi}||\slashed{\nabla}\mathbf{A}||\phi|} + \underset{(1,0,2)\checkmark}{|\slashed{\nabla}\dot{\phi}||\mathbf{A}||\phi|} + \underset{(1,2,0)\checkmark}{|\slashed{\nabla}\dot{\phi}||\slashed{\nabla}\phi||\slashed{\nabla}\mathbf{A}|} \\
&\quad + \underset{(1,1,2)\checkmark}{|\slashed{\nabla}\dot{\phi}||\slashed{\nabla}\mathbf{A}||\mathbf{A}||\phi|} + \underset{(0,3,1)\checkmark}{|\slashed{\nabla}\phi||\dot{\phi}||\slashed{\nabla} A_0||A_0|} + \underset{(1,1,2)\checkmark}{|\slashed{\nabla}\dot{A}_0||\slashed{\nabla}\phi||\phi||A_0|} + \underset{(0,2,3)\checkmark}{|\slashed{\nabla} A_0||\slashed{\nabla}\phi||\phi||A_0|^2} \\
&\quad + \underset{(1,1,1)\checkmark}{|\slashed{\nabla}^2\phi||\slashed{\nabla}\phi||A_0|} + \underset{(1,1,2)\checkmark}{|\slashed{\nabla}^2\mathbf{A}||\slashed{\nabla}\phi||\phi||A_0|} + \underset{(0,2,2)\checkmark}{|\slashed{\nabla}\phi|^2|A_0||\mathbf{A}|} + \underset{(0,2,1)\checkmark}{|\slashed{\nabla}\phi|^2|A_0|} \\
&\quad + \underset{(0,2,2)\checkmark}{|\slashed{\nabla}\phi||\slashed{\nabla}\mathbf{A}||\phi||A_0|} + \underset{(0,1,3)\checkmark}{|\slashed{\nabla}\phi||\phi||A_0||\mathbf{A}|} + \underset{(0,3,1)\checkmark}{|\slashed{\nabla}\phi|^2|\slashed{\nabla}\mathbf{A}||A_0|} + \underset{(0,2,3)\checkmark}{|\slashed{\nabla}\phi||\slashed{\nabla}\mathbf{A}||\phi||A_0||\mathbf{A}|} \\
&\quad + \underset{(1,2,0)\checkmark}{|\slashed{\nabla}^2\phi||\slashed{\nabla}\phi||\slashed{\nabla} A_0|} + \underset{(0,3,1)\checkmark}{|\slashed{\nabla}\phi|^2|\slashed{\nabla} A_0||\mathbf{A}|} + \underset{(1,2,0)\checkmark}{|\slashed{\nabla}^2\phi||\slashed{\nabla}\phi||\dot{\mathbf{A}}|} + \underset{(0,3,1)\checkmark}{|\slashed{\nabla}\phi|^2|\dot{\mathbf{A}}||\mathbf{A}|} + \underset{(0,3,0)\checkmark}{|\slashed{\nabla}\phi|^2|\slashed{\nabla} A_0|} \\
&\quad + \underset{(0,3,0)\checkmark}{|\slashed{\nabla}\phi|^2|\dot{\mathbf{A}}|}.
\end{aligned}
$$

Altogether this says that

$$\|\mathsf{e}_2\|_{L^1(\mathbb{S}^3)} \lesssim Q^{\mathrm{IV}}\left(\|(\phi, \mathbf{A}, A_0)\|_2\right)$$

for some polynomial Q^{IV} with positive coefficients. An inspection of the triplets (m, k, l) above shows that in fact each error term contains at least one full power of $\|f\|_2^2$, so in fact

$$
\begin{aligned}
\|\mathsf{e}_2\|_{L^1} &\lesssim \|(\phi, \mathbf{A}, A_0)\|_2^2 Q^{\mathrm{III}}\left(\|(\phi, \mathbf{A}, A_0)\|_2\right) \\
&\lesssim \left(\mathrm{S}_2[\phi, A] + \|\dot{A}_0\|_{H^1}^2\right)\left(Q^{\mathrm{II}}\left(\mathrm{S}_2[\phi, A]^{1/2}\right) + Q^{\mathrm{I}}\left(\|\dot{A}_0\|_{H^1}\right)\right)
\end{aligned}
$$

for polynomials $Q^{\mathrm{I,II,III}}$. Now by proposition 4.6.3, $\|\dot{A}_0\|^2_{H^1} \leqslant C\mathrm{S}_2[\phi](1+\mathrm{S}_1[A])^2$. At this point we can either assume the first order estimates (theorem 4.6.1), or bound $\|\dot{A}_0\|^2_{H^1}$ by a polynomial in $\mathrm{S}_2[\phi, A]$ of degree higher than one; both methods are fine, but we will need to assume the first order estimates to close the second order ones anyway, so assuming $\mathrm{S}_1[\phi, A] \lesssim 1$ we have $\|\dot{A}_0\|^2_{H^1} \lesssim \mathrm{S}_2[\phi, A]$. Hence for any fixed τ

$$\|\mathsf{e}_2\|_{L^1}(\tau) \lesssim \mathrm{S}_2[\phi, A](\tau) P\left(\mathrm{S}_2[\phi, A](\tau)^{1/2}\right) \tag{4.6.6}$$

for some polynomial P.

Theorem 4.6.4. *Let I be a fixed compact interval in $\mathbb{R}$ containing zero. There exists $\varepsilon > 0$ such that if $\mathrm{S}_2[\phi, \mathbf{A}](0) \leqslant \varepsilon$, then*

$$\mathrm{S}_2[\phi, A](\tau) \simeq \mathrm{S}_2[\phi, A](0)$$

for all $\tau \in I$.

Proof. Integrating e_2 over the region $\mathbb{S}^3 \times [0, \tau]$, $\tau > 0$,

$$\begin{aligned}\int_{\mathbb{S}^3\times[0,\tau]} \mathsf{e}_2 \,\mathrm{dv} &= \int_0^\tau \int_{\mathbb{S}^3} \mathsf{e}_2(\sigma)\,\mathrm{dv}_{\mathfrak{s}_3}\,\mathrm{d}\sigma \\ &= \sum_i \left(\mathcal{E}_\tau[\nabla\!\!\!\!/_i \phi] + \mathcal{E}_\tau[\nabla\!\!\!\!/_i A]\right) - \sum_i \left(\mathcal{E}_0[\nabla\!\!\!\!/_i \phi] + \mathcal{E}_0[\nabla\!\!\!\!/_i A]\right).\end{aligned} \tag{4.6.7}$$

From theorem 4.6.1 we know that $\mathrm{S}_1[\phi, A](\tau) \simeq \mathrm{S}_1[\phi, A](0)$, and also that $\mathcal{E}_\tau[A] \simeq \mathrm{S}_1[A](\tau)$ and $\mathcal{E}_\tau[\phi] \simeq \mathrm{S}_1[\phi](\tau)$ for all τ. Furthermore, we have that $\mathrm{S}_1[A](\tau)$ is small, so by (4.5.4)

$$\mathcal{E}_\tau[\phi] + \sum_i \mathcal{E}_\tau[\nabla\!\!\!\!/_i \phi] \simeq \mathrm{S}_2[\phi](\tau).$$

By (4.5.3),

$$\mathcal{E}_\tau[A] + \sum_i \mathcal{E}_\tau[\nabla\!\!\!\!/_i A] \simeq \mathrm{S}_2[A](\tau),$$

so adding $\mathcal{E}_\tau[\phi, A] = \mathcal{E}_0[\phi, A]$ to both sides of (4.6.7) we have

$$\begin{aligned}\mathcal{E}_\tau[\phi, A] + \sum_i \left(\mathcal{E}_\tau[\nabla\!\!\!\!/_i \phi] + \mathcal{E}_\tau[\nabla\!\!\!\!/_i A]\right) &= \mathcal{E}_0[\phi, A] + \sum_i \left(\mathcal{E}_0[\nabla\!\!\!\!/_i \phi] + \mathcal{E}_0[\nabla\!\!\!\!/_i A]\right) \\ &\quad + \int_0^\tau \int_{\mathbb{S}^3} \mathsf{e}_2(\sigma)\,\mathrm{dv}_{\mathfrak{s}_3}\,\mathrm{d}\sigma,\end{aligned}$$

or equivalently

$$\mathrm{S}_2[\phi, A](\tau) \simeq \mathrm{S}_2[\phi, A](0) + \int_0^\tau \int_{\mathbb{S}^3} \mathsf{e}_2(\sigma)\,\mathrm{dv}_{\mathfrak{s}_3}\,\mathrm{d}\sigma. \tag{4.6.8}$$

Now (4.6.6) gives

$$\begin{aligned}\mathrm{S}_2[\phi, A](\tau) &\lesssim \mathrm{S}_2[\phi, A](0) + \int_0^\tau \|\mathsf{e}_2\|_{L^1(\mathbb{S}^3)}(\sigma)\,\mathrm{d}\sigma \\ &\lesssim \mathrm{S}_2[\phi, A](0) + \int_0^\tau \mathrm{S}_2[\phi, A](\sigma) P\left(\mathrm{S}_2[\phi, A](\sigma)^{1/2}\right)\mathrm{d}\sigma,\end{aligned}$$

and so by lemma A.2.12 we have

$$\mathrm{S}_2[\phi, A](\tau) \lesssim \mathrm{S}_2[\phi, A](0)$$

for $\tau \in I$. Equation (4.6.8) similarly shows that $\mathrm{S}_2[\phi, A](0) \lesssim \mathrm{S}_2[\phi, A](\tau)$, and so

$$\mathrm{S}_2[\phi, A](\tau) \simeq \mathrm{S}_2[\phi, A](0).$$

for all $\tau \in I$. In particular, picking I large enough to contain $[-\pi/2, \pi/2]$ gives

$$\mathrm{S}_2[\phi, A](\mathscr{I}^-) \simeq \mathrm{S}_2[\phi, A](\mathscr{I}^+).$$

□

4.7 Higher Order Estimates

From here it is straightforward to play the same game for higher order estimates. It is clear that if for a given τ and $m \geqslant 1$ the $(m+1)$-th Sobolev energy $\mathrm{S}_{m+1}[\phi, A](\tau)$ is sufficiently small, then

$$\sum_{k=0}^{m} \mathcal{E}_\tau[\nabla\!\!\!/^k \phi] \simeq \mathrm{S}_{m+1}[\phi](\tau) \quad \text{and} \quad \sum_{k=0}^{m} \mathcal{E}_\tau[\nabla\!\!\!/^k A] \simeq \mathrm{S}_{m+1}[A](\tau),$$

where as before $\mathcal{E}_\tau[\nabla\!\!\!/^k \phi] = \sum_{i_1,\dots,i_k \in \{1,2,3\}} \mathcal{E}_\tau[\nabla\!\!\!/_{i_1} \dots \nabla\!\!\!/_{i_k} \phi]$, and similarly for A_a. We suppress sums over the basis vector fields $\{X_i\}$ from now on. It is clear that to prove that

$$\mathrm{S}_{m+1}[\phi, A](\tau) \simeq \mathrm{S}_{m+1}[\phi, A](0) \tag{4.7.1}$$

it is enough to prove the estimate

$$\|\mathsf{e}_{m+1}\|_{L^1}(\tau) \lesssim \mathrm{S}_{m+1}[\phi, A](\tau) P\left(\mathrm{S}_{m+1}[\phi, A](\tau)^{1/2}\right) \tag{4.7.2}$$

for a polynomial P, since then the proof of (4.7.1) then goes through exactly as in the proof of theorem 4.6.4. Now because

$$H^{m+1}(\mathbb{S}^3) \hookrightarrow \mathcal{C}^{m-1}(\mathbb{S}^3),$$

in our $(m+1)$-th order estimates we need only track derivatives of order m and higher, since all the others will be L^∞-controlled by S_{m+1}. More precisely, since the S_{m+1} energies control the L^∞ norms of $\nabla\!\!\!/^{m-1}\phi$, $\nabla\!\!\!/^{m-1}A$, $\nabla\!\!\!/^{m-2}\dot{\phi}$ and $\nabla\!\!\!/^{m-2}\dot{\mathbf{A}}$, we will only track terms of order higher than these (and also $\dot{A}_0$, which we will deal with separately as before). As before, one can write down the bounds for the commutators of $\nabla\!\!\!/_i$ with the field equation operators $\mathscr{F}$ and $\mathscr{W}$, acting this time on a general 1-form α_a and a general scalar field ψ,

$$\left|[\nabla\!\!\!/, \mathscr{F}]\alpha\right|_{\mathbb{S}^3} \lesssim |\nabla\!\!\!/^2 \boldsymbol{\alpha}| + |\nabla\!\!\!/ \dot{\alpha}_0| + \text{l.o.t.s},$$

and

$$\left|[\nabla\!\!\!/, \mathscr{W}]\psi\right| \lesssim |\psi \nabla\!\!\!/ \dot{A}_0| + |\dot{\psi} \nabla\!\!\!/ A_0| + |\nabla\!\!\!/^2 \psi| + |\psi \nabla\!\!\!/^2 \mathbf{A}| + \text{l.o.t.s},$$

where the lower order terms are terms that are of order one or zero in derivatives of α_a, A_a, or ψ. Now estimate the $(m+1)$-th error term:

$$\begin{aligned}
\mathsf{e}_{m+1} &:= T^b \left(\nabla^a \mathbf{T}_{ab}[\nabla\!\!\!/^m A] + \nabla^a \mathbf{T}_{ab}[\nabla\!\!\!/^m \phi]\right) \\
&= T^b\Big(\mathscr{F}[\nabla\!\!\!/^m A]^a (\nabla_a (\nabla\!\!\!/^m A)_b - \nabla_b (\nabla\!\!\!/^m A)_a) + \mathrm{Re}\left(\overline{\mathscr{W}[\nabla\!\!\!/^m \phi]} \mathrm{D}_b (\nabla\!\!\!/^m \phi)\right) \\
&+ (\nabla_a A_b - \nabla_b A_a)\,\mathrm{Im}\left(\nabla\!\!\!/^m \bar{\phi} \mathrm{D}^a \nabla\!\!\!/^m \phi\right)\Big) \\
&\leqslant \left|\mathscr{F}[\nabla\!\!\!/^m A]^\mu (\nabla\!\!\!/_\mu (\nabla\!\!\!/^m A_0) - \nabla\!\!\!/^m \dot{\mathbf{A}}_\mu)\right| + \left|\mathrm{Re}\left(\overline{\mathscr{W}[\nabla\!\!\!/^m \phi]} \mathrm{D}_0 (\nabla\!\!\!/^m \phi)\right)\right| \\
&+ \left|(\nabla\!\!\!/_\mu A_0 - \dot{\mathbf{A}}_\mu)\,\mathrm{Im}\left(\nabla\!\!\!/^m \bar{\phi} \mathrm{D}\!\!\!/^\mu \nabla\!\!\!/^m \phi\right)\right| \\
&\lesssim \left|\mathscr{F}[\nabla\!\!\!/^m A]\right|_{\mathbb{S}^3} \left[|\nabla\!\!\!/^{m+1} A_0| + |\nabla\!\!\!/^m \dot{\mathbf{A}}|\right] + \left|\mathscr{W}[\nabla\!\!\!/^m \phi]\right| \left[|\nabla\!\!\!/^m \dot{\phi}| + |A_0||\nabla\!\!\!/^m \phi|\right] \\
&+ \left[|\nabla\!\!\!/ A_0| + |\dot{\mathbf{A}}|\right] \left[|\nabla\!\!\!/^m \phi||\nabla\!\!\!/^{m+1} \phi| + |\nabla\!\!\!/^m \phi||\mathbf{A}||\nabla\!\!\!/^m \phi|\right] + \text{l.o.t.s} \\
&\lesssim \left[|\nabla\!\!\!/^{m+1} A_0| + |\nabla\!\!\!/^m \dot{\mathbf{A}}|\right] \left[\left|\nabla\!\!\!/^m \mathscr{F}[A]\right|_{\mathbb{S}^3} + \left|[\nabla\!\!\!/^m, \mathscr{F}]A\right|_{\mathbb{S}^3}\right]
\end{aligned}$$

$$
\begin{aligned}
&+\Big[\big|\nabla^m\mathscr{W}[\phi]\big|+\big|[\nabla^m,\mathscr{W}]\phi\big|\Big]|\nabla^m\dot\phi|+|\nabla^m\phi||\nabla^{m+1}\phi|\Big[|\nabla A_0|+|\dot{\mathbf{A}}|\Big]+\text{l.o.t.s}\\
&\lesssim\Big[|\nabla^{m+1}A_0|+|\nabla^m\dot{\mathbf{A}}|\Big]\Big[\big|\nabla^m(\phi\not{D}\phi)\big|+\big|\nabla^{m-1}[\nabla,\mathscr{F}]A\big|+\big|[\nabla,\mathscr{F}]\nabla^{m-1}A\big|\Big]\\
&+|\nabla^m\dot\phi|\Big[|\nabla^{m-1}[\nabla,\mathscr{W}]\phi|+|[\nabla,\mathscr{W}]\nabla^{m-1}\phi|\Big]+|\nabla^m\phi||\nabla^{m+1}\phi|\Big[|\nabla A_0|+|\dot{\mathbf{A}}|\Big]\\
&+\text{l.o.t.s}\\
&\lesssim\Big[|\nabla^{m+1}A_0|+|\nabla^m\dot{\mathbf{A}}|\Big]\Big[\big|\nabla^m(\phi\nabla\phi+\mathbf{A}\phi^2)\big|+\big|\nabla^{m-1}(\nabla^2\mathbf{A}+\nabla\dot A_0+\text{l.o.t.s})\big|\\
&+|\nabla^{m+1}\mathbf{A}|+|\nabla^m\dot A_0|+\text{l.o.t.s}\Big]\\
&+|\nabla^m\dot\phi|\Big[\big|\nabla^{m-1}(\phi\nabla\dot A_0+\dot\phi\nabla A_0+\nabla^2\phi+\phi\nabla^2\mathbf{A}+\text{l.o.t.s})\big|+|\nabla^{m-1}\phi||\nabla\dot A_0|\\
&+|\nabla^{m-1}\dot\phi||\nabla A_0|+|\nabla^{m+1}\phi|+|\nabla^{m-1}\phi||\nabla^2\mathbf{A}|\Big]\\
&+|\nabla^m\phi||\nabla^{m+1}\phi|\Big[|\nabla A_0|+|\dot{\mathbf{A}}|\Big]+\text{l.o.t.s}\\
&\lesssim\Big[|\nabla^{m+1}A_0|+|\nabla^m\dot{\mathbf{A}}|\Big]\Big[\sum_{k=0}^{m}|\nabla^{m-k}\phi||\nabla^{k+1}\phi|+|\nabla^m(\mathbf{A}\phi^2)|\\
&+|\nabla^{m+1}\mathbf{A}|+|\nabla^m\dot A_0|\Big]\\
&+|\nabla^m\dot\phi|\Big[\sum_{k=0}^{m-1}|\nabla^{m-1-k}\phi||\nabla^{k+1}\dot A_0|+\sum_{k=0}^{m-1}|\nabla^{m-1-k}\dot\phi||\nabla^{k+1}A_0|\\
&+|\nabla^{m+1}\phi|+\sum_{k=0}^{m-1}|\nabla^{m-1-k}\phi||\nabla^{k+2}\mathbf{A}|\Big]\\
&+|\nabla^m\phi||\nabla^{m+1}\phi|\Big[|\nabla A_0|+|\dot{\mathbf{A}}|\Big]+\text{l.o.t.s}\\
&\lesssim_{\mathrm{S}^{1/2}_{m+1}}\Big[|\nabla^{m+1}A_0|+|\nabla^m\dot{\mathbf{A}}|\Big]\Big[|\nabla^m\phi||\nabla\phi|+|\phi||\nabla^{m+1}\phi|+|\phi|^2|\nabla^m\mathbf{A}|\\
&+|\phi||\mathbf{A}||\nabla^m\phi|+|\nabla^{m+1}\mathbf{A}|+|\nabla^m\dot A_0|\Big]\\
&+|\nabla^m\dot\phi|\Big[|\nabla^m\dot A_0|+|\nabla^{m-1}\dot\phi||\nabla A_0|+|\dot\phi||\nabla^m A_0|+|\nabla^{m+1}\phi|+|\nabla^{m+1}\mathbf{A}|\Big]\\
&+|\nabla^m\phi||\nabla^{m+1}\phi|+\text{l.o.t.s}\\
&\lesssim_{\mathrm{S}^{1/2}_{m+1}}\Big[|\nabla^{m+1}A_0|+|\nabla^m\dot{\mathbf{A}}|\Big]\Big[|\nabla^m\phi|+|\nabla^{m+1}\phi|+|\nabla^m\mathbf{A}|+|\nabla^{m+1}\mathbf{A}|+|\nabla^m\dot A_0|\Big]\\
&+|\nabla^m\dot\phi|\Big[|\nabla^m\dot A_0|+|\nabla^{m-1}\dot\phi|+|\nabla^m A_0|+|\nabla^{m+1}\phi|+|\nabla^{m+1}\mathbf{A}|\Big]\\
&+|\nabla^m\phi||\nabla^{m+1}\phi|+\text{l.o.t.s}\\
&\lesssim_{\mathrm{S}^{1/2}_{m+1}}\Big[|\nabla^{m+1}A_0|+|\nabla^m\dot{\mathbf{A}}|\Big]\Big[|\nabla^{m+1}\phi|+|\nabla^{m+1}\mathbf{A}|+|\nabla^m\dot A_0|\Big]\\
&+|\nabla^m\dot\phi|\Big[|\nabla^m\dot A_0|+|\nabla^{m-1}\dot\phi|+|\nabla^m A_0|+|\nabla^{m+1}\phi|+|\nabla^{m+1}\mathbf{A}|\Big]\\
&+|\nabla^m\phi||\nabla^{m+1}\phi|+\text{l.o.t.s}
\end{aligned}
$$

$$\begin{aligned}
&\lesssim_{\mathrm{S}_{m+1}^{1/2}} |\nabla^{m+1} A_0||\nabla^{m+1}\phi| + |\nabla^{m+1} A_0||\nabla^{m+1}\mathbf{A}| + |\nabla^{m+1} A_0||\nabla^{m}\dot{A}_0| \\
&+ |\nabla^{m}\dot{\mathbf{A}}||\nabla^{m+1}\phi| + |\nabla^{m}\dot{\mathbf{A}}||\nabla^{m+1}\mathbf{A}| + |\nabla^{m}\dot{\mathbf{A}}||\nabla^{m}\dot{A}_0| + |\nabla^{m}\dot{\phi}||\nabla^{m}\dot{A}_0| \\
&+ |\nabla^{m}\dot{\phi}||\nabla^{m-1}\dot{\phi}| + |\nabla^{m}\dot{\phi}||\nabla^{m} A_0| + |\nabla^{m}\dot{\phi}||\nabla^{m+1}\phi| + |\nabla^{m}\dot{\phi}||\nabla^{m+1}\mathbf{A}| \\
&+ |\nabla^{m}\phi||\nabla^{m+1}\phi| + \text{l.o.t.s},
\end{aligned}$$

where by $\lesssim_{\mathrm{S}_{m+1}^{1/2}}$ we mean *bounded up to a polynomial in* $\mathrm{S}_{m+1}^{1/2}$. Note also that, like in the estimate of section 4.6.3 where the triplets (m, k, l) sum to at least two, the lower order terms in the above are at least quadratic in the fields so that one can control them by a full power of S_{m+1}. Furthermore, inspecting the leading order terms in the above one sees that, with the exception of $\nabla^m \dot{A}_0$, they are all readily controlled by S_{m+1}:

$$\begin{aligned}
\|\mathsf{e}_{m+1}\|_{L^1} &\lesssim_{\mathrm{S}_{m+1}^{1/2}} \mathrm{S}_{m+1} + \|\nabla^{m+1} A_0 \nabla^m \dot{A}_0\|_{L^1} + \|\nabla^m \dot{\mathbf{A}} \nabla^m \dot{A}_0\|_{L^1} + \|\nabla^m \dot{\phi} \nabla^m \dot{A}_0\|_{L^1} \\
&\lesssim_{\mathrm{S}_{m+1}^{1/2}} \mathrm{S}_{m+1} + \|\dot{A}_0\|_{H^m}^2.
\end{aligned}$$

As in proposition 4.6.3, standard elliptic and wave equation estimates inductively show that for sufficiently small S_m,

$$\|\dot{A}_0\|_{H^m}^2 \lesssim_{\mathrm{S}_m^{1/2}} \mathrm{S}_{m+1}, \tag{4.7.3}$$

so altogether we have

$$\|\mathsf{e}_{m+1}\|_{L^1} \lesssim \mathrm{S}_{m+1} P(\mathrm{S}_{m+1}^{1/2})$$

for some polynomial P.

4.8 Proofs of Main Theorems

4.8.1 Proof of Theorem 4.2.2

The $m = 1$ case is trivial, while for $m = 2$ we have already proved the estimates $\mathrm{S}_m[\phi, A](\tau) \simeq \mathrm{S}_m[\phi, A](0)$ and $\|\dot{A}_0\|_{H^{m-1}}^2(\tau) \lesssim \mathrm{S}_m[\phi, A](\tau)$ for small initial data. We proceed by induction. Suppose the estimates

$$\mathrm{S}_m[\phi, A](\tau) \simeq \mathrm{S}_m[\phi, A](0) \quad \text{and} \quad \|\dot{A}_0\|_{H^{m-1}}^2(\tau) \lesssim \mathrm{S}_m[\phi, A](\tau)$$

hold for some $m \in \mathbb{N}$ provided $\mathrm{S}_m[\phi, A](0)$ is small enough. The second of these is immediate for $m + 1$ by (4.7.3), which then implies (4.7.2). Arguing as in the proof of theorem 4.6.4 and applying lemma A.2.12 then gives (4.7.1). □

4.8.2 Proof of Theorem 4.2.3

We restrict ourselves to the case of $\mathscr{I}^+$, the case of $\mathscr{I}^-$ being analogous. Pick admissible initial data u_0 on Σ such that $\mathrm{S}_m[\phi, \mathbf{A}](\Sigma)$ is sufficiently small. Then $\mathrm{S}_m[\phi, A](\Sigma) < \varepsilon_0$ for some small $\varepsilon_0 > 0$, so by corollary 4.4.2 there exists a solution (ϕ, A_a) in $E_m = \bigcap_{k=0}^m \mathcal{C}_b^k(I; H^{m-k})$ to the system (4.3.6) unique up to trivial gauge transformations such that I contains $[-\pi/2, \pi/2]$. Since the solution (ϕ, A_a) is at least $\mathcal{C}^1$ in τ for $m \geqslant 2$, $u = (\phi, \mathbf{A}, \dot{\phi}, \dot{\mathbf{A}}, A_0)$ has a well-defined restriction to $\mathscr{I}^+$. This defines the future trace operator

$$\begin{aligned} \mathfrak{T}_m^+ : \mathrm{S}^0_{m,\varepsilon_0} &\longrightarrow \mathrm{S}_m^+, \\ u_0 &\longmapsto u^+ = (\phi, \mathbf{A}, \dot{\phi}, \dot{\mathbf{A}}, A_0)|_{\mathscr{I}^+}. \end{aligned}$$

By theorem 4.2.2, whenever ε_0 is sufficiently small we have the estimate

$$\mathrm{S}_m[\phi, A](\mathscr{I}^+) \leqslant C \mathrm{S}_m[\phi, A](\Sigma) \leqslant C\varepsilon_0 =: \varepsilon_1, \tag{4.8.1}$$

so the operator $\mathfrak{T}_m^+$ is bounded. The data u^+ on $\mathscr{I}^+$ has size at most $\varepsilon_1 = C\varepsilon_0$, so, reducing ε_0 if necessary, we can evolve u^+ backwards in time to find data $\tilde{u}_0$ on Σ. But by uniqueness, $u_0 = \tilde{u}_0$. Thus the map $\mathfrak{T}_m^+$ is injective for ε_0 small enough. To be able to invert $\mathfrak{T}_m^+$, we restrict the co-domain of $\mathfrak{T}_m^+$ to its image:

$$\mathfrak{T}_m^+ : \mathrm{S}^0_{m,\varepsilon_0} \longrightarrow \mathfrak{T}_m^+(\mathrm{S}^0_{m,\varepsilon_0}) =: \mathscr{D}^+_{m,\varepsilon_1}.$$

By definition, $\mathfrak{T}_m^+$ is now surjective and so bijective, and from the estimate (4.8.1) it is clear that $\mathscr{D}^+_{m,\varepsilon_1} \subset \mathrm{S}^+_{m,\varepsilon_1}$. The operator $\mathfrak{T}_m^+$ is thus invertible and satisfies the bounds

$$\|\mathfrak{T}_m^+ u_0\|^2_{\mathrm{S}_m} \lesssim \|u_0\|^2_{\mathrm{S}_m} \quad \text{and} \quad \|(\mathfrak{T}_m^+)^{-1} u^+\|^2_{\mathrm{S}_m} \lesssim \|u^+\|^2_{\mathrm{S}_m}$$

for $u_0 \in \mathrm{S}^0_{m,\varepsilon_0}$, $u^+ \in \mathscr{D}^+_{m,\varepsilon_1}$. Furthermore, the set $\mathscr{D}^+_{m,\varepsilon_1}$ contains a small ball around the origin in S_m^+. Indeed, if $v^+ \in \mathrm{S}_m^+$ has small enough norm, say $\|v^+\|^2_{\mathrm{S}_m} < \delta \ll \varepsilon_0$, then $\|(\mathfrak{T}_m^+)^{-1} v^+\|^2_{\mathrm{S}_m} \leqslant C\|v^+\|^2_{\mathrm{S}_m} < C\delta < \varepsilon_0$, and so $(\mathfrak{T}_m^+)^{-1} v^+ \in \mathrm{S}^0_{m,\varepsilon_0}$.

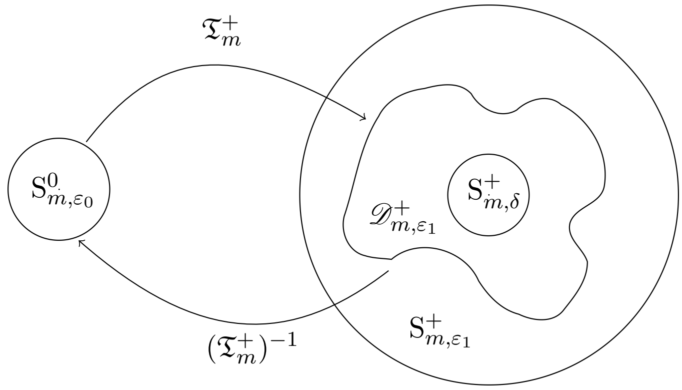

Figure 4.2: The image of a small ball under the future trace operator $\mathfrak{T}_m^+$.

Constructing the scattering operator is now simply a matter of composing the inverse past trace operator and the future trace operator. We define

$$\begin{aligned} \mathscr{S}_m : \mathscr{D}^-_{m,\varepsilon_1} &\longrightarrow \mathscr{D}^+_{m,\varepsilon_1}, \\ \mathscr{S}_m &:= \mathfrak{T}_m^+ \circ (\mathfrak{T}_m^-)^{-1}. \end{aligned}$$

Then $\mathscr{S}_m$ is invertible with inverse $\mathscr{S}_m^{-1} = \mathfrak{T}_m^- \circ (\mathfrak{T}_m^+)^{-1}$, and the estimates

$$\|u^+\|^2_{\mathrm{S}_m} \simeq \|u^-\|^2_{\mathrm{S}_m}$$

for $u^\pm \in \mathscr{D}^\pm_{m,\varepsilon_1}$ follow from the estimates for $\mathfrak{T}^\pm_m$. □

Remark 4.8.1. It is not immediately clear what the set $\mathscr{D}^+_{m,\varepsilon_1}$ looks like, for two reasons. Firstly, the sets $\mathrm{S}^{\pm,0}_m$ are not vector spaces since admissible initial data is not additive. Secondly, the fact that $\mathfrak{T}^+_m$ is a nonlinear operator precludes any straightforward application of the open mapping theorem, so it is not even obvious that $\mathscr{D}^+_{m,\varepsilon_1}$ is open and connected. Nonetheless, by symmetry it is clear that the set of past asymptotic data $\mathscr{D}^-_{m,\varepsilon_1}$ and the set of future asymptotic data $\mathscr{D}^+_{m,\varepsilon_1}$ are of the same 'size' in the sense that they are contained in balls of the same radius in S^-_m and S^+_m respectively.

Remark 4.8.2. The lack of vector space structure on the domains of definition of the operators $\mathfrak{T}^\pm_m$ and $\mathscr{S}_m$ makes it difficult to discuss their regularity beyond boundedness. This lack of vector space structure stems, most importantly, from the constraint equations in the system (4.3.6). It is fairly easy to see that any extension of e.g. $\mathscr{S}_m$ off the constraint surface that preserves boundedness will automatically be continuous at the zero solution, but continuity at more general solutions will require a more careful analysis of (4.3.6) linearized around a general solution, as well as a choice of extension. Differentiability will pose further complications.

4.8.3 Proof of Theorem 4.2.4

Suppose $\mathrm{S}_m[\tilde{\phi}, \tilde{\mathbf{A}}](\alpha = 0)$ is small. We derive the asymptotics for $\mathscr{I}^+$, the ones for $\mathscr{I}^-$ being analogous. By (4.2.3), $\mathrm{S}_m[\phi, \mathbf{A}](\tau = 0)$ is small too, and the elliptic estimates for A_0 imply that the full Sobolev energy $\mathrm{S}_m[\phi, A](\tau = 0)$ is small. Then according to our estimates and Sobolev embedding, ϕ, $\mathbf{A}$ and A_0 are continuous on all of $\widehat{\mathrm{dS}}_4$ with a $\mathcal{C}^{m-2}$ trace on $\mathscr{I}^+$.

Let $m = 2$. Then $\phi = \Omega^{-1}\tilde{\phi}$ has a continuous limit on $\mathscr{I}^+$, so

$$|\tilde{\phi}| \lesssim \Omega \lesssim \frac{1}{\cosh(H\alpha)} \lesssim \mathrm{e}^{-H\alpha}$$

as $\alpha \to +\infty$. The timelike component of A_a is $A_0 = \partial^a_\tau A_a = H^{-1}\cosh(H\alpha)\partial^a_\alpha \tilde{A}_a = H^{-1}\cosh(H\alpha)\tilde{A}_\alpha$ and has a continuous limit on $\mathscr{I}^+$, so similarly

$$|\tilde{A}_\alpha| \lesssim \mathrm{e}^{-H\alpha}$$

as $\alpha \to +\infty$. Finally the $\mathbb{S}^3$ components of A_a are

$$|\mathbf{A}|^2_{\mathfrak{s}_3} = -\mathfrak{e}^{ab}\mathbf{A}_a\mathbf{A}_b = -\Omega^{-2}\tilde{g}^{ab}\tilde{\mathbf{A}}_a\tilde{\mathbf{A}}_b = \Omega^{-2}\frac{H^2}{\cosh^2(H\alpha)}\mathfrak{s}_3^{\mu\nu}\tilde{\mathbf{A}}_\mu\tilde{\mathbf{A}}_\nu = |\tilde{\mathbf{A}}|^2_{\mathfrak{s}_3},$$

so $|\tilde{\mathbf{A}}|_{\mathfrak{s}_3} \lesssim 1$.

Next let us work in the static coordinates (4.1.2). These coordinates are only appropriate in region I of fig. 2.6 since they become singular on the horizons $r = 1/H$, and ∂_t becomes spacelike in regions II and IV and past-pointing in region III. Following the flow of the vector field ∂_t in region I, one is forced to the top right corner of fig. 4.1 as

$t \to +\infty$. A preferred point on $\mathscr{I}^+$ has therefore been singled out for an observer following the flow of ∂_t; this point is the timelike infinity i^+ for observers living in region I of fig. 2.6.

In these coordinates the conformal factor Ω is given by

$$\Omega = \frac{H}{\cosh(Ht)} \frac{1}{\sqrt{F_t(r)}},$$

where $F_t(r) = 1 - \tanh^2(Ht) H^2 r^2$. Keeping r fixed, for the scalar field we then have

$$|\tilde{\phi}| \lesssim \Omega \lesssim_r \mathrm{e}^{-Ht}$$

as $t \to +\infty$. For the Maxwell potential, we find the relations

$$\tilde{A}_t = H^2 \operatorname{sech}^2(Ht) F_t(r)^{-1} \left(-rF(r)^{1/2} \sinh(Ht) A_\zeta + H^{-1} F(r)^{1/2} \cosh(Ht) A_\tau \right),$$
$$\tilde{A}_r = H^2 \operatorname{sech}^2(Ht) F_t(r)^{-1} \left(H^{-1} F(r)^{-1/2} \cosh(Ht) A_\zeta - rF(r)^{-1/2} \sinh(Ht) A_\tau \right).$$

Since A_ζ and A_τ have continuous limits as $t \to +\infty$ for fixed r, we have

$$|\tilde{A}_t| \lesssim_r \mathrm{e}^{-Ht} \quad \text{and} \quad |\tilde{A}_r| \lesssim_r \mathrm{e}^{-Ht}.$$

Expanding the 3-sphere norm $|\tilde{\mathbf{A}}|^2_{\mathfrak{s}_3}$,

$$|\tilde{\mathbf{A}}|^2_{\mathfrak{s}_3} = \tilde{A}_\zeta^2 + \frac{1}{\sin^2\zeta} |\tilde{A}|^2_{\mathfrak{s}_2} \lesssim 1,$$

we see that $|\tilde{A}|_{\mathfrak{s}_2} \lesssim \sin\zeta$, where one computes $\sin\zeta = \operatorname{sech}(Ht) Hr F_t(r)^{-1/2}$. Thus

$$\frac{1}{r} |\tilde{A}|_{\mathfrak{s}_2} \lesssim_r \mathrm{e}^{-Ht}$$

as $t \to +\infty$.

Now suppose $m = 3$. This in particular means that

$$|\nabla\!\!\!\!/ \phi|^2 = (\partial_\zeta \phi)^2 + \frac{1}{\sin^2\zeta} |\nabla_{\mathbb{S}^2} \phi|^2$$

has a continuous limit on $\mathscr{I}^+$, and so $\partial_\zeta\phi$ and $(\sin\zeta)^{-1} |\nabla_{\mathbb{S}^2}\phi|$ do too. Since ϕ scales conformally as $\phi = \Omega^{-1}\tilde{\phi}$, one computes

$$\begin{aligned} \partial_\zeta \phi = H^{-1} \cosh(Ht) F_t(r)^{1/2} \Big(& rF(r)^{-1/2} \sinh(Ht) \partial_t \tilde{\phi} \\ & + H^{-1} F(r)^{1/2} \cosh(Ht) \partial_r \tilde{\phi} \Big) \end{aligned} \tag{4.8.2}$$

and

$$\begin{aligned} \partial_\tau \phi + (\partial_\tau \Omega) \Omega^{-1} \phi = H^{-1} \cosh(Ht) F_t(r)^{1/2} \Big(& H^{-1} F(r)^{-1/2} \cosh(Ht) \partial_t \tilde{\phi} \\ & + rF(r)^{1/2} \sinh(Ht) \partial_r \tilde{\phi} \Big). \end{aligned} \tag{4.8.3}$$

Because $\Omega \partial_\zeta \phi$ and $\Omega \partial_\tau \phi + (\partial_\tau \Omega)\phi$ have continuous limits on $\mathscr{I}^+$, one therefore sees that

$$|\partial_t \tilde{\phi}| \lesssim_r \mathrm{e}^{-Ht} \quad \text{and} \quad |\partial_r \tilde{\phi}| \lesssim_r \mathrm{e}^{-Ht}$$

as $t \to +\infty$. For the $\mathbb{S}^2$ derivatives, the fact that $(\sin\zeta)^{-1}|\nabla_{\mathbb{S}^2}\phi| = \Omega^{-1}(\sin\zeta)^{-1}|\nabla_{\mathbb{S}^2}\tilde{\phi}|$ has a continuous limit on $\mathscr{I}^+$ implies that

$$\left|\frac{1}{r}\nabla_{\mathbb{S}^2}\tilde{\phi}\right| \lesssim_r \mathrm{e}^{-2Ht} \tag{4.8.4}$$

as $t \to +\infty$. Let us study the e^{-Ht} component of $\tilde{\phi}$,

$$\tilde{\varphi} := \mathrm{e}^{Ht}\tilde{\phi}.$$

Rewriting (4.8.2) and (4.8.3) in terms of $\tilde{\varphi}$, one has

$$\mathcal{O}\left(\mathrm{e}^{-Ht}\right) = rF(r)^{-1/2}\sinh(Ht)\mathrm{e}^{-Ht}\left(\partial_t\tilde{\varphi} - H\tilde{\varphi}\right) + H^{-1}F(r)^{1/2}\cosh(Ht)\mathrm{e}^{-Ht}\partial_r\tilde{\varphi}$$

and

$$\begin{aligned}\mathcal{O}\left(\mathrm{e}^{-Ht}\right) - F(r)^{1/2}\sinh(Ht)\mathrm{e}^{-Ht}\tilde{\varphi} = {} & H^{-1}F(r)^{-1/2}\cosh(Ht)\mathrm{e}^{-Ht}\left(\partial_t\tilde{\varphi} - H\tilde{\varphi}\right) \\ & + rF(r)^{1/2}\sinh(Ht)\mathrm{e}^{-Ht}\partial_r\tilde{\varphi},\end{aligned}$$

which, taking the limit $t \to +\infty$, become

$$\begin{aligned} 0 &\mathrel{\widehat{=}} Hr\partial_t\tilde{\varphi} - H^2 r\tilde{\varphi} + F(r)\partial_r\tilde{\varphi}, \\ -HF(r)\tilde{\varphi} &\mathrel{\widehat{=}} \partial_t\tilde{\varphi} - H\tilde{\varphi} + HrF(r)\partial_r\tilde{\varphi}, \end{aligned}$$

where $\mathrel{\widehat{=}}$ denotes equality at $t = +\infty$. Solving these algebraically shows that $\partial_t\tilde{\varphi} \mathrel{\widehat{=}} 0$ and

$$H^2 r\tilde{\varphi} \mathrel{\widehat{=}} F(r)\partial_r\tilde{\varphi}.$$

Further, the bound (4.8.4) shows that at $t = +\infty$ the function $\tilde{\varphi}$ is independent of the $\mathbb{S}^2$ coordinates, so the above equation is an ODE in r, with solution

$$\tilde{\varphi}(r) \mathrel{\widehat{=}} \frac{1}{\sqrt{F(r)}}\tilde{\varphi}(0).$$

We conclude that there exists a constant c such that

$$\tilde{\phi} \sim cF(r)^{-1/2}\mathrm{e}^{-Ht} + \mathcal{O}\left(\mathrm{e}^{-2Ht}\right) \quad \text{as} \quad t \to +\infty.$$

One can check by hand that $\tilde{\Phi}_1(t,r) = F(r)^{-1/2}\mathrm{e}^{-Ht}$ is the eigenfunction of the uncharged ($A_a \equiv 0$) spherically symmetric conformally invariant wave operator

$$\tilde{\Box} + \frac{1}{6}\tilde{\mathrm{R}} = F(r)^{-1}\partial_t^2 - \frac{1}{r}\partial_r\left(rF(r)\partial_r\right) + 2H^2$$

with eigenvalue H^2. $\square$

5

Large Data Decay of Yang–Mills–Higgs Fields

5.1 Introduction

The analytical study of the full classical Yang–Mills–Higgs equations goes back to at least the late 1970s, with Segal's local existence proof [107, 108] on Minkowski space of pure SU(2) Yang–Mills fields with initial data $(\mathbf{A}_i, \partial_t \mathbf{A}_i)$ in $H^3 \times H^2$. A short time after Segal's proof, in 1981, Ginibre and Velo [54], Choquet-Bruhat and Christodoulou [23], and Eardley and Moncrief [36, 37] all published proofs of similarly major results, though using profoundly different techniques. Ginibre and Velo's work extended Segal's work to coupled Yang–Mills–Higgs equations in arbitrary dimension, in particular proving global existence in two and three spacetime dimensions. In four dimensions, Choquet-Bruhat and Christodoulou made use of the conformal invariance of the Yang–Mills–Higgs–Dirac equations and a short-time existence theorem on the Einstein cylinder to prove the global existence of solutions on Minkowski space for sufficiently small $H^2 \times H^1$ initial data[1] (cf. [21]). Eardley and Moncrief, on the other hand, instead developed a physical space technique for extracting remarkable a priori estimates that allowed them to prove the global existence of solutions for *large* $H^2 \times H^1$ initial data. A short time later, Goganov and Kapitanskii published a proof of global unique solvability [55] for only locally $H^2 \times H^1$ data on Minkowski space, in particular allowing arbitrary magnetic charge at spatial infinity. Their proof shows that the equations are well-posed in local lightcones, with solutions determined only by the data at the base of the lightcone. Further improvements have been obtained by Klainerman and Machedon [64, 65, 66] and others [92, 109, 115].

The strategy of Eardley and Moncrief is to write down wave equations for the fields F_{ab} and $\mathrm{D}_a \phi$, treat the nonlinear terms in these equations as sources, and express F_{ab}

[1]Here and throughout the rest of the chapter, we refer to the regularity of the initial Yang–Mills potential and its time derivative, $(\mathbf{A}_i, \partial_t \mathbf{A}_i) \in H^{k+1} \times H^k$. For the coupled Yang–Mills–Higgs system we shall also have the initial data for the scalar field, $(\mathbf{A}_i, \partial_t \mathbf{A}_i, \phi, \partial_t \phi) \in H^{k+1} \times H^k \times H^{k+1} \times H^k$, which we will sometimes abbreviate as simply $H^{k+1} \times H^k$ data. For the Yang–Mills–Higgs–Dirac equations one also has the initial Dirac field ψ, $(\mathbf{A}_i, \partial_t \mathbf{A}_i, \phi, \partial_t \phi, \psi) \in H^{k+1} \times H^k \times H^{k+1} \times H^k \times H^{k+1}$.

and $\mathrm{D}_a\phi$ at a point p as integrals over the backward lightcone of p. Their key observation is that these lightcone integrals can be estimated by expressions of the form

$$E_0 \int_0^t (\|F(s)\|_{L^\infty} + \|\mathrm{D}\phi(s)\|_{L^\infty})\, \mathrm{d}s,$$

which implies, via Grönwall's inequality, that the L^∞ norms cannot blow up in finite time. Part of the trick is to define the L^∞ norms in a gauge-independent manner, and use the Crönstrom gauge in intermediate calculations. Equipped with this estimate, it is then straightforward to show that the $H^2 \times H^1$ norm of the solution does not blow up in finite time. An incarnation of this method has been adapted, for pure Yang–Mills equations, to arbitrary smooth globally hyperbolic four dimensional spacetimes by Chruściel and Shatah [28], by replacing the lightcone integrals with Friedlander's representation formula [46] for the covariant wave equation. However, Chruściel and Shatah require effectively $H^3 \times H^2$ data to deal with a term that causes difficulties in curved space[2]. Though the system has been well-studied, Eardley and Moncrief's method with $H^2 \times H^1$ data for coupled Yang–Mills and Higgs equations does not seem to have been explicitly adapted to curved space, even in the case of the Einstein cylinder. The scalar field part scales differently under a conformal transformation, putting it on unequal footing with the Yang–Mills potential. In particular, this upsets the conformal invariance of the system somewhat, breaking the covariance of the canonical energy-momentum tensor. And although formally the field equations remain conformally invariant, the scalar field introduces a boundary term in the conformal variation of the action that has a non-trivial dependence on the decay of the scalar field. This is expected to be of some importance in path integral formulations of interacting quantum field theories.

In this chapter we extend the L^∞ estimates of Eardley and Moncrief to the Einstein cylinder. Our method is inspired by and combines the techniques of [23, 36, 55]: we first work on Minkowski space and localize Eardley and Moncrief's estimates, removing the requirement of the global finiteness of the energy. Then, using a conformal transformation, we glue a small conical patch of Minkowski space onto the Einstein cylinder, and show that L^∞ estimates in the Minkowskian patch imply local L^∞ estimates on the cylinder. By patching a finite number of such cones all the way around the Einstein cylinder, we deduce L^∞ bounds on any finite section of the cylinder. This allows us to show that Choquet-Bruhat and Christodoulou's small data result on the Einstein cylinder [23] extends to large data, and consequently removes the small data restriction in the scattering theory of chapter 4. Finally, by using an inverse conformal transformation, we deduce large data decay rates for Yang–Mills–Higgs fields on Minkowski and de Sitter spacetimes.

The structure of this chapter is as follows. In the rest of this section we outline the notation used in this chapter. In section 5.2 we sketch the method of Eardley and Moncrief and show how their estimates can be localized. In sections 5.3 and 5.4 we glue the Minkowskian L^∞ estimates onto the Einstein cylinder and use them deduce the global existence of Yang–Mills–Higgs fields on $\mathbb{R} \times \mathbb{S}^3$. Finally, in section 5.5 we deduce decay rates for the fields on Minkowski space and de Sitter space.

[2] Eardley and Moncrief's result has been re-proven by Klainerman and Rodnianski by applying their newly developed Kirchoff–Sobolev parametrix for the wave equation [70]. A similar method has since been used by Ghanem [52] to give another proof of the a priori estimates, for $H^2 \times H^1$ data, for pure Yang–Mills on curved spacetimes.

5.1.1 Notation

On any globally hyperbolic spacetime $(\mathscr{M}, g)$ we can choose a global smooth time function t such that $\nabla^a t$ is uniformly timelike on $\mathscr{M}$ and the metric g_{ab} takes the form

$$g_{ab} = T_a T_b - h_{ab}, \quad \text{i.e.} \quad g = N^2 \mathrm{d}t^2 - h,$$

where T^a is a smooth future-oriented uniformly timelike vector field with lapse function $N(t)$, $T^a = N^{-1}\partial_t$, and h_{ab} is a smooth Riemannian metric for each fixed t. The vector field T^a defines a foliation of $\mathscr{M}$ by hypersurfaces Σ_t of constant t, and identifies $\mathscr{M} = \mathbb{R}\times\Sigma$, where each Σ_t is diffeomorphic to Σ. On the hypersurfaces Σ_t, we define Sobolev spaces with respect to the Riemannian metric $h(t)$. To be able to work with Sobolev spaces in spacetime, we define the four dimensional Riemannian metric

$$\Gamma_{ab} \coloneqq 2T_a T_b - g_{ab} = T_a T_b + h_{ab},$$

and define Sobolev norms on general subsets of $\mathscr{M}$ with respect to Γ. For example, for a matrix-valued 2-form $F_{ab} = F^{\alpha}_{ab}\theta_\alpha$ on $\mathscr{M}$ we set

$$|F|^2_\Gamma \coloneqq \sum_\alpha F^\alpha_{ab} F^\alpha_{cd} \Gamma^{ac}\Gamma^{bd},$$

and define

$$\|F\|_{L^\infty(K)} \coloneqq \sup_K |F|_\Gamma$$

for any $K \subset \mathscr{M}$.

In this chapter we denote Minkowski space $(\mathbb{R}^4, \eta)$ by $\mathbb{M}$, and specifically work on three conformally related spacetimes of the above form: Minkowski space, the Einstein cylinder $(\mathfrak{E}, \mathfrak{e})$, and de Sitter space $(\mathrm{dS}_4 = \mathbb{R}\times\mathbb{S}^3, \tilde{g})$, where $\tilde{g} = \mathrm{d}\alpha^2 - (\cosh^2\alpha)\mathfrak{s}_3$. Unless stated otherwise, we will denote the Levi–Civita connection on $\mathbb{M}$ by ∇_a, the Levi–Civita connection on $\mathfrak{E}$ by $\hat{\nabla}_a$, and the Levi–Civita connection on dS_4 by $\tilde{\nabla}_a$. We also denote the Levi–Civita connection on $\mathbb{R}^3$ by $\boldsymbol{\nabla}$ and the Levi–Civita connection on $\mathbb{S}^3$ by $\not{\nabla}$. In each of the three spacetimes one has a standard uniformly timelike vector field: ∂_t in $\mathbb{M}$, ∂_τ in $\mathfrak{E}$, and ∂_α in dS_4. We shall use these to define foliations of $\mathbb{M}$, $\mathfrak{E}$ and dS_4, as described above. Given a solution (ϕ, A_a) to the Yang–Mills–Higgs equations on Minkowski space, we will denote the corresponding conformally related solution on the Einstein cylinder by $(\hat{\phi}, \hat{A}_a)$, and the corresponding solution on de Sitter space by $(\tilde{\phi}, \tilde{A}_a)$. The timelike components (corresponding to the time coordinate in each spacetime) of the Yang–Mills potential will be denoted with the index 0, *i.e.* $A_0 = (\partial_t)^a A_a$, $\hat{A}_0 = (\partial_\tau)^a \hat{A}_a$, and $\tilde{A}_0 = (\partial_\alpha)^a \tilde{A}_a$. We will denote by $\mathbf{A}$ (or $\hat{\mathbf{A}}$, or $\tilde{\mathbf{A}}$) the projection of A_a onto the spacelike slice Σ_t (or Σ_τ, or Σ_α respectively), and define the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ on $\mathbb{M}$ by

$$\mathbf{E}_i = F_{0i}, \quad \text{and} \quad \mathbf{B}^i = \frac{1}{2}\varepsilon^{ijk}F_{jk}.$$

The electric and magnetic fields on $\mathfrak{E}$ and dS_4, denoted $\hat{\mathbf{E}}$ and $\hat{\mathbf{B}}$, and $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{B}}$ respectively, are defined similarly. Here the Roman indices i, j, k run over $\{1,2,3\}$ and denote contractions with the spatial basis vectors $\partial_i = \partial/\partial x^i$, $i = 1,2,3$. We also define

$$\pi = \mathrm{D}_0\phi,$$

where $\mathrm{D}_a\phi = \nabla_a\phi + A_a\phi$, and define $\hat{\pi}$ and $\tilde{\pi}$ analogously. In intermediate calculations we shall want to manipulate the components of the Yang–Mills field F_{ab} relative to a null tetrad (l, n, e_A), $A \in \{\theta, \phi\}$, and will denote by $F_{ln} = l^a F_{ab} n^b$, $F_{lA} = l^a F_{ab}(e_A)^b$, and so on.

Finally, in the analysis we shall use the letter $p(t)$ to denote an arbitrary positive 'generalized' polynomial in t perhaps containing positive fractional powers of t.

5.2 Localized L^∞ Estimates on Minkowski Space

On Minkowski space the field equations (2.2.27) simplify to

$$\mathrm{D}^b F_{ab} = -((\mathrm{D}_a\phi)\cdot\theta_\alpha\phi)\theta_\alpha \quad \text{and} \quad \mathrm{D}^a\mathrm{D}_a\phi + \lambda_0|\phi|^2\phi = 0. \tag{5.2.1}$$

In the temporal gauge $A_0 = 0$ they further split into

$$\dot{\mathbf{E}}_i + \boldsymbol{\nabla}_j F_{ij} + [\mathbf{A}_j, F_{ij}] = ((\mathbf{D}_i\phi)\cdot\theta_\alpha\phi)\theta_\alpha, \qquad \dot{\pi} - \mathbf{D}_i\mathbf{D}_i\phi + \lambda_0|\phi|^2\phi = 0, \tag{5.2.2}$$

and the constraint equation

$$\boldsymbol{\nabla}\cdot\mathbf{E} + [\mathbf{A}_i, \mathbf{E}_i] = (\pi\cdot\theta_\alpha\phi)\theta_\alpha, \tag{5.2.3}$$

where $\mathbf{D}_i\phi$ is $\mathrm{D}_a\phi$ projected to Σ_t. Of course, the constraint (5.2.3) is propagated in the sense that it is satisfied for all time if it is satisfied initially. We will ultimately consider the system (5.2.2)–(5.2.3), but shall use the Cronström gauge to derive the intermediate a priori L^∞ estimates.

By differentiating the Bianchi identity (2.2.28) and using the field equations (5.2.1), one derives a wave equation for the curvature F_{ab}, which turns out to be

$$\begin{aligned}\Box F_{ab} = {}& ((F_{ab}\phi)\cdot\theta_\alpha\phi)\,\theta_\alpha + ((\mathrm{D}_b\phi)\cdot\theta_\alpha(\mathrm{D}_a\phi) - (\mathrm{D}_a\phi)\cdot\theta_\alpha(\mathrm{D}_b\phi))\,\theta_\alpha \\ & - 2\nabla^c([A_c, F_{ab}]) + [\nabla_c A^c, F_{ab}] - [A^c,[A_c,F_{ab}]] - 2[F_b{}^c, F_{ac}].\end{aligned} \tag{5.2.4}$$

By differentiating the wave equation for ϕ and using the field equation for F_{ab}, one also derives

$$\mathrm{D}^a\mathrm{D}_a(\mathrm{D}_b\phi) = ((\mathrm{D}_b\phi)\cdot\theta_\alpha\phi)\,\theta_\alpha\phi - 2F_b{}^a\mathrm{D}_a\phi - \lambda_0\mathrm{D}_b(|\phi|^2\phi),$$

which can be written as

$$\begin{aligned}\Box(\mathrm{D}_b\phi) = {}& -2\nabla^a(A_a\mathrm{D}_b\phi) + (\nabla^a A_a)\mathrm{D}_b\phi - A^a A_a \mathrm{D}_b\phi \\ & + ((\mathrm{D}_b\phi)\cdot\theta_\alpha\phi)\,\theta_\alpha\phi - 2F_b{}^a\mathrm{D}_a\phi - \lambda_0\mathrm{D}_b(|\phi|^2\phi).\end{aligned} \tag{5.2.5}$$

Here $\Box$ denotes the standard wave operator on Minkowski space. It is worth observing that temporal gauge initial data $(\mathbf{A}, \mathbf{E}, \phi, \pi)$ for the equations (5.2.2) defines initial data for the wave equations (5.2.4) and (5.2.5). Indeed, the data for F_{ab} is given by

$$\begin{aligned}F_{0i}|_{t=0} &= \mathbf{E}_i, \\ \partial_t F_{0i}|_{t=0} &= -\boldsymbol{\nabla}_j F_{ij} - [\mathbf{A}_j, F_{ij}] + ((\mathbf{D}_i\phi)\cdot\theta_\alpha\phi)\,\theta_\alpha, \\ F_{ij}|_{t=0} &= \boldsymbol{\nabla}_i\mathbf{A}_j - \boldsymbol{\nabla}_j\mathbf{A}_i + [\mathbf{A}_i, \mathbf{A}_j], \\ \partial_t F_{ij}|_{t=0} &= \boldsymbol{\nabla}_i\mathbf{E}_j - \boldsymbol{\nabla}_j\mathbf{E}_i + [\mathbf{E}_i, \mathbf{A}_j] + [\mathbf{A}_i, \mathbf{E}_j],\end{aligned}$$

while data for $\mathrm{D}_a\phi$ is given by

$$\begin{aligned}\mathrm{D}_0\phi|_{t=0} &= \pi, \\ \partial_t(\mathrm{D}_0\phi)|_{t=0} &= \mathbf{D}_i\mathbf{D}_i\phi - \lambda_0|\phi|^2\phi, \\ \mathbf{D}_i\phi|_{t=0} &= \boldsymbol{\nabla}_i\phi + \mathbf{A}_i\phi, \\ \partial_t(\mathbf{D}_i\phi)|_{t=0} &= \boldsymbol{\nabla}_i\pi + \mathbf{E}_i\phi + \mathbf{A}_i\pi.\end{aligned}$$

We will use the wave equations (5.2.4) and (5.2.5) to write down the crucial integral expressions for F_{ab} and $\mathrm{D}_a\phi$. Before we do that, however, we need a couple of preliminary tools.

5.2.1 Conservation of Energy

In standard spherical coordinates on Minkowski space, the vector field ∂_t is a global uniformly timelike Killing field. Furthermore, the stress-energy tensor (2.2.30) is conserved, and reads

$$\mathbf{T}_{ab} = -\langle F_{ac}, F_b{}^c\rangle + \frac{1}{4}\eta_{ab}\langle F_{cd}, F^{cd}\rangle + (\mathrm{D}_a\phi)\cdot(\mathrm{D}_b\phi) - \frac{1}{2}\eta_{ab}(\mathrm{D}_c\phi)\cdot(\mathrm{D}^c\phi) + \frac{1}{4}\lambda_0\eta_{ab}|\phi|^4.$$

Contracting $\mathbf{T}_{ab}$ with the Killing field $(\partial_t)^b$ defines a conserved current whose timelike component is

$$\mathbf{T}_{00} = \frac{1}{2}\langle \mathbf{E}_i, \mathbf{E}_i\rangle + \frac{1}{2}\langle \mathbf{B}_i, \mathbf{B}_i\rangle + \frac{1}{2}\pi\cdot\pi + \frac{1}{2}(\mathbf{D}_i\phi)\cdot(\mathbf{D}_i\phi) + \frac{1}{4}\lambda_0|\phi|^4.$$

It follows that the energy

$$E_0(t) = \frac{1}{2}\int_{\mathbb{R}^3}\left(|\mathbf{E}|^2 + |\mathbf{B}|^2 + |\pi|^2 + |\mathbf{D}\phi|^2 + \frac{1}{2}\lambda_0|\phi|^4\right)\mathrm{d}^3x$$

is conserved, where $|\mathbf{E}|^2 = \langle \mathbf{E}_i, \mathbf{E}_i\rangle$, and so on. More generally, one may contract $\mathbf{T}_{ab}$ with any timelike Killing field K^a to get a conserved current

$$J_a \coloneqq \mathbf{T}_{ab}K^b,$$

and derive energy identities by integrating the identity $\nabla_a J^a = 0$ over bounded regions of spacetime. We will do so shortly to derive an energy identity on a lightcone. To do this, we equip ourselves with the following basis of vector fields,

$$l^a = -\partial_t + \partial_r, \qquad n^a = \partial_t + \partial_r, \qquad e^a_\theta = \frac{1}{r}\partial_\theta, \qquad e^a_\phi = \frac{1}{r\sin\theta}\partial_\phi.$$

The vector fields (l, n, e_A), $A \in \{\theta, \phi\}$, satisfy

$$l_a l^a = 0 = n_a n^a, \qquad l_a n^a = -2, \qquad (e_A)_a(e_B)^a = -\delta_{AB},$$

and the Minkowski metric can be written in terms of the basis (l^a, n^a, e^a_A) as

$$\eta_{ab} = -\frac{1}{2}(l_a n_b + l_b n_a) + (e_A)_a(e_A)_b,$$

where the index A is summed over $\{\theta, \phi\}$. Similarly, the spacetime volume form can be written as

$$\mathrm{d}t\wedge\mathrm{d}^3x = \frac{1}{2}l^\flat\wedge n^\flat\wedge e^\flat_\theta\wedge e^\flat_\phi.$$

Putting $K^a = \partial_t$ and integrating $\nabla_a J^a = 0$ over the region bounded by the past lightcone of the origin $K = \{t = -r\}$ and the surface $\Sigma = \{t = -t_0\}$, $t_0 > 0$, we get

$$\frac{1}{2}\int_{B(r_0)}\left(|\mathbf{E}|^2 + |\mathbf{B}|^2 + |\pi|^2 + |\mathbf{D}\phi|^2 + \frac{1}{2}\lambda_0|\phi|^4\right)\mathrm{d}^3x = -\int_{K(t_0)}(J_a l^a)\Big|_{t=-r} r^2\,\mathrm{d}r\,\mathrm{dv}_{\mathfrak{s}_2},$$

where $K(t_0)$ is the past lightcone of the origin up to $t = -t_0$, and $B(r_0)$ is the solid ball in Σ of radius $r_0 = t_0$. Expressing $K^a = \frac{1}{2}(n^a - l^a)$, we have

$$\begin{aligned}&\frac{1}{2}\int_{B(r_0)} \left(|\mathbf{E}|^2 + |\mathbf{B}|^2 + |\pi|^2 + |\mathbf{D}\phi|^2 + \frac{1}{2}\lambda_0|\phi|^4\right) \mathrm{d}^3x \\ &= \frac{1}{2}\int_{K(t_0)} \left(\frac{1}{4}|F_{ln}|^2 + |F_{lA}|^2 + \frac{1}{2}|F_{AB}|^2 + |\mathrm{D}_l\phi|^2 + |\mathrm{D}_A\phi|^2 + \frac{1}{2}\lambda_0|\phi|^4\right)\Bigg|_{t=-r} r^2\,\mathrm{d}r\,\mathrm{dv}_{\mathbb{S}^2}\,.\end{aligned} \tag{5.2.6}$$

Let us denote the left-hand side of the energy identity (5.2.6), the energy in $B(r_0)$ at time $-t_0$, by $E_{B(r_0)}(-t_0)$.

Definition 5.2.1. We define the *local energy* $E_{\mathrm{loc}}(p)$ of a point $p = (t, x)$ by

$$\begin{aligned}E_{\mathrm{loc}}(p) &:= \sup_{s\in[0,t]} \frac{1}{2}\int_{B(x,t-s)} \left(|\mathbf{E}|^2 + |\mathbf{B}|^2 + |\pi|^2 + |\mathbf{D}\phi|^2 + \frac{1}{2}\lambda_0|\phi|^4\right) \mathrm{d}^3x(s) \\ &= \sup_{s\in[0,t]} E_{B(x,t-s)}(s),\end{aligned}$$

where $B(x, r)$ is the ball of radius r centred at $x \in \mathbb{R}^3$.

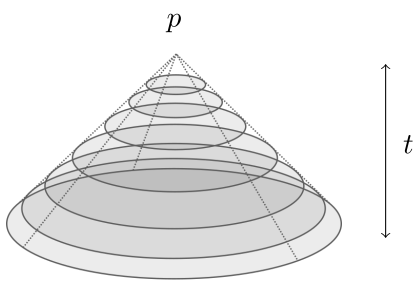

Figure 5.1: The local energy $E_{\mathrm{loc}}(p)$ is the supremum of the integrals of the energy density over the leaves of a foliation of the backward lightcone of p.

5.2.2 The Cronström Gauge

If $K(p)$ is the backward lightcone from p to the initial surface Σ as above, we can choose an open set S_p containing the set bounded by $K(p)$ and Σ and impose the Cronström gauge in S_p. The Cronström gauge is defined by

$$(x^a - x_p^a)A_a(x) = 0 \quad \text{and} \quad A_a(x_p) = 0 \quad \text{in } S_p, \tag{5.2.7}$$

and it can be shown [36] that on Minkowski space a given pair of fields (ϕ, A_a) can always be transformed to the Crönstrom gauge in any star-shaped region (within the domain of existence of the solution). Furthermore, the associated gauge transformation is trivial at p, $U(x_p) = \mathbb{1}$. An extremely useful feature of the Cronström gauge is that it allows one to express the Yang–Mills potential A_a entirely in terms of the field F_{ab}. Translating the origin to the point p as before, one has

$$A_b(x) = \int_0^1 s x^a F_{ab}(sx)\,\mathrm{d}s. \tag{5.2.8}$$

From this one also derives

$$(\nabla_a A^a)(x) = \int_0^1 \Big(s^2 x^a [F_{ab}(sx), A^b(sx)] - s^2 x^a \left((\mathrm{D}_a \phi)(sx) \cdot \theta_\alpha \phi(sx) \right) \theta_\alpha \Big) \, \mathrm{d}s. \tag{5.2.9}$$

In the following estimates we will translate an arbitrary point $p = (t_0, x_0)$ to the origin for convenience, so that the initial data will sit at $\{t = -t_0\}$. We will also write E_{loc} to denote $E_{\mathrm{loc}}(0)$, the local energy of the origin, where the lightcone considered will be of height t_0 to make contact with the initial data.

5.2.3 Integral Representations and Localization

We recall that on Minkowski space $(\mathbb{R}^4, \eta)$, $\eta = \mathrm{d}t^2 - \mathrm{d}r^2 - r^2 \mathfrak{s}_2$, the retarded Green's function G for the wave operator $\Box$ is given by

$$G(t, r) = \frac{1}{4\pi r} \delta(t - r),$$

so that any solution u to $\Box u = f$ can be written as

$$u(t_0, x_0) = u^{(0)}(t_0, x_0) + (G * f)(t_0, x_0),$$

where $u^{(0)}$ is the solution to the free wave equation $\Box u^{(0)} = 0$ determined by the data for u. The convolution $G * f$ can be expressed as an integral over the past lightcone of $p = (t_0, x_0)$: translating (t_0, x_0) to the origin, we have

$$\begin{aligned}
(G * f)(0) &= \int_{\mathbb{R}} \mathrm{d}t \int_{\mathbb{R}^3} r^2 \, \mathrm{d}r \, \mathrm{dv}_{\mathfrak{s}_2} \, G(-t, -x) f(t, x) \\
&= \int_{\mathbb{R}} \mathrm{d}t \int_{\mathbb{R}^3} r^2 \, \mathrm{d}r \, \mathrm{dv}_{\mathfrak{s}_2} \, \frac{1}{4\pi r} \delta(t + r) f(t, x) \\
&= \frac{1}{4\pi} \int_K r \, \mathrm{d}r \, \mathrm{dv}_{\mathfrak{s}_2} \, f(-r, x),
\end{aligned}$$

where $r = |x|$ and K is the past lightcone of the origin.

Suppose p is a point in the domain of local existence of some solution (ϕ, A_a) in temporal gauge. We now impose the Cronström gauge in an open set S_p containing the past lightcone $K(p)$ from p to the initial surface Σ, as described above. Note that the gauge transformation taking the temporal gauge solution (ϕ, A_a) to the Cronström gauge has $U(p) = 1$, so it follows that $F_{ab}(p)$, $\phi(p)$, and $(\mathrm{D}_a \phi)(p)$ are invariant under the gauge transformation. Using the above observation, we express the solutions to the wave equations (5.2.4) and (5.2.5) at p as integrals of the nonlinearities (in Cronström gauge) over the past lightcone $K(p)$ of p up to the initial surface Σ. As before, translating the point $p = (t_0, x_0)$ to the origin for convenience, the initial surface ends up at $\Sigma = \{t = -t_0\}$ and we find

$$\begin{aligned}
F_{\mu\nu}(0) &= F^{(0)}_{\mu\nu}(0) \\
&+ \frac{1}{4\pi} \int_{K(t_0)} r \, \mathrm{d}r \, \mathrm{dv}_{\mathfrak{s}_2} \Big\{ -2\nabla^c([A_c, F_{\mu\nu}]) + [\nabla_c A^c, F_{\mu\nu}] - [A^c, [A_c, F_{\mu\nu}]] \\
&+ \left((\mathrm{D}_\nu \phi) \cdot \theta_\alpha (\mathrm{D}_\mu \phi) - (\mathrm{D}_\mu \phi) \cdot \theta_\alpha (\mathrm{D}_\nu \phi) \right) \theta_\alpha - 2[F_\nu{}^c, F_{\mu c}] \\
&+ \left((F_{\mu\nu} \phi) \cdot \theta_\alpha \phi \right) \theta_\alpha \Big\} \Big|_{t=-r}
\end{aligned} \tag{5.2.10}$$

and

$$\begin{aligned}(\mathrm{D}_\nu\phi)(0) &= (\mathrm{D}_\nu\phi)^{(0)}(0)\\ &+ \frac{1}{4\pi}\int_{K(t_0)} r\,\mathrm{d}r\,\mathrm{dv}_{\mathfrak{s}_2}\Big\{-2\nabla^c(A_c\mathrm{D}_\nu\phi) + (\nabla^c A_c)\mathrm{D}_\nu\phi - A^c A_c \mathrm{D}_\nu\phi\\ &- 2F_\nu{}^c\mathrm{D}_c\phi + ((\mathrm{D}_\nu\phi)\cdot\theta_\alpha\phi)\,\theta_\alpha\phi - \lambda_0\mathrm{D}_\nu(|\phi|^2\phi)\Big\}\Big|_{t=-r},\end{aligned} \tag{5.2.11}$$

where the indices μ,ν indicate contraction with the basis vectors $\partial/\partial x^\mu$, $\partial/\partial x^\nu$, so that $F_{\mu\nu}$ transforms as a scalar.

Lemma 5.2.2. *The L^∞ estimates of Eardley and Moncrief can be localized entirely to the lightcone. Specifically, one has the estimate*

$$N(t) \leqslant p(t) + q(t)\int_0^t N(s)\,\mathrm{d}s,$$

where

$$N(s) = \|F(s)\|^2_{L^\infty(B(t-s))} + \|\mathrm{D}\phi(s)\|^2_{L^\infty(B(t-s))},$$

and $p(t)$ and $q(t)$ are positive polynomials (perhaps containing positive fractional powers) in t, with coefficients depending on the $(H^2(B(t))\times H^1(B(t)))^2$ norm of the temporal gauge initial data, the local energy E_{loc} in the lightcone from p to Σ, and the L^2 norm of ϕ on $B(t)\cap\Sigma$.

Proof. The terms on the right-hand sides of (5.2.10) and (5.2.11) are categorized by colour according to the types of techniques, due to Eardley and Moncrief [36], required to estimate them. The olive-coloured terms in each equation (the linear part of the solution and the first term inside the integral) can be expressed explicitly in terms of the initial data; the blue terms (the second and third terms in each integral) are dealt with by using the Cronström gauge expressions (5.2.8) and (5.2.9); the purple terms (the fourth and fifth terms in the integral for F_{ab} and the fourth term in the integral for $\mathrm{D}_a\phi$) may be estimated by observing that they all contain exactly one factor encoding the flux across the lightcone; finally, the orange terms (the last term in the integral for F_{ab} and the last two terms in the integral for $\mathrm{D}_a\phi$) are estimated by relatively simple applications of the Hölder inequality and the Sobolev embedding theorems. We briefly show how to localize one term from each colour class. The following estimates follow the original techniques of Eardley and Moncrief [36], while keeping track of the locality of spacetime quantities wherever appropriate.

The olive terms

$$I_1^{\mathrm{olive}} := F^{(0)}_{\mu\nu}(0) - \frac{1}{2\pi}\int_{K(t_0)} \nabla^c([A_c, F_{\mu\nu}])\,r\,\mathrm{d}r\,\mathrm{dv}_{\mathfrak{s}_2}$$

may be expressed explicitly, using the method of spherical means for the first term and by integrating by parts and using the condition $x^a A_a = 0$ for the second term, in terms of the temporal gauge initial data on the 2-sphere defined by $\Sigma\cap K(t_0)$. Likewise for the terms

$$I_2^{\mathrm{olive}} := (\mathrm{D}_\nu\phi)^{(0)}(0) - \frac{1}{2\pi}\int_{K(t_0)} \nabla^c(A_c\mathrm{D}_\nu\phi)\,r\,\mathrm{d}r\,\mathrm{dv}_{\mathfrak{s}_2}\,.$$

The details are contained in equation (2.39) of [36].

For the blue terms, let us consider

$$I^{\text{blue}} := \int_{K(t_0)} (\nabla^c A_c)(\mathrm{D}_\nu\phi)\, r\, \mathrm{d}r\, \mathrm{dv}_{\mathfrak{s}_2}\,.$$

Using the Cronström gauge expression (5.2.9) and the fact that $x^a F_{ab} = rl^a F_{ab} = rF_{lb}$ for $x \in K$, we find

$$\begin{aligned}
I^{\text{blue}} &= \int_0^{r_0} \mathrm{d}r \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2}\, r \int_0^1 \mathrm{d}s \Big\{ s^2 \left[x^a F_{ab}(sx), A^b(sx)\right]\Big|_K \\
&\quad - s^2 \left(x^a(\mathrm{D}_a\phi)(x)|_K \cdot (\theta_\alpha\phi)(sx)|_K\theta_\alpha\right) \Big\} (\mathrm{D}_\nu\phi)(x)|_K \\
&= \int_0^{r_0} \mathrm{d}r \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2}\, r \int_0^1 \mathrm{d}s\, s^2 r \left[F_{lb}(sx), A^b(sx)\right]\Big|_K (\mathrm{D}_\nu\phi)(x)|_K \\
&\quad - \int_0^{r_0} \mathrm{d}r \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2}\, r \int_0^1 \mathrm{d}s\, s^2 r \left((\mathrm{D}_l\phi)(sx)|_K \cdot (\theta_\alpha\phi)(sx)|_K\theta_\alpha\right)(\mathrm{D}_\nu\phi)(x)|_K \\
&=: I_1^{\text{blue}} - I_2^{\text{blue}}.
\end{aligned}$$

Consider the above summands separately. Using (5.2.8) and making the change of variables $(sr, ur) = (r', \bar{r})$, for the first one we have

$$\begin{aligned}
I_1^{\text{blue}} &= \int_0^{r_0} \mathrm{d}r \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2}\, r^3 \int_0^1 \mathrm{d}s\, s^2 \int_0^1 \mathrm{d}u\, u \left[F_{lb}(sx), {F_l}^b(ux)\right]\Big|_K (\mathrm{D}_\nu\phi)(x)|_K \\
&= \int_0^{r_0} \mathrm{d}r \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2} \int_0^1 \mathrm{d}s \int_0^1 \mathrm{d}u\, r^3 s^2 u [F_{lA}(-sr, sr, \omega), F_{lA}(-ur, ur, \omega)](\mathrm{D}_\nu\phi)(-r, r, \omega) \\
&= \int_0^{r_0} \mathrm{d}r \frac{1}{r^2} \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2} \int_0^r \mathrm{d}r' \int_0^r \mathrm{d}\bar{r}\, (r')^2 \bar{r} [F_{lA}(-r', r', \omega), F_{lA}(-\bar{r}, \bar{r}, \omega)](\mathrm{D}_\nu\phi)(-r, r, \omega) \\
&\lesssim \int_0^{r_0} \mathrm{d}r \frac{1}{r^2} \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2} \int_0^r \mathrm{d}r' \int_0^r \mathrm{d}\bar{r}\, (r')^2 \bar{r} |F_{lA}(-r', r', \omega)||F_{lA}(-\bar{r}, \bar{r}, \omega)| \|\mathrm{D}\phi(-r)\|_{L^\infty(B(r))} \\
&\lesssim \int_0^{r_0} \mathrm{d}r \frac{1}{r} \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2} \left(\int_0^r \mathrm{d}r'\, r' |F_{lA}(-r', r', \omega)|\right)^2 \|\mathrm{D}\phi(-r)\|_{L^\infty(B(r))} \\
&\lesssim \int_0^{r_0} \mathrm{d}r\, \|F_{lA}\|^2_{L^2(K(r))} \|\mathrm{D}\phi(-r)\|_{L^\infty(B(r))},
\end{aligned}$$

where $|F|$ denotes the Frobenius norm of F_{ab}, $K(r)$ is the subcone of $K(t_0)$ of height r, and we have used the Cauchy–Schwarz inequality in the last line. Using the energy identity (5.2.6), we thus have the estimate

$$I_1^{\text{blue}} \lesssim E_{\text{loc}} \int_0^{t_0} \|\mathrm{D}\phi(-t)\|_{L^\infty(B(t))}\, \mathrm{d}t.$$

To estimate I_2^{blue}, we make the same change of variables $sr = r'$ to get

$$\begin{aligned}
I_2^{\text{blue}} &= \int_0^{r_0} \mathrm{d}r \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2} \int_0^r \mathrm{d}r' \frac{1}{r} (r')^2 \left((\mathrm{D}_l\phi)(-r', r', \omega) \cdot (\theta_\alpha\phi)(-r', r', \omega)\right)(\theta_\alpha \mathrm{D}_\nu\phi)(-r, r, \omega) \\
&\lesssim \int_0^{r_0} \mathrm{d}r\, \|\mathrm{D}\phi(-r)\|_{L^\infty(B(r))} \frac{1}{r} \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2} \int_0^r \mathrm{d}r' (r')^2 |\mathrm{D}_l\phi|(-r', r', \omega)|\phi|(-r', r', \omega).
\end{aligned}$$

Using Hölder's inequality with exponents $(3, 2, 6)$, one has

$$\begin{aligned}
I_2^{\text{blue}} &\lesssim \int_0^{r_0} \mathrm{d}r\, \|\mathrm{D}\phi(-r)\|_{L^\infty(B(r))} \frac{1}{r} \left(\int_0^r (r')^2\, \mathrm{d}r'\right)^{1/3} \\
&\quad \times \left(\int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2} \int_0^r \mathrm{d}r'\, (r')^2 |\mathrm{D}_l\phi|^2(-r', r', \omega)\right)^{1/2} \left(\int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2} \int_0^r \mathrm{d}r'\, (r')^2 |\phi|^6(-r', r', \omega)\right)^{1/6} \\
&\lesssim \int_0^{r_0} \mathrm{d}r\, \|\mathrm{D}\phi(-r)\|_{L^\infty(B(r))} \|\phi\|_{L^6(K(r))} \|\mathrm{D}_l\phi\|_{L^2(K(r))}.
\end{aligned}$$

Now $\|\mathrm{D}_l\phi\|_{L^2(K(r))} \lesssim E_{\mathrm{loc}}^{1/2}$ is immediate by (5.2.6), and since

$$\frac{\mathrm{d}}{\mathrm{d}r'}(\phi(-r',r',\omega)) = (l^a\nabla_a\phi)(-r',r',\omega),$$

by the gauge-invariant Sobolev estimate of Jaffe–Taubes (see §6 of [59]) one has

$$\|\phi\|_{L^6(K(r))} \lesssim \left(\|\mathrm{D}^{\|}\phi\|_{L^2(K(r))} + \|\phi\|_{L^2(K(r))}\right),$$

where $\mathrm{D}^{\|} = (\mathrm{D}_l, \mathrm{D}_A)$. If we can show that the L^2 norm of ϕ on the cone can be controlled by the local energy and the L^2 norm of ϕ at the base of the cone, we can conclude that

$$I_2^{\mathrm{blue}} \lesssim E_{\mathrm{loc}}^{1/2}\left(2t_0E_{\mathrm{loc}}^{1/2} + \|\phi\|_{L^2(B(r_0))}\right)\int_0^{t_0}\|\mathrm{D}\phi(-t)\|_{L^\infty(B(t))}\,\mathrm{d}t.$$

Lemma 5.2.3. *The L^2 norm of ϕ on the cone $K(t_0)$ satisfies the bound*

$$\|\phi\|_{L^2(K(t_0))} \leqslant \|\phi\|_{L^2(B(r_0))}(-t_0) + 2E_{\mathrm{loc}}^{1/2}t_0.$$

If moreover $\lambda_0 \neq 0$, then

$$\|\phi\|_{L^2(K(t_0))} \lesssim E_{\mathrm{loc}}^{1/2}t_0^{3/4}(1+t_0^{1/4}).$$

Proof. Since the bound is gauge independent, it suffices to prove it in the temporal gauge. Integrate $\nabla_a(|\phi|^2K^a)$, $K^a = \partial_t$, over the region $\mathbf{K}(t_0)$ bounded by the past lightcone K of the origin and the initial surface $\Sigma = \{t = -t_0\}$:

$$\int_{\mathbf{K}(t_0)}\nabla_a(|\phi|^2K^a)\,\mathrm{d}t\wedge\mathrm{d}^3x = \int_{K(t_0)}|\phi|^2\,r^2\,\mathrm{d}r\,\mathrm{dv}_{\mathbb{S}^2} - \int_{B(r_0)\cap\Sigma}|\phi|^2\,\mathrm{d}^3x.$$

Now

$$\begin{aligned}\left|\int_{\mathbf{K}(t_0)}\nabla_a(|\phi|^2K^a)\,\mathrm{d}t\wedge\mathrm{d}^3x\right| &= \left|\int_{\mathbf{K}(t_0)}\partial_t(|\phi|^2)\,\mathrm{d}t\wedge\mathrm{d}^3x\right| \\ &\leqslant 2\int_{-t_0}^0\int_{\mathbb{S}^2}\int_0^{-t}|\phi\cdot\pi|(t,r,\omega)\,r^2\,\mathrm{d}r\,\mathrm{dv}_{\mathbb{S}^2}\,\mathrm{d}t \\ &\leqslant 2\int_0^{t_0}\|\phi\|_{L^2(B(t))}(-t)\|\pi\|_{L^2(B(t))}(-t)\,\mathrm{d}t \\ &\leqslant 2E_{\mathrm{loc}}^{1/2}\int_0^{t_0}\|\phi\|_{L^2(B(r_0))}(-t)\,\mathrm{d}t,\end{aligned}$$

where we estimate the L^2 norm of ϕ on $B(r_0)$ by

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\phi\|^2_{L^2(B(r_0))} = 2\int_{B(r_0)}\phi\cdot\pi\,\mathrm{d}^3x \leqslant 2\|\phi\|_{L^2(B(r_0))}E_{B(r_0)}^{1/2}.$$

This implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\phi\|_{L^2(B(r_0))}(t) \leqslant E_{\mathrm{loc}}^{1/2},$$

and so for $-t_0 \leqslant t \leqslant 0$

$$\|\phi\|_{L^2(B(r_0))}(t) \leqslant \|\phi\|_{L^2(B(r_0))}(-t_0) + E_{\mathrm{loc}}^{1/2}t_0.$$

Altogether then

$$\begin{aligned}\|\phi\|^2_{L^2(K(t_0))} &\leqslant \|\phi\|^2_{L^2(B(r_0))}(-t_0) + 2E_{\mathrm{loc}}^{1/2}\int_0^{t_0}(\|\phi\|_{L^2(B(r_0))}(-t_0) + E_{\mathrm{loc}}^{1/2}t_0)\,\mathrm{d}t \\ &\leqslant \|\phi\|^2_{L^2(B(r_0))}(-t_0) + 4E_{\mathrm{loc}}^{1/2}\|\phi\|_{L^2(B(r_0))}(-t_0)t_0 + 4E_{\mathrm{loc}}t_0^2 \\ &\leqslant \left(\|\phi\|_{L^2(B(r_0))}(-t_0) + 2E_{\mathrm{loc}}^{1/2}t_0\right)^2,\end{aligned}$$

which implies the first inequality. Now if $\lambda_0 \neq 0$, since $B(r_0)$ is bounded we have

$$\|\phi\|^2_{L^2(B(r_0))} \leqslant \|\phi\|^2_{L^4(B(r_0))}\frac{2}{\sqrt{3}}\sqrt{\pi}r_0^{3/2},$$

so by (5.2.6)

$$\|\phi\|_{L^2(B(r_0))}(-t_0) \lesssim E_{\mathrm{loc}}^{1/2}t_0^{3/4}.$$

Putting this into the first estimate completes the proof of the lemma. □

For the purple terms, we consider as an example the term

$$I^{\mathrm{purple}} := \int_{K(t_0)} (F_\nu{}^c\mathrm{D}_c\phi)\, r\,\mathrm{d}r\,\mathrm{dv}_{\mathfrak{s}_2}.$$

Expanding the product, we have

$$F_\nu{}^c\mathrm{D}_c\phi = -\frac{1}{2}F_{\nu l}\mathrm{D}_n\phi - \frac{1}{2}F_{\nu n}\mathrm{D}_l\phi + F_{\nu A}\mathrm{D}_A\phi,$$

so the last two terms can be estimated by

$$\begin{aligned}&\int_0^{r_0}\mathrm{d}r\int_{\mathbb{S}^2}\mathrm{dv}_{\mathfrak{s}_2}\,\|F(-r)\|_{L^\infty(B(r))}\,r|\mathrm{D}^{\|}\phi|(-r,r,\omega) \\ &\lesssim \left(\int_0^{r_0}\mathrm{d}r\,\|F(-r)\|^2_{L^\infty(B(r))}\right)^{1/2}\left(\int_0^{r_0}\mathrm{d}r\int_{\mathbb{S}^2}\mathrm{dv}_{\mathfrak{s}_2}\,r^2|\mathrm{D}^{\|}\phi|^2(-r,r,\omega)\right)^{1/2} \\ &\lesssim E_{\mathrm{loc}}^{1/2}\left(\int_0^{t_0}\|F(-t)\|^2_{L^\infty(B(t))}\,\mathrm{d}t\right)^{1/2}.\end{aligned}$$

To estimate the first term, we introduce the basis consisting of $e_0 = \partial_t$, $e_1 = \partial_r$, and e_A. One then has

$$e_0 = \frac{1}{2}(n-l) \quad \text{and} \quad e_1 = \frac{1}{2}(n+l),$$

and that the Cartesian basis $\partial/\partial x^j$ for $\mathbb{R}^3$ is related to the basis $\{e_1, e_A\}$ by an orthogonal transformation O,

$$\frac{\partial}{\partial x^j} = O_{jk}e_k, \qquad e_k = O_{jk}\frac{\partial}{\partial x^j}.$$

If $\nu = t$, using $\partial_t = \frac{1}{2}(n-l)$ the first term then reads

$$F_{tl}\mathrm{D}_n\phi = \frac{1}{2}F_{nl}\mathrm{D}_n\phi.$$

One can thus estimate

$$\begin{aligned}\int_{K(t_0)} r\,\mathrm{d}r\,\mathrm{dv}_{\mathfrak{s}_2}\,|F_{tl}\mathrm{D}_n\phi| &\lesssim \int_0^{r_0}\mathrm{d}r\int_{\mathbb{S}^2}\mathrm{dv}_{\mathfrak{s}_2}\,\|\mathrm{D}\phi(-r)\|_{L^\infty(B(r))}\,r|F_{nl}|(-r,r,\omega) \\ &\lesssim E_{\mathrm{loc}}^{1/2}\left(\int_0^{t_0}\|\mathrm{D}\phi(-t)\|^2_{L^\infty(B(t))}\,\mathrm{d}t\right)^{1/2}.\end{aligned}$$

If, on the other hand, $\nu = i$, then

$$F_{il} = O_{im} F_{e_m l} = O_{i1} F_{e_1 l} + O_{iA} F_{Al} = O_{i1} \frac{1}{2} F_{nl} + O_{iA} F_{Al},$$

so a similar estimate can be deduced.

Finally, for the orange terms let us consider as an example the term

$$I^{\text{orange}} := \int_{K(t_0)} ((\mathrm{D}_\nu \phi) \cdot \theta_\alpha \phi) (\theta_\alpha \phi) \, r \, \mathrm{d}r \, \mathrm{dv}_{\mathfrak{s}_2} .$$

Applying Cauchy–Schwarz, we have

$$\begin{aligned} I^{\text{orange}} &\lesssim \int_0^{r_0} \mathrm{d}r \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2} \, r \, \|\mathrm{D}\phi(-r)\|_{L^\infty(B(r))} |\phi|^2(-r, r, \omega) \\ &\lesssim \left(\int_0^{r_0} \mathrm{d}r \int_{\mathbb{S}^2} \mathrm{dv}_{\mathfrak{s}_2} \, r^2 |\phi|^4(-r, r, \omega) \right)^{1/2} \left(\int_0^{r_0} \mathrm{d}r \, \|\mathrm{D}\phi(-r)\|^2_{L^\infty(B(r))} \right)^{1/2} \\ &\lesssim \|\phi\|^2_{L^4(K(t_0))} \left(\int_0^{r_0} \|\mathrm{D}\phi(-r)\|^2_{L^\infty(B(r))} \right)^{1/2} . \end{aligned}$$

By Gagliardo–Nirenberg interpolation and the Jaffe–Taubes invariance argument, we have

$$\|\phi\|_{L^4(K(t_0))} \lesssim \left(\|\mathrm{D}^{\|}\phi\|^{3/4}_{L^2(K(t_0))} \|\phi\|^{1/4}_{L^2(K(t_0))} + \|\phi\|_{L^2(K(t_0))} \right),$$

so it follows that $\|\phi\|^2_{L^4(K(t_0))}$ can be estimated by a polynomial (perhaps containing positive fractional powers) in E_{loc}, t_0, and the L^2 norm of ϕ on the base of the cone $K(t_0)$.

Returning to (5.2.4) and (5.2.5), altogether the above estimates imply the bounds

$$\begin{aligned} \|F(0)\|^2_{L^\infty(B(0))} &\leqslant p_1(t_0) + q_1(t_0) \int_0^{t_0} \left(\|\mathrm{D}\phi(-t)\|^2_{L^\infty(B(t))} + \|F(-t)\|^2_{L^\infty(B(t))} \right) \mathrm{d}t, \\ \|\mathrm{D}\phi(0)\|^2_{L^\infty(B(0))} &\leqslant p_2(t_0) + q_2(t_0) \int_0^{t_0} \left(\|\mathrm{D}\phi(-t)\|^2_{L^\infty(B(t))} + \|F(-t)\|^2_{L^\infty(B(t))} \right) \mathrm{d}t, \end{aligned}$$

where $p_{1,2}(t_0)$, $q_{1,2}(t_0)$ are positive polynomials in t_0 with coefficients depending only on E_{loc} and the temporal gauge initial data (including $\|\phi\|_{L^2(B(r_0))}(-t_0)$) on $\Sigma \cap \mathbf{K}(t_0)$. Translating the origin so that p has coordinates $(t, 0)$, the lemma follows. □

Given the result of lemma 5.2.2, one now wishes to apply Grönwall's lemma to deduce that the uniform norm N does not blow up. Some care is required at this point, since the function $N(s)$ may not be continuous in s. Indeed, continuity may fail in the second variable of the function

$$f(s_1, s_2) = \|F(s_1)\|_{L^\infty(B(t-s_2))}$$

if one considers a function $F(s_1)$ with multiple maxima in $\overline{B(t)}$. But to apply Grönwall's lemma one only needs to show that $|N(s)| \, \mathrm{d}s$ defines a locally finite measure,

$$\int_0^t |N(s)| \, \mathrm{d}s < \infty.$$

But $\|F(s)\|_{L^\infty(B(t-s))} \leqslant \|F(s)\|_{L^\infty(B(t))}$, and $\|F(s)\|_{L^\infty(B(t))}$ is a continuous function in s for any fixed t, since

$$\begin{aligned} \left| \|F(s+\varepsilon)\|_{L^\infty} - \|F(s)\|_{L^\infty} \right| &\leqslant \|F(s+\varepsilon) - F(s)\|_{L^\infty} \\ &\lesssim \|F(s+\varepsilon) - F(s)\|_{H^2} \xrightarrow[\varepsilon \to 0]{} 0. \end{aligned}$$

Applying the same argument to $\|\mathrm{D}\phi(s)\|_{L^\infty(B(t-s))}$, we see that $N(s)$ is bounded by a continuous function, and is therefore a locally finite measure. We thus obtain

$$N(t) < \infty \text{ for all } t > 0.$$

The construction can be repeated for any point $p \in \mathbb{M}$, so we can package the above work into the following theorem.

Theorem 5.2.4. *Consider temporal gauge initial data* $(\mathbf{A}, \mathbf{E}, \phi, \pi) \in (H^2_{\mathrm{loc}}(\mathbb{R}^3) \times H^1_{\mathrm{loc}}(\mathbb{R}^3))^2$ *for the system* (5.2.1) *satisfying the constraint* (5.2.3). *Then the fields* F *and* $\mathrm{D}\phi$ *are* $L^\infty_{\mathrm{loc}}(\mathbb{R} \times \mathbb{R}^3)$ *in the domain of existence of the solution.*

5.3 Gluing onto the Einstein Cylinder

In this section we explain how the local uniform estimates on Minkowski space can be used to deduce global uniform estimates on the Einstein cylinder. It pays to state clearly what we shall be doing: we will prescribe $(H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3))^2$ initial data on the Einstein cylinder $\mathfrak{E}$, and consider a copy of Minkowski space $\mathbb{M}$ conformally embedded in $\mathfrak{E}$ in such a way that the initial surface in $\mathfrak{E}$ coincides with the initial surface in $\mathbb{M}$, as depicted in fig. 5.2 below. Initial data on $\mathfrak{E}$ prescribed in this way will define locally $(H^2(\mathbb{R}^3) \times H^1(\mathbb{R}^3))^2$ initial data for the system on $\mathbb{M}$, allowing us to deduce local L^∞ estimates in $\mathbb{M}$ as per the previous section. We shall then transport these local estimates back to $\mathfrak{E}$, and patch them all the way around the 3-sphere.

We conformally embed Minkowski space $(\mathbb{M}, \eta = \mathrm{d}t^2 - \mathrm{d}r^2 - r^2\mathfrak{s}_2)$ into the Einstein cylinder $(\mathfrak{E}, \mathfrak{e} = \mathrm{d}\tau^2 - \mathfrak{s}_3)$ using the conformal factor

$$\Omega = 2\cos\left(\frac{\tau+\zeta}{2}\right)\cos\left(\frac{\tau-\zeta}{2}\right) = \frac{2}{\sqrt{1+(t-r)^2}\sqrt{1+(t+r)^2}},$$

as described in section 2.3.1. One has $\Omega^2\eta = \mathfrak{e}$, and $\mathbb{M}$ is the subset of $\mathfrak{E} = \mathbb{R}_\tau \times \mathbb{S}^3$ given by

$$\mathbb{M} = \{(\tau, \zeta) : |\tau| + \zeta < \pi, \zeta \geqslant 0\} \times \mathbb{S}^2.$$

A picture of this embedding (for $t \geqslant 0$) is shown below.

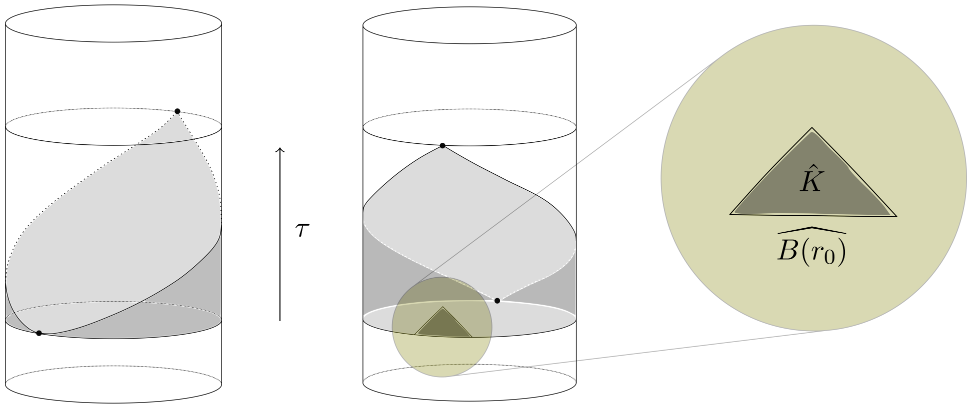

Figure 5.2: The embedding of $\mathbb{M}$ (for $t \geqslant 0$) into $\mathfrak{E}$.

Instead of considering the whole of $\mathbb{M}$ embedded into $\mathfrak{E}$, we only consider the domain of dependence of a small ball in $\mathbb{M}$ glued onto $\mathfrak{E}$. Let $B(r_0)$ be the ball of radius r_0 centred at the origin $O \in \mathbb{M}$, and consider the cone $K = D^+(B(r_0))$. We consider the image $\hat{K}$ of K under the embedding $\mathbb{M} \hookrightarrow \mathfrak{E}$; as conformal transformations preserve the causal structure, $\hat{K}$ is the domain of dependence of $\widehat{B(r_0)}$, where $\widehat{B(r_0)}$ is the image of $B(r_0)$ under the embedding.

5.3.1 Conformal Transport of Estimates

The weights

$$A_a = \hat{A}_a \quad \text{and} \quad \phi = \Omega\hat{\phi} \tag{5.3.1}$$

leave the system (2.2.27) invariant under the conformal transformation $g_{ab} \rightsquigarrow \hat{g}_{ab} = \Omega^2 g_{ab}$. As a result, the fields F_{ab} and $\mathrm{D}_a\phi$ transform according to $F_{ab} = \hat{F}_{ab}$ and $\hat{\mathrm{D}}_a\hat{\phi} = \Omega^{-1}(\mathrm{D}_a\phi - \Upsilon_a\phi)$, where $\Upsilon_a = \partial_a \log\Omega$. Consider a cone K with image $\hat{K}$ under the embedding $\mathbb{M} \hookrightarrow \mathfrak{E}$, as described above. It is clear that $0 < C_1 \leqslant |\Omega| \leqslant C_2 < \infty$ in K, so $\|\phi\|_{L^\infty(K)} \simeq \|\hat{\phi}\|_{L^\infty(\hat{K})}$ is immediate. Indeed, for example

$$|\Omega^{-1}| \leqslant \frac{1}{2}\left|\sqrt{1+(t-r)^2}\sqrt{1+(t+r)^2}\right| \leqslant \frac{1}{2}(1+4r_0^2),$$

and

$$|\Omega| \leqslant 2.$$

To deduce the same type of equivalence for tensor fields, one needs to check that the norms defined by the Riemannian metrics

$$\Gamma^{ab} = 2T^aT^b - \eta_{ab} \quad \text{and} \quad \hat{\Gamma}^{ab} = 2\hat{T}^a\hat{T}^b - \mathfrak{e}^{ab},$$

where $T^a = \partial_t$ and $\hat{T}^a = \partial_\tau$, are equivalent, at least in K.

Proposition 5.3.1. *For any* 1*-form* X_a *one has* $|X|_\Gamma \simeq |X|_{\hat{\Gamma}}$ *in* K*.*

Proof. By a direct calculation using the chain rule, one finds

$$T^a = \frac{1}{4}\Omega^2\left((2+u^2+v^2)\hat{T}^a + (u^2-v^2)\hat{Z}^a\right),$$

where $\hat{Z}^a = \partial_\zeta$, $u = t-r$, and $v = t+r$. A further calculation then shows that

$$\begin{aligned}\Omega^{-2}|X|^2_\Gamma = {}& \frac{1}{8}\Omega^2\left((2+u^2+v^2)^2 - 1\right)(\hat{T}^aX_a)^2 \\ &+ \frac{1}{4}\Omega^2(2+u^2+v^2)(u^2-v^2)(\hat{T}^aX_a)(\hat{Z}^aX_a) + (u^2-v^2)^2(\hat{Z}^aX_a)^2 + |X|^2_{\mathfrak{s}_3}.\end{aligned}$$

It is clear that $|X|^2_\Gamma \lesssim |X|^2_{\hat{\Gamma}}$, while for the lower bound it is enough to observe that

$$(2+u^2+v^2)(u^2-v^2)(\hat{T}^aX_a)(\hat{Z}^aX_a) \geqslant -\frac{1}{4}(2+u^2+v^2)^2(\hat{T}^aX_a)^2 - (u^2-v^2)^2(\hat{Z}^aX_a)^2,$$

so that

$$\begin{aligned}\Omega^{-2}|X|^2_\Gamma &\geqslant \frac{1}{8}\Omega^2\left(\frac{1}{2}(2+u^2+v^2)^2-1\right)(\hat{T}^aX_a)^2\\ &+\left((u^2-v^2)^2\left(1-\frac{\Omega^2}{4}\right)\right)(\hat{Z}^aX_a)^2+|X|^2_{\mathfrak{s}_3}\\ &\geqslant \frac{1}{8}\Omega^2(\hat{T}^aX_a)^2+|X|^2_{\mathfrak{s}_3}\\ &\geqslant \frac{1}{8}\Omega^2|X|^2_{\hat{\Gamma}},\end{aligned}$$

as $\Omega^2/4 \leqslant 1$. □

It follows that

$$\|F\|_{L^\infty(K)} \simeq \|\hat{F}\|_{L^\infty(\hat{K})} \quad \text{and} \quad \|\hat{\mathrm{D}}\hat{\phi}\|_{L^\infty(\hat{K})} \lesssim \|\mathrm{D}\phi\|_{L^\infty(K)} + \|\Upsilon\phi\|_{L^\infty(K)}.$$

Note that these are gauge-independent. This demonstrates that local L^∞ estimates on Minkowski space imply local L^∞ estimates on the Einstein cylinder. We show below how initial data on the Einstein cylinder defines initial data on Minkowski space, and use this to complete our construction.

Consider temporal gauge (with respect to ∂_τ) initial data $(\hat{\mathbf{A}}, \hat{\mathbf{E}}, \hat{\phi}, \hat{\pi}) \in (H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3))^2$ for the Yang–Mills–Higgs equations on $\mathfrak{E}$,

$$\dot{\hat{\mathbf{E}}}_i + \not{\nabla}^j \hat{F}_{ij} + [\hat{\mathbf{A}}^j, \hat{F}_{ij}] = ((\hat{\not{\mathbf{D}}}_i\hat{\phi})\cdot\theta_\alpha\hat{\phi})\theta_\alpha, \quad \dot{\hat{\pi}} - \hat{\not{\mathbf{D}}}^j\hat{\not{\mathbf{D}}}_j\hat{\phi} + \hat{\phi} + \lambda_0|\hat{\phi}|^2\hat{\phi} = 0, \tag{5.3.2}$$

satisfying the constraint

$$\not{\nabla}^j\hat{\mathbf{E}}_j + [\hat{\mathbf{A}}^j, \hat{\mathbf{E}}_j] = (\hat{\pi}\cdot\theta_\alpha\hat{\phi})\theta_\alpha, \tag{5.3.3}$$

where $\hat{\not{\mathbf{D}}}_i\hat{\phi}$ is $\hat{\mathrm{D}}_a\hat{\phi}$ projected to $\mathbb{S}^3$. Since $\hat{T}^a = \partial_\tau$ is not everywhere parallel to $T^a = \partial_t$, the temporal gauge on $\mathfrak{E}$ is of course not the same as the temporal gauge on $\mathbb{M}$. However, $\hat{T}^a$ and T^a *are* parallel on the initial surface $\Sigma_0 = \{\tau = 0\} = \{t = 0\}$,

$$r_+^2\frac{\partial}{\partial t}\bigg|_{t=0} = \frac{\partial}{\partial\tau}\bigg|_{\tau=0},$$

where $r_+^2 = (1+r^2)/2$. Thus on the initial surface Σ_0 one has $A_0 = 0$ a.e. $\iff$ $\hat{A}_0 = 0$ a.e.. The data $(\hat{\mathbf{A}}, \hat{\mathbf{E}}) \in H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3)$ then gives rise to temporal gauge initial data $(\mathbf{A}, \mathbf{E}) \in H^2_{\mathrm{loc}}(\mathbb{R}^3) \times H^1_{\mathrm{loc}}(\mathbb{R}^3)$ on Minkowski space: one has

$$\hat{\mathbf{A}}_a\Big|_{\tau=0} = \mathbf{A}_a\Big|_{t=0} \quad \text{and} \quad \hat{\mathbf{E}}_a\Big|_{\tau=0} = r_+^2\mathbf{E}_a\Big|_{t=0}.$$

For the scalar field part, one similarly has

$$\hat{\phi}\Big|_{\tau=0} = r_+^2\phi\Big|_{t=0},$$

and (since $(\partial_t\Omega)|_{t=0} = 0$),

$$(\partial_\tau\hat{\phi})\Big|_{\tau=0} = (\Omega^{-1}\partial_t\hat{\phi})\Big|_{t=0} = (\Omega^{-2}\partial_t\phi)\Big|_{t=0} = r_+^4(\partial_t\phi)\Big|_{t=0},$$

i.e.

$$\hat{\pi}\Big|_{\tau=0} = r_+^4 \pi\Big|_{t=0}.$$

Thus $(\hat{\phi}, \hat{\pi}) \in H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3)$ similarly gives rise to temporal gauge initial data $(\phi, \pi) \in H^2_{\mathrm{loc}}(\mathbb{R}^3) \times H^1_{\mathrm{loc}}(\mathbb{R}^3)$. Furthermore, that the Minkowskian initial data satisfies the constraint equation (5.2.3) as a consequence of the constraint equation (5.3.3) on the Einstein cylinder follows from the conformal invariance of the field equations and the fact that ∂_t and ∂_τ are parallel initially. In summary, $(H^2 \times H^1)^2$ temporal gauge initial data on $\mathfrak{E}$ gives rise to $(H^2_{\mathrm{loc}} \times H^1_{\mathrm{loc}})^2$ temporal gauge initial data on $\mathbb{M}$.

Remark 5.3.2. The locality is necessary. Indeed, the measures on $\{t = 0\}$ and $\{\tau = 0\}$ are related by

$$\mathrm{dv}_{\mathfrak{S}^3} = r_+^{-6}\, \mathrm{dv}_{\mathbb{R}^3},$$

so the L^2 norms of the initial data scale as

$$\int_{\mathbb{S}^3} |\hat{\phi}|^2 \,\mathrm{dv}_{\mathfrak{S}^3} = \int_{\mathbb{R}^3} \frac{1}{r_+^2} |\phi|^2 \,\mathrm{dv}_{\mathbb{R}^3}, \qquad \int_{\mathbb{S}^3} |\hat{\mathbf{A}}|^2 \,\mathrm{dv}_{\mathfrak{S}^3} = \int_{\mathbb{R}^3} \frac{1}{r_+^2} |\mathbf{A}|^2 \,\mathrm{dv}_{\mathbb{R}^3},$$
$$\int_{\mathbb{S}^3} |\hat{\pi}|^2 \,\mathrm{dv}_{\mathbb{S}^3} = \int_{\mathbb{R}^3} r_+^2 |\pi|^2 \,\mathrm{dv}_{\mathbb{R}^3}, \qquad \int_{\mathbb{S}^3} |\hat{\mathbf{E}}|^2 \,\mathrm{dv}_{\mathbb{S}^3} = \int_{\mathbb{R}^3} r_+^2 |\mathbf{E}|^2 \,\mathrm{dv}_{\mathbb{R}^3},$$

where $|\hat{\mathbf{A}}|^2$ and $|\hat{\mathbf{E}}|^2$ are computed with respect to the metric on $\mathbb{S}^3$, while $|\mathbf{A}|^2$ and $|\mathbf{E}|^2$ are computed with respect to the metric on $\mathbb{R}^3$ as appropriate. One sees that $H^1(\mathbb{S}^3) \times L^2(\mathbb{S}^3)$ initial data on $\mathbb{S}^3$ does not define $H^1(\mathbb{R}^3) \times L^2(\mathbb{R}^3)$ data on $\mathbb{R}^3$, particularly because the finiteness of the weighted norm

$$\int_{\mathbb{R}^3} \frac{1}{r_+^2} \left(|\phi|^2 + |\mathbf{A}|^2 \right) \mathrm{dv}_{\mathbb{R}^3}$$

does not imply $\|\phi\|_{L^2(\mathbb{R}^3)} + \|\mathbf{A}\|_{L^2(\mathbb{R}^3)} < \infty$.

Consider any local solution $(\hat{\phi}, \hat{A}_a)$ on $\mathfrak{E}$ with $(H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3))^2$ initial data. Then the conformally related fields $(\phi, A_a) = (\Omega\hat{\phi}, \hat{A}_a)$ restricted to $\mathbb{M}$ are a solution to the Yang–Mills–Higgs equations on $\mathbb{M}$ with $(H^2_{\mathrm{loc}} \times H^1_{\mathrm{loc}})^2$ initial data, so by the local L^∞ estimates of section 5.2 satisfy

$$\|F\|_{L^\infty(K)} + \|\mathrm{D}\phi\|_{L^\infty(K)} < \infty.$$

To show that this implies

$$\|\hat{F}\|_{L^\infty(\hat{K})} + \|\hat{\mathrm{D}}\hat{\phi}\|_{L^\infty(\hat{K})} < \infty,$$

it only remains to check that $\|\Upsilon\phi\|_{L^\infty}$ is bounded in K. We have $\|\Upsilon\phi\|_{L^\infty(K)} \leqslant \|\Upsilon\|_{L^\infty(K)} \|\phi\|_{L^\infty(K)}$, and can estimate $\Upsilon_a = \partial_a \log \Omega$ easily by, for example,

$$|\Upsilon_t| \leqslant \left| \frac{(t-r)}{1+(t-r)^2} + \frac{(t+r)}{1+(t+r)^2} \right| \leqslant 2t \leqslant 2r_0,$$

for the Υ_t component, and similarly for the Υ_r component. To estimate $\|\phi\|_{L^\infty}$, we make use of the temporal gauge condition on $\mathbb{M}$,

$$\phi(t) = \phi(0) + \int_0^t \pi(s)\, \mathrm{d}s,$$

so that

$$\begin{aligned}\|\phi\|_{L^\infty(K)} &\leqslant \|\phi(0)\|_{L^\infty(B(r_0))} + t\|\pi\|_{L^\infty(K)} \\ &\leqslant \|\phi(0)\|_{H^2(B(r_0))} + r_0\|\pi\|_{L^\infty(K)} \\ &< \infty.\end{aligned}$$

Since the $\|\phi\|_{L^\infty}$ norm is gauge-independent, this does not present any issues with respect to gauge choice. Thus $\|\Upsilon\phi\|_{L^\infty(K)} < \infty$, and we have

$$\|\hat{F}\|_{L^\infty(\hat{K})} + \|\hat{\mathrm{D}}\hat{\phi}\|_{L^\infty(\hat{K})} < \infty.$$

Since the position of the cone $\hat{K}$ on the Einstein cylinder was arbitrary (inasmuch as the position of the embedded copy of Minkowski space was arbitrary in $\mathfrak{E}$), we have proven the following.

Theorem 5.3.3. *For given temporal gauge initial data* $(\hat{\mathbf{A}}, \hat{\mathbf{E}}, \hat{\phi}, \hat{\pi}) \in (H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3))^2$ *for the system* (5.3.2) *satisfying the constraint* (5.3.3)*, the fields* $\hat{F}$ *and* $\hat{\mathrm{D}}\hat{\phi}$ *are* $L^\infty([0,\tau_0] \times \mathbb{S}^3)$ *for some* τ_0 *independent of the size of the initial data.*

5.4 Global Existence on the Einstein Cylinder

5.4.1 Local Existence à la Choquet-Bruhat and Christodoulou

Theorem 5.4.1 (Choquet-Bruhat and Christodoulou, 1981, [23])**.** *Let* $(\hat{\mathbf{a}}, \hat{\mathbf{e}}, \hat{\phi}_0, \hat{\phi}_1) \in (H^s(\mathbb{S}^3) \times H^{s-1}(\mathbb{S}^3))^2$ *and* $\hat{a}_0 \in H^s(\mathbb{S}^3)$*,* $s \geqslant 2$*, be initial data for the Yang–Mills–Higgs equations*

$$\hat{\mathrm{D}}^b \hat{F}_{ab} = -((\hat{\mathrm{D}}_a\hat{\phi}) \cdot \theta_\alpha \hat{\phi})\theta_\alpha, \qquad \hat{\mathrm{D}}^a\hat{\mathrm{D}}_a\hat{\phi} + \hat{\phi} + \lambda_0|\hat{\phi}|^2\hat{\phi} = 0 \tag{5.4.1}$$

on $\mathfrak{E}$ *satisfying the constraint*

$$\hat{\nabla}^j \hat{\mathbf{e}}_j + [\hat{\mathbf{a}}^j, \hat{\mathbf{e}}_j] = (\hat{\pi} \cdot \theta_\alpha \hat{\phi}_0)\theta_\alpha, \tag{5.4.2}$$

where $\hat{\pi} = \hat{\phi}_1 + \hat{a}_0\hat{\phi}_0$*. Then there exists* $\varepsilon > 0$ *such that there exists a solution*

$$\hat{\phi},\ \hat{A}_a \in E_s((-\varepsilon,\varepsilon) \times \mathbb{S}^3) := \bigcap_{k=0}^{s} \mathcal{C}^k((-\varepsilon,\varepsilon); H^{s-k}(\mathbb{S}^3))$$

to (5.4.1) *in Lorenz gauge* $\hat{\nabla}_a \hat{A}^a = 0$*, with*

$$\hat{\mathbf{A}}\Big|_{\tau=0} = \hat{\mathbf{a}}, \qquad \hat{A}_0\Big|_{\tau=0} = \hat{a}_0, \qquad \hat{\mathbf{E}}\Big|_{\tau=0} = \hat{\mathbf{e}}, \qquad \hat{\phi}\Big|_{\tau=0} = \hat{\phi}_0, \qquad \dot{\hat{\phi}}\Big|_{\tau=0} = \hat{\phi}_1.$$

The largest such number ε *depends continuously on the size* M *of the data, where*

$$M = \|\hat{\phi}_0\|_{H^s} + \|\hat{\mathbf{a}}\|_{H^s} + \|\hat{\phi}_1\|_{H^{s-1}} + \|\hat{\mathbf{e}}\|_{H^{s-1}} + \|\hat{a}_0\|_{H^s},$$

and tends to infinity as M *tends to zero. Furthermore, the solution is unique*[3] *up to gauge transformations preserving the Lorenz gauge.*

[3] It is worth recalling here that we work with a *compact* connected gauge group G.

Remark 5.4.2. The component $\hat{A}_0$ is non-dynamical and the $\hat{a}_0$ component of the initial data can in fact be chosen to be zero without restricting the class of solutions (cf. §4 of [23]).

Corollary 5.4.3. *Let* $(\hat{\mathbf{a}}, \hat{\mathbf{e}}, \hat{\phi}_0, \hat{\phi}_1) \in (H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3))^2$ *be temporal gauge initial data for the system* (5.4.1) *on* $\mathfrak{E}$*, satisfying the constraint* (5.4.2)*. Then there exists* $\varepsilon > 0$ *such that there exists a solution* $(\hat{\phi}, \hat{A}_a)$ *in* $E_2((-\varepsilon, \varepsilon) \times \mathbb{S}^3)$ *to* (5.4.1) *in temporal gauge, with*

$$\hat{\mathbf{A}}\Big|_{\tau=0} = \hat{\mathbf{a}}, \qquad \hat{\mathbf{E}}\Big|_{\tau=0} = \hat{\mathbf{e}}, \qquad \hat{\phi}\Big|_{\tau=0} = \hat{\phi}_0, \qquad \hat{\pi}\Big|_{\tau=0} = \hat{\phi}_1.$$

The largest such number ε *depends continuously on the size* M' *of the data, where*

$$M' = \|\hat{\phi}_0\|_{H^2} + \|\hat{\mathbf{a}}\|_{H^2} + \|\hat{\phi}_1\|_{H^1} + \|\hat{\mathbf{e}}_1\|_{H^1},$$

and tends to infinity as M' *tends to zero. Furthermore, the solution is unique up to gauge transformations preserving the temporal gauge.*

Proof. This is immediate from theorem 5.4.1, if one can demonstrate that there exists a gauge transformation from the Lorenz gauge to the temporal gauge preserving the requisite regularity. A general gauge transformation U of the system (5.4.1) takes

$$\hat{A}_a \rightsquigarrow U\hat{A}_aU^{-1} + U\partial_aU^{-1},$$

so to set $\hat{A}_0 = 0$ one needs to solve $U\hat{A}_0U^{-1} + U\partial_\tau U^{-1} = 0$, or equivalently

$$\hat{A}_0 = U^{-1}\partial_\tau U.$$

Since G is a compact connected matrix Lie group, there exists $u \in \mathfrak{g}$ such that $U = \mathrm{e}^u$, so in terms of u the above equation becomes $\partial_\tau u = \hat{A}_0$. This has the solution

$$u(\tau) = u(0) + \int_0^\tau \hat{A}_0(\sigma)\,\mathrm{d}\sigma,$$

so choosing $u(0) = 0$ (and $\hat{a}_0 = 0$) gives the required gauge transformation. □

Remark 5.4.4. It is implicit in theorem 5.4.1 that if the largest time of existence is finite, $\varepsilon_{\max} < \infty$, then

$$\|\hat{\phi}(\tau)\|_{H^2(\mathbb{S}^3)} + \|\hat{\mathbf{A}}(\tau)\|_{H^2(\mathbb{S}^3)} + \|\hat{\mathbf{E}}(\tau)\|_{H^1(\mathbb{S}^3)} + \|\hat{\pi}(\tau)\|_{H^1(\mathbb{S}^3)} \longrightarrow \infty$$

as $\tau \to \varepsilon_{\max}$. We shall show that the time of existence is in fact infinite by showing that the above norm does not blow up in finite time.

5.4.2 Energy Estimates

On the Einstein cylinder $\mathfrak{E}$ we use the alternative stress-energy tensor (2.2.30) for the system (2.2.27),

$$\begin{aligned}\hat{\mathbf{T}}_{ab} = &- \langle \hat{F}_{ac}, \hat{F}_b{}^c \rangle + \frac{1}{4}\mathfrak{e}_{ab}\langle \hat{F}_{cd}, \hat{F}^{cd}\rangle \\ &+ (\hat{\mathrm{D}}_a\hat{\phi})\cdot(\hat{\mathrm{D}}_b\hat{\phi}) - \frac{1}{2}\mathfrak{e}_{ab}(\hat{\mathrm{D}}_c\hat{\phi})\cdot(\hat{\mathrm{D}}^c\hat{\phi}) + \frac{1}{2}\mathfrak{e}_{ab}|\hat{\phi}|^2 + \frac{1}{4}\lambda_0\mathfrak{e}_{ab}|\hat{\phi}|^4.\end{aligned} \tag{5.4.3}$$

This differs from the canonical stress-energy tensor (2.2.29) on $\mathfrak{E}$ by a term proportional to $\hat{\mathrm{R}}_{ab}|\hat{\phi}|^2$, but satisfies the exact conservation law

$$\hat{\nabla}^a\hat{\mathbf{T}}_{ab} = 0.$$

It thus defines a conserved energy on $\mathfrak{E}$,

$$\begin{aligned}\hat{E}_0 &= \int_{\mathbb{S}^3} \hat{\mathbf{T}}_{ab}(\partial_\tau)^a(\partial_\tau)^b \,\mathrm{dv}_{\mathfrak{s}_3} \\ &= \frac{1}{2}\int_{\mathbb{S}^3}\left(|\hat{\mathbf{E}}|^2 + |\hat{\mathbf{B}}|^2 + |\hat{\pi}|^2 + |\hat{\not{\mathbf{D}}}\hat{\phi}|^2 + |\hat{\phi}|^2 + \frac{1}{2}\lambda_0|\hat{\phi}|^4\right)\mathrm{dv}_{\mathfrak{s}_3},\end{aligned}$$

satisfying

$$\frac{\mathrm{d}\hat{E}_0}{\mathrm{d}\tau} = 0,$$

where $\hat{\not{\mathbf{D}}}\hat{\phi}$ is the projection onto $\mathbb{S}^3$ of $\mathrm{D}_a\phi$. We also define the approximate energies

$$\hat{\mathcal{E}}_1(\tau) := \frac{1}{2}\int_{\mathbb{S}^3}\left(|\hat{\mathbf{E}}|^2 + |\not{\nabla}\hat{\mathbf{A}}|^2 + |\hat{\mathbf{A}}|^2 + |\hat{\pi}|^2 + |\not{\nabla}\hat{\phi}|^2 + |\hat{\phi}|^2\right)\mathrm{dv}_{\mathfrak{s}_3} \tag{5.4.4}$$

and

$$\hat{\mathcal{E}}_2(\tau) := \frac{1}{2}\int_{\mathbb{S}^3}\left(|\not{\nabla}\hat{\mathbf{E}}|^2 + |\not{\nabla}^2\hat{\mathbf{A}}|^2 + |\not{\nabla}\hat{\pi}|^2 + |\not{\nabla}^2\hat{\phi}|^2\right)\mathrm{dv}_{\mathfrak{s}_3}\,. \tag{5.4.5}$$

It is clear that $(\hat{\mathcal{E}}_1 + \hat{\mathcal{E}}_2)^{1/2}$ is equivalent to the $(H^2 \times H^1)^2$ norm of the solution (in temporal gauge) on $\{\tau\}\times\mathbb{S}^3$. By differentiating $\hat{\mathcal{E}}_1$ in τ, integrating by parts and using the equations (5.3.2) and (5.3.3), one arrives at the estimate

$$\begin{aligned}\left|\frac{\mathrm{d}\hat{\mathcal{E}}_1}{\mathrm{d}\tau}\right| &\lesssim \left(1 + \|\hat{F}(\tau)\|_{L^\infty} + \|\hat{\mathrm{D}}\hat{\phi}(\tau)\|_{L^\infty}\right)\hat{\mathcal{E}}_1 + \lambda_0\|\hat{\pi}(\tau)\|_{L^\infty}\|\hat{\phi}\|^3_{L^3} \\ &\lesssim \left(1 + \|\hat{F}(\tau)\|_{L^\infty} + \|\hat{\mathrm{D}}\hat{\phi}(\tau)\|_{L^\infty} + \lambda_0\|\hat{\mathrm{D}}\hat{\phi}(\tau)\|_{L^\infty}\|\hat{\phi}(\tau)\|_{L^\infty}\right)\hat{\mathcal{E}}_1\end{aligned} \tag{5.4.6}$$

where the constants depend only on the structure group G and the geometry of $\mathbb{S}^3$. One similarly finds that

$$\left|\frac{\mathrm{d}\hat{\mathcal{E}}_2}{\mathrm{d}\tau}\right| \lesssim \left(1 + \|\hat{F}(\tau)\|_{L^\infty} + \|\hat{\mathrm{D}}\hat{\phi}(\tau)\|_{L^\infty} + \|\hat{\phi}(\tau)\|_{L^\infty} + \|\hat{\mathbf{A}}(\tau)\|_{L^\infty}\right)^2(\hat{\mathcal{E}}_1 + \hat{\mathcal{E}}_2). \tag{5.4.7}$$

Putting together (5.4.6) and (5.4.7), it follows that

$$\left|\frac{\mathrm{d}}{\mathrm{d}\tau}(\hat{\mathcal{E}}_1 + \hat{\mathcal{E}}_2)\right| \lesssim \left(1 + \|\hat{F}(\tau)\|_{L^\infty} + \|\hat{\mathrm{D}}\hat{\phi}(\tau)\|_{L^\infty} + \|\hat{\phi}(\tau)\|_{L^\infty} + \|\hat{\mathbf{A}}(\tau)\|_{L^\infty}\right)^2(\hat{\mathcal{E}}_1 + \hat{\mathcal{E}}_2). \tag{5.4.8}$$

To estimate $\|\hat{\phi}\|_{L^\infty}$ and $\|\hat{\mathbf{A}}\|_{L^\infty}$, notice that $\partial_\tau \hat{\phi} = \hat{\pi}$ and $\partial_\tau \hat{\mathbf{A}} = \hat{\mathbf{E}}$ imply

$$\hat{\phi}(\tau) = \hat{\phi}(0) + \int_0^\tau \hat{\pi}(\sigma)\,\mathrm{d}\sigma \quad \text{and} \quad \hat{\mathbf{A}}(\tau) = \hat{\mathbf{A}}(0) + \int_0^\tau \hat{\mathbf{E}}(\sigma)\,\mathrm{d}\sigma,$$

which give the estimates

$$\begin{aligned} &\|\hat{\phi}(\tau)\|_{L^\infty} \lesssim \|\hat{\phi}_0\|_{H^2} + \int_0^\tau \|\hat{\pi}(\sigma)\|_{L^\infty}\,\mathrm{d}\sigma \quad \text{and} \\ &\|\hat{\mathbf{A}}(\tau)\|_{L^\infty} \lesssim \|\hat{\mathbf{a}}\|_{H^2} + \int_0^\tau \|\hat{\mathbf{E}}(\sigma)\|_{L^\infty}\,\mathrm{d}\sigma. \end{aligned} \tag{5.4.9}$$

We thus have the following.

Theorem 5.4.5. *Let $(\hat{\mathbf{a}}, \hat{\mathbf{e}}, \hat{\phi}_0, \hat{\phi}_1) \in (H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3))^2$ be temporal gauge initial data for the system (5.4.1) on $\mathfrak{E}$ satisfying the constraint (5.4.2). Then there exists a global solution $(\hat{\phi}, \hat{A}_a)$ in $E_2(\mathbb{R} \times \mathbb{S}^3)$ to (5.4.1) in temporal gauge with*

$$\hat{\mathbf{A}}\Big|_{\tau=0} = \hat{\mathbf{a}}, \qquad \hat{\mathbf{E}}\Big|_{\tau=0} = \hat{\mathbf{e}}, \qquad \hat{\phi}\Big|_{\tau=0} = \hat{\phi}_0, \quad \textit{and} \quad \hat{\pi}\Big|_{\tau=0} = \hat{\phi}_1.$$

Furthermore, the solution is unique up to gauge transformations preserving the temporal gauge.

Proof. Let $\varepsilon_{\max} > 0$ be the maximal time of existence guaranteed by theorem 5.4.1. As per remark 5.4.4, either $\varepsilon_{\max} = \infty$ or the $(H^2 \times H^1)^2$ norm of the solution blows up as $\tau \to \varepsilon_{\max}$. We show that the former is true by assuming that $\varepsilon_{\max} < \infty$ and deriving a contradiction. We work with $\tau \geqslant 0$; the following argument applies equally well in the case $\tau < 0$. The local solution $(\hat{A}_a, \hat{\phi})$ satisfies

$$\|\hat{\mathbf{A}}(\tau)\|_{H^2} + \|\hat{\mathbf{E}}(\tau)\|_{H^1} + \|\hat{\phi}(\tau)\|_{H^2} + \|\hat{\pi}(\tau)\|_{H^1} < \infty$$

for all $\tau < \varepsilon_{\max}$, and in particular at $\tau = \varepsilon_{\max} - \tau_0/2$, where τ_0 is as in theorem 5.3.3. By considering the fields $(\hat{\mathbf{A}}, \hat{\mathbf{E}}, \hat{\phi}, \hat{\pi})$ restricted to $\tau = \varepsilon_{\max} - \tau_0/2$ as initial data and applying theorem 5.3.3, one has that

$$\|\hat{F}(\tau)\|_{L^\infty} + \|\hat{\mathrm{D}}\hat{\phi}(\tau)\|_{L^\infty} < \infty$$

for $\tau \leqslant \varepsilon_{\max} + \tau_0/2$. But then the estimates (5.4.9) show that

$$\|\hat{\phi}(\tau)\|_{L^\infty} + \|\hat{\mathbf{A}}(\tau)\|_{L^\infty} < \infty$$

for $\tau \leqslant \varepsilon_{\max} + \tau_0/2$, and so by (5.4.8) one deduces that $(\hat{\mathcal{E}}_1 + \hat{\mathcal{E}}_2)(\tau) < \infty$ up to $\tau = \varepsilon_{\max} + \tau_0/2$. Since $(\hat{\mathcal{E}}_1 + \hat{\mathcal{E}}_2)^{1/2}$ is equivalent to the $(H^2 \times H^1)^2$ norm of $(\hat{\mathbf{A}}, \hat{\mathbf{E}}, \hat{\phi}, \hat{\pi})$, this contradicts the assumption that $\varepsilon_{\max}$ was the maximal time of existence. Thus $\varepsilon_{\max} = \infty$. □

5.5 Asymptotics

5.5.1 De Sitter Space

Recall that de Sitter space dS_4 is the manifold $\mathbb{R} \times \mathbb{S}^3$ equipped with the metric

$$\tilde{g} = \mathrm{d}\alpha^2 - (\cosh^2 \alpha)\mathfrak{s}_3. \tag{5.5.1}$$

The vector field $\tilde{T}^a = \partial_\alpha$ is uniformly timelike and normal to surfaces of constant α; we define the associated Riemannian metric $\tilde{\Gamma}$ on dS_4 by

$$\tilde{\Gamma}_{ab} = 2\tilde{T}_a \tilde{T}_b - \tilde{g}_{ab}.$$

By making the change of variables $\tan(\tau/2) = \tanh(\alpha/2)$, one finds that the de Sitter metric is conformal to the metric on the Einstein cylinder,

$$\tilde{g} = \frac{1}{\cos^2 \tau}(\mathrm{d}\tau^2 - \mathfrak{s}_3),$$

with the associated conformal factor $\omega = \cos\tau$. Under this conformal transformation dS_4 is mapped to the section $(-\pi/2, \pi/2) \times \mathbb{S}^3$ of the Einstein cylinder, which puts the past and future null infinities of dS_4 at

$$\mathscr{I}^- = \left\{\tau = -\frac{\pi}{2}\right\} \times \mathbb{S}^3 \quad \text{and} \quad \mathscr{I}^+ = \left\{\tau = \frac{\pi}{2}\right\} \times \mathbb{S}^3.$$

Let us denote by $\tilde{\phi}$ and $\tilde{A}_a$ the scalar field and the Yang–Mills potential on de Sitter space. These are conformally related to the corresponding fields $\hat{\phi}$ and $\hat{A}_a$ on the Einstein cylinder by

$$\hat{\phi} = \omega^{-1}\tilde{\phi}, \qquad \hat{A}_a = \tilde{A}_a.$$

It is clear that $(H^2 \times H^1)^2$ initial data on the hypersurface $\{\alpha = 0\}$ in de Sitter space defines $(H^2 \times H^1)$ initial data on $\{\tau = 0\}$ in the Einstein cylinder. This follows from the fact that ∂_α is everywhere parallel to ∂_τ, and the form of the conformal factor ω. By theorem 5.4.5, we thus have a temporal gauge solution $(\hat{\phi}, \hat{A}_a)$ in $E_2(\mathbb{R} \times \mathbb{S}^3)$ on $\mathfrak{E}$, which is uniformly continuous on $I \times \mathbb{S}^3$ for any compact interval I. Indeed, this follows from the Sobolev embedding $H^2(\mathbb{S}^3) \hookrightarrow \mathcal{C}^{0,\frac{1}{2}}(\mathbb{S}^3)$, which implies the inclusion

$$E_2(I \times \mathbb{S}^3) \subset \mathcal{C}^0(I \times \mathbb{S}^3).$$

Fixing the residual gauge freedom if necessary, we thus deduce that there exists a constant $c > 0$ such that

$$|\tilde{\phi}| \leqslant c\omega \lesssim \mathrm{e}^{-|\alpha|},$$

and, since $|\tilde{A}|^2_{\tilde{\Gamma}} = \omega^2 |\hat{A}|^2_{\hat{\Gamma}}$, also that

$$|\tilde{A}|_{\tilde{\Gamma}} \leqslant c\omega \lesssim \mathrm{e}^{-|\alpha|}.$$

5.5.2 Minkowski Space

Here we denote by (ϕ, A_a) the fields on Minkowski space $\mathbb{M}$, with the corresponding conformally related fields on the Einstein cylinder still denoted $(\hat{\phi}, \hat{A}_a)$. Let $(\mathbf{a}, \mathbf{e}, \phi_0, \phi_1)$ be temporal gauge initial data for (5.2.2) satisfying the constraint (5.2.3), such that

$$(\hat{\mathbf{a}}, \hat{\mathbf{e}}, \hat{\phi}_0, \hat{\phi}_1) = (\mathbf{a}, r_+^2\mathbf{e}, r_+^2\phi_0, r_+^4\phi_1) \in H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3) \times H^2(\mathbb{S}^3) \times H^1(\mathbb{S}^3).$$ [4]

By construction, the data is such that it satisfies the hypotheses of theorem 5.4.5, giving a global temporal gauge solution $(\hat{\phi}, \hat{A}_a)$ on the Einstein cylinder. This solution is related to the solution on Minkowski space by the usual scaling $\phi = \Omega\hat{\phi}$ and $A_a = \hat{A}_a$, where $\Omega = 2(1+(t-r)^2)^{-1/2}(1+(t+r)^2)^{-1/2}$. Set $u = t-r$, $v = t+r$, $u_+ = \sqrt{1+u^2}$, and $v_+ = \sqrt{1+v^2}$. On $\mathbb{M}$ we have the tetrad

$$l^a = -\partial_t + \partial_r = -2\partial_u, \qquad n^a = \partial_t + \partial_r = 2\partial_v, \qquad e^a_\theta = \frac{1}{r}\partial_\theta, \qquad e^a_\phi = \frac{1}{r\sin\theta}\partial_\phi,$$

with the metric expressed as

$$\eta_{ab} = -\frac{1}{2}(l_a n_b + n_a l_b) + (e_A)_a (e_A)_b.$$

On $\mathfrak{E}$ we define the variables $\hat{u} = \tau - \zeta$, $\hat{v} = \tau + \zeta$, and the tetrad

$$\hat{l}^a = -\partial_\tau + \partial_\zeta = -2\partial_{\hat{u}}, \quad \hat{n}^a = \partial_\tau + \partial_\zeta = 2\partial_{\hat{v}}, \quad \hat{e}^a_\theta = \frac{1}{\sin\zeta}\partial_\theta, \quad \hat{e}^a_\phi = \frac{1}{\sin\zeta\sin\theta}\partial_\phi,$$

in which the metric $\mathfrak{e}$ takes the form

$$\mathfrak{e}_{ab} = -\frac{1}{2}(\hat{l}_a\hat{n}_b + \hat{n}_a\hat{l}_b) + (\hat{e}_A)_a(\hat{e}_A)_b.$$

The relation between the two tetrads is

$$l^a = \frac{2}{u_+^2}\hat{l}^a, \qquad n^a = \frac{2}{v_+^2}\hat{n}^a, \qquad \hat{e}^a_\theta = \Omega e^a_\theta, \qquad \hat{e}^a_\phi = \Omega e^a_\phi, \tag{5.5.2}$$

where the Minkowski conformal factor is

$$\Omega = \frac{2}{u_+ v_+}.$$

Using the conformal scaling of ϕ, we then immediately deduce that

$$|\phi| \lesssim u_+^{-1} v_+^{-1}.$$

On the other hand, fixing the residual gauge freedom if necessary and using the relations (5.5.2), for the Yang–Mills potential we deduce

$$|A_l| \lesssim u_+^{-2}|\hat{A}_{\hat{l}}| \lesssim u_+^{-2}, \qquad |A_n| \lesssim v_+^{-2}|\hat{A}_{\hat{n}}| \lesssim v_+^{-2},$$

and

$$|A|_{\mathfrak{s}_2} \lesssim \Omega \lesssim u_+^{-1} v_+^{-1}.$$

The above decay rates reproduce the decay rates of Yang and Yu [125], requiring one fewer order of differentiability in the data. However, as a consequence of the use of the conformal method, our results do not apply to the case of arbitrary charge at spatial infinity.

[4] Note that a sufficient condition is that $(\mathbf{a}, \mathbf{e}, \phi_0, \phi_1) \in (H^2_1(\mathbb{R}^3) \times H^1_2(\mathbb{R}^3))^2$, in the notation of [23].

6
Conclusions and Further Work

By finding a suitable gauge, in chapter 3 we constructed a conformal scattering theory for Maxwell potentials. We discovered that an appropriate gauge choice was a modification of the physical temporal gauge, and that it neatly reduced the natural energy on null infinity to a norm in the component $\hat{A}_2$. In specific highly symmetric cases such as flat space, we found that additionally imposing the Coulomb gauge leads to a reduction on null infinity in the form of an ODE. This ODE serves as a reduction of the Maxwell constraint equation, and allows one to recover the component $\hat{A}_1$ from $\hat{A}_2$ on $\mathscr{I}$. Furthermore, we found that using the Morawetz vector field[1] for the scattering theory results in stronger decay conditions on the initial and characteristic data than using the standard static Killing field ∂_t.

In chapter 4 we proved that small data solutions of the Maxwell-scalar field system obey the peeling property and decay exponentially on de Sitter space. By carefully modifying the Coulomb gauge and using the Poincaré inequality, we were able to control all components of the system and close the estimates between $\mathscr{I}^-$ and $\mathscr{I}^+$. In addition, we proved the existence of small data nonlinear scattering operators for each order of differentiability of the data. If the solution was sufficiently smooth, we found that the slowest-decaying component of the scalar field asymptotically decoupled from the charge.

In chapter 5 we proved uniform estimates for solutions of the full Yang–Mills–Higgs equations in temporal gauge on the Einstein cylinder, and extended classical small data existence theorems on Minkowski and de Sitter spacetimes to large data. By conformally embedding Minkowski and de Sitter spacetimes in the Einstein cylinder, we deduced polynomial and exponential large data decay rates.

We have therefore carried out a thorough analysis of the Yang–Mills–Higgs equations on the Einstein cylinder and de Sitter space, proving new large data well-posedness and decay results, as well as constructing conformal scattering operators. We note that working with a positive cosmological constant ensured that solving the system from $\mathscr{I}$ required the resolution of a spacelike Cauchy problem. In the asymptotically flat case, our work

[1] An important feature of the Morawetz vector field which is responsible for the introduction of the angular derivatives in the energy (3.3.35) is the fact that K_0^a is transverse to $\mathscr{I}^+$.

suggests that the temporal gauge—perhaps with a strengthening on the initial surface—might be an appropriate gauge condition for constructing a conformal scattering theory for nonlinear gauge theories such as the Maxwell-scalar field system. There are currently two obstructions to constructing a natural conformal scattering theory for the Maxwell-scalar field system. The first is that solutions for finite energy initial data have not been proven to remain finite on null infinity, even on flat space. One might hope to use techniques that have proven successful in the flat case to extract finite energy well-posedness of the Maxwell-scalar field system on the Einstein cylinder, and therefore deduce finiteness on $\mathscr{I}$. An alternative, simpler, approach might be to use the conformal patching construction of chapter 5. This is currently being addressed in [116]. The strategy is to glue two copies of Minkowski space situated antipodally around the Einstein cylinder, as depicted below.

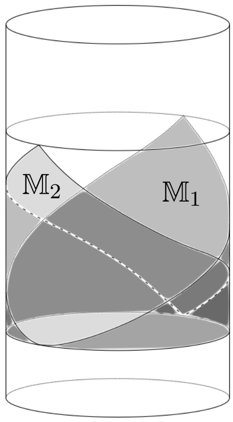

Figure 6.1: We patch two copies of Minkowski space around the Einstein cylinder.

It is then possible to use local existence theorems for finite energy data on Minkowski space [65, 109] to prove that a solution having finite energy data on the Einstein cylinder exists. There are technical issues one needs to address, in particular the cutting-off of data near the spatial infinities of the two copies $\mathbb{M}_{1,2}$ of Minkowski space, and the patching of the gauges of the respective solutions on the overlap $\mathbb{M}_1 \cap \mathbb{M}_2$. This result will further improve on chapters 4 and 5 and allow us to construct a conformal scattering theory for large finite-energy data on de Sitter space.

On Minkowski space, there is the second obstruction of resolving the Goursat problem for finite energy characteristic data for the nonlinear Maxwell-scalar field system. This has to be done without loss of regularity, and existing results for symmetric hyperbolic systems are insufficient [19, 25]. This is currently being investigated.

In summary, we have made good progress towards understanding the conformal scattering of genuinely nonlinear gauge theories, with upcoming results likely to bring the frontiers of QFT and mathematical relativity closer together. Eventually we hope to shed light on aspects of the memory effect and the related infrared divergences mentioned in section 1.1. It will also eventually be of interest to construct (conformal) scattering theories for Einstein's equations, however the situation is much more complex there and will require more work.

Appendices

A
Some Standard Results in PDE Theory

A.1 Elliptic PDEs

A.1.1 The Divergence Theorem

Theorem A.1.1 (The Divergence Theorem)**.** *Let ω be a smooth $(n-1)$-form on a smooth compact n-dimensional orientable manifold $\mathscr{M}$ with boundary $\partial\mathscr{M}$. Then*

$$\int_{\mathscr{M}} \mathrm{d}\omega = \int_{\partial\mathscr{M}} \omega.$$

Remark A.1.2. If the manifold $\mathscr{M}$ is equipped with a metric g, we may rewrite theorem A.1.1 as follows. Let dv be the volume form on $\mathscr{M}$ and let X be a vector field on $\mathscr{M}$. Put $\omega = X \lrcorner \,\mathrm{dv}$. Then by Cartan's identity

$$\mathrm{d}\omega = \mathscr{L}_X \,\mathrm{dv} = (\operatorname{div} X)\,\mathrm{dv},$$

where $\operatorname{div} X = \nabla_a X^a$, and so the divergence theorem reads

$$\int_{\mathscr{M}} (\operatorname{div} X)\,\mathrm{dv} = \int_{\partial\mathscr{M}} X \lrcorner \,\mathrm{dv}\,. \tag{A.1.1}$$

For our purposes (A.1.1) will be a useful intermediary between the fully differential form-theoretic formulation of the divergence theorem as above, and the more basic description of the divergence theorem in terms of normal vector fields, for we will frequently be working with null hypersurfaces.

A.1.2 Elliptic Regularity

Definition A.1.3. Let

$$P = \sum_{|\alpha|\leqslant k} a_\alpha(x)\partial^\alpha$$

be a differential operator of order k with $\mathcal{C}^\infty$ coefficients a_α. We say P is *elliptic* at x_0 if

$$\sigma_P(\xi) = \sum_{|\alpha|=k} a_\alpha(x_0)\xi^\alpha \neq 0$$

for all non-zero $\xi \in \mathbb{R}^n$. We say P is elliptic on a domain $\Omega \subset \mathbb{R}^n$ if $\sigma_P(\xi) \neq 0$ for all non-zero $\xi \in \mathbb{R}^n$ for all $x \in \Omega$.

Theorem A.1.4. *Suppose Ω is an open set in $\mathbb{R}^n$ and P is an elliptic operator of order k with $\mathcal{C}^\infty$ coefficients on Ω. Let ϕ and f be distributions on Ω satisfying*

$$P\phi = f.$$

If $f \in H^s_{\mathrm{loc}}(\Omega)$ for some $s \in \mathbb{R}$, then $\phi \in H^{s+k}_{\mathrm{loc}}(\Omega)$.

A.1.3 Geometry of $\mathbb{S}^3$

The 3-sphere $\mathbb{S}^3$ with its standard metric

$$\mathfrak{s}_3 = \mathrm{d}\zeta^2 + \sin^2\zeta(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)$$

is a maximally symmetric Riemannian manifold. The Ricci and Riemann tensors of $\mathfrak{s}_3$ are expressible entirely in terms of the metric $\mathfrak{s}_3$,

$$\mathrm{R}_{\mu\nu} = -2(\mathfrak{s}_3)_{\mu\nu}$$

and

$$\mathrm{R}_{\mu\nu\rho\sigma} = (\mathfrak{s}_3)_{\rho\nu}(\mathfrak{s}_3)_{\mu\sigma} - (\mathfrak{s}_3)_{\nu\sigma}(\mathfrak{s}_3)_{\mu\rho}.$$

The scalar curvature is therefore $\mathrm{R} = -6$.

Let $*$ denote the Hodge star operator on $\mathbb{S}^3$ and d the exterior derivative on $\mathbb{S}^3$. Let $\mathbf{A}$ be a 1-form and f a function on $\mathbb{S}^3$, and write

$$\begin{aligned}
\operatorname{curl}\mathbf{A} &\coloneqq *\,\mathrm{d}\mathbf{A},\\
\operatorname{div}\mathbf{A} &\coloneqq *\,\mathrm{d}*\mathbf{A},\\
\operatorname{grad} f &\coloneqq \mathrm{d}f.
\end{aligned}$$

It is easy to check that the definitions of $\operatorname{div}\mathbf{A}$ and $\operatorname{grad} f$ coincide with the notions of div and grad in terms of the Levi–Civita connection $\not{\nabla}$ on $\mathbb{S}^3$, that is $\operatorname{div}\mathbf{A} = \not{\nabla}_\mu \mathbf{A}^\mu$ and $(\operatorname{grad} f)_\mu = \not{\nabla}_\mu f$. With these definitions

$$\operatorname{curl}(\operatorname{curl}\mathbf{A}) - \operatorname{grad}(\operatorname{div}\mathbf{A}) = *\,\mathrm{d}*\mathrm{d}\mathbf{A} - \mathrm{d}*\mathrm{d}*\mathbf{A} = \delta\mathrm{d}\mathbf{A} + \mathrm{d}\delta\mathbf{A} \eqqcolon -\not{\Delta}^{(1)}\mathbf{A},$$

where $\delta \coloneqq (-1)^{3k} *\mathrm{d}*$ is the *codifferential* acting on k-forms on $\mathbb{S}^3$ and the operator

$$\begin{aligned}
-\not{\Delta}^{(1)} &: \Gamma(\Lambda^1\mathbb{S}^3) \longrightarrow \Gamma(\Lambda^1\mathbb{S}^3),\\
-\not{\Delta}^{(1)} &\coloneqq \delta\mathrm{d} + \mathrm{d}\delta,
\end{aligned}$$

is the *Hodge Laplacian* on 1-forms on $\mathbb{S}^3$. The operator $\not{\Delta}^{(1)}$ can be extended to act on arbitrary k-forms in the obvious way (giving a sequence of operators $\not{\Delta}^{(k)}$, if one wishes to distinguish between their domains), but it is important to note that if $k \neq 0$ the action of $\not{\Delta}^{(k)}$ differs from the connection Laplacian $\not{\Delta} \coloneqq \not{\nabla}^\mu\not{\nabla}_\mu$ in a way that depends on the degree of the forms it is acting on. The difference is given by the Weitzenböck formula, which in the case of 1-forms is known as Bochner's theorem (see §2.2.2 of [101]).

Theorem A.1.5 (Bochner's Theorem)**.** *Let $(\mathscr{M}, g)$ be a Riemannian manifold with a positive definite metric g and let ∇ be the Levi–Civita connection of g. Considered as operators $\Gamma(\Lambda^1\mathscr{M}) \to \Gamma(\Lambda^1\mathscr{M})$, the Hodge Laplacian $\Delta^{(1)}$ and the connection Laplacian $\Delta = \nabla_\mu\nabla^\mu$ are related by*

$$-\Delta^{(1)} = \Delta + \mathrm{R},$$

where R is the scalar curvature of g.

If $\mathscr{M} = \mathbb{S}^3$, we thus have

$$-\not\Delta^{(1)} = \not\Delta - 6.$$

Now suppose that $\mathbf{A} \in \Gamma(\Lambda^1\mathbb{S}^3)$ satisfies $\operatorname{div}\mathbf{A} = 0$. Then

$$\operatorname{curl}(\operatorname{curl}\mathbf{A}) = -\not\Delta^{(1)}\mathbf{A} = (\not\Delta - 6)\mathbf{A}.$$

Given any $\mathbf{A} \in \Gamma(\Lambda^1\mathbb{S}^3)$, the elliptic equation

$$(\not\Delta - 6)\mathbf{B} = \operatorname{curl}(\operatorname{curl}\mathbf{A}) \tag{A.1.2}$$

on $\mathbb{S}^3$ has a unique solution $\mathbf{B} \in \Gamma(\Lambda^1\mathbb{S}^3)$. This allows us to define the projection onto divergence free 1-forms $\mathcal{P} : \Gamma(\Lambda^1\mathbb{S}^3) \to \Gamma(\Lambda^1\mathbb{S}^3)$,

$$\mathcal{P}\mathbf{A} := (\not\Delta - 6)^{-1}\operatorname{curl}(\operatorname{curl}\mathbf{A}).$$

By construction, for any $\mathbf{A}$ satisfying $\operatorname{div}\mathbf{A} = 0$, $\mathcal{P}\mathbf{A} = \mathbf{A}$, and $\operatorname{div}\mathcal{P}\mathbf{B} = 0$ for any $\mathbf{B}$. This second identity follows by commuting the div operator into the equation (A.1.2). Furthermore, for any function f

$$\mathcal{P}(\operatorname{grad} f) = (\not\Delta - 6)^{-1}(\operatorname{curl}(\operatorname{curl}(\operatorname{grad} f))) = (\not\Delta - 6)^{-1}(\mathbf{0}) = \mathbf{0}.$$

A.2 Hyperbolic PDEs

A.2.1 The Cauchy Problem

In this section we present standard results in hyperbolic PDE theory that are used throughout various parts of this thesis. The proofs are classical and can be found, for example, in [15], [75], or [110].

Definition A.2.1. Let $(\mathscr{M}, g)$ be a smooth Lorentzian manifold and let $E \to \mathscr{M}$ be a real or complex vector bundle. A linear second order differential operator $P : \mathcal{C}^\infty(\mathscr{M}; E) \to \mathcal{C}^\infty(\mathscr{M}; E)$ is said to be *normally hyperbolic* if its principal symbol is given by the metric,

$$\sigma_P(\xi) = g^{ab}\xi_a\xi_b \quad \forall \xi \in T^*\mathscr{M}.$$

That is, in local coordinates on $\mathscr{M}$ and a local trivialization of E the operator P is given by

$$P = g^{ab}\partial_a\partial_b + A^a\partial_a + B,$$

where A^a is a smooth E-valued vector field on $\mathscr{M}$ and B is a smooth E-valued function on $\mathscr{M}$.

Theorem A.2.2. *Let $\mathscr{M}$ be a smooth globally hyperbolic Lorentzian manifold and let Σ be a spacelike Cauchy surface in $\mathscr{M}$. Let T be the future-directed timelike unit normal to Σ and let $E \to \mathscr{M}$ be a vector bundle over $\mathscr{M}$. Let P be a smooth linear normally hyperbolic operator acting on sections of E. Then for each $\phi_0, \phi_1 \in \mathcal{C}_c^\infty(\Sigma; E)$ and each $f \in \mathcal{C}_c^\infty(\mathscr{M}; E)$ there exists a unique $\phi \in \mathcal{C}^\infty(\mathscr{M}; E)$ satisfying*

$$P\phi = f, \qquad (\phi, \nabla_T \phi)|_\Sigma = (\phi_0, \phi_1).$$

Moreover, the data propagates at finite speed in the sense that $\operatorname{supp}(\phi) \subset J(K)$*, where*

$$K = \operatorname{supp}(\phi_0) \cup \operatorname{supp}(\phi_1) \cup \operatorname{supp}(f).$$

Furthermore, the map

$$\begin{aligned} \mathcal{C}_c^\infty(\Sigma; E) \times \mathcal{C}_c^\infty(\Sigma; E) \times \mathcal{C}_c^\infty(\mathscr{M}; E) &\longrightarrow \mathcal{C}^\infty(\mathscr{M}; E); \\ (\phi_0, \phi_1, f) &\longmapsto \phi \end{aligned}$$

is linear and continuous.

Theorem A.2.3. *Let the setup be as in theorem A.2.2 and let $s \in \mathbb{N}_0$. Let $\{\Sigma_t\}_{t\in\mathbb{R}}$, be a foliation of $\mathscr{M}$ by smooth spacelike hypersurfaces such that $\Sigma_0 = \Sigma$. Then for each $(\phi_0, \phi_1) \in H^{s+1}(\Sigma; E) \oplus H^s(\Sigma; E)$ and $f \in L^1([0,T]; H^s(\Sigma; E)$ there exists a unique*

$$\phi \in \mathcal{C}^0([0,T]; H^{s+1}(\Sigma_t; E)) \cap \mathcal{C}^1([0,T]; H^s(\Sigma_t; E)) =: E_{s+1}((0,T) \times \Sigma_t; E)$$

satisfying

$$P\phi = f \text{ in } (0,T) \times \Sigma_t, \qquad (\phi, \nabla_T \phi)|_\Sigma = (\phi_0, \phi_1).$$

As before, the data propagates at finite speed, $\operatorname{supp}(\phi) \subset J(K)$*, and the map*

$$\begin{aligned} H^{s+1}(\Sigma; E) \times H^s(\Sigma; E) \times L^1([0,T]; H^s(\Sigma; E)) &\longrightarrow E_{s+1}((0,T) \times \Sigma_t; E) \\ (\phi_0, \phi_1, f) &\longmapsto \phi \end{aligned}$$

is linear and continuous.

A.2.2 The Goursat Problem

As resolving the Goursat problem is a key step in the construction of a conformal scattering theory from $\mathscr{I}^-$ to $\mathscr{I}^+$ in the zero cosmological constant case (chapter 3), in this section we give a brief overview of Hörmander's resolution of the linear characteristic Cauchy problem [58], including the proof. In fact, Hörmander resolves a much wider class of initial value problems termed *weakly spacelike* Cauchy problems. We work on a smooth globally hyperbolic Lorentzian manifold $\tilde{X}$ of the form $\tilde{X} = \mathbb{R} \times X$, where X is smooth, compact, and has dimension 3. We suppose that the metric on $\tilde{X}$ is

$$\tilde{g} = \mathrm{d}t^2 - g_t,$$

where g_t is a Riemannian metric on X which in local coordinates is given by

$$g_t = g_{jk}(t,x)\, \mathrm{d}x^j\, \mathrm{d}x^k.$$

We assume that the metrics g_t are uniformly equivalent to g_0 and that the manifolds (X, g_t) have a positive radius of injectivity (at least for t belonging to any compact interval). We will frequently drop the subscript t and simply write $g(t,x)$ to denote $g_t(t,x)$. Let $\widetilde{\mathrm{dv}}$ be the spacetime volume form associated to $\tilde{g}$, and let dv be the smooth volume form on X induced by g_0. Since g_t is uniformly equivalent to g_0, the volume form dv_t associated to g_t is uniformly equivalent to dv. We denote by $\Box$ the wave operator on $\tilde{X}$, $\Box = \nabla^a\nabla_a = \tilde{g}^{ab}\nabla_a\nabla_b$, where ∇ is the Levi–Civita connection of $\tilde{g}$. We also denote by $\not\nabla$ the Levi–Civita connection of g_t, and by $\not\Delta$ the Laplace–Beltrami operator associated to g_t.

Consider the hypersurface $\Sigma = \{t = \varphi(x)\}$ in $\tilde{X}$. The co-normal to Σ is $n^\flat = \mathrm{d}t - (\partial_j\varphi)\,\mathrm{d}x^j$, so that the normal is

$$n = \partial_t + (\partial_j\varphi)g^{jk}(\varphi(x), x)\partial_k$$

and has norm $n^a n_a = 1 - g^{jk}(\varphi(x), x)\partial_j\varphi(x)\partial_k\varphi(x) = 1 - |\not\nabla\varphi|^2(\varphi(x), x)$.

Definition A.2.4. Let $\varphi : X \to \mathbb{R}$ be a Lipschitz continuous function X. Then φ is differentiable almost everywhere, and we say that the hypersurface $\Sigma = \{(\varphi(x), x) : x \in X\} \subset \tilde{X}$ is *weakly spacelike* if

$$\inf_X \left(1 - |\not\nabla\varphi|^2\right) \geqslant 0, \tag{A.2.1}$$

where $|\not\nabla\varphi|^2 = g^{jk}(\varphi(x), x)\partial_j\varphi(x)\partial_k\varphi(x)$. We say Σ is *spacelike* if $1 - |\not\nabla\varphi|^2 > 0$ almost everywhere, *uniformly spacelike* if $\inf_X (1 - |\not\nabla\varphi|^2) > 0$, and *characterstic* if $|\not\nabla\varphi|^2 = 1$ almost everywhere.

Consider the equation

$$\Box\phi + L_1\phi = f, \tag{A.2.2}$$

where L_1 is a linear first order differential operator with smooth coefficients and f is at least $L^1_{\mathrm{loc}}(\mathbb{R}; L^2(X))$. By Theorem A.2.3, (A.2.2) has a unique solution

$$\phi \in \mathcal{C}^0(\mathbb{R}; H^1(X)) \cap \mathcal{C}^1(\mathbb{R}; L^2(X)) = E_1(\mathbb{R} \times X)$$

for finite energy Cauchy data (ϕ_0, ϕ_1) given on X. For the sake of brevity, we consider the homogeneous equation

$$\Box\phi + L_1\phi = 0. \tag{A.2.3}$$

Let $\mathscr{E}$ be the subspace of the Banach space $E_1(\mathbb{R} \times X)$ of solutions to (A.2.3). We show that the finite energy Goursat problem can be resolved to get a unique solution ϕ in the finite energy space $\mathscr{E}$.

Theorem A.2.5 (Hörmander 1990, [58]). *Let $\Sigma = \{(\varphi(x), x) : x \in X\}$ be a weakly spacelike hypersurface in $\tilde{X}$ and let $\mathrm{dv}^0_\Sigma = (1 - |\not\nabla\varphi|^2)\,\mathrm{dv}_\Sigma$, where dv_Σ is the volume form induced on Σ by the spacetime volume form $\widetilde{\mathrm{dv}}$. Then there exists an isomorphism*[1] *of Hilbert spaces*

$$\begin{aligned} \mathscr{E} &\longrightarrow H^1(\Sigma) \times L^2(\Sigma, \mathrm{dv}^0_\Sigma); \\ \phi &\longmapsto (\phi, \partial_t\phi). \end{aligned} \tag{A.2.4}$$

[1] Isomorphism here means a continuous linear bijection whose inverse is also continuous.

Remark A.2.6. Note that when Σ is characteristic, $\mathrm{dv}^0_\Sigma \equiv 0$, so the space $L^2(\Sigma, \mathrm{dv}^0_\Sigma)$ is of dimension zero. Nothing is stated about $\partial_t\phi$ in this case, and simply the values of ϕ on Σ are sufficient to recover the full solution $\phi \in \mathscr{E}$. These values $\phi|_\Sigma$ are the *characteristic data* for the wave equation (A.2.3).

Proof. Consider the stress-energy tensor

$$\mathbf{T}_{ab} = \nabla_a\phi\nabla_b\phi - \frac{1}{2}\tilde{g}_{ab}\nabla_c\phi\nabla^c\phi + \frac{1}{2}\tilde{g}_{ab}\phi^2$$

for the equation (A.2.3). As a consequence of (A.2.3), it satisfies the approximate conservation law

$$\begin{aligned}\nabla^a\mathbf{T}_{ab} &= (\Box\phi + \phi)\nabla_b\phi \\ &= L_1'\phi\nabla_b\phi,\end{aligned}$$

where $L_1' = -L_1 + 1$. Let $T^a = \partial_t$ be the unit normal to X in $\tilde{X}$ and integrate $\nabla^a(\mathbf{T}_{ab}T^b)$ over the region $B = \{(t,x) \,:\, T \leqslant t \leqslant \varphi(x)\}$ for some $T < \min\varphi(x)$:

$$\int_B \nabla^a(\mathbf{T}_{ab}T^b)\,\widetilde{\mathrm{dv}} = -\int_{t=T}\mathbf{T}_{ab}T^aT^b\,\mathrm{dv}_t(x) + \int_\Sigma \mathbf{T}_{ab}T^an^b\,\mathrm{dv}_\Sigma\,.$$

We have $\mathbf{T}_{ab}T^aT^b = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}|\nabla\!\!\!/\phi|^2 + \frac{1}{2}\phi^2$, and

$$\begin{aligned}\mathbf{T}_{ab}T^an^b &= \frac{1}{2}\dot{\phi}^2 + \dot{\phi}\nabla\!\!\!/\varphi\cdot\nabla\!\!\!/\phi + \frac{1}{2}|\nabla\!\!\!/\phi|^2 + \frac{1}{2}\phi^2 \\ &= \frac{1}{2}\dot{\phi}^2(1 - |\nabla\!\!\!/\varphi|^2) + \frac{1}{2}|\dot{\phi}\nabla\!\!\!/\varphi + \nabla\!\!\!/\phi|^2 + \frac{1}{2}\phi^2,\end{aligned}$$

where $\dot{\phi}$ denotes $\nabla_T\phi$ and $\nabla\!\!\!/\phi$ is the gradient of ϕ along directions tangent to X so that $|\nabla\!\!\!/\phi|^2 = g^{jk}(t,x)\partial_j\phi(x)\partial_k\phi(x)$ and $\nabla\!\!\!/\varphi\cdot\nabla\!\!\!/\phi = g^{jk}(t,x)\partial_j\varphi(x)\partial_k\phi(x)$. Notice that $\nabla\!\!\!/^{\|}\phi := \nabla\!\!\!/(\phi(\varphi(x),x)) = \dot{\phi}\nabla\!\!\!/\varphi + \nabla\!\!\!/\phi$ is the gradient of ϕ on Σ, so, denoting

$$\mathcal{E}(t) = \frac{1}{2}\int_{X\times\{t\}}\left(\dot{\phi}^2 + |\nabla\!\!\!/\phi|^2 + \phi^2\right)\mathrm{dv}_t(x),$$

altogether we have

$$\frac{1}{2}\int_\Sigma \dot{\phi}^2\,\mathrm{dv}^0_\Sigma + \frac{1}{2}\int_\Sigma\left(|\nabla\!\!\!/^{\|}\phi|^2 + \phi^2\right)\mathrm{dv}_\Sigma = \mathcal{E}(T) + \int_B \nabla^a(\mathbf{T}_{ab}T^b)\,\widetilde{\mathrm{dv}}.$$

Now $\nabla^a(\mathbf{T}_{ab}T^b) = (\nabla^a\mathbf{T}_{ab})T^b + \mathbf{T}_{ab}\nabla^{(a}T^{b)} = (L_1'\phi)\dot{\phi} + \mathbf{T}_{ab}\nabla^{(a}T^{b)}$, so one has the bound

$$|\nabla^a(\mathbf{T}_{ab}T^b)| \leqslant C\left(\dot{\phi}^2 + |\nabla\!\!\!/\phi|^2 + \phi^2\right),$$

where C depends on the L^∞ norms of the coefficients of L_1' and $\nabla^{(a}T^{b)}$. Consequently

$$\frac{1}{2}\int_\Sigma \dot{\phi}^2\,\mathrm{dv}^0_\Sigma + \frac{1}{2}\int_\Sigma\left(|\nabla\!\!\!/^{\|}\phi|^2 + \phi^2\right)\mathrm{dv}_\Sigma \lesssim \sup_{t\in I}\mathcal{E}(t) \lesssim \|\phi\|_{\mathscr{E}}, \tag{A.2.5}$$

where I is an interval containing $[T, \max\varphi]$. Estimate (A.2.5) shows that the forward map (A.2.4) is continuous. To show that the map (A.2.4) is an isomorphism it remains to prove that an estimate opposite to (A.2.5) holds, and prove surjectivity. To prove the opposite estimate, integrate $\nabla^a(\mathbf{T}_{ab}T^b)$ over the region $C_t = \{(s,x) : \max\{\varphi(x), T\} \leqslant s \leqslant t\}$ for some $T > 0$: this gives

$$\int_{C_t} \nabla^a(\mathbf{T}_{ab}T^b)\,\widetilde{\mathrm{dv}} = \mathcal{E}(t) - \int_{\Sigma_T} \mathbf{T}_{ab}T^a n^b\,\mathrm{dv}_\Sigma - \mathcal{E}(T)^c,$$

where $\Sigma_T = \Sigma \cap \{\varphi(x) \geqslant T\}$ and $\mathcal{E}(T)^c$ is the integral of $\mathbf{T}_{ab}T^aT^b$ over $\{x \in X \times \{T\} : \varphi(x) \leqslant T\}$. Choosing $T < \min\varphi$ so that $\Sigma_T = \Sigma$ and $\mathcal{E}(T)^c = 0$, we have

$$\mathcal{E}(t) = \int_{C_t} \nabla^a(\mathbf{T}_{ab}T^b)\,\widetilde{\mathrm{dv}} + \underbrace{\int_\Sigma \mathbf{T}_{ab}T^a n^b\,\mathrm{dv}_\Sigma}_{=:M} \leqslant C\int_T^t \mathcal{E}(s)\,\mathrm{d}s + M.$$

From Grönwall's lemma it follows that $\mathcal{E}(t) \leqslant \mathrm{e}^{C(t-T)}M$, which gives

$$\|\phi\|^2_{\mathscr{E}} \lesssim \frac{1}{2}\int_\Sigma \dot{\phi}^2\,\mathrm{dv}^0_\Sigma + \frac{1}{2}\int_\Sigma \left(|\nabla\!\!\!\!/^{\,\|}\phi|^2 + \phi^2\right)\mathrm{dv}_\Sigma, \tag{A.2.6}$$

where the norm $\|\cdot\|_{\mathscr{E}}$ is understood to be taken over a compact interval I in time which contains $[\min\varphi, \max\varphi]$. Together, (A.2.5) and (A.2.6) say that

$$\|\phi\|^2_{\mathscr{E}} \simeq \|\dot{\phi}\|^2_{L^2(\Sigma,\mathrm{dv}^0_\Sigma)} + \|\phi\|^2_{H^1(\Sigma)}.$$

This implies that the map (A.2.4) is an isomorphism from $\mathscr{E}$ onto its image, which is a closed subspace of $H^1(\Sigma) \oplus L^2(\Sigma, \mathrm{dv}^0_\Sigma)$. To prove that the map (A.2.4) is a bijection it remains to prove surjectivity. For a smooth spacelike surface Σ this is classical (theorem A.2.3). We extend this to arbitrary uniformly spacelike surfaces by approximation with smooth ones, and then pass to arbitrary weakly spacelike surfaces by approximating the equation (A.2.3) with equations for which they are uniformly spacelike. This original method of Hörmander [58] is equivalent to a similar method used by Mason and Nicolas [78], where they instead view this latter approximation as slowing down the speed of propagation of the equation (A.2.3). We shall first need a lemma.

Lemma A.2.7 (Hörmander 1990, [58])**.** *Let $\varphi : X \to \mathbb{R}$ be Lipschitz. If the graph of φ is uniformly spacelike, then there is a sequence $\varphi_\mu \in \mathcal{C}^\infty(X)$ converging uniformly to φ such that for some $\lambda < 1$*

$$g^{jk}(\varphi(x), x)\partial_j\varphi_\mu(x)\partial_k\varphi_\mu(x) < \lambda.$$

Furthermore, $\nabla\!\!\!\!/\,\varphi_\mu \to \nabla\!\!\!\!/\,\varphi$ almost everywhere.

Since we are working with finite energy solutions of (A.2.3), they do not possess enough derivatives to be interpreted classically. For this reason we need a weak formulation of (A.2.3). For any $\chi \in \mathcal{C}^\infty_c(\tilde{X})$ and $\phi \in \mathcal{C}^\infty(\tilde{X})$ integration by parts gives

$$\iint_{t\leqslant\varphi(x)} \chi\Box\phi\,\widetilde{\mathrm{dv}} = \iint_{t\leqslant\varphi(x)} \phi\Box\chi\,\widetilde{\mathrm{dv}} + \int_\Sigma \chi\nabla_n\phi\,\mathrm{dv}_\Sigma - \int_\Sigma \phi\nabla_n\chi\,\mathrm{dv}_\Sigma,$$

where $\nabla_n = n^a\nabla_a = \partial_t + \nabla\!\!\!\!/\,\varphi\cdot\nabla\!\!\!\!/$. We thus have

$$\iint_{t\leqslant\varphi(x)} \chi\Box\phi\,\widetilde{\mathrm{dv}} = \iint_{t\leqslant\varphi(x)} \phi\Box\chi\,\widetilde{\mathrm{dv}} + \int_\Sigma \chi\dot{\phi}\,\mathrm{dv}^0_\Sigma + \int_\Sigma \chi\nabla\!\!\!\!/\,\varphi\cdot\nabla\!\!\!\!/(\phi|_\Sigma)\,\mathrm{dv}_\Sigma - \int_\Sigma \phi(\dot{\chi} + \nabla\!\!\!\!/\,\varphi\cdot\nabla\!\!\!\!/\,\chi)\,\mathrm{dv}_\Sigma\,.$$

Similarly, if L_1^* denotes the adjoint of $L_1 = c\partial_t + \mathbf{b}\cdot\nabla\!\!\!/ + b_0$ with respect to $\widetilde{\mathrm{dv}}$, then

$$\iint_{t\leqslant\varphi(x)} \chi L_1\phi\,\widetilde{\mathrm{dv}} = \iint_{t\leqslant\varphi(x)} \phi L_1^*\chi + \int_\Sigma \phi\chi\underbrace{(c+\mathbf{b}\cdot n)}_{=:b}\,\mathrm{dv}_\Sigma\,.$$

Definition A.2.8. We say $\phi\in\mathscr{E}$ is a weak solution of (A.2.3) with data $(\phi,\dot\phi)|_\Sigma \in H^1(\Sigma)\oplus L^2(\Sigma,\mathrm{dv}_\Sigma^0)$ on a weakly spacelike surface $\Sigma = \{(\varphi(x),x)\}$ if for all $\chi\in\mathcal{C}_c^\infty(\tilde X)$

$$\begin{aligned}&\iint_{t\leqslant\varphi(x)} \phi(\Box\chi + L_1^*\chi)\,\widetilde{\mathrm{dv}} + \int_\Sigma \chi\dot\phi\,\mathrm{dv}_\Sigma^0 \\ &+\int_\Sigma \chi\nabla\!\!\!/\varphi\cdot\nabla\!\!\!/(\phi|_\Sigma)\,\mathrm{dv}_\Sigma - \int_\Sigma \phi(\dot\chi + \nabla\!\!\!/\varphi\cdot\nabla\!\!\!/\chi)\,\mathrm{dv}_\Sigma + \int_\Sigma b\chi\phi\,\mathrm{dv}_\Sigma = 0.\end{aligned} \tag{A.2.7}$$

Equipped with the weak formulation (A.2.7), we may now proceed to prove the surjectivity of the map (A.2.4).

Surjectivity: the Uniformly Spacelike Case

Let φ be uniformly spacelike and Lipshitz, let $\phi_0,\phi_1\in\mathcal{C}^\infty(X)$ be given, and pick φ_μ smooth and uniformly spacelike as in lemma A.2.7. By theorem A.2.3, for each μ there exists a solution $\phi_\mu\in\mathcal{C}^\infty(\tilde X)$ satisfying (A.2.3) such that

$$\phi_\mu|_{\Sigma_\mu} = \phi_0 \quad\text{and}\quad \dot\phi_\mu|_{\Sigma_\mu} = \phi_1.$$

The estimate (A.2.6) shows that

$$\|\phi_\mu\|_{\mathscr{E}}^2 \lesssim \left(\|\phi_0\|_{H^1(X)}^2 + \|\phi_1\|_{L^2(X)}^2\right),$$

so the sequence ϕ_μ is uniformly bounded in $\mathscr{E}$.

Lemma A.2.9. *Up to a subsequence, one has that*

$$\phi_\mu \longrightarrow \phi \ \ \text{in}\ \ \mathcal{C}^0(I;L^2(X)),$$

where I is a compact interval containing $[\min\varphi,\max\varphi]$.

Proof. Since ϕ_μ is uniformly bounded in $\mathscr{E}$, we have in particular that ϕ_μ is uniformly bounded in $\mathcal{C}^0(I;H^1(X))\subset\mathcal{C}^0(I;L^2(X))$. Moreover, it is also uniformly bounded in $\mathcal{C}^1(I;L^2(X))$, which means that for any $t,s\in I$

$$\phi_\mu(t)-\phi_\mu(s) = \int_s^t \dot\phi_\mu\,\mathrm{d}t,$$

and so

$$\begin{aligned}\|\phi_\mu(t)-\phi_\mu(s)\|_{L^2(X)} &\leqslant |t-s|\sup_t\|\dot\phi_\mu\|_{L^2(X)}\\ &\lesssim |t-s|\|\phi_\mu\|_{\mathscr{E}}\\ &\lesssim |t-s|,\end{aligned}$$

so that ϕ_μ is equicontinuous in t. By Arzelá–Ascoli (see for example theorem 11.28 of [102]), $\{\phi_\mu\}_\mu\subset\mathcal{C}^0(I;L^2(X))$ is compact in the compact-open topology on $\mathcal{C}^0(I;L^2(X))$; since the interval I is compact, it follows that ϕ_μ has a subsequence that converges in $\mathcal{C}^0(I;L^2(X))$, whose limit we call ϕ. ∎

This implies that for all $t \in I$ the sequence $\phi_\mu(t,\cdot)$ converges to $\phi(t,\cdot)$ strongly in $L^2(X)$, so in particular has a subsequence that converges almost everywhere on X. That is, taking a further subsequence if necessary, we also know that $\phi_\mu \to \phi$ almost everywhere on $I \times X$. Moreover, since ϕ_μ is uniformly bounded in $\mathscr{E}$, Fatou's Lemma implies that the limit ϕ has $\|\phi\|_{\mathscr{E}} < \infty$, i.e. $\phi \in \mathscr{E}$.

Now consider the sequence $\phi_\mu(\varphi(x),x)$. By (A.2.5), one knows that smooth solutions ϕ to (A.2.3) satisfy $\|\phi|_\Sigma\|_{H^1(\Sigma)} \lesssim \|\phi\|_{\mathscr{E}}$, so

$$\|\phi_\mu(\varphi(x),x)\|_{H^1(X)} \simeq \|\phi_\mu|_\Sigma\|_{H^1(\Sigma)} \lesssim \|\phi_\mu\|_{\mathscr{E}} \leqslant C(\phi_0,\phi_1).$$

The sequence $\phi_\mu(\varphi(x),x)$ is therefore uniformly bounded in $H^1(X)$, and we can extract a weakly convergent subsequence in $H^1(X)$, the limit of which as we have seen must be $\phi(\varphi(x),x)$,

$$\phi_\mu(\varphi(\cdot),\cdot) \rightharpoonup \phi(\varphi(\cdot),\cdot) \text{ in } H^1(X).$$

By Rellich–Kondrachov, $\phi_\mu(\varphi(\cdot),\cdot)$ converges to $\phi(\varphi(\cdot),\cdot)$ strongly in $L^2(X)$. But we have

$$\phi_\mu(\varphi(x),x) - \phi_0(x) = \int_{\varphi_\mu(x)}^{\varphi(x)} \dot{\phi}_\mu(t,x)\,\mathrm{d}t,$$

so

$$\begin{aligned}\int_X |\phi_\mu(\varphi(x),x) - \phi_0(x)|^2\,\mathrm{dv}(x) &= \int_X \left|\int_{\varphi_\mu(x)}^{\varphi(x)} \dot{\phi}_\mu(t,x)\,\mathrm{d}t\right|^2 \mathrm{dv}(x)\\ &\leqslant \int_X |\varphi(x)-\varphi_\mu(x)| \int_{\varphi_\mu(x)}^{\varphi(x)} |\dot{\phi}_\mu(t,x)|^2\,\mathrm{d}t\,\mathrm{dv}(x)\\ &\lesssim \|\phi_\mu\|_{\mathscr{E}}^2 \|\varphi-\varphi_\mu\|_{L^\infty(X)} \longrightarrow 0,\end{aligned}$$

and hence $\phi|_\Sigma(x) = \phi(\varphi(x),x) = \phi_0(x)$. Now apply (A.2.7) to ϕ_μ and φ_μ. We have

$$\begin{aligned}\iint_{t\leqslant\varphi_\mu(x)} \phi_\mu(\Box\chi + L_1^*\chi)\,\widetilde{\mathrm{dv}} = &-\int_{\Sigma_\mu} \chi\phi_1\,\mathrm{dv}^0_{\Sigma_\mu} - \int_{\Sigma_\mu} \chi\bar{\nabla}\varphi_\mu\cdot\bar{\nabla}\phi_0\,\mathrm{dv}_\Sigma\\ &+\int_{\Sigma_\mu} \phi_0(\dot{\chi}+\bar{\nabla}\varphi_\mu\cdot\bar{\nabla}\chi)\,\mathrm{dv}_\Sigma - \int_{\Sigma_\mu} b\chi\phi_0\,\mathrm{dv}_\Sigma,\end{aligned}$$

where we have noted that since $n_\mu = \partial_t + (\bar{\nabla}\varphi_\mu)^j\partial_j$, we can choose the measures dv_{Σ_μ} to all be equal to dv_Σ, for example by choosing the vector fields l_μ, $(l_\mu)^a(n_\mu)_a = 1$, defining $\mathrm{dv}_{\Sigma_\mu} = l_\mu \lrcorner \widetilde{\mathrm{dv}}$ to be equal to ∂_t for all μ. Note that the measure $\mathrm{dv}^0_{\Sigma_\mu}$ is then given by $(1-|\bar{\nabla}\varphi_\mu|^2)\,\mathrm{dv}_\Sigma$. Subtracting the above identity from (A.2.7), we get

$$\begin{gathered}\iint_{t\leqslant\varphi_\mu(x)} \phi_\mu(\Box\chi + L_1^*\chi)\,\widetilde{\mathrm{dv}} - \iint_{t\leqslant\varphi(x)} \phi(\Box\chi + L_1^*\chi)\,\widetilde{\mathrm{dv}}\\ +\int_{\Sigma_\mu} \chi\bar{\nabla}\varphi_\mu\cdot\bar{\nabla}\phi_0\,\mathrm{dv}_\Sigma - \int_\Sigma \chi\bar{\nabla}\varphi\cdot\bar{\nabla}\phi_0\,\mathrm{dv}_\Sigma\\ +\int_\Sigma \phi(\dot{\chi}+\bar{\nabla}\varphi\cdot\bar{\nabla}\chi)\,\mathrm{dv}_\Sigma - \int_{\Sigma_\mu} \phi_0(\dot{\chi}+\bar{\nabla}\varphi_\mu\cdot\bar{\nabla}\chi)\,\mathrm{dv}_\Sigma\\ +\int_{\Sigma_\mu} b\chi\phi_0\,\mathrm{dv}_\Sigma - \int_\Sigma b\chi\phi\,\mathrm{dv}_\Sigma\\ =\\ \int_\Sigma \chi\dot{\phi}\,\mathrm{dv}^0_\Sigma - \int_{\Sigma_\mu} \chi\phi_1\,\mathrm{dv}^0_{\Sigma_\mu}.\end{gathered} \tag{A.2.8}$$

We show that the pairs of terms arranged into separate lines in (A.2.8) go to zero as $\mu \to \infty$, with the exception of the last pair. Denoting $P^*\chi = \Box\chi + L_1^*\chi$, the first line in (A.2.8) is

$$\iint_{t\leqslant\varphi_\mu} \phi_\mu P^*\chi \,\widetilde{\mathrm{dv}} - \iint_{t\leqslant\varphi} \phi P^*\chi \,\widetilde{\mathrm{dv}} = \int_{t\leqslant\varphi_\mu} (\phi_\mu - \phi) P^*\chi \,\widetilde{\mathrm{dv}} + \iint_{\varphi_\mu\leqslant t\leqslant\varphi} \phi P^*\chi \,\widetilde{\mathrm{dv}},$$

where

$$\begin{aligned}\left|\iint_{\varphi_\mu\leqslant t\leqslant\varphi} \phi P^*\chi \,\widetilde{\mathrm{dv}}\right| &\leqslant \iint_{\varphi_\mu\leqslant t\leqslant\varphi} |\phi P^*\chi| \,\widetilde{\mathrm{dv}} \\ &\lesssim \|\varphi_\mu - \varphi\|_{L^\infty(X)} \sup_t \|\phi P^*\chi\|_{L^1(X)}(t) \\ &\lesssim \|\varphi_\mu - \varphi\|_{L^\infty(X)} \|\phi\|_{\mathscr{E}}\end{aligned}$$

and

$$\left|\iint_{t\leqslant\varphi_\mu} (\phi_\mu - \phi) P^*\chi \,\widetilde{\mathrm{dv}}\right| \lesssim \sup_t \|\phi_\mu - \phi\|_{L^1(X)}(t) \lesssim \|\phi_\mu - \phi\|_{C^0(I;L^2(X))}.$$

Thus the first pair of terms tends to zero as $\mu \to \infty$ by virtue of lemma A.2.9 and the fact that φ_μ converges to φ uniformly. The second, third and fourth pairs of terms can be dealt with using the dominated convergence theorem. Indeed, for example

$$\begin{aligned}&\left|\int_{\Sigma_\mu} \phi_0(\dot\chi + \nabla\!\!\!\!/\,\varphi_\mu \cdot \nabla\!\!\!\!/\,\chi)\,\mathrm{dv}_\Sigma - \int_\Sigma \phi(\dot\chi + \nabla\!\!\!\!/\,\varphi \cdot \nabla\!\!\!\!/\,\chi)\,\mathrm{dv}_\Sigma\right| \\ &\leqslant \left|\int_{\Sigma_\mu} \phi(\dot\chi + \nabla\!\!\!\!/\,\varphi_\mu \cdot \nabla\!\!\!\!/\,\chi)\,\mathrm{dv}_\Sigma - \int_\Sigma \phi(\dot\chi + \nabla\!\!\!\!/\,\varphi_\mu \cdot \nabla\!\!\!\!/\,\chi)\,\mathrm{dv}_\Sigma\right| \\ &+ \left|\int_{\Sigma} \phi(\dot\chi + \nabla\!\!\!\!/\,\varphi_\mu \cdot \nabla\!\!\!\!/\,\chi)\,\mathrm{dv}_\Sigma - \int_\Sigma \phi(\dot\chi + \nabla\!\!\!\!/\,\varphi \cdot \nabla\!\!\!\!/\,\chi)\,\mathrm{dv}_\Sigma\right|,\end{aligned}$$

and since $\varphi_\mu \to \varphi$ almost everywhere, $\phi \in \mathcal{C}^0(I; H^1(X))$, and the sequence φ_μ is bounded in $W^{1,\infty}(X)$ (since φ Lipschitz), the dominated convergence theorem shows that both terms on the right got to zero as $\mu \to \infty$. The other terms are similar, and so taking the limit $\mu \to \infty$ in (A.2.8) shows that

$$\int_\Sigma \chi(\dot\phi - \phi_1)\,\mathrm{dv}^0_\Sigma = 0.$$

In other words, $\dot\phi|_\Sigma = \phi_1$ dv^0_Σ-almost everywhere. This completes the proof of surjectivity in the uniformly spacelike case.

Surjectivity: the Weakly Spacelike Case

Now suppose $\Sigma = \{(\varphi(x), x) : x \in X\}$ is weakly spacelike. If $c^2 < 1$, then Σ is uniformly spacelike with respect to the modified equation

$$\Box_{c^2}\phi + L_1\phi = 0,$$

where $\Box_{c^2}$ is the wave operator associated to $g_{c^2} = \mathrm{d}t^2 - c^2 g_t$. Given $\phi_0, \phi_1 \in \mathcal{C}^\infty(X)$, we have proved that there is a unique solution ϕ of this equation with $\phi|_\Sigma = \phi_0$ and $\dot\phi|_\Sigma = \phi_1$.

Furthermore, the energy of ϕ is uniformly bounded as $c^2 \to 1$, since the definition of energy is independent of c^2 up to equivalence. Thus we can find sequences

$$c_\mu^2 \longrightarrow 1$$

and, as in Lemma A.2.9,

$$\phi_\mu \longrightarrow \phi \text{ in } \mathcal{C}^0(I; L^2(X))$$

(and also weakly in $L^2(I; H^1(X))$ and almost everywhere on $I \times X$), such that

$$\Box_{c_\mu^2} \phi_\mu + L_1 \phi_\mu = 0$$

with $\phi_\mu|_\Sigma = \phi_0$ and $\dot{\phi}_\mu|_\Sigma = \phi_1$ (note that since ϕ_0, ϕ_1 are smooth, so are the ϕ_μ and restriction to Σ is well-defined). The energy $\|\phi_\mu\|^2_{\mathscr{E}}$ of ϕ_μ is uniformly bounded, so as before

$$\phi \in \mathcal{C}^0(I; H^1(X)) \cap \mathcal{C}^1(I; L^2(X)).$$

To see that ϕ satisfies the equation in the weak sense, we only need to take the limit as $\mu \to \infty$ in

$$\begin{aligned}\iint_{t \leqslant \varphi(x)} \phi_\mu P^* \chi \, \widetilde{\mathrm{dv}} = -\int_\Sigma \chi \dot{\phi}_\mu \, \mathrm{dv}^0_\Sigma - \int_\Sigma \chi \bar{\nabla} \varphi \cdot \bar{\nabla}(\phi_\mu|_\Sigma) \, \mathrm{dv}_\Sigma \\ + \int_\Sigma \phi_\mu (\dot{\chi} + \bar{\nabla}\varphi \cdot \bar{\nabla}\chi) \, \mathrm{dv}_\Sigma - \int_\Sigma b \chi \phi_\mu \, \mathrm{dv}_\Sigma \,.\end{aligned}$$

The same arguments as in the case of a uniformly spacelike Σ go through here to take the limit, and we get that ϕ satisfies (A.2.3) weakly with data $(\phi, \dot{\phi})|_\Sigma = (\phi_0, \phi_1)$ on Σ, where ϕ_1 is defined almost everywhere with respect to dv^0_Σ. □

The above proof is a slightly modified version of the original proof by Hörmander, in which the metric $\tilde{g}$ and the coefficients of the first order operator L_1 are assumed to be smooth. These regularity assumptions are far from sharp, as noted (but not proven) by Hörmander himself in his original paper [58]. Sharper versions of Theorem A.2.5 have been obtained by Nicolas.

Theorem A.2.10 (Nicolas 2007, [90])**.** *The uniformly spacelike Cauchy problem for* (A.2.3) *has a unique solution in* $\mathscr{E}$ *for* $(\phi_0, \phi_1) \in H^1(\Sigma) \oplus L^2(\Sigma)$ *as long as* $\tilde{g} \in \mathcal{C}^0(\tilde{X}) \cap W^{1,\infty}_{\mathrm{loc}}(\tilde{X})$ *and the coefficients of* L_1 *are in* $L^\infty_{\mathrm{loc}}(\tilde{X})$*.*

For the fully characteristic Cauchy problem, Nicolas requires slightly more regularity.

Theorem A.2.11 (Nicolas 2007, [90])**.** *The map* (A.2.4) *is a well-defined isomorphism as long as* $\tilde{g} \in L^\infty_{\mathrm{loc}}(\mathbb{R}; \mathcal{C}^0(X)) \cap W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}; \mathcal{C}^0(X))$*, the coefficients of the first-order terms of* L_1 *are in* $L^\infty_{\mathrm{loc}}(\mathbb{R}; \mathcal{C}^0(X))$*, and the coefficients of the zeroth-order terms in* L_1 *are* $L^\infty_{\mathrm{loc}}(\tilde{X})$*.*

A.2.3 A Nonlinear Grönwall Inequality

Lemma A.2.12. *Let $\tau \in [0,1]$ and $f : [0,1] \to \mathbb{R}$ be a continuous non-negative function. Suppose f satisfies the inequality*

$$f(\tau) \leqslant f(0) + \int_0^\tau f(\sigma)P(f(\sigma)^{1/2})\,\mathrm{d}\sigma$$

for some polynomial P with positive coefficients. Then there exists $\varepsilon > 0$ such that if $f(0) \leqslant \varepsilon$, then

$$f(\tau) \leqslant Cf(0)$$

for some $C > 1$ and all $\tau \in [0,1]$.

Proof. The case when P has order zero is trivial, so assume that $P(x) = \sum_{k=0}^d P_k x^k$ for some $d > 0$ and some non-negative real numbers $\{P_k\}_k$. We have

$$\begin{aligned}
f(\tau) &\leqslant f(0) + \int_0^\tau f(\sigma)P(f(\sigma)^{1/2})\,\mathrm{d}\sigma \\
&\leqslant f(0) + \int_0^\tau \sum_{k=0}^d P_k f(\sigma)^{k/2+1}\,\mathrm{d}\sigma \\
&\leqslant f(0) + \int_{\{0<\sigma<\tau\,:\,f(\sigma)<1\}} \sum_{k=0}^d P_k f(\sigma)^{k/2+1}\,\mathrm{d}\sigma + \int_{\{0<\sigma<\tau\,:\,f(\sigma)>1\}} \sum_{k=0}^d P_k f(\sigma)^{k/2+1}\,\mathrm{d}\sigma \\
&\leqslant f(0) + \int_0^\tau \sum_{k=0}^d P_k f(\sigma)\,\mathrm{d}\sigma + \int_0^\tau \sum_{k=0}^d P_k f(\sigma)^{d/2+1}\,\mathrm{d}\sigma \\
&\leqslant f(0) + \int_0^\tau \tilde{C} f(\sigma)\,\mathrm{d}\sigma + \int_0^\tau \tilde{C} f(\sigma)^{d/2+1}\,\mathrm{d}\sigma,
\end{aligned}$$

where $\tilde{C} = (d+1)\max_k P_k$. Now set

$$g(\tau) := f(0) + \int_0^\tau \tilde{C} f(\sigma)\,\mathrm{d}\sigma + \int_0^\tau \tilde{C} f(\sigma)^{d/2+1}\,\mathrm{d}\sigma.$$

Then $f(\tau) \leqslant g(\tau)$, $f(0) = g(0)$, and $g'(\tau) \leqslant \tilde{C}f(\tau) + \tilde{C}f(\tau)^{d/2+1} \leqslant \tilde{C}g(\tau)(1 + g(\tau)^{d/2})$. Defining

$$G(\tau) := g(\tau)^{1/\tilde{C}}\tilde{C}^{-2/(\tilde{C}d)}\left(1 + g(\tau)^{d/2}\right)^{-2/(\tilde{C}d)}$$

and differentiating, one obtains

$$\begin{aligned}
G'(\tau) &= g'(\tau)g(\tau)^{1/\tilde{C}-1}\tilde{C}^{-2/(\tilde{C}d)-1}\left(1 + g(\tau)^{d/2}\right)^{-2/(\tilde{C}d)-1} \\
&\leqslant g(\tau)^{1/\tilde{C}}\tilde{C}^{-2(\tilde{C}d)}\left(1 + g(\tau)^{d/2}\right)^{-2/(\tilde{C}d)},
\end{aligned}$$

so that $G'(\tau) \leqslant G(\tau)$. Since τ is contained in a compact interval, this gives $G(\tau) \lesssim G(0)$, or equivalently

$$\begin{aligned}
g(\tau)^{1/\tilde{C}}\left(1 + g(\tau)^{d/2}\right)^{-2/(\tilde{C}d)} &\lesssim g(0)^{1/\tilde{C}}\left(1 + g(0)^{d/2}\right)^{-2/(\tilde{C}d)} \\
&\lesssim g(0)^{1/\tilde{C}}.
\end{aligned}$$

Rearranging gives

$$g(\tau)^{d/2} \lesssim g(0)^{d/2}(1 + g(\tau)^{d/2}),$$

so if $g(0) = f(0)$ is small enough one has $g(\tau)^{d/2} \lesssim g(0)^{d/2}$, and so

$$f(\tau) \leqslant g(\tau) \leqslant Cg(0) \leqslant Cf(0).$$

□

A.3 The Sobolev Embedding Theorem

Various forms of the Sobolev embedding theorem are used throughout this thesis. The following definitions and theorems may be found in standard texts such as [6], [40], or [41].

Definition A.3.1. Let $(\mathscr{M}, g)$ be a smooth Riemannian manifold of dimension n and ∇ the Levi–Civita connection of g. For a real function ϕ belonging to $\mathcal{C}^k(\mathscr{M})$, $k \geqslant 0$ an integer, we define

$$|\nabla^k \phi|^2 := (\nabla^{a_1}\nabla^{a_2} \dots \nabla^{a_k}\phi)\,(\nabla_{a_1}\nabla_{a_2} \dots \nabla_{a_k}\phi)\,,$$

and denote by $\mathfrak{C}^{k,p}$ the vector space of $\mathcal{C}^\infty$ functions ϕ such that $|\nabla^l \phi| \in L^p(\mathscr{M})$ for all $0 \leqslant l \leqslant k$ and $p \geqslant 1$ a real number.

Definition A.3.2. The Sobolev space $W^{k,p}(\mathscr{M})$ is the completion of $\mathfrak{C}^{k,p}$ with respect to the norm

$$\|\phi\|_{W^{k,p}} := \sum_{l=0}^{k} \|\nabla^l \phi\|_{L^p}.$$

It turns out that the space $W^{k,p}(\mathscr{M})$ does not depend on the Riemannian metric g (theorem 2.20, [6]).

Theorem A.3.3. *Let $\mathscr{M}$ be a smooth compact Riemannian manifold of dimension n, let k, l be integers with $k > l \geqslant 0$, and let p, q be real numbers with $1 \leqslant q < p$ satisfying*

$$\frac{1}{p} = \frac{1}{q} - \frac{(k-l)}{n}.$$

Then

$$W^{k,q}(\mathscr{M}) \subset W^{l,p}(\mathscr{M}),$$

and the identity operator is continuous (the embedding is compact).

Moreover, if

$$\frac{(k-r-\alpha)}{n} \geqslant \frac{1}{q},$$

then

$$W^{k,q}(\mathscr{M}) \subset \mathcal{C}^{r,\alpha}(\mathscr{M}),$$

and the identity operator is continuous (the embedding is compact). Here $r \geqslant 0$ is an integer, α is a real number satisfying $0 < \alpha \leqslant 1$, $\mathcal{C}^{r,\alpha}$ is the space of $\mathcal{C}^r$ functions the r-th derivatives of which belong to $\mathcal{C}^\alpha$, $\mathcal{C}^r$ is the space of functions ϕ of finite $\|\phi\|_{\mathcal{C}^r} := \max_{0\leqslant l\leqslant r} \sup |\nabla^l u|$ norm, and $\mathcal{C}^\alpha$ is the space of functions of finite $\|\phi\|_{\mathcal{C}^\alpha} := \sup|\phi| + \sup_{P\neq Q}\{|\phi(P) - \phi(Q)|d(P,Q)^{-\alpha}\}$ norm.

Remark A.3.4. Theorem A.3.3 holds in the same form with $\mathscr{M}$ an open subset of $\mathbb{R}^n$ with a Lipschitz boundary $\partial\mathscr{M}$. Most generally, $\mathscr{M}$ can be a complete Riemannian manifold with bounded sectional curvature and a positive injectivity radius $\delta > 0$.

The following theorem provides a sharp form of the embedding $W^{1,q} \subset L^p$, which is occasionally useful when working with precisely known constants.

Theorem A.3.5. *Let $\mathscr{M}$ be a smooth compact Riemannian manifold of dimension n and let the real numbers p, q satisfy*

$$\frac{1}{p} = \frac{1}{q} - \frac{1}{n} > 0.$$

Then for every $\varepsilon > 0$ there exists a constant $A_q(\varepsilon)$ such that every $\phi \in W^{1,q}(\mathscr{M})$ satisfies

$$\|\phi\|_{L^p} \leqslant (\mathsf{K}(n,q) + \varepsilon)\,\|\nabla\phi\|_{L^q} + A_q(\varepsilon)\|\phi\|_{L^q},$$

where $\mathsf{K}(n,q)$ is the smallest constant having this property and is given by

$$\mathsf{K}(n,q) = \left(\frac{q-1}{n-q}\right)\left(\frac{n-q}{n(q-1)}\right)^{\frac{1}{q}}\left(\frac{\Gamma(n+1)}{\Gamma(n/q)\Gamma(n+1-n/q)\omega_{n-1}}\right)^{\frac{1}{n}}$$

for $1 < q < n$ and

$$\mathsf{K}(n,1) = \frac{1}{n}\left(\frac{n}{\omega_{n-1}}\right)^{\frac{1}{n}}.$$

B
Conformal Transformations

This appendix collects some results concerning the transformation of various quantities under the conformal rescaling of the metric $\hat{g}_{ab} = \Omega^2 g_{ab}$. Should inverse identities be needed, the roles of $\hat{g}_{ab}$ and g_{ab} in the below identities may be interchanged by employing the changes $g_{ab} \leftrightarrow \hat{g}_{ab}$, $\Omega \leftrightarrow \Omega^{-1}$, and consequently $\Upsilon_a \leftrightarrow -\Upsilon_a$. We recall that we use the Penrose convention for curvature tensors, so that the Riemann tensor R_{abcd} is defined by

$$[\nabla_a, \nabla_b]X^c = -\mathrm{R}^c{}_{dab}X^d.$$

The Ricci tensor and the scalar curvature are then defined by

$$\mathrm{R}_{ab} := \mathrm{R}^c{}_{acb}, \qquad \mathrm{R} := \mathrm{R}^a{}_a,$$

the trace-free part of the Ricci tensor is defined by

$$\Phi_{ab} := -\frac{1}{2}\left(\mathrm{R}_{ab} - \frac{1}{4}\mathrm{R}g_{ab}\right),$$

and the Einstein tensor G_{ab} is defined by

$$\mathrm{G}_{ab} := \mathrm{R}_{ab} - \frac{1}{2}g_{ab}\mathrm{R}.$$

The Weyl tensor is given by

$$\mathrm{C}_{abcd} = \mathrm{R}_{abcd} + \left(\mathrm{R}_{a[d}g_{c]b} + \mathrm{R}_{b[c}g_{d]a}\right) + \frac{1}{3}\mathrm{R}g_{a[c}g_{d]b},$$

which also provides a decomposition of the Riemann tensor into its irreducible parts. In the following α_a is a generic 1-form.

B.1 Curvature and Other

$$\hat{\mathrm{C}}^a{}_{bcd} = \mathrm{C}^a{}_{bcd} \tag{B.1.1}$$

$$\hat{\mathrm{R}}_{ab} = \mathrm{R}_{ab} + 2\left(\nabla_a \Upsilon_b + \frac{1}{2} g_{ab} \nabla^c \Upsilon_c\right) - 2(\Upsilon_a \Upsilon_b - g_{ab} g^{cd} \Upsilon_c \Upsilon_d) \tag{B.1.2}$$

$$\hat{\mathrm{R}} = \Omega^{-2}(\mathrm{R} + 6\nabla^a \Upsilon_a + 6g^{ab}\Upsilon_a\Upsilon_b) \tag{B.1.3}$$

$$\hat{\Phi}_{ab} = \Phi_{ab} - \nabla_a \Upsilon_b + \frac{1}{4} g_{ab} \nabla^c \Upsilon_c + \Upsilon_a \Upsilon_b - \frac{1}{4} g_{ab} g^{cd} \Upsilon_c \Upsilon_d \tag{B.1.4}$$

$$\hat{\mathrm{G}}_{ab} = \mathrm{G}_{ab} + 2(\nabla_a \Upsilon_b - g_{ab} \nabla^c \Upsilon_c) - 2\left(\Upsilon_a \Upsilon_b + \frac{1}{2} g_{ab} g^{cd} \Upsilon_c \Upsilon_d\right) \tag{B.1.5}$$

$$\hat{\Gamma}^a_{bc} = \Gamma^a_{bc} + \Upsilon_c \delta^a_b + \Upsilon_b \delta^a_c - \Upsilon_d g^{ad} g_{bc} \tag{B.1.6}$$

$$\hat{\nabla}_a \alpha_b = \nabla_a \alpha_b - \Upsilon_a \alpha_b - \Upsilon_b \alpha_a + g_{ab} g^{cd} \Upsilon_c \alpha_d \tag{B.1.7}$$

C
The Newman–Penrose Formalism

C.1 Spin Coefficient Formalism

In this thesis we employ the *Newman–Penrose formalism*, which is described in detail in [95] and [96]. For a metric of signature $(+,-,-,-)$ we pick a complex null tetrad $(l, n, m, \bar{m})$ of vectors such that l, n are real and m, $\bar{m}$ are complex null vectors,

$$l_a l^a = n_a n^a = m_a m^a = \bar{m}_a \bar{m}^a = 0,$$

with the orthogonality properties

$$l_a m^a = l_a \bar{m}^a = n_a m^a = n_a \bar{m}^a = 0$$

and

$$l_a n^a = 1, \qquad m_a \bar{m}^a = -1,$$

and such that the metric takes the form

$$g_{ab} = l_a n_b + n_a l_b - m_a \bar{m}_b - \bar{m}_a m_b.$$

The volume form is then given by

$$\mathrm{dv} = l^\flat \wedge n^\flat \wedge (i m^\flat \wedge \bar{m}^\flat),$$

where $l^\flat$, $n^\flat$, $m^\flat$ and $\bar{m}^\flat$ denote the 1-forms corresponding to the vector fields l, n, m and $\bar{m}$. We define the four directional derivatives

$$D := l^a \nabla_a, \qquad \Delta := n^a \nabla_a, \qquad \delta := m^a \nabla_a, \qquad \bar{\delta} := \bar{m}^a \nabla_a,$$

and the *spin coefficients*

$$
\begin{aligned}
\kappa &:= m^a D l_a, \\
\varepsilon &:= \frac{1}{2}\left(n^a D l_a + m^a D \bar{m}_a\right), \\
\pi &:= -\bar{m}^a D n_a, \\
\rho &:= m^a \bar{\delta} l_a, \\
\alpha &:= \frac{1}{2}\left(n^a \bar{\delta} l_a + m^a \bar{\delta} \bar{m}_a\right), \\
\lambda &:= -\bar{m}^a \bar{\delta} n_a, \\
\sigma &:= m^a \delta l_a, \\
\beta &:= \frac{1}{2}\left(n^a \delta l_a + m^a \delta \bar{m}_a\right), \\
\mu &:= -\bar{m}^a \delta n_a, \\
\tau &:= m^a \Delta l_a, \\
\gamma &:= \frac{1}{2}\left(n^a \Delta l_a + m^a \Delta \bar{m}_a\right), \\
\nu &:= -\bar{m}^a \Delta n_a.
\end{aligned}
$$

With these definitions the vectors $(l, n, m, \bar{m})$ satisfy the transport equations

$$
\begin{aligned}
Dl^a &= (\varepsilon + \bar{\varepsilon}) l^a - \bar{\kappa} m^a - \kappa \bar{m}^a, & Dn^a &= -(\varepsilon + \bar{\varepsilon}) n^a + \pi m^a + \bar{\pi} \bar{m}^a, \\
\delta l^a &= (\beta + \bar{\alpha}) l^a - \bar{\rho} m^a - \sigma \bar{m}^a, & \delta n^a &= -(\beta + \bar{\alpha}) n^a + \mu m^a + \bar{\lambda} \bar{m}^a, \\
\bar{\delta} l^a &= (\alpha + \bar{\beta}) l^a - \bar{\sigma} m^a - \rho \bar{m}^a, & \bar{\delta} n^a &= -(\alpha + \bar{\beta}) n^a + \lambda m^a + \bar{\mu} \bar{m}^a, \\
\Delta l^a &= (\gamma + \bar{\gamma}) l^a - \bar{\tau} m^a - \tau \bar{m}^a, & \Delta n^a &= -(\gamma + \bar{\gamma}) n^a + \nu m^a + \bar{\nu} \bar{m}^a,
\end{aligned}
$$

and

$$
\begin{aligned}
Dm^a &= (\varepsilon - \bar{\varepsilon}) m^a + \bar{\pi} l^a - \kappa n^a, & D\bar{m}^a &= (\bar{\varepsilon} - \varepsilon) \bar{m}^a + \pi l^a - \bar{\kappa} n^a, \\
\delta m^a &= (\beta - \bar{\alpha}) m^a + \bar{\lambda} l^a - \sigma n^a, & \delta \bar{m}^a &= (\bar{\alpha} - \beta) \bar{m}^a + \mu l^a - \bar{\rho} n^a, \\
\bar{\delta} m^a &= (\alpha - \bar{\beta}) m^a + \bar{\mu} l^a - \rho n^a, & \bar{\delta} \bar{m}^a &= (\bar{\beta} - \alpha) \bar{m}^a + \lambda l^a - \bar{\sigma} n^a, \\
\Delta m^a &= (\gamma - \bar{\gamma}) m^a + \bar{\nu} l^a - \tau n^a, & \Delta \bar{m}^a &= (\bar{\gamma} - \gamma) \bar{m}^a + \nu l^a - \bar{\tau} n^a.
\end{aligned}
$$

The operators D, Δ, δ and $\bar{\delta}$ satisfy the commutation relations

$$
\begin{aligned}
[\Delta, D] &= (\gamma + \bar{\gamma}) D + (\varepsilon + \bar{\varepsilon}) \Delta - (\tau + \bar{\pi}) \bar{\delta} - (\pi + \bar{\tau}) \delta, \\
[\delta, D] &= (\beta + \bar{\alpha} - \bar{\pi}) D + \kappa \Delta - \sigma \bar{\delta} - (\varepsilon - \bar{\varepsilon} + \bar{\rho}) \delta, \\
[\delta, \Delta] &= -\bar{\nu} D + (\tau - \bar{\alpha} - \beta) \Delta + \bar{\lambda} \bar{\delta} + (\bar{\gamma} - \gamma + \mu) \delta, \\
[\bar{\delta}, \delta] &= (\bar{\mu} - \mu) D + (\bar{\rho} - \rho) \Delta + (\beta - \bar{\alpha}) \bar{\delta} + (\alpha - \bar{\beta}) \delta.
\end{aligned}
$$

Finally, defining the curvature components

$$
\Lambda := \frac{1}{24} \mathrm{R},
$$

$$\Psi_0 := \mathrm{C}_{abcd} l^a m^b l^c m^d, \qquad \Psi_1 := \mathrm{C}_{abcd} l^a m^b l^c n^d,$$
$$\Psi_2 := \mathrm{C}_{abcd} l^a m^b \bar{m}^c n^d, \qquad \Psi_3 := \mathrm{C}_{abcd} l^a n^b \bar{m}^c n^d,$$
$$\Psi_4 := \mathrm{C}_{abcd} \bar{m}^a n^b \bar{m}^c n^d,$$

and

$$\Phi_{00} := -\frac{1}{2}\mathrm{R}_{ab} l^a l^b, \qquad \Phi_{01} := -\frac{1}{2}\mathrm{R}_{ab} l^a m^b,$$
$$\Phi_{02} := -\frac{1}{2}\mathrm{R}_{ab} m^a m^b, \qquad \Phi_{10} := -\frac{1}{2}\mathrm{R}_{ab} l^a \bar{m}^b,$$
$$\Phi_{11} := -\frac{1}{2}\mathrm{R}_{ab} l^a n^b + 3\Lambda, \qquad \Phi_{12} := -\frac{1}{2}\mathrm{R}_{ab} m^a n^b,$$
$$\Phi_{20} := -\frac{1}{2}\mathrm{R}_{ab} \bar{m}^a \bar{m}^b, \qquad \Phi_{21} := -\frac{1}{2}\mathrm{R}_{ab} \bar{m}^a n^b,$$
$$\Phi_{22} := -\frac{1}{2}\mathrm{R}_{ab} n^a n^b,$$

the spin coefficients satisfy the curvature equations

$$\begin{aligned}
&D\rho - \bar{\delta}\kappa = \rho^2 + |\sigma|^2 - \bar{\kappa}\tau - \kappa(-\pi + 3\alpha + \bar{\beta}) + \rho(\varepsilon + \bar{\varepsilon}) + \Phi_{00},\\
&\Delta\mu - \delta\nu = -\mu^2 - |\lambda|^2 + \bar{\nu}\pi - \nu(\tau - 3\beta - \bar{\alpha}) - \mu(\gamma + \bar{\gamma}) - \Phi_{22},\\
&D\sigma - \delta\kappa = \sigma(\rho + \bar{\rho} - \bar{\varepsilon} + 3\varepsilon) - \kappa(\tau - \bar{\pi} + \bar{\alpha} + 3\beta) + \Psi_0,\\
&\Delta\lambda - \bar{\delta}\nu = \lambda(-\mu - \bar{\mu} + \bar{\gamma} - 3\gamma) - \nu(-\pi + \bar{\tau} - \bar{\beta} - 3\alpha) - \Psi_4,\\
&D\tau - \Delta\kappa = \rho(\tau + \bar{\pi}) + \sigma(\bar{\tau} + \pi) + \tau(-\bar{\varepsilon} + \varepsilon) - \kappa(\bar{\gamma} + 3\gamma) + \Psi_1 + \Phi_{01},\\
&\Delta\pi - D\nu = -\mu(\pi + \bar{\tau}) - \lambda(\bar{\pi} + \tau) + \pi(\bar{\gamma} - \gamma) + \nu(\bar{\varepsilon} + 3\varepsilon) - \Psi_3 - \Phi_{21},\\
&\delta\rho - \bar{\delta}\sigma = \tau(\rho - \bar{\rho}) + \kappa(-\bar{\mu} + \mu) + \rho(\bar{\alpha} + \beta) - \sigma(-\bar{\beta} + 3\alpha) - \Psi_1 + \Phi_{01},\\
&\bar{\delta}\mu - \delta\lambda = \pi(-\mu + \bar{\mu}) + \nu(\bar{\rho} - \rho) - \mu(\bar{\beta} + \alpha) - \lambda(\bar{\alpha} - 3\beta) + \Psi_3 - \Phi_{21},\\
&\delta\tau - \Delta\sigma = \mu\sigma + \bar{\lambda}\rho + \tau^2 - \kappa\bar{\nu} + \tau(\beta - \bar{\alpha}) - \sigma(3\gamma - \bar{\gamma}) + \Phi_{02},\\
&\bar{\delta}\pi - D\lambda = -\rho\lambda - \bar{\sigma}\mu - \pi^2 + \nu\bar{\kappa} + \pi(-\alpha + \bar{\beta}) - \lambda(-3\varepsilon + \bar{\varepsilon}) - \Phi_{20},\\
&\Delta\rho - \bar{\delta}\tau = -\rho\bar{\mu} - \sigma\lambda - |\tau|^2 + \kappa\nu + \rho(\gamma + \bar{\gamma}) - \tau(\alpha - \bar{\beta}) - \Psi_2 - 2\Lambda,\\
&D\mu - \delta\pi = \mu\bar{\rho} + \sigma\lambda + |\pi|^2 - \kappa\nu - \mu(\varepsilon + \bar{\varepsilon}) - \pi(-\beta + \bar{\alpha}) + \Psi_2 + 2\Lambda,\\
&\Delta\beta - \delta\gamma = -\tau\mu + \nu\sigma + \bar{\nu}\varepsilon - \alpha\bar{\lambda} + \beta(-\mu - \bar{\gamma} + \gamma) - \gamma(-\bar{\alpha} - \beta + \tau) - \Phi_{12},\\
&D\alpha - \bar{\delta}\varepsilon = \pi\rho - \kappa\lambda - \bar{\kappa}\gamma + \beta\bar{\sigma} + \alpha(\rho + \bar{\varepsilon} - \varepsilon) - \varepsilon(\bar{\beta} + \alpha - \pi) + \Phi_{10},\\
&D\beta - \delta\varepsilon = -\kappa(\mu + \gamma) + \sigma(\pi + \alpha) + \beta(\bar{\rho} - \bar{\varepsilon}) - \varepsilon(-\bar{\pi} + \bar{\alpha}) + \Psi_1,\\
&\Delta\alpha - \bar{\delta}\gamma = \nu(\rho + \varepsilon) - \lambda(\tau + \beta) + \alpha(-\bar{\mu} + \bar{\gamma}) - \gamma(\bar{\tau} - \bar{\beta}) - \Psi_3,\\
&D\gamma - \Delta\varepsilon = -\kappa\nu + \tau\pi + \beta(\pi + \bar{\tau}) + \alpha(\bar{\pi} + \tau) - \varepsilon(\gamma + \bar{\gamma}) - \gamma(\varepsilon + \bar{\varepsilon}) + \Psi_2 + \Phi_{11} - \Lambda,\\
&\bar{\delta}\beta - \delta\alpha = -\rho\mu + \sigma\lambda - |\alpha|^2 - |\beta|^2 + 2\alpha\beta + \gamma(\bar{\rho} - \rho) + \varepsilon(-\mu + \bar{\mu}) + \Psi_2 - \Phi_{11} - \Lambda.
\end{aligned}$$

It will be useful to know how the spin coefficients $\{\kappa, \varepsilon, \pi, \rho, \alpha, \lambda, \sigma, \beta, \mu, \tau, \gamma, \nu\}$ behave under the conformal change

$$\hat{l}^a = l^a, \qquad \hat{n}^a = \Omega^{-2} n^a, \qquad \hat{m}^a = \Omega^{-1} m^a, \qquad \hat{\bar{m}}^a = \Omega^{-1} \bar{m}^a$$

of the associated tetrad. The transformation rules are

$$\begin{bmatrix} \hat{\varepsilon} & \hat{\kappa} & \hat{\pi} \\ \hat{\alpha} & \hat{\rho} & \hat{\lambda} \\ \hat{\beta} & \hat{\sigma} & \hat{\mu} \\ \hat{\gamma} & \hat{\tau} & \hat{\nu} \end{bmatrix} = \begin{bmatrix} \varepsilon + \Upsilon_a l^a & \Omega\kappa & \Omega^{-1}(\pi + \bar{m}^a \Upsilon_a) \\ \Omega^{-1}\alpha & \rho - \Upsilon_a l^a & \Omega^{-2}\lambda \\ \Omega^{-1}\beta & \sigma & \Omega^{-2}(\mu + \Upsilon_a n^a) \\ \Omega^{-2}\gamma & \Omega^{-1}(\tau - \Upsilon_a m^a) & \Omega^{-3}\nu \end{bmatrix}. \tag{C.1.1}$$

C.2 Compacted Spin Coefficient Formalism

The spin coefficient formalism described in appendix C.1 has its advantages in the fact that one operates entirely with scalar quantities, and that the spin coefficients are in general complex quantities and thus encode twice the amount of information compared to real quantities. The formalism is most useful when the physical or geometrical problem in question possesses symmetries, for then many of the spin coefficients vanish or coincide and the equations in appendix C.1 simplify considerably.

However, there may in general be significant freedom in choosing the frame $(l, m, \bar{m}, n)$. In calculations this may introduce quantities that do not in themselves have direct geometrical meaning—'gauge quantities'—and detract from the intrinsic geometrical or physical content of the problem. These issues are of course avoided entirely when using a fully covariant approach, but in explicit problems this may not always be convenient. The compacted spin coefficient formalism provides a sort of partially covariant approach, which turns out to be ideally suited to studying problems in which two null directions are singled out by the nature of the problem. The problem of studying field asymptotics on $\mathscr{I}^+$ and $\mathscr{I}^-$ is of course one such example.

Suppose we have chosen a tetrad $(l, m, \bar{m}, n)$ satisfying the properties of appendix C.1. Any transformation of the form

$$l^a \mapsto \lambda\bar{\lambda} l^a, \qquad m^a \mapsto \lambda\bar{\lambda}^{-1} m^a, \qquad \bar{m}^a \mapsto \lambda^{-1}\bar{\lambda}\bar{m}^a, \qquad n^a \mapsto \lambda^{-1}\bar{\lambda}^{-1} n^a \tag{C.2.1}$$

for a nowhere vanishing complex scalar field λ leaves these properties unchanged; there is a 'gauge group', which at each point is seen to be just the multiplicative group of complex numbers λ. The compacted spin coefficient formalism is concerned with dealing with scalars (or tensors) η that under (C.2.1) transform according to

$$\eta \mapsto \lambda^p \bar{\lambda}^q \eta.$$

We say such a scalar (tensor) has type (or weight) $\{p, q\}$, and call η a $\{p, q\}$-scalar (tensor). The vectors l, m, $\bar{m}$, n are of course themselves vectors of weight $\{1, 1\}$, $\{1, -1\}$, $\{-1, 1\}$, and $\{-1, -1\}$ respectively.

Spin coefficients may be divided into two classes according to whether or not they are weighted quantities. It turns out that the coefficients κ, ρ, σ, τ, ν, μ, λ, π are such quantities, whereas ε, α, β, γ are not. Indeed, one can check that, for example,

$$\sigma \mapsto \lambda^3\bar{\lambda}^{-1}\sigma,$$

but

$$\beta \mapsto \lambda\bar{\lambda}^{-1}\beta + \bar{\lambda}^{-1}\delta\lambda.$$

The types of the weighted spin coefficients are summarised below.

κ	$\{3, 1\}$	ν	$\{-3, -1\}$
σ	$\{3, -1\}$	λ	$\{-3, 1\}$
ρ	$\{1, 1\}$	μ	$\{-1, -1\}$
τ	$\{1, -1\}$	π	$\{-1, 1\}$.

For any tensor field there exists a set of weighted scalars defining the tensor field obtained by contracting the tensor field with various combinations of the vector fields l^a, m^a, $\bar{m}^a$, n^a. Clearly the product of a $\{p, q\}$-scalar and a $\{p', q'\}$-scalar is a $\{p + p', q + q'\}$-scalar. We next wish to introduce derivative operators in such a way that acting on weighted scalars they produce (differently) weighted scalars. We define for a $\{p, q\}$-scalar (or tensor) η

$$
\begin{aligned}
\text{þ}\eta &:= (D - p\varepsilon - q\bar{\varepsilon})\eta, \\
\text{þ}'\eta &:= (\Delta - p\gamma - q\bar{\gamma})\eta, \\
\eth\eta &:= (\delta - p\beta - q\bar{\alpha})\eta, \\
\eth'\eta &:= (\bar{\delta} - p\alpha - q\bar{\beta})\eta.
\end{aligned}
$$

One can check that these operators have the following types:

$$
\begin{array}{ll}
\text{þ} & \{1, 1\} \\
\eth & \{1, -1\} \\
\eth' & \{-1, 1\} \\
\text{þ}' & \{-1, -1\},
\end{array}
$$

in the sense that, for example, acting on a $\{p, q\}$-scalar, þ produces a $\{p+1, q+1\}$-scalar. Notice that it is precisely the *unweighted* spin coefficients ε, α, β, γ that get incorporated in the definitions of þ, $\eth$, $\eth'$, þ$'$, and so they get withdrawn from our formalism, leaving only weighted objects. It can be easily verified that the operators þ, $\eth$, $\eth'$, þ$'$ are additive and satisfy the Leibniz rule. Since complex conjugation changes the type of a $\{p, q\}$ scalar to $\{q, p\}$, in order to have the relations

$$
\overline{\text{þ}\eta} = \bar{\text{þ}}\bar{\eta}, \qquad \overline{\eth\eta} = \bar{\eth}\bar{\eta},
$$

we define

$$
\bar{\text{þ}} := \text{þ}, \qquad \bar{\text{þ}}' := \text{þ}', \qquad \bar{\eth} := \eth', \qquad \bar{\eth}' := \eth.
$$

The transport equations of appendix C.1 now become

$$
\begin{aligned}
\text{þ}l^a &= -\bar{\kappa}m^a - \kappa\bar{m}^a, & \text{þ}n^a &= \pi m^a + \bar{\pi}\bar{m}^a, \\
\eth l^a &= -\bar{\rho}m^a - \sigma\bar{m}^a, & \eth n^a &= \mu m^a + \bar{\lambda}\bar{m}^a, \\
\eth' l^a &= -\bar{\sigma}m^a - \rho\bar{m}^a, & \eth' n^a &= \lambda m^a + \bar{\mu}\bar{m}^a, \\
\text{þ}' l^a &= -\bar{\tau}m^a - \tau\bar{m}^a, & \text{þ}' n^a &= \nu m^a + \bar{\nu}\bar{m}^a,
\end{aligned}
$$

and

$$
\begin{aligned}
\text{þ}m^a &= \bar{\pi}l^a - \kappa n^a, & \text{þ}\bar{m}^a &= \pi l^a - \bar{\kappa}n^a, \\
\eth m^a &= \bar{\lambda}l^a - \sigma n^a, & \eth\bar{m}^a &= \mu l^a - \bar{\rho}n^a, \\
\eth' m^a &= \bar{\mu}l^a - \rho n^a, & \eth'\bar{m}^a &= \lambda l^a - \bar{\sigma}n^a \\
\text{þ}' m^a &= \bar{\nu}l^a - \tau n^a, & \text{þ}'\bar{m}^a &= \nu l^a - \bar{\tau}n^a.
\end{aligned}
$$

The operators þ, ð, ð′, þ′ satisfy the commutation relations

$$
\begin{aligned}
[þ, þ'] &= (\bar{\tau} + \pi)ð + (\tau + \bar{\pi})ð' - p(-\kappa\nu + \tau\pi + \Psi_2 + \Phi_{11} - \Lambda) - q(-\bar{\kappa}\bar{\nu} + \bar{\tau}\bar{\pi} + \bar{\Psi}_2 + \Phi_{11} - \Lambda),\\
[þ, ð] &= \bar{\rho}ð + \sigma ð' + \bar{\pi}þ - \kappa þ' - p(-\mu\kappa + \pi\sigma + \Psi_1) - q(-\bar{\lambda}\bar{\kappa} + \bar{\rho}\bar{\pi} + \Phi_{01}),\\
[ð, ð'] &= (\mu - \bar{\mu})þ + (\rho - \bar{\rho})þ' + p(-\mu\rho + \sigma\lambda + \Psi_2 - \Phi_{11} - \Lambda) - q(-\bar{\mu}\bar{\rho} + \bar{\sigma}\bar{\lambda} + \bar{\Psi}_2 - \Phi_{11} - \Lambda),\\
[þ, ð'] &= \rho ð' + \bar{\sigma}ð + \pi þ - \bar{\kappa}þ' - p(-\bar{\mu}\bar{\kappa} + \bar{\pi}\bar{\sigma} + \bar{\Psi}_1) - q(-\lambda\kappa + \pi\rho + \bar{\Phi}_{01}),\\
[þ', ð] &= -\mu ð - \bar{\lambda}ð' - \tau þ' + \bar{\nu}þ - p(-\bar{\rho}\bar{\nu} + \bar{\tau}\bar{\lambda} + \bar{\Psi}_3) - q(-\sigma\nu + \mu\tau + \bar{\Phi}_{21}),\\
[þ', ð'] &= -\bar{\mu}ð' - \lambda ð - \bar{\tau}þ' + \nu þ - p(-\rho\nu + \tau\lambda + \Psi_3) - q(-\bar{\sigma}\bar{\nu} + \bar{\mu}\bar{\tau} + \Phi_{21}).
\end{aligned}
$$

Finally, the curvature equations in terms of the compacted spin coefficient derivative operators are

$$
\begin{aligned}
þ\rho - \bar{ð}\kappa &= \rho^2 + |\sigma|^2 - \bar{\kappa}\tau + \pi\kappa + \Phi_{00},\\
þ\sigma - ð\kappa &= (\rho + \bar{\rho})\sigma - (\tau - \bar{\pi})\kappa + \Psi_0,\\
þ\tau - þ'\kappa &= (\tau + \bar{\pi})\rho + (\bar{\tau} + \pi)\sigma + \Psi_1 + \Phi_{01},\\
ð\rho - ð'\sigma &= (\rho - \bar{\rho})\tau + (-\bar{\mu} + \mu)\kappa - \Psi_1 + \Phi_{01},\\
ð\tau - þ'\sigma &= \mu\sigma + \bar{\lambda}\rho + \tau^2 - \kappa\bar{\nu} + \Phi_{02},\\
þ'\rho - ð'\tau &= -\rho\bar{\mu} - \sigma\lambda - |\tau|^2 + \kappa\nu - \Psi_2 - 2\Lambda.
\end{aligned}
$$

C.3 Specific Christoffel Symbols, Spin Coefficients, and Curvature Quantities

C.3.1 The Physical Schwarzschild Metric

The non-zero Christoffel symbols associated to the physical Schwarzschild metric

$$g_{ab} = F(r)\,\mathrm{d}t^2 - F(r)^{-1}\mathrm{d}r^2 - r^2\mathfrak{s}_2,$$

where $F(r) = 1 - 2mr^{-1}$, are

$$
\begin{aligned}
&\Gamma^t_{tr} = \Gamma^t_{rt} = \frac{m}{r^2}F(r)^{-1}, && \Gamma^r_{tt} = \frac{m}{r^2}F(r), && \Gamma^r_{rr} = -\frac{m}{r^2}F(r)^{-1},\\
&\Gamma^r_{\theta\theta} = -rF(r), && \Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r}, && \Gamma^\phi_{r\phi} = \Gamma^\phi_{\phi r} = \frac{1}{r},\\
&\Gamma^r_{\phi\phi} = -rF(r)\sin^2\theta, && \Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta, && \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta.
\end{aligned}
$$

In the coordinates (u, r, θ, ϕ) the Schwarzschild metric takes the form

$$g_{ab} = F(r)\,\mathrm{d}u^2 + 2\,\mathrm{d}u\,\mathrm{d}r - r^2\mathfrak{s}_2,$$

and has the associated Christoffel symbols

$$
\begin{aligned}
&\Gamma^u_{uu} = -\frac{m}{r^2}, && \Gamma^u_{\theta\theta} = r, && \Gamma^u_{\phi\phi} = r\sin^2\theta, && \\
&\Gamma^r_{uu} = \frac{m}{r^2}F(r), && \Gamma^r_{ur} = \Gamma^r_{ru} = \frac{m}{r^2}, && \Gamma^r_{\theta\theta} = -rF(r), && \Gamma^r_{\phi\phi} = -r\sin^2\theta F(r),\\
&\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r}, && \Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta, && \Gamma^\phi_{r\phi} = \Gamma^\phi_{\phi r} = \frac{1}{r}, && \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta.
\end{aligned}
$$

The NP spin coefficients associated to the null tetrad

$$(l^a, n^a, m^a, \bar{m}^a) = \left(\partial_u - \frac{1}{2}F(r)\partial_r,\ \partial_r,\ \frac{1}{\sqrt{2}r}\left(\partial_\theta + \frac{i}{\sin\theta}\partial_\phi\right),\ \frac{1}{\sqrt{2}r}\left(\partial_\theta - \frac{i}{\sin\theta}\partial_\phi\right)\right) \tag{C.3.1}$$

for the metric $g_{ab} = F(r)\,\mathrm{d}u^2 + 2\,\mathrm{d}u\,\mathrm{d}r - r^2\mathfrak{s}_2$ are

$$\begin{bmatrix} \varepsilon & \kappa & \pi \\ \alpha & \rho & \lambda \\ \beta & \sigma & \mu \\ \gamma & \tau & \nu \end{bmatrix} = \begin{bmatrix} -\dfrac{m}{2r^2} & 0 & 0 \\ -\dfrac{\sqrt{2}}{4r}\cot\theta & \dfrac{1}{2r}F(r) & 0 \\ \dfrac{\sqrt{2}}{4r}\cot\theta & 0 & \dfrac{1}{r} \\ 0 & 0 & 0 \end{bmatrix}. \tag{C.3.2}$$

With respect to this tetrad the NP curvature scalars for g_{ab} are given by

$$\Phi_{ab} = 0, \qquad \Lambda = 0, \qquad \Psi_2 = -\frac{m}{r^3}, \qquad \Psi_{0,1,3,4} = 0.$$

C.3.2 The Rescaled Schwarzschild Metric

The non-zero Christoffel symbols associated to the rescaled Schwarzschild metric

$$\hat{g}_{ab} = R^2F(R)\,\mathrm{d}u^2 - 2\,\mathrm{d}u\,\mathrm{d}R - \mathfrak{s}_2,$$

where $F(R) = 1 - 2mR$, are

$$\hat{\Gamma}^u_{uu} = R(1-3mR), \quad \hat{\Gamma}^R_{uu} = R^3F(R)(1-3mR), \quad \hat{\Gamma}^R_{uR} = \hat{\Gamma}^R_{Ru} = -R(1-3mR),$$
$$\hat{\Gamma}^\theta_{\phi\phi} = -\sin\theta\,\cos\theta, \quad \hat{\Gamma}^\phi_{\theta\phi} = \hat{\Gamma}^\phi_{\phi\theta} = \cot\theta.$$

The NP spin coefficients associated to the null tetrad

$$(\hat{l}^a, \hat{n}^a, \hat{m}^a, \bar{\hat{m}}^a) = \left(\partial_u + \frac{1}{2}R^2F(R)\partial_R,\ -\partial_R,\ \frac{1}{\sqrt{2}}\left(\partial_\theta + \frac{i}{\sin\theta}\partial_\phi\right),\ \frac{1}{\sqrt{2}}\left(\partial_\theta - \frac{i}{\sin\theta}\partial_\phi\right)\right) \tag{C.3.3}$$

for the metric $\hat{g}_{ab} = R^2F(R)\,\mathrm{d}u^2 - 2\,\mathrm{d}u\,\mathrm{d}R - \mathfrak{s}_2$ are

$$\begin{bmatrix} \hat{\varepsilon} & \hat{\kappa} & \hat{\pi} \\ \hat{\alpha} & \hat{\rho} & \hat{\lambda} \\ \hat{\beta} & \hat{\sigma} & \hat{\mu} \\ \hat{\gamma} & \hat{\tau} & \hat{\nu} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{2}R(1-3mR) & 0 & 0 \\ -\dfrac{\sqrt{2}}{4}\cot\theta & 0 & 0 \\ \dfrac{\sqrt{2}}{4}\cot\theta & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{C.3.4}$$

With respect to this tetrad the NP curvature scalars for $\hat{g}_{ab}$ are given by

$$\hat{\Phi}_{00,01,02,12,22} = 0, \quad \hat{\Phi}_{11} = \frac{1}{2}(1-3mR), \quad \hat{\Lambda} = \frac{1}{2}mR, \quad \hat{\Psi}_2 = -mR, \quad \hat{\Psi}_{0,1,3,4} = 0.$$

C.3.3 The Einstein Cylinder

The non-zero Christoffel symbols associated to the metric on the Einstein cylinder

$$\mathfrak{e}_{ab} = \mathrm{d}\tau^2 - \mathrm{d}\zeta^2 - \sin^2\zeta(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)$$

are

$$\begin{aligned}
&\Gamma^{\zeta}_{\theta\theta} = -\sin\zeta\,\cos\zeta, && \Gamma^{\zeta}_{\theta\theta} = -\sin^2\theta\,\sin\zeta\,\cos\zeta,\\
&\Gamma^{\theta}_{\zeta\theta} = \Gamma^{\theta}_{\theta\zeta} = \cot\zeta, && \Gamma^{\theta}_{\phi\phi} = -\sin\theta\,\cos\theta,\\
&\Gamma^{\phi}_{\zeta\phi} = \Gamma^{\phi}_{\phi\zeta} = \cot\zeta, && \Gamma^{\phi}_{\theta\phi} = \Gamma^{\phi}_{\phi\theta} = \cot\theta.
\end{aligned}$$

The non-zero components of the Riemann tensors of $\mathfrak{e}_{ab}$ are

$$\begin{aligned}
&\mathrm{R}^{\zeta}{}_{\theta\zeta\theta} = -\mathrm{R}^{\zeta}{}_{\theta\theta\zeta} = -\sin^2\zeta, && \mathrm{R}^{\zeta}{}_{\phi\zeta\phi} = -\mathrm{R}^{\zeta}_{\phi\phi\zeta} = -\sin^2\theta\,\sin^2\zeta,\\
&\mathrm{R}^{\theta}{}_{\zeta\zeta\theta} = -\mathrm{R}^{\theta}{}_{\zeta\theta\zeta} = 1, && \mathrm{R}^{\theta}{}_{\phi\theta\phi} = -\mathrm{R}^{\theta}{}_{\phi\phi\theta} = -\sin^2\theta\,\sin^2\zeta,\\
&\mathrm{R}^{\phi}{}_{\zeta\zeta\phi} = -\mathrm{R}^{\phi}{}_{\zeta\phi\zeta} = 1, && \mathrm{R}^{\phi}{}_{\theta\theta\phi} = -\mathrm{R}^{\phi}{}_{\theta\phi\theta} = \sin^2\zeta.
\end{aligned}$$

The non-zero components of the Ricci tensor of $\mathfrak{e}_{ab}$ are

$$\mathrm{R}_{\zeta\zeta} = -2, \qquad \mathrm{R}_{\theta\theta} = -2\sin^2\zeta, \qquad \mathrm{R}_{\phi\phi} = -2\sin^2\theta\,\sin^2\zeta,$$

and the scalar curvature is

$$\mathrm{R} = 6.$$

A work of art is never finished;
it is simply abandoned.

— Thomas W. Körner

Bibliography

[1]B. P. Abbott and LIGO Scientific Collaboration and Virgo Collaboration, "Observation of gravitational waves from a binary black hole merger", Phys. Rev. Lett. **116** (2016).

[2]B. P. Abbott and LIGO Scientific Collaboration and Virgo Collaboration, "Properties of the binary black hole merger GW150914", Phys. Rev. Lett. **116** (2016).

[3]D. Anninos, G. S. Ng, and A. Strominger, "Asymptotic symmetries and charges in de Sitter space", arXiv:1009.4730 (2010).

[4]A. Ashtekar, B. Bonga, and A. Kesavan, "Asymptotics with a positive cosmological constant: I. basic framework", arXiv:1409.3816 (2014).

[5]A. Ashtekar, "Geometry and physics of null infinity", arXiv:1409.1800 (2014).

[6]T. Aubin, *Some Nonlinear Problems in Riemannian Geometry* (Springer-Verlag, 1998).

[7]A. Bachelot, "Asymptotic completeness for the Klein–Gordon equation on the Schwarzschild metric", Ann. Inst. Henri Poincaré, Physique théorique **61**, 411–441 (1994).

[8]A. Bachelot, "Gravitational scattering of electromagnetic field by a Schwarzschild black hole", Ann. Inst. Henri Poincaré, Physique théorique **54**, 261–320 (1991).

[9]A. Bachelot, "The Hawking effect", Ann. Inst. Henri Poincaré, Physique théorique **70**, 41–99 (1999).

[10]J. Baez, Journal of Functional Analysis **83**, 317–332 (1989).

[11]J. C. Baez, "Conserved quantities for the Yang–Mills equations", Advances in Mathematics **82**, 126–131 (1990).

[12]J. C. Baez, "Scattering for the Yang–Mills equations", Transactions of the American Mathematical Society **315**, 823–832 (1989).

[13]J. C. Baez, I. E. Segal, and Z.-F. Zhou, "The global Goursat problem and scattering for nonlinear wave equations", Journal of Functional Analysis **93**, 239–269 (1990).

[14]J. C. Baez and Z. Zhou, "The global Goursat problem on $\mathbb{R}\times\mathbb{S}^1$", Journal of Functional Analysis **83**, 364–382 (1989).

[15]C. Bär, N. Ginoux, and F. Pfäffle, *Wave equations on Lorentzian manifolds and Quantization* (European Mathematical Society, 2007).

[16]J. D. Barrow and D. J. Shaw, "The value of the cosmological constant", arXiv:1105.3105 (2011).

[17]L. Bieri and D. Garfinkle, "An electromagnetic analogue of gravitational wave memory", Classical and Quantum Gravity **30** (2013).

[18]P. Bizoń and A. Rostworowski, "Weakly turbulent instability of anti–de Sitter spacetime", Phys. Rev. Lett. **107**, 031102 (2011).

[19]A. Cabet, P. T. Chruściel, and R. T. Wafo, "On the characteristic initial value problem for nonlinear symmetric hyperbolic systems, including Einstein equations", arXiv:1406.3009 (2016).

[20]T. Candy, C. Kauffman, and H. Lindblad, "Asymptotic behaviour of the Maxwell–Klein–Gordon system", arXiv:1803.11086 (2018).

[21]Y. Choquet-Bruhat, S. M. Paneitz, and I. E. Segal, "The Yang–Mills equations on the universal cosmos", Journal of Functional Analysis **53**, 112–150 (1983).

[22]Y. Choquet-Bruhat, "Solution globale des equations de Maxwell–Dirac–Klein–Gordon", Rendiconti del Circolo Matematico di Palermo **31**, 267–288 (1982).

[23]Y. Choquet-Bruhat and D. Christodoulou, "Existence of global solutions of the Yang–Mills, Higgs and spinor field equations in $3 + 1$ dimensions", Annales scientifiques de l'École Normale Supérieure **Ser. 4, 14**, 481–506 (1981).

[24]D. Christodoulou and S. Klainerman, *The global nonlinear stability of Minkowski space*, PMS-41 (Princeton University Press, 1994).

[25]P. T. Chruściel and T.-T. Paetz, "The many ways of the characteristic Cauchy problem", Classical and Quantum Gravity **29** (2012).

[26]P. T. Chruściel and E. Delay, "Existence of non-trivial, vacuum, asymptotically simple spacetimes", Classical and Quantum Gravity **19**, 3389 (2002).

[27]P. T. Chruściel and E. Delay, "On mapping properties of the general relativistic constraints operator in weighted function spaces, with applications", Mem. Soc. Math. France **94**, 1–103 (2003).

[28]P. T. Chruściel and J. Shatah, "Global existence of solutions of the Yang–Mills equations on globally hyperbolic four dimensional Lorentzian manifolds", Asian J. Math. **1**, 530–548 (1997).

[29]J. Corvino, "Scalar curvature deformation and a gluing construction for the Einstein constraint equations", Communications in Mathematical Physics **214**, 137–189 (2000).

[30]J. Corvino and R. M. Schoen, "On the asymptotics for the vacuum Einstein constraint equations", J. Differential Geom. **73**, 185–217 (2006).

[31]M. Dafermos and G. Holzegel, "Dynamic instability of solitons in $4 + 1$-dimensional gravity with negative cosmological constant", (2006).

[32]M. Dafermos, I. Rodnianski, and Y. Shlapentokh-Rothman, "A scattering theory for the wave equation on Kerr black hole exteriors", arXiv:1412.8379 (2014).

[33]J. Dimock, "Scattering for the wave equation on the Schwarzschild metric", General Relativity and Gravitation **17**, 353–369 (1985).

[34]J. Dimock and B. S. Kay, "Scattering for massive scalar fields on Coulomb potentials and Schwarzschild metrics", Classical and Quantum Gravity **3**, 71 (1986).

[35]F. Dyson, Phys. Rev. **85**, 631 (1952).

[36]D. M. Eardley and V. Moncrief, "The global existence of Yang–Mills–Higgs fields in 4-dimensional Minkowski space. II. Completion of proof", Comm. Math. Phys. **83**, 193–212 (1982).

[37]D. M. Eardley and V. Moncrief, "The global existence of Yang–Mills–Higgs fields in 4-dimensional Minkowski space. I. Local existence and smoothness properties", Comm. Math. Phys. **83**, 171–191 (1982).

[38]H. Elvang and Y.-T. Huang, *Scattering amplitudes in gauge theory and gravity* (Cambridge University Press, 2015).

[39]F. Esposito and L. Witten, *Asymptotic structure of space-time* (Springer, Boston, MA, 1977).

[40]L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Textbooks in Mathematics (CRC Press, 2015).

[41]G. B. Folland, *Introduction to partial differential equations*, 2nd ed. (Princeton University Press, Chichester, West Sussex, 1995).

[42]F. G. Friedlander, "On the radiation field of pulse solutions of the wave equation", Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **269**, 53–65 (1962).

[43]F. G. Friedlander, "On the radiation field of pulse solutions of the wave equation II", Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **279**, 386–394 (1964).

[44]F. G. Friedlander, "On the radiation field of pulse solutions of the wave equation III", Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **299**, 264–278 (1967).

[45]F. G. Friedlander, "Radiation fields and hyperbolic scattering theory", Mathematical Proceedings of the Cambridge Philosophical Society **88**, 483–515 (1980).

[46]F. G. Friedlander, *The wave equation on a curved space-time* (Cambridge University Press, Cambridge, UK, 1975), p. 282.

[47]H. Friedrich, "Existence and structure of past asymptotically simple solutions of Einstein's field equations with positive cosmological constant", Journal of Geometry and Physics **3**, 101–117 (1986).

[48]H. Friedrich, "On the global existence and the asymptotic behavior of solutions to the Einstein–Maxwell–Yang–Mills equations", J. Differential Geom. **34**, 275–345 (1991).

[49]G. Gabrielse, D. Hanneke, T. Kinoshita, M. Nio, and B. Odom, "Erratum to New determination of the fine structure constant from the electron g value and QED", Phys. Rev. Lett. **99** (2007).

[50]G. Gabrielse, D. Hanneke, T. Kinoshita, M. Nio, and B. Odom, "New determination of the fine structure constant from the electron g value and QED", Phys. Rev. Lett. **99** (2006).

[51]V. Georgiev and P. P. Schirmer, "Global existence of low regularity solutions of non-linear wave equations", Mathematische Zeitschrift **219**, 1–19 (1995).

[52]S. Ghanem, "The global non-blow-up of the Yang–Mills curvature on curved space-times", Journal of Hyperbolic Differential Equations **13**, 603–631 (2016).

[53]J. Ginibre and G. Velo, "Regularity of solutions of critical and subcritical nonlinear wave equations", Nonlinear Analysis: Theory, Methods & Applications **22**, 1–19 (1994).

[54]J. Ginibre and G. Velo, "The Cauchy problem for coupled Yang–Mills and scalar fields in the temporal gauge", Comm. Math. Phys. **82**, 1–28 (1981).

[55]M. V. Goganov and L. V. Kapitanskii, "Global solvability of the Cauchy problem for the Yang–Mills–Higgs equations", Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **37**, 18–48 (1985).

[56]D. Häfner and J.-P. Nicolas, "Scattering of massless Dirac fields by a Kerr black hole", Reviews in Mathematical Physics **16**, 29–123 (2004).

[57]P. Hintz and A. Vasy, "The global non-linear stability of the Kerr–de Sitter family of black holes", arXiv:1606.04014.

[58]L. Hörmander, "A remark on the characteristic Cauchy problem", Journal of Functional Analysis **93**, 270–277 (1990).

[59]A. Jaffe and C. Taubes, *Vortices and monopoles: structure of static gauge theories* (Birkhäuser Verlag, Boston-Basel-Stuttgart, 1980).

[60]J. Joudioux, "Conformal scattering for a nonlinear wave equation on a curved background", arXiv:1004.1464 (2010).

[61]J. Joudioux, "Hörmander's method for the characteristic Cauchy problem and conformal scattering for a nonlinear wave equation", arXiv:1903.12591 (2019).

[62]C. Kehle and Y. Shlapentokh-Rothman, "A scattering theory for linear waves on the interior of Reissner–Nordström black holes", arXiv:1804.05438 (2018).

[63]D. Kenneflick, "Who is afraid of the referee? Einstein and gravitational waves, talk at Stony Brook", http://130.184.202.6/Physics/Referee.pdf (2005).

[64]S. Klainerman and M. Machedon, "Finite energy solutions of the Yang–Mills equations in $\mathbb{R}^{3+1}$", Annals of Mathematics **142**, 39–119 (1995).

[65]S. Klainerman and M. Machedon, "On the Maxwell–Klein–Gordon equation with finite energy.", Duke Math. J. **74**, 19–44 (1994).

[66]S. Klainerman and M. Machedon, "Space-time estimates for null forms and the local existence theorem", Communications on Pure and Applied Mathematics **46**, 1221–1268 (1993).

[67]S. Klainerman and F. Nicolò, "Corrigendum for peeling properties of asymptotically flat solutions to the Einstein vacuum equations", Classical and Quantum Gravity **21**, 1925 (2004).

[68]S. Klainerman and F. Nicolò, "On local and global aspects of the Cauchy problem in general relativity", Classical and Quantum Gravity **16**, R73–R157 (1999).

[69]S. Klainerman and F. Nicolò, "Peeling properties of asymptotically flat solutions to the Einstein vacuum equations", Classical and Quantum Gravity **20**, 3215–3257 (2003).

[70]S. Klainerman and I. Rodnianski, "A Kirchoff–Sobolev parametrix for the wave equation and applications", arXiv:math/0603009 (2006).

[71]J. Krieger and J. Luhrmann, "Concentration compactness for the critical Maxwell–Klein–Gordon equation", arXiv:1503.09101.

[72]P. P. Kulish and L. D. Faddeev, "Asymptotic conditions and infrared divergences in quantum electrodynamics", Russian, Theor. Math. Phys. **4**, 153–170 (1970).

[73]P. D. Lax and R. S. Phillips, "Scattering theory", English, Bull. Amer. Math. Soc. **70**, 130–142 (1964).

[74]P. D. Lax and R. S. Phillips, *Scattering theory*, English (Academic Press, New York and London, 1967), p. 276.

[75]J. Leray, *Hyperbolic Differential Equations* (unpublished, 1953).

[76]H. Lindblad and J. Sterbenz, "Global stability for charged scalar fields on Minkowski space", arXiv:math/0410499.

[77]L. Mason and J.-P. Nicolas, "Peeling of Dirac and Maxwell fields on a Schwarzschild background", Journal of Geometry and Physics **62**, 867–889 (2012).

[78]L. J. Mason and J.-P. Nicolas, "Conformal scattering and the Goursat problem", Journal of Hyperbolic Differential Equations **01**, 197–233 (2004).

[79]L. J. Mason and J.-P. Nicolas, "Regularity at spacelike and null infinity", arXiv:gr-qc/0701049 (2007).

[80]F. Melnyk, "The Hawking effect for spin 1/2 fields", Communications in Mathematical Physics **244**, 483–525 (2004).

[81]R. Melrose, A. Sá Barreto, and A. Vasy, "Asymptotics of solutions of the wave equation on de Sitter–Schwarzschild space", Commun. Partial Differ. Equ. **39**, 512–529 (2014).

[82]M. Mokdad, "Conformal scattering of Maxwell fields on Reissner–Nordström–de Sitter black hole spacetimes", arXiv:1706.06993 (2017).

[83]C. S. Morawetz, "The decay of solutions of the exterior initial-boundary value problem for the wave equation", Communications on Pure and Applied Mathematics **14**, 561–568 (1961).

[84]G. Moschidis, "A proof of the instability of AdS for the Einstein–massless Vlasov system", arXiv:1812.04268.

[85]G. Moschidis, "A proof of the instability of AdS for the Einstein–null dust system with an inner mirror", arXiv:1704.08681.

[86]C. Nash and R. L. Stuller, "Infrared behaviour of Yang–Mills theories", Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences **78**, 217–233 (1978).

[87]E. Newman and R. Penrose, "An approach to gravitational radiation by a method of spin coefficients", Journal of Mathematical Physics **3**, 566–578 (1962).

[88]J.-P. Nicolas and T. X. Pham, "Peeling on Kerr spacetime: linear and non linear scalar fields", arXiv:1801.08996 (2018).

[89]J.-P. Nicolas, "Conformal scattering on the Schwarzschild metric", arXiv:1312.1386 (2013).

[90]J.-P. Nicolas, "On Lars Hörmander's remark on the characteristic Cauchy problem", Comptes Rendus Mathematique **344**, 621–626 (2007).

[91]J.-P. Nicolas, "The conformal approach to asymptotic analysis", arXiv:1508.02592 (2015).

[92]S.-J. Oh, "Finite energy global well-posedness of the Yang–Mills equations on $\mathbb{R}^{1+3}$: an approach using the Yang–Mills heat flow", Duke Math.J. **164**, 1669–1732 (2015).

[93]R. Penrose, "Asymptotic properties of fields and space-times", Physical Review Letters **10**, 66–68 (1963).

[94]R. Penrose, "Zero rest-mass fields including gravitation: asymptotic behaviour", Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **284**, 159–203 (1965).

[95]R. Penrose and W. Rindler, *Spinors and space-time*, Vol. 1 (Cambridge University Press, 1984).

[96]R. Penrose and W. Rindler, *Spinors and space-time*, Vol. 2 (Cambridge University Press, 1986).

[97]S. Perlmutter et al., "Measurements of Ω and Λ from 42 high-redshift supernovae", The Astrophysical Journal **517**, 565 (1999).

[98] S. Perlmutter, "Supernovae, dark energy, and the accelerating universe: the status of the cosmological parameters", Int. J. Mod. Phys. A **15S1**, Proc. of the 19th Intl. Symp. on Photon and Lepton Interactions at High Energy LP99, ed. J.A. Jaros and M.E. Peskin, 715 (2000).

[99] A. G. Riess et al., "Observational evidence from supernovae for an accelerating universe and a cosmological constant", The Astronomical Journal **116**, 1009 (1998).

[100] H. Ringström, "Future stability of the Einstein-non-linear scalar field system", Inventiones Mathematicae **173**, 123–208 (2008).

[101] S. Rosenberg, *The Laplacian on a Riemannian Manifold: an Introduction to Analysis on Manifolds*, London Mathematical Society Student Texts (London Mathematical Society, 1997).

[102] W. Rudin, *Real and complex analysis*, McGraw–Hill International Editions: Mathematics Series (McGraw–Hill Education, 1987).

[103] R. K. Sachs and B. Hermann, "Gravitational waves in general relativity VI. The outgoing radiation condition", Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **264**, 309–338 (1961).

[104] R. K. Sachs and B. Hermann, "Gravitational waves in general relativity VIII. Waves in asymptotically flat space-time", Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **270**, 103–126 (1962).

[105] G. Satishchandran and R. M. Wald, "The asymptotic behavior of massless fields and the memory effect", arXiv:1901.05942 (2019).

[106] B. P. Schmidt et al., "The High Z supernova search: measuring cosmic deceleration and global curvature of the universe using type Ia supernovae", Astrophys. J. **507**, 46–63 (1998).

[107] I. Segal, Ann. Math. **78** (1963).

[108] I. Segal, J. Funct. Anal. **33** (1979).

[109] S. Selberg and A. Tesfahun, "Null structure and local well-posedness in the energy class for the Yang–Mills equations in Lorenz gauge", arXiv:1309.1977 (2013).

[110] C. D. Sogge, *Lectures on non-linear wave equations*, 2nd Edition (International Press, 1995).

[111] C. D. Sogge, "On local existence for nonlinear wave equations satisfying variable coefficient null conditions", Communications in Partial Differential Equations **18**, 1795–1821 (1993).

[112] M. Spradlin, A. Strominger, and A. Volovich, "Les Houches lecture notes on de Sitter Space", arXiv:hep-th/0110007 (2001).

[113] F. Strocchi, *An introduction to non-perturbative foundations of quantum field theory*, International Series of Monographs on Physics (Oxford University Press, 2013).

[114] C. Svedberg, "Future stability of the Einstein-Maxwell-scalar field system", Annales Henri Poincaré **12**, 849 (2011).

[115] T. Tao, "Local well-posedness of the Yang–Mills equation in the temporal gauge below the energy norm", Journal of Differential Equations **189**, 366–382 (2003).

[116] G. Taujanskas, J.-P. Nicolas, and L. Mason, "Finite energy well-posedness of the Maxwell-scalar field system on the Einstein cylinder", In preparation.

[117]G. Taujanskas, "Conformal scattering of the Maxwell-scalar field system on de Sitter space", J. Hyperbolic Differ. Equ. **16**, 743–791 (2019).

[118]G. Taujanskas, "Large data decay of Yang–Mills–Higgs fields on Minkowski and de Sitter spacetimes", J. Math. Phys. **60**, 121504 (2019).

[119]The LIGO Scientific Collaboration and The Virgo Collaboration, "An improved analysis of GW150914 using a fully spin-precessing waveform model", Phys. Rev. X **6** (2016).

[120]The LIGO Scientific Collaboration and The Virgo Collaboration, "Astrophysical implications of the binary black-hole merger GW150914", The Astrophysical Journal **818** (2016).

[121]The LIGO Scientific Collaboration and The Virgo Collaboration, "Tests of general relativity with GW150914", Phys. Rev. Lett. **116** (2016).

[122]M. Van de Moortel, "Stability and instability of the sub-extremal Reissner–Nordström black hole interior for the Einstein–Maxwell–Klein–Gordon equations in spherical symmetry", Comm. Math. Phys. **360**, 103–168 (2017).

[123]E. T. Whittaker, "On the partial differential equations of mathematical physics", Math. Ann. **57**, 333 (1903).

[124]C. M. Will, "The confrontation between general relativity and experiment", arXiv:1403.7377 (2014).

[125]S. Yang and P. Yu, "On global dynamics of the Maxwell–Klein–Gordon equations", arXiv:1804.00078 (2018).