

<https://helda.helsinki.fi>

---

## Detection of Hermitian connections in wave equations with cubic non-linearity

Chen, Xi

2022

---

Chen , X , Lassas , M , Oksanen , L & Paternain , G P 2022 , ' Detection of Hermitian connections in wave equations with cubic non-linearity ' , European Mathematical Society. Journal , vol. 24 , no. 7 , pp. 2191-2232 . <https://doi.org/10.4171/JEMS/1136>

---

<http://hdl.handle.net/10138/344708>

<https://doi.org/10.4171/JEMS/1136>

---

cc\_by

publishedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*



Xi Chen · Matti Lassas · Lauri Oksanen · Gabriel P. Paternain

# Detection of Hermitian connections in wave equations with cubic non-linearity

*To the memory of our friend and colleague Slava Kurylev*

Received May 8, 2019; revised October 8, 2020

**Abstract.** We consider the geometric non-linear inverse problem of recovering a Hermitian connection  $A$  from the source-to-solution map of the cubic wave equation  $\square_A \phi + \kappa |\phi|^2 \phi = f$ , where  $\kappa \neq 0$  and  $\square_A$  is the connection wave operator in the Minkowski space  $\mathbb{R}^{1+3}$ . The equation arises naturally when considering the Yang–Mills–Higgs equations with Mexican hat type potentials. Our proof exploits the microlocal analysis of non-linear wave interactions, but instead of employing information contained in the geometry of the wave front sets as in previous literature, we study the principal symbols of waves generated by suitable interactions. Moreover, our approach relies on inversion of a novel non-abelian broken light ray transform, a result interesting in its own right.

**Keywords.** Inverse problems, non-linear hyperbolic equation, Hermitian connection

---

## Contents

1. Introduction	2192
1.1. The Yang–Mills–Higgs equations	2193
1.2. Formulation of the main results	2193
1.3. Previous literature	2196
1.4. A conjecture on higher order non-linearities	2198
1.5. Outline of the paper	2198
2. Preliminaries	2199
2.1. Direct problem	2199

---

Xi Chen: Shanghai Center for Mathematical Sciences, Fudan University, Shanghai 200438, China; [xi\\_chen@fudan.edu.cn](mailto:xi_chen@fudan.edu.cn)

Matti Lassas: Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland; [Matti.Lassas@helsinki.fi](mailto:Matti.Lassas@helsinki.fi)

Lauri Oksanen: Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland; [lauri.oksanen@helsinki.fi](mailto:lauri.oksanen@helsinki.fi)

Gabriel P. Paternain: Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge CB3 0WB, UK; [g.p.paternain@dpmms.cam.ac.uk](mailto:g.p.paternain@dpmms.cam.ac.uk)

*Mathematics Subject Classification (2020):* 35R30, 58J45

2.2. Notations for microlocal analysis . . . . .	2200
2.3. Microlocal analysis of the wave operator . . . . .	2201
2.4. Flowout from a point in the Minkowski space . . . . .	2203
2.5. Trivialization of the half-density bundle over the flowout . . . . .	2205
2.6. Parallel transport equation for the principal symbol . . . . .	2207
2.7. Positively homogeneous symbols . . . . .	2208
3. Microlocal analysis of the interaction of three waves . . . . .	2210
3.1. The linear span of three lightlike vectors . . . . .	2211
3.2. Three-fold linearization of cubic connection wave equations . . . . .	2213
3.3. Wave interactions as products of conormal distributions . . . . .	2214
3.4. From the source-to-solution map to the broken light ray transform . . . . .	2217
4. Inversion of the broken light ray transform . . . . .	2219
A. Conormal distributions and IPL distributions . . . . .	2222
B. Theorem 3 in the context of Example 1 . . . . .	2226
B.1. Trivialization of the Maslov bundle over the flowout . . . . .	2228
C. Flowout from the triple intersection . . . . .	2228
References . . . . .	2230

## 1. Introduction

This paper considers an inverse problem for a non-linear wave equation motivated by theoretical physics and differential geometry. The main problem we wish to address is the following: can the geometric structures governing the wave propagation be globally determined from local information, or more physically, can an observer do local measurements to determine the geometric structures in the maximal region where the waves can propagate and return back? There has been recent progress on this question when the geometric structure is space-time itself and the relevant PDEs are the Einstein equations [22].

Here we propose the study of a natural non-linear wave equation when the Lorentzian background is fixed and the goal is the reconstruction of a Hermitian connection. The main difference between the inverse problems for the Einstein equations and the equation considered here is that, in the former case, the geometric structure (the metric) to be reconstructed appears in the leading order terms, and in the latter case, the structure (the connection) appears in the lower order terms. This difference poses novel challenges, since a perturbation in the leading order affects the wave front sets of solutions whereas lower order perturbations do not.

The leading order terms can frequently be reconstructed via study of distances (or time separations/earliest arrival times), whereas lower order terms often require reductions to light ray transform questions. Nevertheless, our approach exploits the recent philosophy that non-linear interaction of waves creates new singularities and enriches the dynamics [20, 22, 27]. As we shall see, this interaction leads to a broken non-abelian ray transform on lightlike geodesics that has not been previously studied.

Our main long-term goal is the study of inverse problems for the Yang–Mills–Higgs equation. The present paper is the first step in this direction and our objective here is to start exposing the main features that this problem will have by considering a simplified, but non-trivial model case.

1.1. The Yang–Mills–Higgs equations

Let  $(M, g)$  be a Lorentzian manifold of dimension  $1 + 3$  and consider a compact Lie group  $G$  with Lie algebra  $\mathfrak{g}$ . We choose a positive definite inner product on  $\mathfrak{g}$  invariant under the adjoint action. To simplify the exposition we discuss the case of the trivial bundle  $M \times \mathfrak{g}$  over  $M$ . The Yang–Mills–Higgs equations are PDEs on a pair  $(A, \Phi)$ , where

$$(A, \Phi) \in C^\infty(M; T^*M \otimes \mathfrak{g}) \oplus C^\infty(M; \mathfrak{g}).$$

Since the bundle  $M \times \mathfrak{g}$  is trivial,  $A$  is a connection, and it is called the Yang–Mills potential;  $\Phi$  is the Higgs field. The Yang–Mills–Higgs equations are

$$D_A^* F_A + [\Phi, D_A \Phi] = 0; \tag{1}$$

$$D_A^* D_A \Phi + V'(|\Phi|^2)\Phi = 0, \tag{2}$$

where  $F_A := dA + A \wedge A$  is the curvature of  $A$ ,  $D_A \Phi := d\Phi + [A, \Phi]$  is the associated covariant derivative, and  $V'$  is the derivative of a smooth function  $V : [0, \infty) \rightarrow \mathbb{R}$ . The adoint  $D_A^*$  is taken with respect to  $g$  and hence  $\square_A := D_A^* D_A$  is the wave operator associated with  $g$  and  $A$ .

An extensively studied case is the Yang–Mills–Higgs equations with the Mexican hat type potential,

$$V(|\Phi|^2) = \frac{1}{2}\kappa(|\Phi|^2 - b)^2, \tag{3}$$

where  $\kappa, b \in \mathbb{R}$ ; see e.g. [8, (10.5)] where the Lagrangian formulation of the problem is used. We will consider the potential (3) with  $\kappa \neq 0$ , and to simplify the notations, with  $b = 0$ . The case  $b \neq 0$  is not substantially different. Our choice can be viewed as the simplest potential introducing a non-linearity. We refer also to [37] where Yang–Mills–Higgs equations, with the potential (3), are discussed in a purely mathematical context,  $(M, g)$  being a Riemannian manifold there.

As is well known, equations (1)–(2) are invariant under the group of gauge transformations which in this case coincides with the set of maps  $\mathbf{u} \in C^\infty(M; G)$  and the action on pairs is

$$(A, \Phi) \mapsto (\mathbf{u}^{-1} d\mathbf{u} + \mathbf{u}^{-1} A \mathbf{u}, \mathbf{u}^{-1} \Phi \mathbf{u}).$$

When  $\Phi = 0$  we obtain the pure Yang–Mills equation  $D_A^* F_A = 0$ .

1.2. Formulation of the main results

Dealing with the equations (1)–(2) from the outset might be too ambitious, so here we propose a simplified model. We shall suppose that we have a trivial bundle  $E = M \times \mathbb{C}^n$  and a Hermitian connection  $A$  on  $E$  giving rise to a covariant derivative  $d + A$ . In this case, the gauge group is  $U(n)$ . We take  $V$  to be the Mexican hat type potential (3) with  $b = 0$ , discard equation (1) completely and focus on the analogue of equation (2) with  $M \times \mathfrak{g}$  replaced by  $E$ . That is, we consider the equation

$$\square_A \phi + \kappa |\phi|^2 \phi = 0, \tag{4}$$

where  $\phi$  is a section of  $E$ ,  $\square_A = (d + A)^*(d + A)$  and  $|\phi|$  is the norm with respect to the standard Hermitian inner product of  $\mathbb{C}^n$ . We shall further simplify matters by assuming that  $M$  is  $\mathbb{R}^{1+3}$  and  $g$  is the Minkowski metric.

Let us consider Cartesian coordinates  $(t = x^0, x^1, x^2, x^3)$  on  $\mathbb{R}^{1+3}$ . Let  $\epsilon_0 > 0$  and define  $B(\epsilon_0) = \{y \in \mathbb{R}^3 : |y| < \epsilon_0\}$  and

$$\mathfrak{U} = (0, 1) \times B(\epsilon_0). \tag{5}$$

We will define a source-to-solution map, associated to (4), that corresponds physically to measurements gathered in  $\mathfrak{U}$ . We can think that the measurements are performed by an observer travelling along the path

$$\mu : [0, 1] \rightarrow \mathbb{R}^{1+3}, \quad \mu(t) = (t, 0). \tag{6}$$

In what follows, we consider only finite time intervals, and write  $\mathcal{M} = (-1, 2) \times \mathbb{R}^3$ . and let  $\mathcal{C}$  be a small neighbourhood of the zero section in  $C_0^4(\mathfrak{U}; E)$ . Then the source-to-solution map

$$L_A f := \phi|_{\mathfrak{U}}, \quad f \in \mathcal{C}, \tag{7}$$

is well-defined, where  $\phi$  is the solution of

$$\begin{aligned} \square_A \phi + \kappa |\phi|^2 \phi &= f \quad \text{in } \mathcal{M}, \\ \phi|_{t < 0} &= 0. \end{aligned} \tag{8}$$

We discuss the existence of  $L_A$  in more detail in Section 2 below.

The goal of the observer is to determine the Yang–Mills potential  $A$  up to the natural obstructions, given the source-to-solution map  $L_A$ . The causal structure of  $(M, g)$  encodes the finite speed of propagation for the wave equation (4). Given  $x, y \in M$  we say that  $x \leq y$  if  $x = y$  or  $x$  can be joined to  $y$  by a future pointing causal curve, and denote the causal future of  $x \in M$  by  $\mathcal{J}^+(x) = \{y \in M : x \leq y\}$ . The causal future  $\mathcal{J}^+(x)$  is the largest set that waves generated at  $x$  can reach. The causal past of a point  $z \in M$  is denoted by  $\mathcal{J}^-(z) = \{y \in M : y \leq z\}$ . If waves generated at  $x$  are recorded at  $z$ , the finite speed of propagation dictates that no information on the potential  $A$  outside the causal diamond  $\mathcal{J}^+(x) \cap \mathcal{J}^-(z)$  is obtained.

The *model inverse problem* is to determine  $A$  given  $L_A$  in the largest domain possible, that is, in

$$\mathbb{D} = \bigcup_{x, z \in \mathfrak{U}} (\mathcal{J}^+(x) \cap \mathcal{J}^-(z)), \tag{9}$$

up to the natural gauge,

$$A \mapsto \mathbf{u}^{-1} d\mathbf{u} + \mathbf{u}^{-1} A \mathbf{u}, \tag{10}$$

where  $\mathbf{u} \in C^\infty(\mathbb{D}; U(n))$  and  $\mathbf{u}|_{\mathfrak{U}} = \text{id}$ . The sets  $\mathfrak{U}$  and  $\mathbb{D}$  are visualized in Figure 2.

Observe that if we have two connections  $A$  and  $B$  on  $M$  such that there exists a smooth map  $\mathbf{u} : \mathcal{M} \rightarrow U(n)$  with the property that  $B = \mathbf{u}^{-1} d\mathbf{u} + \mathbf{u}^{-1} A \mathbf{u}$  and  $\mathbf{u}|_{\mathfrak{U}} = \text{id}$ ,

then  $\square_B = \mathbf{u}^{-1}\square_A\mathbf{u}$  and  $|\mathbf{u}\phi| = |\phi|$ . Moreover, as  $f$  has compact support in  $\mathfrak{U}$  we have  $\mathbf{u}f = f$ . Therefore  $\phi$  solves (8) for  $B$  if and only if  $\mathbf{u}\phi$  solves (8) for  $A$ , and it follows that  $L_A = L_B$ . This shows that the gauge (10) is indeed natural.

Our main theorem asserts that the model inverse problem has a unique solution, or in more physical terms, the measurements performed on  $\mathfrak{U}$ , as encoded by  $L_A$ , determine the gauge equivalence class of the Yang–Mills potential  $A$ , in the largest possible causal diamond  $\mathbb{D}$ . As  $\mathbb{D}$  is strictly larger than  $\mathfrak{U}$ , we can view the determination of the equivalence class of  $A$  as a form of remote sensing. We emphasize that the gauge equivalence classes of Yang–Mills potentials, not the potentials themselves, correspond to physically distinct configurations.

**Theorem 1.** *Let  $A$  and  $B$  be two connections in  $\mathbb{R}^{1+3}$  such that  $L_A = L_B$  where the source-to-solution map  $L_A$  is defined as above, and  $L_B$  is defined analogously, with  $A$  replaced by  $B$  in (8). Suppose that  $\kappa \neq 0$  in (8). Then there exists a smooth  $\mathbf{u} : \mathbb{D} \rightarrow U(\mathbb{N})$  such that  $\mathbf{u}|_{\mathfrak{U}} = \text{id}$  and  $B = \mathbf{u}^{-1}d\mathbf{u} + \mathbf{u}^{-1}A\mathbf{u}$ .*

It is straightforward to see that  $L_A = L_B$  implies  $A = B$  on  $\mathfrak{U}$ . The non-trivial content of the theorem is the gauge equivalence away from  $\mathfrak{U}$ . To see that  $A$  and  $B$  coincide on  $\mathfrak{U}$ , we fix  $y \in \mathfrak{U}$  and choose  $\phi \in C_0^\infty(\mathfrak{U}; E)$  such that  $\phi(y) = 0$ . Then for small  $\epsilon > 0$  we have  $f := \epsilon(\square_A\phi + \kappa|\phi|^2\phi) \in \mathcal{C}$ . Since  $L_A = L_B$  we see that at  $y$ ,

$$-A_0\partial_t\phi + \sum_{j=1}^3 A_j\partial_{x^j}\phi = -B_0\partial_t\phi + \sum_{j=1}^3 B_j\partial_{x^j}\phi$$

(cf. (16) below), and since  $d\phi$  at  $y$  is arbitrary, we have  $A = B$  at  $y$ .

The proof of Theorem 1 is based on analysis of interactions of three singular waves. To explain this briefly, consider a source of the form  $f = \epsilon_1 f_1 + \epsilon_2 f_2 + \epsilon_3 f_3$  where  $\epsilon_j > 0$  are small and  $f_j$  are conormal distributions, and let  $\phi = \phi(\epsilon)$ , with  $\epsilon = (\epsilon_1, \epsilon_2, \epsilon_3)$ , be the corresponding family of solutions to (8). Then the cross-derivative

$$\partial_{\epsilon_1}\partial_{\epsilon_2}\partial_{\epsilon_3}\phi|_{\epsilon=0} \tag{11}$$

satisfies a linear wave equation with a right-hand side that corresponds to a certain product of  $\partial_{\epsilon_j}\phi|_{\epsilon=0}$ ,  $j = 1, 2, 3$ . As also the functions  $\partial_{\epsilon_j}\phi|_{\epsilon=0}$  satisfy the linear wave equation, we can view the cross-derivative as a result of their interaction. In a suitable microlocal sense, the cross-derivative has a principal symbol, and we use information contained in the principal symbol in order to recover a novel broken non-abelian light ray transform of the connection  $A$ , which we will define next.

For a lightlike geodesic segment  $\gamma : [0, \ell] \rightarrow \mathbb{D}$ , denote by  $\mathbf{P}_\gamma^A$  the parallel transport map with respect to  $A$  from  $\gamma(0)$  to  $\gamma(\ell)$  along  $\gamma$ . We recall that  $\mathbf{P}_\gamma^A$  is obtained as the solution map to a differential equation (see (35) below). The broken non-abelian light ray transform of the connection  $A$  along lightlike geodesics is the composition  $\mathbf{S}_{\gamma_2\gamma_1}^A = \mathbf{P}_{\gamma_2}^A\mathbf{P}_{\gamma_1}^A$  parametrized by two lightlike geodesic segments  $\gamma_j : [0, \ell_j] \rightarrow \mathbb{D}$ ,  $j = 1, 2$ , satisfying

$$\gamma_1(0), \gamma_2(\ell_2) \in \mathfrak{U}, \quad \gamma_1(\ell_1) = \gamma_2(0). \tag{12}$$

We prove the following theorem:

**Theorem 2.** *Let  $A$  and  $B$  be two connections in  $\mathbb{R}^{1+3}$  such that  $\mathbf{S}_{\gamma_2\gamma_1}^A = \mathbf{S}_{\gamma_2\gamma_1}^B$  for all lightlike geodesic segments  $\gamma_j : [0, \ell_j] \rightarrow \mathbb{D}$ ,  $j = 1, 2$ , satisfying (12). Then there exists a smooth  $\mathbf{u} : \mathbb{D} \rightarrow U(n)$  such that*

$$\mathbf{u}|_{\mathfrak{U}} = \text{id} \quad \text{and} \quad B = \mathbf{u}^{-1}d\mathbf{u} + \mathbf{u}^{-1}A\mathbf{u}.$$

In more physical terms, we can say that the interaction of the three waves  $\partial_{\epsilon_j}\phi|_{\epsilon=0}$ ,  $j = 1, 2, 3$ , produces an artificial source, which can be viewed either as two moving point sources or as a filament in spacetime, and which emits a wave encoded by the cross-derivative  $\partial_{\epsilon_1}\partial_{\epsilon_2}\partial_{\epsilon_3}\phi|_{\epsilon=0}$ . We show that when the sources  $f_j$ ,  $j = 1, 2, 3$ , are chosen carefully, the singular wave front emitted by the artificial source returns to  $\mathfrak{U}$ . This wave front is visualized in Figure 1. Stretching the physical analogy further, we can say the leading amplitude of this singular wave front gives us the information that allows us to apply Theorem 2 in the proof of Theorem 1.

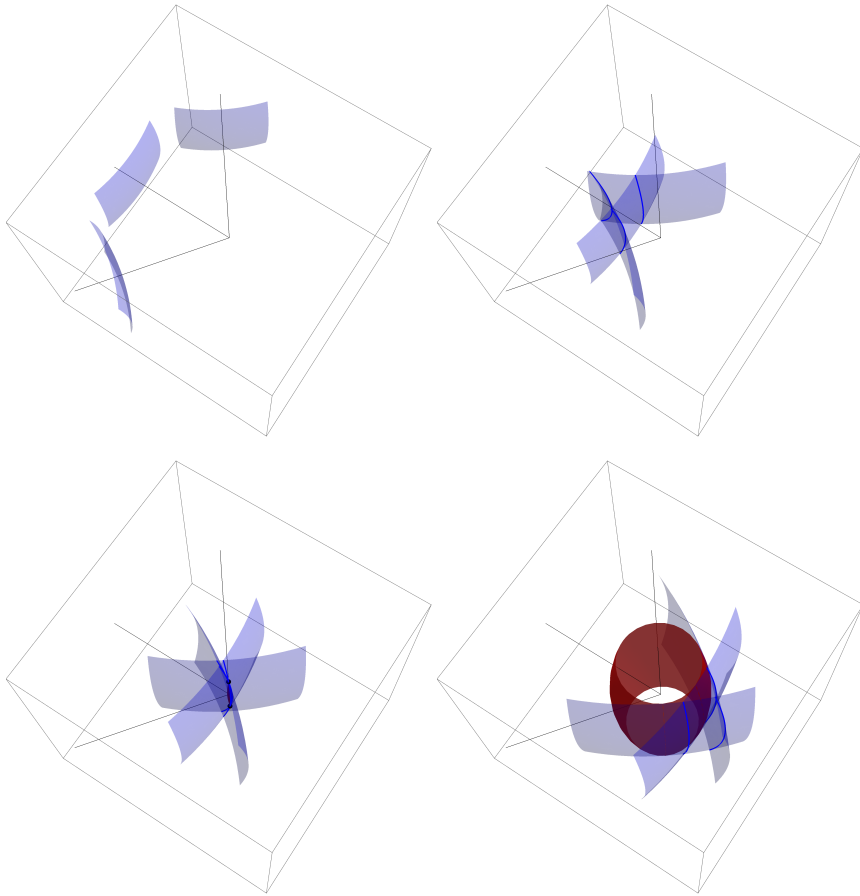
### 1.3. Previous literature

Interaction of singular waves has been studied outside the context of inverse problems. In particular, the wave front set of the cross-derivative (11) was studied in the case of the 1 + 2-dimensional Minkowski space by Rauch and Reed [32]. We also mention [6, 30] and the recent preprint [33] for later results of similar nature. The previous results on inverse problems for non-linear wave equations [20, 22, 27] are based on the study of the wave front set resulting from the interaction of four singular waves.

In contrast to all the above mentioned works, in the present paper we employ more precise information on the singular structure of the interaction, encoded by (11), than just its wave front set. Namely, as already mentioned above, we compute the principal symbol of (11). The principal symbol contains information on the connection whereas the wave front set does not.

Paradoxically, Theorem 1 is open for the linear case,  $\kappa = 0$ , but a positive solution is known if  $A$  and  $B$  are supposed to be time-independent [21]. In the time-dependent case there are results [35] available only in the abelian case of a line bundle,  $n = 1$ , and it is an open problem if recovery of  $A$  in the optimal causal diamond  $\mathbb{D}$  is possible in this case. The result [35] assumes that the region where the measurements are available, corresponding to our  $\mathfrak{U}$ , encloses the region where  $A$  is to be determined, corresponding to our  $\mathbb{D}$ , in the sense that the latter is a cylinder in the Minkowski space, and the former is the timelike part of the boundary of the cylinder. This can be viewed as an additional (pseudo)convexity assumption, not imposed in the present paper.

Let us also mention that the linear, abelian, time-independent case has been studied extensively (see e.g. [2, 3, 17]), but these results do not carry over to the time-dependent case. The reason is that the results (and also those in [21]) are based on Tataru’s unique continuation principle [36], which again is known to fail for equations with time-dependent coefficients [1].



**Fig. 1.** The interaction of three pieces of spherical waves, the blue surfaces. *Top left:* The pieces propagate along the black lines, and do not intersect yet. *Top right:* The pieces intersect along the blue curves. Pairwise intersections do not produce new propagating wave fronts. *Bottom left:* All the three pieces intersect in the two black points, moving along the vertical axis over the point where the black lines intersect. The points act as artificial sources that produce a new propagating wave, the red surface. The line segment traced by the two points can also be viewed as the projection of a one-dimensional filament acting as an artificial source. The filament curves in spacetime since the two points move with a non-constant speed. *Bottom right:* As time progresses, the red propagating wave front grows. Eventually it will reach the points where the pieces of the spherical waves originate from.

We emphasize once more that the focus of the current paper is on the recovery of the lower order terms in a non-linear wave equation. This makes definite progress towards Open Problem 5 in [25], and is different from the previous results where only the determination of the leading order terms is considered. See for instance [20, 27] where the determination of the metric tensor (or its conformal class) is studied for scalar-valued



non-linear equations, or [22, 28] where the determination of the metric tensor is studied for the Einstein equations coupled with different matter field equations. The difference between recovery of leading and lower order terms is reflected in the key novelty of our approach, the study of principal symbols instead of wave front sets.

The difference is also apparent in the existing theory of inverse problems for linear wave equations. The case of a linear wave equation with time-independent coefficients, and with sources and observations in disjoint sets, illustrates this. In this case the theory is still under active development, and the best results available are very different for leading and lower order terms: the recovery of the metric [26] is based on distance functions, whereas the recovery of the lower order terms [19] is based on focussing of waves. The latter also requires additional convexity assumptions, not present in the former case.

#### 1.4. A conjecture on higher order non-linearities

One outcome of the current paper is the following emergent principle for dealing with inverse problems for waves with polynomial non-linearities using an approach similar to ours. Assume for simplicity that we are in the line bundle case (i.e.  $n = 1$ ) and consider an equation of the form

$$\square_A \phi + \phi^N = f, \quad (13)$$

where  $\kappa \neq 0$ . Then to recover  $A$  (up to gauge) from a source-to-solution map it is necessary to consider the  $J$ -fold linearization of (13), where

$$J \geq \max \{3, N\}. \quad (14)$$

The necessity is discussed further in Remarks 2 and 4 below. We conjecture that (14) is also a sufficient condition, but the present paper establishes this only in the case  $N = 3$ .

As mentioned above, the previous results on inverse problems for non-linear wave equations are based on analysis of four singular, interacting waves. The use of only three waves in the present paper leads to a more economic proof, and is particularly well-suited to the cubic non-linearity in (8). A more detailed comparison of interaction three versus four waves is given at the beginning of Section 3.1.

#### 1.5. Outline of the paper

This paper is organized as follows. Section 2 contains preliminaries mostly having to do with the direct problem (8). Section 3 contains the microlocal analysis for the interaction of three waves and shows that we can recover the broken non-abelian light ray transform along lightlike geodesics from the knowledge of  $L_A$ . Section 4 solves the geometric inverse problem of determining  $A$  up to gauge from the broken non-abelian light ray transform and completes the proof of Theorem 1. Appendix A recalls the theory of conormal and Intersecting Pair of Lagrangian (IPL) distributions; Appendix B contains certain technical details concerning symplectic transformations to a model pair of intersecting Lagrangians, and Maslov bundles; and Appendix C gives a full description of the wave front set of the cross-derivative  $\partial_{\epsilon_1} \partial_{\epsilon_2} \partial_{\epsilon_3} \phi|_{\epsilon=0}$ , that is, the red surface in Figure 1.

We would like to dedicate this paper to the memory of our friend and colleague Slava Kurylev who was instrumental in initiating the present line of research on inverse problems for the Yang–Mills–Higgs equations.

## 2. Preliminaries

In this section, to accommodate further work, we let  $(M, g)$  be an arbitrary, globally hyperbolic Lorentzian manifold of dimension  $1 + m$ . Also  $E$  can be an arbitrary Hermitian vector bundle over  $M$ . Recall that a Lorentzian manifold  $(M, g)$  is globally hyperbolic if there are no closed causal paths in  $M$ , and the causal diamond  $\mathcal{J}^+(x) \cap \mathcal{J}^-(z)$  is compact for any pair of points  $x, z \in M$  (see [5]). A globally hyperbolic manifold  $(M, g)$  is isometric to a product manifold  $\mathbb{R} \times M_0$  with the Lorentzian metric given by

$$g = -c(t, x')dt^2 + g_0(t, x'), \quad (t, x') \in \mathbb{R} \times M_0, \tag{15}$$

where  $c : \mathbb{R} \times M_0 \rightarrow \mathbb{R}^+$  is smooth and  $g_0$  is a Riemannian metric on  $M_0$  depending smoothly on  $t$  (see [4]). Moreover, the vector field  $\partial_t$  gives time-orientation on  $M$ . To simplify the discussion, we make the further assumption that all the geodesics of  $(M, g)$  are defined on the whole  $\mathbb{R}$ .

### 2.1. Direct problem

We write occasionally  $\nabla = d + A$  for the covariant derivative associated to the connection  $A$ , and view it as a map

$$\nabla : C^\infty(M; E) \rightarrow C^\infty(M; T^*M \otimes E).$$

Writing  $g = g_{ij}dx^i dx^j$  in coordinates, we denote by  $|g|$  and  $g^{ij}$  the determinant and inverse of  $g_{ij}$ , respectively. Moreover,  $A = A_j dx^j$  is a 1-form, and each  $A_j$  is a skew-Hermitian matrix. Let us now write the wave operator  $\square_A = \nabla^* \nabla$  in coordinates. Consider compactly supported sections  $\phi$  and  $\psi = \psi_j dx^j$  of  $E$  and  $T^*M \otimes E$ , respectively. Then

$$\langle \nabla \phi, \psi \rangle_{L^2(M; T^*M \otimes E)} = \int_M g^{ij} \langle (\partial_{x^i} \phi + A_i \phi), \psi_j \rangle_E dV_g,$$

where  $\langle \cdot, \cdot \rangle_E$  is the inner product on  $E$ , and  $dV_g$  denotes the volume form on  $(M, g)$ . Integrating by parts and using the fact that  $A$  is skew-Hermitian, we have

$$\begin{aligned} \langle d\phi, \psi \rangle_{L^2(M; T^*M \otimes E)} &= - \int_M \langle \phi, |g|^{-1/2} \partial_{x^i} (|g|^{1/2} g^{ij} \psi_j) \rangle_E dV_g, \\ \langle A\phi, \psi \rangle_{L^2(M; T^*M \otimes E)} &= - \int_M \langle \phi, g^{ij} A_i \psi_j \rangle_E dV_g. \end{aligned}$$

Consequently,  $\square_A \phi$  takes the form

$$-|g|^{-1/2} \partial_{x^i} (|g|^{1/2} g^{ij} \partial_{x^j} \phi) - 2g^{ij} A_i \partial_{x^j} \phi - |g|^{-1/2} \partial_{x^i} (|g|^{1/2} g^{ij} A_j) \phi - g^{ij} A_i A_j \phi.$$

**Remark 1.** To prove Theorem 1, we will need the operator  $\square_A$  exclusively in Minkowski space where we can explicitly write

$$\square_A \phi = \square \phi - 2 \left( -A_0 \partial_t \phi + \sum_{j=1}^3 A_j \partial_{x_j} \phi \right) + \left( \operatorname{div} A + A_0^2 - \sum_{j=1}^3 A_j^2 \right) \phi, \tag{16}$$

where  $\square$  and  $\operatorname{div}$  are the usual wave operator and divergence,

$$\square \phi = \partial_t^2 \phi - \sum_{j=1}^3 \partial_{x_j}^2 \phi, \quad \operatorname{div} A = \partial_t A_0 - \sum_{j=1}^3 \partial_{x_j} A_j.$$

Let  $T > 0$  and let us consider the non-linear Cauchy problem

$$\begin{cases} \square_A \phi(t, x) = H(t, x, \phi(t, x)) + f(t, x) & \text{on } (-\infty, T) \times M_0, \\ \phi = 0, & \text{for } t < 0, \end{cases} \tag{17}$$

where  $H : (\mathbb{R} \times M_0) \times E \rightarrow E$  is a smooth map operating sectionwise such that  $H(t, x, 0) = 0$ , and  $f$  is a section of  $E$ . We will now give sufficient conditions on  $f$  in order for (17) to have a unique solution.

As the leading term of  $\square_A$  is simply the canonical wave operator on  $(M, g)$ , acting on each component of  $\phi$ , we can use the standard results for quasilinear hyperbolic equations to show existence and uniqueness of solutions to this Cauchy problem. See, for example, [18, Theorem 6 and its proof] (with the notations explained in detail in [23, Appendix C]) or [15, Theorems I–III and proof of Lemma 2.7]. By these results, we know that for an integer  $r > m/2 + 2$  and any compact set  $\mathcal{K} \subset (0, T) \times M_0$ , there is  $\epsilon_0 > 0$  such that for any  $f \in C_0^r(\mathcal{K})$  satisfying  $\|f\|_{C^r(\mathcal{K})} < \epsilon_0$ , the initial value problem (17) has a unique solution. Recall that  $m$  is the dimension of the underlying space  $M_0$ . In particular, in the case of the Minkowski space  $\mathbb{R}^{1+3}$ , we may take  $r = 4$ , and see that the source-to-solution map  $L_A$  is well-defined by (7).

### 2.2. Notations for microlocal analysis

For a conic Lagrangian submanifold  $\Lambda_0 \subset T^*M \setminus 0$  and a vector bundle  $\mathbb{E}$  over  $M$ , we denote by  $I^p(M; \Lambda_0; \mathbb{E})$  the space of Lagrangian distributions of order  $p \in \mathbb{R}$  associated to  $\Lambda_0$ , and taking values in  $\mathbb{E}$ . If  $\Lambda_1 \subset T^*M \setminus 0$  is another conic Lagrangian submanifold intersecting  $\Lambda_0$  cleanly, we denote by  $I^p(M; \Lambda_0, \Lambda_1; \mathbb{E})$  the space of Intersecting Pair of Lagrangian (IPL) distributions of order  $p \in \mathbb{R}$  associated to  $(\Lambda_0, \Lambda_1)$ , and taking values in  $\mathbb{E}$ .

We occasionally use the notation  $\langle \cdot, \cdot \rangle$  for the duality pairing between covectors and vectors, and

$$I(M; \Lambda_0; \mathbb{E}) = \bigcup_{p \in \mathbb{R}} I^p(M; \Lambda_0; \mathbb{E}), \quad I(M; \Lambda_0, \Lambda_1; \mathbb{E}) = \bigcup_{p \in \mathbb{R}} I^p(M; \Lambda_0, \Lambda_1; \mathbb{E}).$$

If  $\Lambda_0$  coincides with the conormal bundle

$$N^*K = \{(x, \xi) \in T^*M : x \in K, \langle \xi, v \rangle = 0 \text{ for all } v \in T_x K\}$$

of a submanifold  $K \subset M$  in the sense that  $\Lambda_0 = N^*K \setminus 0$ , then the distributions in  $I(M; \Lambda_0; \mathbb{E})$  are called *conormal distributions*. Although removing the zero section from  $N^*K$ , when considering it as a conic Lagrangian manifold, is somewhat awkward notationally, it is natural to consider  $N^*K$  as a submanifold of  $T^*M$ , since then the fibres  $N_x^*K \subset T_x^*M$ ,  $x \in K$ , are linear subspaces. We recall the basic properties of conormal and IPL distributions in Appendix A below.

The wave front set of a distribution  $u \in \mathcal{D}'(M)$  is denoted by  $\text{WF}(u)$  (see [11, Def. 2.5.2]). It is a subset of  $T^*M \setminus 0$ , and its projection on  $M$  is called the *singular support*  $\text{singsupp}(u)$  of  $u$ . The wave front set  $\text{WF}(u)$  is conical and closed in  $T^*M \setminus 0$ , and it is occasionally convenient to use the notation

$$\text{ccl } \mathcal{B} = \{(x, \lambda \xi) \in T^*M \setminus 0 : (x, \xi) \in \overline{\mathcal{B}}, \lambda > 0\}$$

for the conical closure of a set  $\mathcal{B} \subset T^*M \setminus 0$ . For  $u \in I(M; \Lambda_0; \mathbb{E})$  we have  $\text{WF}(u) \subset \Lambda_0$ .

If  $\mathcal{K}$  is the Schwartz kernel of a pseudodifferential operator  $\chi$  on  $M$ , then the projection of  $\text{WF}(\mathcal{K}) \subset (T^*M \setminus 0)^2$  on the first factor  $T^*M \setminus 0$  is called the *essential support* of  $\chi$ . (As  $\text{WF}(\mathcal{K})$  is contained in the conormal bundle of the diagonal  $\{(x, y) \in M^2 : x = y\}$ , the choice between the first and second factor makes no difference.) Following [11, p. 124] we write  $\text{WF}(\chi)$  for this set.

We denote by  $\Omega^{1/2}$  the half-density bundle over  $M$ . When  $\Lambda_0$  and  $\Lambda_1 \setminus \Lambda_0$  coincide with conormal bundles, and  $\mathbb{E} = E \otimes \Omega^{1/2}$ , there is a coordinate invariant way to define the principal symbol  $\sigma[u]$  of  $u \in I(M; \Lambda_0; \mathbb{E})$ , respectively  $u \in I(M; \Lambda_0, \Lambda_1; \mathbb{E})$ , as an equivalence class of sections of  $E \otimes \Omega^{1/2}$  over  $\Lambda_0$ , respectively  $\Lambda_1 \setminus \Lambda_0$ . We will not emphasize the difference between the equivalence class  $\sigma[u]$  and a representative of it, and we will also use the same notation for the half-density bundles over  $M$  and  $\Lambda_j$ ,  $j = 0, 1$ . Let us remark that there is typically no natural way to relate these bundles. For example, while it is natural to use  $|g|^{1/4}$  to trivialize  $\Omega^{1/2}$  over  $M$ , the Lorentzian metric  $g$  on  $M$  typically does not induce a natural trivialization of  $\Omega^{1/2}$  over  $\Lambda_j$ .

For IPL distributions in  $I(M; \Lambda_0, \Lambda_1; \mathbb{E})$ , there is also a refined notion of principal symbol, with components on both  $\Lambda_0$  and  $\Lambda_1$ . We will use the refined principal symbol only in Appendix A. The notation  $\sigma[\chi]$  is used also for the principal symbol of a pseudodifferential operator  $\chi$  on  $M$ . In this case,  $\sigma[\chi]$  is represented by a section of  $T^*M \setminus 0$ .

### 2.3. Microlocal analysis of the wave operator

It is convenient to rescale (8), and consider the non-linear wave operator

$$Q_0(\phi) = \frac{1}{2}(\square_A \phi + \kappa|\phi|^2 \phi),$$

where  $|\cdot| = |\cdot|_E$  is the norm with respect to the inner product  $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_E$ . In order to make use of the microlocal machinery developed in [9], we conjugate the operator  $Q_0$

with the half-density  $|g|^{1/4}$  and consider the operator

$$Q(u) := |g|^{1/4} Q_0(|g|^{-1/4}u)$$

acting on the sections of  $E \otimes \Omega^{1/2}$ . Writing  $Pu = |g|^{1/4} \square_A(|g|^{-1/4}u)/2$ , the operator  $Q$  reads

$$Q(u) = Pu + \frac{\kappa}{2} ||g|^{-1/4}u|^2 u. \tag{18}$$

For the sake of convenience, we will slightly abuse the notation, and write

$$uv = |g|^{1/4} ( (|g|^{-1/4}u) (|g|^{-1/4}v) )$$

for products of half-densities as functions. Then  $Q(u) = Pu + \kappa|u|^2u/2$ .

Writing  $\iota = \sqrt{-1}$ , the full symbol of the operator  $P$  reads

$$P(x, \xi) = g^{ij} \xi_i \xi_j / 2 + \iota^{-1} (\partial_{x^i} g^{ij} / 2 + g^{ij} A_i) \xi_j + \text{zeroth order terms,}$$

and we see that  $P$  has the following principal symbol  $\sigma[P]$  and subprincipal symbol  $\sigma_{\text{sub}}[P]$ , in the sense of Duistermaat and Hörmander [9, p. 189 (5.2.8)]:

$$\sigma[P](x, \xi) = g^{ij} \xi_i \xi_j / 2, \quad \sigma_{\text{sub}}[P](x, \xi) = \iota^{-1} g^{ij} A_i \xi_j, \quad (x, \xi) \in T^*M. \tag{19}$$

We write also

$$\sigma[P] = \langle \xi, \xi \rangle_g / 2$$

where  $\langle \xi, \xi \rangle_g$  denotes the inner product with respect to  $g$ . Let us remark that the subprincipal symbol transforming as a connection is discussed in [16] in the more general context of pseudodifferential operators on vector bundles.

We denote by  $H_P$  the Hamiltonian vector field associated to  $\sigma[P]$ , and by  $\Sigma(P)$  the characteristic set of  $P$ , that is,

$$\begin{aligned} H_P &= g^{ij} \xi_j \partial_{x^i} - \frac{1}{2} (\partial_{x^i} g^{jk}) \xi_j \xi_k \partial_{\xi^i}, \\ \Sigma(P) &= \{(x, \xi) \in T^*M \setminus 0 : \langle \xi, \xi \rangle_g = 0\}. \end{aligned} \tag{20}$$

The covectors  $\xi$  satisfying  $\langle \xi, \xi \rangle_g = 0$  are called *lightlike*. We denote by  $\Phi_s, s \in \mathbb{R}$ , the flow of  $H_P$ , and we define for a set  $\mathcal{B} \subset \Sigma(P)$  the *future flowout* of  $\mathcal{B}$  by

$$\{(y, \eta) \in \Sigma(P) : (y, \eta) = \Phi_s(x, \xi), s \in \mathbb{R}, (x, \xi) \in \mathcal{B}, y \geq x\}. \tag{21}$$

Let us recall the parametrix construction for the linear wave equation

$$\begin{cases} Pu = f & \text{in } \mathbb{R} \times M_0, \\ u|_{t < 0} = 0, \end{cases} \tag{22}$$

which originates from [9]. We will follow the purely symbolic construction from [31], the only difference being that  $u$  is vector-valued in our case. For the convenience of the reader, we give a proof of the theorem below in Appendix A.

**Theorem 3.** *Let  $f \in I^k(M; \Lambda_0; E \otimes \Omega^{1/2})$  for a conic Lagrangian  $\Lambda_0 \subset T^*M \setminus 0$ . Suppose that  $H_P$  is nowhere tangent to  $\Lambda_0$ , write  $\mathcal{B} = \Lambda_0 \cap \Sigma(P)$ , and denote by  $\Lambda_1$  the future flowout of  $\mathcal{B}$ . Then the solution  $u$  of (22) is in  $I^{k-2+1/2}(M; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2})$ . Moreover,*

$$(\mathcal{L}_{H_P} + \iota \sigma_{\text{sub}}[P])\sigma[u] = 0 \quad \text{on } \Lambda_1 \setminus \Lambda_0, \tag{23}$$

$$\sigma[u] = \mathcal{R}((\sigma[P])^{-1}\sigma[f]) \quad \text{on } \mathcal{B}, \tag{24}$$

where  $\mathcal{L}_{H_P}$  is the Lie derivative with respect to  $H_P$ ,  $\sigma[u]$  and  $\sigma[f]$  are the principal symbols of  $u$  and  $f$  on  $\Lambda_1$  and  $\Lambda_0$ , respectively, and  $\mathcal{R}$  is a map, defined by (74) in Appendix A below,

$$\mathcal{R} : S^{k-1/2+n/4}(\Lambda_0 \setminus \partial\Lambda_1; \Omega^{1/2}) \rightarrow S^{k+n/4}(\Lambda_1; \Omega^{1/2})|_{\partial\Lambda_1},$$

that acts as multiplication by a scalar on  $E$ . Here it is assumed that  $\Lambda_0$  and  $\Lambda_1 \setminus \Lambda_0$  coincide with conormal bundles.

#### 2.4. Flowout from a point in the Minkowski space

The following case will be of particular importance for us. We have also included a detailed discussion of Theorem 3 in the context of this example case in Appendix B.

**Example 1.** Let  $(M, g)$  be the 1 + 3-dimensional Minkowski space, and consider a light-like vector  $\xi^0 \in T_0\mathbb{R}^{1+3} \setminus 0$  of the form  $\xi^0 = (1, \theta_0)$  where  $\theta_0$  is in the unit sphere

$$S^2 = \{\theta \in \mathbb{R}^3 : |\theta| = 1\}.$$

Let  $c \in E_0 \setminus 0$ , that is,  $c$  is a non-zero vector in the fibre of  $E$  over the origin, and let  $\chi$  be a pseudodifferential operator such that  $\sigma[\chi] \neq 0$  near  $\text{ccl}\{(0, \xi^0)\}$ . We define

$$f \in I(M; \Lambda_0; E \otimes \Omega^{1/2}), \quad f = c|g|^{1/4}\chi\delta,$$

where  $\Lambda_0 = T_0\mathbb{R}^{1+3} \setminus 0 = N^*\{0\} \setminus 0$  and  $\delta$  is the Dirac delta distribution at the origin. The corresponding future flowout  $\Lambda_1$  satisfies  $\Lambda_1 \setminus \Lambda_0 \subset N^*K \setminus 0$  where

$$K = \{(t, t\theta) \in \mathbb{R}^{1+3} : t > 0, \theta \in S^2\} \tag{25}$$

is the future light-cone in the spacetime  $\mathbb{R}^{1+3}$  emanating from the origin. Letting  $u$  be the solution of (22), its restriction on  $\mathbb{R}^{1+3} \setminus 0$  is a conormal distribution in  $I(\mathbb{R}^{1+3} \setminus 0; N^*K \setminus 0; E \otimes \Omega^{1/2})$ . As  $\sigma[f](0, \xi^0) \neq 0$ , Theorem 3 implies that for all  $s > 0$  we have  $\sigma[u](\gamma(s), \xi^0) \neq 0$  where  $\gamma(s) = (s, s\theta_0)$ , and  $\xi^0$  is also viewed as an element of  $T_{\gamma(s)}\mathbb{R}^{1+3}$ . The smaller the essential support  $\text{WF}(\chi)$  is chosen around

$$\text{ccl}(0, \pm\xi^0) := \text{ccl}\{(0, \xi^0), (0, -\xi^0)\},$$

the smaller is  $\text{singsupp}(u) \subset \bar{K} = K \cup \{0\}$  around

$$\gamma(\mathbb{R}_+) := \{\gamma(s) : s \geq 0\}.$$

The pseudodifferential operator  $\chi$  can be chosen for example as follows. Choose functions  $\chi_1 \in C^\infty(S^2)$ ,  $\chi_2 \in C^\infty(\mathbb{R})$  and  $\chi_3 \in C^\infty(\mathbb{R}^{1+3})$  such that  $\chi_1(\theta_0) = 1$ ,  $\chi_2(1) = 1$  and  $\chi_3(0) = 1$ . Let also  $q \in \mathbb{R}$ . Then, writing  $\xi = (\xi_0, \xi')$  and  $c(\xi) = (\sqrt{2}|\xi|)^{-1}$ , with  $|\xi|$  the Euclidean norm of  $\xi$ , we define the function

$$\chi_0(x, \xi) = \chi_3(x)\chi_2(c(\xi)\xi_0)\chi_1(c(\xi)\xi')|\xi|^q. \tag{26}$$

Now  $\chi_0$  is positively homogeneous of degree  $q$ . Choose furthermore  $\chi_4 \in C_0^\infty(\mathbb{R}^{1+4})$  such that  $\chi_4 = 1$  near the origin. Then  $(1 - \chi_4(\xi))\chi_0(x, \xi)$  is smooth also near  $\xi = 0$ , and it is a symbol in the sense of [11, Def. 1.1.1]. Now we define a pseudodifferential operator by

$$\chi u = \int_{\mathbb{R}^{1+4}} e^{i\xi x} (1 - \chi_4(\xi))\chi_0(x, \xi)\hat{u}(\xi) d\xi.$$

Ignoring  $2\pi$  factors, the full symbol of  $\chi$  is simply  $(1 - \chi_4(\xi))\chi_0(x, \xi)$ , and the principal symbol  $\sigma[\chi]$  is the corresponding equivalence class modulo symbols of one degree lower order.

Choose now another  $\tilde{\chi}_4 \in C_0^\infty(\mathbb{R}^{1+4})$  such that  $\tilde{\chi}_4 = 1$  near the origin. Then

$$(1 - \chi_4(\xi))\chi_0(x, \xi) - (1 - \tilde{\chi}_4(\xi))\chi_0(x, \xi) = (\tilde{\chi}_4(\xi) - \chi_4(\xi))\chi_0(x, \xi)$$

is smooth near  $\xi = 0$  since  $\tilde{\chi}_4 - \chi_4 = 0$  there. Moreover, it is a symbol of order  $-\infty$  as both  $\chi_4$  and  $\tilde{\chi}_4$  are compactly supported. Therefore also  $(1 - \tilde{\chi}_4(\xi))\chi_0(x, \xi)$  is a representative of  $\sigma[\chi]$ . As the support of  $\chi_4$  can be chosen arbitrarily small, it makes sense to say that  $\sigma[\chi] = \chi_0$  although this does not quite hold for any representative. Also for  $q = 0$  we have, in this sense,  $\sigma[\chi] = 1$  in  $\text{ccl}\{(0, \xi^0)\}$ .

Observe that  $\text{WF}(\chi)$  becomes small around  $\text{ccl}\{(0, \xi^0)\}$  when the supports of  $\chi_j$ ,  $j = 1, 2, 3$ , are small around  $\theta_0, 1$ , and  $0$  respectively. Also  $\chi u(x) = 0$  for any  $x$  outside the support of  $\chi_3$ , and for any  $k \in \mathbb{N}$  we can choose  $q < 0$  negative enough, so that  $\chi \delta \in H^k(\mathbb{R}^{1+3})$ .

Let us check that indeed  $\Lambda_1 \setminus \Lambda_0 = N^*K \setminus 0$ . The lightlike vectors in  $T_0\mathbb{R}^{1+3} \setminus 0$  are given by  $(\lambda, \lambda\theta)$  with  $\lambda \in \mathbb{R} \setminus 0$  and  $\theta \in S^2$ . In the Minkowski space, the tangent-cotangent isomorphism corresponds to changing the sign of the first component. Thus,

$$\Lambda_1 \setminus \Lambda_0 = \{(s\lambda, s\lambda\theta; -\lambda, \lambda\theta) \in \mathbb{R}^{1+3} \times \mathbb{R}^{1+3} : \lambda \in \mathbb{R} \setminus 0, \theta \in S^2, s\lambda > 0\}.$$

We can also reparametrize

$$\Lambda_1 \setminus \Lambda_0 = \{(t, t\theta; -\lambda, \lambda\theta) \in \mathbb{R}^{1+3} \times \mathbb{R}^{1+3} : \lambda \in \mathbb{R} \setminus 0, \theta \in S^2, t > 0\}.$$

Clearly the projection of  $\Lambda_1 \setminus \Lambda_0$  on the base space  $\mathbb{R}^{1+3}$  is  $K$ .

For the convenience of the reader, we will explicitly compute the conormal bundle of  $K$ . Toward that end, we choose local coordinates  $\mathbb{R}^2 \supset B \ni a \mapsto \Theta(a) \in S^2$  on  $S^2$  and see that the tangent space of  $K$  at  $(t, \theta_0)$  is given by the range of

$$\begin{pmatrix} 1 & 0 \\ \theta_0 & t \frac{d\Theta}{da} \end{pmatrix}. \tag{27}$$

Writing  $\theta_0^\perp = \{v \in \mathbb{R}^3 : v \cdot \theta_0 = 0\}$  where  $\cdot$  denotes the Euclidean inner product, we find that  $\theta_0^\perp$  is the range of  $d\Theta/da$ . Writing  $(\xi_0, \xi') \in N_{(t, \theta_0)}^* K$ , this fibre is characterized by

$$\xi_0 \delta t + \xi' \theta_0 \delta t + \xi' \delta \theta = 0, \quad \delta t \in \mathbb{R}, \delta \theta \in \theta_0^\perp.$$

Taking first  $\delta t = 1$  and  $\delta \theta = 0$  we have  $\xi_0 = -\xi' \theta_0$ . Letting then  $\delta \theta$  vary we see that  $\xi' \perp \theta_0^\perp$ , that is,  $\xi' = \lambda \theta_0$  where  $\lambda \in \mathbb{R}$ . Here we can view  $\theta_0 \in \mathbb{R}^3$  as a covector since the tangent-cotangent isomorphism in  $\mathbb{R}^3$  is the identity. Hence

$$N_{(t, \theta_0)}^* K \setminus 0 = \{(-\lambda, \lambda \theta_0) : \lambda \in \mathbb{R} \setminus 0\},$$

and indeed  $\Lambda_1 \setminus \Lambda_0 = N^* K \setminus 0$ .

Observe that  $\Lambda_1 \setminus \Lambda_0$  is embedded in the following smooth submanifold of  $T^* M \setminus 0$ , which is the flowout to both past and future:

$$\hat{\Lambda}_1 = \{(t, t\theta; -\lambda, \lambda\theta) \in \mathbb{R}^{1+3} \times \mathbb{R}^{1+3} : \lambda \in \mathbb{R} \setminus 0, \theta \in S^2, t \in \mathbb{R}\}. \tag{28}$$

Note that while  $K$  is singular at  $t = 0$ , we see that  $\hat{\Lambda}_1$  is not by considering the derivative analogous to (27),

$$\begin{pmatrix} 1 & 0 & 0 \\ \theta_0 & t \frac{d\Theta}{da} & 0 \\ 0 & 0 & -1 \\ 0 & \lambda \frac{d\Theta}{da} & \theta_0 \end{pmatrix}.$$

This matrix is injective since  $\lambda \neq 0$ .

We will next write the transport equation (23) for the principal symbol  $\sigma[u]$  as a parallel transport equation with respect to the covariant derivative  $\nabla$ , and we begin by discussing  $\Omega^{1/2}$  over the flowout  $\Lambda_1$ .

### 2.5. Trivialization of the half-density bundle over the flowout

We want to trivialize  $\Omega^{1/2}$  in a way that preserves homogeneity properties, as possessed for example by  $\chi_0$  in (26). Let us point out that, even in the context of Example 1, there appears to be no canonical choice of a non-vanishing section of  $\Omega^{1/2}$  over the conormal bundle  $\Lambda_1 \setminus \Lambda_0 = N^* K \setminus 0$ . For example, the Sasaki metric on  $T^* \mathbb{R}^{1+3}$ , associated with the Minkowski metric, is degenerate when restricted on  $N^* K \setminus 0$ .

The submanifold  $K$  in Example 1 is of codimension 1 in  $\mathbb{R}^{1+3}$ . This holds in general in the sense that if  $\Lambda_0$  and  $\Lambda_1$  in Theorem 3 satisfy

$$\Lambda_1 \setminus \Lambda_0 = N^* K \setminus 0 \tag{29}$$

for a submanifold  $K \subset M$ , then  $K$  is of codimension 1. This can be seen as follows. Observe first that  $\Lambda_1 \subset \Sigma(P)$  simply because  $\Lambda_1$  is the future flowout from  $\Lambda_0 \cap \Sigma(P)$ . Therefore, for any  $x \in K$ , the fibre  $N_x^* K$  can contain only lightlike vectors with respect



to  $g$ . On the other hand, if  $\xi^1, \xi^2 \in T_x^*M$  are lightlike and linearly independent, then their linear span satisfies (see e.g. [34, Cor. 1.1.5])

$$\{\xi \in \text{span}(\xi^1, \xi^2) : \xi \text{ is lightlike}\} = \text{span}(\xi^1) \cup \text{span}(\xi^2). \tag{30}$$

In particular  $\text{span}(\xi^1, \xi^2)$  contains vectors that are not lightlike, and therefore at most one of  $\xi^j, j = 1, 2$ , can belong to  $N_x^*K$ . This shows that  $N_x^*K$  is of dimension 1, or equivalently,  $K$  is of codimension 1 in  $M$ .

We will trivialize  $\Omega^{1/2}$  over  $\Lambda_1$  by choosing a strictly positive half-density  $\omega$  in  $C^\infty(\Lambda_1; \Omega^{1/2})$  that is positively homogeneous of degree  $1/2$ . We begin by recalling the definition of positive homogeneity following [13, p. 13]. Let  $\lambda \in \mathbb{R} \setminus 0$  and define

$$m_\lambda : T^*M \setminus 0 \rightarrow T^*M \setminus 0, \quad m_\lambda(x, \xi) = (x, \lambda\xi).$$

Then  $m_\lambda$  restricts to a map on  $\Lambda_1$  in a natural way, and we denote the restriction still by  $m_\lambda$ . The half-density  $\omega$  is said to be *positively homogeneous* of degree  $q \in \mathbb{R}$  if  $m_\lambda^*\omega = \lambda^q\omega$  for all  $\lambda > 0$ .

If in local coordinates

$$K = \{(x^0, x') \in \mathbb{R}^{1+m} : x^0 = 0\}, \tag{31}$$

then  $N^*K = \{(x^0, x'; \xi_0, \xi') \in \mathbb{R}^{1+m} \times \mathbb{R}^{1+m} : x^0 = 0, \xi' = 0\}$ . We emphasize that since the conormal bundle of  $K$  coincides with the flowout  $\Lambda_1$  in the sense of (29), the local coordinate  $x^0$  is not the time coordinate  $t$  in (15). Here  $(x, \xi) = (x^0, x'; \xi_0, \xi')$  are the induced coordinates on  $T^*M$ . Considering the restriction

$$m_\lambda : N^*K \setminus 0 \rightarrow N^*K \setminus 0,$$

the pullback  $m_\lambda^*\omega$  is given by

$$m_\lambda^*\omega(x', \xi_0) = \left| \frac{dm_\lambda}{d(x', \xi_0)} \right|^{1/2} \omega(x', \lambda\xi_0) = \lambda^{1/2}\omega(x', \lambda\xi_0).$$

Thus  $\omega$  being positively homogeneous of degree  $1/2$  means that  $\omega(x', \lambda\xi_0) = \omega(x', \xi_0)$ ; moreover,  $\omega$  being strictly positive means that  $\omega(x', \xi_0) > 0$ . We will, in fact, choose a half-density  $\omega$  that is also symmetric in the sense that

$$\omega(x', \lambda\xi_0) = \omega(x', \xi_0), \quad \lambda \in \mathbb{R} \setminus 0, \tag{32}$$

a coordinate invariant formulation of which reads  $m_\lambda^*\omega = |\lambda|^{1/2}\omega$  for all  $\lambda \in \mathbb{R} \setminus 0$ .

In general, a strictly positive half-density  $\omega \in C^\infty(\Lambda_1; \Omega^{1/2})$  satisfying (32) can be constructed by choosing an auxiliary Riemannian metric on  $M$ , restricting the associated Sasaki metric  $h$  on  $T^*M \setminus 0$  to  $\Lambda_1$ , and taking  $\omega = |h|^{1/4}$ . In particular, when  $(M, g)$  is the Minkowski space, it feels natural to choose the Euclidean metric on  $M$ .

2.6. Parallel transport equation for the principal symbol

Let us fix a strictly positive half-density  $\omega$  over  $\overline{N^*K}$  satisfying (32). Here the closure is taken in  $T^*M$ , and we remark that, in view of (29),  $\Lambda_1 \subset \overline{N^*K}$ .

We write  $\sigma[u] = a\omega$  where  $a$  is a section in  $C^\infty(N^*K \setminus 0; E)$ . A computation in coordinates shows that  $\mathcal{L}_{H_P}(a\omega) = (H_P a)\omega + a\mathcal{L}_{H_P}\omega$ . Introducing the notation  $\operatorname{div}_\omega H_P = \omega^{-1}\mathcal{L}_{H_P}\omega$ , we write

$$\omega^{-1}\mathcal{L}_{H_P}(a\omega) = H_P a + a \operatorname{div}_\omega H_P.$$

We want to further rewrite this as a conjugate differentiation along bicharacteristics, that is, along the flow curves of  $H_P$ . Recall that  $\Phi_s, s \in \mathbb{R}$ , denotes the flow of  $H_P$ . Writing  $\beta(s) = \Phi_s(x_0, \xi^0)$  for the bicharacteristic through  $(x_0, \xi^0) \in N^*K \setminus 0$ , we have

$$(H_P a) \circ \beta(s) = \partial_r a(\Phi_r(\beta(s)))|_{r=0} = \partial_s(a \circ \beta)(s).$$

We further write  $\alpha = a \circ \beta$ , and define

$$\rho(s) = \int_0^s \rho'(s') ds', \quad \rho' = \operatorname{div}_\omega H_P \circ \beta. \tag{33}$$

Then

$$\omega^{-1}\mathcal{L}_{H_P}(a\omega) \circ \beta = \partial_s \alpha + \rho' \alpha = e^{-\rho} \partial_s(e^\rho \alpha).$$

We denote by  $\gamma$  the projection of the bicharacteristic  $\beta$  to the base manifold  $M$ , and by  $\dot{\gamma}^*$  the tangent vector of  $\gamma$  as a covector, that is,  $\dot{\gamma}_i^* = g_{ij} \dot{\gamma}^j$ . It follows from (20) that  $\gamma$  is a geodesic of  $(M, g)$ , and that  $\beta(s) = (\gamma(s), \dot{\gamma}^*(s))$ . As  $\beta(s) \in \Sigma(P)$ , the geodesic  $\gamma$  is lightlike. Moreover, using (19), we obtain

$$\iota(\sigma_{\text{sub}}[P] \circ \beta)(s) = g^{ij} A_i \dot{\gamma}_j^*(s) = A_i \dot{\gamma}^i(s) = \langle A, \dot{\gamma}(s) \rangle,$$

where  $\langle \cdot, \cdot \rangle$  is again the duality pairing between covectors and vectors.

The covariant derivative on the bundle  $E$  along the geodesic  $\gamma(s)$  is given by

$$\nabla_{\dot{\gamma}} = \partial_s + \langle A, \dot{\gamma}(s) \rangle.$$

Therefore (23) along  $\beta$  is equivalent to  $e^{-\rho} \nabla_{\dot{\gamma}}(e^\rho \alpha) = 0$ . Define the following symbol along  $\beta$  for conormal distributions in  $I(M; N^*K \setminus 0; E \otimes \Omega^{1/2})$ :

$$\sigma_{\omega, \beta}[\cdot] = e^\rho (\omega^{-1} \sigma[\cdot]) \circ \beta.$$

Then we can rewrite (23) along  $\beta$  as

$$\nabla_{\dot{\gamma}} \sigma_{\omega, \beta}[u] = 0. \tag{34}$$

This is the parallel transport equation along  $\gamma$  with respect to the connection  $A$ .

If  $x = \gamma(s_0)$  and  $y = \gamma(s_1)$  for some  $s_0, s_1 \in \mathbb{R}$ , then we write

$$\mathbf{P}_{y \leftarrow x}^A : E_x \rightarrow E_y$$

for the parallel transport map from  $x$  to  $y$  along  $\gamma$ . That is,  $\mathbf{P}_{y \leftarrow x}^A v = w(s_1)$  where  $w$  is the solution of

$$\begin{cases} \dot{w} + \langle A, \dot{\gamma} \rangle w = 0 & \text{on } (s_0, s_1), \\ w(s_0) = v. \end{cases} \tag{35}$$

In the introduction we used the notation  $\mathbf{P}_\gamma^A$ , but in what follows it will be more convenient to emphasize the end points  $x$  and  $y$  rather than  $\gamma$ . In general, the map  $\mathbf{P}_{y \leftarrow x}^A$  depends on the geodesic  $\gamma$  joining  $x$  and  $y$ , but not on the parametrization of  $\gamma$ ; however, we are mainly interested in the Minkowski case, and in this case  $\mathbf{P}_{y \leftarrow x}^A$  depends only on the points  $x$  and  $y$ . To summarize, writing  $\xi = \dot{\gamma}^*(s_0)$  and  $\eta = \dot{\gamma}^*(s_1)$ , it follows from (34) that

$$e^{\rho(s_1)}(\omega^{-1}\sigma[u])(y, \eta) = e^{\rho(s_0)}\mathbf{P}_{y \leftarrow x}^A((\omega^{-1}\sigma[u])(x, \xi)). \tag{36}$$

2.7. Positively homogeneous symbols

We will now consider how (36) changes under rescaling of  $\xi \in N_x^*K \setminus 0$ , assuming that  $\sigma[u]$  is positively homogeneous of degree  $q + 1/2 \in \mathbb{R}$  at  $(x, \xi)$ , that is,

$$(\omega^{-1}\sigma[u])(x, \lambda\xi) = \lambda^q (\omega^{-1}\sigma[u])(x, \xi), \quad \lambda > 0. \tag{37}$$

**Proposition 1.** *Let  $\Lambda_j, j = 0, 1$ , and  $u$  be as in Theorem 3, and suppose that (29) holds. Let  $(x, \xi) \in N^*K \setminus 0$  and  $(y, \eta) = \Phi_{s_1}(x, \xi)$  for some  $s_1 \in \mathbb{R}$ . Suppose that (37) holds at  $(x, \pm\xi)$  and that  $(y, \eta) \in N^*K \setminus 0$ . Then*

$$e^{\rho(s_1)}(\omega^{-1}\sigma[u])(y, \pm\lambda\eta) = \lambda^q \mathbf{P}_{y \leftarrow x}^A((\omega^{-1}\sigma[u])(x, \pm\xi)), \quad \lambda > 0. \tag{38}$$

Recall that  $\mathcal{B} = \Lambda_0 \cap \Sigma(P) = \Lambda_0 \cap \Lambda_1$ . As the symbol  $\sigma[u]$  on  $\Lambda_1 \setminus \Lambda_0$  is smooth up to  $\mathcal{B}$ , equation (38) also holds when  $(x, \pm\xi) \in \mathcal{B}$ . Then  $(y, \eta) \in N^*K \setminus 0$  implies  $x < y$ , that is, the causal relation  $x \leq y$  holds and  $x \neq y$ . Writing again  $\gamma(s)$  for the projection of  $\beta(s) = \Phi_s(x, \xi)$  to  $M$ , we have

$$\gamma(0) = x < y = \gamma(s_1).$$

Hence, if  $\xi$  is future pointing then  $s_1 > 0$ , and if  $\xi$  is past pointing then  $s_1 < 0$ . We emphasize that also past pointing singularities are propagated forward in time by the wave equation (22).

*Proof of Proposition 1.* For  $\lambda \in \mathbb{R} \setminus 0$ , the bicharacteristic through  $(x, \lambda\xi)$  is given by

$$\beta_\lambda(s) = (\gamma(\lambda s), \lambda \dot{\gamma}^*(\lambda s)),$$

where  $\gamma$  still denotes the projection of  $\beta = \beta_1$  to  $M$ . Analogously to (33), we write  $\rho'_\lambda = \text{div}_\omega H_P \circ \beta_\lambda$  and  $\rho_\lambda(s) = \int_0^s \rho'_\lambda(s') ds'$ , with the usual convention that  $\int_0^s \rho'_\lambda(s') ds' = -\int_s^0 \rho'_\lambda(s') ds'$  for  $s < 0$ . By applying (36) to  $\beta_\lambda$ , we obtain

$$\begin{aligned} e^{\rho_\lambda(s_1/\lambda)}(\omega^{-1}\sigma[u])(y, \lambda\eta) &= \mathbf{P}_{y \leftarrow x}^A((\omega^{-1}\sigma[u])(x, \lambda\xi)) \\ &= |\lambda|^q \mathbf{P}_{y \leftarrow x}^A((\omega^{-1}\sigma[u])(x, \pm\xi)), \end{aligned}$$

where the sign is the one in  $\lambda = \pm|\lambda|$ .

It remains to show that  $e^{\rho_\lambda(s_1/\lambda)} = e^{\rho(s_1)}$ . We denote the restriction of the flow  $\Phi_s$  on  $N^*K \setminus 0$  still by  $\Phi_s$ . In coordinates satisfying (31), the Lie derivative  $\mathcal{L}_{H_P}\omega$  is

$$\begin{aligned} \mathcal{L}_{H_P}\omega(x', \xi_0) &= \partial_s \Phi_s^* \omega(x', \xi_0)|_{s=0} = \partial_s \left( \left| \frac{d\Phi_s}{d(x', \xi_0)} \right|^{1/2} \omega(\Phi_s(x', \xi_0)) \right) \Big|_{s=0} \\ &= \partial_s \left( \left| \frac{d\Phi_s}{d(x', \xi_0)} \right|^{1/2} \right) \Big|_{s=0} \omega(x', \xi_0) + \partial_s (\omega(\Phi_s(x', \xi_0))) \Big|_{s=0}, \end{aligned}$$

where we have used the fact that  $\Phi_0 = \text{id}$ , and therefore  $|d\Phi_0/d(x', \xi_0)|^{1/2} = 1$ . We write

$$H_P(x', \xi_0) = H^0(x', \xi_0)\partial_{\xi_0} + \sum_{i=1}^n H^i(x', \xi_0)\partial_{x'^i}$$

for the Hamiltonian vector field  $H_P$  on  $N^*K \setminus 0$ . Then (see e.g. [29, p. 418])

$$\partial_s \left( \left| \frac{d\Phi_s}{d(x', \xi_0)} \right|^{1/2} \right) \Big|_{s=0} = \frac{1}{2} \partial_s \left( \left| \frac{d\Phi_s}{d(x', \xi_0)} \right| \right) \Big|_{s=0} = \frac{1}{2} (\partial_{\xi_0} H^0 + \sum_{i=1}^n \partial_{x'^i} H^i).$$

Equation (20), together with the fact that  $\xi' = 0$  on  $N^*K$ , implies that

$$\partial_{\xi_0} H^0 + \sum_{j=1}^n \partial_{x'^j} H^j = -(\partial_{x^0} g^{00})\xi_0 + \sum_{i=1}^n (\partial_{x^i} g^{0i})\xi_0.$$

In particular,

$$(\partial_{\xi_0} H^0)(x', \lambda\xi_0) + \sum_{j=1}^n (\partial_{x'^j} H^j)(x', \lambda\xi_0) = \lambda(\partial_{\xi_0} H^0)(x', \xi_0) + \lambda \sum_{j=1}^n (\partial_{x'^j} H^j)(x', \xi_0).$$

Moreover, as  $\omega$  satisfies (32),

$$\begin{aligned} \partial_s \omega(\Phi_s(x', \lambda\xi_0))|_{s=0} &= \partial_s \omega(\gamma(\lambda s), \lambda\dot{\gamma}^*(\lambda s))|_{s=0} = \partial_s \omega(\gamma(\lambda s), \dot{\gamma}^*(\lambda s))|_{s=0} \\ &= \lambda \partial_s \omega(\gamma(s), \dot{\gamma}^*(s))|_{s=0} = \lambda \partial_s \omega(\Phi_s(x', \xi_0))|_{s=0}. \end{aligned}$$

The above six equations imply that  $\mathcal{L}_{H_P}\omega(x', \lambda\xi_0) = \lambda \mathcal{L}_{H_P}\omega(x', \xi_0)$ . This again implies

$$\text{div}_\omega H_P(x', \lambda\xi_0) = \omega^{-1}(x', \lambda\xi_0) \mathcal{L}_{H_P}\omega(x', \lambda\xi_0) = \lambda \text{div}_\omega H_P(x', \xi_0),$$

and therefore the change of variables  $s' = \lambda s$  gives

$$\begin{aligned} \rho_\lambda(s_1/\lambda) &= \int_0^{s_1/\lambda} \operatorname{div}_\omega H_P(\gamma(\lambda s), \lambda \dot{\gamma}^*(\lambda s)) ds = \int_0^{s_1} \operatorname{div}_\omega H_P(\gamma(s'), \dot{\gamma}^*(s')) ds' \\ &= \rho(s_1). \end{aligned} \quad \blacksquare$$

### 3. Microlocal analysis of the interaction of three waves

The core idea of the proof of Theorem 1 is to choose the source  $f$  in (8) as the weighted superposition of three singular sources,

$$f(\epsilon) = \sum_{j=1}^3 \epsilon_j f_j, \quad \epsilon = (\epsilon_1, \epsilon_2, \epsilon_3), \quad \epsilon_j \in \mathbb{R}, \tag{39}$$

where  $f_j, j = 1, 2, 3$ , are conormal distributions supported in  $\mathfrak{U}$  satisfying

$$\operatorname{supp}(f_j) \cap \mathcal{J}^+(\operatorname{supp}(f_k)) = \emptyset \quad \text{for } j \neq k. \tag{40}$$

Recall that  $\mathfrak{U}$  is the set where the measurements are gathered; see the definition (7) of the source-to-solution map  $L_A$ . Each  $f_j$  will be similar to the source  $f$  in Example 1. We denote by  $\phi = \phi(\epsilon)$  the solution of (8) with source  $f = f(\epsilon)$ , and write

$$\phi_{\text{in}} = \partial_{\epsilon_1} \phi|_{\epsilon=0}, \quad \phi_{\text{out}} = \partial_{\epsilon_1} \partial_{\epsilon_2} \partial_{\epsilon_3} \phi|_{\epsilon=0}. \tag{41}$$

We will choose  $f_1$  so that  $\phi_{\text{in}}$  is singular along a lightlike geodesic  $\gamma_{\text{in}}$  intersecting  $\mathfrak{U}$ . Then we choose  $f_2$  and  $f_3$  to be perturbations of  $f_1$  such that  $\phi_{\text{out}}$  is singular along another lightlike geodesic  $\gamma_{\text{out}}$  that intersects both  $\gamma_{\text{in}}$  and  $\mathfrak{U}$ . Let us fix the parametrization of  $\gamma_{\text{in}}$  and  $\gamma_{\text{out}}$  so that

$$\gamma_{\text{in}}(0) = x \in \mathfrak{U}, \quad \gamma_{\text{in}}(s_{\text{in}}) = \gamma_{\text{out}}(0) = y, \quad \gamma_{\text{out}}(s_{\text{out}}) = z \in \mathfrak{U} \tag{42}$$

for some  $s_{\text{in}}, s_{\text{out}} \in \mathbb{R}$ . Taking into account the fact that the wave equation (8) propagates singularities forward in time, we consider only the case  $x < y < z$  (see Figure 2, left).

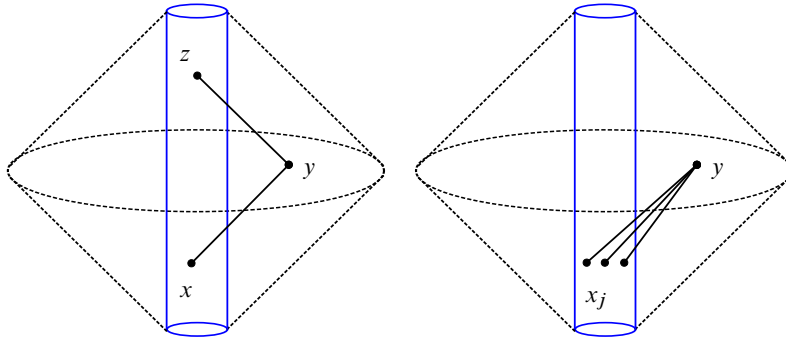
We will use the following shorthand notations related to (42):

$$\begin{aligned} \mathbb{L} &= \{(x, y) \in M^2 : \text{there is a lightlike geodesic joining } x \text{ and } y\}, \\ \mathbb{S}^+(\mathfrak{U}) &= \{(x, y, z) \in M^3 : (x, y), (y, z) \in \mathbb{L}, x < y < z, x, z \in \mathfrak{U}, y \notin \mathfrak{U}\}. \end{aligned} \tag{43}$$

Assume from now on that  $(M, g)$  is the 1 + 3-dimensional Minkowski space, and denote the geodesic joining a pair  $(x, y) \in \mathbb{L}$  by  $\gamma_{y \leftarrow x}$ . What follows is independent of the parametrization of  $\gamma_{y \leftarrow x}$ , but let us fix it by requiring that  $\gamma_{y \leftarrow x}(0) = x$  and  $\dot{\gamma}_{y \leftarrow x}(0) = (1, \theta)$  where  $\theta \in S^2$ . We will also use the notation

$$\mathbf{S}_{z \leftarrow y \leftarrow x}^A = \mathbf{P}_{z \leftarrow y}^A \mathbf{P}_{y \leftarrow x}^A, \quad (x, y, z) \in \mathbb{S}^+(\mathfrak{U}), \tag{44}$$

for the broken non-abelian light ray transform of the connection  $A$ .



**Fig. 2.** The set  $\mathcal{U}$  (see (5)) is the blue cylinder, and the set  $\mathbb{D}$  (see (9)) is the diamond-like region drawn with dashed curves. *Left:* Points  $x, z \in \mathcal{U}$  and  $y \in \mathbb{D}$  as in (42). The geodesic segment from  $x$  to  $y$  is  $\gamma_{\text{in}}$  and the segment from  $y$  to  $z$  is  $\gamma_{\text{out}}$ . We also write  $\gamma_{\text{in}} = \gamma_{y \leftarrow x}$  and  $\gamma_{\text{out}} = \gamma_{z \leftarrow y}$ . *Right:* Points  $x_j, j = 1, 2, 3$  (see (52)), together with the geodesic segments  $\gamma_{y \leftarrow x_j}$ . Here  $x_1 = x$ .

Now we can summarize the above discussion as follows: for each  $(x, y, z) \in \mathbb{S}^+(\mathcal{U})$  we want to choose  $f_j, j = 1, 2, 3$ , so that  $\phi_{\text{in}}$  is singular along  $\gamma_{y \leftarrow x}$  and that  $\phi_{\text{out}}$  is singular along  $\gamma_{z \leftarrow y}$ . Moreover, we will choose  $f_j, j = 1, 2, 3$ , in such a way that, in a suitable microlocal sense,  $\phi_{\text{out}}$  is a conormal distribution. Then we will show using Proposition 1 that  $\mathbb{S}_{z \leftarrow y \leftarrow x}^A$  is determined by the principal symbol of  $\phi_{\text{out}}|_{\mathcal{U}}$  when  $f_j, j = 1, 2, 3$ , are varied. As each  $f_j$  is supported in  $\mathcal{U}$ , this will show that the source-to-solution map  $L_A$  determines  $\mathbb{S}_{z \leftarrow y \leftarrow x}^A$  for all  $(x, y, z) \in \mathbb{S}^+(\mathcal{U})$ .

### 3.1. The linear span of three lightlike vectors

We will begin with a lemma in linear algebra that will be a key component when proving that singularities propagating along any  $\gamma_{z \leftarrow y}$ , with the pair  $(y, z)$  in

$$\mathbb{S}^{\text{out}}(\mathcal{U}) = \{(y, z) : (x, y, z) \in \mathbb{S}^+(\mathcal{U}) \text{ for some } x \in \mathcal{U}\},$$

can be generated by choosing suitable  $f_j, j = 1, 2, 3$ , supported in  $\mathcal{U}$ . We believe that this lemma can also be used to simplify the proofs of previous results, such as those in [20, 24, 27]. These results are based on using a superposition of four singular sources, which leads to more complicated computations.

To highlight the difference, we recall that all the previous results consider non-linear wave equations on 1 + 3-dimensional Lorentzian manifolds, and use the fact that if  $N^*K_j \setminus 0, j = 1, 2, 3, 4$ , are flowouts analogous to  $N^*K \setminus 0$  in Example 1, then generically  $\bigcap_{j=1}^4 K_j$  is discrete. It follows that  $N_y^*(\bigcap_{j=1}^4 K_j) = T_y^*M$  for a point  $y \in \bigcap_{j=1}^4 K_j$ . Then, using analogous notation, it is shown that  $\phi_{\text{out}} = \partial_{\epsilon_1} \partial_{\epsilon_2} \partial_{\epsilon_3} \partial_{\epsilon_4} \phi|_{\epsilon=0}$  can be made singular along any  $\gamma_{z \leftarrow y}$  with  $(y, z) \in \mathbb{S}^{\text{out}}(\mathcal{U})$ . As  $\dot{\gamma}_{z \leftarrow y}(0) \in N_y^*(\bigcap_{j=1}^4 K_j)$  trivially, there is no geometric obstruction to  $\phi_{\text{out}}$  being singular on  $\dot{\gamma}_{z \leftarrow y}$  in view of the product calculus for conormal distributions (see Lemma 2 below).

On the other hand, we will use only three flowouts, and the intersection  $\bigcap_{j=1}^3 K_j$  will be of dimension 1. Lemma 1 implies, however, that for any fixed  $y \in K_1$  and any fixed

lightlike  $\eta \in T_y^*M$  we can guarantee that  $(y, \eta) \in N^*(\bigcap_{j=1}^3 K_j)$  by choosing  $K_2$  and  $K_3$  carefully.

**Remark 2.** Consider the solution  $\phi(\epsilon)$  of (13) with vanishing initial conditions and the source  $f(\epsilon) = \epsilon_1 f_1 + \epsilon_2 f_2$ , with  $f_j$  a conormal distribution supported in  $\mathcal{U}$  satisfying (40). Regardless of the degree of non-linearity  $N \geq 2$ , it is not possible to make  $\partial_{\epsilon_1} \partial_{\epsilon_2} \phi|_{\epsilon=0}$  singular along arbitrary  $\gamma_{z \leftarrow y}$  with  $(y, z) \in \mathbb{S}^{\text{out}}(\mathcal{U})$ . Indeed, using (30) it can be shown that

$$\text{singsupp}(\partial_{\epsilon_1} \partial_{\epsilon_2} \phi|_{\epsilon=0}) \subset \bigcup_{j=1}^2 \text{singsupp}(\partial_{\epsilon_j} \phi|_{\epsilon=0}).$$

This explains the lower bound  $J \geq 3$  in (14).

The following lemma is formulated for the Minkowski space but, being of pointwise nature, it generalizes to an arbitrary Lorentzian manifold.

**Lemma 1.** *Let  $y$  be a point in  $\mathbb{R}^{1+3}$ , and let  $\xi_1, \eta \in T_y \mathbb{R}^{1+3} \setminus 0$  be lightlike. In any neighbourhood of  $\xi_1$  in  $T_y \mathbb{R}^{1+3}$ , there exist two lightlike vectors  $\xi_2, \xi_3$  such that  $\eta$  is in  $\text{span}(\xi_1, \xi_2, \xi_3)$ .*

*Proof.* The statement is invariant with respect to non-vanishing rescaling of  $\xi_1$  and  $\eta$ , and we assume without loss of generality that  $\xi_1 = (1, \xi'_1)$  and  $\eta = (1, \eta')$ . Let  $v \in S^2$  be such that  $\xi'_1, \eta' \perp v$ . After a rotation in  $\mathbb{R}^3$ , we may assume that  $v = (0, 0, 1)$ . Then  $\xi'_1 = (\theta_1, 0)$  and  $\eta' = (\theta_0, 0)$  with  $\theta_1, \theta_0 \in S^1$ . After a rotation in  $\mathbb{R}^2$ , we may assume that  $\theta_1 = (1, 0)$ . We write

$$a(r) = \sqrt{1 - r^2} \quad \text{for } r \in [-1, 1].$$

Then there is  $r_0 \in [-1, 1]$ , and a choice of sign, such that  $\theta_0 = (\pm a(r_0), r_0)$ . To summarize, we may assume without loss of generality that

$$\xi_1 = (1, 1, 0, 0), \quad \eta = (1, \pm a(r_0), r_0, 0). \tag{45}$$

It is enough to show that in any neighbourhood  $U_0 \subset S^1$  of  $\theta_1$  there are  $\theta_2, \theta_3 \in U_0$  such that  $(1, \theta_j), j = 1, 2, 3, \text{span } \mathbb{R}^3$ . We set  $\theta_2 = (a(r), r)$  and  $\theta_3 = (a(r), -r)$  where  $r > 0$  is chosen to be small enough so that  $\theta_j \in U_0, j = 2, 3$ . Now  $(1, \theta_j), j = 1, 2, 3$ , are linearly independent since

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & a(r) & a(r) \\ 0 & r & -r \end{vmatrix} = 2r(1 - a(r)) \neq 0.$$

To summarize, we may choose

$$\xi_2 = (1, a(r), r, 0), \quad \xi_3 = (1, a(r), -r, 0). \tag{46}$$

■

We will also need an explicit version of the above lemma. As in the proof, we assume without loss of generality that (45) holds, and choose  $\xi_2$  and  $\xi_3$  according to (46). Then

$$r^2\eta = (-2b + \mathcal{O}(r))\xi_1 + (b + \mathcal{O}(r))\xi_2 + (b + \mathcal{O}(r))\xi_3, \tag{47}$$

where  $b = 1 \mp a(r_0) = 1 \mp \sqrt{1 - r_0^2}$ .

**Remark 3.** In the context of (42), we will take  $\xi_1 = \dot{\gamma}_{\text{in}}(s_{\text{in}})$  and  $\eta = \dot{\gamma}_{\text{out}}(0)$ . Assuming that  $y \notin \mathfrak{U}$ , the condition  $\gamma_{\text{out}}(s_{\text{out}}) \in \mathfrak{U}$  implies that the second component of  $\eta$  in (45), that is,  $\pm a(r_0)$ , must have negative sign. Then  $b \geq 1$ , and in particular  $\eta$  is not in  $\text{span}(\xi_j, \xi_k)$  for any  $j, k \in \{1, 2, 3\}$ . This fact will be important in what follows, and the set  $\mathfrak{U}$  was chosen to be of the particular form (5) to avoid technicalities related, for example, to geodesics that return to  $\mathfrak{U}$  after exiting it.

### 3.2. Three-fold linearization of cubic connection wave equations

Let us consider a source  $f(\epsilon)$  of the form (39), and the half-density  $u(\epsilon) = |g|^{1/4}\phi(\epsilon)$  where  $\phi(\epsilon)$  denotes again the solution of (8) with source  $f = f(\epsilon)$ . Recalling the definition of  $Q$  (see (18)), we observe that  $u = u(\epsilon)$  satisfies the wave equation

$$Q(u) = |g|^{1/4}f. \tag{48}$$

To simplify the notation, we assume that  $\kappa = -1$  in (4). The general case is analogous. We introduce

$$u_{\mathbf{j}_1, \dots, \mathbf{j}_l} = \partial_{\epsilon_{\mathbf{j}_1}} \cdots \partial_{\epsilon_{\mathbf{j}_l}} u, \quad v_{\mathbf{j}_1, \dots, \mathbf{j}_l} = u_{\mathbf{j}_1, \dots, \mathbf{j}_l}|_{\epsilon=0},$$

and write  $\text{Re}$  for the real part of a complex number. With these notations, we differentiate (48), and obtain

$$Pu_{\mathbf{j}} - \text{Re}(\langle u_{\mathbf{j}}, u \rangle)u - \frac{1}{2}|u|^2u_{\mathbf{j}} = |g|^{1/4}f_{\mathbf{j}}.$$

Together with the fact that  $u|_{\epsilon=0} = 0$ , this implies  $Pv_{\mathbf{j}} = |g|^{1/4}f_{\mathbf{j}}$ . Differentiating the equation (48) twice yields

$$Pu_{\mathbf{j}\mathbf{k}} - \text{Re}(\langle u_{\mathbf{j}\mathbf{k}}, u \rangle + \langle u_{\mathbf{j}}, u_{\mathbf{k}} \rangle)u - \text{Re}(\langle u_{\mathbf{j}}, u \rangle)u_{\mathbf{k}} - \text{Re}(\langle u_{\mathbf{k}}, u \rangle)u_{\mathbf{j}} - \frac{1}{2}|u|^2u_{\mathbf{j}\mathbf{k}} = 0, \tag{49}$$

and  $v_{\mathbf{j}\mathbf{k}} = 0$ . Finally, one more differentiation gives the desired linear equation, which we call the *three-fold linearization*,

$$Pv_{123} = \text{Re}(\langle v_1, v_2 \rangle)v_3 + \text{Re}(\langle v_1, v_3 \rangle)v_2 + \text{Re}(\langle v_2, v_3 \rangle)v_1. \tag{50}$$

The right-hand side of (50) is the sum of products of  $v_1, v_2, v_3$ , and we call it the *three-fold interaction* of these three solutions to the linear wave equation (22).

**Remark 4.** Consider the solution  $\phi(\epsilon)$  of (13) with vanishing initial conditions and the source  $f(\epsilon) = \sum_{j=1}^N \epsilon_j f_j$ , with  $f_j$  a distribution supported in  $\mathfrak{U}$ . Analogously to (49), we see that  $\partial_{\epsilon}^{\alpha} \phi|_{\epsilon=0} = 0$  for any multi-index  $\alpha \in \mathbb{N}^{1+3}$  satisfying  $|\alpha| < N$  where  $N$  is the degree of non-linearity in equation (13). This explains the lower bound  $J \geq N$  in (14).



3.3. Wave interactions as products of conormal distributions

Consider a triple  $(x, y, z) \in \mathbb{S}^+(\mathcal{U})$  and let us construct  $f_j, j = 1, 2, 3$ , such that  $\phi_{\text{out}}$ , defined by (41), is singular on  $\gamma_{z \leftarrow y}$ . We write

$$\eta = \dot{\gamma}_{z \leftarrow y}^*(0), \quad \xi_1 = \dot{\gamma}_{y \leftarrow x}^*(s_{\text{in}}), \tag{51}$$

where  $s_{\text{in}} > 0$  satisfies  $\gamma_{y \leftarrow x}(s_{\text{in}}) = y$ .

The geodesic  $\gamma$  with initial condition  $(x, \xi)$  is denoted by  $\gamma(\cdot; x, \xi)$ . Let  $\mathcal{V} \subset T_y^*M$  be a small enough neighbourhood of  $\xi_1$  so that for all  $\xi \in \mathcal{V}$  we have  $\gamma(-s_{\text{in}}; y, \xi) \in \mathcal{U}$  and  $\xi$  is future pointing. Let  $\xi_2, \xi_3 \in \mathcal{V}$  be two lightlike vectors as in Lemma 1 and write

$$x_j = \gamma(-s_{\text{in}}; y, \xi_j), \quad j = 1, 2, 3. \tag{52}$$

With a slight abuse of notation, we also write  $\xi_j = \dot{\gamma}(-s_{\text{in}}; y, \xi_j)$  (see Figure 2, right) for the geometric setup.

Analogously to Example 1, we let  $c_j \in E_{x_j} \setminus 0$ , and let  $\chi_j$  be a pseudodifferential operator with the following properties:

- ( $\chi 1$ )  $\sigma[\chi_j]$  is positively homogeneous of degree  $q$ , symmetric with respect to  $0 \in T^*M$ , and real-valued,
- ( $\chi 2$ )  $\sigma[\chi_j] \neq 0$  near  $(x_j, \xi_j)$ ,
- ( $\chi 3$ )  $\text{WF}(\chi_j)$  is contained in a small neighbourhood of  $\text{ccl}(x_j, \pm \xi_j)$ .

The degree  $q$  of  $\chi_j$  is chosen to be small enough so that  $f_j \in C^4(M)$  where

$$f_j \in I(M; N^*\{x_j\} \setminus 0; E), \quad f_j = c_j \chi_j \delta_{x_j}. \tag{53}$$

Here  $\delta_{x_j}$  is the Dirac delta distribution at  $x_j$ . Moreover, we choose  $\chi_j$  so that  $\text{supp}(\chi_j \delta_j) \subset \mathcal{U}$  and the support condition (40) is satisfied. Recalling that  $\mathcal{C} \subset C_0^4(\mathcal{U}; E)$  is the domain of the source-to-solution map  $L_A$  (see (7)), we find that the linear combination  $\epsilon_1 f_1 + \epsilon_2 f_2 + \epsilon_3 f_3$  is in  $\mathcal{C}$  for small enough  $\epsilon_j \geq 0$ .

Recall that  $v_j$  denotes the solution of

$$P v_j = |g|^{1/4} f_j, \quad v_j|_{t < 0} = 0. \tag{54}$$

Moreover, we denote the restriction of  $v_j$  on  $\mathbb{R}^{1+3} \setminus \{x_j\}$  still by  $v_j$ , and write  $K_j = x_j + K$  where  $K$  is as in (25). That is, writing  $x_j = (t_j, x'_j)$ ,

$$K_j = \{(t_j + s, x'_j + s\theta) \in \mathbb{R}^{1+3} : s > 0, \theta \in S^2\}. \tag{55}$$

Then  $v_j \in I(\mathbb{R}^{1+3} \setminus x_j; N^*K_j \setminus 0; E \otimes \Omega^{1/2})$ . As in Example 1, property ( $\chi 3$ ) implies that  $\text{singsupp}(v_j)$  is contained in a small neighbourhood  $\mathcal{S}_j \subset \overline{K_j}$  of  $\gamma_{y \leftarrow x_j}(\mathbb{R}_+)$ .

Clearly  $y \in \bigcap_{j=1}^3 \gamma_{y \leftarrow x_j}(\mathbb{R}_+) \subset \bigcap_{j=1}^3 K_j$ . It follows from Remark 3 that the covectors  $\xi_j, j = 1, 2, 3$ , are linearly independent. This again implies that

$$N_y^*(K_1 \cap K_2 \cap K_3) = \text{span}(\xi_1, \xi_2, \xi_3).$$

When  $\text{WF}(\chi_{\mathbf{j}})$ ,  $\mathbf{j} = 1, 2, 3$ , are small enough, this also guarantees for distinct  $\mathbf{j}, \mathbf{k}, \mathbf{l}$  that  $K_{\mathbf{j}}$  and  $K_{\mathbf{k}}$ , as well as  $K_{\mathbf{j}}$  and  $K_{\mathbf{k}} \cap K_{\mathbf{l}}$ , are transversal in  $\mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$ .

Let us now consider products of the form  $\text{Re}(\langle v_{\mathbf{j}}, v_{\mathbf{k}} \rangle) v_{\mathbf{l}}$ , with distinct  $\mathbf{j}, \mathbf{k}, \mathbf{l}$ , that appear on the right-hand side of (50). As  $s_{\text{in}} > 0$ , we have  $y \neq x_{\mathbf{j}}$  for each  $\mathbf{j} = 1, 2, 3$ , and  $v_{\mathbf{j}}$  are conormal distributions near  $y$ . Thus their products can be analysed using the product calculus for conormal distributions. We recall this calculus in the next lemma, which is a variant of [10, Lemma 1.1]. With obvious modifications it also holds when the conormal distributions  $u_{\mathbf{j}}$ ,  $\mathbf{j} = 1, 2$ , take values on  $E \otimes \Omega^{1/2}$  and the product is defined in terms of  $\langle \cdot, \cdot \rangle_E$ , and also when  $u_1$  takes values on  $\Omega^{1/2}$  and  $u_2$  on  $E \otimes \Omega^{1/2}$  and the product is defined in terms of the scalar-vector product on  $E$ .

**Lemma 2.** *Let  $K_{\mathbf{j}} \subset M$ ,  $\mathbf{j} = 1, 2$ , be transversal submanifolds, and let  $u_{\mathbf{j}}$  be a conormal distribution in  $I(M; N^*K_{\mathbf{j}} \setminus 0; \Omega^{1/2})$ . For a fixed nowhere vanishing half-density  $\mu \in C^\infty(M; \Omega^{1/2})$ , define the product of  $u_1$  and  $u_2$  by*

$$u_1 u_2 = \mu((\mu^{-1}u_1)(\mu^{-1}u_2)).$$

Let  $\chi$  be a pseudodifferential operator with  $\text{WF}(\chi)$  disjoint from both  $N^*K_{\mathbf{j}}$ ,  $\mathbf{j} = 1, 2$ . Then

$$\chi(u_1 u_2) \in I(M; N^*(K_1 \cap K_2) \setminus 0; \Omega^{1/2})$$

with principal symbol, ignoring the  $2\pi$  and  $i$  factors,

$$\sigma[\chi(u_1 u_2)](x, \xi) = \mu^{-1}(x)\sigma[\chi](x, \xi)\sigma[u_1](x, \xi^{(1)})\sigma[u_2](x, \xi^{(2)}),$$

where  $\xi = \xi^{(1)} + \xi^{(2)}$  with  $\xi^{(i)} \in N_x^*K_{\mathbf{j}} \setminus 0$  and  $x \in K_1 \cap K_2$ .

Let us remark that, due to the transversality of  $K_1$  and  $K_2$ ,

$$N_x^*(K_1 \cap K_2) = N_x^*K_1 \oplus N_x^*K_2, \quad x \in K_1 \cap K_2.$$

Therefore the decomposition  $\xi = \xi^{(1)} + \xi^{(2)}$ ,  $\xi^{(i)} \in N_x^*K_{\mathbf{j}}$ , is unique for any covector  $\xi \in N_x^*(K_1 \cap K_2)$ . The condition that both  $\xi^{(1)}$  and  $\xi^{(2)}$  are non-zero is equivalent to

$$(x, \xi) \in N^*(K_1 \cap K_2) \setminus (N^*K_1 \cup N^*K_2).$$

We return to our study of  $v_{\mathbf{j}}$ ,  $\mathbf{j} = 1, 2, 3$ . Due to Remark 3, we can choose a pseudodifferential operator  $\chi$  such that  $\chi = 1$  near  $\text{ccl}(y, \eta)$ , and  $\text{WF}(\chi)$  is contained in a small conical neighbourhood of  $\text{ccl}(y, \eta)$ . Then applying Lemma 2 twice, we obtain

$$\chi(\text{Re}(\langle v_{\mathbf{j}}, v_{\mathbf{k}} \rangle) v_{\mathbf{l}}) \in I(\mathbb{R}^{1+3}; \Lambda_0; E \otimes \Omega^{1/2}), \tag{56}$$

where

$$\Lambda_0 = N^*(K_1 \cap K_2 \cap K_3) \setminus 0$$

and  $\mathbf{j}, \mathbf{k}, \mathbf{l} \in \{1, 2, 3\}$  are pairwise distinct. For the convenience of the reader, we give a detailed proof of this.

*Proof of (56).* To simplify the notation, we will consider the product  $\text{Re}(\langle v_1, v_2 \rangle)v_3$ , the other cases being analogous. Write

$$\eta = \eta^{(1)} + \eta^{(2)} + \eta^{(3)}, \quad \eta^{(j)} \in N_y^* K_j = \text{span}(\xi_j), \tag{57}$$

and let  $\chi_0$  and  $\chi_1$  be pseudodifferential operators such that

$$\text{WF}(\chi_0) \cap (N^* K_1 \cup N^* K_2) = \emptyset, \quad \text{WF}(\chi_1) \cap (N^*(K_1 \cap K_2) \cup N^* K_3) = \emptyset,$$

$\chi_0 = 1$  near  $\text{ccl}(y, \eta^{(1)} + \eta^{(2)})$  and  $\chi_1 = 1$  near  $\text{ccl}(y, \eta)$ . Recall that  $\eta$  is not in  $\text{span}(\xi_j, \xi_k)$  for any  $j, k \in \{1, 2, 3\}$  (see Remark 3), and this guarantees that  $\chi_0$  and  $\chi_1$  with the above properties exist. Furthermore, as  $s_{\text{in}} > 0$ , we can choose  $\chi_0$  so that  $\text{WF}(\chi_0)$  is also disjoint from  $\{x_1, x_2, x_3\}$ . Now Lemma 2 implies that

$$\begin{aligned} \chi_0(\text{Re}(v_1, v_2)) &\in I(\mathbb{R}^{1+3}; N^*(K_1 \cap K_2) \setminus 0; \Omega^{1/2}), \\ \chi_1(\chi_0(\text{Re}(v_1, v_2))v_3) &\in I(\mathbb{R}^{1+3}; \Lambda_0; E \otimes \Omega^{1/2}). \end{aligned} \tag{58}$$

We write  $w = (1 - \chi_0)(\text{Re}(v_1, v_2))v_3$  and will show that  $(y, \eta) \notin \text{WF}(w)$ . As  $\text{WF}(w)$  is conical and closed, it then follows that  $\text{WF}(\chi_1) \cap \text{WF}(w) = \emptyset$  whenever  $\text{WF}(\chi_1)$  is contained in a small enough conical neighbourhood of  $\text{ccl}(y, \eta)$ . In this case  $\chi_1 w$  is smooth, and (58), together with

$$\chi_1(\text{Re}(\langle v_1, v_2 \rangle)v_3) = \chi_1(\chi_0(\text{Re}(v_1, v_2))v_3) + \chi_1 w,$$

implies that  $\chi_1(\text{Re}(\langle v_1, v_2 \rangle)v_3) \in I(\mathbb{R}^{1+3}; \Lambda_0; E \otimes \Omega^{1/2})$ .

To simplify the notation, we write

$$\text{WF}_y(w) = \text{WF}(w) \cap T_y^* \mathbb{R}^{1+3}.$$

It remains to show that  $\eta \notin \text{WF}_y(w)$ . We have  $\text{WF}_y(v_j) \subset N_y^* K_j = \text{span}(\xi_j)$ . As the vectors  $\xi_1$  and  $\xi_2$  are linearly independent, [14, Th. 8.2.10] implies that

$$\text{WF}_y(\langle v_1, v_2 \rangle) \subset N_y^*(K_1 \cap K_2) = \text{span}(\xi_1, \xi_2).$$

Moreover, as  $\chi_0 = 1$  near  $\text{ccl}(y, \eta^{(1)} + \eta^{(2)})$ ,

$$\text{WF}_y((1 - \chi_0)(\text{Re}(v_1, v_2))) \subset \text{span}(\xi_1, \xi_2) \setminus \text{span}(\eta^{(1)} + \eta^{(2)}).$$

Using again [14, Th. 8.2.10], we have

$$\text{WF}_y(w) \subset \{\zeta^1 + \zeta^2 : \zeta^1 \in (\text{span}(\xi_1, \xi_2) \setminus \text{span}(\eta^{(1)} + \eta^{(2)})) \cup \{0\}, \zeta^2 \in \text{span}(\xi_3)\}.$$

As the direct sum decomposition (57) is unique, and as  $\eta^{(1)} + \eta^{(2)} \neq 0$  by Remark 3, it follows that  $\eta \notin \text{WF}_y(w)$ . ■

Let us denote the right-hand side of (50) by  $f_{\text{out}}$ , that is,

$$f_{\text{out}} = \text{Re}(\langle v_1, v_2 \rangle)v_3 + \text{Re}(\langle v_1, v_3 \rangle)v_2 + \text{Re}(\langle v_2, v_3 \rangle)v_1.$$

Furthermore, we write  $v_{\text{out}}$  and  $w$  for the solutions of the following two wave equations:

$$\begin{cases} P v_{\text{out}} = \chi f_{\text{out}}, \\ v_{\text{out}}|_{t < 0} = 0, \end{cases} \quad \begin{cases} P w = (1 - \chi) f_{\text{out}}, \\ w|_{t < 0} = 0, \end{cases} \tag{59}$$

where  $\chi$  is as in (56). Then the solution  $v_{123}$  of (50) satisfies  $v_{123} = v_{\text{out}} + w$ , and due to (56), it follows that  $\chi f_{\text{out}} \in I(\mathbb{R}^{1+3}; \Lambda_0; E \otimes \Omega^{1/2})$ . We will treat  $w$  as a remainder term.

Recall that  $y = \gamma_{z \leftarrow y}(0)$  and  $\eta = \gamma_{z \leftarrow y}^*(0)$ . As  $\chi = 1$  near  $\text{ccl}(y, \eta)$ , we have

$$(y, \eta) \notin \text{WF}((1 - \chi) f_{\text{out}}).$$

As  $\text{singsupp}(f_{\text{out}}) \subset \mathcal{S}$  and as  $\mathcal{S}$  is a small neighbourhood of  $\bigcup_{j=1}^3 \gamma_{y \leftarrow x_j}(\mathbb{R}_+)$  in  $\bigcup_{j=1}^3 \overline{K_j}$ , we see that the bicharacteristic

$$\beta_{\text{out}}(s) = (\gamma_{z \leftarrow y}(s), \gamma_{z \leftarrow y}^*(s)) \tag{60}$$

does not intersect  $\text{WF}((1 - \chi) f_{\text{out}})$ . Writing

$$(z, \zeta) = \beta_{\text{out}}(s_{\text{out}})$$

for some  $s_{\text{out}} > 0$  and  $\zeta \in T_z^* M \setminus 0$ , it follows from [12, Th. 23.2.9] that  $(z, \zeta) \notin \text{WF}(w)$ .

Let us now consider the future flowout  $\Lambda_1$  from  $\Lambda_0 \cap \Sigma(P)$ , with  $\Lambda_0$  as in (56). We will show in Appendix C that  $\Lambda_1 \setminus \Lambda_0$  coincides with a conormal bundle, and thus we can apply Theorem 3 to the equation for  $v_{\text{out}}$  in (59). Let us point out that this study of the structure of  $\Lambda_1$  is not essential for the proof. We could alternatively treat  $v_{\text{out}}$  as a Lagrangian distribution, but then its symbol should be viewed as a section of  $E \otimes \Omega^{1/2} \otimes \mathcal{M}$ , with  $\mathcal{M}$  the Maslov bundle over  $\Lambda_1$ . However, we prefer to avoid technicalities related to general Lagrangian distributions in the present paper.

Theorem 3 implies that  $v_{\text{out}} \in I(M; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2})$ , and as  $(z, \zeta) \notin \text{WF}(w)$ , we may choose a pseudodifferential operator  $\tilde{\chi}$  acting on  $\mathcal{D}'(\mathcal{V})$  such that  $\sigma[\tilde{\chi}] = 1$  near  $(z, \zeta)$  and  $\tilde{\chi}(v_{123}|_{\mathcal{V}}) \in I(\mathcal{V}; \Lambda_1; E \otimes \Omega^{1/2})$ . Then  $\sigma[\tilde{\chi}(v_{123}|_{\mathcal{V}})](z, \zeta) = \sigma[v_{\text{out}}](z, \zeta)$  and in terms of the source-to-solution map this reads

$$\sigma[\tilde{\chi}|g|^{1/4} \partial_{\epsilon_1} \partial_{\epsilon_2} \partial_{\epsilon_3} L_A(\epsilon_1 f_1 + \epsilon_2 f_2 + \epsilon_3 f_3)|_{\epsilon=0}](z, \zeta) = \sigma[v_{\text{out}}](z, \zeta). \tag{61}$$

This shows that  $L_A$  determines  $\sigma[v_{\text{out}}](z, \zeta)$ . We will recover  $\mathbf{S}_{z \leftarrow y \leftarrow x}^A$  by considering  $\sigma[v_{\text{out}}]$  at  $(z, \zeta) \in \Lambda_1$ .

### 3.4. From the source-to-solution map to the broken light ray transform

Having understood the propagation and interaction of the linear waves, we now prove that the source-to-solution map determines the broken non-abelian light ray transform.

**Theorem 4.** *Let  $A$  and  $B$  be two connections in  $\mathbb{R}^{1+3}$  such that  $L_A = L_B$ , as in Theorem 1. Then  $\mathbf{S}_{z \leftarrow y \leftarrow x}^A = \mathbf{S}_{z \leftarrow y \leftarrow x}^B$  for all  $(x, y, z) \in \mathbb{S}^+(\mathcal{V})$ .*

We give a constructive proof of Theorem 4 by applying a method that recovers  $S_{z \leftarrow y \leftarrow x}^A$  for any  $(x, y, z) \in \mathbb{S}^+(\mathcal{U})$  given  $L_A$ . Therefore we continue considering only a single connection  $A$ .

Let  $(x, y, z) \in \mathbb{S}^+(\mathcal{U})$  and let  $\eta, \xi_1 \in T_y^*M$  be defined by (51). Without loss of generality we may assume that  $\eta$  and  $\xi_1$  are of the form (45), and letting again  $\xi_2, \xi_3 \in T_y^*M$  be as in Lemma 1, we have (47). Recall that  $K_j$  is defined by (55). Decomposing  $\eta$  as in (57), for small  $r > 0$  we have

$$\eta^{(1)} = -r^{-2}\lambda_1\xi_1, \quad \eta^{(j)} = r^{-2}\lambda_j\xi_j, \quad j = 2, 3, \tag{62}$$

where

$$\lambda_1 = 2b + \mathcal{O}(r), \quad \lambda_j = b + \mathcal{O}(r), \quad j = 2, 3.$$

We also recall that  $b \geq 1$  (see Remark 3). In particular,  $\lambda_j > 0, j = 1, 2, 3$ , for small  $r > 0$ . Let  $v_j$  be again the solution of (54) with  $f_j$  as in (53), and write

$$\hat{v}_j = \sigma[v_j](y, \eta^{(j)}), \quad j = 1, 2, 3.$$

The product calculus in Lemma 2 implies that  $\chi f_{\text{out}}$  in (59) satisfies

$$|g(y)|^{1/2}\sigma[\chi f_{\text{out}}](y, \eta) = \text{Re}(\langle \hat{v}_1, \hat{v}_2 \rangle)\hat{v}_3 + \text{Re}(\langle \hat{v}_1, \hat{v}_3 \rangle)\hat{v}_2 + \text{Re}(\langle \hat{v}_2, \hat{v}_3 \rangle)\hat{v}_1, \tag{63}$$

where we have used  $\chi(y, \eta) = 1$ .

We will apply Proposition 1 to (54). Toward that end, we choose a strictly positive  $\omega_j \in C^\infty(N^*K_j; \Omega^{1/2})$  satisfying (32), and denote by  $\rho_j$  the factor (33) along the bicharacteristic  $\beta_j(s) = (\gamma_{y \leftarrow x_j}(s), \dot{\gamma}_{y \leftarrow x_j}^*(s))$ . It follows from (53) and  $(\chi 1)$  that in any coordinates, ignoring  $2\pi$  factors, we have

$$\sigma[|g|^{1/4}f_j](x_j, \pm\lambda\xi_j) = c_j|g(x_j)|^{1/4}\sigma[\chi_j](x_j, \pm\lambda\xi_j) = \lambda^q c_j \alpha_j, \quad \lambda > 0, \tag{64}$$

where  $\alpha_j = |g(x_j)|^{1/4}\sigma[\chi_j](x_j, \xi_j)$ . Of course,  $|g(x_j)|^{1/4} = 1$  in the Cartesian coordinates, but  $\alpha_j$  transforms as a half-density. Observe, however, that  $\alpha_j > 0$  in any coordinate system.

Let us translate the origin to  $x_1$  and pass to the coordinates  $(\tilde{x}^0, \tilde{x}', \tilde{\xi}_0, \tilde{\xi}')$  in (83) in Appendix B below. In these coordinates  $(x_1, -\lambda\xi_1) = (0; 0, -\lambda\xi'_1)$  for  $\lambda > 0$ . Moreover, (84) and (64) give

$$\sigma[v_1](0, -\lambda\xi'_1) = -\iota_0\lambda^{q-1}c_1\alpha_1, \quad \lambda > 0,$$

where  $\iota_0 \in \mathbb{C} \setminus 0$  is a constant that contains all the  $2\pi$  and  $\iota$  factors. In particular, the positive homogeneity (37) holds for  $v_1$  at  $(x_1, -\xi_1)$ . The coordinates (83) can be used in a neighbourhood of the bicharacteristic segment  $\beta_1^-(s) = (\gamma_{y \leftarrow x_1}(s), -\dot{\gamma}_{y \leftarrow x_1}^*(s))$ ,  $s \in [0, s_{\text{in}}]$ , and in these coordinates  $(y, -\xi_1) = (s_{\text{in}}0; 0, -\xi'_1)$ . It follows from Proposition 1 and (62) that

$$e^{\rho_1(s_{\text{in}})}(\omega_1^{-1}\sigma[v_1])(s_{\text{in}}, -r^{-2}\lambda_1\xi'_1) = (r^{-2}\lambda_1)^{q-1}\mathbf{P}_{y \leftarrow x_1}^A((\omega_1^{-1}\sigma[v_1])(0, -\xi'_1)),$$

and therefore in the coordinates (83),

$$\hat{v}_1 = -(r^{-2}\lambda_1)^{q-1}\iota_0\tilde{\alpha}_1\mathbf{P}_{y\leftarrow x_1}^A c_1, \tag{65}$$

where  $\tilde{\alpha}_1 := e^{-\rho_1(s_{\text{in}})}\omega_1(s_{\text{in}}, \xi'_1)\omega_1^{-1}(0, \xi'_1)\alpha_1$  contains all the volume factors.

Analogously, we may use coordinates of the form (81) in a neighbourhood of the bicharacteristic segment  $\beta_j(s)$ ,  $s \in [0, s_{\text{in}}]$ ,  $\mathbf{j} = 2, 3$ , to obtain

$$\hat{v}_j = (r^{-2}\lambda_j)^{q-1}\iota_0\tilde{\alpha}_j\mathbf{P}_{y\leftarrow x_j}^A c_j, \quad \mathbf{j} = 2, 3, \tag{66}$$

where  $\tilde{\alpha}_j = e^{-\rho_j(s_{\text{in}})}\omega_j(s_{\text{in}}, \xi'_j)\omega_j^{-1}(0, \xi'_j)\alpha_j$ . We have given the volume factors  $\tilde{\alpha}_j$ ,  $\mathbf{j} = 1, 2, 3$ , in different coordinate systems, but the way they transform under changes of coordinates is inconsequential. For our present purposes it is enough to observe that they are independent of the connection  $A$  and satisfy  $\tilde{\alpha}_j > 0$  in any coordinate system. Let us also emphasize that  $\hat{v}_j$ , as the value of a section of  $E \otimes \Omega^{1/2}$  over  $K_j$  at the point  $(y, \eta^{(\mathbf{j})})$ , is a coordinate invariant quantity.

Combining (63), (65) and (66) yields  $|g(y)|^{1/2}\sigma[\chi f_{\text{out}}](y, \eta) = \alpha_{\text{in}}c_{\text{in}}$  where

$$\begin{aligned} c_{\text{in}} &= \text{Re}(\langle c_{1,\text{in}}, c_{2,\text{in}} \rangle)c_{3,\text{in}} + \text{Re}(\langle c_{1,\text{in}}, c_{3,\text{in}} \rangle)c_{2,\text{in}} + \text{Re}(\langle c_{2,\text{in}}, c_{3,\text{in}} \rangle)c_{1,\text{in}}, \\ \alpha_{\text{in}} &= -r^{-6(q-1)}(\lambda_1\lambda_2\lambda_3)^{q-1}|\iota_0|^2\tilde{\alpha}_1\tilde{\alpha}_2\tilde{\alpha}_3, \quad c_{\mathbf{j},\text{in}} = \mathbf{P}_{y\leftarrow x_j}^A c_j. \end{aligned}$$

Similarly to the above, we may apply Proposition 1 to (59) and the bicharacteristic  $\beta_{\text{out}}$  defined by (60). This gives, ignoring  $\iota$  and  $2\pi$  factors,

$$\sigma[v_{\text{out}}](z, \zeta) = \alpha_{\text{out}}\alpha_{\text{in}}\mathbf{P}_{z\leftarrow y}^A c_{\text{in}}$$

for some volume factor  $\alpha_{\text{out}} > 0$  independent of the connection  $A$ .

Recall that the source-to-solution map  $L_A$  determines  $\sigma[v_{\text{out}}](z, \zeta)$  via (61), and observe that both the factors  $\alpha_{\text{in}}$  and  $\alpha_{\text{out}}$  are independent of  $A$ . Therefore  $L_A$  determines the parallel transport  $\mathbf{P}_{z\leftarrow y}^A c_{\text{in}}$ . Letting  $x_2, x_3 \rightarrow x_1 = x$  and  $c_2, c_3 \rightarrow c_1$ , we have

$$\mathbf{P}_{z\leftarrow y}^A c_{\text{in}} \rightarrow 3|\mathbf{P}_{y\leftarrow x}^A c_1|^2\mathbf{P}_{z\leftarrow y}^A\mathbf{P}_{y\leftarrow x}^A c_1 = 3|c_1|^2\mathbf{S}_{z\leftarrow y\leftarrow x}^A c_1,$$

where we have used the fact that  $\mathbf{P}_{y\leftarrow x}^A$  is unitary. Thus  $L_A$  determines  $\mathbf{S}_{z\leftarrow y\leftarrow x}^A$  after varying  $c_1 \in E_x \setminus 0$ , and we have shown Theorem 4.

#### 4. Inversion of the broken light ray transform

From now on we assume that  $E$  is the trivial bundle  $M \times \mathbb{C}^n$ . Recall the definition (5) of  $\mathfrak{U}$ . More generally, we write

$$\mathfrak{U}(\epsilon) = (0, 1) \times B(\epsilon)$$

where  $B(\epsilon)$  is the open ball of radius  $\epsilon > 0$ , centred at the origin of  $\mathbb{R}^3$ . We also use the shorthand notation

$$\mathbb{D}(\mathfrak{U}(\epsilon)) = \{y \in M : \text{there is } (x, y, z) \in \mathbb{S}^+(\mathfrak{U}(\epsilon))\}, \quad 0 < \epsilon < \epsilon_0,$$

where  $\mathbb{S}^+(\mathfrak{U}(\epsilon))$  is defined by (43).

We recall that  $S^A_{z \leftarrow y \leftarrow x}$  is defined by (44), that is,

$$S^A_{z \leftarrow y \leftarrow x} = P^A_{z \leftarrow y} P^A_{y \leftarrow x}, \quad (x, y, z) \in S^+(\mathcal{U}),$$

and that the parallel transport map  $P^A_{y \leftarrow x}$  is the fundamental solution to the ordinary differential equation (35). In this section we will prove the following variation of Theorem 2.

**Theorem 5.** *Let  $A$  and  $B$  be two connections in  $\mathbb{R}^{1+3}$  such that for all  $0 < \epsilon < \epsilon_0$ ,*

$$S^A_{z \leftarrow y \leftarrow x} = S^B_{z \leftarrow y \leftarrow x} \quad \text{for all } (x, y, z) \in S^+(\mathcal{U}(\epsilon)). \tag{67}$$

*Then there exists a smooth  $\mathbf{u} : \mathbb{D}(\Omega(\epsilon_0)) \rightarrow U(n)$  such that*

$$\mathbf{u}|_{\mathcal{U}(\epsilon_0)} = \text{id} \quad \text{and} \quad B = \mathbf{u}^{-1} d\mathbf{u} + \mathbf{u}^{-1} A \mathbf{u}.$$

The causal diamond defined by (9) satisfies  $\mathbb{D} = \mathbb{D}(\mathcal{U}(\epsilon_0))$ , and therefore Theorem 2 follows immediately from Theorem 5. The latter is, however, more convenient when proving Theorem 1. In fact, Theorem 1 follows simply by combining Theorems 4 and 5. Similarly to  $S^{\text{out}}(\mathcal{U})$  we define

$$S^{\text{in}}(\mathcal{U}) = \{(x, y) : (x, y, z) \in S^+(\mathcal{U}) \text{ for some } z \in \mathcal{U}\}.$$

**Lemma 3.** *Let  $A$  and  $B$  be two connections in  $\mathbb{R}^{1+3}$ . Define  $\mathbf{u}(y, x) = P^A_{y \leftarrow x} P^B_{x \leftarrow y}$  for  $(x, y) \in S^{\text{in}}(\mathcal{U})$ . If (67) holds then  $\mathbf{u}(y, x_1) = \mathbf{u}(y, x_2)$  for all  $(x_j, y) \in S^{\text{in}}(\mathcal{U})$ ,  $j = 1, 2$ .*

*Proof.* Note that  $P^A_{y \leftarrow x}$  is a linear isomorphism and  $(P^A_{y \leftarrow x})^{-1} = P^A_{x \leftarrow y}$ . In particular,

$$(S^A_{z \leftarrow y \leftarrow x})^{-1} = S^A_{x \leftarrow y \leftarrow z},$$

and (67) implies that  $S^A_{x \leftarrow y \leftarrow z} = S^B_{x \leftarrow y \leftarrow z}$  for all  $(x, y, z) \in S^+(\mathcal{U})$ .

Consider now  $y, x_1$  and  $x_2$  as in the claim. Then there is  $z \in \mathcal{U}$  such that  $(x_j, y, z) \in S^+(\mathcal{U})$  for  $j = 1, 2$ . We have

$$S^A_{x_2 \leftarrow y \leftarrow z} S^A_{z \leftarrow y \leftarrow x_1} = P^A_{x_2 \leftarrow y} P^A_{y \leftarrow z} P^A_{z \leftarrow y} P^A_{y \leftarrow x_1} = P^A_{x_2 \leftarrow y} P^A_{y \leftarrow x_1}.$$

But  $S^A_{x_2 \leftarrow y \leftarrow z} S^A_{z \leftarrow y \leftarrow x_1} = S^B_{x_2 \leftarrow y \leftarrow z} S^B_{z \leftarrow y \leftarrow x_1}$ , and so  $P^A_{x_2 \leftarrow y} P^A_{y \leftarrow x_1} = P^B_{x_2 \leftarrow y} P^B_{y \leftarrow x_1}$ . We apply  $P^A_{y \leftarrow x_2}$  on the left and  $P^B_{x_1 \leftarrow y}$  on the right, and obtain

$$P^A_{y \leftarrow x_1} P^B_{x_1 \leftarrow y} = P^A_{y \leftarrow x_2} P^B_{x_2 \leftarrow y}. \quad \blacksquare$$

*Proof of Theorem 5.* For  $y \in \mathbb{D}(\mathcal{U})$  there are  $x, z \in \mathcal{U}$  such that  $(x, y, z) \in S^+(\mathcal{U})$  and we define  $\mathbf{u}(y, x)$  as in Lemma 3. It follows from Lemma 3 that  $\mathbf{u}(y, x) = \mathbf{u}(y)$  and  $\mathbf{u}$  can be viewed as a function of  $y \in \mathbb{D}(\mathcal{U})$ . The parallel transport map takes values in  $U(n)$  and therefore  $\mathbf{u}$  is a section of the trivial principal bundle over  $\mathbb{D}(\mathcal{U})$  with group  $U(n)$ .

Observe that  $\mathcal{U} \cap \mathbb{D}(\mathcal{U}) = \emptyset$  by the definition of  $S^+(\mathcal{U})$  (see (43)). For this reason we shrink  $\mathcal{U}$ , that is, we define  $\mathbf{u}$  as above but replace  $\mathcal{U}$  with  $\mathcal{U}(\epsilon)$ ,  $0 < \epsilon < \epsilon_0$ . This allows us to define  $\mathbf{u}$  on  $\mathcal{U}(\epsilon_0) \setminus \mu([0, 1])$  (see (6) for the definition of the path  $\mu$ ). By the continuity of the parallel transport map, we can define  $\mathbf{u}$  on the whole  $\mathcal{U}(\epsilon_0)$ . Using the

continuity again, we let  $x \rightarrow y \in (0, 1) \times \partial B(\epsilon)$  in  $\mathbf{u}(y, x)$ , and see that  $\mathbf{u}(y) = \text{id}$  for any  $y \in (0, 1) \times \partial B(\epsilon)$  and any  $0 < \epsilon < \epsilon_0$ . Using the continuity once more, we see that  $\mathbf{u} = \text{id}$  in the whole  $\mathfrak{U} = \mathfrak{U}(\epsilon_0)$ .

We write  $\tilde{A} = \mathbf{u}^{-1}d\mathbf{u} + \mathbf{u}^{-1}A\mathbf{u}$ . Let  $(x, y) \in \mathbb{S}^{\text{in}}(\mathfrak{U}(\epsilon))$ ,  $0 < \epsilon < \epsilon_0$ , and consider the geodesic  $\gamma_{y \leftarrow x}$  from  $x$  to  $y$ . Let  $s_{\text{in}} > 0$  satisfy  $y = \gamma_{y \leftarrow x}(s_{\text{in}})$ . By using the choice  $\mathbf{u}(y) = \mathbf{u}(y, x)$  we will show that

$$\langle \tilde{A}, \dot{\gamma}_{y \leftarrow x} \rangle = \langle B, \dot{\gamma}_{y \leftarrow x} \rangle. \tag{68}$$

We write  $\gamma = \gamma_{y \leftarrow x}$  and consider the fundamental matrix solution

$$\mathbf{U}^A : \mathbb{R}^2 \rightarrow U(n)$$

of (35). That is, for fixed  $s \in \mathbb{R}$ , the function  $\mathbf{U}^A(t, s)$  in  $t$  is the solution of

$$\begin{cases} \partial_t \mathbf{U}^A(t, s) + \langle A, \dot{\gamma}(t) \rangle \mathbf{U}^A(t, s) = 0 & \text{on } \mathbb{R}, \\ \mathbf{U}^A(s, s) = \text{id}. \end{cases} \tag{69}$$

Clearly  $\mathbf{P}_\gamma^A = \mathbf{U}^A(s_{\text{in}}, 0)$ . We define  $\mathbf{U}^B$  analogously.

By (69) we have

$$\partial_t \mathbf{U}^A(s, s) = -\langle A, \dot{\gamma}(s) \rangle \mathbf{U}^A(s, s) = -\langle A, \dot{\gamma}(s) \rangle.$$

Moreover, differentiating (69) in  $s$  gives

$$\begin{cases} \partial_t \partial_s \mathbf{U}^A(t, s) + \langle A, \dot{\gamma}(t) \rangle \partial_s \mathbf{U}^A(t, s) = 0, \\ \partial_t \mathbf{U}^A(s, s) + \partial_s \mathbf{U}^A(s, s) = 0. \end{cases}$$

Therefore  $\mathbf{W}(t, s) = \partial_s \mathbf{U}^A(t, s)$  satisfies

$$\begin{cases} \partial_t \mathbf{W}(t, s) + \langle A, \dot{\gamma}(t) \rangle \mathbf{W}(t, s) = 0, \\ \mathbf{W}(s, s) = \langle A, \dot{\gamma}(s) \rangle, \end{cases}$$

and writing  $\mathbf{W}$  in terms of the fundamental solution  $\mathbf{U}^A$  gives

$$\partial_s \mathbf{U}^A(t, s) = \mathbf{U}^A(t, s) \langle A, \dot{\gamma}(s) \rangle. \tag{70}$$

We have  $\mathbf{u}(\gamma(t), y) = \mathbf{U}^A(t, 0) \mathbf{U}^B(0, t)$ , and using (69) for  $A$  and (70) for  $B$  we get

$$\begin{aligned} \langle d\mathbf{u}, \dot{\gamma}(t) \rangle &= \partial_t (\mathbf{U}^A(t, 0) \mathbf{U}^B(0, t)) \\ &= -\langle A, \dot{\gamma}(t) \rangle \mathbf{U}^A(t, 0) \mathbf{U}^B(0, t) + \mathbf{U}^A(0, t) \mathbf{U}^B(0, t) \langle B, \dot{\gamma}(t) \rangle \\ &= -\langle A\mathbf{u}, \dot{\gamma}(t) \rangle + \langle \mathbf{u}B, \dot{\gamma}(t) \rangle. \end{aligned}$$

We obtain (68) after rearranging and multiplying both sides with  $\mathbf{u}^{-1}$ ,

$$\langle \mathbf{u}^{-1}d\mathbf{u} + \mathbf{u}^{-1}A\mathbf{u}, \dot{\gamma}(t) \rangle = \langle B, \dot{\gamma}(t) \rangle.$$

We have shown that (68) holds for any  $(x, y) \in \mathbb{S}^{\text{in}}(\mathfrak{U}(\epsilon))$ . As  $\mathfrak{U}(\epsilon)$  is open, Lemma 1 shows that the vectors  $\dot{\gamma}_{y \leftarrow x}$  span  $T_y M$  as  $x$  varies in the set  $\{x \in \mathfrak{U} : (x, y) \in \mathbb{S}^{\text{in}}(\mathfrak{U}(\epsilon))\}$ . Hence  $\tilde{A} = B$  at  $y$  for each  $y \in \mathbb{D}(\mathfrak{U}(\epsilon))$  and each  $0 < \epsilon < \epsilon_0$ . ■



### Appendix A. Conormal distributions and IPL distributions

We formulate the framework of  $E \otimes \Omega^{1/2}$ -valued conormal and IPL distributions for the application to the connection wave equation. The contents of the appendix is modelled on the pioneering works of Hörmander [11], Duistermaat–Hörmander [9] and Melrose–Uhlmann [31]. Additionally, we refer the reader to Hörmander’s books [12, 14] for the basics of distributions and half-densities.

Let us begin with the notion of conormal distribution.

**Definition 1.** Let  $X$  be an  $n$ -dimensional manifold and  $Y$  an  $(n - N)$ -dimensional closed submanifold. We say a section-valued distribution  $u \in \mathcal{D}'(X, E \otimes \Omega^{1/2})$  is a *conormal distribution*,  $u \in I^m(X; N^*Y; E \otimes \Omega^{1/2})$ , if in local coordinates  $(x', x'') \in \mathbb{R}^{N+(n-N)}$  on  $X$  such that  $Y$  is defined by  $x' = 0$ , the distribution  $u$  takes the form

$$u = (2\pi)^{-(n+2N)/4} e^{-i\pi N/4} \int_{\mathbb{R}^N} e^{i\theta x'} a(x, \theta) d\theta, \tag{71}$$

where  $a \in S^{m+n/4}(\mathbb{R}^n \times (\mathbb{R}^N \setminus 0); E \otimes \Omega^{1/2})$ .

From the viewpoint of Lagrangian distributions,  $u$  is associated with the conormal bundle  $N^*Y \setminus 0$ , and  $\text{WF}(u) \subset N^*Y \setminus 0$ . If  $\xi = (\xi', \xi'')$  denotes the induced coordinates on the cotangent space of  $X$ ,  $N^*Y$  is defined by  $\{x' = 0, \xi'' = 0\}$ . The principal symbol  $\sigma[u] \in S^{m+n/4}(N^*Y \setminus 0; E \otimes \Omega^{1/2})$  is defined as  $\sigma[u](x'', \xi') = a(0, x'', \xi')$ . We have the exact sequence

$$\begin{aligned} 0 \hookrightarrow I^{m-1}(X; N^*Y \setminus 0; E \otimes \Omega^{1/2}) \hookrightarrow I^m(X; N^*Y \setminus 0; E \otimes \Omega^{1/2}) \\ \xrightarrow{\sigma} S^{m+n/4}(N^*Y \setminus 0; E \otimes \Omega^{1/2}) \rightarrow 0. \end{aligned}$$

When the source term for a linear wave equation is a conormal distribution, the solution is not in the same class, and the wider class of IPL distributions was introduced in [31] to tackle this problem.

We begin with the model case where  $\tilde{\Lambda}_0 = T_0^*\mathbb{R}^n \setminus 0$  and

$$\tilde{\Lambda}_1 = \{(x, \xi) \in T^*\mathbb{R}^n \setminus 0 : \xi_1 = 0, x^1 \geq 0, x^2 = \dots = x^n = 0\}.$$

We write  $x = (x^1, x')$  and  $\xi = (\xi_1, \xi')$ .

**Definition 2.** We say that a compactly supported distribution  $u \in \mathcal{E}'(\mathbb{R}^n; E \otimes \Omega^{1/2})$  belongs to the space  $I^m(\mathbb{R}^n; \tilde{\Lambda}_0, \tilde{\Lambda}_1; E \otimes \Omega^{1/2})$  if modulo compactly supported smooth functions, it can be expressed as

$$u = \int_0^\infty \int_{\mathbb{R}^n} \exp(i\xi_1(x^1 - s) + \xi'x') a(s, x, \xi) d\xi ds,$$

where the amplitude  $a \in S^{m+1/2-n/4}(\mathbb{R}^n \times (\mathbb{R}^n \setminus 0); E \otimes \Omega^{1/2})$  is supported on the conical closure of a compact set.

Observe that  $\tilde{\Lambda}_0 = N^*\{0\} \setminus 0$ , and that  $\tilde{\Lambda}_1 \setminus \partial\tilde{\Lambda}_1 \subset N^*Y \setminus 0$  where  $Y = \{x' = 0\}$ . The wavefront set of  $A$  is contained in the union  $\tilde{\Lambda}_0 \cup \tilde{\Lambda}_1$  and away from the intersection  $\partial\tilde{\Lambda}_1 = \tilde{\Lambda}_0 \cap \tilde{\Lambda}_1$ , the IPL distribution  $A$  is a conormal distribution in the following sense (see [31, pp. 486–487, Proposition 2.3 and Remark 2.7]).

**Proposition 2.** *Suppose  $u \in I^m(\mathbb{R}^n; \tilde{\Lambda}_0, \tilde{\Lambda}_1; E \otimes \Omega^{1/2})$  and  $\chi$  is a zero-th order pseudodifferential operator. We have*

$$\chi u \in \begin{cases} I^m(\mathbb{R}^n; N^*Y \setminus 0; E \otimes \Omega^{1/2}) & \text{when } \text{WF}(\chi) \cap \tilde{\Lambda}_0 = \emptyset, \\ I^{m-1/2}(\mathbb{R}^n; N^*\{0\} \setminus 0; E \otimes \Omega^{1/2}) & \text{when } \text{WF}(\chi) \cap \tilde{\Lambda}_1 = \emptyset. \end{cases}$$

This microlocalization leads to the following local principal symbol maps, ignoring  $2\pi$  and  $i$  factors related to the normalization in (71):

$$\begin{aligned} \sigma^{(1)}[u] &= a(x^1, (x^1, 0), (0, \xi')) && \text{on } \tilde{\Lambda}_1 \setminus \partial\tilde{\Lambda}_1, \\ \sigma^{(0)}[u] &= -i a(0, 0, \xi) / \xi_1 && \text{on } \tilde{\Lambda}_0 \setminus \partial\tilde{\Lambda}_1, \\ \xi_1 \sigma^{(0)}[u] &= -i \sigma^{(1)}[u] && \text{on } \partial\tilde{\Lambda}_1. \end{aligned} \tag{72}$$

IPL distributions can be defined on any pair of Lagrangians  $(\Lambda_0, \Lambda_1)$ ,  $\Lambda_j \subset T^*X \setminus 0$ , with clean intersection. By *clean intersection* of two Lagrangians, we mean

$$T_\lambda(\Lambda_0) \cap T_\lambda(\Lambda_1) = T_\lambda(\partial\Lambda_1) \quad \text{for any } \lambda \in \partial\Lambda_1,$$

given two Lagrangians  $\Lambda_0$  and  $\Lambda_1$  with  $\Lambda_0 \cap \Lambda_1 = \partial\Lambda_1$ . However, we will omit discussion of Lagrangian distributions, and make the additional assumption that

$$\Lambda_j \setminus \partial\Lambda_1 \subset N^*Y_j \setminus 0, \quad j = 0, 1, \tag{73}$$

for some submanifolds  $Y_j \subset X$ .

**Definition 3.** Suppose  $(\Lambda_0, \Lambda_1)$ , a pair of Lagrangians over a smooth  $n$ -manifold  $X$ , intersect cleanly at  $\Lambda_1$ . The space  $I^m(X; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2})$  consists of distributions of the form

$$u_0 + u_1 + \sum_j F_j v_j,$$

where  $u_0 \in I^{m-1/2}(X; \Lambda_0; E \otimes \Omega^{1/2})$ ;  $u_1 \in I^m(X; \Lambda_1 \setminus \partial\Lambda_1; E \otimes \Omega^{1/2})$ ;  $\{F_j\}$  is a family of zero-th order Fourier integral operators associated with the inverse of the homogeneous symplectic transformation from  $V_j$  to  $T^*\mathbb{R}^n$ , where  $\{V_j\}$  is a locally finite, countable covering of  $\partial\Lambda_1$ ; and  $v_j \in I^m(\mathbb{R}^n; \tilde{\Lambda}_0, \tilde{\Lambda}_1; E \otimes \Omega^{1/2})$ .

To symbolically construct the parametrix of the wave operator, the key thing is to understand the principal symbols of IPL distributions. As in the model case, away from the intersection of  $\Lambda_0$  and  $\Lambda_1$ , the corresponding IPL distributions are conormal distributions assuming (73).

**Proposition 3.** *Suppose that (73) holds. Let  $u \in I^m(X; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2})$  and let  $\chi$  be a properly supported zero-th order pseudodifferential operator. We have*

$$\chi u \in \begin{cases} I^m(X; N^*Y_1 \setminus 0; E \otimes \Omega^{1/2}) & \text{when } \text{WF}(\chi) \cap \Lambda_0 = \emptyset, \\ I^{m-1/2}(X; N^*Y_0 \setminus 0; E \otimes \Omega^{1/2}) & \text{when } \text{WF}(\chi) \cap \Lambda_1 = \emptyset. \end{cases}$$

Choosing  $\chi$  so that  $\sigma[\chi](x, \xi) \neq 0$  at  $(x, \xi) \in (N^*Y_0 \setminus 0 \cup N^*Y_1 \setminus 0) \setminus \partial\Lambda_1$ , we can define the principal symbols of  $u$  away from the intersection  $\partial\Lambda_1$ ,

$$\begin{aligned} \sigma^{(1)}[u] &= \sigma[\chi u] / \sigma[\chi] \in S^{m+n/4}(N^*Y_1 \setminus 0; E \otimes \Omega^{1/2}), \\ \sigma^{(0)}[u] &= \sigma[\chi u] / \sigma[\chi] \in S^{m-1/2+n/4}(N^*Y_0 \setminus 0; E \otimes \Omega^{1/2}). \end{aligned}$$

However, the map from the IPL distributions  $I^m(X; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2})$  to the symbols

$$S^{m+n/4}(\Lambda_1 \setminus \partial\Lambda_1; E \otimes \Omega^{1/2}) \times S^{m-1/2+n/4}(\Lambda_0 \setminus \partial\Lambda_1; E \otimes \Omega^{1/2})$$

cannot be defined as the principal symbol map, since it is not surjective. Indeed, Proposition 2 implies that  $\sigma^{(1)}[u]$  extends to a smooth section of  $E \otimes \Omega^{1/2}$  up to  $\partial\Lambda_1$ , whilst  $h\sigma^{(0)}[u]$  extends to a smooth section over  $\Lambda_0$  if  $h$  is a smooth function vanishing on  $\partial\Lambda_1$ . Hence the principal symbol space should be a proper subspace of the product space of symbols.

Moreover, the fact that the Lagrangians  $\Lambda_j$ ,  $j = 0, 1$ , may not be conormal bundles near the intersection  $\partial\Lambda_1$  causes additional complications. In the case of applications that we are interested in,  $\Lambda_1$  fails to be a conormal bundle near  $\partial\Lambda_1$ . The principal symbol of a general Lagrangian distribution is not a section of  $E \otimes \Omega^{1/2}$  but of  $E \otimes \Omega^{1/2} \otimes L$  where  $L$  is the Maslov bundle. To avoid discussion of Maslov bundles over  $\Lambda_j$ ,  $j = 0, 1$ , we make the assumption that they are trivial. In all the cases we are interested in,  $\Lambda_0$  is a conormal bundle and  $\Lambda_1$  is the future flowout of  $\Lambda_0 \cap \Sigma(P)$ , with  $P$  the wave operator in the Minkowski space. We will show that the Maslov bundles over the flowouts of interest are trivial in Section B.1 below.

To understand the principal symbol space, assuming triviality of the Maslov bundles, we now elucidate the relationship between  $\sigma^{(1)}[u]$  and  $\sigma^{(0)}[u]$ . Following [31], we define a map

$$R : S^{m-1/2+n/4}(\Lambda_0 \setminus \partial\Lambda_1; E \otimes \Omega^{1/2}) \rightarrow S^{m+n/4}(\Lambda_1; E \otimes \Omega^{1/2})|_{\partial\Lambda_1}$$

that encodes the relation (72). Then the sections

$$\left\{ (a^{(1)}, a^{(0)}) : \begin{aligned} &a^{(0)} \in S^{m-1/2+n/4}(\Lambda_0 \setminus \partial\Lambda_1; E \otimes \Omega^{1/2}), \\ &a^{(1)} \in S^{m+n/4}(\Lambda_1; E \otimes \Omega^{1/2}), \\ &a^{(1)}|_{\partial\Lambda_1} = \mathcal{R}a^{(0)}, \\ &ha^{(0)} \text{ is smooth on } \partial\Lambda_1 \text{ if } h \text{ vanishes on } \partial\Lambda_1 \end{aligned} \right\},$$

where

$$\mathcal{R} = e^{i\pi/4}(2\pi)^{1/4}R, \tag{74}$$

will be the desired principal symbol space, which we will call  $S^m(\Lambda_1, \Lambda_0; E \otimes \Omega^{1/2})$ , and we have the exact sequence

$$0 \hookrightarrow I^{m-1/2}(X; \Lambda_0; E \otimes \Omega^{1/2}) + I^{m-1}(X; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2}) \hookrightarrow I^m(X; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2}) \xrightarrow{\sigma} S^{m+n/4}(\Lambda_0, \Lambda_1; E \otimes \Omega^{1/2}) \rightarrow 0. \tag{75}$$

We remark that the constant  $e^{i\pi/4}(2\pi)^{1/4}$  in (74) results from the fact that  $\Lambda_0$  is parametrized with  $n + 1$  phase variables whilst  $\Lambda_1$  is parametrized with  $n$  phase variables (cf. the normalization in (71)).

Apart from the Hermitian bundle factor, the map  $R$  was constructed by Melrose–Uhlmann [31, pp. 491–493]. But that factor is harmless. Indeed, after passing to the model case, we can simply define  $R$  by (72), that is,  $Ra^{(0)} = \iota(\xi_1 a^{(0)})|_{\partial\Lambda_1}$ . This definition entails, of course, that  $R$  does not depend on the choice of a homogeneous symplectic transformation that maps  $(\Lambda_0, \Lambda_1)$  locally to the model case  $(\tilde{\Lambda}_0, \tilde{\Lambda}_1)$ . Such coordinate invariance was shown in [31].

Having the principal symbol map (75), we are ready to give a proof of Theorem 3, following [31].

*Proof of Theorem 3.* We will construct a parametrix for (22), and the claim will then follow from solving (22) with a smooth function  $f$  on the right-hand side. The first step of the construction is to find  $u_{(0)} \in I^{k-2+1/2}(M; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2})$  such that

$$Pu_{(0)} = f + f_{(1)} + e_{(1)}, \tag{76}$$

where  $f_{(1)} \in I^{k-1}(M; \Lambda_0; E \otimes \Omega^{1/2})$  and  $e_{(1)} \in I^{k-3/2}(M; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2})$ .

To do so, we use the symbol calculus. First of all, since  $P$  is elliptic on  $\Lambda_0 \setminus \partial\Lambda_1$ , we choose

$$\sigma[u_{(0)}] = p^{-1}\sigma[f] \quad \text{on } \Lambda_0 \setminus \partial\Lambda_1,$$

where we use the shorthand notation  $p = \sigma[P]$  for the principal symbol of  $P$ . As an element of  $I^{k-2+1/2}(M; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2})$ ,  $u_{(0)}$  has the following principal symbol on  $\partial\Lambda_1$ :

$$\sigma[u_{(0)}]|_{\partial\Lambda_1} = \mathcal{R}(p^{-1}\sigma[f]). \tag{77}$$

This can be viewed as the initial condition of the bicharacteristic flow on  $\Lambda_1$  emanating from  $\Lambda_0$ . On the other hand, the symbol calculus on  $\Lambda_1$  obeys the transport equation

$$\sigma[Pu] = (i^{-1}\mathcal{L}_{H_P} + c)\sigma[u] \quad \text{on } \Lambda_1,$$

where  $c = \sigma_{\text{sub}}[P]$  is the subprincipal symbol of  $P$ . Noting that  $\text{WF}(f)$  does not intersect  $\Lambda_1 \setminus \partial\Lambda_1$ , we have

$$(\mathcal{L}_{H_P} + ic)\sigma[u_{(0)}] = 0 \quad \text{on } \Lambda_1 \setminus \partial\Lambda_1. \tag{78}$$

Combining (77) and (78), we have solved (76). Next, we iteratively solve the equations

$$Pu_{(j-1)} = f_{(j-1)} + e_{(j-1)} + f_{(j)} + e_{(j)}, \tag{79}$$

where  $f_{(j)} \in I^{k-j}(M; \Lambda_0; E \otimes \Omega^{1/2})$  and  $e_{(j)} \in I^{k-j-1/2}(M; \Lambda_0, \Lambda_1; E \otimes \Omega^{1/2})$ . This can be done by choosing  $u_{(j)}$  that obey

$$\begin{aligned} \sigma[u_{(j)}] &= p^{-1}\sigma[f_{(j)}] && \text{on } \Lambda_0, \\ (\mathcal{L}_{H_p} + \iota c)\sigma[u_{(j)}] &= \iota\sigma[e_{(j)}] && \text{on } \Lambda_1 \setminus \partial\Lambda_1, \\ \sigma[u_{(j)}] &= \mathcal{R}(p^{-1}\sigma[f_{(j)}]) && \text{on } \partial\Lambda_1. \end{aligned}$$

We complete the proof by adding up the equations (78) and (79) for  $j = 1, \dots, N$ , and letting  $N \rightarrow \infty$ . ■

**Appendix B. Theorem 3 in the context of Example 1**

Let  $\Lambda_j, j = 0, 1$ , be as in Example 1, that is,  $\Lambda_0 = N^*\{0\} \setminus 0$  and, taking into account the microlocal cutoff  $\chi$ ,

$$\Lambda_1 = \{(t, t\theta; -\lambda, \lambda\theta) \in \mathbb{R}^{1+3} \times \mathbb{R}^{1+3} : \lambda \in \mathbb{R} \setminus 0, \theta \in \mathcal{V}, t \geq 0\}, \tag{80}$$

where  $\mathcal{V}$  is a small neighbourhood of  $\theta_0$  in  $S^2$ . The smaller  $\text{WF}(\chi)$  is, the smaller we can choose  $\mathcal{V}$ .

Observe that  $\Lambda_1$  has two components

$$\Lambda_1^\pm = \{(t, t\theta; -\lambda, \lambda\theta) \in \mathbb{R}^{1+3} \times \mathbb{R}^{1+3} : \pm\lambda > 0, \theta \in \mathcal{V}, t \geq 0\}.$$

Let us give an explicit choice of two homogeneous symplectic transformations  $F^\pm$  taking  $(\Lambda_0, \Lambda_1^\pm)$  to the model case  $(\tilde{\Lambda}_0, \tilde{\Lambda}_1)$ . In a neighbourhood of  $\Lambda_1^+$  in  $T^*\mathbb{R}^{1+3} \setminus 0$  we define the map

$$F^+(x^0, x'; \xi_0, \xi') = (x^0, x' - x^0\xi'/|\xi'|; \xi_0 + |\xi'|, \xi').$$

Then  $F^+$  takes  $N^*\{0\} \setminus 0$  to itself and  $\Lambda_1^+$  to  $\tilde{\Lambda}_1$ . Indeed,

$$F^+(0, 0; \xi_0, \xi') = (0, 0; \xi_0 + |\xi'|, \xi'),$$

and for  $\lambda > 0$ ,

$$F^+(t, t\theta; -\lambda, \lambda\theta) = (t, t\theta - t\theta; -\lambda + \lambda, \lambda\theta) = (t, 0; 0, \lambda\theta).$$

Analogously, we define

$$F^-(x^0, x'; \xi_0, \xi') = (x^0, x' + x^0\xi'/|\xi'|; \xi_0 - |\xi'|, \xi')$$

taking  $N^*\{0\} \setminus 0$  to itself and  $\Lambda_1^-$  to  $\tilde{\Lambda}_1$ . We will show that  $F^-$  is symplectic, the proof for  $F^+$  being analogous.

We have

$$\frac{dF^-}{d(x^0, x'; \xi_0, \xi')} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \xi'/|\xi'| & \text{id} & 0 & B \\ 0 & 0 & 1 & -(\xi')^T/|\xi'| \\ 0 & 0 & 0 & \text{id} \end{pmatrix},$$

where  $B$  is a symmetric matrix, the precise form of which is inconsequential. Writing briefly  $dF^-$  for the above derivative and  $J$  for the symplectic form on  $T^*\mathbb{R}^{1+3}$  as a matrix, we assert that  $(dF^-)^T J dF^- = J$ .

We write  $\eta = \xi'/|\xi'|$  and

$$J = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \text{id} \\ -1 & 0 & 0 & 0 \\ 0 & -\text{id} & 0 & 0 \end{pmatrix}.$$

Then it follows that

$$\begin{aligned} (dF^-)^T J dF^- &= \begin{pmatrix} 1 & \eta^T & 0 & 0 \\ 0 & \text{id} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & B & -\eta & \text{id} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & -\eta^T \\ 0 & 0 & 0 & \text{id} \\ -1 & 0 & 0 & 0 \\ -\eta & -\text{id} & 0 & -B \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 1 & -\eta^T + \eta \\ 0 & 0 & 0 & \text{id} \\ -1 & 0 & 0 & 0 \\ \eta - \eta & -\text{id} & 0 & B - B \end{pmatrix} = J. \end{aligned}$$

Let us now write the initial condition (24) in this context. We denote by  $(\tilde{x}, \tilde{\xi}) = (\tilde{x}^0, \tilde{x}'; \tilde{\xi}_0, \tilde{\xi}')$  the local coordinates on  $T^*\mathbb{R}^{1+3}$  given by  $F^+$ , that is,

$$(\tilde{x}^0, \tilde{x}'; \tilde{\xi}_0, \tilde{\xi}') = F^+(x^0, x'; \xi_0, \xi'). \tag{81}$$

Recall that  $p = \sigma[P]$  is a section of  $T^*\mathbb{R}^{1+3} \setminus 0$ , and that  $2p(x, \xi) = -\xi_0^2 + |\xi'|^2$  in the Cartesian coordinates  $(x, \xi) = (x^0, x'; \xi_0, \xi')$ . Therefore,

$$p(\tilde{x}, \tilde{\xi}) = -(\tilde{\xi}_0 - |\tilde{\xi}'|)^2/2 + |\tilde{\xi}'|^2/2 = \tilde{\xi}_0(|\tilde{\xi}'| - \tilde{\xi}_0/2),$$

and (24) can be written

$$\begin{aligned} \sigma[u](0, \tilde{\xi}') &= \mathcal{R}(p^{-1}\sigma[f])(0, \tilde{\xi}') = \iota_0(\tilde{\xi}_0(\tilde{\xi}_0(|\tilde{\xi}'| - \tilde{\xi}_0/2))^{-1}\sigma[f](0, \tilde{\xi}))|_{\tilde{\xi}_0=0} \\ &= \iota_0|\tilde{\xi}'|^{-1}\sigma[f](0; 0, \tilde{\xi}'), \end{aligned} \tag{82}$$

where  $\iota_0 \in \mathbb{C} \setminus 0$  is a constant containing the  $2\pi$  and  $\iota$  factors. Analogously, setting

$$(\tilde{x}^0, \tilde{x}'; \tilde{\xi}_0, \tilde{\xi}') = F^-(x^0, x'; \xi_0, \xi'), \tag{83}$$

we obtain

$$p(\tilde{x}, \tilde{\xi}) = -(\tilde{\xi}_0 + |\tilde{\xi}'|)^2/2 + |\tilde{\xi}'|^2/2 = -\tilde{\xi}_0(|\tilde{\xi}'| + \tilde{\xi}_0/2),$$

and (24) reads

$$\sigma[u](0, \tilde{\xi}') = -\iota_0|\tilde{\xi}'|^{-1}\sigma[f](0; 0, \tilde{\xi}'), \tag{84}$$

with the same  $\iota_0$ .

*B.1. Trivialization of the Maslov bundle over the flowout*

It is well-known [11, Th. 3.3.4] that the Maslov bundle over a conormal bundle  $N^*K \setminus 0$  is trivial. However, even in the case of the flowout  $\Lambda_1$  in Example 1, more care is needed, since  $\Lambda_1$  fails to coincide with a conormal bundle near  $\mathcal{B} = \Lambda_1 \cap \Lambda_0$ . We will discuss the Maslov bundle in detail only in the context of Example 1, this being the most important case for the purposes of the present paper.

When the essential support  $\text{WF}(\chi)$  is small around  $\text{ccl}(0, \pm\xi^0)$ , the future flowout  $\Lambda_1$  is embedded in the following flowout to both past and future, a localized version of (28):

$$\hat{\Lambda}_1 = \{(t, t\theta; -\lambda, \lambda\theta) \in \mathbb{R}^{1+3} \times \mathbb{R}^{1+3} : \lambda \in \mathbb{R} \setminus 0, \theta \in \mathcal{V}, t \in \mathbb{R}\},$$

where  $\mathcal{V} \subset S^2$  is as in (80). We suppose that  $\mathcal{V}$  is chosen so that there is a diffeomorphism  $\Theta : B \rightarrow \mathcal{V}$  where  $B$  is an open ball in  $\mathbb{R}^2$ , with centre at the origin, and  $\Theta(0) = \theta_0$ .

Analogously to  $\Lambda_1$ , also  $\hat{\Lambda}_1$  has two components

$$\hat{\Lambda}_1^\pm = \{(t, t\theta; -\lambda, \lambda\theta) \in \mathbb{R}^{1+3} \times \mathbb{R}^{1+3} : \pm\lambda > 0, \theta \in \mathcal{V}, t \in \mathbb{R}\}.$$

We will show that  $\hat{\Lambda}_1^+$  is contractible. The same holds for  $\hat{\Lambda}_1^-$  with an analogous proof. It then follows that any vector bundle over  $\hat{\Lambda}_1 = \hat{\Lambda}_1^+ \cup \hat{\Lambda}_1^-$  is trivial [7]. In particular, the Maslov bundle over  $\Lambda_1$  is trivial. As  $\Lambda_0 = N^*\{0\} \setminus 0$ , also the Maslov bundle over  $\Lambda_0$  is trivial.

Recall that a manifold being contractible means that the identity map is smoothly homotopic to the constant map. The manifold  $\hat{\Lambda}_1^+$  is contractible, since the map

$$H : \hat{\Lambda}_1^+ \times [0, 1] \rightarrow \hat{\Lambda}_1^+, \\ H(t, t\theta; -\lambda, \lambda\theta; s) = (st, sth_1(s, \theta); -h_2(s, \lambda), h_2(s, \lambda)h_1(s, \theta)),$$

with  $h_1(s, \theta) = \Theta(s\Theta^{-1}(\theta))$  and  $h_2(s, \lambda) = s\lambda + 1 - s$ , satisfies

$$H(t, t\theta; -\lambda, \lambda\theta; 1) = (t, t\theta; -\lambda, \lambda\theta), \quad H(t, t\theta; -\lambda, \lambda\theta; 0) = (0, 0, -1, \theta_0).$$

**Appendix C. Flowout from the triple intersection**

We consider the intersection of the three cones, defined by (55), that is,

$$K_j = \{(t_j + t, x'_j + t\theta) \in \mathbb{R}^{1+3} : t > 0, \theta \in S^2\}, \quad \mathbf{j} = 1, 2, 3,$$

where, for small  $r > 0$ ,

$$(t_j, x'_j) = x_j = \gamma(-s_{\text{in}}; y, \xi_j), \quad \xi_j = (1, \sqrt{1 - r_j^2}, r_j, 0), \quad r_1 = 0, r_2 = r, r_3 = -r$$

(see (45), (46) and (52)). After a translation, we may assume without loss of generality that  $y = 0$ . Then  $t_j = -s_{\text{in}}$  for all  $\mathbf{j} = 1, 2, 3$ , and

$$x'_j = (-s_{\text{in}}\sqrt{1 - r_j^2}, -s_{\text{in}}r_j, 0), \quad \mathbf{j} = 1, 2, 3.$$

The points  $(t, x') \in \mathbb{R}^{1+3}$  in  $K_1 \cap K_2 \cap K_3$  satisfy  $|x' - x'_j|^2 = |t + s_{\text{in}}|^2$  for each  $\mathbf{j} = 1, 2, 3$ , or equivalently, writing  $x' = (x^1, x^2, x^3)$ ,

$$|x^1 + s_{\text{in}}\sqrt{1 - r_j^2}|^2 + |x^2 + s_{\text{in}}r_j|^2 + |x^3|^2 = |t + s_{\text{in}}|^2, \quad \mathbf{j} = 1, 2, 3.$$

To simplify the notation, we write  $z = x^3$ . Taking  $x^1 = x^2 = 0$ , the above three equations simplify to the single equation  $z^2 = t^2 + 2s_{\text{in}}t$ . This equation defines the filament in spacetime that acts as an artificial source, as discussed in the introduction. Solving for  $z$  gives an equation for the two moving point sources in Figure 1. In the figure, we have taken  $s_{\text{in}} = 2$  and  $r = 0.8$ .

For the study of the flowout from the filament it is more convenient to solve for  $t$  rather than for  $z$ . Near  $y = 0$  this yields

$$t = T(z) = -s_{\text{in}} + \sqrt{s_{\text{in}}^2 + z^2}.$$

As  $K_1 \cap K_2 \cap K_3$  is a smooth one-dimensional manifold near  $y$ , locally we have

$$K_1 \cap K_2 \cap K_3 = \{(T(z), 0, 0, z) : z \in \mathbb{R}\}.$$

For a covector  $(\zeta_0, \dots, \zeta_4) \in \mathbb{R}^4$  in the fibre of  $\Lambda_0 = N^*(K_1 \cap K_2 \cap K_3) \setminus 0$  at  $(T(z), 0, 0, z)$  it follows that  $\zeta_0 T'(z) + \zeta_4 = 0$  where  $T'$  is the derivative of  $T$ . Therefore

$$\Lambda_0 = \{(T(z), 0, 0, z; \zeta_0, \zeta_1, \zeta_2, -\zeta_0 T'(z)) : z \in \mathbb{R}, (\zeta_0, \zeta_1, \zeta_2) \in \mathbb{R}^3 \setminus 0\}.$$

We will proceed to compute the future flowout  $\Lambda_1$  of  $\Lambda_0 \cap \Sigma(P)$ . To simplify the notation, we write  $\zeta_0 = -\lambda$ . Then  $(-\lambda, \zeta_1, \zeta_2, \lambda T'(z))$  is lightlike if and only if

$$\lambda^2 = \zeta_1^2 + \zeta_2^2 + \lambda^2 |T'(z)|^2.$$

We write  $(\zeta_1, \zeta_2) = \mu\theta$  with  $\mu \geq 0$  and  $\theta \in S^1$ . Then  $\mu^2 = (1 - |T'(z)|^2)\lambda^2$ , and

$$\Lambda_1 \setminus \Lambda_0 = \{(T(z) + s\lambda, s\lambda\tilde{\theta}, z + s\lambda\varepsilon; -\lambda, \lambda\tilde{\theta}, \lambda\varepsilon) : \tilde{\theta} = \sqrt{1 - \varepsilon^2}\theta, \varepsilon = T'(z), \theta \in S^1, \lambda \in \mathbb{R} \setminus 0, z \in \mathbb{R}, s\lambda > 0\}.$$

Observe that  $\varepsilon = T'(z) = z/\sqrt{s_{\text{in}}^2 + z^2}$ , and this is small when we localize near  $y = 0$  as the notation suggests.

Let us show that  $\Lambda_1 \setminus \Lambda_0$  coincides with a conormal bundle. It is enough to verify that its projection to the base space  $\mathbb{R}^{1+3}$  is a smooth manifold. For our purposes it is enough to consider  $\Lambda_1 \setminus \Lambda_0$  only over the compact diamond  $\mathbb{D}$  and therefore it is enough verify that the map  $F(t, \theta, z) = (T(z) + t, t\theta, z + t\varepsilon)$  is injective and has injective differential for  $t > 0, \theta \in S^1$  and small  $|z|$ .

Suppose that

$$F(t_1, \theta_1, z_1) = F(t_2, \theta_2, z_2). \tag{85}$$



As  $t > 0$  and  $\varepsilon$  is small, the second component of (85) implies  $\theta_1 = \theta_2$  and  $t_1 \sqrt{1 - |T'(z_1)|^2} = t_2 \sqrt{1 - |T'(z_2)|^2}$ . Writing  $Z_j = \sqrt{s_{\text{in}}^2 + z_j^2}$ , and using

$$1 - |T'(z_j)|^2 = \frac{s_{\text{in}}^2 + z_j^2 - z_j^2}{s_{\text{in}}^2 + z_j^2} = \frac{s_{\text{in}}^2}{Z_j^2},$$

the latter equation reduces to  $t_1/Z_1 = t_2/Z_2$ . On the other hand, the last component of (85) gives  $z_1(1 + t_1/Z_1) = z_2(1 + t_2/Z_2)$ . As  $1 + t_1/Z_1 = 1 + t_2/Z_2 > 0$ , we get  $z_1 = z_2$ . This together with  $t_1/Z_1 = t_2/Z_2$  implies that also  $t_1 = t_2$ . We have shown that  $F$  is injective.

Observe that, writing  $Z = Z_j$  when  $z_1 = z$ , we have

$$\frac{d\varepsilon}{dz} = T''(z) = \frac{1}{Z} - \frac{z^2}{Z^3} = \frac{1 - \varepsilon^2}{Z}.$$

Therefore, letting  $\mathbb{R} \supset B \ni a \mapsto \Theta(a) \in S^1$  be local coordinates on  $S^1$ ,

$$dF = \begin{pmatrix} 1 & 0 & \varepsilon \\ \tilde{\theta} & t\sqrt{1 - \varepsilon^2} \Theta' & -t \frac{\varepsilon}{\sqrt{1 - \varepsilon^2}} \frac{1 - \varepsilon^2}{Z} \theta \\ \varepsilon & 0 & 1 + t \frac{1 - \varepsilon^2}{Z} \end{pmatrix}.$$

When  $z = 0$ , this reduces to

$$dF = \begin{pmatrix} 1 & 0 & 0 \\ \tilde{\theta} & t\Theta' & 0 \\ 0 & 0 & 1 + \frac{t}{Z} \end{pmatrix},$$

which is invertible since  $t > 0$ ,  $Z > 0$  and  $\Theta' \neq 0$ . The same holds when  $|z|$  is small enough, and we have shown that  $\Lambda_1 \setminus \Lambda_0$  coincides with a conormal bundle. Let us also remark that an argument similar to that in Section B.1 shows that the Maslov bundle over  $\Lambda_1$  is trivial.

*Acknowledgments.* LO thanks Allan Greenleaf, Alexander Strohmaier and Gunther Uhlmann for discussions on microlocal analysis.

*Funding.* ML was supported by Academy of Finland grants 320113 and 312119. LO was supported by EPSRC grants EP/P01593X/1 and EP/R002207/1, XC and GPP were supported by EPSRC grant EP/R001898/1, and XC was supported by NSFC grant 11701094. GPP thanks the University of Washington for hospitality while this work was in progress and the Leverhulme trust for financial support.

**References**

[1] Alinhac, S.: Non-unicité du problème de Cauchy. *Ann. of Math. (2)* **117**, 77–108 (1983) [Zbl 0516.35018](#) [MR 683803](#)

- [2] Anderson, M., Katsuda, A., Kurylev, Y., Lassas, M., Taylor, M.: Boundary regularity for the Ricci equation, geometric convergence, and Gel'fand's inverse boundary problem. *Invent. Math.* **158**, 261–321 (2004) Zbl [1177.35245](#) MR [2096795](#)
- [3] Belishev, M. I., Kurylev, Y. V.: To the reconstruction of a Riemannian manifold via its spectral data (BC-method). *Comm. Partial Differential Equations* **17**, 767–804 (1992) Zbl [0812.58094](#) MR [1177292](#)
- [4] Bernal, A. N., Sánchez, M.: Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes. *Comm. Math. Phys.* **257**, 43–50 (2005) Zbl [1081.53059](#) MR [2163568](#)
- [5] Bernal, A. N., Sánchez, M.: Globally hyperbolic spacetimes can be defined as ‘causal’ instead of ‘strongly causal’. *Classical Quantum Gravity* **24**, 745–749 (2007) Zbl [1177.53065](#) MR [2294243](#)
- [6] Bony, J.-M.: Second microlocalization and propagation of singularities for semilinear hyperbolic equations. In: *Hyperbolic Equations and Related Topics (Katata/Kyoto, 1984)*, Academic Press, Boston, MA, 11–49 (1986) Zbl [0669.35073](#) MR [925240](#)
- [7] Bott, R., Tu, L. W.: *Differential Forms in Algebraic Topology*. Grad. Texts in Math. 82, Springer, New York (1982) Zbl [0496.55001](#) MR [658304](#)
- [8] Cottingham, W. N., Greenwood, D. A.: *An Introduction to the Standard Model of Particle Physics*. 2nd ed., Cambridge Univ. Press, Cambridge (2007) Zbl [1126.81002](#) MR [2358902](#)
- [9] Duistermaat, J. J., Hörmander, L.: Fourier integral operators. II. *Acta Math.* **128**, 183–269 (1972) Zbl [0232.47055](#) MR [388464](#)
- [10] Greenleaf, A., Uhlmann, G.: Recovering singularities of a potential from singularities of scattering data. *Comm. Math. Phys.* **157**, 549–572 (1993) Zbl [0790.35112](#) MR [1243710](#)
- [11] Hörmander, L.: Fourier integral operators. I. *Acta Math.* **127**, 79–183 (1971) Zbl [0212.46601](#) MR [388463](#)
- [12] Hörmander, L.: *The Analysis of Linear Partial Differential Operators*. III. Grundlehren Math. Wiss. 274, Springer, Berlin (1985) Zbl [0601.35001](#) MR [781536](#)
- [13] Hörmander, L.: *The Analysis of Linear Partial Differential Operators*. IV. Grundlehren Math. Wiss. 275, Springer, Berlin (1985) Zbl [0612.35001](#) MR [781537](#)
- [14] Hörmander, L.: *The Analysis of Linear Partial Differential Operators*. I. 2nd ed., Springer, Berlin (1990) Zbl [0712.35001](#) MR [1065136](#)
- [15] Hughes, T. J. R., Kato, T., Marsden, J. E.: Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity. *Arch. Ration. Mech. Anal.* **63**, 273–294 (1977) (1976) Zbl [0361.35046](#) MR [420024](#)
- [16] Jakobson, D., Strohmaier, A.: High energy limits of Laplace-type and Dirac-type eigenfunctions and frame flows. *Comm. Math. Phys.* **270**, 813–833 (2007) Zbl [1114.58014](#) MR [2276467](#)
- [17] Katchalov, A., Kurylev, Y., Lassas, M.: *Inverse Boundary Spectral Problems*. Chapman & Hall/CRC Monogr. Surveys Pure Appl. Math. 123, Chapman & Hall/CRC, Boca Raton, FL (2001) Zbl [1037.35098](#) MR [1889089](#)
- [18] Kato, T.: Quasi-linear equations of evolution, with applications to partial differential equations. In: *Spectral Theory and Differential Equations (Dundee, 1974; dedicated to Konrad Jörgens)*, Lecture Notes in Math. 448, Springer, 25–70 (1975) Zbl [0315.35077](#) MR [0407477](#)
- [19] Kian, Y., Kurylev, Y., Lassas, M., Oksanen, L.: Hyperbolic inverse problem with data on disjoint sets. *arXiv:1602.0362v2* (2018)

- [20] Kurylev, Y., Lassas, M., Uhlmann, G.: Inverse problems for Lorentzian manifolds and non-linear hyperbolic equations. *Invent. Math.* **212**, 781–857 (2018) Zbl [1396.35074](#) MR [3802298](#)
- [21] Kurylev, Y., Oksanen, L., Paternain, G. P.: Inverse problems for the connection Laplacian. *J. Differential Geom.* **110**, 457–494 (2018) Zbl [1415.53024](#) MR [3880231](#)
- [22] Kurylev, Y., Lassas, M., Oksanen, L., Uhlmann, G.: Inverse problem for Einstein-scalar field equations. arXiv:[1406.4776v2](#) (2018)
- [23] Kurylev, Y., Lassas, M., Uhlmann, G.: Determination of structures in the space-time from local measurements: a detailed exposition. arXiv:[1305.1739v2](#) (2013)
- [24] Kurylev, Y., Lassas, M., Uhlmann, G.: Inverse problems in spacetime I: Inverse problems for Einstein equations – Extended preprint version. arXiv:[1405.4503](#) (2014)
- [25] Lassas, M.: Inverse problems for linear and non-linear hyperbolic equations. In: Proc. International Congress of Mathematicians (Rio de Janeiro, 2018), Vol. IV, World Sci., Hackensack, NJ, 3751–3771 (2018) Zbl [1447.35006](#) MR [3966550](#)
- [26] Lassas, M., Oksanen, L.: Inverse problem for the Riemannian wave equation with Dirichlet data and Neumann data on disjoint sets. *Duke Math. J.* **163**, 1071–1103 (2014) Zbl [1375.35634](#) MR [3192525](#)
- [27] Lassas, M., Uhlmann, G., Wang, Y.: Inverse problems for semilinear wave equations on Lorentzian manifolds. *Comm. Math. Phys.* **360**, 555–609 (2018) Zbl [1409.58019](#) MR [3800791](#)
- [28] Lassas, M., Uhlmann, G., Wang, Y.: Determination of vacuum space-times from the Einstein–Maxwell equations. arXiv: [1703.10704](#) (2017)
- [29] Loomis, L. H., Sternberg, S.: *Advanced calculus*. Jones and Bartlett, Boston, MA (1990) Zbl [0782.46041](#) MR [1140004](#)
- [30] Melrose, R., Ritter, N.: Interaction of nonlinear progressing waves for semilinear wave equations. *Ann. of Math. (2)* **121**, 187–213 (1985) Zbl [0575.35063](#) MR [782559](#)
- [31] Melrose, R. B., Uhlmann, G. A.: Lagrangian intersection and the Cauchy problem. *Comm. Pure Appl. Math.* **32**, 483–519 (1979) Zbl [0396.58006](#) MR [528633](#)
- [32] Rauch, J., Reed, M. C.: Singularities produced by the nonlinear interaction of three progressing waves; examples. *Comm. Partial Differential Equations* **7**, 1117–1133 (1982) Zbl [0502.35060](#) MR [673827](#)
- [33] Sá Barreto, A., Wang, Y.: Singularities generated by the triple interaction of semilinear conormal waves. *Anal. PDE* **14**, 135–170 (2021) Zbl [07365667](#) MR [4229201](#)
- [34] Sachs, R. K., Wu, H. H.: *General Relativity for Mathematicians*. Grad. Texts in Math. 48, Springer, New York (1977) Zbl [0373.53001](#) MR [0503498](#)
- [35] Salazar, R.: Determination of time-dependent coefficients for a hyperbolic inverse problem. *Inverse Problems* **29**, art. 095015, 17 pp. (2013) Zbl [1292.35336](#) MR [3104931](#)
- [36] Tataru, D.: Unique continuation for solutions to PDE’s; between Hörmander’s theorem and Holmgren’s theorem. *Comm. Partial Differential Equations* **20**, 855–884 (1995) Zbl [0846.35021](#) MR [1326909](#)
- [37] Uhlenbeck, K. K.: Variational problems for gauge fields. In: Proc. International Congress of Mathematicians (Warszawa, 1983), Vol. 1, PWN, Warszawa, 585–591 (1984) Zbl [0562.53059](#) MR [0804715](#)