INVERSE PROBLEM FOR THE BOLTZMANN EQUATION IN COSMOLOGY

YIRAN WANG

Abstract. We study the inverse problem of recovering primordial perturbations from anisotropies of Cosmic Microwave Background (CMB) using the kinetic model. Mathematically, the problem in concern is the inverse source problem for the linear Boltzmann equation with measurements on some Cauchy surface. We obtain two stable determination results for generic absorption coefficients and scattering kernels.

Contents

1. Introduction

Consider the source problem for the linear Boltzmann equation (or non-stationary transport equation) on $\mathcal{M} = (0, T) \times \mathbb{R}^3, T > 0$:

$$
\begin{aligned}
&\partial_t u(t, x, \theta) + \theta \cdot \nabla_x u(t, x, \theta) + \sigma(t, x, \theta)u(t, x, \theta) \\
&= \int_{\mathbb{S}^2} k(t, x, \theta, \theta')u(t, x, \theta')d\theta' + f(t, x),
\end{aligned} \tag{1}
$$

where $t \in (0, T), x \in \mathbb{R}^3, \theta \in \mathbb{S}^2$. Here, σ is the absorption coefficient, k is the scattering kernel and f is the source term. We consider the zero initial condition

$$
u(0, x, \theta) = 0. \tag{2}
$$

Date: August 5, 2021.

In this work, we study the inverse problem of determining the source term f from the measurement of u at $t = T > 0$

$$u(T, x, \theta) = u_T(x, \theta). \tag{3}$$

The inverse problem for (1) and its stationary version has a rich history, see Section 7.4 of [10]. Both the determination of σ, k and the source term f have been investigated. In particular, there are lots of interest due to its application in optical imaging, see review papers [1, 15]. Our perspective is somewhat different from previous works as our motivation comes from inverse problems in cosmology. We are interested in the determination of primordial gravitational perturbations from the anisotropies of the Cosmic Microwave Background (CMB), see [11]. The physics background will be discussed in Section 2. The pure transport regime (namely without σ, K in (1)) serves as a good model for the standard universe after the decoupling time or the "surface of last scattering". This was studied by Vasy and the author in [20] using the light ray transform. Before the decoupling time, photon interactions cannot be ignored and a kinetic model based on the Boltzmann equation is appropriate. As is well-known in cosmology literatures e.g. [4, 6] (see also Section 2), the linearization of the Boltzmann equation on a Friedman-Lemaîte-Robertson-Walker (FLRW) universe with respect to small metric perturbations naturally leads to a source problem for the Boltzmann equation as (1) in which the source term is related to the metric perturbation. In fact, similar problems can also be considered in the relativistic kinetic theory such as the Nordström-Vlasov system (see Section 4 of [7]) on the linearization level, or even the Einstein-Boltzmann equations.

In this work, we obtain two results on the stable determination of the source term in (1). To state our first result, we need to introduce some notations. Consider the Minkowski spacetime $(\mathbb{R}^{3+1}, g)$ with signature $(-, +, +, +)$. On the dual space $\mathbb{R}^{3+1}_{(\tau,\xi)}$, we let $\Gamma^{tm}_{\pm} = \{(\tau, \xi) \in \mathbb{R}^{3+1} : \tau^2 > |\xi|^2, \pm\tau > 0\}$ be the set of future/past pointing time-like vectors. Let $\Gamma^{sp} = \{(\tau, \xi) \in \mathbb{R}^{3+1} : \tau^2 < |\xi|^2\}$ be the set of space-like vectors. Finally, let $\Gamma^{lt}_{\pm} = \{(\tau, \xi) \in \mathbb{R}^{3+1} : \tau^2 = |\xi|^2, \pm\xi_0 > 0\}$ be the set of future/past pointing light like vectors. We also let $\Gamma^{lt} = \Gamma^{lt}_+ \cup \Gamma^{lt}_-$. Let ϕ be the characteristic function of Γ^{sp}. We define $\phi(D)$ to be a Fourier multiplier

$$\phi(D)f = \mathcal{F}^{-1}(\phi \mathcal{F} f), \quad f \in L^2(\mathbb{R}^4)$$

where $\mathcal{F}, \mathcal{F}^{-1}$ denote the Fourier and inverse Fourier transform in t, x variables. Throughout the work, we let $\mathcal{V} = (0, T) \times \Omega$ where Ω is a relatively compact set of $\mathbb{R}^3$ (so that $\mathcal{V}$ is a relatively compact set of $\overline{\mathcal{M}}$). *Throughout the paper, we assume that σ, k and f are supported in $\mathcal{V}$.* Our first result is

Theorem 1.1. *Let $\sigma \in C^6$ be independent of the x and θ variable. There exists a dense subset $\mathcal{U}$ of $C^6(\mathcal{V} \times \mathbb{S}^2 \times \mathbb{S}^2)$ such that the following is true. Consider the source problem (1) and (2) with $k \in \mathcal{U}$ and $f \in H^2_{\rm comp}(\mathcal{M})$. Then f is uniquely determined by u_T in (3). Moreover, we have the following* *stability estimate*

$$\|\phi(D)f\|_{H^2(\mathcal{M})} \leq C\|u_T\|_{H^{5/2}(\mathbb{R}^3\times\mathbb{S}^2)} \tag{4}$$

for some $C > 0$ depending on σ, k.

The type of stability estimate (4) seems to be new and it is particularly important for our analysis. In fact, we will use the stability estimate to recover $\phi(D)f$ then use the analyticity of the Fourier transform of f to prove the uniqueness.

Next, for the CMB inverse problem, the metric perturbations that describe the evolution of the universe are not arbitrary. In fact, they are solutions of the linearized Einstein equations, see Section 2. This leads us to study the inverse problem of (1) when the source f is a solution of certain wave equations. For $s \geq 0$, we denote $\mathcal{M}_s = \{s\} \times \mathbb{R}^n$. Consider

$$P(z, \partial) = \Box + \sum_{j=0}^{n} A_j(z)\partial_j + B(z) \tag{5}$$

where A_j, B are real or complex valued smooth functions in z. Consider the Cauchy problem

$$\begin{aligned} P(z,\partial)f(z) = 0, \text{ on } \mathcal{M} \\ u = f_1, \quad \partial_t u = f_2 \text{ on } \mathcal{M}_0 \end{aligned} \tag{6}$$

Our second result is the stable determination of f from u_T. Below, we take $\mathcal{V}$ sufficiently large so that the solution of (6) with initial data supported in a fixe compact set $\mathcal{X}$ of $\mathbb{R}^3$ is contained in $\mathcal{V}$.

Theorem 1.2. *Let f be the solution of* (6) *on $\mathcal{M}$ with Cauchy data $f_1 \in H^2(\mathcal{M}_0), f_2 \in H^1(\mathcal{M}_0)$ supported in a compact set $\mathcal{X}$ of $\mathcal{M}_0$ such that f is supported in $\mathcal{V}$. Suppose that the coefficients $A_j(z)$ in* (5) *are real valued smooth functions. Let u be the solution of* (1)*,* (2) *with source f.*

Then there exists an open dense set $\mathcal{U}$ of $C^\infty(\mathcal{V}\times\mathbb{S}^2)\times C^6(\mathcal{V}\times\mathbb{S}^2\times\mathbb{S}^2)$ such that for $(\sigma,k)\in\mathcal{U}$, f_1, f_2 is uniquely determined by u_T and there exists $C > 0$ such that

$$\|f\|_{H^2(\mathcal{M})} \leq C\|(f_1,f_2)\|_{H^2(\mathbb{R}^3)\times H^1(\mathbb{R}^3)} \leq C\|u_T\|_{H^{5/2}(\mathbb{R}^3\times\mathbb{S}^2)} \tag{7}$$

In Section 8, we will prove a stronger version of the theorem to include certain pseudo-differential operators which are motivated by the CMB inverse problem, see Section 2. For $\sigma = k = 0$, Theorem 1.2 was proved in [20] and further generalized in [22].

In proving both Theorem 1.1 and 1.2, we follow the spirit in Stefanov and Uhlmann [17] for the stationary transport equation to treat the map $f \to u_T$ as a perturbation of the light ray transform on the Minkowski spacetime. The difficulty is that, unlike the geodesic ray transform in the Riemmanian setting, the normal operator of the light ray transform is not an elliptic pseudo-differential operator. In fact, the Schwartz kernel belongs to the class of paired

Lagrangian distributions, see [21]. The key of our approach is to restore the ellipticity by using either $\phi(D)$ or the parametrix of the Cauchy problem.

We end the introduction with a few remarks. First, our results should hold for general dimensions, however we study $\mathbb{R}^{3+1}$ for its physical relevance. Second, as we assume that f is compactly supported in $[0,T]\times\Omega$, one can consider the problem with measurements on the lateral boundary $[0,T]\times\partial\Omega$ using the method we develop here. The problem then is the time-dependent version of the inverse source problem studied in [17]. Finally, the stability estimates suggest that our results can be generalized via perturbation arguments to other scenarios such as small metric perturbations of the Minkowski spacetime as in [20], small perturbations of σ for Theorem 1.1 and nonlinear perturbations in the Boltzmann equation.

2. The kinetic theory for CMB

In this section, we discuss the inverse problem of determining primordial perturbations from the anisotropies of CMB. Our goal is to show how the source problem for the Boltzmann equation naturally appears and how the source term is connected to the metric perturbations. We will consider a simple setup, emphasizing more on the mathematical structure of the problem. The perturbation theory for CMB anisotropies has been well-developed in cosmology literatures, which can be found in [4, 6] for instance.

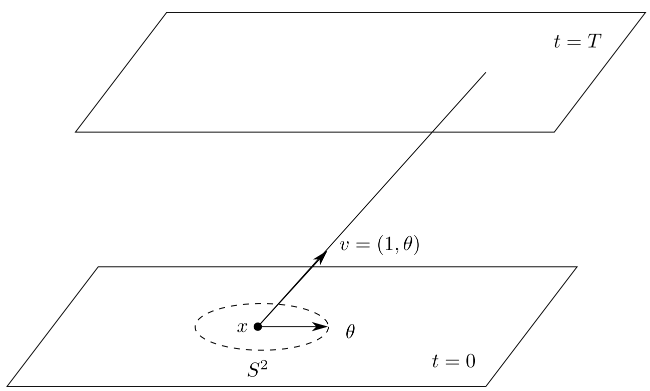

Figure 1. Parametrization of the light rays

Consider the FLRW spacetime $(\mathcal{M}, g)$ as the Universe model, where $\mathcal{M} = [0,\infty)\times\mathbb{R}^3$ and $g = -dt^2+a^2(t,x)dx^2$ with $a > 0$. However, $(\mathcal{M}, g)$ is conformal to the Minkowski spacetime and the conformal transformation does not change much of the analysis. So we work with the Minkowski spacetime $(\mathcal{M}, g)$ below (by simply taking $a = 1$). Let $\Phi, \Psi \in C^\infty(\mathcal{M})$. For $\epsilon > 0$ small, we consider a smooth family of Lorentzian metrics on $\mathcal{M}$

$$g_\epsilon = (-1+\epsilon\Phi)dt^2 + (1+\epsilon\Psi)dx^2 + \epsilon^2 h_\epsilon \tag{8}$$

where h_ϵ is a family of symmetric two tensors on $\mathcal{M}$ smooth for $\epsilon \in [0, \epsilon_0), \epsilon_0 > 0$. Later, we also use $z = (z_0, \cdots, z_3) = (t, x_1, x_2, x_3)$ for local coordinates. Note that g_0 is the Minkowski metric, and we regard $(\mathcal{M}, g_\epsilon)$ as the perturbation of the Minkowski spacetime.

Consider light-like geodesics $\gamma_\epsilon(s), s \geq 0$ on $(\mathcal{M}, g_\epsilon)$ originating from $\mathcal{M}_0 = \{0\} \times \mathbb{R}^3$ which we think of as photon trajectories. They satisfy the geodesic equation

$$\ddot{\gamma}_\epsilon^k(s) + \Gamma^k_{\epsilon,ij}(s)\dot{\gamma}_\epsilon^i(s)\dot{\gamma}_\epsilon^j(s) = 0 \tag{9}$$

with initial conditions

$$\gamma_\epsilon(0) = \tilde{z}, \quad \dot{\gamma}_\epsilon(0) = \tilde{\zeta}.$$

Here $\Gamma^k_{\epsilon,ij}(s)$ denotes the Christoffel symbols for g_ϵ along $\gamma_\epsilon(s)$. Let $p^i_\epsilon(s) = \dot{\gamma}^i_\epsilon(s), i = 1, 2, 3$ be the momentum and $p^0_\epsilon(s) = \dot{\gamma}^0_\epsilon(s)$ be the energy of the photon. In particular, $p_\epsilon = (p^i_\epsilon)^3_{i=0}$ is a vector field along γ_ϵ. As we consider light-like geodesics for (massless) photons, p_ϵ are (future pointing) null vectors, namely $g_\epsilon(p_\epsilon, p_\epsilon) = 0$ along γ_ϵ. It is convenient to denote $v = p^0_\epsilon > 0$ the energy and $\theta^i = p^i_\epsilon/p^0_\epsilon, i = 1, 2, 3$. In particular, we have $p_\epsilon = v(1, \theta)$.

Now let f_ϵ be the photon distribution function which is a function of z, p variables where $z \in \mathbb{R}^{3+1}$ and p is on the mass shell

$$\Sigma_z = \{p \in T_z\mathbb{R}^{3+1} : g_\epsilon(p, p) = 0\}.$$

We assume that f_ϵ satisfies the linear Boltzmann equation, see [6, Section 4.5]. This means that along γ_ϵ

$$\frac{d}{ds} f_\epsilon(\gamma_\epsilon(s), p_\epsilon(s)) = C[f_\epsilon] \tag{10}$$

where $C[f]$ denotes the interaction term

$$C[f] = -\sigma(z)f(z, p) + \int k(z, \theta, \theta')f(z, v(1, \theta'))d\theta' \tag{11}$$

where σ denotes absorption coefficients, k is the scattering kernel and the integration is over $\{\theta : v(1, \theta) \in \Sigma_z \text{ for } v > 0\}$. The terms in (11) accounts for photon interactions in Thomson scattering for example. We get from (10) and (11) the equation

$$\sum_{i=0}^{3} \frac{\partial f_\epsilon}{\partial z^i}(z, p)\frac{\partial \gamma^i_\epsilon}{\partial s} + \frac{\partial f_\epsilon}{\partial p}(z, p)\frac{\partial p_\epsilon}{\partial s} = -\sigma(z)f_\epsilon(z, p) + \int k(z, \theta, \theta')f_\epsilon(z, v(1, \theta'))d\theta' \tag{12}$$

Now we consider f_ϵ as a perturbation of some background distribution with an expansion

$$f_\epsilon(z, p) = f_0(v) + \epsilon f_1(z, v, \theta) + O(\epsilon^2) \tag{13}$$

Here, f_0 is the background photon distribution. When modeling the cosmic microwave background, one often assumes that f_0 satisfies the Planck distribution

$$f_0(v) = (e^{v/\mathcal{T}} + 1)^{-1}$$

see page 149 of [6]. Here, $\mathcal{T} > 0$ be the background temperature of the universe. f_1 in (13) is the first order perturbation term and θ is taken over $\mathbb{S}^2$. In particular, $(1, \theta)$ is a future pointing light-like vector for the background Minkowski metric g_0.

We find the ϵ derivative of the equation (12) at $\epsilon = 0$.

$$
\begin{aligned}
&v\frac{\partial f_1}{\partial t}(z, v, \theta) + v\sum_{j=1}^{3}\theta_j\frac{\partial f_1}{\partial z^j}(z, v, \theta) + \frac{\partial f_1}{\partial p_0^j}(z, v, \theta)\frac{\partial p_0^j}{\partial s} \\
&+\frac{\partial f_0}{\partial v}(v)\partial_\epsilon\frac{\partial p_\epsilon^0}{\partial s}|_{\epsilon=0} = -\sigma(z)f_1(z, v, \theta) + \int_{\mathbb{S}^2} k(z, \theta, \theta')f_1(z, v, \theta')d\theta'
\end{aligned} \tag{14}
$$

Here, $p_0(s) = \dot\gamma_0(s)$ is the vector filed along the geodesic γ_0 for the background metric. We observe that $\partial_s p_0^j(s) = 0$ for $j = 0, 1, 2, 3$ which can be seen from the geodesic equation on $(\mathbb{R}^{3+1}, g_0)$ and the fact that g_0 is a constant metric. It remains to find $\partial_\epsilon\frac{\partial p_\epsilon^0}{\partial s}|_{\epsilon=0}$ in (2). We use the geodesic equation (9) on $(\mathbb{R}^{3+1}, g_\epsilon)$

$$
\dot p_\epsilon^0(s) + \Gamma^0_{\epsilon,ij}(\gamma_\epsilon(s))p_\epsilon^i(s)p_\epsilon^j(s) = 0
$$

thus

$$
\partial_\epsilon(\frac{\partial p_\epsilon^0}{\partial s})|_{\epsilon=0} = \partial_\epsilon\Gamma^0_{\epsilon,ij}|_{\epsilon=0}p_0^ip_0^j
$$

In the calculation, we used the fact that the Christoffel symbols for the Minkowski spacetime all vanish. To find the linearization of the Christoffel symbol, recall that

$$
\Gamma^i_{jk} = \frac{1}{2}g^{i\mu}(\frac{\partial g_{\mu k}}{\partial z^j} - \frac{\partial g_{jk}}{\partial z^\mu} + \frac{\partial g_{j\mu}}{\partial z^k})
$$

Therefore,

$$
\partial_\epsilon\Gamma^0_{\epsilon,ij}|_{\epsilon=0} = 0 \text{ if } i \neq j, \text{ and } i, j \neq 0
$$

$$
\partial_\epsilon\Gamma^0_{\epsilon,jj}|_{\epsilon=0} = +\frac{1}{2}\frac{\partial\Psi}{\partial z^0} \text{ if } j \neq 0 \text{ and } \partial_\epsilon\Gamma^0_{\epsilon,0j}|_{\epsilon=0} = -\frac{1}{2}\frac{\partial\Phi}{\partial z^j}
$$

We deduce that

$$
\partial_\epsilon\frac{\partial p_\epsilon^0}{\partial s}|_{\epsilon=0} = \frac{1}{2}\frac{\partial\Psi}{\partial z^0}p_0^jp_0^j - \frac{1}{2}\frac{\partial\Phi}{\partial z^j}p_0^0p_0^j
$$

Using this in (14) we get

$$
\begin{aligned}
&v\frac{\partial f_1}{\partial t}(z, v, \theta) + v\theta^j\frac{\partial f_1}{\partial x^j}(z, v, \theta) + \sigma(z)f_1(z, v, \theta) \\
&-\int_{\mathbb{S}^2} k(z, \theta, \theta')f_1(z, v, \theta')d\theta' = -\frac{\partial f_0}{\partial v}(v)v(\frac{1}{2}\frac{\partial\Psi}{\partial t}v - \frac{1}{2}\frac{\partial\Phi}{\partial z^j}v\theta^j)
\end{aligned} \tag{15}
$$

At this point, we will take $v > 0$ to be a fixed constant and ignore it in f_1. We get from (15) that

$$
\begin{aligned}
&\frac{\partial f_1}{\partial t}(z, \theta) + \theta^j\frac{\partial f_1}{\partial z^j}(z, \theta) + \sigma(z)f_1(z, \theta) - \int_{\mathbb{S}^2} k(z, \theta, \theta')f_1(z, \theta')d\theta' \\
&= C(\frac{1}{2}\frac{\partial\Psi}{\partial t} - \frac{1}{2}\frac{\partial\Phi}{\partial z^j}\theta^j)
\end{aligned} \tag{16}
$$

where C is a non-zero constant and σ, k are changed by a scalar factor. This is essentially the Boltzmann equation we considered in the introduction, and the inverse problem is to determine Φ, Ψ from the observation of f_1 at $t = T$. We remark that in cosmology literatures, one integrates (15) in v and obtain an equation for a quantity independent of v. The quantity is related to the temperature perturbation or the redshift of the CMB, see Section 4.5.1 of [6] for instance. However, the mathematical structure of the so-derived equation is identical to (16). Actually, it is more convenient to change f_1 in (16) to $\tilde{f}_1 = f_1 + \frac{1}{2}C\Phi$. Then we find from (16) that

$$\begin{aligned}\frac{\partial \tilde{f}_1}{\partial t}(z,\theta) + \theta^j \frac{\partial \tilde{f}_1}{\partial z^j}(z,\theta) + \sigma(z)\tilde{f}_1(z,\theta) - \int_{\mathbb{S}^2} k(z,\theta,\theta')\tilde{f}_1(z,\theta')d\theta' \\ = C(\frac{1}{2}\frac{\partial \Psi}{\partial t} + \frac{1}{2}\frac{\partial \Phi}{\partial t} + B(z)\Phi)\end{aligned} \tag{17}$$

where $B(z)$ is a smooth function of z related to σ, k. When $\sigma = k = 0$, $B(z) = 0$. The difference of f_1 and $\tilde{f}_1$ is independent of the direction and is in fact not measurable from CMB anisotropies, see [6, Chapter 4].

Finally, let's consider the metric perturbations Φ, Ψ in (8). When modeling the evolution of the universe, one assumes that g_ϵ satisfies the Einstein equations with matters. In case of scalar fields matter, the linearized term Φ, Ψ are known to be equal and satisfy the Bardeen's equation, which is of the form

$$\partial_t^2 \Psi + A_0(t)\partial_t \Psi + \Delta \Psi + B_0(t)\Psi = 0. \tag{18}$$

See equation (6.48) of [13]. Here, $A_0(t), B_0(t)$ are smooth functions and Δ is the Laplacian on $\mathbb{R}^3$. The inverse problem now is to determine Ψ in $\mathcal{M}$ satisfying (18) with measurement of $\tilde{f}_1$ of (17) at $t = T$. Note that this problem fits Theorem 1.2 except that the source term in (17) involves an extra differential operator. This will be treated in the stronger version Theorem 8.1 of Theorem 1.2 in Section 8.

3. Solvability of the direct problem

We consider the solvability of the source problem for the linear Boltzmann equation. We will follow the approach in [17] to give a proof based on analytic Fredholm theory. Compared with the results for the stationary transport equation in [17], we need higher regularity requirements for σ, k and f. We assume $n \geq 2$ in this section.

Theorem 3.1. *For $\sigma \in C^5(\mathcal{V})$, there exists an open and dense subset $\mathscr{U}$ of $C^5(\mathcal{V} \times \mathbb{S}^{n-1} \times \mathbb{S}^{n-1})$ such that for any $k \in \mathscr{U}$ and $f \in H^2_{\text{comp}}(\mathcal{M})$, the equation* (1) *with initial condition* (2) *has a unique solution $u \in H^2(\mathcal{M} \times \mathbb{S}^{n-1})$.*

We remark that the theorem can also be stated for (σ, k) in a open dense subset $\mathscr{U}$ of $C^5(\mathcal{V}) \times C^5(\mathcal{V} \times \mathbb{S}^{n-1} \times \mathbb{S}^{n-1})$.

For the proof, we let

$$T_0 = \partial_t + \theta \cdot \nabla_x, \quad T_1 = T_0 + \sigma, \quad T = T_1 - K$$

where σ is regarded as the multiplication operator and K is the integral operator in (1). For $k = 0$, the equation $T_1 u = f$ with $u = 0$ at $t = 0$ can be solved explicitly. For $\theta \in \mathbb{S}^{n-1}, t > 0, x \in \mathbb{R}^n$, consider $u(t, x, \theta) = u(t, x + t\theta)$ which satisifes

$$\frac{d}{dt} u(t, x + t\theta) + \sigma(t, x + t\theta) u(t, x + t\theta) = f(t, x + t\theta) \tag{19}$$

An integrating factor is $E(t, x, \theta) = e^{\int_0^t \sigma(s, x+s\theta) ds}$. We solve (19) that

$$\begin{aligned} u(t, x + t\theta) &= e^{-\int_0^t \sigma(s,x+s\theta)ds} \int_0^t e^{\int_0^s \sigma(\tilde{s},x+\tilde{s}\theta)d\tilde{s}} f(s, x + s\theta) ds \\ &= \int_0^t e^{-\int_s^t \sigma(\tilde{s},x+\tilde{s}\theta)d\tilde{s}} f(s, x + s\theta) ds \end{aligned}$$

Thus we can write T_1^{-1} as

$$\begin{aligned} T_1^{-1} f(t, x, \theta) &= \int_0^t \kappa(t, x, s, \theta) f(s, x + s\theta) ds, \\ \text{with } \kappa(t, x, s, \theta) &= e^{-\int_s^t \sigma(\tilde{s},x+\tilde{s}\theta)d\tilde{s}} \end{aligned} \tag{20}$$

Next, for $Tu = (T_1 - K)u = f$, we apply T_1^{-1} and get $(\mathrm{Id} - T_1^{-1} K)u = T_1^{-1} f$. The main part of the proof is to show that $\mathrm{Id} - T_1^{-1} K$ is invertible for suitable k so that

$$u = (\mathrm{Id} - T_1^{-1} K)^{-1} T_1^{-1} f \tag{21}$$

Notice that this can be written as

$$u = T_1^{-1} (\mathrm{Id} - K T_1^{-1})^{-1} f. \tag{22}$$

We will show that $\mathrm{Id} - K T_1^{-1}$ is invertible. As in [17], we introduce

$$A = (\mathrm{Id} - (K T_1^{-1})^2)^{-1}$$

and write

$$(\mathrm{Id} - K T_1^{-1})^{-1} = (\mathrm{Id} + K T_1^{-1}) A.$$

We will show that $(K T_1^{-1})^2$ is compact and apply analytic Fredholm theory to conclude that A is invertible.

For the proof, we will need the following proposition and its variant about singular operators.

Proposition 3.2 (Proposition 3.4 of [17]). *Let A be the operator*

$$A f(x) = \int \frac{\alpha(x, y, |x - y|, \frac{x-y}{|x-y|})}{|x - y|^{n-1}} f(y) dy$$

with $\alpha(x, y, r, \theta)$ compactly supported in $x, y \in \mathbb{R}^n$.

(i) *If $\alpha \in C^2$, then $A : L^2(\mathbb{R}^n) \to H^1(\mathbb{R}^n)$ is continuous with a norm not exceeding $C\|\alpha\|_{C^2}$.*

(ii) Let $\alpha(x, y, r, \theta) = \alpha'(x, y, r, \theta)\phi(\theta)$*, then*

$$\|A\|_{L^2 \to H^1} \leq C\|\alpha'\|_{C^2}\|\phi\|_{H^1(\mathbb{S}^{n-1})}$$

We remark that the constant C is independent of $\|\alpha\|_{C^2}$ but depends on the support of α. It can be made uniform if $\alpha(x, y, \cdot, \cdot)$ is supported in a fixed compact set in x, y variables. The proposition can be slightly improved for $H^m, m \geq 0$ functions.

Proposition 3.3. *Consider operator* A *in Proposition 3.2.*

(i) If $\alpha \in C^{m+2}, m = 0, 1, 2, \cdots,$ *then* $A : H^m(\mathbb{R}^n) \to H^{m+1}(\mathbb{R}^n)$ *is continuous with a norm not exceeding* $C\|\alpha\|_{\tilde{C}^{m+2}}$ *where*

$$\|\alpha\|_{\tilde{C}^{m+2}} = \sup \sum_{|\gamma|=2, |\beta|=m} |\partial^\beta_{x,y}\partial^\gamma_{x,y,r,\theta}\alpha|$$

(ii) Let $\alpha(x, y, r, \theta) = \alpha'(x, y, r, \theta)\phi(\theta)$*, then*

$$\|A\|_{H^m \to H^{m+1}} \leq C\|\alpha'\|_{\tilde{C}^{m+2}}\|\phi\|_{H^1(\mathbb{S}^{n-1})}$$

Proof. Assume that $f \in H^m_{\text{comp}}(\mathcal{M})$. We prove for $m = 1$ and the other cases are similar. Using polar coordinate, we write

$$Af(x) = \int \alpha(x, x + r\theta, r, \theta)f(x + r\theta)drd\theta$$

For $i = 1, 2, \cdots, n$, we get

$$\partial_{x_i}Af(x) = \int \partial_{x^i}\alpha(x, x + r\theta, r, \theta)f(x + r\theta) + \alpha(x, x + r\theta, r, \theta)\partial_{x^i}f(x + r\theta)drd\theta$$

$$= \int \frac{(\partial_{x_i}\alpha + \partial_{y_i}\alpha)(x, y, |x - y|, \frac{x-y}{|x-y|})}{|x - y|^{n-1}}f(y)dy + \int \frac{\alpha(x, y, |x - y|, \frac{x-y}{|x-y|})}{|x - y|^{n-1}}\partial_{y^i}f(y)dy$$

Now we can apply Proposition 3.2 to finish the proof. □

Lemma 3.4. *The operator* $KT_1^{-1}K$ *is compact on* $H^2(\mathcal{V} \times \mathbb{S}^{n-1})$*.*

Proof. We find that

$$KT_1^{-1}f(t, x, \theta) = \int_{\mathbb{S}^{n-1}} k(t, x, \theta, \theta') \int_0^t \kappa(t, x, s, \theta')f(s, x + s\theta', \theta')dsd\theta'$$

where κ is defined in (20). Set $y = x+s\theta'$. We get $s = |y-x|, \theta' = (y-x)/|y-x|$ and

$$KT_1^{-1}f(t, x, \theta) = \int_{\mathcal{V}} \frac{k(t, x, \theta, \frac{y-x}{|y-x|})\kappa(t, x, |y - x|, \frac{y-x}{|y-x|})}{|y - x|^{n-1}}f(|y - x|, y, \frac{y - x}{|y - x|})dy$$

Next, we compute

$$KT_1^{-1}Kf(t, x, \theta) = \int \frac{\alpha(t, x, y, \theta, \theta')}{|y - x|^{n-1}}f(|y - x|, y, \theta')dyd\theta' \tag{23}$$

where

$$\alpha(t, x, y, \theta, \theta') = k(t, x, \theta, \frac{y - x}{|y - x|})\kappa(t, x, |y - x|, \frac{y - x}{|y - x|})k(|y - x|, y, \frac{y - x}{|y - x|}, \theta')$$

Note that α is C^5. Since f is compactly supported in the t variable, we can write f in Fourier sine series in t as

$$f(t,x,\theta) = \sum_{n=-\infty}^{\infty} f_n(x,\theta) e^{i2\pi nt/T},$$

where $f_n(x,\theta) = \frac{1}{T}\int_0^T f(t,x,\theta)e^{-i2\pi nt/T}dt$. Also, for $f \in H^2(\mathcal{V}\times\mathbb{S}^{n-1})$, Plancherel's theorem tells

$$\|f\|^2_{H^2(\mathcal{M}\times\mathbb{S}^{n-1})} = \sum_{n=-\infty}^{\infty} \|f_n\|^2_{H^2(\mathbb{R}^n\times\mathbb{S}^{n-1})}$$

and

$$\|\partial_t^2 f\|^2_{H^2(\mathcal{M}\times\mathbb{S}^{n-1})} = \sum_{n=-\infty}^{\infty} T^{-4}n^4\|f_n\|^2_{H^2(\mathbb{R}^n\times\mathbb{S}^{n-1})}$$

See for example [18]. Here, f_n are functions on $\Omega\times\mathbb{S}^{n-1}$ and we extended them trivially to $\mathbb{R}^n\times\mathbb{S}^{n-1}$ for convenience. Let $g_n(t,x,\theta) = f_n(x,\theta)e^{i2\pi nt/T}$. We have

$$KT_1^{-1}Kg_n(t,x,\theta) = \int \frac{\alpha(t,x,y,\theta,\theta')}{|y-x|^{n-1}} e^{i2\pi n\frac{|y-x|}{T}} f_n(y,\theta')dyd\theta'$$

For fixed t, it follows from Proposition 3.3 (i) that the operator is bounded from $H^2(\mathbb{R}^n\times\mathbb{S}^{n-1})$ to $H^3(\mathbb{R}^n\times\mathbb{S}^{n-1})$ with norm not exceeding $Cn^2\|\alpha\|_{C^4}$ with C depending on $\mathcal{V}$, namely

$$\|KT_1^{-1}Kg_n\|_{H^3(\mathbb{R}^n\times\mathbb{S}^{n-1})} \leq C\|\alpha\|_{C^4}n^2\|f_n\|_{H^2(\mathbb{R}^n\times\mathbb{S}^{n-1})}$$

Note that when applying Proposition 3.2, we need k to be compactly supported in t, x variable. Summing up in n, we get

$$\|KT_1^{-1}Kf(t,\cdot,\cdot)\|^2_{H^3(\mathbb{R}^n\times\mathbb{S}^{n-1})} \leq C\sum_{n=-\infty}^{\infty} n^4\|f_n\|^2_{H^2(\mathbb{R}^n\times\mathbb{S}^{n-1})} \leq C\|f\|^2_{H^2(\mathcal{M}\times\mathbb{S}^{n-1})}$$

under our regularity assumption on f. This shows in particular that $KT_1^{-1}Kf \in L^2([0,T], H^3(\mathbb{R}^3\times\mathbb{S}^{n-1}))$. By considering $\partial_t^\beta(KT_1^{-1}Kf)$ for $|\beta|\leq 3$, we see that $KT_1^{-1}K$ is bounded from $H^2(\mathcal{V}\times\mathbb{S}^{n-1})$ to $H^3(\mathcal{V}\times\mathbb{S}^{n-1})$ by using Proposition 3.3 and $\sigma, k \in C^5$. Thus $KT_1^{-1}K$ is compact on $H^2(\mathcal{V}\times\mathbb{S}^{n-1})$. $\square$

Proof of Theorem 3.1. We aim to find k such that T^{-1} exists. Let $\lambda\in\mathbb{C}$. We replace the scattering kernel k in (1) by λk and denote the corresponding operator by λK. Formally, we consider

$$A(\lambda) = (\mathrm{Id} - (\lambda KT_1^{-1})^2)^{-1} \tag{24}$$

so that

$$(\mathrm{Id} - \lambda KT_1^{-1})^{-1} = (\mathrm{Id} + \lambda KT_1^{-1})A(\lambda). \tag{25}$$

We need to justify the invertibility of the operator in $A(\lambda)$. Since $(\lambda KT_1^{-1})^2$ is compact from Lemma 3.4, by analytic Fredholm theorem [14, Theorem VI.14], we know that there exist a discrete set $\mathcal{S}$ of $\mathbb{C}$ such that for $\lambda\notin\mathcal{S}$, $A(\lambda)$ exits and for such λ, (25) is justified. We now use (22) and that T_1^{-1} is bounded on $H^2(\mathcal{M})$ which might not be optimal. But this shows that the operator T is invertible on $H^2(\mathcal{M})$ for scattering kernel λk where $\lambda \in \mathbb{C}\backslash\mathcal{S}$ which implies that the set of such k is dense in $C^5(\mathcal{M} \times \mathbb{S}^{n-1} \times \mathbb{S}^{n-1})$. $\square$

Let u be the solution of (1) with initial condition (2). We set

$$Xf = u|_{t=T}. \tag{26}$$

We can use Theorem 3.1 to obtain a representation for X. Let ρ_T be the restriction operator to $t = T$. It follows from (21) that

$$X = \rho_T T_1^{-1}(\mathrm{Id} - KT_1^{-1})^{-1} \tag{27}$$

We use that $\rho_T T_1^{-1}$ is a light ray transform with weight and is bounded from $H^2(\mathcal{M})$ to $H^{5/2}(\mathbb{R}^n \times \mathbb{S}^{n-1})$, see Section 6. We conclude that $X : H^2(\mathcal{M}) \to H^{5/2}(\mathbb{R}^n \times \mathbb{S}^{n-1})$ is bounded.

4. The Minkowski light ray transform

To prove Theorem 1.1, we will treat X in (27) as a perturbation of the light ray transform on Minkowski spacetime by compact operators. In fact, when $\sigma = k = 0$, we see that

$$Xf = \rho_T T_0^{-1} f = \int_0^T f(s, x + s\theta) ds$$

which is basically the light ray transform on the Minkowski spacetime $(\mathbb{R}^{3+1}, g)$ where $g = -dt^2 + dx_1^2 + dx_2^2 + dx_3^2$. We parametrize the light ray transform using null vectors at $\mathcal{M}_0$ as follows, see Figure 1. For $\theta \in \mathbb{S}^2, x \in \mathbb{R}^3$, the light like geodesics from $(0, x)$ in the direction $(1, \theta)$ is given by $l_{x,\theta}(s) = (s, x + s\theta), s \in \mathbb{R}$. The set of light rays are parametrized by the set $\mathcal{C} \stackrel{\text{def}}{=} \mathbb{R}^3 \times \mathbb{S}^2$. We parametrize the light ray transform as

$$Lf(x, \theta) = \int_{\mathbb{R}} f(s, x + s\theta) ds \tag{28}$$

When $\sigma \neq 0, k = 0$, X is a light ray transform with weight which we study later in Section 6. If $\sigma(z) = \sigma(t)$ only depends on the t variable, we have

$$Xf = \int_0^T \kappa(s) f(s, x + s\theta) ds = L(\kappa f) \text{ where } \kappa(s) = e^{-\int_s^T \sigma(\tilde{s}) d\tilde{s}} \tag{29}$$

In this case, it suffices to look at the light ray transform L. This is why we impose the assumption on σ in Theorem 1.1.

It is known that L is injective on C_0^∞ functions. However, when acting on say Schwartz functions, L has a non-trivial kernel consisting of functions whose Fourier transform is supported in Γ^{tm}, see for instance [9]. Recall the operator $\phi(D)$ defined in the introduction. Let $\phi \in \mathcal{D}'(\mathbb{R}^{3+1})$ such that $\phi(\zeta) = 1$ if $\zeta \in \Gamma^{sp} \cup \Gamma^{lt} \cup \{0\}$ and $\phi(\zeta) = 0$ if $\zeta \in \Gamma^{tm}$. Then ϕ is continuous in Γ^{sp} up to the boundary of Γ^{sp}. It is easy to see that $\phi(D) : H^s(\mathbb{R}^{3+1}) \to H^s(\mathbb{R}^{3+1}), s \in \mathbb{R}$ is bounded. Also, $\phi^2(D) = \phi(D)$ so $\phi(D)$ is a projection on $H^s(\mathbb{R}^{3+1})$. We denote the range of $\phi(D)$ on $H^s(\mathbb{R}^{3+1})$ by $\mathscr{H}^s$ which is a closed subspace of $H^s(\mathbb{R}^{3+1})$, hence a Hilbert space. The following simple result plays an important role.

Proposition 4.1. *For $f \in L^2_{\text{comp}}(\mathbb{R}^{3+1})$, we have*

$$L(\kappa f) = L(\phi(D)(\kappa f)). \tag{30}$$

Proof. We use the Fourier slice theorem

$$\begin{aligned} \mathcal{F}_y(L(\kappa f))(\xi, \theta) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}} e^{-iy\xi} \kappa(s) f(s, y + s\theta) ds dy \\ &= \mathcal{F}_{(t,x)}(\kappa f)(-\theta \cdot \xi, \xi, \theta) \end{aligned} \tag{31}$$

Here, $\kappa f \in L^2_{\text{comp}}$. We apply the Fourier slice theorem to $L\phi(D)\kappa f$ and get

$$\mathcal{F}_y(L(\phi(D)\kappa f))(\xi, \theta) = \phi(-\theta \cdot \xi, \xi) \mathcal{F}_{(t,x)}(\kappa f)(-\theta \cdot \xi, \xi, \theta) \tag{32}$$

For any $\theta \in \mathbb{S}^2$, we see that $(-\theta \cdot \xi, \xi)$ is non time-like for all $\xi \in \mathbb{R}^3 \backslash 0$. So $\phi(-\theta \cdot \xi, \xi) = 1$. Thus (31) is equal to (32) for $\xi \in \mathbb{R}^3$ and $\theta \in \mathbb{S}^2$ by the continuity of $\mathcal{F}_{(t,x)}(\kappa f)$ and that ϕ is continuous in Γ^{sp} up to the boundary of Γ^{sp}. Taking the inverse Fourier transform gives (30). □

Next, let L^* be the adjoint of L and $N = L^* L$ be the normal operator. We recall from [12] that (for general dimension $n \geq 2$)

$$Nf = \int_{\mathbb{R}^{n+1}} K_N(t, x, t', x') f(t', x') dt' dx' \tag{33}$$

where the Schwartz kernel

$$K_N(t, x, t', x') = \frac{\delta(t - t' - |x - x'|) + \delta(t - t' + |x - x'|)}{|x - x'|^{n-1}} \tag{34}$$

In particular, N can be written as a Fourier multiplier

$$Nf(t, x) = \int_{\mathbb{R}^{n+1}} e^{i(t,x)\cdot(\tau,\xi)} k(\tau, \xi) \hat{f}(\tau, \xi) d\tau d\xi \tag{35}$$

where

$$k(\tau, \xi) = C_n \frac{(|\xi|^2 - \tau^2)_+^{\frac{n-3}{2}}}{|\xi|^{n-2}}, \quad C_n = 2\pi |\mathbb{S}^{n-2}|. \tag{36}$$

Here, for $t \in \mathbb{R}, \alpha \in \mathbb{R}$, we denote t^α_+ the homogeneous distribution. We can write the normal operator as

$$Nf(t, x) = \int e^{i(t-t', x-x')\cdot(\tau,\xi)} k(\tau, \xi) f(t', x') d\tau d\xi dt' dx'$$

This is a pseudo-differential operator away from $\tau^2 = |\xi|^2$ where the symbol $k(\tau, \xi)$ is singular. We will discuss later that the kernel belongs to the class of paired Lagrangian distributions, see also [21]. However, it is not needed for Theorem 1.1. We recall the Sobolev estimate for L, L^*.

Lemma 4.2 (Theorem 1.2 of [21]). *For $n \geq 3, s \in \mathbb{R}$, the light ray transform $L : H^s_{\text{comp}}(\mathbb{R}^{n+1}) \to H^{s+1/2}_{\text{loc}}(\mathbb{R}^n \times \mathbb{S}^{n-1}), s \in \mathbb{R}$ and its adjoint $L^* : H^s_{\text{comp}}(\mathbb{R}^n \times \mathbb{S}^{n-1}) \to H^{s+1/2}_{\text{loc}}(\mathbb{R}^{n+1})$ are continuous.*

For $n = 3$, the analysis of L and N are simpler. Note that

$$k(\tau, \xi) = 4\pi^2 \frac{\phi(\tau, \xi)}{|\xi|} \tag{37}$$

It follows from (35) that $Nf = N\phi(D)f$. Let Q be defined by a Fourier multiplier

$$\mathcal{F}(Qf)(\tau, \xi) = q(\tau, \xi)\hat{f}(\tau, \xi)$$

where

$$q(\tau, \xi) = (4\pi^2)^{-1}\phi(\tau, \xi)|\xi|^{-1} \tag{38}$$

Then we see that

$$QN\phi(D)f = \phi(D)f \tag{39}$$

which means that N is invertible on $\mathscr{H}^s$. We need the mapping properties of Q on Sobolev spaces. From (38), we can write $Q = \phi(D)Q_1$ where Q_1 is defined by Fourier multiplier as

$$Q_1 f = \mathcal{F}^{-1}((2\pi)^{-2}|\xi|\hat{f}(\xi))$$

We see that Q_1 is an pseudo-differential operator of order 1 on $\mathbb{R}^4$ (modulo a smoothing operator). We know that $\phi(D)$ is bounded on $H^s(\mathbb{R}^4), s \in \mathbb{R}$ thus

$$Q : H^s_{\text{comp}}(\mathbb{R}^{3+1}) \to H^{s-1}_{\text{loc}}(\mathbb{R}^{3+1})$$

is bounded for $s \in \mathbb{R}$.

5. Proof of Theorem 1.1

We outline the proof of Theorem 1.1. We write (27) as

$$X = \rho_T(\text{Id} - \text{Id} + (\text{Id} + T_1^{-1}K)^{-1})T_1^{-1} = L\kappa + E$$

where

$$E = \rho_T(-\text{Id} + (\text{Id} + T_1^{-1}K)^{-1})T_1^{-1} = \rho_T T_1^{-1}K(\text{Id} - T_1^{-1}K)^{-1}T_1^{-1} \tag{40}$$

To "invert" X, we apply L^* to X

$$L^*X = L^*L\kappa + L^*E\kappa$$

The operator $N = L^*L$ is invertible on $\mathscr{H}^s$. In fact, it suffices to consider L^*X on $\mathscr{H}^s$ as follows.

Consider the operator T_1^{-1} defined in the proof of Theorem 3.1. For σ depending only on t, we have

$$T_1^{-1}f(t, x, \theta) = \int_0^t \kappa(s)f(s, x + s\theta)ds$$

For any $t \in [0, T]$, let ρ_t be the restriction operator to $\mathcal{M}_t = \{t\} \times \mathbb{R}^3$. We see that $\rho_t T_1^{-1}$ is also a light ray transform with weight κ. The argument in Proposition 4.1 implies that

$$\rho_t T_1^{-1} f = \rho_t T_1^{-1} \phi(D)\kappa f.$$

Because it holds for all t, we have

$$Xf = L\phi(D)\kappa f + E\phi(D)\kappa f$$

which further gives

$$\phi(D)L^* Xf = \phi(D)N\phi(D)\kappa f + \phi(D)L^* E\phi(D)\kappa f \tag{41}$$

Now we can apply Q and get

$$Q\phi(D)L^* Xf = \phi(D)\kappa f + Q\phi(D)L^* E\phi(D)\kappa f \tag{42}$$

Recall that κ is non-vanishing. It suffices to consider the right hand side of (42) as acting on functions in $\mathscr{H}^s$, that is to consider an operator

$$\mathrm{Id} + Q\kappa\phi(D)L^* E$$

on $\mathscr{H}^s$. The main part of the proof is to show that $Q\kappa\phi(D)L^*E : \mathscr{H}^2 \to \mathscr{H}^2$ is compact. We know that $\kappa\phi(D)$ is bounded on $H^2(\mathcal{M})$ and Q is a pseudo-differential operator of order 1. Thus, it suffices to show that $\partial_t(L^*E), \partial_{x_i}(L^*E) : H^2 \to H^2, i = 1, 2, 3$ is compact which we prove below.

We write E in (40) as

$$E = \rho_T T_1^{-1} K T_1^{-1}(\mathrm{Id} - T_1^{-1}K)^{-1}$$

so that

$$L^*E = L^*\rho_T T_1^{-1} K T_1^{-1}(\mathrm{Id} - T_1^{-1}K)^{-1} \tag{43}$$

We observe that $\rho_T T_1^{-1}$ is almost L, however it is acting on $\mathcal{M} \times \mathbb{S}^2$ instead of $\mathcal{M}$. As in [17], we let J be defined as $Jf(z, \theta) = f(z)$. Then $\rho_T T_1^{-1} J = L$. Also, we note that J is bounded $L^2(\mathcal{M}) \to L^2(\mathcal{M} \times \mathbb{S}^2)$ for example. To show the compactness, we adopt the idea in [17] to decompose the operator (43).

For $k \in C^6(\mathcal{M} \times \mathbb{S}^2_\theta \times \mathbb{S}^2_{\theta'})$, we write

$$k(z, \theta, \theta') = \sum_{j=1}^{\infty} \Theta_j(\theta) k_j(z, \theta') \tag{44}$$

where Θ_j are spherical harmonics on $\mathbb{S}^2$ and k_j are the corresponding Fourier coefficients. With our regularity assumptions, the series converges, see page 121 of [17]. We remark that it suffices to impose weaker regularity assumptions on k to get such expansion. But this is not our concern.

Now we let K_j, B_j be defined as

$$K_j f(t, x, \theta) = \int \Theta_j(\theta) k_j(t, x, \theta') f(t, x, \theta') d\theta',$$

$$B_j f(t, x) = \int k_j(t, x, \theta)\kappa(t, x, s, \theta) f(s, x + s\theta, \theta) ds d\theta.$$

Using (43), we get

$$
\begin{aligned}
L^*E &= \sum_{j=1}^{\infty} L^*\rho_T T_1^{-1} K_j T_1^{-1}(\mathrm{Id} - T_1^{-1}K)^{-1}J \\
&= \sum_{j=1}^{\infty} \underbrace{L^*\rho_T T_1^{-1}\Theta_j J}_{P_j^1} \underbrace{B_j(\mathrm{Id} - T_1^{-1}K)^{-1}J}_{P_j^2}
\end{aligned} \tag{45}
$$

We analyze P_j^1 and P_j^2 separately. First, we find the kernel of $P_j^1 = L^*L\Theta_j J$. We use the expression of L^* in [12] (see also Proposition 6.2 in Section 6 later) to get

$$
L^*L\Theta_j Jf(t,x) = \int_{\mathbb{S}^2} L\Theta_j Jf(x - t\theta, \theta)d\theta = \int_{\mathbb{S}^2}\int_{\mathbb{R}} \Theta_j(\theta)f(s, x + s\theta - t\theta)dsd\theta
$$

Here, we split the integral in s to $s \geq t$ and $s \leq t$. For $s \geq t$, we let $r = s - t$ and make a change of variable $y = x + r\theta$. The $s \leq t$ part can be treated similarly. We get

$$
\begin{aligned}
L^*L\Theta_j Jf(t,x) = \int_{\mathbb{R}^3} [\Theta_j(\frac{y-x}{|y-x|})f(t + |x-y|, y) \\
+\Theta_j(\frac{x-y}{|x-y|})f(t - |x-y|, y)]|x-y|^{-2}dy
\end{aligned}
$$

Note that $L^*L\Theta_j f$ is compactly supported in $\mathcal{M}$ because f is. Again, we write f in Fourier sine series in t as in Lemma 3.4 and use the notations there. We look at

$$
\begin{aligned}
L^*L\Theta_j Jg_n(t,x) &= \int_{\mathbb{R}^3} \Theta_j(\frac{y-x}{|y-x|})e^{i2\pi n\frac{(t-|y-x|)}{T}} f_n(y)|x-y|^{-2}dy \\
&+ \int_{\mathbb{R}^3} \Theta_j(\frac{x-y}{|x-y|})e^{i2\pi n\frac{(t+|y-x|)}{T}} f_n(y)|x-y|^{-2}dy
\end{aligned} \tag{46}
$$

From Proposition 3.3 (ii), we deduce that for fixed t

$$
\|L^*L\Theta_j Jg_n(t,\cdot)\|_{H^3(\mathbb{R}^3)} \leq C\|\Theta_j\|_{H^1} n^2 \|f_n\|_{H^2(\mathbb{R}^3)}.
$$

thus

$$
\|L^*L\Theta_j Jf(t,\cdot)\|_{H^3(\mathbb{R}^3)} \leq C\|\Theta_j\|_{H^1}\|f\|_{H^2(\mathcal{M})}
$$

which implies that $L^*L\Theta_j Jf \in L^2([0,T], H^3(\mathbb{R}^3))$. We see that $\partial_{x^i} P_j^1, i = 1,2,3$ is bounded on $H^2(\mathcal{M})$. We need to consider $\partial_t P_j^1$ on H^2 for which we need three t derivatives of P_j^1. For this purpose, note that each t derivative of (46) results in a factor n. So we need $f \in H^5([0,T], H^2(\mathbb{R}^3))$ to get

$$
\|L^*L\Theta_j Jf\|_{H^3(\mathcal{M})} \leq C\|\Theta_j\|_{H^1}\|f\|_{H^5([0,T],H^2(\mathbb{R}^3))}.
$$

Thus $\partial_t P_j^1, \partial_x P_j^1$ are bounded from $H^5([0,T], H^2(\mathbb{R}^3))$ to $H^2(\mathcal{M})$. This suggests us prove the following mapping property for P_j^2.

Lemma 5.1. *$P_j^2 : H^2(\mathcal{M}) \to H^5([0,T], H^2(\mathbb{R}^3))$ is compact.*

Proof. We decompose P_j^2 as follows

$$P_j^2 = B_j J + B_j(\mathrm{Id} - T_1^{-1}K)^{-1} K T_1^{-1} J \tag{47}$$

We first show that $KT_1^{-1}J$ is compact from $H^2(\mathcal{M})$ to $H^2(\mathcal{M} \times \mathbb{R}^2)$. We compute

$$KT_1^{-1}Jf(t,x,\theta) = \int_{\mathbb{S}^{n-1}} k(t,x,\theta,\theta') \int_0^t e^{-\int_s^t \sigma(\tilde{s}, x+\tilde{s}\theta', \theta')d\tilde{s}} f(s, x+s\theta') ds d\theta'$$

Here, we included a general weight in T_1^{-1} instead of $\kappa(s)$ because we need it later for Section 8. Set $y = x + s\theta'$. We get $s = |y-x|, \theta' = (y-x)/|y-x|$ and

$$KT_1^{-1}Jf(t,x,\theta) = \int_{\mathcal{V}} \frac{k(t,x,\theta,\frac{y-x}{|y-x|})\kappa(t,x,|y-x|,\frac{y-x}{|y-x|})}{|y-x|^{n-1}} f(|y-x|, y) dy$$

We use the Fourier sine series of f in t as in Lemma 3.4 and the notations there to get

$$KT_1^{-1}Jg_n(t,x,\theta) = \int_{\mathcal{V}} \frac{\alpha(t,x,y,\theta)}{|y-x|^{n-1}} e^{i2\pi n \frac{|y-x|}{T}} f_n(y) dy$$

For fixed t, θ, it follows from Proposition 3.3 (i) that

$$\|KT_1^{-1}Jg_n\|_{H^1} \leq Cn^2 \|f_n\|_{L^2(\mathbb{R}^n \times \mathbb{S}^{n-1})}$$

Summing up in n, we get for fixed t, θ

$$\|KT_1^{-1}Jf(t,\cdot,\theta)\|^2_{H^3(\mathbb{R}^3)} \leq C \sum_{n=-\infty}^{\infty} n^4 \|f_n\|^2_{L^2(\mathbb{R}^n \times \mathbb{S}^{n-1})} \leq C\|f\|^2_{H^2(\mathcal{M})}.$$

By considering $\partial^\beta_{t,\theta}(KT_1^{-1}Jf)$ for $|\beta| \leq 3$, we see that $KT_1^{-1}J$ is bounded from $H^2(\mathcal{M})$ to $H^3(\mathcal{M} \times \mathbb{S}^2)$ thus compact to $H^2(\mathcal{M} \times \mathbb{S}^2)$. We know that $(\mathrm{Id} - T_1^{-1}K)^{-1}$ is bounded on H^2. It remains to consider B_j on $H^2(\mathcal{M} \times \mathbb{S}^2)$ which is

$$\begin{aligned} B_j f(t,x) &= \int k_j(t,x,\theta)\kappa(t,x,s,\theta) f(s, x+s\theta, \theta) ds d\theta \\ &= \int \frac{k_j(t,x,\frac{y-x}{|y-x|})\kappa(t,x,|x-y|,\frac{y-x}{|y-x|})}{|x-y|^2} f(|x-y|, y, \frac{y-x}{|y-x|}) dy \end{aligned}$$

By considering the t derivative which hits k_j, κ, we can show that B_j is bounded from $H^2(\mathcal{M} \times \mathbb{S}^2)$ to $H^5([0,T] \times H^2(\mathbb{R}^3))$ using Proposition 3.3. The bound depends on $\|k_j\|_{C^5}$. It is very important that the t derivatives do not produce n factors as in the analysis for P_j^1. This is why we can gain regularity in t. We thus proved that the second term of (47) is compact.

Next, consider $B_j J$ in (47) which is

$$B_j J f(t,x) = \int \frac{k_j(t,x,\frac{y-x}{|y-x|})\kappa(t,x,|y-x|,\frac{y-x}{|y-x|})}{|y-x|^2} f(|y-x|, y) dy$$

We use Proposition 3.2 and consider the t derivative of $B_j J$ as above to conclude that $B_j J$ is bounded from $H^2(\mathcal{M})$ to $H^6([0,T], H^3(\mathbb{R}^3))$, hence compact from $H^2(\mathcal{M})$ to $H^5([0,T], H^2(\mathbb{R}^3))$. This shows the compactness of P_2^j with a bound $C\|k_j\|_{C^6}$. This completes the proof of the lemma. □

Using these estimates, we see that P_j^1, P_j^2 are bounded on $H^2(\mathcal{M})$ with a bound $\|\kappa_j\|_{C^6}\|\Theta_j\|_{H^1}$. Now consider derivatives of k in t, x, θ variables in the series (44). Using the argument in the end of page 121 of [17], we see that the series

$$\sum_{j=1}^{\infty} \|\kappa_j\|_{C^6}\|\Theta_j\|_{H^1} < \infty$$

with the regularity assumption of k. Thus the series in (43) converges uniformly. This shows that L^*E is compact on H^2.

Completion of the proof of Theorem 1.1. We use (41)

$$L^*Xf = \phi(D)L^*L\phi(D)\kappa f + L^*E\phi(D)\kappa f$$

and apply Q to get

$$QL^*Xf = \phi(D)\kappa f + Q\kappa^{-1}L^*E\phi(D)\kappa f$$

We recall that the function κ depends on σ. For $\lambda \in \mathbb{C}$ in a neighborhood of $[0,1]$, we replace the scattering kernel k by λk and denote the operator K by $K(\lambda)$. Then we see that the operator in (40), now denoted by $E(\lambda)$

$$E(\lambda) = \rho_T T_1^{-1}K(\lambda)(\mathrm{Id} - T_1^{-1}K(\lambda))^{-1}T_1^{-1} \tag{48}$$

is a family of operators meromorphic in λ. Suppose $E(\lambda)$ is holomorphic in $\mathcal{U} \subset \mathbb{C}$. For $\lambda = 0$, we know that $\mathrm{Id} + \mathcal{E}(0)$ is invertible. Thus by analytic Fredholm theorem [14, Theorem VI.14], we conclude that $\mathrm{Id} + \mathcal{E}(\lambda)$ with $\mathcal{E}(\lambda) = QL^*E(\lambda)$ is invertible for λ in $\mathcal{U}\backslash\mathcal{S}$ where $\mathcal{S}$ is a discrete set. This implies that QL^*X is invertible for an open dense subset of k. Also, we obtain the stability estimate

$$\|\phi(D)\kappa f\|_{H^2(\mathcal{M})} \leq \|L^*Xf\|_{H^3(\mathcal{M})} \leq C\|Xf\|_{H^{5/2}(\mathcal{C})}$$

using the estimate of L^* and Q in Section 4.

Finally, for $f \in H^2_{\mathrm{comp}}(\mathcal{M})$, we know that $Xf \in H^{5/2}(\mathcal{C})$. If $Xf = 0$, we get $\phi(D)\kappa f = 0$. By taking Fourier transform, we see that $\mathcal{F}(\kappa f)(\zeta) = 0$ for $\zeta \in \Gamma^{sp}$. But κf is compactly supported so $\mathcal{F}(\kappa f)(\zeta)$ is analytic in ζ. We conclude that $\kappa f = 0$ so $f = 0$. This proves the uniqueness. □

6. The light ray transform with weights

The rest of the paper is devoted to the proof of Theorem 1.2. It is worth pointing out the differences of the proof to Theorem 1.1. Following the notations in Section 3, we can represent X as a perturbation of a weighted light ray transform L_κ but it cannot be reduced to the Minkowski light ray transform when σ depends on the x variable. There are some results on the injectivity of the weighted light ray transform for analytic weights, see [16]. But the description of the kernel is unclear so the argument for Theorem 1.1, especially Proposition 4.1 does not seem to work.

The way we prove Theorem 1.2 is to use the constraint of the Cauchy problem to obtain some stability estimate for the weighted light ray transform. Roughly speaking, for the Cauchy problem (6), we can find a parametrix and represent the solution as $f = E(f_1, f_2)$. We then show that the composition LE can be microlocally inverted with a compact remainder. Then the analytic Fredholm argument can be applied. We remark that for the light ray transform without weights, similar stability estimates have been studied in [20, 21].

The desired stability estimate relies on the microlocal structure of the light ray transform which we study in this section. Let $\kappa(t, x, \theta) \in C^\infty(\mathbb{R} \times \mathbb{R}^3 \times \mathbb{S}^2)$ be real valued with compact support in t, x variables. We consider a weighted light ray transform

$$L_\kappa f(x, \theta) = \int_{-\infty}^{\infty} \kappa(s, x + s\theta, \theta) f(s, x + s\theta) ds \tag{49}$$

for $f \in C_0^\infty(\mathbb{R}^4)$. Note that (20) is of the form (49). We first compute the Schwartz kernel of the normal operator.

Proposition 6.1. *Let $N_\kappa = L_\kappa^* L_\kappa$. The Schwartz kernel of N_κ as a distribution on $\mathbb{R}^4 \times \mathbb{R}^4$ is given by*

(50)

$$K(t, x, s, y) = \frac{\kappa(t, x, \frac{y-x}{|y-x|})\kappa(s, y, \frac{x-y}{|x-y|})}{|y - x|^2}(\delta(t - s - |y - x|) + \delta(t - s + |y - x|)).$$

Proof. We start by computing L_κ^*. Let $g \in C_0^\infty(\mathbb{R}^3 \times \mathbb{S}^2)$. We have

$$\begin{aligned}\langle L_\kappa f, g\rangle &= \int_{\mathbb{S}^2}\int_{\mathbb{R}^3}\int_{\mathbb{R}} \kappa(s, x + s\theta, \theta) f(s, x + s\theta) g(x, \theta) ds dx d\theta \\ &= \int_{\mathbb{R}^3}\int_{\mathbb{R}}\int_{\mathbb{S}^2} \kappa(s, y, \theta) f(s, y) g(y - s\theta, \theta) d\theta ds dy\end{aligned}$$

where we made a change of variable via $x + s\theta = y$. Thus, we get

$$L_\kappa^* g(s, y) = \int_{\mathbb{S}^2} \kappa(s, y, \theta) g(y - s\theta, \theta) d\theta$$

Next, we compute

$$\begin{aligned}L_\kappa^* L_\kappa f(t, x) &= \int_{\mathbb{S}^2} \kappa(t, x, \theta) L_\kappa f(x - t\theta, \theta) d\theta \\ &= \int_{\mathbb{S}^2}\int_{\mathbb{R}} \kappa(t, x, \theta)\kappa(s, x - t\theta + s\theta, \theta) f(s, x - t\theta + s\theta) ds d\theta\end{aligned} \tag{51}$$

Then we split the integral to $s \geq t$ and $s \leq t$. For $s \geq t$, we let $r = s - t \geq 0$ and use polar coordinates for $\mathbb{R}^3$ centered at x: $y = x + r\theta$. We see that $\theta = (y - x)/|y - x|$ and $r = |y - x|$. In this case, the integral (51) denoted by

I_+ below is

$$\begin{aligned} I_+(t,x) &= \int_{\mathbb{S}^2}\int_0^\infty \kappa(t,x,\theta)\kappa(r+t,x+r\theta,\theta)f(r+t,x+r\theta)drd\theta \\ &= \int_{\mathbb{R}^3} \frac{\kappa(t,x,\frac{y-x}{|y-x|})\kappa(t+|y-x|,y,\frac{y-x}{|y-x|})}{|y-x|^2} f(t+|y-x|,y)dy \\ &= \int_{\mathbb{R}^4} \frac{\kappa(t,x,\frac{y-x}{|y-x|})\kappa(s,y,\frac{y-x}{|y-x|})}{|y-x|^2}\delta(s-t-|y-x|)f(s,y)dsdy \end{aligned}$$

For $s \leq t$ in (51), we follow the same procedure. Let $r = t-s \geq 0$ and we change θ to $-\theta$ in the integral. Then we use polar coordinate $y = x + r\theta$. If we denote the integral of (51) by I_-, we get

$$\begin{aligned} I_-(t,x) &= \int_{\mathbb{S}^2}\int_0^\infty \kappa(t,x,-\theta)\kappa(s,x+r\theta,-\theta)f(s,x+r\theta)drd\theta \\ &= \int_{\mathbb{R}^3} \frac{\kappa(t,x,\frac{x-y}{|x-y|})\kappa(t-|x-y|,y,\frac{x-y}{|x-y|})}{|y-x|^2} f(t-|y-x|,y)dy \\ &= \int_{\mathbb{R}^4} \frac{\kappa(t,x,\frac{x-y}{|x-y|})\kappa(s,y,\frac{x-y}{|x-y|})}{|y-x|^2}\delta(s-t+|x-y|)f(s,y)dsdy \end{aligned}$$

Adding up I_+ and I_+, we obtain the kernel K. $\square$

Next, we will show that the Schwartz kernel of N_κ is a paired Lagrangian distribution from which we will derive Sobolev estimate for N_κ and L_κ. For globally hyperbolic Lorentzian manifolds without conjugate points, this was recently proved in [22] for the light ray transform without weight using the calculation of the Schwartz kernel of the normal operator. The method works the same for the weighted light ray transform given Proposition 6.1.

We briefly recall the notion of paired Lagrangian distribution, see e.g. [3]. Let $\mathcal{X}$ be a C^∞ manifold of dimension n. Let Λ_0, Λ_1 be conic Lagrangian submanifolds of $T^*(\mathcal{X}\times\mathcal{X})\backslash 0$. Suppose that Λ_1 intersects Λ_0 cleanly at a co-dimension k, $1 \leq k \leq 2n-1$ submanifold $\Sigma = \Lambda_0 \cap \Lambda_1$, namely

$$T_p(\Lambda_0\cap\Lambda_1) = T_p(\Lambda_0)\cap T_p(\Lambda), \quad \forall p \in \Sigma.$$

It is known that all such intersecting pairs (Λ_0, Λ_1) are locally symplectic diffeomorphic to each other. It suffices to consider the following model problem. Let $\tilde{\mathcal{X}} = \mathbb{R}^n = \mathbb{R}^k\times\mathbb{R}^{n-k}, 1 \leq k \leq n-1$, and use coordinates $x = (x', x''), x' \in \mathbb{R}^k, x'' \in \mathbb{R}^{n-k}$. Let $\tilde\Lambda_0 = \{(x,\xi,x,-\xi) \in T^*(\tilde{\mathcal{X}}\times\tilde{\mathcal{X}})\backslash 0 : \xi \neq 0\}$ be the punctured conormal bundle of Diag in $T^*(\tilde{\mathcal{X}}\times\tilde{\mathcal{X}})$, and

$$\tilde\Lambda_1 = \{(x,\xi,y,\eta) \in T^*(\tilde{\mathcal{X}}\times\tilde{\mathcal{X}})\backslash 0 : x'' = y'', \xi' = \eta' = 0, \xi'' = \eta'' \neq 0\}$$

which is the punctured conormal bundle to $\{(x,y) \in \tilde{\mathcal{X}}\times\tilde{\mathcal{X}} : x'' = y''\}$. The two Lagrangians intersect cleanly at $\tilde\Sigma = \{(x,\xi,y,\eta) \in T^*(\tilde{\mathcal{X}}\times\tilde{\mathcal{X}})\backslash 0 : x'' = y'', \xi'' = \eta'', x' = y', \xi' = \eta' = 0\}$ which is of co-dimension k. For this model pair, the paired Lagrangian distribution $I^{p,l}(\mathbb{R}^n \times \mathbb{R}^n; \tilde{\Lambda}_0, \tilde{\Lambda}_1)$ consists of oscillatory integrals

$$u(x,y) = \int e^{i[(x'-y'-s)\cdot\eta' + (x''-y'')\cdot\eta'' + s\cdot\sigma]} a(s,x,y,\eta,\sigma) d\eta d\sigma ds \tag{52}$$

where a is a product type symbol which is a C^∞ function and satisfies

$$|\partial_\eta^\alpha \partial_\sigma^\beta \partial_s^\theta \partial_x^\gamma \partial_y^\delta a(s,x,y,\eta,\sigma)| \leq C(1+|\eta|)^{p+k/2-|\alpha|}(1+|\sigma|)^{l-k/2-|\beta|} \tag{53}$$

for multi-indices $\alpha, \beta, \theta, \gamma, \delta$ over each compact set $\mathcal{K}$ of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^k$. The constant C depends on the indices and $\mathcal{K}$. The set of product type symbols is denoted by $S^{p,l}(\mathbb{R}^n \times \mathbb{R}^n; \mathbb{R}^n; \mathbb{R}^k)$. We use the notation $I^{p,l}(\mathbb{R}^n \times \mathbb{R}^n; \tilde{\Lambda}_0, \tilde{\Lambda}_1)$ to denote the space of operators $A : \mathcal{E}'(\mathbb{R}^n; \Omega^{\frac{1}{2}}_{\mathbb{R}^n}) \to \mathcal{D}'(\mathbb{R}^n; \Omega^{\frac{1}{2}}_{\mathbb{R}^n})$ where $\Omega^{\frac{1}{2}}_{\mathbb{R}^n}$ denotes the line bundle of half-densities on $\mathbb{R}^n$, whose Schwartz kernel K_A is a paired Lagrangian distribution with values in $\Omega^{\frac{1}{2}}_{\mathbb{R}^n \times \mathbb{R}^n}$. Away from $\tilde{\Sigma} = \tilde{\Lambda}_0 \cap \tilde{\Lambda}_1$, $u \in I^{p+l}(\mathbb{R}^n \times \mathbb{R}^n; \tilde{\Lambda}_0)$ and $u \in I^p(\mathbb{R}^n \times \mathbb{R}^n; \tilde{\Lambda}_1)$ using Hörmander's notion of Lagrangian distributions, see [8, Section 25.1].

There is an equivalent description of the paired Lagrangian distribution (52) introduced in [3, Section 5] that is convenient for our purpose. Modulo $C_0^\infty(\mathbb{R}^n \times \mathbb{R}^n)$, (52) can be written as

$$u(x,y) = \int e^{i[(x'-y')\cdot\eta' + (x''-y'')\cdot\eta'']} b(x,y,\eta) d\eta \tag{54}$$

where b satisfies the following estimates. First, in the region $|\eta'| \leq C|\eta''|, |\eta''| \geq 1$, b satisfies

$$|(Qb)(x,y,\eta)| \leq C\langle\eta''\rangle^{p+k/2}\langle\eta'\rangle^{l-k/2}$$

for all Q which is a finite product of differential operators of the form $D_{\eta'}, \eta'_j D_{\eta'_m}$, $\eta''_j D_{\eta''_m}$. Second, in the region $|\eta''| \leq C|\eta'|, |\eta'| \geq 1$, b satisfies the standard regularity estimate

$$|(Qb)(x,y,\eta)| \leq C\langle\eta'\rangle^{p+l}$$

for all Q which is a finite product of differential operators of the form $\eta'_j D_{\eta'_m}$, $\eta'_j D_{\eta''_m}$. We refer the readers to [3, Section 5] for the argument of the equivalence of (52) and (54), and a description of the principal symbols [3, Lemma 5.3].

We will show that N_κ is a paired Lagrangian distribution associated with two Lagrangians Λ_0, Λ_1 described as follows. First, let

$$\begin{gathered} \Lambda_0 = \{(t,x,\tau,\xi;t',x',\tau',\xi') \in T^*\mathbb{R}^{n+1}\backslash 0 \times T^*\mathbb{R}^{n+1}\backslash 0 : \\ t = t', x = x', \tau = -\tau', \xi = -\xi'\} \end{gathered} \tag{55}$$

which is the punctured conormal bundle of the diagonal in $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ and

$$\begin{gathered} \Lambda_1 = \{(t,x,\tau,\xi;t',x',\tau',\xi') \in T^*\mathbb{R}^{n+1}\backslash 0 \times T^*\mathbb{R}^{n+1}\backslash 0 : \\ x = x' + (t-t')\xi/|\xi|, \tau = \pm|\xi|, \tau' = -\tau, \xi' = -\xi\}. \end{gathered} \tag{56}$$

The two Lagrangians intersect cleanly at

$$\Sigma = \{(t,x,\tau,\xi;t',x',\tau',\xi') \in T^*\mathbb{R}^{n+1}\backslash 0 \times T^*\mathbb{R}^{n+1}\backslash 0 : t = t', x = x', \tag{57}$$
$$\tau = -\tau', \xi = -\xi', \tau^2 = |\xi|^2\}$$

In fact, Λ_1 is the flow out of Σ under the Hamilton vector field H_f of $f(\tau,\xi) = \frac{1}{2}(\tau^2 - |\xi|^2)$. Indeed, we have

$$H_f = \tau\frac{\partial}{\partial t} + \sum_{i=1}^{3}\xi_i\frac{\partial}{\partial x^i}$$

Let $\gamma(s) = (t(s), x(s), \tau(s), \xi(s))$ be a null bi-characteristic which satisfy

$$\dot{t}(s) = \tau, \quad \dot{x}_i(s) = \xi_i, \quad \dot{\tau}(s) = 0, \quad \dot{\xi}_i(s) = 0 \quad s \geq 0$$
$$t(0) = t', \quad x(0) = x', \quad \tau(0) = \tau', \quad \xi(0) = \xi'$$

with $f(\tau',\xi') = 0$. We solve that

$$t(s) = t' + s\tau', \quad x(s) = x' + s\xi', \quad \tau(s) = \tau', \quad \xi(s) = \xi'.$$

which up to a re-parametrization gives (56).

Proposition 6.2. *For the weighted light ray transform L_κ in* (49)*, the Schwartz kernel of the normal operator $N_\kappa = L_\kappa^* L_\kappa$ belongs to $I^{-3/2,1/2}(\mathbb{R}^{3+1}\times\mathbb{R}^{3+1};\Lambda_0,\Lambda_1)$, in which Λ_0, Λ_1 are two cleanly intersection Lagrangians defined in* (55)*,* (56)*. The principal symbol of N_κ on $\Lambda_1\backslash\Sigma$ is non-vanishing if κ is a positive smooth function.*

Proof. To see that the kernel is a paired Lagrangian distribution, we will find its oscillatory integral representation. We start from

$$N_\kappa f(t,x) = \int K(t,x,t',x')f(t',x')dt'dx'$$

where K is given by Proposition 6.1. Using the Fourier transform of the delta distribution, we have

$$\begin{aligned}
N_\kappa f(t,x) &= C\int_{\mathbb{R}^3}\int_{\mathbb{R}}\int_{\mathbb{R}}(e^{i(t-t'-|x'-x|)\tau} + e^{i(t-t'+|x'-x|)\tau}) \\
&\quad\cdot\frac{\kappa(t,x,\frac{x'-x}{|x'-x|})\kappa(t',x',-\frac{x'-x}{|x'-x|})}{|x'-x|^2}f(t',x')d\tau dt'dx' \\
&= \int_{\mathbb{R}^3}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{i(t-t')\tau}B(t,x,t',x',\tau)f(t',x')d\tau dt'dx'
\end{aligned}$$

where C is a non-zero constant due to the Fourier transform which we do not keep track of, and

$$\begin{aligned}
B(t,x,t',x',\tau) &= (e^{i|x'-x|\tau} + e^{-i|x'-x|\tau})\frac{\kappa(t,x,\frac{x'-x}{|x'-x|})\kappa(t',x',-\frac{x'-x}{|x'-x|})}{|x'-x|^2} \\
&= (e^{i|z|\tau} + e^{-i|z|\tau})\frac{\kappa(t,x,\frac{z}{|z|})\kappa(t',x+z,-\frac{z}{|z|})}{|z|^2}
\end{aligned}$$

Here, we used $z = x' - x$. We take Fourier transform in z variable to get
(58)
$$N_\kappa f(t,x) = C\int_{\mathbb{R}^3}\int_{\mathbb{R}}\int_{\mathbb{R}^3}\int_{\mathbb{R}} e^{i(t-t')\tau}e^{i(x'-x)\xi}A(t,x,t',\tau,\xi)f(t',x')d\tau d\xi dt'dx'$$
where
$$A(t,x,t',\tau,\xi) = \int_{\mathbb{R}^3}(e^{-iz'\xi - i|z'|\tau} + e^{-iz'\xi+i|z'|\tau})\frac{\kappa(t,x,\frac{z'}{|z'|})\kappa(t',x+z',-\frac{z'}{|z'|})}{|z'|^2}dz'$$
Let $z' = r\theta$ in polar coordinate. We get
(59)
$$A(t,x,t',\tau,\xi) = \int_{\mathbb{S}^2}\int_0^\infty (e^{-ir\theta\cdot\xi - ir\tau} + e^{-ir\theta\cdot\xi+ir\tau})\kappa(t,x,\theta)\kappa(t',x+r\theta,-\theta)drd\theta$$

Without loss of generality, we assume that $\kappa(t',x+r\theta,-\theta)$ is compactly supported in r. Using Fourier transform in r variable and integration by parts, we see that

$$A(t,x,t',\tau,\xi) = \int_{\mathbb{S}^2} F(t,t',x,\theta\cdot\xi+\tau,\theta) + F(t,t',x,\theta\cdot\xi-\tau,-\theta)d\theta \tag{60}$$

where $F(t,t',x,\sigma,\theta)$ is smooth and compactly supported in t,t',x variables. Furthermore, for $|\sigma|$ large and $k \geq 0$, we have

$$|\partial_\sigma^k F| \leq C_k|\sigma|^{-k-1} \tag{61}$$

where C_k depends on k. We compute the integral in (60) in θ. It suffices to consider $\xi = |\xi|(0,0,1)$ then $\theta\cdot\xi = |\xi|\theta_3$. We let $v = |\xi|\theta_3$, and use $w^2+\theta_3^2 = 1$ to get

$$\begin{aligned} A(t,x,t',\tau,\xi) = &\int_{-1}^1\int_{-1}^1(F(t,t',x,\theta_3|\xi|+\tau,\theta) \\ &+F(t,t',x,\theta_3|\xi|-\tau,-\theta))(1-\theta_3^2)^{\frac{1}{2}}d\theta_3 dw \end{aligned} \tag{62}$$

Now that we computed the amplitude function in (58), we are ready to show that the kernel of N_κ is a paired Lagrangian distribution. We will transform (58) to the model form (54). We let $s = \tau - |\xi|$ and write

$$K(t,x,t',x') = \int_{\mathbb{R}^{n+1}} e^{i(t-t')(s+|\xi|)+i(x-x')\cdot\xi}A(t,x,t',s+|\xi|,\xi)dsd\xi \tag{63}$$

Consider the simplectic change of variables on $T^*\mathbb{R}^{3+1}$
$$\tilde{x} = x - (t-t')\xi/|\xi|, \quad \tilde{t} = t-t', \quad s = s, \quad \xi = \xi.$$
We can choose an Fourier integral operator with symbol of order 0 which quantizes the symplectic change of variable to transform K to

$$K(\tilde{t},\tilde{x},t',x') = \int_{\mathbb{R}^{n+1}} e^{i\tilde{t}s+i\tilde{x}\cdot\xi}A(t,x,t',s+|\xi|,\xi)dsd\xi \tag{64}$$

modulo a smooth term. We claim that $A(t,x,t',s+|\xi|,\xi)$ satisfies the product type estimate with $p=-3/2, l=1/2$. From (62), we have

(65)

$$A(t,x,t',s+|\xi|,\xi) = \int_{\mathbb{S}^1}\int_{-1}^{1} F(t,t',x,(\alpha+1)|\xi|+s,\alpha,w)(1-\alpha^2)^{\frac{1}{2}}d\alpha dw$$
$$+\int_{\mathbb{S}^1}\int_{-1}^{1} F(t,t',x,(\alpha-1)|\xi|-s,-\alpha,-w))(1-\alpha^2)^{\frac{1}{2}}d\alpha dw$$

First, for $|\xi|\le C|s|, |s|\ge 1$, we use the estimate (61) to get

$$|A(t,x,t',s+|\xi|,\xi)|\le C|s|^{-1}$$

One can verify the same estimate for Qk where Q is the finite product of differential operators of the form sD_s, sD_{ξ_m}. For $|s|\le C|\xi|, |\xi|\ge 1$, we have

$$|A(t,x,t',s+|\xi|,\xi)|\le C|\xi|^{-1}$$

and one can verify the estimate for Qk where Q is the finite product of differential operators of the form $D_s, sD_s, \xi_j D_{\xi_m}$. So K is a paired Lagrangian distribution according to (54).

Finally, we compute the principal symbol on $\Lambda_1\backslash\Lambda_0$. For this purpose, we can actually use the kernel representation in Proposition 6.1 and find that

(66)
$$N_\kappa(t,x,t',x') = \int_{\mathbb{R}} \frac{\kappa(t,x,\frac{x'-x}{|x'-x|})\kappa(t',x',\frac{x-x'}{|x-x'|})}{|t-t'|^2} \cdot(e^{i(t-t'-|x-x'|)\tau}+e^{i(t-t'+|x-x'|)\tau})d\tau$$

which is valid for $t'\neq t$. This gives an oscillatory representation of the Fourier integral operator on Λ_1 with a real phase function, see [19, Chapter VI]. We see that the symbol is non-vanishing and positive because κ is positive. □

We derive the Sobolev estimates for L_κ and L_κ^*.

Proposition 6.3. *The weighted light ray transform L_κ in* (49) *is bounded from $H^s_{\text{comp}}(\mathbb{R}^{3+1})$ to $H^{s+1/2}_{\text{loc}}(\mathbb{R}^3\times\mathbb{S}^2)$. Its adjoint L_κ^* is bounded from $H^s_{\text{comp}}(\mathbb{R}^3\times\mathbb{S}^2)$ to $H^{s+1/2}_{\text{loc}}(\mathbb{R}^{3+1})$.*

Proof. Because Λ_1 is the flow out of $\Lambda_0\cap\Lambda_1$ see above Proposition 6.2, we can apply Proposition 5.6 of [3] to get the estimates for N_κ which further gives estimates for L_κ, L_κ^*. □

7. Analysis of the composition

We follow the approach in [22] to analyze the composition of the weighted light ray transform and the parametrix of the Cauchy problem. We briefly recall the parametrix of the Cauchy problem (6). The operator $P(z,\partial)$ in (5) is strictly hyperbolic of multiplicity one with respect to the Cauchy hypersurfaces $\mathcal{M}_t$, see Definition 5.1.1 of [5]. This means that all bi-characteristic curves of P are transversal to $\mathcal{M}_t$ and for $(\bar z,\bar\zeta)\in T^*\mathcal{M}_t\backslash 0$

$$\mathcal{P}(\bar z,\zeta)=0,\quad \zeta|_{T_{\bar z}\mathcal{M}_t}=\bar\zeta$$

has exactly one solution. It is convenient to use $D_j = -\imath\partial_j, j = 0, 1, 2, \cdots, n$ in which $\imath^2 = -1$. Also, it is convenient to work on a larger set. We let $\mathcal{N} = (0, \tilde{T}) \times \mathbb{R}^3$ for $\tilde{T} > T$ and consider the Cauchy problem on $\mathcal{N}$

$$\begin{aligned} P(z, D)u(z) = 0, \text{ on } \mathcal{N} \\ u = f_1, D_t u = f_2 \text{ on } \mathcal{N}_0. \end{aligned} \tag{67}$$

We use Duistermaat-Hörmander's parametrix construction, see e.g. [5]. The restriction operator $\rho_0 : C^\infty(\mathcal{N}) \to C^\infty(\mathcal{N}_0)$ is an FIO in $I^{1/4}(\mathcal{N}, \mathcal{N}_0; C_0)$ with canonical relation

$$C_0 = \{(z, \zeta, \bar{z}, \bar{\zeta}) \in T^*\mathcal{N}\backslash 0 \times T^*\mathcal{N}_0\backslash 0 : \bar{z} = z, \bar{\zeta} = \zeta|_{T_{\bar{z}}\mathcal{N}_0}\} \tag{68}$$

We consider the canonical relation C_{wv} defined by

$$\begin{aligned} C_{wv} = \{(w, \iota, \bar{z}, \bar{\zeta}) \in T^*\mathcal{N}\backslash 0 \times T^*\mathcal{N}_0\backslash 0 : (w, \iota) \text{ is on the bicharacteristic} \\ \text{strip through some } (\bar{z}, \zeta) \text{ such that } \bar{\zeta} = \zeta|_{T_{\bar{z}}\mathcal{N}_0} \text{ and } \mathcal{P}(\bar{z}, \zeta) = 0\} \end{aligned} \tag{69}$$

The next result is straight forward from Theorem 5.1.2 of [5].

Proposition 7.1. *There exists $E_1 \in I^{-1/4}(\mathcal{N}, \mathcal{N}_0; C_{wv}), E_2 \in I^{-5/4}(\mathcal{N}, \mathcal{N}_0; C_{wv})$ such that*

$$\begin{aligned} P(z, D)E_k \in C^\infty(\mathcal{N}), \quad k = 1, 2 \\ \rho_0 E_1 - \mathrm{Id} \in C^\infty(\mathcal{N}_0), \quad \rho_0 E_2 \in C^\infty(\mathcal{N}_0) \\ \rho_0 D_t E_1 \in C^\infty(\mathcal{N}_0), \quad \rho_0 D_t E_2 - \mathrm{Id} \in C^\infty(\mathcal{N}_0) \end{aligned} \tag{70}$$

Now we represent the solution of (67) as $u = E_1 f_1 + E_2 f_2$ modulo a smooth term. It is natural to decompose C_{wv} as the disjoint union of C^+_{wv} and C^-_{wv} which are

$$\begin{aligned} C^\pm_{wv} = \{(w, \iota, \bar{z}, \bar{\zeta}) \in T^*\mathcal{N}\backslash 0 \times T^*\mathcal{N}_0\backslash 0 : \iota \text{ is future/past} \\ \text{pointing light-like and lies on the bicharacteristic strip through} \\ \text{some } (\bar{z}, \zeta) \text{ such that } \bar{\zeta} = \zeta|_{T_{\bar{z}}\mathcal{N}_0} \text{ and } \mathcal{P}(\bar{z}, \zeta) = 0\} \end{aligned} \tag{71}$$

We can decompose (for $k = 1, 2$)

$$E_k = E^+_k + E^-_k, \quad E^\pm_k \in I^{1-k-1/4}(\mathcal{N}, \mathcal{N}_0; C^\pm_{wv}).$$

We need the relation of the principal symbols of $E^\pm_1, E^\pm_2$. We remark that the Maslov bundle and the half density bundle can be trivialized because the Lagrangians involved allow global parametrization. We will not show these factors in the notations.

Lemma 7.2 (Lemma of [22])**.** *Let $e^\pm_k, k = 1, 2$ be the principal symbol of $E^\pm_k$ on $\Lambda^\pm = (C^\pm_{wv})'$ respectively. Suppose that the sub-principal symbol of $P(z, \partial)$ is purely imaginary, in which case $P(z, \partial)$ is of the form*

$$P(z, \partial) = \Box + \sum_{j=0}^{n} A_j(z)\partial_j + B(z) \tag{72}$$

where $A_j(z)$ are real valued smooth functions. Then $e_k^\pm, k = 1,2$ are real valued and

$$e_1^+ > 0, \quad e_2^+ > 0, \quad e_1^- > 0, \quad e_2^- < 0.$$

Now we can outline the steps for inverting the composition $L_\kappa E$ where $E = E_k^\pm, k = 1,2$ in Proposition 7.1. The idea is to consider the normal operator $(L_\kappa E)^*(L_\kappa E)$ and fine-tune it so the operator is well-behaved. Let $\chi_{[0,T]}(t), t \in \mathbb{R}$ be the characteristic function for $[0,T]$ in $\mathbb{R}$. Let χ be a smooth function supported in $(T,T') \times \mathbb{R}^3$. See Figure 2. We analyze $E^*\chi N_\kappa \chi_{[0,T]} E$ and show it is an elliptic pseudo-differential operator on $\mathcal{N}_0$. There are three main ingredients.

(i) As $\chi \cdot \chi_{[0,T]} = 0$, we know that $\chi N_\kappa \chi_{[0,T]} \in I^{-3/2}(\mathbb{R}^4, \mathbb{R}^4; \Lambda_1)$ at least when the characteristic function $\chi_{[0,T]}$ were smooth. Note that the role of χ is to keep the kernel of N_κ away from the diagonal Λ_0 where the principal symbol is singular.

(ii) Let $\Lambda_\pm = (C_{wv}^\pm)'$. It was shown in [22] that Λ_1 intersect $\Lambda_\pm$ cleanly with excess one so the composition $\chi N_\kappa \chi_{[0,T]} E \in I^*(\mathcal{N}, \mathcal{N}_0; C_{wv})$ as a result of Duistermaat-Guillemin's clean FIO calculus with the order $*$ to be determined. For this, we need to address some issue caused by the characteristic function.

(iii) We can compose the operator in (ii) with E^* by using clean FIO calculus again to conclude that $E^*\chi N_\kappa \chi_{[0,T]} E \in \Psi^*(\mathcal{N}_0)$. The operator can be shown to be elliptic and a parametrix can be constructed.

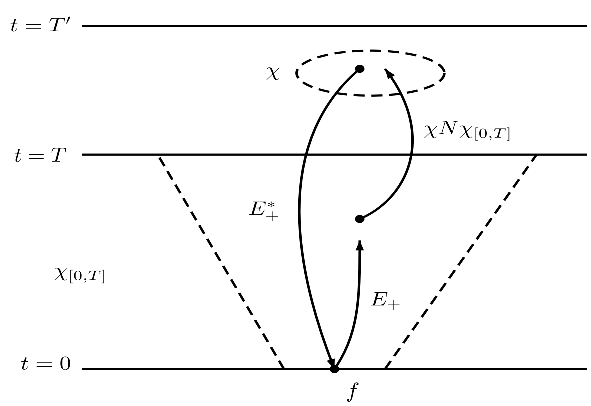

FIGURE 2. Illustration of the composition of operators

We consider the compositions in (ii) and (iii). In fact, we will include some pseudo-differential operators such as those showed up in Section 2.

Lemma 7.3. *Let $A(D)$ be a pseudo-differential operator on $\mathcal{M}$ of order m. We assume that the principal symbol of $A(D)$ is non-vanishing on $\Lambda_\pm$. Then the composition $\chi N_\kappa \chi_{[0,T]} A(D) E_k^\pm \in I^{-1/4-k+m}(\mathcal{N}, \mathcal{N}_0; C_{wv}^\pm)$ and the principal symbol is non-vanishing.*

Proof. We explain the idea by replacing $\chi_{[0,T]}$ by a smooth function $\tilde{\chi}$ compactly supported in $[0,T]$. Note that $A(D)E_k^\pm \in I^{1-k+m}(\mathcal{N}, \mathcal{N}_0; C_{wv}^\pm)$ with non-vanishing principal symbol. Because $\chi(t)\tilde{\chi}(t) = 0$, we know that $\chi N_\kappa \tilde{\chi} \in I^{-3/2}(\mathcal{N}, \mathcal{N}; \Lambda_1)$. It is proved in Lemma 6.1 of [22] that Λ_1 intersects $\Lambda_\pm$ cleanly. One can apply the clean FIO calculus [8, Theorem 25.2.3] directly to see that $\chi N_\kappa \tilde{\chi} A(D)E_k^\pm \in I^{-3/2+1/4+1-k+m}(\mathcal{N}, \mathcal{N}_0; C_{wv}^\pm)$. For $p = (t, x, \tau, \xi, y, \eta) \in \Lambda^\pm$, let C_p be the fiber over p in $T^*\mathcal{M} \times T^*\mathcal{M} \times T^*\mathcal{N}_0$ which is connected and compact. Then the principal symbol of the composition at p is given by

$$\int_{C_p} \sigma(\chi N_\kappa \tilde{\chi})(t, x, \tau, \xi, t', x', \tau', \xi')\sigma(A(D)E_k^\pm)(t', x', \tau', \xi', y, \eta) \tag{73}$$

where $\sigma(\chi N_\kappa \tilde{\chi}), \sigma(A(D)E_k^\pm)$ denote the principal symbols of $\chi N_\kappa \tilde{\chi}, E_k^\pm$ respectively and the integration is over the fiber C_p, see [8, Theorem 25.2.3]. With proper choice of the phase function (modulo the Maslov factor), both symbols are real valued and non-vanishing on the fiber, see Lemma 7.2 and the proof of Proposition 6.2. We see that the principal symbol of the composition is non-vanishing.

Next, for the characteristic function, the difference is that the fiber C_p is connected but not compact. The arguments in [8, Theorem 25.2.3] still work provided that the integral (73) is finite. This was justified in [22, Lemma 6.2] without the operator $A(D)$. But the proof there can be repeated line by line by changing $E_k^\pm$ to $A(D)E_k^\pm$. □

Lemma 7.4. *Let $A(D)$ be as in Lemma 7.3. For $j, k = 1, 2$, we have*

(1) $E_j^{\pm,*}\chi N_\kappa \chi_{[0,T]} A(D)E_k^\pm \in \Psi^{1-j-k+m}(\mathcal{N}_0)$ are elliptic.

(2) $E_j^{+,*}\chi N_\kappa \chi_{[0,T]} A(D)E_k^-, E_j^{-,*}\chi N_\kappa \chi_{[0,T]} A(D)E_k^+$ are smoothing operators on $\mathcal{N}_0$.

Proof. First, $E_j^{\pm,*} \in I^{-1/4+1-j}(\mathcal{N}_0, \mathcal{N}; C_{wv}^{\pm,-1})$ and $\chi N_\kappa \chi_{[t_0,t_1]} A(D)E_k^\pm \in I^{-1/4-k+m}(\mathcal{N}, \mathcal{N}_0; C_{wv}^\pm)$. Let $\Lambda^\pm = (C_{wv}^\pm)'$ and $\Lambda^{\pm,-1} = (C_{wv}^{\pm,-1})'$. It is proved in Lemma 6.3 of [22] that $\Lambda^{\pm,-1}$ intersect $\Lambda^\pm$ cleanly with excess one. Now we can use the clean FIO calculus [8, Theorem 25.2.3] to conclude that $E_j^{\pm,*}\chi N_\kappa \chi_{[0,T]} A(D)E_k^\pm \in \Psi^{1-j-k+m}(\mathcal{N})$. As both principal symbols of $E_j^{\pm,*}$ and $\chi N_\kappa \chi_{[0,T]} A(D)E_k^\pm$ are real and non-vanishing (modulo the Maslov factor), the principal of the composition is the integration of the product of principal symbols so is also non-vanishing. This proves part (1). Part (2) can be seen from a wave front set analysis using e.g. [5, Theorem 1.3.7]. □

8. Proof of Theorem 1.2

We will prove a stronger version of Theorem 1.2.

Theorem 8.1. (Assumption) *Let f be the solution of* (6) *on $\mathcal{M}$ with Cauchy data $f_1 \in H^2(\mathcal{M}_0), f_2 \in H^1(\mathcal{M}_0), s \in \mathbb{R}$ supported in a compact set $\mathcal{X}$ of $\mathcal{M}_0$ such that f is supported in $\mathcal{V}$. Suppose that*

(1) the coefficients $A_j(z)$ in (5) *are real valued smooth functions.*

(2) $A(D)$ is a pseudo-differential operator as in Lemma 7.3

(3) When $\sigma = k = 0$, $LA(D)f = 0$ implies $f = 0$.

Let u be the solution of (1) *with zero initial condition and source $A(D)f$.*

(Conclusion) *There exists an open dense set $\mathcal{U}$ of $C^\infty(\mathcal{V} \times \mathbb{S}^2) \times C^6(\mathcal{V} \times \mathbb{S}^2 \times \mathbb{S}^2)$ such that for $(\sigma, k) \in \mathcal{U}$, f is uniquely determined by u_T. Furthermore, there exists $C > 0$ such that*

$$\|u\|_{H^{2+m}(\mathcal{M})} \leq C\|(f_1, f_2)\|_{H^{2+m}(\mathcal{M}_0)\times H^{1+m}(\mathcal{M}_0)} \leq C\|u_T\|_{H^{5/2}(\mathcal{C})}$$

We make a few remarks. First, because the parametrix construction in Section 7 is microlocal, it is convenient to work with smooth σ. However, it is likely to lower the regularity requirement if one is willing to keep track of the dependency of the constant C on σ throughout the arguments. Next, about assumption (3). This condition only involves the Minkowski light ray transform or the transport regime ($\sigma = k = 0$). Let's consider the example in Section 2 in which $A(D) = \frac{\partial}{\partial t} + B(z)$. When $\sigma = k = 0$, $B(z) = 0$ so $A(D) = \partial_t$. Note that f satisfies (18). If $L(A(D)f) = 0$, because $A(D)f$ is compactly supported in $\mathbb{R}^4$ and say $A(D)f \in L^2_{\text{comp}}(\mathcal{M})$, we know from the injectivity of Minkowski light ray transform that $A(D)f = 0$ so $\partial_t f = 0$. Using the Bardeen's equation (18), we get

$$\Delta f + B_0(t)f = 0$$

for $t \in [0, T]$ almost everywhere. Fixed t, using the fact that f is compactly supported in x and that Δ has no L^2 eigenvalues, we conclude that $f(t, x) = 0$. Thus (3) of Theorem 8.1 is satisfied.

We outline the proof of Theorem 8.1. We start with (26)

$$Xf = \rho(\mathrm{Id} - \mathrm{Id} + (\mathrm{Id} + T_1^{-1}K)^{-1})T_1^{-1}f = L_\kappa f + Ef$$

where

$$E = \rho(-\mathrm{Id} + (\mathrm{Id} + T_1^{-1}K)^{-1})T_1^{-1}$$

Now we write

$$f = E_1 f_1 + E_2 f_2$$

where E_1, E_2 are the parametrix of the Cauchy problem. We consider

$$XAf = L_\kappa A E_1 f_1 + L_\kappa A E_2 f_2 + EA(E_1 f_1 + E_2 f_2) \tag{74}$$

Our goal is to convert the right hand side to identity plus compact operators. Roughly speaking, we will apply $E_k^{\pm,*}\chi L_\kappa^*, k = 1, 2$ to (74) and manipulate to get f_1, f_2. As a result, we need to analyze operators of two types. First

$$P_1 = E_i^{\pm,*}\chi L_\kappa^* L_\kappa A E_j^{\pm}, \quad i, j = 1, 2, \tag{75}$$

and second

$$P_2 = E_i^{\pm,*}\chi L_\kappa^* EAE_j, \quad i, j = 1, 2. \tag{76}$$

For P_1, we will use results in Section 7 to find a parametrix which are pseudo-differential operators of order -1 on $\mathcal{M}_0$. In fact, we will also need to replace σ by $\lambda\sigma, \lambda \in \mathbb{C}$ and consider the holomorphic dependency on λ in order to apply analytic Fredholm theory, because we do not know the injectivity of the weighted light ray transform in general. For P_2, we show the compactness as in Section 5. Finally, we finish the proof using analytic Fredholm theorem. We split the proof into three subsections.

8.1. **Analysis of P_1.** We first ignore the terms with operator E in (74) and focus on the parametrix construction. We apply χL_κ^* to $L_\kappa Af$ to get

$$\begin{aligned}\chi N_\kappa Af &= \chi N_\kappa \chi_{[0,T]} AE_1^+ f_1 + \chi N_\kappa \chi_{[0,T]} AE_2^+ f_2 \\ &+ \chi N_\kappa \chi_{[0,T]} AE_1^- f_1 + \chi N_\kappa \chi_{[0,T]} AE_2^- f_2.\end{aligned} \tag{77}$$

Now we apply $E_1^{+,*}$ and use Lemma 7.4 to get

$$E_1^{+,*}\chi N_\kappa Af = E_1^{+,*}\chi N_\kappa \chi_{[0,T]} AE_1^+ f_1 + E_1^{+,*}\chi N_\kappa \chi_{[0,T]} AE_2^+ f_2 + R_1 f_1 + R_2 f_2 \tag{78}$$

with R_1, R_2 smoothing operators. In the following, we use $R_j, j = 1, 2$ to denote generic smoothing operators which may change line by line. From Lemma 7.4 part (1), we see that $E_1^{+,*}\chi N_\kappa \chi_{[0,T]} AE_1^+ \in \Psi^{-1+m}(\mathcal{M}_0)$ and $E_1^{+,*}\chi N_\kappa \chi_{[0,T]} AE_2^+ \in \Psi^{-2+m}(\mathcal{M}_0)$.

On the other hand, we apply $E_1^{-,*}$ to (77) to get

$$E_1^{-,*}\chi N_\kappa Af = E_1^{-,*}\chi N_\kappa \chi_{[0,T]} AE_1^- f_1 + E_1^{-,*}\chi N_\kappa \chi_{[0,T]} AE_2^- f_2 + R_1 f_1 + R_2 f_2 \tag{79}$$

From Lemma 7.4 part (1), we see that $E_1^{-,*}\chi N_\kappa \chi_{[0,T]} AE_1^- \in \Psi^{-1+m}(\mathcal{M}_0)$ and $E_1^{-,*}\chi N_\kappa \chi_{[0,T]} AE_2^- \in \Psi^{-2+m}(\mathcal{M}_0)$. Without loss of generality, we assume that the principal symbol of A on $\Lambda_\pm$ is positive. Then it follows from Lemma 7.2, the composition results Lemma 7.3 and 7.4, and the positivity of the symbol of N_κ on $\Lambda_1\backslash\Lambda_0$ in Proposition 6.1 that

$$\begin{aligned}&\sigma(E_1^{+,*}\chi N_\kappa \chi_{[0,T]} AE_1^+) > 0, \quad \sigma(E_1^{-,*}\chi N_\kappa \chi_{[0,T]} AE_1^-) > 0\\ &\sigma(E_1^{+,*}\chi N_\kappa \chi_{[0,T]} AE_2^+) > 0, \quad \sigma(E_1^{-,*}\chi N_\kappa \chi_{[0,T]} AE_2^-) < 0.\end{aligned}$$

Let $Q^+, Q^- \in \Psi^{1-m}(\mathcal{M}_0)$ be parametrices for $E_1^{+,*}\chi N_\kappa \chi_{[0,T]} AE_1^+$ and $E_1^{-,*}\chi N_\kappa \chi_{[0,T]} AE_1^-$ respectively. We know that the principal symbols of $Q^\pm$ are positive. Applying $Q^\pm$ to (78), (79), we get

$$Q^+ E_1^{+,*}\chi N_\kappa Af = f_1 + B_+ f_2 + R_1 f_1 + R_2 f_2 \tag{80}$$

$$Q^- E_1^{-,*}\chi N_\kappa Au = f_1 + B_- f_2 + R_1 f_1 + R_2 f_2 \tag{81}$$

where

$$B_+ = Q_+ E_1^{+,*}\chi N_\kappa \chi_{[0,T]} AE_2^+, \quad B_- = Q_- E_1^{-,*}\chi N_\kappa \chi_{[0,T]} AE_2^-$$

From (80), (81), we get

$$Q^+ E_1^{+,*}\chi N_\kappa Af - Q^- E_1^{-,*}\chi N_\kappa Af = (B_+ - B_-) f_2 + R_1 f_1 + R_2 f_2$$

Note that $B_\pm \in \Psi^{-1}(\mathcal{M}_0)$ are elliptic. Also, the principal symbol of B_+ is positive but the principal symbol of B_- is negative. Thus $B_+ - B_- \in \Psi^{-1}(\mathcal{M}_0)$

is elliptic. Let $W \in \Psi^1(\mathcal{M}_0)$ be a parametrix for $B_+ - B_-$. We get

$$WQ^+E_1^{+,*}\chi N_\kappa Af - WQ^-E_1^{-,*}\chi N_\kappa Af = f_2 + R_1 f_1 + R_2 f_2 \tag{82}$$

So we solved f_2 up to smooth terms. We can use f_2 for example in (80) to get

$$\begin{aligned} Q^+E_1^{+,*}\chi N_\kappa Af - B_+(WQ^+E_1^{+,*}\chi N_\kappa Af - WQ^-E_1^{-,*}\chi N_\kappa Af) \\ = f_1 + R_3 f_1 + R_4 f_2 \end{aligned} \tag{83}$$

where R_3, R_4 are smoothing operators. We are done with the parametrix construction.

8.2. **Compactness of P_2.** We repeat the construction with the terms involving operator E in (74). Using (74) and (82), we arrive at

$$\begin{aligned} WQ^+E_1^{+,*}\chi N_\kappa Af - WQ^-E_1^{-,*}\chi N_\kappa Af = f_2 + R_1 f_1 + R_2 f_2 \\ +WQ^+E_1^{+,*}\chi L_\kappa^* EA(E_1 f_1 + E_2 f_2) - WQ^-E_1^{-,*}\chi L_\kappa^* EA(E_1 f_1 + E_2 f_2) \\ = f_2 + R_1 f_1 + R_2 f_2 + R_1' f_1 + R_2' f_2 \end{aligned} \tag{84}$$

where

$$\begin{aligned} R_1' &= WQ^+E_1^{+,*}\chi L_\kappa^* EAE_1 - WQ^-E_1^{-,*}\chi L_\kappa^* EAE_1 \\ R_2' &= WQ^+E_1^{+,*}\chi L_\kappa^* EAE_2 - WQ^-E_1^{-,*}\chi L_\kappa^* EAE_2 \end{aligned} \tag{85}$$

Using (83) and (84), we get

$$\begin{aligned} Q^+E_1^{+,*}\chi N_\kappa Af - B_+(WQ^+E_1^{+,*}\chi N_\kappa Af - WQ^-E_1^{-,*}\chi N_\kappa Af) \\ = f_1 + R_3 f_1 + R_4 f_2 + Q^+E_1^{+,*}\chi L_\kappa^* EA(E_1 f_1 + E_2 f_2) \\ -B_+R_1' f_1 + B_+R_2' f_2 = f_1 + R_3 f_1 + R_4 f_2 + R_3' f_1 + R_4'' f_2 \end{aligned} \tag{86}$$

where

$$\begin{aligned} R_3' &= -B_+R_1' + Q^+E_1^{+,*}\chi L_\kappa^* EAE_1 \\ R_4' &= B_+R_2' + Q^+E_1^{+,*}\chi L_\kappa^* EAE_2 \end{aligned} \tag{87}$$

We write (84) and (86) in the matrix form as

$$\begin{aligned} \begin{pmatrix} Q^+E_1^{+,*}\chi N_\kappa Af - B_+(WQ^+E_1^{+,*}\chi N_\kappa Af - WQ^-E_1^{-,*}\chi N_\kappa Af) \\ WQ^+E_1^{+,*}\chi N_\kappa Af - WQ^-E_1^{-,*}\chi N_\kappa Af \end{pmatrix} \\ = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} + \begin{pmatrix} R_3 + R_3' & R_4 + R_4' \\ R_1 + R_1' & R_2 + R_2' \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \end{aligned} \tag{88}$$

We already know that $R_j, j = 1, 2, 3, 4$ are smoothing operators. We show that $R_j', j = 1, 2, 3, 4$ are compact operators among suitable spaces.

Among the operators $R_j', j = 1, 2, 3, 4$, there is a common operator $L_\kappa^* E$ see (85), (87). We can prove as in Section 5 that $L_\kappa^* E$ is compact from $H^2(\mathcal{M})$ to $H^3(\mathcal{M})$ (or $\partial_t L_\kappa^* E, \partial_x L_\kappa^* E$ are compact on $H^2(\mathcal{M})$). Note that the additional weight factor κ in T_1^{-1} was considered in Section 5. For the other components of $R'_j, j = 1,2,3,4$, we know that $WQ^\pm \in \Psi^{2-m}(\mathcal{M}_0)$ and $Q^+, B_+WQ^\pm \in \Psi^{1-m}(\mathcal{M}_0)$ so they are bounded operators as

$$\begin{aligned} WQ^\pm &: H^s_{\text{comp}}(\mathcal{M}_0) \to H^{s-2+m}_{\text{loc}}(\mathcal{M}_0) \\ Q^+, B_+WQ^\pm &: H^s_{\text{comp}}(\mathcal{M}_0) \to H^{s-1+m}_{\text{loc}}(\mathcal{M}_0) \end{aligned} \tag{89}$$

Also, for $j = 1, 2$

$$E_j^* : H^s_{\text{comp}}(\mathcal{M}) \to H^{s+j-1}_{\text{loc}}(\mathcal{M}_0), \quad E_j : H^s_{\text{comp}}(\mathcal{M}_0) \to H^{s+j-1}_{\text{loc}}(\mathcal{M})$$

Now we look at R'_1 on $H^{2+m}(\mathcal{M}_0)$ which can be decomposed as

$$H^{2+m}(\mathcal{M}_0) \xrightarrow{AE_1} H^2(\mathcal{M}) \xrightarrow{L^*_\kappa E} H^3(\mathcal{M}) \xrightarrow{WQ^\pm E_1^{+,*}\chi} H^{1+m}(\mathcal{M}_0)$$

in which all the operators are bounded and $L^*_\kappa E$ is compact in addition. So R'_1 is compact. We use similar diagram below. Next, for R'_2 on $H^{1+m}(\mathcal{M}_0)$, we have

$$H^{1+m}(\mathcal{M}_0) \xrightarrow{AE_2} H^2(\mathcal{M}) \xrightarrow{L^*_\kappa E} H^3(\mathcal{M}) \xrightarrow{WQ^\pm E_1^{+,*}\chi} H^{1+m}(\mathcal{M}_0)$$

Similarly, for R'_3, R'_4, we have

$$R'_3 : H^{2+m}(\mathcal{M}_0) \xrightarrow{AE_1} H^2(\mathcal{M}) \xrightarrow{L^*_\kappa E} H^3(\mathcal{M}) \xrightarrow{Q^\pm E_1^{+,*}\chi} H^{m+2}(\mathcal{M}_0)$$

$$R'_4 : H^{1+m}(\mathcal{M}_0) \xrightarrow{AE_2} H^2(\mathcal{M}) \xrightarrow{L^*_\kappa E} H^3(\mathcal{M}) \xrightarrow{Q^\pm E_1^{+,*}\chi} H^{m+2}(\mathcal{M}_0)$$

We reached the conclusion that $R'_j, j = 1,2,3,4$ are compact operators.

8.3. **Completion of the proof.** From (88), we obtain an operator $\mathbf{Id} + \mathbf{R}$ on $H^{s+1}(\mathcal{X}) \times H^s(\mathcal{X})$ where $\mathcal{X}$ is a compact subset of $\mathcal{M}$ and

$$\mathbf{R} = \begin{pmatrix} R_3 + R'_3 & R_4 + R'_4 \\ R_1 + R'_1 & R_2 + R'_2 \end{pmatrix} \tag{90}$$

is compact. To apply the analytic Fredholm theorem, we let $\lambda \in \mathbb{C}$ and replace σ, k in (1) by $\lambda\sigma, \lambda k$. We denote the corresponding operator by $\mathbf{R}(\lambda)$.

Observe that the weight κ in the light ray transform L_κ (49) is now holormophic in λ. Then the calculation of kernel in Proposition 6.1 shows that the kernel of the normal operator $N_\kappa = L^*_\kappa L_\kappa$ is holomorphic in λ. In particular, the proof of Proposition 6.2 shows that N_κ is an FIO on Λ_1 with symbol holomorphic in λ. If we follow the construction in Section 7, we see that the operators in Lemma 7.3 and Lemma 7.4 are holomorphic in λ. Finally, going through the construction in Subsection 8.1, we see that the operators on the left hand side of (88) are holomorphic in λ and the remainder terms R_1, R_2, R_3, R_4 are holomorphic in λ. We remark that it suffices to consider the parametrix construction up to a finite order residue term, that is with $R_j, j = 1,2,3,4$ belonging to $\Psi^{-N}(\mathcal{M}_0)$ for N sufficiently large. Then it is clear that R_j are holomorphic in λ and compact on suitable Sobolev spaces.

Next, in the expression (85) and (87), we know that $L^*_\kappa E$ is meromorphic in λ as shown in Section 5. The other operators in (85) and (87) are holomorphic in λ, and $R'_j, j = 1,2,3,4$ are meromorphic in λ. This proves that $\mathbf{R}(\lambda)$ is meromorphic in λ.

When $\lambda = 0$, we see that (88) is reduced to

$$\begin{pmatrix} Q^+E_1^{+,*}\chi NAf - B_+(WQ^+E_1^{+,*}\chi NAf - WQ^-E_1^{-,*}\chi NAf) \\ WQ^+E_1^{+,*}\chi NAf - WQ^-E_1^{-,*}\chi NAf \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} + \begin{pmatrix} R_3 & R_4 \\ R_1 & R_2 \end{pmatrix}\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \tag{91}$$

where $N = L^*L$ with L the Minkowski light ray transform. We know that $L^* : H^s_{\text{comp}}(\mathcal{C}) \to H^{s+\frac{1}{2}}_{\text{loc}}(\mathbb{R}^{3+1})$ is bounded and

$$E_1^{\pm,*}\chi L^* : H^s_{\text{comp}}(\mathcal{C}) \to H^{s+1/2}_{\text{loc}}(\mathcal{M})$$

is bounded. Thus using estimates (89) and (91), we get for $\rho \in \mathbb{R}$ that

$$\begin{aligned} \|f_1\|_{H^{2+m}(\mathcal{M}_0)} &\leq C\|LA(D)f\|_{H^{1+3/2}(\mathcal{C})} + C_\rho\|f_1\|_{H^{s+1+\rho}} + C_\rho\|f_2\|_{H^{s+\rho}} \\ \|f_2\|_{H^{1+m}(\mathcal{M}_0)} &\leq C\|LA(D)f\|_{H^{1+3/2}(\mathcal{C})} + C_\rho\|f_1\|_{H^{s+1+\rho}} + C_\rho\|f_2\|_{H^{s+\rho}} \end{aligned} \tag{92}$$

From assumption (3) of Theorem 8.1, we know that $LA(D)$ is injective. We can use the argument of Theorem 1.1 of [20] to remove the last two terms in each of the equations in (92) and obtain

$$\|f_1\|_{H^{2+m}(\mathcal{M}_0)} \leq C\|LA(D)f\|_{H^{1+3/2}(\mathcal{C})}, \quad \|f_2\|_{H^{1+m}(\mathcal{M}_0)} \leq C\|LA(D)f\|_{H^{1+3/2}(\mathcal{C})}.$$

This shows the invertiblity of $\mathbf{Id}+\mathbf{R}(0)$. Thus we can apply analytic Fredholm theorem to conclude that $\mathbf{Id} + \mathbf{R}(\lambda)$ is invertible on $H^{s+1}(\mathcal{X}) \times H^s(\mathcal{X})$ for $\lambda \in \mathbb{C}\backslash\mathcal{S}$ where $\mathcal{S}$ is a discrete set. Therefore, $\mathbf{Id} + \mathbf{R}$ with $\mathbf{R}$ in (90) is invertible for σ, k in an open dense set of $C^\infty \times C^6$.

To get the stability estimate, we examine the operators in the left hand side of (88). We know that $L^*_\kappa : H^s_{\text{comp}}(\mathcal{C}) \to H^{s+\frac{1}{2}}_{\text{loc}}(\mathbb{R}^{3+1})$ is bounded, see Proposition 6.3. We obtain that

$$E_1^{\pm,*}\chi L^*_\kappa : H^s_{\text{comp}}(\mathcal{C}) \to H^{s+1/2}_{\text{loc}}(\mathcal{M})$$

is bounded. Thus using estimates (89), the invertibility of $\mathbf{Id} + \mathbf{R}$ and the estimate of $A(D)$, we get

$$\|f_1\|_{H^{2+m}(\mathcal{M}_0)} \leq C\|u_T\|_{H^{1+3/2}(\mathcal{C})}, \quad \|f_2\|_{H^{1+m}(\mathcal{M}_0)} \leq C\|u_T\|_{H^{1+3/2}(\mathcal{C})}.$$

This completes the proof of Theorem 8.1.

Proof of Theorem 1.2. We need to check (3) of Theorem 8.1 which follows from the injectivity of the Minkowski light ray transform on $L^1_{\text{comp}}(\mathbb{R}^4)$ (hence on $L^2_{\text{comp}}(\mathbb{R}^4)$), see [20, Theorem 8.1]. □

References

[1] G. Bal. *Inverse transport theory and applications.* Inverse Problems 25.5 (2009): 053001.

[2] M. Choulli, P. Stefanov. *Inverse scattering and inverse boundary value problems for the linear Boltzmann equation.* Communications in Partial Differential Equations 21.5-6 (1996): 763-785.

[3] M. de Hoop, G. Uhlmann, A. Vasy. *Diffraction from conormal singularities.* Ann. Sci. Éc. Norm. Supér.(4) 48.2 (2015): 351-408.

[4] S. Dodelson. *Modern cosmology.* Amsterdam (Netherlands): Academic Press. (2003).

[5] J. J. Duistermaat. *Fourier Integral Operators.* Vol. 130. Springer Science & Business Media, 1995.

[6] R. Durrer. *The Cosmic Microwave Background.* Cambridge University Press, Cambridge, UK, 2008

[7] D. Fajman, J. Joudioux, J. Smulevici. *A vector field method for relativistic transport equations with applications.* Analysis & PDE 10.7 (2017): 1539-1612.

[8] L. Hörmander. *The Analysis of Linear Partial Differential Operators IV: Fourier Integral Operators.* Classics in Mathematics (2009).

[9] J. Ilmavirta. *X-ray transforms in pseudo-Riemannian geometry.* The Journal of Geometric Analysis 28.1 (2018): 606-626.

[10] V. Isakov. *Inverse problems for partial differential equations.* Vol. 127. New York: Springer, 2006.

[11] L. M. Krauss, S. Dodelson, S. Meyer. *Primordial gravitational waves and cosmology.* Science 328.5981 (2010): 989-992.

[12] M. Lassas, L. Oksanen, P. Stefanov, G. Uhlmann. *The light ray transform on Lorentzian manifolds.* Communications in Mathematical Physics, 1-31. (2020)

[13] V. Mukhanov, H. Feldman, R. Brandenberger. *Theory of cosmological perturbations.* Physics reports 215.5-6 (1992): 203-333.

[14] M. Reed, B. Simon. *Modern Methods of Mathematical Physics, Vol. 1.* New York: Academic (1979).

[15] P. Stefanov. *Inverse problems in transport theory.* Inside Out: Inverse Problems and Applications, Math. Sci. Res. Inst. Publ 47 (2003): 111-131.

[16] P. Stefanov. *Support theorem for the light ray transform on analytic Lorentzian manifolds.* Proceedings of the American Mathematical Society.

[17] P. Stefanov, G. Uhlmann. *An inverse source problem in optical molecular imaging.* Analysis & PDE 1.1 (2008): 115-126.

[18] A. Torchinsky. *Real-variable methods in harmonic analysis.* Academic Press, 1986.

[19] F. Tréves. *Introduction to pseudodifferential and Fourier integral operators Volume 2: Fourier integral operators.* Springer Science & Business Media, 1980.

[20] A. Vasy, Y. Wang. *On the light ray transform of wave equation solutions.* Communications in Mathematical Physics, 2021.

[21] Y. Wang. *Microlocal analysis of the light ray transform on globally hyperbolic Lorentzian manifolds.* arXiv:2104.08576, (2021).

[22] Y. Wang. *Some integral geometry problems for wave equations.* preprint, 2021.

Yiran Wang
Department of Mathematics, Emory University
Email address: yiran.wang@emory.edu