# INTRODUCTION TO NONLINEAR WAVE EQUATIONS

JONATHAN LUK

Contents



## 1. Introduction

**Warning: This is a first draft of the lecture notes and should be used with a lot of care! The notes are updated quite frequently as the lectures progress - please check the website for updates. Also, I would appreciate any comments or corrections.**

In this course, we consider nonlinear wave equations. This can take the form of a scalar equation or a system of equations. A scalar wave equation takes the form

$$\sum_{\mu,\nu=0}^{n} \frac{1}{\sqrt{-\det g}}\partial_\mu(\sqrt{-\det g}(g^{-1})^{\mu\nu}\partial_\nu\phi) = F(\phi, \partial\phi), \tag{1.1}$$

where $g$ is a Lorentzian metric on $I \times \mathbb{R}^n$ and the unknown function $\phi : I \times \mathbb{R}^n \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval (i.e., a connected subset of $\mathbb{R}$ with at least two points), which will be thought of as a time interval. We will call this an equation in $(n+1)$-dimensions or in $n$ spatial dimensions. We will also frequently say that this is an equation in $\mathbb{R}^{n+1}$.

We say that $g$ is a Lorentzian metric if it is a symmetric $(n+1)\times(n+1)$ matrix (thus having real eigenvalues) with eigenvalues $\lambda_0 < 0 < \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$. It thus makes sense to divide by $\sqrt{-\det\, g}$ and also to define the inverse of $g$.

We say that the equation is **linear** if $g$ is independent of $\phi$ and $F$ is a linear function of both of its arguments; the equation is **semilinear** if $g$ is independent of $\phi$ and $F$ is a nonlinear function; and the equation is **quasilinear** if $g$ is a function of $\phi$ and/or $\partial\phi$.

We will be mostly interested in the *Cauchy problem* for (1.1). Suppose we label the coordinate of the interval $I$ by $t$ or $x^0$ and the coordinates of $\mathbb{R}^n$ by $(x_1, ..., x_n)$ and assume that $g$ has the property that $g(\partial_{x_i}, \partial_{x_j})$ is a strictly positive definite matrix for $i, j = 1, ...n$. We can then pose data on the $\{t = 0\}$

hypersurface, i.e., we prescribe $(\phi, \partial_t\phi)\restriction_{\{t=0\}}$ to be some appropriate functions. The *Cauchy problem* asks for a solution to (1.1) such that the restrictions of $\phi$ and $\partial_t\phi$ to the hypersurface $\{t=0\}$ are precisely the prescribed functions.

More generally, we are also interested in system of equations. Consider $\phi : I \times \mathbb{R}^n \to \mathbb{R}^m$ satisfying the equation

$$\sum_{\mu,\nu=0}^{n} \frac{1}{\sqrt{-\det g_I}}\partial_\mu(\sqrt{-\det g_I}(g_I^{-1})^{\mu\nu}\partial_\nu\phi_I) = F_I(\phi, \partial\phi), \tag{1.2}$$

where $I = 1, ..., m$.

Wave equations occurs naturally in many physical theories. We now look at a number of examples, many of which we will consider in the course.

**Example 1.1** (Linear wave equation, D'Alembert 1749)**.** Let $\phi : I \times \mathbb{R}^n \to \mathbb{R}$.

$$\Box_m\phi = -\partial_t^2\phi + \sum_{i=1}^{m}\partial_{x_i}^2\phi = 0.$$

**Example 1.2** (Maxwell's equation, Maxwell 1861-1862)**.** Let $E : I \times \mathbb{R}^3 \to \mathbb{R}^3$ and $B : I \times \mathbb{R}^3 \to \mathbb{R}^3$ be (time-dependent) vector fields representing the electric and magnetic field. The Maxwell's equations are given by

$$\begin{cases}\partial_t E = \nabla \times B\\ \partial_t B = -\nabla \times E\\ \nabla \cdot E = 0\\ \nabla \cdot B = 0.\end{cases}$$

A priori, they may not look like wave equations (they are not even second order!). However, if we differentiate the first equation by $\partial_t$, use the second and third equations, we get

$$\Box E_i = 0$$

for $i = 1, 2, 3$. Similarly,

$$\Box B_i = 0.$$

Thus, given initial data $(E, B)\restriction_{\{t=0\}}= (E^0, B^0)$ which are divergence-free, we can solve

$$\begin{cases}\Box E_i = 0, \quad \Box B_i = 0\\ (E_i, \partial_t E_i)\restriction_{\{t=0\}}= (E_i^0, (\nabla \times B^0)_i), \quad (B_i, \partial_t B_i)\restriction_{\{t=0\}}= (B_i^0, -(\nabla \times E^0)_i).\end{cases}$$

**Exercise:** Show that the solution indeed satisfies the Maxwell equations.

**Example 1.3** (Irrotational conpressible fluids, Euler 1752)**.** A fluid in $I \times \mathbb{R}^3$ is described by a vector field $v : I \times \mathbb{R}^3 \to \mathbb{R}^3$ describing the velocity of the fluid and a non-negative function $h : I \times \mathbb{R}^3 \to \mathbb{R}_{\geq 0}$ describing the enthalpy.

Define the pressure $p$ to be a function of the enthalpy $p = p(h)$ such that

(1) $p > 0$,
(2) $\rho := \frac{dp}{dh} > 0$,
(3) $\eta^2 := \rho(\frac{d^2p}{dh^2})^{-1} > 0$.

We call $\rho$ the density of the fluid and $\eta$ the speed of sound. The Euler equations are given by

$$\begin{cases}\frac{\partial}{\partial t}v_i + (v \cdot \nabla)v_i = -\frac{\partial h}{\partial x_i}\\ \frac{\partial}{\partial t}\rho + \nabla \cdot (\rho v) = 0,\end{cases} \tag{1.3}$$

for $i = 1, 2, 3$. We say that a flow is **irrotational** if $\nabla \times v = 0$. In that case, we can write $v = -\nabla\phi$, where $\phi$ is defined up to adding a function of time.

The first equation in (1.3) gives

$$\nabla\left(\frac{\partial\phi}{\partial t} - \frac{1}{2}|\nabla\phi|^2 - h\right) = 0.$$

Since we have the freedom to add a function of time to $\phi$ (which does not change $v$), we can choose

$$\frac{\partial\phi}{\partial t}-\frac{1}{2}|\nabla\phi|^2-h=0.$$

Then the second equation in (1.3) gives

$$\eta^{-2}\frac{\partial^2\phi}{\partial t^2}-2\eta^{-2}\frac{\partial\phi}{\partial x_i}\frac{\partial^2\phi}{\partial t\partial x_i}+\eta^{-2}\frac{\partial\phi}{\partial x_i}\frac{\partial\phi}{\partial x_j}\frac{\partial^2\phi}{\partial x_i\partial x_j}-\frac{\partial^2\phi}{\partial x_i\partial x_i}=0.$$

Here, we have use the convention that repeated indices are summed over.

**Example 1.4** (Einstein vacuum equations, Einstein 1915)**.** The Einstein vacuum equations describe the propagate of gravitational waves in the absence of matter and take the form

$$Ric(g)=0$$

where the Lorentzian metric $g$ is the unknown. In a coordinate system, these equations take the form (**Exercise**)

$$Ric(g)_{\mu\nu}=-\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\alpha\beta}g_{\mu\nu}-\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\mu\nu}g_{\alpha\beta}+\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\alpha\nu}g_{\beta\mu}+\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\beta\mu}g_{\alpha\nu}+F_{\mu\nu}(g,\partial g),$$

where $F_{\mu\nu}(g,\partial g)$ is a function of $g$ and its derivatives. Here, we summed over repeated indices, where the indices run through $0,1,2,3$.

This does not look like a wave equation (because of the second to fourth terms)! However, as we will see later in the course that a more careful choice of coordinates allows one to rewrite this system as a system of nonlinear wave equations.

**Example 1.5** (Wave map equations)**.** Let $\phi : I\times\mathbb{R}^n\to\mathbb{S}^m=\{x\in\mathbb{R}^{m+1}:|x|=1\}$. The wave map equation is given by the following system of $(m+1)$ equations:

$$\Box\phi=\phi(\partial_t\phi^{\mathbf{T}}\partial_t\phi-\sum_{i=1}^n\partial_i\phi^{\mathbf{T}}\partial_i\phi),$$

where $\mathbf{T}$ denotes the transpose of a vector in $\mathbb{R}^{m+1}$.

**Exercise:** Show that this is well-defined, i.e., suppose that $|\phi_0|^2=1$ and $\phi_0^t\phi_1=0$. Then, if a solution $\phi$ exists in $I\times\mathbb{R}^n$, then $|\phi|^2=1$, i.e., $\phi$ is indeed a map to the sphere.

We will begin the course by studying the linear wave equation. As we will see, the solution for the Cauchy problem for the linear wave equation can in fact be written down explicitly. Nevertheless, we will single out $3+1$ properties here:

(0) Existence and uniqueness of solutions for the Cauchy problem.

(1) (Conservation of energy)

$$E(t)=\frac{1}{2}\int_{\mathbb{R}^n}((\partial_t\phi)^2+\sum_{i=1}^n(\partial_i\phi)^2)(t,x)dx.$$

is independent of time.

(2) (Dispersion) For smooth and compactly supported initial data, the solution $\phi$ satisfies a decay estimate

$$\sup_x|\phi|(t,x)\leq C(1+t)^{-\frac{n-1}{2}}$$

for some $C>0$. (One may immediately ask: how can conservation and decay happen at the same time? The answer is that the support of the function grows and therefore even the pointwise value of $\phi$ is decaying, the $L^2$ norm is conserved. This is why the phenomenon is called *dispersion*.)

(3) (Finite speed of propagation) If $(\phi_0,\phi_1)=(0,0)$ in $\{y\in\mathbb{R}^n:|y-x|\leq t\}$, then $\phi(t,x)=0$.

On the other hand, for nonlinear equations, things are very different!

(0) One has a *general* theory for existence and uniqueness of *local* solutions.

(1) There is a wide range of large time behaviour:
   (a) Some solutions (to some equations) behave like linear equation and disperse;
   (b) Some are global-in-time but do not disperse;
   (c) Some are not global-in-time.

(2) Restricting to small and localized initial data, the solutions behave like their linear counterpart for a long time. However, even in this case, the global behaviour can be very different from that of linear solutions.

## 2. Linear wave equation via Fourier transform

We use the Fourier transform to study the wave equation on $I \times \mathbb{R}^n$:

$$\begin{cases} \Box\phi = 0 \\ (\phi, \partial_t\phi) \restriction_{\{t=0\}} = (\phi_0, \phi_1). \end{cases} \tag{2.1}$$

### 2.1. Review of Fourier theory.

Let us briefly review some elements of Fourier theory. To begin, let us define some function spaces:

**Definition 2.1.** Define the following function spaces:

(1) Let $U \subseteq \mathbb{R}^n$ be an open set. Define $C^\infty(U)$ be the space of all functions $f : U \to \mathbb{R}$ which are smooth.
(2) Let $U \subseteq \mathbb{R}^n$ be an open set. Define $C_c^\infty(U) = \{f \in C^\infty(U) : supp(f) \subseteq U \text{ is a compact set}\}$.
(3) Let the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ be

$$\mathcal{S}(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}^n) : \sup_x |x^\alpha \partial_x^\beta f| \leq C_{\alpha,\beta} \text{ for all multi-indices } \alpha, \beta\}.$$

Here, we have used the multi-index notation, i.e., for $\alpha = (\alpha_1, \alpha_2, ...\alpha_n)$, we denote by $x^\alpha$ the function $x_1^{\alpha_1} \cdot ... \cdot x_n^{\alpha_n}$ and $\partial_x^\alpha$ the differential operator $\partial_{x_1}^{\alpha_1} ... \partial_{x_n}^{\alpha_n}$.
(4) For $p \in [1, \infty)$, let $L^p(\mathbb{R}^n)$ be the completion of $\mathcal{S}(\mathbb{R}^n)$ under the norm $\|f\|_{L^p(\mathbb{R}^n)} := (\int_{\mathbb{R}^n} |f|^p(x)\, dx)^{\frac{1}{p}}$.
(5) In the $p = \infty$ case, let $L^\infty(\mathbb{R}^n)$ be the space of functions $f : \mathbb{R}^n \to \mathbb{R}$ which are bounded except on a Lebesgue measure 0 set. Define the norm to be $\|f\|_{L^\infty(\mathbb{R}^n)} := esssup_{x \in \mathbb{R}^n} |f(x)|$.

**Definition 2.2.** Given a function $f \in L^1(\mathbb{R}^n)$, define the *Fourier transform* by the formula

$$\hat{f}(\xi) := \int_{\mathbb{R}^n} f(x) e^{-2\pi i x\cdot\xi}\, dx.$$

For convenience, we will also use the notation $\mathcal{F}(f) := \hat{f}$. The following are the basic results that we will use:

**Theorem 2.3.** *(1) (Fourier inversion formula) Define the* inverse Fourier transform *by*

$$\check{g}(x) := \int_{\mathbb{R}^n} g(\xi) e^{2\pi i x\cdot\xi}\, d\xi.$$

*For $f$ such that $f, \hat{f} \in L^1(\mathbb{R}^n)$, we have*

$$(\hat{f})\check{} = f \tag{2.2}$$

*almost everywhere.*
*(2) The Fourier transform maps the Schwartz space onto itself.*
*(3) (Plancheral theorem) Let $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Then*

$$\|\hat{f}\|_{L^2(\mathbb{R}^n)} = \|f\|_{L^2(\mathbb{R}^n)}.$$

*As a consequence, the Fourier transform extends (by density) to an isometric isomorphism* $: L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$.

### 2.2. Returning to the linear wave equation.

We now apply this to the linear wave equation. For the purpose of this discussion, let us from now on only consider initial data $(\phi_0, \phi_1) \in C_c^\infty(\mathbb{R}^n) \times C_c^\infty(\mathbb{R}^n) \subseteq \mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n)$. Taking the Fourier transform of the wave equation and noting that

$$\mathcal{F}(\partial_i\phi)(\xi) = 2\pi i \xi_i \hat{\phi}(\xi),$$

we get

$$-\partial_t^2 \hat{\phi}(\xi) - 4\pi^2 |\xi|^2 \hat{\phi}(\xi) = 0.$$

The general solution to this ODE is given by

$$\hat{\phi}(t,\xi) = A(\xi)\sin(2\pi t|\xi|) + B(\xi)\cos(2\pi t|\xi|).$$

The initial conditions imply that

$$A(\xi) = \frac{\hat{\phi}_1(\xi)}{2\pi|\xi|}, \quad B(\xi) = \hat{\phi}_0(\xi),$$

i.e.

$$\hat{\phi}(t,\xi) = \frac{\hat{\phi}_1(\xi)}{2\pi|\xi|}\sin(2\pi t|\xi|) + \hat{\phi}_0(\xi)\cos(2\pi t|\xi|). \tag{2.3}$$

We have thus shown the existence of a solution! In fact, as long as we require the solution to be sufficiently regular such that all the above operations make sense, this is in fact the unique solution. At the very least, it is easy to see that if the initial data are such that

$$(\phi, \partial_t\phi)\restriction_{\{t=0\}} \in H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n),$$

then this is the unique solution such that

$$(\phi, \partial_t\phi) \in L^\infty([0,T]; H^1(\mathbb{R}^n)) \times L^\infty([0,T]; L^2(\mathbb{R}^n)).$$

We will now show the properties of the solution. First, the conservation of energy:

**Proposition 2.4.** *For $(\phi_0, \phi_1) \in C_c^\infty(\mathbb{R}^n)\times C_c^\infty(\mathbb{R}^n)$ and $\phi$ the unique solution to (2.1),*

$$E(t) = \frac{1}{2}\int_{\mathbb{R}^n} ((\partial_t\phi)^2 + \sum_{i=1}^n (\partial_i\phi)^2)(t,x)dx$$

*is independent of time.*

*Proof.* By Plancherel Theorem,

$$E(t) = \frac{1}{2}\int_{\mathbb{R}^n} ((\partial_t\hat{\phi})^2 + \sum_{i=1}^n (2\pi\xi_i\hat{\phi})^2)(t,\xi)d\xi.$$

We shall show that in fact the integrand is independent of time. First,

$$|\partial_t\hat{\phi}|^2 = \Big(\hat{\phi}_1(\xi)\cos(2\pi t|\xi|) - \hat{\phi}_0(\xi)2\pi|\xi|\sin(2\pi t|\xi|)\Big)^2$$

Also,

$$\begin{aligned}\sum_{i=1}^n (2\pi\xi_i\hat{\phi})^2 =& 4\pi^2|\xi|^2\left(\frac{\hat{\phi}_1(\xi)}{2\pi|\xi|}\sin(2\pi t|\xi|) + \hat{\phi}_0(\xi)\cos(2\pi t|\xi|)\right)^2 \\ =& |\hat{\phi}_1(\xi)|^2\sin^2(2\pi|\xi|t) + 4\pi^2|\xi|^2|\hat{\phi}_0(\xi)|^2\cos^2(2\pi|\xi|t) + 4\pi|\xi|\hat{\phi}_0(\xi)\hat{\phi}_1(\xi)\sin(2\pi t|\xi|)\cos(2\pi t|\xi|).\end{aligned}$$

Summing, we get

$$|\partial_t\hat{\phi}|^2 + \sum_{i=1}^n (2\pi\xi_i\hat{\phi})^2 = |\hat{\phi}_1(\xi)|^2 + 4\pi^2|\xi|^2|\hat{\phi}_0(\xi)|^2,$$

which is manifestly independent of $t$. □

We then look at the decay of $\phi$ in $t$. Why do we expect the solution to decay[1]? The reason is that as $t$ becomes large, there are more oscillations in $\sin(2\pi t|\xi|)$ and $\cos(2\pi t|\xi|)$. Before we proceed, we first rewrite the solution:

$$\hat{\phi}(t,\xi) = e^{2\pi it|\xi|}(\frac{\hat{\phi}_0(\xi)}{2} + \frac{\hat{\phi}_1(\xi)}{2\pi i|\xi|}) + e^{-2\pi it|\xi|}(\frac{\hat{\phi}_0(\xi)}{2} - \frac{\hat{\phi}_1(\xi)}{2\pi i|\xi|}). \tag{2.4}$$

We first prove a slightly simpler decay result which hold in the region $\{|x| \leq R\}$, but the estimate degenerates as $R$ becomes large.

[1] Of course we argued previously that this is due to dispersion in physical space, but we want to see this in Fourier variables.

**Proposition 2.5.** *Fix any $R > 0$. For $(\phi_0, \phi_1) \in C_c^\infty(\mathbb{R}^n) \times C_c^\infty(\mathbb{R}^n)$ and $\phi$ the unique solution to* (2.1)*, there exists $C = C(\phi_0, \phi_1, R) > 0$ such that*

$$|\phi(t,x)| \leq \frac{C}{(1+t)^{n-1}}$$

*whenever $|x| \leq R$ and $t \geq 0$.*

*Proof.* We will only consider the case $n = 3$ and the integral

$$I = \int_{\mathbb{R}^3} e^{2\pi i(t|\xi|+x\cdot\xi)} \frac{\hat{\phi}_1(\xi)}{2\pi i|\xi|} d\xi.$$

(Of course, according to the (2.4), there are four such terms to consider.) We make a further simplication that we only consider $t \geq 1$, since it is clear that $|\phi(t,x)| \leq C$ using the fact that $\hat{\phi}_1(\xi) \in \mathcal{S}$.

To proceed, we rewrite $\xi$ in polar coordinates $(|\xi|, \xi_\theta, \xi_\varphi)$ defined by the relations

$$|\xi|^2 = \xi_1^2 + \xi_2^2 + \xi_3^2, \quad \xi_3 = |\xi|\cos\xi_\theta, \quad \xi_1 = |\xi|\sin\xi_\theta\cos\xi_\varphi.$$

In this coordinate system, the volume form is $d\xi = |\xi|^2 \sin\xi_\theta d|\xi| d\xi_\theta d\xi_\varphi$.

The key point is to notice that

$$(\frac{1}{2\pi it}\frac{\partial}{\partial|\xi|})^N e^{2\pi it|\xi|} = e^{2\pi it|\xi|}$$

for every non-negative integer $N$. Using this, we can integrate by parts in $|\xi|$ to get

$$\begin{aligned}
I &= -\int_{\mathbb{R}^3} e^{2\pi it|\xi|}\frac{1}{2\pi it}\frac{\partial}{\partial|\xi|}\left(e^{2\pi ix\cdot\xi}\frac{\hat{\phi}_1(\xi)}{2\pi i}|\xi|\right)\sin\xi_\theta d|\xi|d\xi_\theta d\xi_\varphi \\
&= \int_{\mathbb{R}^3} e^{2\pi it|\xi|}\frac{1}{4\pi^2 t}e^{2\pi ix\cdot\xi}\left((2\pi ix\cdot\xi)\hat{\phi}_1(\xi) + \hat{\phi}_1(\xi) + |\xi|\frac{\partial}{\partial|\xi|}\hat{\phi}_1(\xi)\right)\sin\xi_\theta d|\xi|d\xi_\theta d\xi_\varphi
\end{aligned}$$

Notice that there are no boundary terms arising from the integration by parts since $\hat{\phi}_1(\xi)$ is in Schwartz class and $|\xi|\hat{\phi}_1(\xi)$ vanishes at $|\xi| = 0$. Note that we have thus gained a power of $t$! We integrate by parts one more time to gain another power of $t$. Notice that this time we have a boundary term at $|\xi| = 0$:

$$\begin{aligned}
I = &-\int_{\mathbb{R}^3} e^{2\pi it|\xi|}\frac{1}{8\pi^3 it^2}\frac{\partial}{\partial|\xi|}\left(e^{2\pi ix\cdot\xi}\left((2\pi ix\cdot\xi)\hat{\phi}_1(\xi) + \hat{\phi}_1(\xi) + |\xi|\frac{\partial}{\partial|\xi|}\hat{\phi}_1(\xi)\right)\right)\sin\xi_\theta d|\xi|d\xi_\theta d\xi_\varphi \\
&+ \frac{1}{2\pi^2 it^2}\hat{\phi}_1(\xi = 0) \\
= &-\int_{\mathbb{R}^3} e^{2\pi it|\xi|}\frac{1}{8\pi^3 it^2}e^{2\pi ix\cdot\xi}\left(\frac{(2\pi ix\cdot\xi)^2}{|\xi|}\hat{\phi}_1(\xi) + \frac{4\pi ix\cdot\xi}{|\xi|}\hat{\phi}_1(\xi) + (4\pi ix\cdot\xi)\frac{\partial}{\partial|\xi|}\hat{\phi}_1(\xi)\right. \\
&\left. + \frac{\partial}{\partial|\xi|}\hat{\phi}_1(\xi) + \frac{\partial}{\partial|\xi|}(|\xi|\frac{\partial}{\partial|\xi|}\hat{\phi}_1(\xi))\right)\sin\xi_\theta d|\xi|d\xi_\theta d\xi_\varphi + \frac{1}{2\pi^2 it^2}\hat{\phi}_1(\xi = 0).
\end{aligned}$$

The boundary term in particular does not allow us to integrate by parts again. We thus simply bound each of these terms. Using the fact the $\hat{\phi}_1 \in \mathcal{S}$, it is easy to see that

$$|I| \leq \frac{C(1+R^2)}{t^2},$$

as desired. □

Of course, the above decay bound is not uniform for all $x \in \mathbb{R}^n$. The following theorem, on the other hand, gives a uniform decay rate. Notice that the decay rate obtained is weaker[2]:

**Theorem 2.6.** *For $(\phi_0, \phi_1) \in C_c^\infty(\mathbb{R}^n) \times C_c^\infty(\mathbb{R}^n)$ and $\phi$ the unique solution to* (2.1)*, there exists $C = C(\phi_0, \phi_1) > 0$ such that*

$$\sup_{x\in\mathbb{R}^n}|\phi(t,x)| \leq \frac{C}{(1+t)^{\frac{n-1}{2}}}$$

*for $t \geq 0$.*

[2]Nevertheless, as we will see in the Example Sheet, it is sharp!

*Proof.* We give a proof that works when $n$ is odd. As before, we will again only consider $n = 3$ and the term

$$I = \int_{\mathbb{R}^3} e^{2\pi i(t|\xi|+x\cdot\xi)} \frac{\hat{\phi}_1(\xi)}{2\pi i|\xi|} d\xi.$$

Without loss of generality, we assume that $x = (0, 0, \alpha)$ for some $\alpha \in \mathbb{R}$. Let $\rho = \sqrt{\xi_1^2 + \xi_2^2}$. We use the coordinate system $(\xi_3, \rho, \xi_\theta)$, where $\xi_\theta$ is given by

$$\rho \cos \xi_\theta = \xi_1.$$

We now try to repeat the argument in the proof of Proposition 2.5 with $\frac{\partial}{\partial|\xi|}$ replaced by $\frac{\partial}{\partial\rho}$. The key point is of course that $\frac{\partial}{\partial\rho} e^{2\pi i x\cdot\xi} = 0$ so we do not get a dependence on $x$. We first note that

$$\frac{\partial|\xi|}{\partial\rho} = \frac{\rho}{|\xi|}$$

and thus for every non-negative integer $N$, we have

$$(\frac{|\xi|}{2\pi i\rho t}\frac{\partial}{\partial\rho})^N e^{2\pi i t|\xi|} = e^{2\pi i t|\xi|}.$$

Using this, we can integrate by parts in $\rho$. Notice that the volume form is $d\xi = \rho d\rho d\xi_\theta d\xi_3$. We thus have

$$I = -\int_{\mathbb{R}^3} e^{2\pi i t|\xi|+2\pi i x\cdot\xi} \frac{1}{2\pi i t}\frac{\partial}{\partial\rho}\left(\frac{\hat{\phi}_1(\xi)}{2\pi i}\right) d\rho d\xi_\theta d\xi_3 + \int e^{2\pi i t|\xi_3|+2\pi i\alpha\xi_3} \frac{1}{it}\frac{\hat{\phi}_1(\rho = 0, \xi_3)}{2\pi i} d\xi_3,$$

where the last term is a boundary term at $\rho = 0$. It is easy to see that both terms are bounded in magnitude by $\frac{C}{t}$, uniformly in $x$. □

*Remark* 2.7. The proof we gave above would not give the desired result when $n$ is even. In fact, in that case the same proof only gives the weaker decay (**Exercise**)

$$\sup_{x\in\mathbb{R}^n} |\phi(t, x)| \leq \frac{C}{(1 + t)^{\frac{n-2}{2}}}.$$

Nevertheless, the decay rate as stated in Theorem 2.6 still holds with a more refined analysis.

## 3. Linear wave equation via fundamental solution and representation formula

[The treatment of this section is largely inspired by the very nice lecture notes of Oh, which are available at https://math.berkeley.edu/∼sjoh/pdfs/linearWave.pdf]

In this section, we present a different point of view in solving the linear wave equation (2.1), namely we consider the fundamental solution, i.e., we look for a distribution $E \in \mathcal{D}'(\mathbb{R}^n)$ such that

$$\Box E = \delta_0. \tag{3.1}$$

For reasons that will become clear later, we will look for the *forward* fundamental solution $E_+$, which in addition to satisfying (3.1), also verifies

$$supp(E_+) \subseteq \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : 0 \leq |x| \leq t\}.$$

The goal will be to write down a representation formula in physical space[3] for the solution to the Cauchy problem for the linear wave equation (2.1). Before that, let us briefly review some elements of distribution theory.

[3]In view of (2.3), we can of course write

$$\phi(t, x) = \phi_1 * (\frac{\sin(2\pi t|\xi|)}{2\pi|\xi|})\check{} + \phi_0 * (\cos(2\pi t|\xi|))\check{}.$$

The point, however, is to derive this in physical space.

3.1. **Review of distribution theory.** We begin with some definitions:

**Definition 3.1.** (1) A *distribution* in an open set $U \subseteq \mathbb{R}^n$ is a linear map $u : C_c^\infty(U) \to \mathbb{R}$ such that for every compact set $K \subseteq U$, there exists $C > 0$ and $k \in \mathbb{N} \cup \{0\}$ such that

$$|u(\varphi)| \leq C \sum_{|\alpha| \leq k} \sup_{x \in K} |\partial^\alpha \varphi(x)|$$

for all $\varphi \in C_c^\infty(K)$. We will also write the pairing as $< u, \varphi >$. Denote the set of all distributions on $U$ by $\mathcal{D}'(U)$.

(2) Let $V \subseteq U \subseteq \mathbb{R}^n$ and $u \in \mathcal{D}'(U)$. Define *the restriction of* $u$ *to* $V$, denoted $u_V$, by $u_V(\varphi) := u(\varphi)$ for every $\varphi \in C_c^\infty(V)$.

(3) Let $U \subseteq \mathbb{R}^n$ and $u \in \mathcal{D}'(U)$. Define *the support of* $u$, denoted *supp*$(u)$, as the set of all points in $U$ having no open neighbourhood to which the restriction of $u$ is 0.

(4) Let $U \subseteq \mathbb{R}^n$ and $u \in \mathcal{D}'(U)$. Define the distributional derivative $\partial^\alpha u$ as the distribution such that $< \partial^\alpha u, \varphi > := (-1)^{-|\alpha|} < u, \partial^\alpha \varphi >$.

We will need some basic facts about distributions, which we will not prove.

**Proposition 3.2.** *(1) (Approximation by $C_c^\infty$ functions) If $U \subseteq \mathbb{R}^n$ is open and $u \in \mathcal{D}'(U)$, then there exists a sequence $u_j \in C_c^\infty(U)$ such that $u_j \to u$ in $\mathcal{D}'(U)$, i.e., for every $\varphi \in C_c^\infty(U)$, it holds that $\int_U u_j \varphi \to < u, \varphi >$.*

*(2) (Composition with smooth maps) Let $U \in \mathbb{R}^n$ and $f : U \to \mathbb{R}$ be a smooth map such that $df \neq 0$. There exists a unique continuous map $f^* : \mathcal{D}'(\mathbb{R}) \to \mathcal{D}'(U)$ such that $f^* u = u \circ f$ whenever $u$ is a continuous function. Given a distribution $u \in \mathcal{D}'(\mathbb{R})$, this can be defined as the limit of $u_j \circ f$ where $u_j \to u$ and $u_j \in C_c^\infty(\mathbb{R})$ (whose existence is guaranteed by the previous part). We will abuse notation to write $f^* u$ as $u \circ f$ even for general $u \in \mathcal{D}'(\mathbb{R})$.*

*(3) (Chain rule) Let $u \in \mathcal{D}'(\mathbb{R})$, $U \in \mathbb{R}^n$ and $f : U \to \mathbb{R}$ be a smooth map such that $df \neq 0$. Then the distributional derivative satisfies the chain rule $\partial(u \circ f) = (\partial f)(u' \circ f)$.*

*Proof.* For proofs of the first two statements, we refer the readers to [L. Hörmander, *The Analysis of Linear Partial Differential Operators I*, Theorem 4.1.5, 6.1.2]. The last statement is easily justifiable using the second statement. □

Our next goal is to review the convolutions of distributions. First, define

**Definition 3.3.** Let $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\varphi \in C_c^\infty(\mathbb{R}^n)$. Define $(u * \varphi)(x) := < u, \varphi(x - \cdot) >$.

Extrapolating from the definitions for functions, the idea is to define the convolution of distributions $u_1 * u_2$ as the (unique) distribution such that $(u_1 * u_2) * \varphi = u_1 * (u_2 * \varphi)$ for every $\varphi \in C_c^\infty(\mathbb{R}^n)$. This may not always be well-defined, but it is well-defined if appropriate support properties are satisfied:

**Lemma 3.4.** *(1) Let $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\varphi \in C_c^\infty(\mathbb{R}^n)$. Then $u * \varphi \in C^\infty(\mathbb{R}^n)$ with $\partial(u * \varphi) = (\partial u) * \varphi = u * (\partial \varphi)$ and supp$(u * \varphi) \subseteq$ supp$(u) +$ supp$(\varphi)$.*

*(2) Let $U : C_c^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ be a linear map such that for every $\varphi_j \to 0$, we have $U(\varphi_j) \to 0$. Assume moreover that $U$ commutes with all translations, i.e., for every $h \in \mathbb{R}^n$ and every $\varphi \in C_c^\infty(\mathbb{R}^n)$, we have $U(\tau_h(\varphi)) = \tau_h(U(\varphi))$, where $\tau_h : C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ is defined by $\tau_h(\psi)(x) := \psi(x - h)$. Then there exists a unique distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ such that $u * \varphi = U(\varphi)$ for all $\varphi \in C_c^\infty(\mathbb{R}^n)$.*

*(3) Let $u_1, u_2 \in \mathcal{D}'(\mathbb{R}^n)$. If $(-$supp$(u_1)) \cap ($supp$(u_2) + K)$ is compact for any compact set $K$, then there exists a unique distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ such that $u * \varphi = u_1 * (u_2 * \varphi)$ for every $\varphi \in C_c^\infty(\mathbb{R}^n)$.*

*Proof.* Part (1) is straightforward and will be omitted. For part (2), define $u$ by $u(\varphi) := (U(\tilde{\varphi}))(0)$ for $\varphi \in C_c^\infty(\mathbb{R}^n)$, where $\tilde{\varphi} \in C_c^\infty(\mathbb{R}^n)$ is defined by $\tilde{\varphi}(x) := \varphi(-x)$. It is now straightforward to check that for every $h \in \mathbb{R}^n$,

$$(U(\varphi))(-h) = (\tau_h(U(\varphi)))(0) = (U(\tau_h(\varphi)))(0) = (u * (\tau_h(\varphi)))(0) = (u * \varphi)(-h).$$

For part (3), the support properties are used to justify that $u_1 * (u_2 * \varphi)$ is well-defined. More precisely, in order to justify that, we need $(-supp(u_1)) \cap (supp(u_2 * \varphi)) = (-supp(u_1)) \cap (supp(u_2) + supp(\varphi))$ to be compact, which is guaranteed by the assumption. The existence of $u$ follows then by part (2). □

Next, we briefly discuss the theory of homogeneous distribution, which turns out to play an important role for the fundamental solution to the linear wave equation. First, a definition:

**Definition 3.5.** We say that a distribution $h \in \mathcal{D}'(\mathbb{R} \setminus \{0\})$ is *homogeneous of degree* $a$ if for every $\varphi \in C_c^\infty(\mathbb{R} \setminus \{0\})$,

$$< h, \varphi >= \lambda^a < h, \varphi_\lambda >, \quad \text{where } \varphi_\lambda(x) := \lambda\varphi(\lambda x).$$

For $a \in \mathbb{C}$ such that $Re(a) > -1$,

$$x_+^a = \mathbb{1}_{\{x\geq 0\}} x^a$$

is a homogeneous distribution of degree $a$. To analytically extend this, notice that

$$\frac{d}{dx} x_+^a = a x_+^{a-1}$$

for $Re(a) > -1$. However, there are poles at the negative integers. In order to deal with this, define

$$\chi_+^a(x) := \frac{x_+^a}{\Gamma(a+1)} \tag{3.2}$$

such that (**Exercise**)

$$\frac{d}{dx}\chi_+^a(x) = \chi_+^{a-1}(x). \tag{3.3}$$

(Let us recall here that $\Gamma(x) := \int_0^\infty t^{x-1} e^{-t}\, dt$ and that $\Gamma(a+1) = a\Gamma(a)$ for all $a \in \mathbb{C}$.) We will be most concerned with the cases where $a$ is either an integer or a half integer. Let us note that following properties:

**Lemma 3.6.** *(1) For any $k \in \mathbb{N}$,*

$$\chi_+^{-k}(x) = \delta_0^{(k-1)}(x),$$

*where $\delta_0(x)$ is the distribution $< \delta_0, \varphi >= \varphi(0)$.*

*(2) For any $k \in \mathbb{N}$,*

$$\chi_+^{-\frac{1}{2}-k}(x) = \frac{1}{\sqrt{\pi}}(\frac{d}{dx})^k(\frac{1}{x_+^{\frac{1}{2}}}),$$

*where derivatives on the right hand side is understood as distributional derivatives.*

*Proof.* For (1), it suffices to note that $\chi_+^0(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$ is the Heaviside step function and using (3.3). For (2), recall that $\Gamma(a)\Gamma(1-a) = \frac{\pi}{\sin(\pi a)}$, which implies $(\Gamma(\frac{1}{2}))^2 = \pi$, i.e., $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. This implies $\chi_+^{-\frac{1}{2}}(x) = \frac{1}{\sqrt{\pi}x_+^{\frac{1}{2}}}$. The rest follows from (3.3). ☐

### 3.2. Returning to fundamental solutions to the linear wave equation.

**Definition 3.7.** We say the $E_+$ is a *forward fundamental solution* to the linear wave equation on $\mathbb{R}^{n+1}$ if

(1) $\Box E_+ = \delta_0$ as distributions (where $\delta_0$ is defined as $\delta_0(\varphi) = \varphi(t = 0, x = 0)$);
(2) $supp(E_+) \subseteq \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : 0 \leq |x| \leq t\}$.

**Proposition 3.8.** *A forward fundamental solution on $\mathbb{R}^{n+1}$, if it exists, is unique.*

*Proof.* Let $E_+$ and $E$ both be forward fundamental solutions. The idea is to justify the following chain of formal manipulations:

$$E = E * \delta_0 = E * (\Box E_+) = (\Box E) * E_+ = \delta_0 * E_+ = E_+.$$

Notice that we have not used the support properties of $E$ and $E_+$ yet! Of course, the key point is that thanks to Lemma 3.4 and the support properties of $E$ and $E_+$, $E * (\Box E_+)$ and $(\Box E) * E_+$ are well-defined and that the formal manipulations can indeed be justified. ☐

Once a forward fundamental solution exists, it can be used to derive a representation formula for the solution to the Cauchy problem for the linear wave equation:

**Proposition 3.9.** *Let $E_+$ be a forward fundamental solution to the linear wave equation on $\mathbb{R}^{n+1}$. Then if $(\phi_0, \phi_1) \in C^\infty(\mathbb{R}^n) \times C^\infty(\mathbb{R}^n)$, the unique solution $\phi$ to (2.1) is given by*

$$\phi(t,x) = -E_+ * (\phi_1 \delta_{\{t=0\}}) - (\partial_t E_+) * (\phi_0 \delta_{\{t=0\}}).$$

*More generally, for $\Box\phi = F$, where $F \in C^\infty(\mathbb{R}^n)$, $\phi$ is given for $t \geq 0$ by the following formula*

$$\phi(t,x) = -E_+ * (\phi_1 \delta_{\{t=0\}}) - (\partial_t E_+) * (\phi_0 \delta_{\{t=0\}}) + (F \mathbb{1}_{\{t\geq 0\}}) * E_+.$$

*Proof.* Without loss of generality assume $t > 0$. Let us suppose that we have a solution $\phi$ and derive a representation formula. Of course, once we have derived this formula, it also proves the existence of solutions.

$$\begin{aligned}\phi \mathbb{1}_{\{t\geq 0\}} =&(\phi \mathbb{1}_{\{t\geq 0\}}) * \delta_0 = (\phi \mathbb{1}_{\{t\geq 0\}}) * (\Box E_+) = (\phi \mathbb{1}_{\{t\geq 0\}}) * (-\partial_t^2 E_+) + (\Delta\phi \mathbb{1}_{\{t\geq 0\}}) * E_+ \\ =&(\phi \mathbb{1}_{\{t\geq 0\}}) * (-\partial_t^2 E_+) + (\partial_t^2 \phi \mathbb{1}_{\{t\geq 0\}}) * E_+ \\ =&-(\partial_t \phi \mathbb{1}_{\{t\geq 0\}}) * (\partial_t E_+) - (\phi \delta_{\{t=0\}}) * (\partial_t E_+) + (\partial_t^2 \phi \mathbb{1}_{\{t\geq 0\}}) * E_+ \\ =&-(\partial_t \phi \delta_{\{t=0\}}) * E_+ - (\phi \delta_{\{t=0\}}) * (\partial_t E_+).\end{aligned}$$

The representation formula in the case $\Box\phi = F$ can be derived in an identical manner. □

The forward fundamental solution $E_+$ can be given in terms of $\chi_+^a$ (recall (3.2)). More precisely, we define $\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2)$ first as a distribution on $\mathbb{R}^{n+1} \setminus \{0\}$ using[4] Proposition 3.2.2. We then note that by Lemma 3.6, $\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2)$ can be written as $\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2) = \frac{u(\omega)}{(t^2+|x|^2)^{\frac{n-1}{2}}}$ for some distribution[5] $u \in \mathcal{D}'(\mathbb{S}^{n-1,1})$, where $\mathbb{S}^{n-1,1}$ is the topological $n$-sphere given by $\mathbb{S}^{n-1,1} := \{(t,x) \in \mathbb{R}\times\mathbb{R}^n : t^2 + |x|^2 = 1\}$. Since $\frac{1}{(t^2+|x|^2)^{\frac{n-1}{2}}}$ is integrable with respect to $(|x|^2 + t^2)^{\frac{n}{2}}\, d\sqrt{|x|^2 + t^2}$ in $\mathbb{R}^{n+1}$, we can therefore extend $\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2)$ to a distribution on the whole of $\mathbb{R}^{n+1}$ by $< \chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2), \varphi >:= \int_0^\infty \frac{<u, \varphi_{\sqrt{|x|^2+t^2}}>}{(t^2+|x|^2)^{\frac{n-1}{2}}} (|x|^2 + t^2)^{\frac{n}{2}}\, d\sqrt{|x|^2 + t^2}$, where $u$ is as above and $\varphi_{\sqrt{|x|^2+t^2}}$ is the restriction of $\varphi$ to the constant $\sqrt{|x|^2 + t^2}$-hypersurface.

**Proposition 3.10.** *The unique forward fundamental solution is given by*

$$E_+(t,x) = -\frac{\pi^{\frac{1-n}{2}}}{2} \mathbb{1}_{\{t\geq 0\}} \chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2).$$

*Partial proof.* The support properties are obviously satisfied. It therefore suffices to show that $\Box E_+ = \delta_0$. We will actually not prove this here, but will simply prove $\Box E_+ = c_n \delta_0$ for some constant $c_n \neq 0$. Notice that even if we do not get the precise numerical value of the constant, this would already be enough for us to understand the behaviour of the solutions!

We now compute that away from zero. As a distribution in $\mathcal{D}'(\mathbb{R}^n \setminus \{0\})$, we have

$$\Box\left(\mathbb{1}_{\{t\geq 0\}} \chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2)\right) = \mathbb{1}_{\{t\geq 0\}} \Box\left(\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2)\right)$$

due to the support properties. We then have, using the chain rule,

$$\begin{aligned}&\mathbb{1}_{\{t\geq 0\}} \Box\left(\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2)\right) \\ =&\mathbb{1}_{\{t\geq 0\}}\left(-\partial_t(2t\chi_+^{-\frac{n+1}{2}}(t^2 - |x|^2)) - 2\sum_{i=1}^n \partial_{x_i}(x_i \chi_+^{-\frac{n+1}{2}}(t^2 - |x|^2))\right) \\ =&\mathbb{1}_{\{t\geq 0\}}\left(-2(n+1)\chi_+^{-\frac{n+1}{2}}(t^2 - |x|^2)) - 4(t^2 - |x|^2)\chi_+^{-\frac{n+3}{2}}(t^2 - |x|^2)\right) = 0,\end{aligned}$$

where the last line is due to the identity (which originally holds for $Re(a) > -1$ but then holds for all $a \in \mathbb{C}$ by analyticity)

$$x\chi_+^a = \frac{x x_+^a}{\Gamma(a+1)} = \frac{(a+1)x_+^{a+1}}{\Gamma(a+2)} = (a+1)\chi_+^{a+1}.$$

[4]Notice that we have to remove $\{0\}$ since $df = 0$ there for $f(t,x) = t^2 - |x|^2$.

[5]Strictly speaking, we have not defined distributions on compact manifolds, but it is not difficult to extend the definition of distributions in s straightforward manner.

Therefore, $\Box E_+$ is a distribution supported at $\{0\}$. In particular, it is a linear combination of the delta measure and its finitely many derivatives (**Exercise**). By a scaling argument, we claim that in fact it is a multiple of $\delta_0$. To see this, for any $\lambda \in \mathbb{R}$ and $\varphi \in C_c^\infty(\mathbb{R}^{n+1})$, define $\varphi_\lambda(t,x) := \varphi(\lambda t, \lambda x)$. We then have

$$< \Box E_+, \varphi_\lambda > = < \Box E_+, \varphi > .$$

On the other hand,

$$< \partial^\alpha \delta_0, \varphi_\lambda > = \lambda^{|\alpha|} < \partial^\alpha \delta_0, \varphi_\lambda > .$$

Therefore $\Box E_+ = c_n \delta_0$ for some $c_n \in \mathbb{R}$. To see that $c_n \neq 0$, we claim that if $E$ is a tempered distribution (i.e., it is in the dual space of $\mathcal{S}$) for all $t \in \mathbb{R}$ such that $E = 0$ on $\{t = -1\}$ and $\Box E = 0$, then $E = 0$. This can be justified by the Fourier transform approach in the previous section (**Exercise:** do this).

□

We can now apply the formula for $E_+$ to obtain the representation formulas in dimensions $n = 1$ and $n = 3$. The same method of course applies for all dimensions, but we highlight the $n = 1$ and $n = 3$ cases for their simplicity and physical relevance respectively.

**Proposition 3.11** (D'Alembert's formula). *Let $(\phi_0, \phi_1) \in C_c^\infty(\mathbb{R}) \times C_c^\infty(\mathbb{R})$ be initial data to (2.1) in $\mathbb{R}^{1+1}$. Then the unique solution is given by*

$$\phi(t,x) = \frac{1}{2}(\phi_0(x+t) - \phi_0(x-t)) + \frac{1}{2}\int_{x-t}^{x+t} \phi_1(y)\, dy.$$

*Proof.* We compute using $\chi_+^0(t^2 - |x|^2) = \mathbb{1}_{\{t^2 - |x|^2 \geq 0\}}$ that

$$(\mathbb{1}_{\{t\geq 0\}}\chi_+^0(t^2 - |x|^2)) * (\phi_1 \delta_{\{t=0\}}) = \int_{-\infty}^{\infty} \mathbb{1}_{\{t\geq 0\}}\mathbb{1}_{\{t^2 - |x-y|^2 \geq 0\}}\phi_1(y)\, dy = \int_{x-t}^{x+t} \phi_1(y)\, dy.$$

Using this, it is then easy to see that

$$(\partial_t(\mathbb{1}_{\{t\geq 0\}}\chi_+^0(t^2 - |x|^2)) * (\phi_1 \delta_{\{t=0\}})) * (\phi_0 \delta_{\{t=0\}}) = \partial_t(\int_{x-t}^{x+t} \phi_0(y)\, dy) = \phi_0(x+t) - \phi_0(x-t).$$

The conclusion follows from Propositions 3.9 and 3.10. □

**Proposition 3.12** (Kirchhoff's formula). *Let $(\phi_0, \phi_1) \in C_c^\infty(\mathbb{R}^3) \times C_c^\infty(\mathbb{R}^3)$ be initial data to (2.1) in $\mathbb{R}^{3+1}$. Then the unique solution is given by*

$$\phi(t,x) = \frac{1}{4\pi t}\int_{\{|y-x|=t\}} \phi_1(y) d\sigma_{\{|y-x|=t\}}(y) + \frac{d}{dt}(\frac{1}{4\pi t}\int_{\{|y-x|=t\}} \phi_0(y) d\sigma_{\{|y-x|=t\}}(y)).$$

*Proof.* Recalling that $\chi_+^{-1}(t^2 - |x|^2) = \delta_0(t^2 - |x|^2)$, the key point is therefore to express $\mathbb{1}_{\{t\geq 0\}}\delta_0(t^2 - |x|^2)$ in a more familiar form. For this computation, it is convenient to use polar coordinates $(r, \theta, \phi)$ with $r = |x|$, $r\cos\theta = x_3$, $r \sin\theta\cos\phi = x_1$. Define $d\sigma_{\mathbb{S}^2} := \sin\theta\, d\theta\, d\phi$.

Notice that the support of $\mathbb{1}_{\{t\geq 0\}}\delta_0(t^2 - |x|^2)$ is contained in $\{t = |x|\}$. It is therefore convenient to introduce the null variables $(u, v) = (t - r, t + r)$ so that the support of $\mathbb{1}_{\{t\geq 0\}}\delta_0(t^2 - |x|^2)$ lies in $\{u = 0\}$. Note also that $dx\, dt = \frac{1}{2} du\, dv$ and $r^2 = \frac{(v-u)^2}{4}$.

Consider a sequence $h_j \in C_c^\infty(\mathbb{R})$ such that $h_j \to \delta_0$ as distributions, we have

$$< \mathbb{1}_{\{t\geq 0\}} h_j(t^2 - |x|^2), \varphi(t,x) > = \int_0^\infty \int_0^\infty \int_{\mathbb{S}^2} h_j(t^2 - |x|^2)\varphi(t, r, \omega) r^2 d\sigma_{\mathbb{S}^2}(\omega) dr dt$$

$$= \int_{-\infty}^{\infty}\int_{|u|}^{\infty}\int_{\mathbb{S}^2} h_j(uv)\varphi(u,v,\omega)\frac{(v-u)^2}{8} d\sigma_{\mathbb{S}^2}(\omega) dv du = \int_{-\infty}^{\infty}\int_{|u|}^{\infty}\int_{\mathbb{S}^2} h_j(\bar{u})\varphi(\frac{\bar{u}}{v}, v, \omega)\frac{(v - \frac{\bar{u}}{v})^2}{8} d\sigma_{\mathbb{S}^2}(\omega) dv \frac{d\bar{u}}{v}$$

$$\to \int_0^\infty \int_{\mathbb{S}^2} \varphi(0, v, \omega)\frac{v}{8} d\sigma_{\mathbb{S}^2}(\omega) dv = \frac{1}{2}\int_0^\infty \int_{\mathbb{S}^2} \varphi(t, r = t, \omega) d\sigma_{\mathbb{S}^2}(\omega) r dr.$$

Therefore,

$$\begin{aligned}(\mathbb{1}_{\{t\geq 0\}}\chi_+^{-1}(t^2 - |x|^2)) * (\phi_1 \delta_{\{t=0\}}) =& \frac{t}{2}\int_{\mathbb{S}^2} \phi_1(x + t\omega) d\sigma_{\mathbb{S}^2}(\omega) \\ =& \frac{1}{2t}\int_{\{|y-x|=t\}} \phi_1(y) d\sigma_{\{|y-x|=t\}}(y).\end{aligned}$$

The conclusion follows from Propositions 3.9 and 3.10. ☐

3.3. **Properties of the solution.** Using the representation formula that we have derived above, we now deduce the properties of the solution to the linear wave equation. Again, as mentioned near the end of Section 1, we are most interested in existence and uniqueness, conservation of energy, decay and finite speed of propagation.

Clearly, since we have a representation formula, the solution exists and is unique for sufficiently regular initial data. (**Exercise:** How regular do the initial data need to be? In what class is the solution unique?)

Using the representation formula, we can now prove the finite speed of propagation, which we did not show using Fourier methods:

**Proposition 3.13** (Finite speed of propagation)**.** *Suppose $\phi \in C^\infty(\mathbb{R}^{n+1})$ is a solution to the linear wave equation with data $(\phi_0, \phi_1)$. If $(\phi_0, \phi_1) = (0, 0)$ in $\{y \in \mathbb{R}^n : |x - y| \leq t\}$, then $\phi(t, x) = 0$.*

*Proof.* This is an immediate consequence of the fact that $supp(E_+) \subseteq \{(t, x) : 0 \leq |x| \leq t\}$. ☐

In fact, there is a much stronger statement, which holds in odd dimensions higher than 3:

**Proposition 3.14** (Strong Huygens principle)**.** *Let $n \geq 3$ be an odd integer. Suppose $\phi \in C^\infty(\mathbb{R}^{n+1})$ is a solution to the linear wave equation with data $(\phi_0, \phi_1)$. If $(\phi_0, \phi_1) = (0, 0)$ in $\{y \in \mathbb{R}^n : |x - y| = t\}$, then $\phi(t, x) = 0$.*

*Proof.* For $n \geq 3$ an odd integer, $supp(E_+) \subseteq \{(t, x) : |x| = t\}$. ☐

The decay properties of the solution can also be read off from the representation formula. As an example, we consider the $n = 3$ case, which follows from Kirchhoff's formula (Proposition 3.12).

**Proposition 3.15.** *For $(\phi_0, \phi_1) \in C_c^\infty(\mathbb{R}^3) \times C_c^\infty(\mathbb{R}^3)$ and $\phi$ the unique solution to (2.1) in $\mathbb{R}^{3+1}$, there exists $C = C(\phi_0, \phi_1) > 0$ such that*

$$\sup_{x \in \mathbb{R}^3} |\phi(t, x)| \leq \frac{C}{(1 + t)}$$

*for $t \geq 0$.*

*Proof.* We will prove the proposition in the case $\phi_0 = 0$. The general case can be treated similarly. By Kirchhoff's formula,

$$\phi(t, x) = \frac{1}{4\pi t} \int_{\{|y-x|=t\}} \phi_1(y) \, d\sigma_{\{|y-x|=t\}}(y).$$

For $t \in [0, 1]$, since $Area(\{|y - x| = t\}) = 4\pi t^2$, we have $\phi(t, x) \leq C \sup_y |\phi_1(y)| \leq C$ for some constant $C > 0$. For $t > 1$, notice that since $\phi_1$ has compact support, $Area(\{|y-x| = t\} \cap (supp(\phi_1))) \leq C$. Therefore, $\sup_x |\phi|(t, x) \leq \frac{C}{t}$. ☐

*Remark* 3.16. In general, one can also use the representation formula to show the following decay rate for $t \geq 0$ for the solution arising from smooth and compactly supported initial data in $\mathbb{R}^{n+1}$:

$$\sup_x |\phi|(t, x) \leq \frac{C}{(1 + t)^{\frac{n-1}{2}}},$$

where $C$ depends on the initial data. We leave this to the reader.

*Remark* 3.17. The compact support condition can be relaxed to some sufficiently strong decay as $|x| \to \infty$. However, it cannot be dropped completely. For instance, in $\mathbb{R}^{3+1}$, there exist initial conditions with $(\phi_0, \phi_1) \in C^\infty(\mathbb{R}^3) \times C^\infty(\mathbb{R}^3)$ such that $|\phi_0|(x) + |\phi_1|(x) \to 0$ as $|x| \to \infty$, but the solution $\sup_x |\phi|(t, x) \to \infty$ as $t \to \infty$. (**Exercise**)

Recall that when we study the linear wave equation using Fourier transform, the difficult property to prove was the finite speed of propagation. Now, the difficulty property to prove is the conservation of energy (which as we recall is an easy computation using the Fourier transform). This shows that each of the different methods has its strength.

## 4. Energy estimates

In the previous lectures, we studied the constant coefficient linear wave equation

$$\Box\phi = 0 \tag{4.1}$$

in $\mathbb{R}^{n+1}$ by solving the equation explicitly either via the Fourier transform or via the fundamental solution. In particular, we have shown the existence and uniqueness of solutions, conservation of energy, dispersion and the finite speed of propagation. We also saw that depending on the property that we wanted to prove, one representation of the solution may be preferable to the other.

In this section, we will introduce yet another method to study the equation (4.1), namely the energy method. In fact, we will show that we can prove (almost) all the properties of the solutions we previously showed using the energy methods. More importantly, as we will see, these methods are more robust and can deal with situations when the solutions to the equation may not be explicitly available. In particular, they will make a crucial appearance in the study of nonlinear equations!

We will use the term "energy methods" to refer loosely to methods which are based on $L^2$ type estimates. The most basic example is the conservation of energy:

**Proposition 4.1.** *Let $\phi$ be a (sufficiently regular) solution to* (4.1). *Define $E(t)$ as before as*

$$E(t) = \frac{1}{2}\int_{\mathbb{R}^n} ((\partial_t\phi)^2 + \sum_{i=1}^n (\partial_i\phi)^2)(t,x)dx.$$

*Then $E$ is constant in time.*

*Proof.* Consider

$$\begin{aligned}\frac{d}{dt}E(t) &= \int_{\mathbb{R}^n} \partial_t\phi\partial_t^2\phi + \sum_{i=1}^n \partial_t\partial_i\phi\partial_i\phi \\ &= \int_{\mathbb{R}^n} \partial_t\phi\sum_{i=1}^n \partial_i^2\phi + \sum_{i=1}^n \partial_t\partial_i\phi\partial_i\phi \\ &= \int_{\mathbb{R}^n} \partial_t\phi\sum_{i=1}^n \partial_i^2\phi - \sum_{i=1}^n \partial_t\phi\partial_i^2\phi = 0.\end{aligned}$$

Here, we have used the equation in the second line and have integrated by parts in the third line. □

*Remark* 4.2. Another (completely equivalent) way to look at the proof is to use the identity

$$0 = \partial_t\phi\Box\phi$$

and integrate it in the region $[0,t]\times\mathbb{R}^n$ with respect to the volume form $dxdt$ and integrate by parts. We will take this point of view in some of the results below.

One immediate consequence of the energy identity Proposition 4.1 is the uniqueness of solutions:

**Corollary 4.3.** *Given initial data*

$$(\phi,\partial_t\phi)\restriction_{\{t=0\}} = (\phi_0,\phi_1),$$

*if we have two solutions $\phi^{(1)}$ and $\phi^{(2)}$ to the linear wave equation with the given data such that both of them satisfy $(\phi^{(i)},\partial_t\phi^{(i)})\in L^\infty([0,T];H^1(\mathbb{R}^n))\times L^\infty([0,T];L^2(\mathbb{R}^n))$ for $i=1,2$, then*

$$\phi^{(1)} = \phi^{(2)}$$

*almost everywhere in $[0,T]\times\mathbb{R}^n$.*

The energy identity in Proposition 4.1 can also be localized in the following sense:

**Proposition 4.4.** *Define, for $0\le t\le R$,*

$$E(t;x,R) := \frac{1}{2}\int_{B(x,R-t)} ((\partial_t\phi)^2 + \sum_{i=1}^n (\partial_i\phi)^2)(t,y)\,dy,$$

*Then $E(t;x,R)$ is a non-increasing function of $t$.*

*Proof.* We start with $\partial_t\phi\Box\phi = 0$ and integrate by parts in the region to the future of $\{t_1\}\times B(x,R)$ and to the past of $(\{t_2\}\times B(x,R-t_2))\cup\left(\cup_{s\in[t_1,t_2]}(\{s\}\times\partial B(x,R-s))\right)$. The boundary terms on $\{t_1\}\times B(x,R-t_1)$ and $\{t_2\}\times B(x,R-t_2)$ give $E(t_1;x,R)$ and $E(t_2;x,R)$ respectively in a way similar to Proposition 4.1. The key point is that the boundary term on $\cup_{s\in[t_1,t_2]}(\{s\}\times\partial B(x,R-s))$ has a definite sign. More precisely,

$$\begin{aligned}&\int_{t_1}^{t_2}\int_{B(x,R-t)}\left(\partial_t\phi(-\partial_t^2\phi+\sum_{i=1}^n\partial_i^2\phi)\right)(t,y)\,dydt\\ =&-E(t_2;x,R)+E(t_1;x,R)-\frac{1}{2}\int_{t_1}^{t_2}\int_{\partial B(x,R-t)}((\partial_t\phi)^2+\sum_{i=1}^n(\partial_i\phi)^2-2\partial_t\phi\sum_{i=1}^n\frac{(y-x)_i}{|y-x|}\partial_i\phi)(t,y)d\sigma(y)dt.\end{aligned}$$

Here, $d\sigma$ is the surface measure on $\partial B(x,R-t)$. It remains to check that the last term is non-positive. To see this, notice that a simple application of Cauchy-Schwarz gives

$$\begin{aligned}2\partial_t\phi\sum_{i=1}^n\frac{(y-x)_i}{|y-x|}\partial_i\phi\leq&(\partial_t\phi)^2+(\sum_{i=1}^n\frac{(y-x)_i}{|y-x|}\partial_i\phi)^2\\ \leq&(\partial_t\phi)^2+(\sum_{i=1}^n(\frac{(y-x)_i}{|y-x|})^2)(\sum_{i=1}^n(\partial_i\phi)^2)\leq(\partial_t\phi)^2+\sum_{i=1}^n(\partial_i\phi)^2,\end{aligned}$$

which implies

$$(\partial_t\phi)^2+\sum_{i=1}^n(\partial_i\phi)^2-2\partial_t\phi\sum_{i=1}^n\frac{(y-x)_i}{|y-x|}\partial_i\phi\geq 0.$$

□

An easy consequence of the localized energy estimates is the finite speed of propagation.

**Corollary 4.5.** *If $(\phi_0,\phi_1)=(0,0)$ in $\{y\in\mathbb{R}^n:|y-x|\leq t\}$, then $\phi(t,x)=0$.*

Notice that we have not yet proved the dispersion of the solutions to the linear wave equation using the energy method. This will be an important topic in later lectures. Instead, we now consider linear estimates for a more general class of (non-constant coefficient!) linear wave equation. We first define the class of equations we consider. We study the following equation in $\mathbb{R}^{n+1}$ with $\phi: I\times\mathbb{R}^n\to\mathbb{R}$:

$$\begin{cases}\partial_\alpha(a^{\alpha\beta}\partial_\beta\phi)=F\\ (\phi,\partial_t\phi)\restriction_{\{t=0\}}=(\phi_0,\phi_1).\end{cases}\tag{4.2}$$

We require $a$ to be a symmetric $(n+1)\times(n+1)$ matrix on $I\times\mathbb{R}^n$ which satisfies

$$\sum_{\alpha,\beta}|a^{\alpha\beta}-m^{\alpha\beta}|<\frac{1}{10}.\tag{4.3}$$

We will also need some regularity assumptions on $a^{\alpha\beta}: I\times\mathbb{R}^n\to\mathbb{R}$, $F: I\times\mathbb{R}^n\to\mathbb{R}$, $\phi_0:\mathbb{R}^n\to\mathbb{R}$ and $\phi_1:\mathbb{R}^n\to\mathbb{R}$. We will state the precise conditions we need later on. For this class of equations, the conservation of energy does not hold, but we can nonetheless show that some appropriately defined "energy" has the property that its growth is controlled. To this end, it is convenient to introduce the notation

$$|\partial\phi|^2:=(\partial_t\phi)^2+\sum_{i=1}^n(\partial_{x^i}\phi)^2.$$

Then we have

**Theorem 4.6.** *Let $\phi$ be a solution to (4.2), then for some constant $C=C(n)>0$, the following energy estimates hold:*

$$\begin{aligned}&\sup_{t\in[0,T]}\|\partial\phi\|_{L^2(\mathbb{R}^n)}(t)\\ \leq& C\left(\|(\phi_0,\phi_1)\|_{\dot{H}^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}+\int_0^T\|F\|_{L^2(\mathbb{R}^n)}(t)dt\right)\exp\left(C\int_0^T\|\partial a\|_{L^\infty(\mathbb{R}^n)}(t)\,dt\right).\end{aligned}$$

*Proof.* The proof is in fact the similar to that for the constant coefficient linear wave equation. We use the identity

$$\partial_t\phi\left(\partial_\alpha(a^{\alpha\beta}\partial_\beta\phi)-F\right)=0 \tag{4.4}$$

and integrate the first term by parts. We consider three different contributions. First, we look at the case when $\alpha=\beta=0$:

$$\begin{aligned}&\int_0^T\int_{\mathbb{R}^n}\partial_t\phi\partial_t(a^{tt}\partial_t\phi)\,dx\,dt\\=&\int_0^T\int_{\mathbb{R}^n}(\partial_t a^{tt})(\partial_t\phi)^2+\frac{1}{2}\partial_t(\partial_t\phi)^2a^{tt}\,dx\,dt\\=&\frac{1}{2}\int_{\{T\}\times\mathbb{R}^n}a^{tt}(\partial_t\phi)^2dx-\frac{1}{2}\int_{\{0\}\times\mathbb{R}^n}a^{tt}(\partial_t\phi)^2\,dx+\frac{1}{2}\int_0^T\int_{\mathbb{R}^3}(\partial_t a^{tt})(\partial_t\phi)^2\,dx\,dt.\end{aligned} \tag{4.5}$$

When we only have $i,j=1,2,...n$, we have the following identity (note that we use the convention where $i,j=1,2,...,n$ and repeated indices are summed over):

$$\begin{aligned}&\int_0^T\int_{\mathbb{R}^n}\partial_t\phi\partial_i(a^{ij}\partial_j\phi)\,dx\,dt\\=&-\frac{1}{2}\int_0^T\int_{\mathbb{R}^n}(\partial_t\partial_i\phi a^{ij}\partial_j\phi+\partial_t\partial_j\phi a^{ij}\partial_i\phi)\,dx\,dt\\=&-\frac{1}{2}\int_0^T\int_{\mathbb{R}^n}\partial_t(\partial_i\phi\partial_j\phi)a^{ij}\,dx\,dt\\=&-\frac{1}{2}\int_{\{T\}\times\mathbb{R}^n}a^{ij}\partial_i\phi\partial_j\phi dx-\frac{1}{2}\int_{\{0\}\times\mathbb{R}^n}a^{ij}\partial_i\phi\partial_j\phi\,dx+\frac{1}{2}\int_0^T\int_{\mathbb{R}^3}(\partial_t a^{ij})(\partial_i\phi\partial_j\phi)\,dx\,dt.\end{aligned} \tag{4.6}$$

Notice that in the derivation above, we have used the symmetry of $a$.

Finally, for the term with $t$ and $i=1,...,n$, we have

$$\begin{aligned}&\int_0^T\int_{\mathbb{R}^n}\partial_t\phi(\partial_i(a^{i0}\partial_t\phi)+\partial_t(a^{i0}\partial_i\phi))\,dx\,dt\\=&\int_0^T\int_{\mathbb{R}^n}\left((\partial_i a^{i0})(\partial_t\phi)^2+a^{i0}\partial_i(\partial_t\phi)^2+(\partial_t a^{i0})(\partial_t\phi)(\partial_i\phi)\right)\,dx\,dt\\=&\int_0^T\int_{\mathbb{R}^n}\left((\partial_i a^{i0})(\partial_t\phi)^2-(\partial_i a^{i0})(\partial_t\phi)^2+(\partial_t a^{i0})(\partial_t\phi)(\partial_i\phi)\right)\,dx\,dt.\end{aligned} \tag{4.7}$$

We now combine the equation (4.4) with the integrated identities (4.5), (4.6) and (4.7) to get

$$\begin{aligned}&\left|\frac{1}{2}\int_{\{T\}\times\mathbb{R}^n}a^{tt}(\partial_t\phi)^2dx-\frac{1}{2}\int_{\{T\}\times\mathbb{R}^n}a^{ij}\partial_i\phi\partial_j\phi dx\right|\\\leq&\left|\frac{1}{2}\int_{\{0\}\times\mathbb{R}^n}a^{tt}(\partial_t\phi)^2dx-\frac{1}{2}\int_{\{0\}\times\mathbb{R}^n}a^{ij}\partial_i\phi\partial_j\phi dx\right|\\&+C\int_0^T\left(\|\partial\phi\|_{L^2(\mathbb{R}^n)}(t)\|F\|_{L^2(\mathbb{R}^n)}(t)+\|\partial a\|_{L^\infty(\mathbb{R}^n)}(t)\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(t)\right)\,dt.\end{aligned} \tag{4.8}$$

By the assumption (4.3), there exists a constant $C>0$ such that the left hand side of (4.8) controls the $L^2$ norm of all derivatives of $\phi$, i.e.,

$$\begin{aligned}&\|\partial\phi\|_{L^2(\mathbb{R}^n)}(T)\\\leq&C\left|\frac{1}{2}\int_{\{T\}\times\mathbb{R}^n}a^{tt}(\partial_t\phi)^2dx-\frac{1}{2}\int_{\{T\}\times\mathbb{R}^n}a^{ij}\partial_i\phi\partial_j\phi dx\right|.\end{aligned} \tag{4.9}$$

Combining (4.8) and (4.9), we get

$$\begin{aligned}&\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(T)\\ \leq& C\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(0)+C\int_0^T\left(\|\partial\phi\|_{L^2(\mathbb{R}^n)}(t)\|F\|_{L^2(\mathbb{R}^n)}(t)+\|\partial a\|_{L^\infty(\mathbb{R}^n)}(t)\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(t)\right)dt.\end{aligned}$$

Now, notice that we can in fact have a stronger estimate on the left hand side which control the supremum of the $L^2$ norm of $|\partial\phi|$ over the interval $[0,T]$. This is because we can perform the argument above in a smaller interval of time. Hence,

$$\begin{aligned}&\sup_{t\in[0,T]}\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(t)\\ \leq& C\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(0)+C\int_0^T\left(\|\partial\phi\|_{L^2(\mathbb{R}^n)}(t)\|F\|_{L^2(\mathbb{R}^n)}(t)+\|\partial a\|_{L^\infty(\mathbb{R}^n)}(t)\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(t)\right)dt.\end{aligned}$$

Now, we can use Cauchy-Schwarz to get

$$\begin{aligned}&\int_0^T\|\partial\phi\|_{L^2(\mathbb{R}^n)}\|F\|_{L^2(\mathbb{R}^n)}dt\\ \leq&\delta\sup_{t\in[0,T]}\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(t)+C\delta^{-1}(\int_0^T\|F\|_{L^2(\mathbb{R}^n)}dt)^2.\end{aligned}$$

Choosing $\delta>0$ sufficiently small, we can therefore absorb the term $\delta\sup_{t\in[0,T]}\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(t)$ to the left hand side to get[6]

$$\begin{aligned}&\sup_{t\in[0,T]}\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(t)\\ \leq& C\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(0)+C(\int_0^T\|F\|_{L^2(\mathbb{R}^n)}(t)dt)^2+C\int_0^T\|\partial a\|_{L^\infty(\mathbb{R}^n)}(t)\|\partial\phi\|^2_{L^2(\mathbb{R}^n)}(t)dt.\end{aligned}$$

The conclusion follows from Gronwall's lemma below. □

We recall here the Gronwall's lemma:

**Lemma 4.7.** *Let $f:\mathbb{R}\to\mathbb{R}$ be a positive continuous function and $g:\mathbb{R}\to\mathbb{R}$ be a positive integrable function such that*

$$f(t)\leq A+\int_0^t f(s)g(s)ds$$

*for some $A\geq 0$ for every $t\in[0,T]$. Then*

$$f(t)\leq A\exp(\int_0^t g(s)ds)$$

*for every $t\in[0,T]$.*

*Proof.* We will actually give two proofs. The main point of the second proof is to illustrate a method that is known as the **bootstrap** method. This method will be important later when we consider nonlinear problems.

(1) Differentiating and using the given inequality, we have

$$\frac{d}{dt}(A+\int_0^t f(s)g(s)ds)=f(t)g(t)\leq g(t)(A+\int_0^t f(s)g(s)ds).$$

Integrating, we obtain

$$\frac{A+\int_0^t f(s)g(s)ds}{A}\leq\exp(\int_0^t g(s)ds).$$

Use the given inequality again to get the desired conclusion.

[6] after fixing $\delta>0$, we then absorb $\delta$ into the constant $C$.

(2) For every $\epsilon > 0$, consider the following condition

$$f(t) \leq (1+\epsilon) A \exp\left((1+\epsilon)\int_0^t g(s)ds\right). \tag{4.10}$$

Define the subset of $[0,T]$ by $B := \{t \in [0,T] : (4.10) \text{ holds for every } s \in [0,t]\}$. We will show that $B$ is non-empty, open and close - thus $B = [0,T]$.

- $B$ is obviously non-empty since $0 \in B$.
- $B$ is obviously closed by the continuity of $f$.
- The only difficult part is to show that $B$ is open. By the continuity of $f$, it suffices to show that if $t \in B$, then we have

$$f(t) \leq A \exp\left((1+\epsilon)\int_0^t g(s)ds\right),$$

i.e., a bound that improves over (4.10). To show this, observe that if $t \in B$, then

$$\begin{aligned} f(t) \leq & A + \int_0^t f(s)g(s)ds \leq A + (1+\epsilon)A\int_0^t g(s)\exp\left((1+\epsilon)\int_0^s g(r)dr\right)ds \\ \leq & A\left(1 + \left(\exp\left((1+\epsilon)\int_0^t g(s)ds\right) - 1\right)\right) = A\exp\left((1+\epsilon)\int_0^t g(s)ds\right). \end{aligned}$$

Therefore, $B = [0,T]$, i.e., (4.10) holds for all $t \in [0,T]$ and for all $\epsilon > 0$. Taking $\epsilon \to 0$, we thus obtain the desired conclusion.

□

Using Gronwall's lemma, we have thus concluded the proof of the energy estimates (Theorem 4.6). As in the case for the constant coefficient linear wave equation, an immediate consequence of the energy estimates is the uniqueness of solutions:

**Corollary 4.8.** *The conclusion of Corollary 4.3 also holds for the more general class of (non-constant coefficient) linear wave equation that is being considered!*

For the constant coefficient linear wave equation (4.1), we know that is $\phi$ is a solution, that $\partial_t\phi$ and $\partial_{x_i}\phi$ are also solutions. As a consequence, say if the initial data are smooth and compactly supported, all higher derivatives for $\phi$ are bounded in $L^2$. In general, of course $\phi$ being a solution does not imply that its derivatives are also solutions. However as a consequence of Theorem 4.6, we can also control higher derivatives of $\phi$ in $L^2$. More precisely, we have the following corollary:

**Corollary 4.9.** *Let $\phi$ be a solution to (4.2) and $k$ be a positive integer. Then for some constant $C = C(n,k) > 0$, the following energy estimates hold:*

$$\begin{aligned} & \sup_{t\in[0,T]} \|(\phi, \partial_t\phi)\|_{H^k(\mathbb{R}^n)\times H^{k-1}(\mathbb{R}^n)}(t) \\ \leq & C(1+T)\Bigg(\|(\phi_0,\phi_1)\|_{H^k(\mathbb{R}^n)\times H^{k-1}(\mathbb{R}^n)} + \int_0^T \|F\|_{H^{k-1}(\mathbb{R}^n)}(t)dt \\ & + C\int_0^T \Big(\sum_{|\alpha|+|\beta|\leq k-1} \|\partial\partial_x^\alpha a \partial\partial_x^\beta \phi\|_{L^2(\mathbb{R}^n)}(t) + \sum_{|\alpha|+|\beta|\leq k-1} \|\partial_x^\alpha b \partial\partial_x^\beta\phi\|_{L^2(\mathbb{R}^n)}(t)\Big)dt\Bigg) \\ & \times \exp\left(C\int_0^T \|\partial a\|_{L^\infty(\mathbb{R}^n)}(t)dt\right). \end{aligned}$$

*Proof.* We first differentiate the equation with $\partial_x^\alpha$ to obtain the bounds for

$$\sum_{|\alpha|\leq k} \sup_{t\in[0,T]} \|(\partial_x^\alpha\phi, \partial_t\partial_x^\alpha\phi)\|_{\dot{H}^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}(t)$$

by the right hand side. It then remains to control

$$\sup_{t\in[0,T]} \|\phi\|_{L^2(\mathbb{R}^n)}(t).$$

On the other hand, it is easy to see (**Exercise**) that

$$\sup_{t\in[0,T]} \|\phi\|_{L^2(\mathbb{R}^n)}(t) \leq C\left(\|\phi_0\|_{L^2(\mathbb{R}^n)} + \int_0^T \|\partial_t\phi\|_{L^2(\mathbb{R}^n)}(t)dt\right),$$

which then gives the desired conclusion. □

## 5. Existence of solutions to general (non-constant coefficient) linear wave equation

In the previous lecture, we showed that if a solution exists, then it must obey certain energy estimates. In particular, the existence of solution is assumed. Now, we show that in fact using the energy estimates together with some abstract theory, we can obtain existence of solutions. For technical reasons, it is convenient to look at a smaller subclass than the previous section. Consider the equation in $\mathbb{R}^{n+1}$ for $\phi : I\times\mathbb{R}^n\to\mathbb{R}$

$$\begin{cases}\partial_\alpha(a^{\alpha\beta}\partial_\beta\phi) = F\\ (\phi,\partial_t\phi)\restriction_{\{t=0\}} = (\phi_0,\phi_1)\in H^k(\mathbb{R}^n)\times H^{k-1}(\mathbb{R}^n),\end{cases} \tag{5.1}$$

where $k\in\mathbb{N}$. We require $a^{\alpha\beta} = a^{\beta\alpha}$ for $\alpha,\beta\in\{0,1,...,n\}$ and

$$\sum_{\alpha,\beta}|a^{\alpha\beta} - m^{\alpha\beta}| < \frac{1}{10}.$$

To simplify the analysis below, we assume that $a$ and all of its derivatives (of all orders) are bounded in $[0,T]\times\mathbb{R}^n$. $F : [0,T]\times\mathbb{R}^n\to\mathbb{R}^n$ will be assumed to be in the function space $F\in L^1([0,T];H^{k-1}(\mathbb{R}^n))$.

The main tool that we will use is the Hahn-Banach theorem from functional analysis. Let us recall the following version of the Hahn–Banach theorem:

**Theorem 5.1** (Hahn–Banach theorem)**.** *Let $X$ be a normed vector space and $Y\subseteq X$ be a subspace with the norm $\|y\|_Y = \|y\|_X$ for every $y\in Y$. Suppose $f\in Y^*$ is a bounded linear functional on $Y$, then there exists $\tilde f\in X^*$ such that $\tilde f\restriction_Y = f$ and $\|\tilde f\|_{X^*} = \|f\|_{Y^*}$.*

Before we state and prove the existence result, we need the following lemma:

**Lemma 5.2.** *Let*

$$L^*\psi := \partial_\alpha(a^{\alpha\beta}\partial_\beta\psi).$$

*be defined as the (formal) adjoint of $L$, which is defined by*

$$L\phi := \partial_\alpha(a^{\alpha\beta}\partial_\beta\phi).$$

*Suppose $\psi\in C_c^\infty((-\infty,T)\times\mathbb{R}^n)$, then for every $m\in\mathbb{Z}$, there exists $C = C(m,T,a,b) > 0$ such that*

$$\|\psi\|_{H^m(\mathbb{R}^n)}(t) \leq C\int_t^T \|L^*\psi\|_{H^{m-1}(\mathbb{R}^n)}(s)ds$$

*for every $t\in[0,T]$.*

*Proof.* For $m\geq 1$, this is simply a consequence of the energy estimates (Corollary 4.9). We now carry out an induction for the cases $m\leq 0$. Assume that the result holds for some $m_0+2$, we wish to prove the same result (with a possibly different constant) for $m_0$. To this end, consider

$$\Psi = (1-\Delta)^{-1}\psi,$$

defined using the Fourier transform. Then, there exist $C>0$ depending on $m_0$, $T$, $a$ and $b$ such that

$$|L^*\psi - (1-\Delta)L^*\Psi| = |L^*(1-\Delta)\Psi - (1-\Delta)L^*\Psi| \leq C(\sum_{1\leq|\alpha|\leq 3}|\partial^\alpha\Psi|).$$

Therefore,

$$\|L^*\Psi\|_{H^{m_0+1}(\mathbb{R}^3)}(t) \leq C\left(\|L^*\psi\|_{H^{m_0-1}(\mathbb{R}^3)}(t) + \|\Psi\|_{H^{m_0+2}(\mathbb{R}^3)}(t)\right).$$

Therefore, by the induction hypothesis,

$$\begin{aligned}&\|\Psi\|_{H^{m_0+2}(\mathbb{R}^3)}(t)\\ \leq &C\int_t^T \left(\|L^*\psi\|_{H^{m_0-1}(\mathbb{R}^3)}(s)+\|\Psi\|_{H^{m_0+2}(\mathbb{R}^3)}(s)\right)ds\\ \leq &C\int_t^T \|L^*\psi\|_{H^{m_0-1}(\mathbb{R}^3)}(s)ds,\end{aligned}$$

where in the last line we have used Gronwall's lemma, noting that the constant is allowed to depend on $T$. This implies

$$\|\psi\|_{H^{m_0}(\mathbb{R}^3)}(t)\leq C\|\Psi\|_{H^{m_0+2}(\mathbb{R}^3)}(t)\leq C\int_t^T \|L^*\psi\|_{H^{m_0-1}(\mathbb{R}^3)}(s)ds.$$

□

We are now ready for the main result in this lecture:

**Theorem 5.3** (Existence of solutions to the linear equation)**.** *Let $k\in\mathbb{N}$. Given $F\in L^1([0,T];H^{k-1}(\mathbb{R}^n))$, there exists a solution*

$$(\phi,\partial_t\phi)\in L^\infty([0,T];H^k(\mathbb{R}^n))\times L^\infty([0,T];H^{k-1}(\mathbb{R}^n))$$

*solving* (5.1)*.*

*Proof.* We begin with the case where $(\phi_0,\phi_1)=(0,0)$. Let $L^*(C^\infty_c((-\infty,T)\times\mathbb{R}^n))$ denote the image of $C^\infty_c((-\infty,T)\times\mathbb{R}^n)$ under the map $L^*: C^\infty_c((-\infty,T)\times\mathbb{R}^n)\to C^\infty_c((-\infty,T)\times\mathbb{R}^n)$. For every element $L^*\psi\in L^*(C^\infty_c((-\infty,T)\times\mathbb{R}^n))$, define a map to $\mathbb{R}$ by

$$L^*\psi\mapsto\int_0^T\int_{\mathbb{R}^n}\psi F\,dx\,dt=:<F,\psi>.$$

(Note that this is well-defined because of uniqueness for $L^*\psi=f$ with zero data at $t=T$, a fact that we proved using energy estimates.) Notice that by the assumption on $F$ and Lemma 5.2, we have the bound

$$\begin{aligned}|\int_0^T\int_{\mathbb{R}^n}\psi F\,dx\,dt|\leq &C(\int_0^T\|F\|_{H^{k-1}(\mathbb{R}^n)}(t)dt)(\sup_{t\in[0,T]}\|\psi\|_{H^{-k+1}(\mathbb{R}^n)})\\ \leq &C\int_0^T\|L^*\psi\|_{H^{-k}(\mathbb{R}^n)}(s)ds.\end{aligned}$$

Using the Hahn-Banach theorem (Theorem 5.1), there exists a function $\phi\in(L^1((-\infty,T);H^{-k}(\mathbb{R}^n)))^*=L^\infty((-\infty,T);H^k(\mathbb{R}^n))$ with $\phi=0$ for $t<0$ that extends the map above, i.e.,

$$<F,\psi>=<\phi,L^*\psi>$$

for every $\psi\in C^\infty_c((-\infty,T)\times\mathbb{R}^n)$. Therefore, $\phi$ is a solution in the sense of distribution. Finally, we use the equation to show that $\phi\in C^1([0,T];L^2(\mathbb{R}^n))$ and therefore $(\phi,\partial_t\phi)\restriction_{\{t=0\}}=(0,0)$. This concludes the proof in the special case where $(\phi_0,\phi_1)=(0,0)$.

We now consider the general case where $(\phi_0,\phi_1)\in H^k(\mathbb{R}^n)\times H^{k-1}(\mathbb{R}^n)$ are prescribed. Let $u$ be a function in $[0,T]\times\mathbb{R}^n$ such that

$$(u,\partial_t u)\restriction_{\{t=0\}}=(\phi_0,\phi_1).$$

At the same time, solve

$$L\eta=F-Lu$$

with initial data

$$(\eta,\partial_t\eta)\restriction_{\{t=0\}}=(0,0).$$

Then letting $\phi:=\eta+u$ gives the desired solution. It is easy to check using the energy estimates again that $(\phi,\partial_t\phi)$ is indeed in the desired function space. □

*Remark* 5.4. Notice that the above theorem gives only the existence of solutions in the sense of distribution. Nevertheless, using the Sobolev embedding theorem, if we assume that the initial data is sufficiently regular, then in fact the solution can be understood classically.

We recall here the following standard Sobolev embedding theorem which we referred to in the remark above. It will also make a few more appearances later.

**Theorem 5.5** (Sobolev embedding theorem)**.** *For every $s > \frac{n}{2}$, there exists $C = C(n,s) > 0$ such that*

$$\|\phi\|_{L^\infty(\mathbb{R}^n)} \leq C\|\phi\|_{H^s(\mathbb{R}^n)}$$

*for all $\phi \in H^s(\mathbb{R}^n)$.*

## 6. Local theory for nonlinear wave equations

We consider quasilinear wave equations in $\mathbb{R}^{n+1}$ for $\phi : I \times \mathbb{R}^n \to \mathbb{R}$ of the form

$$\begin{cases} \partial_\alpha(a^{\alpha\beta}(\phi)\partial_\beta\phi) = F(\phi, \partial\phi) \\ (\phi, \partial_t\phi) \restriction_{\{t=0\}} = (\phi_0, \phi_1) \in H^k(\mathbb{R}^n) \times H^{k-1}(\mathbb{R}^n), \end{cases} \tag{6.1}$$

where $k \in \mathbb{N}$ is to be specified later. We require that $a^{\alpha\beta} = a^{\beta\alpha}$ for $\alpha, \beta \in \{0, 1, ..., n\}$ and satisfy

$$\sum_{\alpha,\beta} |a^{\alpha\beta} - m^{\alpha\beta}| \leq \frac{1}{10}, \quad a(0) = 0, \quad F(0,0) = 0 \tag{6.2}$$

and[7]

$$a^{\alpha\beta} : \mathbb{R} \to \mathbb{R}, \quad F : \mathbb{R} \times \mathbb{R}^{n+1} \to \mathbb{R} \text{ are smooth functions of their arguments.} \tag{6.3}$$

As a consequence of (6.3), we have

$$\sum_{\alpha,\beta} \sum_{|\gamma|\leq N} \sup_{|x|\leq A} |\partial_x^\gamma(a^{\alpha\beta})|(x) \leq C_{A,N} \tag{6.4}$$

and

$$\sum_{|\gamma|\leq N} \sup_{|x|,\|p\|\leq A} |\partial_{x,p}^\gamma F|(x,p) \leq C_{A,N}. \tag{6.5}$$

We now state the main theorem of this section, which is a local existence and uniqueness theorem for this general class of equations.

**Theorem 6.1.** *Fix $a$ and $F$ satisfying the assumptions above.*

*(1) (Existence and uniqueness of local-in-time solutions) There exists*[8]

$$T = T(\|\phi_0\|_{H^{n+2}(\mathbb{R}^n)}, \|\phi_1\|_{H^{n+1}(\mathbb{R}^n)}) > 0$$

*such that there exists a (classical) solution $\phi$ to* (6.1) *with*

$$(\phi, \partial_t\phi) \in L^\infty([0,T]; H^{n+2}(\mathbb{R}^n)) \times L^\infty([0,T]; H^{n+1}(\mathbb{R}^n)).$$

*Moreover, the solution in unique in the function space*

$$(\phi, \partial_t\phi) \in L^\infty([0,T]; H^{n+2}(\mathbb{R}^n)) \times L^\infty([0,T]; H^{n+1}(\mathbb{R}^n)).$$

*(2) (Continuous dependence on initial data) Let $\phi_0^{(i)}$, $\phi_1^{(i)}$ be sequences of functions such that $\phi_0^{(i)} \to \phi_0$ in $H^{n+2}(\mathbb{R}^n)$ and $\phi_1^{(i)} \to \phi_1$ in $H^{n+1}(\mathbb{R}^n)$ as $i \to \infty$. Then taking $T > 0$ sufficiently small, we have*

$$\|(\phi^{(i)} - \phi, \partial_t(\phi^{(i)} - \phi))\|_{L^\infty([0,T];H^s(\mathbb{R}^n))\times L^\infty([0,T];H^{s-1}(\mathbb{R}^n))} \to 0$$

*as $i \to \infty$ for every $1 \leq s < n+2$. (Here, $\phi$ is the solution arising from data $(\phi_0, \phi_1)$ and $\phi^{(i)}$ is the solution arising from data $(\phi_0^{(i)}, \phi_1^{(i)})$.)*

*Remark* 6.2. An evolution equation is said to be well-posed in the sense of Hadamard if existence, uniqueness of solutions and continuous dependence on initial data hold. Theorem 6.1 therefore implies that the equation (6.1) is locally well-posed.

We now turn to the proof of the theorem:

[7]Here, it is understood that $a^{\alpha\beta}$ are functions $: \mathbb{R} \to \mathbb{R}$ for every $\alpha, \beta \in \{0, 1, ..., n\}$.

[8]Of course $T$ also depends on $a$ and $F$ but we will considered them fixed for this theorem.

*Proof.* (1) This part of the theorem is proved by Picard's iteration. By a density argument, it suffices to assume that $(\phi_0, \phi_1) \in \mathcal{S} \times \mathcal{S}$. Define a sequence of (smooth) functions $\phi^{(i)}$ for $i \geq 1$ such that

$$\phi^{(1)} = 0$$

and $\phi^{(i)}$ is defined iteratively for $i \geq 2$ as the unique solution to

$$\begin{cases} \partial_\alpha (a^{\alpha\beta}(\phi^{(i-1)})\partial_\beta \phi^{(i)}) = F(\phi^{(i-1)}, \partial\phi^{(i-1)}) \\ (\phi^{(i)}, \partial_t \phi^{(i)}) \restriction_{\{t=0\}} = (\phi_0, \phi_1). \end{cases} \tag{6.6}$$

We will show two properties of the sequence $(\phi^{(i)}, \partial_t\phi^{(i)})$. First, for $T > 0$ sufficiently small (depending only on $\|\phi_0\|_{H^{n+2}(\mathbb{R}^n)}$ and $\|\phi_1\|_{H^{n+1}(\mathbb{R}^n)}$), the sequence is uniformly (in $i$) bounded in $L^\infty([0,T]; H^{n+2}(\mathbb{R}^n)) \times L^\infty([0,T]; H^{n+1}(\mathbb{R}^n))$, i.e.,

$$\|(\phi^{(i)}, \partial_t\phi^{(i)})\|_{L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))} \leq C \tag{6.7}$$

for some $C > 0$ (depending on $n$, $F$, $a$, $\phi_0$ and $\phi_1$, but independent of $i$). Then, we show that for $T > 0$ chosen to be smaller if necessary, $(\phi^{(i)}, \partial_t\phi^{(i)})$ is a Cauchy sequence in $L^\infty([0,T]; H^1(\mathbb{R}^n)) \times L^\infty([0,T]; L^2(\mathbb{R}^n))$.

(a) We begin with the first part, the idea is to use the energy estimates. Obviously, it suffices to prove that there exists some $A > 0$ such that

$$\begin{aligned} &\|(\phi^{(i-1)}, \partial_t\phi^{(i-1)})\|_{L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))} \leq A \\ \implies &\|(\phi^{(i)}, \partial_t\phi^{(i)})\|_{L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))} \leq A. \end{aligned}$$

To this end, we assume

$$\|(\phi^{(i-1)}, \partial_t\phi^{(i-1)})\|_{L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))} \leq A \tag{6.8}$$

for some $A > 0$ to be chosen below. In fact, it is convenient to use the bootstrap method[9] and assume also that

$$\|(\phi^{(i)}, \partial_t\phi^{(i)})\|_{L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))} \leq 4A. \tag{6.9}$$

Of course, we can make this assumption as long as at the end we can improve the constant in (6.9).

Our next goal is to apply the energy estimates in Corollary 4.9 to improve the bound (6.9). Recall that to apply Corollary 4.9, we need bounds on

$$\int_0^T \sum_{0\leq|\alpha|\leq n+1} \|\partial_x^\alpha F(\phi^{(i-1)}, \partial\phi^{(i-1)})\|_{L^2(\mathbb{R}^n)}(t)\, dt, \tag{6.10}$$

$$\int_0^T \sum_{|\gamma|+|\sigma|\leq n+1} \sum_{\alpha,\beta} \|\partial\partial_x^\gamma \left(a^{\alpha\beta}(\phi^{(i-1)})\right) \partial\partial_x^\sigma \phi^{(i)}\|_{L^2(\mathbb{R}^n)}(t)\, dt \tag{6.11}$$

and

$$\int_0^T \sum_{\substack{|\gamma|+|\sigma|\leq n+1 \\ |\sigma|\leq n}} \sum_{\alpha,\beta} \|\partial_x^\gamma \left(a^{\alpha\beta}(\phi^{(i-1)})\right) \partial\partial\partial_x^\sigma \phi^{(i)}\|_{L^2(\mathbb{R}^n)}(t)\, dt. \tag{6.12}$$

To obtain such bounds, we need to use the smoothness of $F$ and $a$ as well as the Sobolev embedding theorem (Theorem 5.5). First, for $F$ (i.e., for the term (6.10)), we want to prove the following bounds

**Claim 6.3.** There exists $B = B(A, n, F) > 0$ such that for $t \in [0, T]$, we have

$$\sum_{0\leq|\alpha|\leq n+1} \|\partial_x^\alpha F(\phi^{(i-1)}, \partial\phi^{(i-1)})\|_{L^2}(t) \leq B. \tag{6.13}$$

[9]Recall the discussions of the bootstrap method in the proof of Lemma 4.7. What it allows us to do is to prove an estimate by first assuming it as a **bootstrap assumption** and then improving the bound.

*Proof of Claim.* To see this, we first expand the derivatives and note that by the chain rule, Theorem 5.5 and (6.5),

$$\begin{aligned}&\sum_{0\le|\alpha|\le n+1}\|\partial_x^\alpha F(\phi^{(i-1)},\partial\phi^{(i-1)})\|_{L^2(\mathbb{R}^n)}(t)\\ \le& C\Bigg(\sum_{0\le|\alpha|\le n+1}\|\partial\partial_x^\alpha\phi^{(i-1)}\|_{L^2(\mathbb{R}^n)}(t)+\sum_{0\le|\alpha_1|+|\alpha_2|\le n+1}\|\partial\partial_x^{\alpha_1}\phi^{(i-1)}\partial\partial_x^{\alpha_2}\phi^{(i-1)}\|_{L^2(\mathbb{R}^n)}(t)\\ &+\text{cubic terms}+...+\text{terms of order } n+1\Bigg).\end{aligned}$$

Consider for instance the quadratic term

$$\sum_{0\le|\alpha_1|+|\alpha_2|\le n+1}\|\partial\partial_x^{\alpha_1}\phi^{(i-1)}\partial\partial_x^{\alpha_2}\phi^{(i-1)}\|_{L^2(\mathbb{R}^n)}(t).$$

Notice now that either $|\alpha_1|\le\frac{n+1}{2}$ or $|\alpha_2|\le\frac{n+1}{2}$. So we can assume without loss of generality that $|\alpha_1|\le\frac{n+1}{2}$. On the other hand, since $\frac{n+1}{2}+\frac{n}{2}<n+1$, we can apply Theorem 5.5 to get

$$\sum_{0\le|\alpha_1|\le\frac{n+1}{2}}\|\partial\partial_x^{\alpha_1}\phi^{(i-1)}\|_{L^\infty(\mathbb{R}^n)}\le C\sum_{0\le|\alpha|\le n+1}\|\partial\partial_x^{\alpha}\phi^{(i-1)}\|_{L^2(\mathbb{R}^n)}(t).$$

Therefore,

$$\sum_{0\le|\alpha_1|+|\alpha_2|\le n+1}\|\partial\partial_x^{\alpha_1}\phi^{(i-1)}\partial\partial_x^{\alpha_2}\phi^{(i-1)}\|_{L^2(\mathbb{R}^n)}(t)\le C\sum_{0\le|\alpha|\le n+1}\|\partial\partial_x^{\alpha}\phi^{(i-1)}\|_{L^2(\mathbb{R}^n)}(t)^2.$$

In a similar manner, we can treat all the higher order terms to get

$$\begin{aligned}&\sum_{0\le|\alpha|\le n+1}\|\partial_x^\alpha F(\phi^{(i-1)},\partial\phi^{(i-1)})\|_{L^2}(t)\\ \le& C\Bigg(1+\|\phi^{(i-1)}\|_{L^2(\mathbb{R}^n)}+\sum_{0\le|\alpha|\le n+1}\|\partial\partial_x^\alpha\phi^{(i-1)}\|_{L^2(\mathbb{R}^n)}(t)\Bigg)^{n+1}\le B,\end{aligned}$$

where in the last line we have used (6.8). □

In a similar fashion, we can also obtain bound the terms (6.11) and (6.12):

**Claim 6.4.**

$$|(6.11)|+|(6.12)|\le B$$

for some $B=B(A,n,F)>0$

We will not carried out the details of the proof of this claim, but just note that it is carried out in a similar manner as the proof of Claim 6.3 and that we need to use the bootstrap assumption (6.9) in addition to the induction hypothesis (6.8). Notice that for the term (6.12), there can potentially be two $\partial_t$-derivatives, which however can be exchanged to either $\partial_t\partial_x$ or $\partial_x\partial_x$ after using the equation and applying the estimates in Claim 6.3.

Moreover, we have the following bound using Proposition 5.5 and (6.8):

**Claim 6.5.**

$$\sum_{|\gamma|=1}\|\partial_x\left(a^{\alpha\beta}(\phi^{(i-1)})\right)\|_{L^\infty(\mathbb{R}^n)}\le B.$$

We now apply the energy estimates in Corollary 4.9 and use the estimates in Claims 6.3, 6.4 and 6.5 to obtain

$$\begin{aligned}&\|(\phi^{(i)},\partial_t\phi^{(i)})\|_{L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))}\\ \le& C\left(\|(\phi_0,\phi_1)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}+BT+CBT^2\right)\times\exp(CBT).\end{aligned}$$

In order to fix notations, let's call the constant in the above inequality $C_0$ and assume without loss of generality that $C_0\ge 2$. Here is the key point: while $B$ can be very large and depends on $A$, we can choose $T>0$ to be sufficiently small (depending on $A$) such that

$$(BT+C_0BT^2)\exp(C_0BT)\le\|(\phi_0,\phi_1)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}$$

and

$$\exp(C_0 BT) \leq 2.$$

With $T$ chosen as above, we thus have

$$\|(\phi^{(i)}, \partial_t\phi^{(i)})\|_{L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))} \leq 4C_0\|(\phi_0,\phi_1)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}.$$

Now, let's choose $A = 4C_0\|(\phi_0,\phi_1)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}$. We have thus shown the desired implication

$$\begin{aligned}&\|(\phi^{(i-1)}, \partial_t\phi^{(i-1)})\|_{L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))} \leq A\\ \implies &\|(\phi^{(i)}, \partial_t\phi^{(i)})\|_{L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))} \leq A.\end{aligned}$$

for this $A$, provided that $T > 0$ is sufficiently small depending on $A$. In particular, we have improved the constant in the bootstrap assumption (6.9). On the other hand, $A$ depends only on $\|(\phi_0,\phi_1)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}$ and therefore $T$ indeed depends only on $\|(\phi_0,\phi_1)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}$, as desired.

We have thus finished the first part of the proof.

(b) We now move to the second part in which we show that the sequence is Cauchy in a larger space $L^\infty([0,T];H^1(\mathbb{R}^n)) \times L^\infty([0,T];L^2(\mathbb{R}^n))$. To this end, for every $i \geq 3$, we consider the equation for $\phi^{(i)} - \phi^{(i-1)}$:

$$\begin{aligned}&\partial_\alpha(a^{\alpha\beta}(\phi^{(i-1)})\partial_\beta(\phi^{(i)} - \phi^{(i-1)}))\\ =&-\partial_\alpha((a^{\alpha\beta}(\phi^{(i-1)}) - a^{\alpha\beta}(\phi^{(i-2)}))\partial_\beta\phi^{(i-1)})\\ &+F(\phi^{(i-1)},\partial\phi^{(i-1)}) - F(\phi^{(i-2)},\partial\phi^{(i-2)}).\end{aligned} \tag{6.14}$$

Since $a$ is continuously differentiable, we can use (6.4), the bound (6.7) obtained in part (a) and mean value theorem to show that there exists some $C > 0$ (depending on initial data but *independent* of $i$ and $T$) such that

$$\left|\partial_\alpha((a^{\alpha\beta}(\phi^{(i-1)}) - a^{\alpha\beta}(\phi^{(i-2)}))\partial_\beta\phi^{(i-1)})\right| \leq C\left|\partial(\phi^{(i-1)} - \phi^{(i-2)})\right|$$

which implies using (6.7) again that

$$\|\partial_\alpha((a^{\alpha\beta}(\phi^{(i-1)}) - a^{\alpha\beta}(\phi^{(i-2)}))\partial_\beta\phi^{(i-1)})\|_{L^2(\mathbb{R}^n)}(t) \leq C\|\partial(\phi^{(i-1)} - \phi^{(i-2)})\|_{L^2(\mathbb{R}^n)}(t)$$

for every $t \in [0,T]$. Similarly, using (6.5), (6.7) and the mean value theorem, we have

$$\|F(\phi^{(i-1)},\partial\phi^{(i-1)}) - F(\phi^{(i-2)},\partial\phi^{(i-2)})\|_{L^2(\mathbb{R}^n)}(t) \leq C\|\partial(\phi^{(i-1)} - \phi^{(i-2)})\|_{L^2(\mathbb{R}^n)}(t)$$

for some $C = C(\|(\phi_0,\phi_1)\|_{H^k(\mathbb{R}^n)\times H^{k-1}(\mathbb{R}^n)}) > 0$. It is important that $C$ is *independent* of $i$. Also, using the bound in part (a), we have,

$$\|\partial(a^{\alpha\beta}(\phi^{(i-1)}))\|_{L^\infty(\mathbb{R}^n)}(t) \leq C.$$

Therefore, applying Corollary 4.9 to (6.14), (noticing that the data for $\phi^{(i)}$ and $\phi^{(i-1)}$ coincide if $i \geq 2$), we get

$$\begin{aligned}&\sup_{t\in[0,T]}\|(\phi^{(i)} - \phi^{(i-1)}, \partial_t\phi^{(i)} - \partial_t\phi^{(i-1)})\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}(t)\\ \leq &CT\sup_{t\in[0,T]}\|(\phi^{(i-1)} - \phi^{(i-2)}, \partial_t\phi^{(i-1)} - \partial_t\phi^{(i-2)})\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}(t).\end{aligned}$$

By (6.7),

$$\sup_{t\in[0,T]}\|(\phi^{(2)} - \phi^{(1)}, \partial_t\phi^{(2)} - \partial_t\phi^{(1)})\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}(t) \leq C_1.$$

Therefore, choosing $T$ to be sufficiently small, we have for $i \geq 3$,

$$\sup_{t\in[0,T]}\|\phi^{(i)} - \phi^{(i-1)}\|_{H^1(\mathbb{R}^n)}(t) \leq \frac{1}{2}\sup_{t\in[0,T]}\|\phi^{(i-1)} - \phi^{(i-2)}\|_{H^1(\mathbb{R}^n)}(t),$$

which implies

$$\sup_{t\in[0,T]}\|\phi^{(i)} - \phi^{(i-1)}\|_{H^1(\mathbb{R}^n)}(t) \leq 2^{-i+2}C_1.$$

From this the Cauchy property follows straightforwardly.
Now since $(\phi^{(i)}, \partial_t\phi^{(i)})$ is Cauchy in $L^\infty([0,T];H^1(\mathbb{R}^n)) \times L^\infty([0,T];L^2(\mathbb{R}^n))$, there exists a limit $(\phi,\partial_t\phi) \in L^\infty([0,T];H^1(\mathbb{R}^n))\times L^\infty([0,T];L^2(\mathbb{R}^n))$. The uniform (in $i$) bounds in $L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))$ implies that the limit in fact lies in the smaller space $(\phi,\partial_t\phi) \in L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))$. (One may argue as follows: for almost every $t \in [0,T]$, $(\phi^{(i)},\partial_t\phi^{(i)})(t) \in H^{n+2}(\mathbb{R}^n)) \times H^{n+1}(\mathbb{R}^n)$ uniform in $i$ and therefore by Banach-Alaoglu, a subsequence has a weak limit in $H^{n+2}(\mathbb{R}^n)) \times H^{n+1}(\mathbb{R}^n)$. By uniqueness of limits, this limit must agree with $(\phi,\partial_t\phi)(t)$.) This concludes the proof of existence. The proof of uniqueness can be carried out easily by considering the equation satisfied by the difference of two solutions. However, we will not carry out this in detail since uniqueness can alternatively be derived from the continuous dependence of initial data that we will prove immediately below.

(2) Pick some $i \in \mathbb{N}$ sufficiently large and we bound the difference $\phi^{(i)} - \phi$ using the energy estimates. The equation for the difference is given as follows:

$$\begin{aligned}&\partial_\alpha(a^{\alpha\beta}(\phi^{(i)})\partial_\beta(\phi^{(i)}-\phi))\\ =&-\partial_\alpha((a^{\alpha\beta}(\phi^{(i)})-a^{\alpha\beta}(\phi))\partial_\beta\phi)+F(\phi^{(i)},\partial\phi^{(i)})-F(\phi,\partial\phi).\end{aligned}$$

Applying the energy estimates, we get that the following holds for some $C > 0$ independent of $i$ and for all $t \in [0,T]$

$$\begin{aligned}&\sup_{s\in[0,t]}\|(\phi^{(i)}-\phi,\partial_t\phi^{(i)}-\partial_t\phi)\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}(s)\\ \leq& C\left(\|(\phi_0^{(i)}-\phi_0,\phi_1^{(i)}-\phi_1)\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}+\int_0^t\|\partial_\alpha((a^{\alpha\beta}(\phi^{(i)})-a^{\alpha\beta}(\phi))\partial_\beta\phi)\|_{L^2(\mathbb{R}^n)}(t)dt\right.\\ &\left.+\int_0^t\|F(\phi^{(i)},\partial\phi^{(i)})-F(\phi,\partial\phi)\|_{L^2(\mathbb{R}^n)}(t)dt\right).\end{aligned}$$

Now, applying the bounds for $\phi^{(i)}$ and $\phi$ that we obtained in part (1) of the theorem, we get

$$\begin{aligned}&\sup_{s\in[0,t]}\|(\phi^{(i)}-\phi,\partial_t\phi^{(i)}-\partial_t\phi)\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}(s)\\ \leq& C\left(\|(\phi_0^{(i)}-\phi_0,\phi_1^{(i)}-\phi_1)\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}+\int_0^t\|(\phi^{(i)}-\phi,\partial_t\phi^{(i)}-\partial_t\phi)\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}(t)dt\right).\end{aligned}$$

By Gronwall's inequality, we thus have that for some $C = C(T) > 0$,

$$\sup_{t\in[0,T]}\|(\phi^{(i)}-\phi,\partial_t\phi^{(i)}-\partial_t\phi)\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}(t)\leq C\|(\phi_0^{(i)}-\phi_0,\phi_1^{(i)}-\phi_1)\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}.$$

Since the right hand side $\to 0$ as $i \to \infty$, we get that as $i \to \infty$,

$$\sup_{t\in[0,T]}\|(\phi^{(i)}-\phi,\partial_t\phi^{(i)}-\partial_t\phi)\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}(t)\to 0.$$

To obtain the result in general for $1 \leq s < n+2$, simply observe that

$$\begin{aligned}&\sup_{t\in[0,T]}\|(\phi^{(i)}-\phi,\partial_t\phi^{(i)}-\partial_t\phi)\|_{H^s(\mathbb{R}^n)\times H^{s-1}(\mathbb{R}^n)}(t)\\ \leq& C\sup_{t\in[0,T]}\left(\|(\phi^{(i)}-\phi,\partial_t\phi^{(i)}-\partial_t\phi)\|_{H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)}(t)\right)^{\frac{n+2-s}{n+1}}\\ &\times\left(\|(\phi^{(i)}-\phi,\partial_t\phi^{(i)}-\partial_t\phi)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}(t)\right)^{\frac{s-1}{n+1}}\to 0.\end{aligned}\tag{6.15}$$

Now that the estimate (6.15) because of the general functional inequality

$$\|\eta\|_{H^s(\mathbb{R}^n)}\leq C\|\eta\|_{H^{s_1}(\mathbb{R}^n)}^{\theta_1}\|\eta\|_{H^{s_2}(\mathbb{R}^n)}^{\theta_2}\tag{6.16}$$

for some $C = C(s_1,s_2,s,n) > 0$, where $0 \leq s_1 \leq s \leq s_2$, $\theta_1+\theta_2 = 1$ and $\theta_1 s_1 + \theta_2 s_2 = s$. The functional inequality (6.16) can be proven by taking the Fourier transform and using Hölder's inequality (**Exercise**).

□

In Theorem 6.1, we have constructed a unique $L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))$ solution for $T>0$ sufficiently small. For physical theories, we in fact want to construct smooth solutions. It is therefore useful to have the following "persistence of regularity" statement, which states that as long as we have a $L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n))$ solution, "all higher regularity of the initial data is propagated".

**Theorem 6.6.** *(Persistence of regularity) Given initial data $(\phi_0,\phi_1)\in H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)$ to (6.1), let*

$$T_* := \sup\{T>0 : \textit{there exists a unique } L^\infty([0,T];H^{n+2}(\mathbb{R}^n))\times L^\infty([0,T];H^{n+1}(\mathbb{R}^n)) \textit{ solution}\}>0$$

*be the maximal time of existence. (Notice that $T_*>0$ by Theorem 6.1.)*

(1) *If $(\phi_0,\phi_1)\in H^m(\mathbb{R}^n)\times H^{m-1}(\mathbb{R}^n)$ for some $m\in\mathbb{N}$ with $m>n+2$, then the solution $(\phi,\partial_t\phi)\in L^\infty([0,T];H^m(\mathbb{R}^n))\times L^\infty([0,T];H^{m-1}(\mathbb{R}^n))$ for every $T<T_*$.*
(2) *If $(\phi_0,\phi_1)\in\cap_{m=1}^\infty(H^m(\mathbb{R}^n)\times H^{m-1}(\mathbb{R}^n))$, then the solution is smooth in $[0,T_*)\times\mathbb{R}^n$.*

*Proof.* We first show the statement for persistence of $H^m$ regularity. We proceed inductively in $m$. Assume that the conclusion holds for some $m-1\geq n+2$. We will show the conclusion for $m$. Again, we use energy estimates:

$$\begin{aligned}&\sup_{t\in[0,T)}\|(\phi,\partial_t\phi)\|_{H^m(\mathbb{R}^n)\times H^{m-1}(\mathbb{R}^n)}\\ \leq& C\left(\|(\phi_0,\phi_1)\|_{H^m(\mathbb{R}^n)\times H^{m-1}(\mathbb{R}^n)}+\int_0^T\big(\|F\|_{H^{m-1}(\mathbb{R}^n)}(t)+\|\partial a\|_{H^{m-1}(\mathbb{R}^n)}(t)\big)\,dt\right)\end{aligned}$$

Our goal is to show that in fact

$$\|F\|_{H^{m-1}(\mathbb{R}^n)}(t)+\sum_{|\gamma|+|\sigma|\leq k-1}\sum_{\alpha,\beta}\|\partial\partial_x^\gamma\left(a^{\alpha\beta}(\phi)\right)\partial\partial_x^\sigma\phi\|_{L^2}(t)$$

can be estimated *linearly* in terms of $\sup_{t\in[0,T)}\|(\phi,\partial_t\phi)\|_{H^m(\mathbb{R}^n)\times H^{m-1}(\mathbb{R}^n)}$, which then allows us to apply Gronwall's lemma to obtain the desired estimates in $H^m(\mathbb{R}^n)\times H^{m-1}(\mathbb{R}^n)$. This is possible since $m$ is sufficiently large, and we can use Sobolev embedding together with the induction hypothesis to control the lower order term. More precisely, we have

$$\begin{aligned}&\sum_{0\leq|\alpha|\leq m-1}\|\partial_x^\alpha F(\phi,\partial\phi)\|_{L^2}(t)\\ \leq& C\Bigg(\sum_{0\leq|\alpha|\leq m}\|\partial_x^\alpha\phi\|_{L^2}(t)+\sum_{0\leq|\alpha_1|+|\alpha_2|\leq m-1}\|\partial\partial_x^{\alpha_1}\phi\partial\partial_x^{\alpha_2}\phi\|_{L^2}(t)\\ &+\text{cubic terms}+\ldots+\text{terms of order }k-1\Bigg)\end{aligned}$$

We look for instance at the quadratic term. We assume without loss of generality that $|\alpha_1|\leq\frac{m-1}{2}$. On the other hand, since $m\geq n+3$, we have $\frac{m-1}{2}+\frac{n}{2}\leq m-2$. Therefore,

$$\begin{aligned}&\sum_{0\leq|\alpha_1|+|\alpha_2|\leq m-1}\|\partial\partial_x^{\alpha_1}\phi\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)\\ \leq& C\left(\sum_{0\leq|\alpha_1|\leq\frac{m-1}{2}}\|\partial\partial_x^{\alpha_1}\phi\|_{L^\infty(\mathbb{R}^n)}(t)\right)\left(\sum_{0\leq|\alpha_2|\leq m-1}\|\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)\right)\\ \leq& C\|(\phi,\partial_t\phi)\|_{H^{m-1}(\mathbb{R}^n)\times H^{m-2}(\mathbb{R}^n)}(t)\|(\phi,\partial_t\phi)\|_{H^m(\mathbb{R}^n)\times H^{m-1}(\mathbb{R}^n)}(t).\end{aligned}$$

By the induction hypothesis, for every $T<T_*$, we have

$$\sup_{t\in[0,T]}\|\phi\|_{H^{m-1}(\mathbb{R}^n)}(t)\leq C$$

and this term is indeed linear in $\|\phi\|_{H^m(\mathbb{R}^n)}(t)$. The term with $a$ can be treated similarly. Therefore, for every fixed $T < T_*$, the following holds for every $t \in [0,T]$:

$$\begin{aligned}&\sup_{s\in[0,t)} \|(\phi,\partial_t\phi)\|_{H^m(\mathbb{R}^n)\times H^{m-1}(\mathbb{R}^n)}\\ \leq& C(T)\left(\|(\phi_0,\phi_1)\|_{H^m(\mathbb{R}^n)\times H^{m-1}(\mathbb{R}^n)} + \int_0^t \sup_{s\in[0,t)} \|(\phi,\partial_t\phi)\|_{H^m(\mathbb{R}^n)\times H^{m-1}(\mathbb{R}^n)}(s)ds\right).\end{aligned}$$

Gronwall's lemma implies the desired conclusion. Finally, the smoothness statement in (2) follows from the Sobolev embedding theorem (Theorem 5.5). □

*Remark* 6.7. From now on, we will use the terminology "maximal time of existence $T_*$" loosely. It will always mean the maximal time of existence such that a $L^\infty([0,T]; H^{n+2}(\mathbb{R}^n)) \times L^\infty([0,T]; H^{n+1}(\mathbb{R}^n))$ solution exists in $[0,T]\times\mathbb{R}^n$ for every $T < T_*$. Nevertheless, by Theorem 6.6, this coincides with the maximal time of existence of a solution with higher regularity.

In general, it is of course a very hard question to determine the maximal time of existence $T_*$ for a given initial data set to a given equation. The local theory, however, allows us to say that if $T_* < \infty$, then it must be the case that some norms blow up as the time $T_*$ is approached. We will give the precise statements in Theorem 6.10. Before that, we state a technical lemma that we will not prove:

**Lemma 6.8.** *(Schauder estimates) Let $s \geq 0$ and $F : \mathbb{R}^m \to \mathbb{R}$ be a smooth function with $F(0) = 0$. Then, if $\phi : \mathbb{R}^n \to \mathbb{R}^m$ belongs to the space $\phi \in L^\infty(\mathbb{R}^n;\mathbb{R}^m) \cap H^s(\mathbb{R}^n;\mathbb{R}^m)$, there exists a constant $C = C(F, \|\phi\|_{L^\infty(\mathbb{R}^n;\mathbb{R}^m)}, s, n, m) > 0$ such that*

$$\|F(\phi)\|_{H^s(\mathbb{R}^n)} \leq C\|\phi\|_{H^s(\mathbb{R}^n;\mathbb{R}^m)}.$$

**Lemma 6.9.** *For $s > 0$, there exists $C = C(n,s) > 0$ such that*

$$\|f\cdot(\partial_{x_i} g)\|_{H^s(\mathbb{R}^n)} \leq C(\|f\|_{H^{s+1}(\mathbb{R}^n)}\|g\|_{L^\infty(\mathbb{R}^n)} + \|g\|_{H^s(\mathbb{R}^n)}\|f\|_{W^{1,\infty}(\mathbb{R}^n)}).$$

We now state and prove our main result on the breakdown criteria, which can be viewed as a characterization of the maximal time of existence $T_*$:

**Theorem 6.10.** *(Breakdown criteria) Given initial data $(\phi_0,\phi_1) \in H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)$ to (6.1), let $T_*$ be the maximal time of existence. If $T_* < \infty$, then all of the following holds:*

*(1)*
$$\liminf_{t\to T_*} \|(\phi,\partial_t\phi)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}(t) \to \infty.$$

*(2)*
$$\limsup_{t\to T_*}\left(\sum_{|\alpha|\leq\lfloor\frac{n+2}{2}\rfloor} \|\partial_x^\alpha\phi\|_{L^\infty(\mathbb{R}^n)}(t) + \sum_{|\alpha|\leq\lfloor\frac{n}{2}\rfloor}\|\partial_t\partial_x^\alpha\phi\|_{L^\infty(\mathbb{R}^n)}(t)\right) \to \infty.$$

*(3)*
$$\limsup_{t\to T_*}\sum_{|\alpha|\leq 1}\|\partial^\alpha\phi\|_{L^\infty(\mathbb{R}^n)}(t) \to \infty.$$

*Remark* 6.11. Clearly, (3) implies (2) in Theorem 6.10. However, we state these separately since the stronger statement (3) requires the use of the Schauder estimates which we have not proved. In the rest of the notes, we will also avoid applying (3) for the same reason.

*Proof.* All three of the breakdown criteria will be proven with the following strategy: We assume that the conclusion fails, which gives use some estimates of the solution. We then apply the energy estimates to gain enough control of the solution to extend it beyond the time $T_*$, which then contradicts the maximality of $T_*$.

(1) We argue by contradiction. Suppose the conclusion fails, then there exists an increasing sequence $\{t_m\}_{m=1}^\infty \subset \mathbb{R}$ with $t_m \to T_*$ such that
$$\|(\phi,\partial_t\phi)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}(t_m) \leq C_*$$
for some $C_* > 0$ independent of $m$. Theorem 6.1 guarantees that we can find a time of existence $T$, depending only on $C_*$ (most importantly independent of $m$!) such that the equation can be solved

in $[t_m, t_m+T]\times\mathbb{R}^n$. By taking $m$ sufficiently large, we can therefore extend the solution beyond $T_*$, contradicting the maximality of $T_*$!

(2) We proceed by contradiction. Assume that $T_*<\infty$ but

$$\limsup_{t\to T_*}\left(\sum_{|\alpha|\leq\lfloor\frac{n+2}{2}\rfloor}\|\partial_x^\alpha\phi\|_{L^\infty(\mathbb{R}^n)}(t)+\sum_{|\alpha|\leq\lfloor\frac{n}{2}\rfloor}\|\partial_t\partial_x^\alpha\phi\|_{L^\infty(\mathbb{R}^n)}(t)\right)<\infty.$$

In particular, this implies that there exists a constant $D>0$ such that

$$\sup_{t\in[0,T_*)}\left(\sum_{|\alpha|\leq\lfloor\frac{n+2}{2}\rfloor}\|\partial_x^\alpha\phi\|_{L^\infty(\mathbb{R}^n)}(t)+\sum_{|\alpha|\leq\lfloor\frac{n}{2}\rfloor}\|\partial_t\partial_x^\alpha\phi\|_{L^\infty(\mathbb{R}^n)}(t)\right)\leq D.$$

Our goal is to show that this assumption implies that the $H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)$ norm of $(\phi,\partial_t\phi)$ remains bounded as $t\to T_*$. We will achieve this in two steps. First, we bound the $H^{n+1}(\mathbb{R}^n)\times H^n(\mathbb{R}^n)$ norm of $(\phi,\partial_t\phi)$ and then we control its $H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)$ norm.

To proceed, we now apply energy estimates to the nonlinear equation for $\phi$ to get

$$\begin{aligned}&\sup_{t\in[0,T)}\|(\phi,\partial_t\phi)\|_{H^k(\mathbb{R}^n)\times H^{k-1}(\mathbb{R}^n)}\\ \leq& C\left(\|(\phi_0,\phi_1)\|_{H^k(\mathbb{R}^n)\times H^{k-1}(\mathbb{R}^n)}+\int_0^T\left(\|F\|_{H^{k-1}(\mathbb{R}^n)}(t)+\sum_{|\gamma|+|\sigma|\leq k-1}\sum_{\alpha,\beta}\|\partial\partial_x^\gamma\left(a^{\alpha\beta}(\phi)\right)\partial\partial_x^\sigma\phi\|_{L^2}(t)\right)dt\right)\\ &\times\exp\left(C\int_0^T\|\partial a(\phi)\|_{L^\infty(\mathbb{R}^n)}(t)dt\right)\end{aligned}\tag{6.17}$$

for every $T\leq T_*$, where $C=C(T_*)$. We first consider (6.17) with $k=n+1$. Now, we bound $F$ and $\partial a$ as in the proof of part (1) of this theorem. More precisely, we have for some $C=C(D)$ such that

$$\begin{aligned}&\sum_{0\leq|\alpha|\leq k-1}\|\partial_x^\alpha F(\phi,\partial\phi)\|_{L^2}(t)\\ \leq& C\Bigg(\sum_{0\leq|\alpha|\leq k-1}\|\partial\partial_x^\alpha\phi\|_{L^2}(t)+\sum_{0\leq|\alpha_1|+|\alpha_2|\leq k-1}\|\partial\partial_x^{\alpha_1}\phi\partial\partial_x^{\alpha_2}\phi\|_{L^2}(t)\\ &+\text{cubic terms}+...+\text{terms of order } k-1)\end{aligned}\tag{6.18}$$

Look for instance at the quadratic term. The key point is that either $|\alpha_1|$ or $|\alpha_2|\leq\lfloor\frac{k-1}{2}\rfloor$. Without loss of generality, let's say $|\alpha_1|\leq\lfloor\frac{k-1}{2}\rfloor$. Since $k=n+1$, $\lfloor\frac{k-1}{2}\rfloor\leq\lfloor\frac{n}{2}\rfloor$, we have

$$\begin{aligned}&\sum_{|\alpha_1|+|\alpha_2|\leq k-1}\|\partial\partial_x^{\alpha_1}\phi\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)\\ \leq& C\left(\sum_{|\alpha_1|\leq\lfloor\frac{n}{2}\rfloor}\|\partial\partial_x^{\alpha_1}\phi\|_{L^\infty(\mathbb{R}^n)}(t)\right)\left(\sum_{|\alpha_2|\leq k-1}\|\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)\right)\\ \leq& CD\left(\sum_{|\alpha_2|\leq k-1}\|\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)\right).\end{aligned}$$

The higher order terms can be treated in an analogous manner to show that every term is at most linear in

$$\sum_{1\leq|\alpha_2|\leq k-1}\|\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t),$$

multiplied by some polynomials in $D$. A similar argument can also be used to treat the terms

$$\int_0^T\sum_{|\gamma|+|\sigma|\leq k-1}\sum_{\alpha,\beta}\|\partial\partial_x^\gamma\left(a^{\alpha\beta}(\phi)\right)\partial\partial_x^\sigma\phi\|_{L^2}(t)dt$$

and

$$\int_0^T \|\partial a(\phi)\|_{L^\infty(\mathbb{R}^n)}(t)dt.$$

Therefore, as a consequence, we have

$$\begin{aligned}&\sup_{t\in[0,T)} \|(\phi,\partial_t\phi)\|_{H^{n+1}(\mathbb{R}^n)\times H^n(\mathbb{R}^n)}\\ \leq& C(D,T_*)\left(\|(\phi_0,\phi_1)\|_{H^{n+1}(\mathbb{R}^n)\times H^n(\mathbb{R}^n)}+\int_0^T (1+D)^{n+1}\|(\phi,\partial_t\phi)\|_{H^{n+1}(\mathbb{R}^n)\times H^n(\mathbb{R}^n)}dt\right).\end{aligned}$$

By Gronwall's lemma, we get

$$\sup_{t\in[0,T_*)} \|(\phi,\partial_t\phi)\|_{H^{n+1}(\mathbb{R}^n)\times H^n(\mathbb{R}^n)} \leq C(D,T_*)\|(\phi_0,\phi_1)\|_{H^{n+1}(\mathbb{R}^n)\times H^n(\mathbb{R}^n)}. \tag{6.19}$$

As mentioned before, the next step is to control the $H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)$ norm of $(\phi,\partial_t\phi)$. We return to the energy estimates (6.17) with $k=n+2$. Again, we need to show that all terms are bounded at most linearly by $\|(\phi,\partial_t\phi)\|_{H^k(\mathbb{R}^n)\times H^{k-1}(\mathbb{R}^n)}$ up to some constant depending on $D$, $T_*$ and the initial data. Now consider the quadratic term in (6.18). The cases where $n$ is odd and $n$ is even will be treated separately. First, we consider $n$ odd. As before, we can assume that $|\alpha_1|\leq|\alpha_2|$. Since $k-1=n+1$, we have either $|\alpha_1|\leq\lfloor\frac{n}{2}\rfloor$ or we have $|\alpha_1|,|\alpha_2|\leq\lfloor\frac{n}{2}\rfloor+1$.

$$\begin{aligned}&\sum_{|\alpha_1|+|\alpha_2|\leq k-1}\|\partial\partial_x^{\alpha_1}\phi\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)\\ \leq& C\left(\sum_{|\alpha_1|\leq\lfloor\frac{n}{2}\rfloor}\|\partial\partial_x^{\alpha_1}\phi\|_{L^\infty(\mathbb{R}^n)}(t)\right)\left(\sum_{|\alpha_2|\leq k-1}\|\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)\right)\\ &+C\left(\sum_{|\alpha_1|\leq\lfloor\frac{n}{2}\rfloor+1}\|\partial\partial_x^{\alpha_1}\phi\|_{L^\infty(\mathbb{R}^n)}(t)\right)\left(\sum_{|\alpha_2|\leq\lfloor\frac{n}{2}\rfloor+1}\|\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)\right)\\ \leq& C(D,T_*,\|(\phi_0,\phi_1)\|_{H^{n+1}(\mathbb{R}^n)\times H^n(\mathbb{R}^n)})\left(\sum_{|\alpha|\leq k-1}\|\partial\partial_x^{\alpha}\phi\|_{L^2(\mathbb{R}^n)}(t)\right).\end{aligned}$$

Here, to deal with the last term, we have used on one hand that $\lfloor\frac{n}{2}\rfloor+2\leq n+1$ so we can control $\sum_{|\alpha_2|\leq\lfloor\frac{n}{2}\rfloor+1}\|\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)$ using (6.19); and on the other hand that $\lfloor\frac{n}{2}\rfloor+2+\frac{n}{2}<n+2$ (using the oddness of $n$) so that we can use Sobolev embedding to control $\sum_{|\alpha_1|\leq\lfloor\frac{n}{2}\rfloor+1}\|\partial\partial_x^{\alpha_1}\phi\|_{L^\infty(\mathbb{R}^n)}(t)$ by $\sum_{|\alpha|\leq k-1}\|\partial\partial_x^{\alpha}\phi\|_{L^2(\mathbb{R}^n)}(t)$. In the case where $n$ is even, if $|\alpha_1|\leq|\alpha_2|$, we must have $|\alpha_1|\leq\lfloor\frac{n}{2}\rfloor$. Therefore, we can estimate the quadratic term in (6.18) simply by

$$\begin{aligned}&\sum_{|\alpha_1|+|\alpha_2|\leq k-1}\|\partial\partial_x^{\alpha_1}\phi\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)\\ \leq& C\left(\sum_{|\alpha_1|\leq\lfloor\frac{n}{2}\rfloor}\|\partial\partial_x^{\alpha_1}\phi\|_{L^\infty(\mathbb{R}^n)}(t)\right)\left(\sum_{|\alpha_2|\leq k-1}\|\partial\partial_x^{\alpha_2}\phi\|_{L^2(\mathbb{R}^n)}(t)\right)\\ \leq& C(D,T_*)\left(\sum_{|\alpha|\leq k-1}\|\partial\partial_x^{\alpha}\phi\|_{L^2(\mathbb{R}^n)}(t)\right).\end{aligned}$$

One checks that we can similarly estimate higher order terms in (6.18) and also the term containing $a$ in (6.17). Therefore, as a consequence, we have

$$\begin{aligned}&\sup_{t\in[0,T)} \|(\phi,\partial_t\phi)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}\\ \leq& C(D,T_*,\|(\phi_0,\phi_1)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)})\left(1+C(D,T_*)\int_0^T\|(\phi,\partial_t\phi)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}dt\right).\end{aligned}$$

By Gronwall's lemma, we get

$$\sup_{t\in[0,T_*)} \|(\phi,\partial_t\phi)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)} \leq C(D,T_*,\|(\phi_0,\phi_1)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}).$$

We have thus obtain a uniform bound for

$$\|(\phi,\partial_t\phi)\|_{H^{n+2}(\mathbb{R}^n)\times H^{n+1}(\mathbb{R}^n)}(t)$$

when $t < T_*$. This contradits part (1) of this theorem.

(3) **Exercise.** (Hint: Using Lemmas 6.8 and 6.9.)

□

## 7. Global regularity for subcritical equations

We showed last time that there is a very general local theory for nonlinear wave equations. In general, of course there are many different phenomena associated to the long time behaviour. Nevertheless, there is a class of so-called subcritical equations, such that there is a coercive monotonic quantity above scaling (see discussions below), for which the problem of global regularity for general data is tractable. We look at one particular example here.

Consider the equation in $\mathbb{R}^{3+1}$ for $\phi : I\times\mathbb{R}^3\to\mathbb{R}$

$$\begin{cases}\Box\phi = |\phi|^2\phi\\ (\phi,\partial_t\phi)\restriction_{\{t=0\}} = (\phi_0,\phi_1),\end{cases} \tag{7.1}$$

where $I\subseteq\mathbb{R}$ is a time interval. We first show that

**Proposition 7.1.** *As long as the solution remains sufficiently regular, the following quantity is independent of time:*

$$E(t) := \int_{\mathbb{R}^3} \frac{1}{2}((\partial_t\phi)^2 + \sum_{i=1}^3(\partial_i\phi)^2)(t,x) + \frac{1}{4}|\phi|^4(t,x)dx.$$

*Proof.* This is a direct computation. □

Notice that the conservation of a positive quantity alone does not necessarily imply that the solution is globally regular. It is important to show that this quantity in fact gives sufficiently strong control over the solution. In the context of this equation, this is in fact manifested in the following Sobolev inequality:

**Proposition 7.2.** *There exists a constant $C > 0$ such that for every $\phi : \mathbb{R}^3\to\mathbb{R}$, we have*

$$\|\phi\|_{L^6(\mathbb{R}^3)} \leq C\|\phi\|_{\dot{H}^1(\mathbb{R}^3)}.$$

We are now ready to show that

**Theorem 7.3.** *Assume $(\phi_0,\phi_1)\in(H^5(\mathbb{R}^3)\times H^4(\mathbb{R}^3))\cap(C^\infty(\mathbb{R}^3)\times C^\infty(\mathbb{R}^3))$. Then (7.1) has a global-in-time smooth solution.*

*Proof.* We apply the local existence theorem to conclude that there is a smooth local-in-time solution. According to Theorem 6.1, it suffices to show that for every $T$, we have the bound

$$\sum_{|\alpha|\leq 2}\|\partial_x^\alpha\phi\|_{L^\infty([0,T]\times\mathbb{R}^3)} + \sum_{|\alpha|\leq 1}\|\partial_t\partial_x^\alpha\phi\|_{L^\infty([0,T]\times\mathbb{R}^3)} \leq C(T).$$

By Sobolev embedding (Theorem 5.5), it in turn suffices to prove

$$\|(\phi,\partial_t\phi)\|_{L^\infty([0,T];H^4(\mathbb{R}^3))\times L^\infty([0,T];H^3(\mathbb{R}^3))} \leq C(T).$$

We do this in two steps. First, we control the $H^2$ norm.

$$\begin{aligned}&\|(\phi,\partial_t\phi)\|_{L^\infty([0,T];H^2(\mathbb{R}^3))\times L^\infty([0,T];H^1(\mathbb{R}^3))}\\ \leq& C(T)\left(\|(\phi_0,\phi_1)\|_{H^2(\mathbb{R}^3)\times H^1(\mathbb{R}^3)}+\int_0^T\|\phi\cdot\phi\cdot\partial\phi\|_{L^2(\mathbb{R}^n)}(t)dt\right)\\ \leq& C(T)\left(\|(\phi_0,\phi_1)\|_{H^2(\mathbb{R}^3)\times H^1(\mathbb{R}^3)}+\int_0^T\|\phi\|_{L^6(\mathbb{R}^n)}(t)\|\phi\|_{L^6(\mathbb{R}^n)}(t)\|\partial\phi\|_{L^6(\mathbb{R}^n)}(t)dt\right)\\ \leq& C(T)\left(\|(\phi_0,\phi_1)\|_{H^2(\mathbb{R}^3)\times H^1(\mathbb{R}^3)}+\int_0^T(E(t))^2\|(\phi,\partial_t\phi)\|_{H^2(\mathbb{R}^3)\times H^1(\mathbb{R}^3)}(t)dt\right).\end{aligned}$$

Since $E$ is a priori controlled using the conservation law, we can apply Gronwall's lemma to obtain

$$\|(\phi,\partial_t\phi)\|_{L^\infty([0,T];H^2(\mathbb{R}^3))\times L^\infty([0,T];H^1(\mathbb{R}^3))}\leq C(T).$$

Here, of course we have abused notation such that this constant $C(T)$ is (exponentially) larger than the constants in the lines above. In particular, this implies that $\phi$ is bounded in $L^\infty([0,T]\times\mathbb{R}^3)$.

We now proceed to controlling the $H^4$ norm.

$$\begin{aligned}&\|(\phi,\partial_t\phi)\|_{L^\infty([0,T];H^4(\mathbb{R}^3))\times L^\infty([0,T];H^3(\mathbb{R}^3))}\\ \leq& C(T)\left(\|(\phi_0,\phi_1)\|_{H^4(\mathbb{R}^3)\times H^3(\mathbb{R}^3)}+\sum_{|\alpha|+|\beta|+|\gamma|=3}\int_0^T\|\partial_x^\alpha\phi\cdot\partial_x^\beta\phi\cdot\partial_x^\gamma\phi\|_{L^2(\mathbb{R}^n)}(t)dt\right)\\ \leq& C(T)\left(\|(\phi_0,\phi_1)\|_{H^2(\mathbb{R}^3)\times H^1(\mathbb{R}^3)}+\int_0^T\left(\sum_{|\alpha|=1}\|\partial_x^\alpha\phi\|_{L^6(\mathbb{R}^n)}(t)\right)^3dt\right.\\ &+\int_0^T\left(\|\phi\|_{L^\infty(\mathbb{R}^n)}(t)\right)^2\sum_{|\alpha|=3}\|\partial_x^\alpha\phi\|_{L^2(\mathbb{R}^n)}(t)dt\\ &\left.+\int_0^T\|\phi\|_{L^6(\mathbb{R}^n)}(t)\left(\sum_{|\alpha|=1}\|\partial_x^\alpha\phi\|_{L^6(\mathbb{R}^n)}(t)\right)\left(\sum_{|\beta|=2}\|\partial_x^\beta\phi\|_{L^6(\mathbb{R}^n)}(t)\right)dt\right).\end{aligned}$$

Now, the key points are that

(1) $\phi$ is in $L^6$ by the conservation law
(2) $\phi$ in $L^\infty$ and $\partial_x\phi$ in $L^6$ using the $H^2$ control that we have obtained above.

Therefore, the above estimate is *linear* in

$$\|(\phi,\partial_t\phi)\|_{H^4(\mathbb{R}^3)\times H^3(\mathbb{R}^3)}(t),$$

i.e.,

$$\begin{aligned}&\|(\phi,\partial_t\phi)\|_{L^\infty([0,T];H^4(\mathbb{R}^3))\times L^\infty([0,T];H^3(\mathbb{R}^3))}\\ \leq& C(E,T)\left(\|(\phi_0,\phi_1)\|_{H^4(\mathbb{R}^3)\times H^3(\mathbb{R}^3)}+\int_0^T\|(\phi,\partial_t\phi)\|_{H^4(\mathbb{R}^3)\times H^3(\mathbb{R}^3)}(t)dt\right).\end{aligned}$$

We can therefore conclude using Gronwall's lemma. □

*Remark* 7.4. Another perhaps more direct way to prove this theorem is to reprove the local existence theorem and show that if we perform Picard's iteration in $L^\infty([0,T];H^1(\mathbb{R}^3))\times L^\infty([0,T];L^2(\mathbb{R}^3))$, we can show that there exists a solution in $[0,T]\times\mathbb{R}^3$, where $T$ **depends only on** $\|\phi_0\|_{\dot{H}^1(\mathbb{R}^3)}$ **and** $\|\phi_1\|_{L^2(\mathbb{R}^3)}$. Moreover, we can also show as in the local existence theorem that as long as an $H^1$ solution exists, then higher regularity is propagated. This allows us to conclude using the conservation law.

We now briefly discuss the concepts of scaling and criticality. Notice that the above PDE has two features. Firstly, it has a conservation law that controls the $\dot{H}^1$ norm. Secondly, it is invariant with respect to the scaling

$$\phi_\lambda(t,x)=\lambda\phi(\lambda t,\lambda x),$$

i.e., if $\phi(t,x)$ is a solution, then $\phi_\lambda(t,x)$ is also a solution. Now, here is the key point: if we scale to smaller length scales, i.e., we take $\lambda \to \infty$, then the $\dot{H}^1$ norm $\to \infty$. This shows that the conservation law (or more generally a monotonic quantity) is all the more useful at smaller length scales. We call this property **subcritical**. On the other hand, we say that an equation is **supercritical** if the opposite is true and is **critical** if the conservation law scales in exactly the same way as the scaling of the equation.

The concept of criticality is extremely important to predict whether an equation has global regular large data solutions. Most of the time, subcritical equations have a global regular solution. Critical equations, lying at the threshold of subcritical and supercritical equations, exhibit a large array of phenomenon - some of them always have globally regular solutions while some possess solutions that exhibit finite time blow up. On the other hand, very little is known at all about supercritical equations!

Returning to the equation that we were considering, if we look at a more general class of equations

$$\Box\phi = |\phi|^{p-1}\phi$$

in $\mathbb{R}\times\mathbb{R}^3$, we see that the $\dot{H}^1$ norm is always controlled by the conservation law while the equation is invariant under the scaling

$$\phi_\lambda(t,x) = \lambda^{\frac{2}{p-1}}\phi(\lambda t, \lambda x).$$

The equation is therefore (**Exercise**) subcritical if $p < 5$; critical if $p = 5$ and supercritical if $p > 5$. Indeed in the first two cases, it is known that regular initial data give rise to globally regular solutions. (We see moreover that the full range of $1 \leq p \leq 3$ can be treated as above[10]. However, the case $3 < p \leq 5$ requires techniques that we will not have time to cover in the lectures. See the Example Sheet for further discussions.) On the other hand, almost nothing is known regarding large data solutions in the supercritical case!

We end this section with a discussion on the wave map equation from $\mathbb{R}\times\mathbb{R}^n$ to $\mathbb{S}^m \subset \mathbb{R}^{m+1}$. Recall that it is given by

$$\Box\phi = \phi(\partial_t\phi^{\mathbf{T}}\partial_t\phi - \sum_{i=1}^n \partial_i\phi^{\mathbf{T}}\partial_i\phi), \tag{7.2}$$

with initial data $(\phi, \partial_t\phi)\restriction_{\{t=0\}} = (\phi_0, \phi_1)$ such that $|\phi_0|^2 = 1$ and $\phi_1^{\mathbf{T}}\phi_0 = 0$. Let us recall the convention that $\mathbf{T}$ is to be understood as the transpose of a vector in $\mathbb{R}^{m+1}$. For all $n \geq 1$ and $m \geq 1$, we have the following conservation law:

**Proposition 7.5.** *As long as the solution to* (7.2) *remains sufficiently regular, the following quantity*

$$E(t) = \frac{1}{2}\int_{\mathbb{R}^n}(|\partial_t\phi|^2 + \sum_{i=1}^n |\partial_i\phi|^2)(t,x)dx$$

*is independent of $t$. Here, we have used the shorthand $|\partial_t\phi|^2 = \partial_t\phi^{\mathbf{T}}\partial_t\phi$ and similarly for $|\partial_i\phi|^2$.*

*Proof.* Using the fact $|\phi|^2 = 1$, we have $\phi^{\mathbf{T}}\partial_\alpha\phi$ for $\alpha = 0, 1, ..., n$. Therefore, using the wave map equations, we have

$$-\frac{1}{2}\partial_t(\partial_t\phi^{\mathbf{T}}\partial_t\phi + \sum_{i=1}^n \partial_i\phi^{\mathbf{T}}\partial_i\phi) + \sum_{i=1}^n \partial_i(\partial_t\phi^{\mathbf{T}}\partial_i\phi) = 0. \tag{7.3}$$

Integrating by parts in the region between any two times, we obtain the desired conclusion. □

On the other hand, the wave map equation is invariant under the scaling

$$\phi_\lambda(t,x) = \phi(\lambda t, \lambda x).$$

Notice that $\phi_\lambda$ is indeed a map to the sphere for all $\lambda$. In particular, the $n = 1$ case is subcritical; the $n = 2$ case is critical and the $n \geq 3$ case is supercritical. In dimensions $n \geq 2$, there exist solutions which arise from smooth initial data but blow up in finite time[11]. On the other hand, global regularity holds in the $n = 1$ case.

---

[10] Note that if $p \in (1,3)$, the nonlinearity is not smooth. So in principle we cannot apply the local theory that we had. On the other hand, it is not difficult to see that one can construct $L^\infty([0,T]; H^1(\mathbb{R}^3)) \times L^\infty([0,T]; L^2(\mathbb{R}^3))$ directly for these nonlinearities.

[11] Notice while in both the critical and supercritical cases, the existence of blow-up solutions has been exhibited, in general the dynamics in much better understood in the $n = 2$ case.

**Proposition 7.6.** $(H^2(\mathbb{R})\times H^1(\mathbb{R}))\cap(C^\infty(\mathbb{R})\times C^\infty(\mathbb{R}))$ *initial data for the* $1+1$*-dimensional wave map equation* $\mathbb{R}^{1+1}\to\mathbb{S}^2$ *give rise to global-in-time smooth solution. (Here,* $H^2(\mathbb{R})\times H^1(\mathbb{R})$ *is understood so that the map approaches a constant map as* $|x|\to\infty$*, i.e., there exists* $y\in\mathbb{S}^2$ *such that each component of* $(\phi_0-y,\phi_1)$ *is in* $H^2(\mathbb{R})\times H^1(\mathbb{R})$*.)*

*Proof.* We first make a remark regarding the applicability of Theorem 6.1. Notice that unlike in the setting in Theorem 6.1, the wave map equations are in fact a *system* of wave equations as opposed to a single scalar wave equation. Nevertheless, it is easy to see that the proof of Theorem 6.1 using energy estimates can also be applied in this case after trivial modifications.

Therefore, using this local existence theorem (Theorem 6.1), it suffices to show that

$$\sum_{|\alpha|\leq 1}\|\partial^\alpha\phi\|_{L^\infty(\mathbb{R})}\leq C.$$

When $|\alpha|=0$, this is already implied by the fact that this is a map to the sphere, i.e., $|\phi|^2=1$. It turns out that in $1+1$ dimensions, the wave map equation admits many more conservation laws - indeed we have infinitely many of them! Notice that

$$\Box(\partial_t\phi^{\mathbf{T}}\partial_t\phi+\partial_x\phi^{\mathbf{T}}\partial_x\phi)=0.$$

This is because, using the fact that $\phi^{\mathbf{T}}\partial\phi=0$, we have

$$\partial_t^2\left(\partial_t\phi^{\mathbf{T}}\partial_t\phi+\partial_x\phi^{\mathbf{T}}\partial_x\phi\right)=2\partial_t\partial_x(\partial_t\phi^{\mathbf{T}}\partial_x\phi)=\partial_x^2\left(\partial_t\phi^{\mathbf{T}}\partial_t\phi+\partial_x\phi^{\mathbf{T}}\partial_x\phi\right).$$

In particular, this shows that

$$\sum_{|\alpha|\leq 1}\|\partial^\alpha\phi\|_{L^\infty(\mathbb{R})}\leq C.$$

□

*Remark* 7.7. Using

$$\Box(\partial_t\phi^{\mathbf{T}}\partial_t\phi+\partial_x\phi^{\mathbf{T}}\partial_x\phi)=0,$$

we conclude that for $v=t+x$ and $u=t-x$, $\partial_v(\partial_t\phi^{\mathbf{T}}\partial_t\phi+\partial_x\phi^{\mathbf{T}}\partial_x\phi)$ is independent of $u$ and $\partial_u(\partial_t\phi^{\mathbf{T}}\partial_t\phi+\partial_x\phi^{\mathbf{T}}\partial_x\phi)$ is independent of $v$! Hence we have uncountably many conservation laws! Equations with infinitely many conserved quantities can typically called **completely integrable**. Many techniques have been developed to give very precise information about the solutions to completely integrable equations. However, of course these techniques are very special and can only be applied to the very restricted class of completely integrable equations.

We now give an alternative proof of the global regularity of the $(1+1)$-dimensional wave map which uses the **characteristic energy**, but not the additional conservation laws.

*Proof.* Let $v=t+x$ and $u=t-x$. We first prove the boundedness of the characteristic energy. We start with (7.3) and integrate by parts in the region

$$\{0\leq t\leq T\}\cap\{v\leq v_0\}.$$

This shows that as long as the solution remains regular in $t\leq T$, the following quantity is a priori bounded:

$$\int_{-v_0}^{2T-v_0}|\partial_u\phi|^2(u,v_0)du\leq C\|(\phi_0,\phi_1)\|_{H^1(\mathbb{R})\times L^2(\mathbb{R})}.$$

Similarly, integrating by parts in the region

$$\{0\leq t\leq T\}\cap\{u\leq u_0\},$$

we get

$$\int_{-u_0}^{2T-u_0}|\partial_v\phi|^2(u_0,v)dv\leq C\|(\phi_0,\phi_1)\|_{H^1(\mathbb{R})\times L^2(\mathbb{R})}.$$

Since $u_0$, $v_0$ are arbitrary, we have

$$\sup_v\int_{-v}^{2T-v}|\partial_u\phi|^2(u,v)du+\sup_u\int_{-u}^{2T-u}|\partial_v\phi|^2(u,v)dv\leq C\|(\phi_0,\phi_1)\|_{H^1(\mathbb{R})\times L^2(\mathbb{R})}. \tag{7.4}$$

We now use this and the equation to control the derivatives of $\phi$ in $L^\infty$. More precisely, we write the equation as

$$-4\partial_u\partial_v\phi = \phi(\partial_t\phi^{\mathbf{T}}\partial_t\phi - \sum_{i=1}^n \partial_i\phi^{\mathbf{T}}\partial_i\phi) = 4\phi\partial_u\phi^{\mathbf{T}}\partial_v\phi.$$

We can integrate in the $u$ direction to control $\partial_v\phi$, namely, for $u+v\leq 2T$, we have

$$\sup_{v\in\mathbb{R}}|\partial_v\phi|(u,v)\leq C\|(\phi_0,\phi_1)\|_{W^{1,\infty}(\mathbb{R})\times L^\infty(\mathbb{R})} + \sup_v \int_{-v}^u |\partial_u\phi||\partial_v\phi|(u',v)du'.$$

Gronwall's inequality implies that

$$\begin{aligned}\sup_{u,v\in\mathbb{R}}|\partial_v\phi|(u,v) \leq& C\|(\phi_0,\phi_1)\|_{W^{1,\infty}(\mathbb{R})\times L^\infty(\mathbb{R})}\exp(C\sup_v\int_{-v}^u|\partial_u\phi|(u',v)du')\\ \leq& C\|(\phi_0,\phi_1)\|_{W^{1,\infty}(\mathbb{R})\times L^\infty(\mathbb{R})}\exp(CT^{\frac12}\|(\phi_0,\phi_1)\|_{H^1(\mathbb{R})\times L^2(\mathbb{R})}),\end{aligned}$$

where in the last step we have used (7.4) together with the Cauchy-Schwarz inequality. In a similar manner, we can also use (7.4) to obtain

$$\sup_{u,v\in\mathbb{R}}|\partial_u\phi|(u,v) \leq C\|(\phi_0,\phi_1)\|_{W^{1,\infty}(\mathbb{R})\times L^\infty(\mathbb{R})}\exp(CT^{\frac12}\|(\phi_0,\phi_1)\|_{H^1(\mathbb{R})\times L^2(\mathbb{R})}).$$

Since we have controlled the first derivatives of $\phi$ on any time interval $[0,T]$, we apply Theorem 6.1 to conclude the proof. □

## 8. Einstein vacuum equations

We now return to the local theory for nonlinear wave equations and show that it gives a local theory for the Einstein vacuum equations

$$Ric(g) = 0.$$

As we mentioned in the introduction, the Einstein vacuum equations a priori do not look like a system of wave equations. We first recall some basic notions in differential geometry and derive the formula. (Note that repeated indices are always summed over!)

**Definition 8.1.** Given a metric $g$, define the Levi-Civita connection $\nabla$ by

$$\nabla_{X^\alpha\frac{\partial}{\partial x^\alpha}}Y^\beta\frac{\partial}{\partial x^\beta} = X^\alpha\frac{\partial Y^\beta}{\partial x^\alpha}\frac{\partial}{\partial x^\beta} + \Gamma^\mu_{\alpha\beta}X^\alpha Y^\beta\frac{\partial}{\partial x^\mu},$$

where $\Gamma^\mu_{\alpha\beta}$ is the Christoffel symbols given by

$$\Gamma^\mu_{\alpha\beta} = \frac12(g^{-1})^{\mu\nu}(\frac{\partial}{\partial x^\beta}g_{\alpha\nu} + \frac{\partial}{\partial x^\alpha}g_{\beta\nu} - \frac{\partial}{\partial x^\nu}g_{\alpha\beta}).$$

**Definition 8.2.** Define the Riemann curvature tensor $R$ by

$$R^\gamma{}_{\alpha\mu\nu}\frac{\partial}{\partial x^\gamma} = \nabla_{\frac{\partial}{\partial x^\mu}}\nabla_{\frac{\partial}{\partial x^\nu}}\frac{\partial}{\partial x^\alpha} - \nabla_{\frac{\partial}{\partial x^\nu}}\nabla_{\frac{\partial}{\partial x^\mu}}\frac{\partial}{\partial x^\alpha}$$

and

$$R_{\beta\alpha\mu\nu} = g_{\beta\gamma}R^\gamma{}_{\alpha\mu\nu}.$$

We recall (without proof) some standard properties of the Riemann curvature tensor:

**Proposition 8.3.** *The Riemann curvature tensor satisfies the following properties:*

*(1)*

$$R_{\mu\nu\alpha\beta} = R_{\alpha\beta\mu\nu} = -R_{\nu\mu\alpha\beta} = -R_{\mu\nu\beta\alpha};$$

*(2)*

$$R_{\mu\nu\alpha\beta} + R_{\alpha\mu\nu\beta} + R_{\nu\alpha\mu\beta} = 0;$$

*(3)*

$$\nabla_\sigma R_{\mu\nu\alpha\beta} + \nabla_\mu R_{\nu\sigma\alpha\beta} + \nabla_\nu R_{\sigma\mu\alpha\beta} = 0.$$

**Definition 8.4.** Define the Ricci curvature tensor by

$$Ric_{\mu\nu} = (g^{-1})^{\alpha\beta}R_{\mu\alpha\nu\beta}.$$

We compute

$$\nabla_{\frac{\partial}{\partial x^\mu}}\nabla_{\frac{\partial}{\partial x^\nu}}\frac{\partial}{\partial x^\alpha}=\nabla_{\frac{\partial}{\partial x^\mu}}(\Gamma^\rho_{\nu\alpha}\frac{\partial}{\partial x^\rho})=((\frac{\partial}{\partial x^\mu}\Gamma^\rho_{\nu\alpha})+\Gamma^\sigma_{\nu\alpha}\Gamma^\rho_{\mu\sigma})\frac{\partial}{\partial x^\rho}.$$

We will only consider terms that has two derivatives of the metric.

$$R_{\mu\nu\alpha}{}^\gamma=-\frac{\partial}{\partial x^\mu}\Gamma^\gamma_{\nu\alpha}+\frac{\partial}{\partial x^\nu}\Gamma^\gamma_{\mu\alpha}+...$$

and thus the Ricci curvature tensor is given by

$$Ric_{\mu\nu}=-\frac{\partial}{\partial x^\mu}\Gamma^\gamma_{\gamma\nu}+\frac{\partial}{\partial x^\gamma}\Gamma^\gamma_{\mu\nu}+...$$

We can show that in a local coordinate system, the Einstein vacuum equations read

$$0=Ric(g)_{\mu\nu}=-\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\alpha\beta}g_{\mu\nu}-\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\mu\nu}g_{\alpha\beta}+\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\alpha\nu}g_{\beta\mu}+\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\beta\mu}g_{\alpha\nu}+F_{\mu\nu}(g,\partial g), \tag{8.1}$$

where $F_{\mu\nu}(g,\partial g)$ is a function of $g$ and its derivatives. In particular, *if the second, third and fourth terms are absent*, then this is a wave equation and the general local theory for quasilinear equations can be applied. On the other hand, for a particular choice of coordinate system which satisfies the so-called **wave coordinate condition**

$$\Box_g x^\alpha=\frac{1}{\sqrt{-g}}\partial_\mu((g^{-1})^{\mu\nu}\sqrt{-g}\partial_\nu x^\alpha)=0 \tag{8.2}$$

then indeed the second, third and fourth terms can be re-written as terms having at most one derivatives of $g$. Here, we have used $g$ to denote the determinant of $g$.

To see this, we use the matrix identities

$$\partial_\alpha(g^{-1})^{\mu\nu}=-(g^{-1})^{\mu\beta}(g^{-1})^{\sigma\nu}\partial_\alpha g_{\sigma\beta}$$

and

$$\partial_\alpha(\log|\det g|)=(g^{-1})^{\mu\nu}\partial_\alpha g_{\mu\nu}.$$

The wave coordinate condition (8.2) implies

$$0=\partial_\mu((g^{-1})^{\mu\nu}\sqrt{-g})=-\sqrt{-g}\left((g^{-1})^{\mu\alpha}(g^{-1})^{\nu\beta}\partial_\mu g_{\alpha\beta}-\frac{1}{2}(g^{-1})^{\mu\nu}(g^{-1})^{\alpha\beta}\partial_\mu g_{\alpha\beta}\right),$$

which, after contracting with $g_{\nu\sigma}$, in turn implies

$$(g^{-1})^{\mu\alpha}\partial_\mu g_{\alpha\sigma}=\frac{1}{2}(g^{-1})^{\alpha\beta}\partial_\sigma g_{\alpha\beta}.$$

Now define $\lambda_\sigma$ to be

$$\lambda_\sigma=(g^{-1})^{\mu\alpha}\partial_\mu g_{\alpha\sigma}-\frac{1}{2}(g^{-1})^{\alpha\beta}\partial_\sigma g_{\alpha\beta}$$

so that $\lambda_\sigma$ vanishes if (8.2) holds. Using this, we derive

$$\begin{aligned}&-\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\mu\nu}g_{\alpha\beta}\\
=&-\frac{1}{2}\partial_\mu((g^{-1})^{\alpha\beta}\partial_\nu g_{\alpha\beta})+\frac{1}{2}(\partial_\mu(g^{-1})^{\alpha\beta})\partial_\nu g_{\alpha\beta}\\
=&-\partial_\mu((g^{-1})^{\alpha\beta}\partial_\beta g_{\alpha\nu})+\partial_\mu\lambda_\sigma+\frac{1}{2}(\partial_\mu(g^{-1})^{\alpha\beta})\partial_\nu g_{\alpha\beta}\\
=&-(g^{-1})^{\alpha\beta}\partial^2_{\beta\mu}g_{\alpha\nu}-(\partial_\mu(g^{-1})^{\alpha\beta})\partial_\beta g_{\alpha\nu}+\frac{1}{2}(\partial_\mu(g^{-1})^{\alpha\beta})\partial_\nu g_{\alpha\beta}+\partial_\mu\lambda_\sigma\\
=&-(g^{-1})^{\alpha\beta}\partial^2_{\beta\mu}g_{\alpha\nu}+(g^{-1})^{\alpha\sigma}(g^{-1})^{\beta\rho}\partial_\mu g_{\sigma\rho}\partial_\beta g_{\alpha\nu}-\frac{1}{2}(g^{-1})^{\alpha\sigma}(g^{-1})^{\beta\rho}\partial_\mu g_{\sigma\rho}\partial_\nu g_{\alpha\beta}+\partial_\mu\lambda_\sigma.\end{aligned}$$

In a completely analogous manner, we also have

$$\begin{aligned}&-\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\mu\nu}g_{\alpha\beta}\\
=&-(g^{-1})^{\alpha\beta}\partial^2_{\alpha\nu}g_{\beta\mu}+(g^{-1})^{\alpha\sigma}(g^{-1})^{\beta\rho}\partial_\nu g_{\sigma\rho}\partial_\alpha g_{\beta\mu}-\frac{1}{2}(g^{-1})^{\alpha\sigma}(g^{-1})^{\beta\rho}\partial_\mu g_{\sigma\rho}\partial_\nu g_{\alpha\beta}+\partial_\sigma\lambda_\mu.\end{aligned}$$

Using the previous calculations, we have thus shown the following result:

**Proposition 8.5.** *Define the reduced Ricci curvature to be*

$$\begin{aligned}\widetilde{Ric}(g)_{\mu\nu} :=& -\frac{1}{2}(g^{-1})^{\alpha\beta}\partial^2_{\alpha\beta}g_{\mu\nu} + \frac{1}{2}(g^{-1})^{\alpha\sigma}(g^{-1})^{\beta\rho}\partial_\mu g_{\sigma\rho}\partial_\beta g_{\alpha\nu} + \frac{1}{2}(g^{-1})^{\alpha\sigma}(g^{-1})^{\beta\rho}\partial_\nu g_{\sigma\rho}\partial_\alpha g_{\beta\mu}\\ &-\frac{1}{2}(g^{-1})^{\alpha\sigma}(g^{-1})^{\beta\rho}\partial_\mu g_{\sigma\rho}\partial_\nu g_{\alpha\beta} + F_{\mu\nu}(g,\partial g),\end{aligned}$$

*where $F$ is as in* (8.1). *Then*

$$\widetilde{Ric}(g)_{\mu\nu} = Ric(g)_{\mu\nu} - \frac{1}{2}\partial_\mu\lambda_\nu - \frac{1}{2}\partial_\nu\lambda_\mu.$$

In other words, if the wave coordinate condition holds, the Einstein vacuum equations become the reduced Einstein vacuum equations, which is a system of nonlinear wave equations. However, we of course need to guarantee that the wave coordinate condition holds! The strategy, introduced by Choquet-Bruhat, is to construct local solutions to the reduced Einstein vacuum equations and show that in fact if the wave coordinate condition (8.2) holds initially, then it is propagated by the flow. One of the key observations is that $\lambda_\sigma$ in fact satisfies a wave equation.

**Proposition 8.6.** *Given a Lorentzian metric $g$ such that the reduced Einstein vacuum equations are satisfied, i.e., $\widetilde{Ric}(g) = 0$. Then $\lambda$ satisfies a wave equation:*

$$\frac{1}{2}(g^{-1})^{\sigma\mu}\partial^2_{\sigma\mu}\lambda_\nu + c_\nu^{\alpha\beta}\partial_\alpha\lambda_\beta = 0,$$

*where $c_\nu^{\alpha\beta}$ are smooth functions of $g$ and its derivatives.*

*Proof.* The proof is based on the following identity:

$$\nabla^\mu(Ric_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R) = 0. \tag{8.3}$$

To see that this identity holds, first we use the second Bianchi identity (part 3 in Proposition 8.3) to get that

$$\begin{aligned}\nabla^\mu R_{\mu\nu\alpha\beta} = (g^{-1})^{\mu\sigma}\nabla_\sigma R_{\mu\nu\alpha\beta} =& -(g^{-1})^{\mu\sigma}\nabla_\alpha R_{\mu\nu\beta\sigma} - (g^{-1})^{\mu\sigma}\nabla_\beta R_{\mu\nu\sigma\alpha}\\ =& \nabla_\alpha Ric_{\beta\nu} - \nabla_\beta Ric_{\alpha\nu}.\end{aligned}$$

Taking the trace of this identity in $\alpha$, $\nu$, we get (8.3).

We now use this to derive an equation for $\lambda$. First, the vanishing of $\widetilde{Ric}(g)$ implies that

$$0 = Ric(g)_{\mu\nu} - \frac{1}{2}\partial_\mu\lambda_\nu - \frac{1}{2}\partial_\nu\lambda_\mu$$

and in particular after taking the trace, we have

$$0 = R - (g^{-1})^{\mu\nu}\partial_\mu\lambda_\nu.$$

Therefore, by (8.3), we have

$$\begin{aligned}0 =& \nabla^\mu(Ric(g)_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R)\\ =& -(g^{-1})^{\sigma\mu}\partial_\sigma(\frac{1}{2}\partial_\mu\lambda_\nu + \frac{1}{2}\partial_\nu\lambda_\mu - \frac{1}{2}g_{\mu\nu}(g^{-1})^{\alpha\beta}\partial_\alpha\lambda_\beta)\\ &-(g^{-1})^{\sigma\mu}\Gamma^\delta_{\sigma\mu}(\frac{1}{2}\partial_\delta\lambda_\nu + \frac{1}{2}\partial_\nu\lambda_\delta - \frac{1}{2}g_{\delta\nu}(g^{-1})^{\alpha\beta}\partial_\alpha\lambda_\beta)\\ &-(g^{-1})^{\sigma\mu}\Gamma^\delta_{\sigma\nu}(\frac{1}{2}\partial_\mu\lambda_\delta + \frac{1}{2}\partial_\delta\lambda_\mu - \frac{1}{2}g_{\mu\delta}(g^{-1})^{\alpha\beta}\partial_\alpha\lambda_\beta)\\ =& -\frac{1}{2}(g^{-1})^{\sigma\mu}\partial^2_{\sigma\mu}\lambda_\nu - c_\nu^{\alpha\beta}\partial_\alpha\lambda_\beta,\end{aligned}$$

for some $c_\nu^{\alpha\beta}$ which are smooth functions of $g$ and its derivatives.

□

Before we proceed to the main theorem, we first discuss the initial data that have to be posed. Given that we are to solve the reduced Einstein vacuum equations, which are nonlinear wave equations for the metric $g$, we need to prescribe the initial $g$ and $\partial_t g$. It turns out that we only need to prescribe the metric intrinsic to the initial hypersurface and the second fundamental form. The second fundamental form is defined as

$$\hat{k}_{ij} := \frac{1}{2}(\mathcal{L}_n g)_{ij} := \frac{1}{2}n^\ell \partial_\ell g_{ij} + \frac{1}{2}g_{j\ell}\partial_i n^\ell + \frac{1}{2}g_{i\ell}\partial_j n^\ell,$$

where $n$ is the normal to the initial hypersurface such that $g(n,n) = -1$. The remaining components for $g$ and $\partial_t g$ can be prescribed as coordinate conditions. More precisely, we choose $\partial_t$ to be of unit length $-1$ and orthogonal to the initial hypersurface, i.e. $g_{00}\restriction_{\{t=0\}} = -1$, $g_{0i}\restriction_{\{t=0\}} = 0$. Under this choice, the second fundamental form $k_{ij}$ becomes

$$\hat{k}_{ij} = \frac{1}{2}\partial_t g_{ij}.$$

Now, the remaining choices of the initial conditions (i.e., $\partial_t g_{00}$ and $\partial_t g_{0i}$) are fixed by the wave coordinate condition. More precisely,

$$0 = \lambda_i = (g^{-1})^{\mu\alpha}\partial_\mu g_{\alpha i} - \frac{1}{2}(g^{-1})^{\alpha\beta}\partial_i g_{\alpha\beta}$$

fixes $\partial_t g_{0i}$ and

$$0 = \lambda_0 = (g^{-1})^{\mu\alpha}\partial_\mu g_{\alpha 0} - \frac{1}{2}(g^{-1})^{\alpha\beta}\partial_0 g_{\alpha\beta}$$

fixes $\partial_t g_{00}$.

Before we proceed, we make a final observation. We compute $(Ric - \frac{1}{2}gR)_{00}$ and $Ric_{0i}$ on the initial hypersurface and note that there are no terms of the form $\partial_t^2 g$! To see that, notice that

$$Ric_{00} = \frac{1}{2}(g^{-1})^{00}\partial_t^2 g_{00} - \frac{1}{2}(g^{-1})^{\alpha\beta}\partial_t^2 g_{\alpha\beta} + \ldots$$

On the other hand,

$$(g^{-1})^{\alpha\beta}Ric_{\alpha\beta} = -(g^{-1})^{00}(g^{-1})^{\alpha\beta}\partial_t^2 g_{\alpha\beta} + (g^{-1})^{00}(g^{-1})^{00}\partial_t^2 g_{00} + \ldots$$

Hence $(Ric - \frac{1}{2}gR)_{00}$ does not depend on the second time derivatives of $g$. Similarly for $Ric_{0i}$.

In other words, if $(Ric - \frac{1}{2}gR)_{00}$ and $Ric_{0i}$ are to vanish, they must be conditions that have to be *imposed* initially. They are called the constraint equations and can be given geometrically as equations on the initial hypersurface:

$$\hat{\nabla}_i \hat{k}^i{}_j - \hat{\nabla}_j \hat{k}^i{}_i = 0,$$
$$\hat{R}(\hat{g}) + (\hat{k}^i{}_i)^2 - \hat{k}^i{}_j \hat{k}^j{}_i = 0.$$

Combining the above results, we have thus proved the local existence theorem of Choquet-Bruhat:

**Theorem 8.7.** *Given initial data $(\hat{g}_{ij}, \hat{k}_{ij})$ on $\mathbb{R}^n$ satisfying the constraint equations and such that $\sum_{i,j}|\hat{g}_{ij} - \delta_{ij}| \leq \frac{1}{20}$, there exists a metric $g$ in $I \times \mathbb{R}^n$ which solves the Einstein vacuum equations such that the induced metric and the induced second fundamental form on $\{0\} \times \mathbb{R}^n$ coincide with $\hat{g}_{ij}$ and $\hat{k}_{ij}$ respectively.*

*Proof.* Given initial data above, we solve the reduced Einstein vacuum equations, which are a system of nonlinear wave equations. By the local existence theory, there exists a local solution to the reduced Einstein vacuum equations.

To show that this solution is indeed a solution to the Einstein vacuum equations, we need to show that $\lambda_\sigma = 0$. In view of the fact that $\lambda_\sigma$ satisfies a wave equation, it suffices to show that $(\lambda_\sigma, \partial_t \lambda_\sigma)\restriction_{\{t=0\}} = (0,0)$ for $\sigma = 0, 1, 2, 3$.

Of course, we have already set $\lambda_\sigma(t=0) = 0$. To show that $\partial_t \lambda_\sigma(t=0) = 0$, we apply Proposition 8.5. The vanishing of $\widetilde{Ric}(g)_{\mu\nu}$ gives

$$0 = Ric(g)_{\mu\nu} - \frac{1}{2}\partial_\mu \lambda_\nu - \frac{1}{2}\partial_\nu \lambda_\mu,$$

which also implies

$$R = (g^{-1})^{\mu\nu}\partial_\mu \lambda_\nu.$$

Since the initial data $(\hat{g}, \hat{k})$ satisfy the constraints $Ric(g)_{i0}\restriction_{\{t=0\}} = 0$ and $(Ric(g)_{00} - \frac{1}{2}Rg_{00})\restriction_{\{t=0\}} = 0$, we have

$$\frac{1}{2}\partial_t \lambda_i \restriction_{\{t=0\}} = -\frac{1}{2}\partial_i \lambda_0 \restriction_{\{t=0\}} = 0$$

and
$$(\partial_t\lambda_0 + \frac{1}{2}(g^{-1})^{\mu\nu}\partial_\mu\lambda_\nu) \restriction_{\{t=0\}} = 0.$$
The latter condition can be rewritte to give
$$\frac{1}{2}\partial_t\lambda_0 \restriction_{\{t=0\}} = -\frac{1}{2}(g^{-1})^{ij}\partial_i\lambda_j \restriction_{\{t=0\}} = 0.$$
This concludes the proof. □

*Remark* 8.8. In the proof, we needed the assumption that $\sum_{i,j}|\hat{g}_{ij} - \delta_{ij}| \leq \frac{1}{20}$. This can be removed by localize in neighbourhoods where the metric $\hat{g}_{ij}$ is close to come constant coefficient metric. We then change coordinates so that this is close to $\delta_{ij}$ and apply the theorem above. We then use finite speed of propagation for the equation.

*Remark* 8.9. In the proof of Theorem 6.1, we required that the coefficients satisfy $\sum_{\mu,\nu}|(g^{-1})^{\mu\nu} - m^{\mu\nu}| \leq \frac{1}{10}$ everywhere while our assumptions in Theorem 8.7 only guarantee that this holds on the initial slice. Nevertheless, one can show that by choosing $T$ sufficiently small in the proof of Theorem 6.1, we can in fact guarantee that $\sum_{\mu,\nu}|(g^{-1})^{\mu\nu} - m^{\mu\nu}| \leq \frac{1}{10}$ everywhere in $I \times \mathbb{R}^n$.

*Remark* 8.10. Of course, more generally, the initial data are not required to be posed on a $n$-dimensional manifold that has the topology of $\mathbb{R}^n$. Indeed, initial data can be posed on any Riemannian manifold.

*Remark* 8.11. One may worry whether there are any non-trivial solutions to the constraint equations. Solving the constraint equations turns out to be a subject in its own. It suffices to say for the purpose of this course that there are many solutions to the constraint equations.

*Remark* 8.12. Finally, let us mention that geometric uniqueness also holds in the sense that given two solutions $(\mathcal{M}_1, g_1)$ and $(\mathcal{M}_2, g_2)$, there exists open subsets $U_i \subset \mathcal{M}_i$ containing $\Sigma$ such that $(U_1, g_1)$ and $(U_2, g_2)$ are isometric.

## 9. Decay of solutions using energy methods

After a discussion of local theory for general nonlinear wave equations, we return to the linear wave equation on Minkowski spacetime:
$$\Box\phi = 0.$$
Recall that we have proved (twice) that the solutions to this equation disperse (in dimensions $n \geq 2$), but we are yet to give a proof using energy methods. This is the goal of this section, which follows some breakthrough ideas of Klainerman. Of course, we mention that one of the important reasons that we prove dispersion using the energy method is that it is more robust and can be applied to nonlinear problems.

We begin with the following simple lemma:

**Lemma 9.1.** *If* $\Box\phi = 0$*, then*
$$\Box(\Gamma\phi) = 0,$$
*where* $\Gamma$ *is one of the vector fields* $\Gamma \in \{\partial_t, \partial_i, \Omega_{ij} := x_i\partial_j - x_j\partial_i, S := t\partial_t + \sum_{i=1}^n x_i\partial_i, \Omega_{0i} := t\partial_i + x_i\partial_t\}$.

*Proof.* This is an easy calculation. In particular for $\Gamma \in \{\partial_t, \partial_i, x_i\partial_j - x_j\partial_i, t\partial_i + x_i\partial_t\}$,
$$\Box(\Gamma\phi) = \Gamma(\Box\phi)$$
and for $S = t\partial_t + \sum_{i=1}^n x_i\partial_i$, we have
$$\Box(S\phi) = S(\Box\phi) + 2\Box\phi.$$
□

**Definition 9.2.** We call the set of admissible vector fields in Lemma 9.1 the **commuting vector fields**.

Before we proceed, let us fix some notations. We will use the multi-index notation $\Gamma^\alpha = \Gamma_1^{\alpha_1}\Gamma_2^{\alpha_2}...\Gamma_{2n+2+\frac{n(n-1)}{2}}^{\alpha_{2n+2+\frac{n(n-1)}{2}}}$ where $\alpha = (\alpha_1, ..., \alpha_{2n+2+\frac{n(n-1)}{2}})$ and $\Gamma_1$, $\Gamma_2$, ... $\Gamma_{2n+2+\frac{n(n-1)}{2}}$ is an arbitrary ordering of the vector fields $\Gamma$'s in Lemma 9.1. Notice that $\sum_{|\alpha|\leq k}|\Gamma^\alpha\phi|$ controls arbitrary $k$ $\Gamma$ derivatives of $\phi$ since the $\Gamma$'s form a Lie algebra, i.e., $[\Gamma_i, \Gamma_j] = \sum_{k=1}^{2n+2+\frac{n(n-1)}{2}} c_k \Gamma_k$, where $c_k$ are constants depending on $i$ and $j$. As in previous sections, we will use $\partial$ to denote either $\partial_t$ or $\partial_x$.

Obviously, using Lemma 9.1, we can apply the result of the conservation of energy for each of these $\Gamma\phi$:

**Corollary 9.3.** *Given smooth and compactly supported initial data $(\phi_0, \phi_1)$ to the linear wave equation in $(n+1)$-dimensions, we have*

$$\sup_{t\in[0,\infty)} \sum_{|\alpha|\leq k} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^n)}(t) \leq C \sum_{|\alpha|\leq k} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^n)}(0)$$

*for some $C = C(n) > 0$.*

The key idea of Klainerman is that one can prove a weighted Sobolev embedding type result which takes advantage of the weights in the vector fields above that have good commutation properties with $\Box$. More precisely, he proved the following global Sobolev inequality:

**Theorem 9.4** (Klainerman–Sobolev inequality)**.** *There exists $C = C(n) > 0$ such that the following holds for all functions $\phi \in H^{\lfloor\frac{n+2}{2}\rfloor}(\mathbb{R}^n)$ for all $t \geq 0$:*

$$\sup_x (1+t+r)^{\frac{n-1}{2}}(1+|t-r|)^{\frac{1}{2}}|\phi|(t,x) \leq C \sum_{|\alpha|\leq\lfloor\frac{n+2}{2}\rfloor} \|\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^n)}(t).$$

Before we proceed to the proof, let us begin with some consequences of the theorem. The first easy corollary is the decay of the derivatives of $\phi$:

**Corollary 9.5.** *Given data $(\phi_0, \phi_1) \in C_c^\infty(\mathbb{R}^n) \times C_c^\infty(\mathbb{R}^n)$ for the wave equation in $(n+1)$-dimensions, we have the following decay estimates for $t \geq 0$:*

$$\sup_x \sum_{|\alpha|=1} |\partial^\alpha\phi|(t,x) \leq \frac{C}{(1+t)^{\frac{n-1}{2}}} \sum_{|\alpha|\leq\lfloor\frac{n+2}{2}\rfloor} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^n)}(0).$$

Of course this is the sharp decay rate for the derivatives that we obtained before using the explicit formula for the solutions. Notice that this method also has the additional advantage that it is now easy to see what norms of the initial data is the decay rate dependent on.

We also obtain an improved decay rate if $|x| \leq (1-\epsilon)t$ or $|x| \geq (1+\epsilon)t$ for some $\epsilon > 0$:

**Corollary 9.6.** *Let $\phi$ be the solution to the $(n+1)$-dimensional wave equation with initial data $(\phi_0, \phi_1) \in C_c^\infty(\mathbb{R}^n) \times C_c^\infty(\mathbb{R}^n)$. For every $\epsilon > 0$, there exists $C = C(n,\epsilon) > 0$ such that the following decay estimate holds in $S_t := \{x : |x| < (1-\epsilon)t \text{ or } |x| > (1+\epsilon)t\}$ for $t \geq 0$:*

$$\sup_{x\in S_t} \sum_{|\alpha|=1} |\partial^\alpha\phi|(t,x) \leq \frac{C}{(1+t)^{\frac{n}{2}}} \sum_{|\alpha|\leq\lfloor\frac{n+2}{2}\rfloor} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^n)}(0).$$

Of course the natural question is what we can say about $\phi$ itself using this method. Notice that if the initial data are compactly supported, we can use finite speed of propagation and integrate in the $u = t - r$ direction to get the the following result:

**Corollary 9.7.** *Let $\phi$ be the solution to the $(n+1)$-dimensional wave equation such that the initial data $(\phi_0, \phi_1)$ are compactly supported in $B(0,R) \subset \mathbb{R}^n$. Then, there exists $C = C(n,R) > 0$ such that*

$$\sup_x (1+t+r)^{\frac{n-1}{2}}(1+|t-r|)^{-\frac{1}{2}}|\phi|(t,x) \leq C \sum_{|\alpha|\leq\lfloor\frac{n+2}{2}\rfloor} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^n)}(0)$$

*for $t \geq 0$.*

Note that this decay rate is worse than the sharp rate for compactly supported data. In $3+1$ dimensions, one can in fact get the sharp decay rate $|\phi| \leq \frac{C}{1+t}$ by not only introducing the commuting vector fields $\Gamma$'s but also introducing a different energy. See example sheet.

Moreover, we see that in fact some derivatives of $\phi$ decay better than the others! This fact is not so easy to see even with the formulas for explicit solutions. However, it is an easy consequence of the methods above.

Before we state the result, let us begin by introducing some notations: Let $v = t + r$ and $u = t - r$, where $r = |x|$. Notice that we have

$$(\partial_t \phi)^2 + \sum_{i=1}^n (\partial_i \phi)^2 = (\partial_t \phi)^2 + (\partial_r \phi)^2 + |\nabla\!\!\!/ \phi|^2 = 2(\partial_v \phi)^2 + 2(\partial_u \phi)^2 + |\nabla\!\!\!/ \phi|^2.$$

Here, $|\nabla\!\!\!/ \phi|$ is the angular part of the derivatives, whose norm is defined by

$$|\nabla\!\!\!/ \phi|^2 := \frac{1}{2} \sum_{i,j=1}^n (\frac{x_i}{r} \partial_j \phi - \frac{x_j}{r} \partial_i \phi)^2.$$

We finally define the "good derivatives" $\bar{\partial}$ to be either the $\partial_v$ derivative or the angular derivatives, i.e.

$$|\bar{\partial}\phi|^2 := 2(\partial_v \phi)^2 + |\nabla\!\!\!/ \phi|^2.$$

**Corollary 9.8.** *Let $\phi$ be the solution to the $(n+1)$-dimensional wave equation such that the initial data $(\phi_0, \phi_1)$ are compactly supported in $B(0, R) \subset \mathbb{R}^n$. Then, there exists $C = C(n, R) > 0$ such that*

$$(1 + t + r)^{\frac{n+1}{2}} (1 + |t - r|)^{-\frac{1}{2}} |\bar{\partial}\phi|(t, x) \leq C \sum_{|\alpha| \leq \lfloor \frac{n+4}{2} \rfloor} \|\partial \Gamma^\alpha \phi\|_{L^2(\mathbb{R}^n)}(0)$$

*for $t \geq 0$.*

*Proof.* It suffices to establish the bound

$$|\bar{\partial}\phi| \leq \frac{C \sum_{|\alpha|=1} |\Gamma \phi|}{1 + t + r} \tag{9.1}$$

so that we can apply Corollary 9.7. Clearly, we only need to consider the case $t + r > 1$ for in the case $t + r \leq 1$, the right hand side controls all derivatives. To conclude, we note that

$$\partial_v = \frac{S + \sum_{i=1}^n \frac{x_i}{r} \Omega_{0i}}{2(t + r)}, \quad \Omega_{ij} = \frac{1}{t}(x_i \Omega_{0j} - x_j \Omega_{0i}).$$

Also, by definition,

$$|\nabla\!\!\!/ \phi| \leq \frac{C}{r} \sum_{i,j=1}^n |\Omega_{ij} \phi|,$$

which implies also that

$$|\nabla\!\!\!/ \phi| \leq \frac{C}{t} \sum_{i=1}^n |\Omega_{0i} \phi|$$

by the above identity. The result follows. □

*Remark* 9.9. In the physical $3 + 1$ dimensions, note the following important consequence of Corollary 9.8: $\sup_x |\partial\phi|(t, x)$ is in general not integrable in time, while $\sup_x |\bar{\partial}\phi|(t, x)$ is!

We now turn to the proof of Theorem 9.4. We begin with two lemmas. First, we have

**Lemma 9.10.** *There exists $C = C(n, k) > 0$ such that for $t \geq 0$, it holds that*

$$\sum_{|\alpha|=k} |\partial^\alpha \phi| \leq \frac{C}{|t - r|^k} \sum_{|\alpha| \leq k} |\Gamma^\alpha \phi|$$

*for every $C^k$ functions $\phi : \mathbb{R}^{n+1} \to \mathbb{R}$.*

*Proof.* This follows from applying iteratively the identities

$$\partial_i = \frac{-x_j \Omega_{ij} + t\Omega_{0i} - x_i S}{t^2 - r^2}$$

and

$$\partial_t = \frac{tS - x_i \Omega_{0i}}{t^2 - r^2}.$$

□

The second lemma is a scale-invariant version of the Sobolev embedding theorem:

**Lemma 9.11.** *For $k \in \mathbb{Z}$, $k > \frac{n}{2}$, there exists $C = C(n,k) > 0$ such that the following holds for all functions $\phi \in H^k(\mathbb{R}^n)$:*

$$\|\phi\|_{L^\infty} \leq C\|\phi\|_{L^2(\mathbb{R}^n)}^{\frac{2k-n}{2k}} (\sum_{|\alpha|=k} \|\partial_x^\alpha \phi\|_{L^2(\mathbb{R}^n)})^{\frac{n}{2k}}. \tag{9.2}$$

*Proof.* We begin with the following standard Sobolev embedding theorem:

$$\|\phi\|_{L^\infty} \leq C\|\phi\|_{H^k(\mathbb{R}^n)} \leq C(\|\phi\|_{L^2(\mathbb{R}^n)} + \sum_{|\alpha|=k} \|\partial_x^\alpha \phi\|_{L^2(\mathbb{R}^n)})$$

for $k \in \mathbb{Z}$, $k > \frac{n}{2}$. Introduce now a scaling parameter $\lambda$ and rescale $\phi$ by

$$\phi_\lambda(x) = \phi(\lambda x).$$

Notice that for every $m \in \mathbb{Z}$, we have

$$\sum_{|\alpha|=m} \|\partial_x^\alpha \phi_\lambda\|_{L^2(\mathbb{R}^n)} = \lambda^{m-\frac{n}{2}} \sum_{|\alpha|=m} \|\partial_x^\alpha \phi_\lambda\|_{L^2(\mathbb{R}^n)}.$$

Therefore, by the Sobolev embedding theorem above, which holds for all $\phi_\lambda$, we have a family of estimates

$$\|\phi\|_{L^\infty} \leq C\left(\lambda^{-\frac{n}{2}}\|\phi\|_{L^2(\mathbb{R}^n)} + \lambda^{k-\frac{n}{2}} \sum_{|\alpha|=k} \|\partial_x^\alpha \phi\|_{L^2(\mathbb{R}^n)}\right).$$

Choosing $\lambda = \left(\sum_{|\alpha|=k} \|\partial_x^\alpha \phi\|_{L^2(\mathbb{R}^n)}\right)^{-\frac{1}{k}} \left(\|\phi\|_{L^2(\mathbb{R}^n)}\right)^{\frac{1}{k}}$, we get the desired conclusion. □

We are now ready to prove Theorem 9.4:

*Proof.* We will give a proof here which requires replacing the sum $\sum_{|\alpha| \leq \lfloor \frac{n+2}{2} \rfloor}$ by $\sum_{|\alpha| \leq \lfloor \frac{n+3}{2} \rfloor}$ on the right hand side of the inequality[12]. For a proof of the inequality with the stated number of derivatives[13], we refer the readers to Chapter 2.1 of *Lectures of nonlinear wave equations* by Sogge.

(1) We first consider the regions $\{r \leq \frac{t}{2}\}$ and $\{r \geq 2t\}$. By standard Sobolev embedding theorem, we can clearly assume that $|t+r| \geq 1$. Now, we can use Lemma 9.10 to exchage $\partial$ with $\Gamma$ while gaining a weight in $|t-r|^{-1}$. In order to ensure that the weight inside the integral is comparable to that outside the integral, we now introduce appropriate cutoff functions and use the estimate above. First, let $\chi(\frac{r}{t})$ be a cutoff function such that

$$\chi(x) = \begin{cases} 1, & \text{if } x \leq \frac{1}{2} \\ 0, & \text{if } x \geq \frac{3}{4}. \end{cases}$$

Using Lemma 9.10, we have

$$\begin{aligned}\sum_{|\alpha|=k} \|\partial^\alpha(\chi\phi)\|_{L^2(\mathbb{R}^n)} \leq & C\sum_{|\alpha|\leq k} t^{-k+|\alpha|}(\int_{\{r\leq \frac{3t}{4}\}} |\partial^\beta \phi|^2(x)\, dx)^{\frac{1}{2}} \\ \leq & C\sum_{|\alpha|\leq k} t^{-k}(\int_{\{r\leq \frac{3t}{4}\}} |\Gamma^\beta \phi|^2(x)\, dx)^{\frac{1}{2}}\end{aligned}$$

Returning to (9.2), we get the desired result when $r \leq \frac{t}{2}$. The case $r \geq 2t$ can be treated similarly by introducing an appropriate cutoff and using (9.2).

(2) With the result in part (1), it remains to consider the region $\{\frac{t}{2} \leq r \leq 2t\}$. We can assume that $|t-r| \geq 1$. Let us first explain our strategy in a heuristic manner. As we saw above in step (1), what we basically did was to gain a weight of $|t-r|^{\frac{1}{2}}$ for the Sobolev embedding in each direction. To sharpen this when $|t-r|$ is small, note that we can separate the angular directions and the radial direction and for each of the angular directions, we in fact gain $r^{\frac{1}{2}}$. Now, in the region that we are interested in, i.e., $\{\frac{t}{2} \leq r \leq 2t\}$, $t$ is comparable of $r$ and therefore the gain in $r$ is equivalent to $t$.

---

[12]Notice that in $n+1$ dimensions where $n$ is even, the following proof in fact give the desired statement.

[13]Notice that in the applications that follow, it will be sufficient to apply a form of Klainerman–Sobolev inequality with more derivatives, as long as we also assume more regularity on the initial data.

We now turn to the details. We write $x = (r, \vartheta)$, where $\vartheta \in \mathbb{S}^{n-1}$. Denote also by $d\sigma_\vartheta$ the standard measure on the sphere $\mathbb{S}^{n-1}$ of radius 1. Notice that the volume form on $\mathbb{R}^n$ can be written as $r^{n-1} dr d\sigma_\vartheta$. Introduce a cutoff function $\chi(\frac{r}{t})$ such that

$$\chi(x) = \begin{cases} 1, & \text{if } \frac{1}{2} \leq x \leq 2 \\ 0, & \text{if } x < \frac{1}{4} \text{ or } x > 4. \end{cases}$$

where $\chi$ is smooth, compactly supported and $0 \leq \chi \leq 1$ everywhere. Obviously it suffices to control $\chi(\frac{r}{t})\phi$. To proceed, notice that we have the Sobolev embedding theorem on the sphere, i.e., there exists $C = C(n) > 0$ such that

$$|\phi| \leq C\Big(\sum_{|\alpha| \leq \lfloor \frac{n+1}{2} \rfloor} \int_{\mathbb{S}^{n-1}} |\Omega_{ij}^\alpha \phi|^2(\vartheta)\, d\sigma_\vartheta\Big)^{\frac{1}{2}}.$$

First, assume that $r \leq t$.

$$\begin{aligned}
&|\chi\phi|(t,x) \\
\leq& C \sum_{|\alpha| \leq \lfloor \frac{n+1}{2} \rfloor} \Big(\int_{\mathbb{S}^{n-1}} |\chi|^2 |\Omega_{ij}^\alpha \phi|^2(r,\vartheta)\, d\sigma_\vartheta\Big)^{\frac{1}{2}} \\
\leq& C \sum_{|\alpha| \leq \lfloor \frac{n+1}{2} \rfloor} \Big(\int_{\mathbb{S}^{n-1}} \Big(\int_{\frac{t}{4}}^r |\partial_r(\chi \Omega_{ij}^\alpha \phi)|(r',\vartheta)\, dr'\Big)^2 d\sigma_\vartheta\Big)^{\frac{1}{2}} \\
\leq& C \sum_{|\alpha| \leq \lfloor \frac{n+1}{2} \rfloor} \sum_{|\beta| \leq 1} \Big(\int_{\mathbb{S}^{n-1}} \Big(\int_{\frac{t}{4}}^r \frac{1}{|t-r'|} |\Gamma^\beta(\chi \Omega_{ij}^\alpha \phi)|(r',\vartheta)\, dr'\Big)^2 d\sigma_\vartheta\Big)^{\frac{1}{2}} \\
\leq& C \sum_{|\alpha| \leq \lfloor \frac{n+1}{2} \rfloor} \sum_{|\beta| \leq 1} \Big(\int_{\mathbb{S}^{n-1}} \Big(\int_{\frac{t}{4}}^r |\Gamma^\beta(\chi \Omega_{ij}^\alpha \phi)|^2(r',\vartheta)\, dr'\Big)\Big(\int_{\frac{t}{4}}^r \frac{1}{|t-r'|^2}\, dr'\Big)\, d\sigma_\vartheta\Big)^{\frac{1}{2}} \\
\leq& C \sum_{|\alpha| \leq \lfloor \frac{n+1}{2} \rfloor} \sum_{|\beta| \leq 1} \Big(\int_{\mathbb{S}^{n-1}} \Big(\int_{\frac{t}{4}}^r |\Gamma^\beta(\chi \Omega_{ij}^\alpha \phi)|^2(r',\vartheta)\, dr'\Big)|t-r|^{-1}\, d\sigma_\vartheta\Big)^{\frac{1}{2}} \\
\leq& \frac{C}{|t+r|^{\frac{n-1}{2}}|t-r|^{\frac{1}{2}}} \sum_{|\alpha| \leq \lfloor \frac{n+1}{2} \rfloor} \sum_{|\beta| \leq 1} \Big(\int_{\mathbb{S}^{n-1}} \int_{\frac{t}{4}}^r |\Gamma^\beta(\chi \Omega_{ij}^\alpha \phi)|^2(r',\vartheta)(r')^{n-1}\, dr'\, d\sigma_\vartheta\Big)^{\frac{1}{2}} \\
\leq& \frac{C}{|t+r|^{\frac{n-1}{2}}|t-r|^{\frac{1}{2}}} \sum_{|\alpha| \leq \lfloor \frac{n+3}{2} \rfloor} \Big(\int_{\mathbb{R}^n} |\Gamma^\alpha \phi|^2(r',\vartheta)(r')^{n-1}\, dr'\, d\sigma_\vartheta\Big)^{\frac{1}{2}},
\end{aligned}$$

where in the last line we have used that for $\frac{t}{4} \leq r \leq t$, we have $\sum_{|\alpha|=1} |\Gamma^\alpha \chi| \leq C$. Finally, the case $r > t$ can be treated similarly by integrating from $r = 4t$ instead of $r = \frac{t}{4}$.

□

## 10. Small data global regularity of solutions

In the rest of the course, we will be concerned with small data solutions to nonlinear equations. As we mentioned in Section 7, very little is known about the global behaviour of solutions to supercritical equations with general data. On the other hand, much more can be said about these equations when restricted to small initial data.

Roughly speaking, this is because when the initial data are small, then the nonlinear terms are even smaller and thus the linear terms dominate so that we can control the nonlinear terms. However, recall that when we estimate the error terms using energy estimates, they have to be integrated in time. In order to actually control these terms globally in time, we must therefore show at the same time that the solution decays sufficiently fast and the error terms are integrable. It turns out that the methods introduced in Section 9 can easily be adapted for nonlinear equations and are very well-suited for the purpose of showing the decay of the solutions.

We begin with a simple example illustrating the method.

**Theorem 10.1.** *Let $k \geq 6$. Consider the wave map equation*

$$\Box\phi = \phi(\partial_t\phi^{\mathbf{T}}\partial_t\phi - \sum_{i=1}^{4}\partial_i\phi^{\mathbf{T}}\partial_i\phi).$$

*in $\mathbb{R}\times\mathbb{R}^4$ with initial data*[14]

$$(\phi_0,\phi_1)\restriction_{\{t=0\}} \in C^\infty_c(B(0,R))\times C^\infty_c(B(0,R))$$

*such that*

$$\sum_{|\alpha|\leq k}\|\partial\partial^\alpha\phi_0\|_{L^2(\mathbb{R}^4)} + \|\partial^\alpha\phi_1\|_{L^2(\mathbb{R}^4)} < \epsilon.$$

*Then for every $R > 0$, there exists $\epsilon_0 = \epsilon_0(R) > 0$ sufficiently small such that if $\epsilon \leq \epsilon_0$, the unique solution remains smooth for all time.*

*Proof.* We prove this using the bootstrap method. We will assume that some weighted energy is bounded by some constant. Then by Klainerman–Sobolev inequality, we can show that the solution decays. Because the initial data are small, this allows us to prove that the weighted energy is bounded by some better constant. Thus by continuity, we conclude that the weighted energy cannot grow to infinity in any finite time interval and hence using the local existence theorem, the solution exists for all time.

We now turn to the details. Assume that

$$\sup_{t\in[0,T]}\sum_{|\alpha|\leq k}\|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^4)}(t) \leq \epsilon^{\frac{3}{4}} \tag{10.1}$$

for $T < T_*$. Notice that $\epsilon$ sufficiently small, we have $\epsilon \ll \epsilon^{\frac{3}{4}}$ and hence this holds initially.

By Klainerman–Sobolev inequality (and the fact that $k \geq 6$), we therefore have that for some constant $C > 0$,

$$\sum_{|\alpha|\leq\frac{k}{2}}\|(1+t+|x|)^{\frac{3}{2}}\partial\Gamma^\alpha\phi\|_{L^\infty(\mathbb{R}^4)}(t) \leq C\epsilon^{\frac{3}{4}}. \tag{10.2}$$

This further implies that we have

$$\sum_{1\leq|\alpha|\leq\frac{k}{2}+1}\|(1+t+|x|)^{\frac{1}{2}}\Gamma^\alpha\phi\|_{L^\infty(\mathbb{R}^4)}(t) \leq C\epsilon^{\frac{3}{4}}. \tag{10.3}$$

This is because we can trivially bound

$$\sum_{1\leq|\alpha|\leq\frac{k}{2}+1}|\Gamma^\alpha\phi| \leq C(1+t+|x|)\sum_{|\alpha|\leq\frac{k}{2}}|\partial\Gamma^\alpha\phi|.$$

We now turn to the energy estimates. For some constant depending on $k$ (but independent of $T$), we have

$$\begin{aligned}&\sup_{t\in[0,T]}\sum_{|\alpha|\leq k}\|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^4)}(t)\\ \leq& C(k)(\sum_{|\alpha|\leq k}\|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^4)}(0) + \int_0^T\sum_{|\alpha|\leq k}\|\Gamma^\alpha\left(\phi(\partial_t\phi^{\mathbf{T}}\partial_t\phi - \sum_{i=1}^{n}\partial_i\phi^{\mathbf{T}}\partial_i\phi)\right)\|_{L^2(\mathbb{R}^4)}(t)dt.\end{aligned}$$

Notice that

$$\begin{aligned}&\sum_{|\alpha|\leq k}\left|\Gamma^\alpha\left(\phi(\partial_t\phi^{\mathbf{T}}\partial_t\phi - \sum_{i=1}^{n}\partial_i\phi^{\mathbf{T}}\partial_i\phi)\right)\right|\\ \leq& C\left(\sum_{|\alpha_1|+|\alpha_2|\leq k}|\partial\Gamma^{\alpha_1}\phi||\partial\Gamma^{\alpha_2}\phi| + \sum_{|\alpha_1|+|\alpha_2|+|\alpha_3|\leq k,|\alpha_1|\geq 1}|\Gamma^{\alpha_1}\phi||\partial\Gamma^{\alpha_2}\phi||\partial\Gamma^{\alpha_3}\phi|\right).\end{aligned}$$

[14]Notice that by compactly supported, we mean that the map coincide with the constant map outside a ball. Moreover, as always, when we prescribe initial data for the wave map problem, we require $|\phi_0|^2 = 1$ and $\phi_1^{\mathbf{T}}\phi_0 = 0$.

We estimate the first term. Clearly, either $|\alpha_1| \leq \frac{k}{2}$ or $|\alpha_2| \leq \frac{k}{2}$. Without loss of generality, we assume $|\alpha_1| \leq \frac{k}{2}$. We can therefore apply (10.2). More precisely,

$$\begin{aligned}&\sum_{|\alpha_1|+|\alpha_2|\leq k}\int_0^T \||\partial\Gamma^{\alpha_1}\phi||\partial\Gamma^{\alpha_2}\phi|\|_{L^2(\mathbb{R}^4)}(t)dt\\ \leq & C\int_0^T (\sum_{|\alpha_1|\leq\frac{k}{2}}\|\partial\Gamma^{\alpha_1}\phi\|_{L^\infty(\mathbb{R}^4)})(\sum_{|\alpha_2|\leq k}\|\partial\Gamma^{\alpha_2}\phi\|_{L^2(\mathbb{R}^4)})(t)dt\\ \leq & C\epsilon^{\frac{3}{4}}\int_0^T \frac{(\sum_{|\alpha_2|\leq k}\|\partial\Gamma^{\alpha_2}\phi\|_{L^2(\mathbb{R}^4)})}{(1+t)^{\frac{3}{2}}}dt \leq C\epsilon^{\frac{3}{2}},\end{aligned}$$

where we have used (10.1) in the last step. The most important point is that $\frac{1}{(1+t)^{\frac{3}{2}}}$ is integrable in time and therefore this term can be bounded independent of $T$! We now look at the cubic term. At least two of $|\alpha_1|$, $|\alpha_2|$ and $|\alpha_3|$ is $\leq \frac{k}{2}$. There are two cases: $|\alpha_1|, |\alpha_2| \leq \frac{k}{2}$ and $|\alpha_2|, |\alpha_3| \leq \frac{k}{2}$. More precisely, we have

$$\begin{aligned}&\sum_{|\alpha_1|+|\alpha_2|\leq k}\int_0^T \||\Gamma^{\alpha_1}\phi||\partial\Gamma^{\alpha_2}\phi||\partial\Gamma^{\alpha_3}\phi|\|_{L^2(\mathbb{R}^4)}(t)dt\\ \leq & C\int_0^T (\sum_{1\leq|\alpha_1|\leq\frac{k}{2}}\|\Gamma^{\alpha_1}\phi\|_{L^\infty(\mathbb{R}^4)})(\sum_{|\alpha_2|\leq\frac{k}{2}}\|\partial\Gamma^{\alpha_2}\phi\|_{L^\infty(\mathbb{R}^4)})(\sum_{|\alpha_3|\leq k}\|\partial\Gamma^{\alpha_3}\phi\|_{L^2(\mathbb{R}^4)})(t)dt\\ &+C\int_0^T (\sum_{1\leq|\alpha_1|\leq k}\|\Gamma^{\alpha_1}\phi\|_{L^\infty(\mathbb{R}^4)})(\sum_{|\alpha_2|\leq\frac{k}{2}}\|\partial\Gamma^{\alpha_2}\phi\|_{L^\infty(\mathbb{R}^4)})(\sum_{|\alpha_3|\leq \frac{k}{2}}\|\partial\Gamma^{\alpha_3}\phi\|_{L^2(\mathbb{R}^4)})(t)dt\\ \leq & C\int_0^T (\sum_{1\leq|\alpha_1|\leq\frac{k}{2}}\|\Gamma^{\alpha_1}\phi\|_{L^\infty(\mathbb{R}^4)})(\sum_{|\alpha_2|\leq\frac{k}{2}}\|\partial\Gamma^{\alpha_2}\phi\|_{L^\infty(\mathbb{R}^4)})(\sum_{|\alpha_3|\leq k}\|\partial\Gamma^{\alpha_3}\phi\|_{L^2(\mathbb{R}^4)})(t)dt\\ &+C\int_0^T (1+t)(\sum_{|\alpha_1|\leq k-1}\|\partial\Gamma^{\alpha_1}\phi\|_{L^2(\mathbb{R}^4)})(\sum_{|\alpha_2|\leq\frac{k}{2}}\|\partial\Gamma^{\alpha_2}\phi\|_{L^\infty(\mathbb{R}^4)})(\sum_{|\alpha_3|\leq \frac{k}{2}}\|\partial\Gamma^{\alpha_3}\phi\|_{L^\infty(\mathbb{R}^4)})(t)dt\\ \leq & C\epsilon^{\frac{3}{2}}\int_0^T \frac{(\sum_{|\alpha|\leq k}\|\partial\Gamma^{\alpha}\phi\|_{L^2(\mathbb{R}^4)})}{(1+t)^2}dt \leq C\epsilon^{\frac{9}{4}}.\end{aligned}$$

Combining the estimates above, we have thus obtained

$$\begin{aligned}&\sup_{t\in[0,T]}\sum_{|\alpha|\leq k}\|\partial\Gamma^{\alpha}\phi\|_{L^2(\mathbb{R}^4)}(t)\\ \leq & C(k)(\sum_{|\alpha|\leq k}\|\partial\Gamma^{\alpha}\phi\|_{L^2(\mathbb{R}^4)}(0)+\epsilon^{\frac{3}{2}}) \leq C\epsilon.\end{aligned}$$

Since $C\epsilon \leq \frac{\epsilon^{\frac{3}{4}}}{2}$, we have thus improved the constant in (10.1) and concluded the argument. □

The above theorem is of course for the wave map equations in $4+1$ dimensions. However, as one can see from the proof, the method can be quite generally applied in a neighbourhood of the zero solution for a large class nonlinear wave equation with quadratic nonlinear in the derivatives of $\phi$. On the other hand, it is important to note that the proof fails in $3+1$ dimensions! Indeed, that case requires a more delicate analysis and will be the subject we turn to in the next section.

## 11. The null condition

As we have seen in the previous section, in order to prove the small data global regularity result for the wave map equations in $4+1$ dimensions, we have crucially used the fact that $\int_0^\infty \frac{dt}{(1+t)^{\frac{3}{2}}} < \infty$. The same proof however fails in $3+1$ dimensions since in that case, we only have a decay rate of $\frac{1}{1+t}$ and $\int_0^\infty \frac{dt}{(1+t)} = \infty$!

Nevertheless, the same result is in fact also true, but requires a more delicate argument. There are two main observations, the first is in the following lemma:

**Lemma 11.1.** *The following identity holds*

$$\partial_t\phi\partial_t\psi - \sum_{i=1}^{3}\partial_i\phi\partial_i\psi = 2\partial_u\phi\partial_v\psi + 2\partial_v\phi\partial_u\psi - \nabla\!\!\!\!/\phi\cdot\nabla\!\!\!\!/\psi,$$

*where* $\nabla\!\!\!\!/\phi$ *is defined as an* $n$*-dimensional vector with components given by*

$$\nabla\!\!\!\!/_i\phi := \partial_i\phi - \frac{x_i}{r}\partial_r\phi = \sum_{j=1}^{n}\frac{x_j}{r^2}\Omega_{ij}\phi.$$

*The dot product* $\cdot$ *here is simply the standard dot product in* $\mathbb{R}^n$*. Note that we have*

$$\nabla\!\!\!\!/\phi\cdot\nabla\!\!\!\!/\psi := \frac{1}{2r^2}\sum_{i,j=1}^{3}\Omega_{ij}\phi\Omega_{ij}\psi.$$

*Proof.* It is easy to check that

$$\sum_{i=1}^{3}\partial_i\phi\partial_i\psi = \partial_r\phi\partial_r\psi + \frac{1}{2r^2}\sum_{i,j=1}^{3}\Omega_{ij}\phi\Omega_{ij}\psi$$

and

$$\partial_t\phi\partial_t\psi - \partial_r\phi\partial_r\psi = \frac{1}{2}(\partial_t\phi+\partial_r\phi)(\partial_t\psi-\partial_r\psi) + \frac{1}{2}(\partial_t\phi-\partial_r\phi)(\partial_t\psi+\partial_r\psi).$$

□

In other words, in any quadratic terms, at least one of the derivatives is a $\overline{\partial}$ derivative, where as before $\overline{\partial}\in\{\partial_v,\nabla\!\!\!\!/\}$. As we have seen previously (see Corollary 9.8), in fact $|\overline{\partial}\phi|$ decays better than $|\partial\phi|$. Most importantly, the decay is integrable in time! This is good news, but before we declare victory, remember that we need to put one of the two factors in the quadratic term in $L^2$ and the other in $L^\infty$. In particular, we need to show that not only is $|\overline{\partial}\phi|$ better in $L^\infty$ norm, but it is also better in some $L^2$ sense. This is provided by the following proposition:

**Proposition 11.2.** *Let*

$$\Box\phi = F.$$

*For every* $\delta>0$*, there exists* $C=C(\delta)>0$ *such that*

$$\left(\int_0^T\int_{\mathbb{R}^3}\frac{|\overline{\partial}\phi|^2}{(1+|t-|x||)^{1+\delta}}dxdt\right)^{\frac{1}{2}} \leq C\left(\|(\phi_0,\phi_1)\|_{\dot{H}^1(\mathbb{R}^3)\times L^2(\mathbb{R}^3)} + \int_0^T\|F\|_{L^2(\mathbb{R}^3)}(t)dt\right).$$

*Proof.* Let $w:\mathbb{R}\to\mathbb{R}$ be a $C^1$ function to be determined.

$$\begin{aligned}
&\int_0^T\int_{\mathbb{R}^3} w(t-|x|)\partial_t\phi F dxdt\\
=&\int_0^T\int_{\mathbb{R}^3} w(t-|x|)\partial_t\phi\Box\phi dxdt\\
=&-\frac{1}{2}\int_{\mathbb{R}^3} w(T-|x|)\left((\partial_t\phi)^2+\sum_{i=1}^{3}(\partial_i\phi)^2\right)dx(T) + \frac{1}{2}\int_{\mathbb{R}^3} w(-|x|)\left((\partial_t\phi)^2+\sum_{i=1}^{3}(\partial_i\phi)^2\right)dx(0)\\
&+\frac{1}{2}\int_0^T\int_{\mathbb{R}^3} w'(t-|x|)\left((\partial_t\phi)^2+\sum_{i=1}^{3}\frac{2x_i}{r}\partial_t\phi\partial_i\phi+\sum_{i=1}^{3}(\partial_i\phi)^2\right)dxdt.
\end{aligned}$$

Observe now that

$$\begin{aligned}
&(\partial_t\phi)^2+\sum_{i=1}^{3}\frac{2x_i}{r}\partial_t\phi\partial_i\phi+\sum_{i=1}^{3}(\partial_i\phi)^2\\
=&(\partial_t\phi)^2+2\partial_t\phi\partial_r\phi+(\partial_r\phi)^2+|\nabla\!\!\!\!/\phi|^2 = 4(\partial_v\phi)^2+|\nabla\!\!\!\!/\phi|^2.
\end{aligned}$$

Now, we choose $w$ such that it is bounded and decreasing and satisfying $-w'(s) \geq c(1+|s|)^{-1-\delta}$ (where $c$ is a constant which can depend on $\delta$). This can be achieved for instance by

$$w(s) = \begin{cases} (1+s)^{-\delta} & \text{if } s \geq 0, \\ 1 + \delta \int_s^0 \frac{dq}{(1+|q|)^{-1-\delta}} & \text{if } s < 0. \end{cases}$$

Combining all these imply that

$$\begin{aligned} &\int_0^T \int_{\mathbb{R}^3} \frac{|\overline{\partial}\phi|^2}{(1+|t-|x||)^{1+\delta}} dxdt \\ \leq & C\left( \|(\phi_0, \phi_1)\|^2_{\dot{H}^1(\mathbb{R}^3)\times L^2(\mathbb{R}^3)} + \int_0^T (\sup_{t\in[0,T]} \|\partial\phi\|_{L^2(\mathbb{R}^3)}) \|F\|_{L^2(\mathbb{R}^3)}(t) dt \right). \end{aligned}$$

and the conclusion follows from standard energy estimates (see Theorem 4.6). □

Let's look at what we have achieved. We now have an $L^2$ estimate which is integrated in time but contains a weight. Notice that the key point is that the weight is in $|t-|x||$ - indeed, the corresponding estimate with $t+|x|$ weight for

$$\int_0^T \int_{\mathbb{R}^3} \frac{|\partial\phi|^2}{(1+t+|x|)^{1+\delta}} dxdt$$

holds trivially since $\frac{1}{(1+t)^{1+\delta}}$ is integrable in $t$. The point is now that we also have a pointwise decay $|\partial\phi| \leq \frac{C}{(1+t+|x|)(1+|t-|x||)^{\frac{1}{2}}}$. The above estimate, which degenerates when $|t-|x||$ is small, allows us to exploit this extra decay in $|t-|x||$, which was previously not used in the $4+1$ dimensional case.

Now we have good estimates in both $L^\infty$ and $L^2$ for the good estimate $\overline{\partial}$. Before we state and prove the main theorem, the remaining thing to make sure is that after we differentiate the nonlinear term, we still preserve the structure that at least one of the factors has a good derivative. To this end, it is convenient to first introduce a larger class of bilinear forms which have the property that at least one of the factors has a good derivative.

**Definition 11.3.** Let $q^{\alpha\beta}$ be constants. We say that $Q(\phi,\psi) = q^{\alpha\beta}\partial_\alpha\phi\partial_\beta\psi$ satisfies the **classical null condition** if

$$q^{\alpha\beta}\xi_\alpha\xi_\beta = 0, \text{ whenever } m^{\alpha\beta}\xi_\alpha\xi_\beta = 0.$$

When $Q$ satisfies the classical null condition, we also say that $Q$ is a **classical null form**.

It is easy to classify all classical null forms and show that it has the desired property:

**Lemma 11.4.** *Let $Q$ be a classical null form. Then it is a linear combination of $Q_0(\phi,\psi) := m^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi$ and $Q_{\mu\nu}(\phi,\psi) := \partial_\mu\phi\partial_\nu\psi - \partial_\nu\phi\partial_\mu\psi$. Moreover, we have*

$$|Q(\phi,\psi)| \leq C(|\partial\phi||\overline{\partial}\psi| + |\overline{\partial}\phi||\partial\psi|).$$

*Proof.* Clearly $Q_{\mu\nu}$ span the space of anti-symmetric bilinear forms (which has dimension $\frac{(n+1)!}{2!(n-1)!}$), all of which are null forms. It remains to show that all symmetric null forms are multiples of $Q_0$. This is obvious since $q^{\alpha\beta}\xi_\alpha\xi_\beta$ determines a homogeneous polynomial in $\xi$ of degree 2 and has the zero set that coincides with the polynomial $\xi_0^2 - \sum_{i=1}^n \xi_i^2$. The final statement can be checked directly for each of $Q_0$ and $Q_{\mu\nu}$. $Q_0$ is checked in Lemma 11.1 and we leave it as an exercise for the readers to show that $Q_{\mu\nu}$ also verifies the desired property. □

As mentioned above, we then show that the null structure is "preserved" after differentiating by $\Gamma$:

**Lemma 11.5.** *Given a classical null form $Q$. Then for every commuting vector field $\Gamma \in \{\partial_t, \partial_i, \Omega_{ij}, \Omega_{0i}, S\}$, we have*

$$\Gamma(Q(\phi,\psi)) = Q(\Gamma\phi,\psi) + Q(\phi,\Gamma\psi) + \tilde{Q}(\phi,\psi),$$

*for some classical null form $\tilde{Q}$.*

*Proof.* We check this for $Q_0$. We write[15] $\Gamma = \Gamma^\alpha \partial_\alpha$ to get

$$\begin{aligned}&\Gamma(m^{\alpha\beta}\partial_\alpha\phi\partial_\beta\psi)\\=&(m^{\alpha\beta}\partial_\alpha(\Gamma\phi)\partial_\beta\psi)+(m^{\alpha\beta}\partial_\alpha\phi\partial_\beta(\Gamma\psi))-(m^{\alpha\beta}(\partial_\alpha\Gamma^\sigma)\partial_\sigma\phi\partial_\beta\psi)-(m^{\alpha\beta}(\partial_\beta\Gamma^\sigma)\partial_\alpha\phi\partial_\sigma\psi)\\=&(m^{\alpha\beta}\partial_\alpha(\Gamma\phi)\partial_\beta\psi)+(m^{\alpha\beta}\partial_\alpha\phi\partial_\beta(\Gamma\psi))-m^{\alpha\beta}(\partial_\alpha\Gamma^\sigma)(\partial_\sigma\phi\partial_\beta\psi+\partial_\beta\phi\partial_\sigma\psi).\end{aligned}$$

Now it is easy to check that for any commuting vector fields $\Gamma$, the last term can be written as a linear combination of null forms. $Q_{\mu\nu}$ can be checked in a similar manner and we leave it as an exercise. □

We are now ready to state and prove the main theorem of this section:

**Theorem 11.6.** *Consider the wave map equation*

$$\Box\phi = \phi(\partial_t\phi^{\mathbf{T}}\partial_t\phi - \sum_{i=1}^3 \partial_i\phi^{\mathbf{T}}\partial_i\phi).$$

*in* $\mathbb{R}\times\mathbb{R}^3$ *with initial data*[16]

$$(\phi_0,\phi_1)\restriction_{\{t=0\}} \in C^\infty_c(B(0,R))\times C^\infty_c(B(0,R))$$

*such that*

$$\sum_{|\alpha|\leq 5} \|\partial\partial^\alpha\phi_0\|_{L^2(\mathbb{R}^3)} + \|\partial^\alpha\phi_1\|_{L^2(\mathbb{R}^3)} < \epsilon.$$

*Then for every* $R>0$*, there exists* $\epsilon_0=\epsilon_0(R)>0$ *sufficiently small such that if* $\epsilon\leq\epsilon_0$*, the unique solution remains smooth for all time.*

*Proof.* We will use a bootstrap argument as in the $(4+1)$-dimensional case. Fix $\delta\in(0,1)$. The constant $C$ in the proof of the theorem is allowed to depend on $\delta$ and $R$ (but is independent of $T$ and $\epsilon$). Assume that

$$\sup_{t\in[0,T]}\sum_{|\alpha|\leq 5}\|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)} + \sum_{|\alpha|\leq 5}\left(\int_0^T\int_{\mathbb{R}^3}\frac{|\bar{\partial}\Gamma^\alpha\phi|^2}{(1+|t-|x||)^{1+\delta}}dxdt\right)^{\frac12} \leq \epsilon^{\frac34} \tag{11.1}$$

for all $T$ such that the solution remains regular and we will show that this bound in fact holds with a better constant. By Theorem 9.4, we have

$$\sup_{t\in[0,T]}\sum_{|\alpha|\leq 3}\|(1+t)(1+|t-|x|)^{\frac12}\partial\Gamma^\alpha\phi\|_{L^\infty(\mathbb{R}^3)} \leq C\epsilon^{\frac34}. \tag{11.2}$$

Using (9.1) together with (11.2), we get

$$\sup_{t\in[0,T]}\sum_{|\alpha|\leq 2}(1+t)^{\frac32}\|\bar{\partial}\Gamma^\alpha\phi\|_{L^\infty(\mathbb{R}^3)} \leq C\epsilon^{\frac34}. \tag{11.3}$$

(11.2) also implies that[17]

$$\sup_{t\in[0,T]}\sum_{|\alpha|\leq 3}\|\Gamma^\alpha\phi\|_{L^\infty(\mathbb{R}^3)} \leq C\epsilon^{\frac34}. \tag{11.4}$$

We now use this pointwise decay bound to control the energy. By Theorem 4.6 and Proposition 11.2, we have

$$\begin{aligned}&\sup_{t\in[0,T]}\sum_{|\alpha|\leq 5}\|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)}(t) + \sum_{|\alpha|\leq 5}\left(\int_0^T\int_{\mathbb{R}^3}\frac{|\bar{\partial}\Gamma^\alpha\phi|^2}{(1+|t-|x||)^{1+\delta}}dxdt\right)^{\frac12}\\ \leq& C\left(\epsilon+\int_0^T\sum_{|\alpha|\leq 5}\|\Gamma^\alpha\left(\phi(\partial_t\phi^{\mathbf{T}}\partial_t\phi-\sum_{i=1}^3\partial_i\phi^{\mathbf{T}}\partial_i\phi)\right)\|_{L^2(\mathbb{R}^3)}(t)dt\right).\end{aligned} \tag{11.5}$$

---

[15]Not to be confused with our usual notation that $\alpha$ is a multi-index!

[16]As before, by compactly supported, we mean that the map coincide with the constant map outside a ball. Moreover, as always, when we prescribe initial data for the wave map problem, we require $|\phi_0|^2=1$ and $\phi_1^{\mathbf{T}}\phi_0=0$.

[17]In fact we even have decay for this term, but we will not need to use that.

The error term can be bounded as follows:

$$
\begin{aligned}
&\int_0^T \sum_{|\alpha|\leq 5} \|\Gamma^\alpha \left(\phi(\partial_t\phi^{\mathbf{T}}\partial_t\phi - \sum_{i=1}^3 \partial_i\phi^{\mathbf{T}}\partial_i\phi)\right)\|_{L^2(\mathbb{R}^3)}(t)dt \\
\leq &\int_0^T (\sum_{|\alpha|\leq 5} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)})(\sum_{|\beta|\leq 2} \|\overline{\partial}\Gamma^\beta\phi\|_{L^\infty(\mathbb{R}^3)})(t)dt \\
&+\int_0^T (\sum_{|\alpha|\leq 5,\, |\beta|\leq 2} \|\overline{\partial}\Gamma^\alpha\phi\partial\Gamma^\beta\phi\|_{L^2(\mathbb{R}^3)})(t)dt \\
&+\int_0^T (\sum_{1\leq|\alpha|\leq 5} \|\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)})(\sum_{|\beta|\leq 2} \|\partial\Gamma^\beta\phi\|_{L^\infty(\mathbb{R}^3)})(\sum_{|\gamma|\leq 2} \|\overline{\partial}\Gamma^\gamma\phi\|_{L^\infty(\mathbb{R}^3)})(t)dt.
\end{aligned}
$$

Notice that in the above, we have used (11.4). We now control each term

$$
\int_0^T (\sum_{|\alpha|\leq 5} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)})(\sum_{|\beta|\leq 2} \|\overline{\partial}\Gamma^\beta\phi\|_{L^\infty(\mathbb{R}^3)})(t)dt \leq C\int_0^T \frac{\epsilon^{\frac{3}{2}}dt}{(1+t)^{\frac{3}{2}}} \leq C\epsilon^{\frac{3}{2}}.
$$

We now control the second term. Here, we use the estimate in the second term of (11.1), which allows us to exploit the good derivative in $L^2$.

$$
\begin{aligned}
&\int_0^T (\sum_{|\alpha|\leq 5,\, |\beta|\leq 2} \|\overline{\partial}\Gamma^\alpha\phi\partial\Gamma^\beta\phi\|_{L^2(\mathbb{R}^3)})(t)dt \\
\leq &C \sum_{|\alpha|\leq 5,\, |\beta|\leq 2} \int_0^T (\int_{\mathbb{R}^3} |\overline{\partial}\Gamma^\alpha\phi\partial\Gamma^\beta\phi|^2 dx)^{\frac{1}{2}} dt \\
\leq &C \sum_{|\alpha|\leq 5,\, |\beta|\leq 2} \int_0^T \left(\int_{\mathbb{R}^3} \frac{|\overline{\partial}\Gamma^\alpha\phi|^2}{(1+|t-|x||)^{1+\delta}}dx\right)^{\frac{1}{2}} \left(\sup_x \left((1+|t-|x|)^{\frac{1+\delta}{2}}|\partial\Gamma^\beta\phi|(t,x)\right)\right) dt \\
\leq &C(\sum_{|\alpha|\leq 5} \int_0^T \int_{\mathbb{R}^3} \frac{|\overline{\partial}\Gamma^\alpha\phi|^2}{(1+|t-|x||)^{1+\delta}}dxdt)^{\frac{1}{2}}(\sum_{|\beta|\leq 2} \left(\int_0^T \left(\sup_x \left((1+|t-|x|)^{\frac{1+\delta}{2}}|\partial\Gamma^\beta\phi|(t,x)\right)\right)^2 dt\right)^{\frac{1}{2}} \\
\leq &C\times\epsilon^{\frac{3}{4}}\times(\int_0^T \frac{\epsilon^{\frac{3}{2}}}{(1+t)^{2-\delta}}dt)^{\frac{1}{2}} \leq C\epsilon^{\frac{3}{2}},
\end{aligned}
$$

for $\delta\in(0,1)$. Finally, the last term can be controlled by noticing that

$$
\sum_{1\leq|\alpha|\leq 5} \|\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)} \leq C(1+t)\sum_{|\alpha|\leq 4} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)},
$$

which implies

$$
\begin{aligned}
&\int_0^T (\sum_{1\leq|\alpha|\leq 5} \|\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)})(\sum_{|\beta|\leq 2} \|\partial\Gamma^\beta\phi\|_{L^\infty(\mathbb{R}^3)})(\sum_{|\gamma|\leq 2} \|\overline{\partial}\Gamma^\gamma\phi\|_{L^\infty(\mathbb{R}^3)})(t)dt \\
\leq &C\epsilon^{\frac{3}{2}} \int_0^T \frac{\sum_{|\alpha|\leq 4} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)}}{(1+t)^{\frac{3}{2}}}dt \leq C\epsilon^{\frac{9}{4}}.
\end{aligned}
$$

Returning to (11.5), we therefore have

$$
\sup_{t\in[0,T]} \sum_{|\alpha|\leq 5} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)}(t) + \sum_{|\alpha|\leq 5} \left(\int_0^T \int_{\mathbb{R}^3} \frac{|\overline{\partial}\Gamma^\alpha\phi|^2}{(1+|t-|x||)^{1+\delta}}dxdt\right)^{\frac{1}{2}} \leq C(\epsilon+\epsilon^{\frac{3}{2}}+\epsilon^{\frac{9}{4}}).
$$

As long as $\epsilon$ is sufficiently small, this improves the constant in (11.1) and we are done. □

As we saw above, the null structure in the nonlinear terms plays a crucial role in establishing global regularity for solutions arising from small data. It is clear from the proof above that it works more generally to semilinear wave equations in $(3+1)$-dimensions with quadratic nonlinearity satisfying the classical null condition.

## 12. Weak null condition

We consider the system of equations

$$\begin{cases} \Box\phi = Q_0(\psi,\psi) \\ \Box\psi = (\partial_t\phi)^2. \end{cases} \tag{12.1}$$

This equation does not obey the classical null condition. More precisely, the nonlinearity $(\partial_t\phi)^2$ is not a null form. In particular, *even if* $\partial_t\phi$ obeys the decay estimates as for the linear wave equation $\partial\psi$ *does not* decay like $\frac{1}{1+t}$ but instead only decays like $\frac{\log(2+t)}{1+t}$ (**Exercise**). Nevertheless, when $\psi$ enters in the nonlinear term, it does so in a null form. The null form provides more decay and therefore global regularity still holds for small data. Such nonlinear structure is an example of what is more generally known as the **weak null condition** of Lindblad-Rodnianski. We will not discuss systematically the weak null condition, but will just study the example (12.1) given above. We note that a very similar form of this structure will appear again when we discuss the stability of Minkowski space.

**Theorem 12.1.** *Consider the equation* (12.1)*. Suppose the initial data obey*

$$(\phi_0,\phi_1,\psi_0,\psi_1)\restriction_{\{t=0\}}\in (C^\infty_c(B(0,R)))^4$$

*such that*

$$\sum_{|\alpha|\leq 5}\|\partial\partial^\alpha(\phi_0,\psi_0)\|_{L^2(\mathbb R^3)\times L^2(\mathbb R^3)}+\|\partial^\alpha(\phi_1,\psi_1)\|_{L^2(\mathbb R^3)\times L^2(\mathbb R^3)}<\epsilon.$$

*Then for every $R>0$, there exists $\epsilon_0=\epsilon_0(R)>0$ sufficiently small such that if $\epsilon\leq\epsilon_0$, the unique solution remains smooth for all time.*

*Proof.* We begin with the bootstrap assumption

$$\sup_{t\in[0,T]}\sum_{|\alpha|\leq 5}\|\partial\Gamma^\alpha\psi\|_{L^2(\mathbb R^3)}(t)+\sum_{|\alpha|\leq 5}\left(\int_0^T\int_{\mathbb R^3}\frac{|\bar\partial\Gamma^\alpha\psi|^2}{(1+|t-|x||)^{1+\delta}}dxdt\right)^{\frac12}\leq \epsilon^{\frac34}(1+T)^{\frac{1}{10}}. \tag{12.2}$$

This implies decay estimates for $\psi$ which are slightly worse than the optimal ones. More precisely, by Theorem 9.4, we have

$$\sup_{t\in[0,T]}\sum_{|\alpha|\leq 3}\|(1+t)(1+|t-|x|)^{\frac12}\partial\Gamma^\alpha\psi\|_{L^\infty(\mathbb R^3)}\leq C\epsilon^{\frac34}(1+t)^{\frac{1}{10}}. \tag{12.3}$$

Using (9.1) together with (12.3), we get

$$\sup_{t\in[0,T]}\sum_{|\alpha|\leq 2}(1+t)^{\frac32}\|\bar\partial\Gamma^\alpha\psi\|_{L^\infty(\mathbb R^3)}\leq C\epsilon^{\frac34}(1+t)^{\frac{1}{10}}. \tag{12.4}$$

The key observation is that when we proved the boundedness of energy for the wave map equation, we did not need to use the full strength of the decay estimates. Morally, the presence of the null structure allows use to gain an extra decay rate. We now control

$$\sum_{|\alpha|\leq 5}\int_0^T\|\Gamma^\alpha Q_0(\psi,\psi)\|_{L^2(\mathbb R^3)}dt$$

as in the proof of Theorem 11.6. We have two terms, the first is when the factor with fewer $\Gamma$'s comes with a good derivative, while the second term is such that the factor with more $\Gamma$'s comes with a good derivative. The first can be bounded by

$$\int_0^T(\sum_{|\alpha|\leq 5}\|\partial\Gamma^\alpha\psi\|_{L^2(\mathbb R^3)})(\sum_{|\beta|\leq 2}\|\bar\partial\Gamma^\beta\psi\|_{L^\infty(\mathbb R^3)})(t)dt\leq C\int_0^T\frac{\epsilon^{\frac32}dt}{(1+t)^{\frac32-\frac15}}\leq C\epsilon^{\frac32}.$$

We now control the second term. Let $T_0 = 0$, $T_i = 2^{i-1}$, for $i = 1, ..., \lfloor \log_2 T \rfloor$, $T_{\lfloor \log_2 T \rfloor + 1} = T$.

$$\begin{aligned}
&\int_0^T \Big( \sum_{|\alpha|\leq 5,\, |\beta|\leq 2} \|\overline{\partial}\Gamma^\alpha\psi \partial\Gamma^\beta\psi\|_{L^2(\mathbb{R}^3)}\Big)(t)dt \\
\leq &C \sum_{i=0}^{\lfloor \log_2 T\rfloor} \sum_{|\alpha|\leq 5,\, |\beta|\leq 2} \int_{T_i}^{T_{i+1}} \left( \int_{\mathbb{R}^3} \frac{|\overline{\partial}\Gamma^\alpha\psi|^2}{(1+|t-|x||)^{1+\delta}} dx\right)^{\frac12} \left( \sup_x \left( (1+|t-|x|)^{\frac{1+\delta}{2}} |\partial\Gamma^\beta\psi|(t,x)\right)\right) dt \\
\leq &C \sum_{i=0}^{\lfloor \log_2 T\rfloor} \Big( \sum_{|\alpha|\leq 5} \int_{T_i}^{T_{i+1}} \int_{\mathbb{R}^3} \frac{|\overline{\partial}\Gamma^\alpha\psi|^2}{(1+|t-|x||)^{1+\delta}} dxdt\Big)^{\frac12} \Big( \sum_{|\beta|\leq 2} \left( \int_{T_i}^{T_{i+1}} \left( \sup_x \left( (1+|t-|x|)^{\frac{1+\delta}{2}} |\partial\Gamma^\beta\psi|(t,x)\right)\right)^2 dt\right)^{\frac12} \\
\leq &C\epsilon^{\frac32} \sum_{i=0}^{\lfloor \log_2 T\rfloor} (1+T_i)^{\frac{1}{10}} \Big( \int_{T_i}^{T_{i+1}} \frac{dt}{(1+t)^{2-\frac15-\delta}}\Big)^{\frac12} \leq C\epsilon^{\frac32} \sum_{i=0}^{\lfloor \log_2 T\rfloor} (1+T_i)^{-\frac12+\frac15+\frac{\delta}{2}} \leq C\epsilon^{\frac32},
\end{aligned}$$

as long as $\delta \in (0, \frac35)$. Using energy estimates for $\phi$, we can then conclude that

$$\sup_{t\in[0,T]} \sum_{|\alpha|\leq 5} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)}(t) \leq C\epsilon \tag{12.5}$$

and also

$$\sup_{x,t} \sum_{|\alpha|\leq 3} (1+t+|x|)(1+|t-|x||)^{\frac12} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)}(t) \leq C\epsilon \tag{12.6}$$

via Klainerman–Sobolev inequality. We now use (12.5) and (12.6) for the energy estimates for $\psi$ to get

$$\begin{aligned}
&\sup_{t\in[0,T]} \sum_{|\alpha|\leq 5} \|\partial\Gamma^\alpha\psi\|_{L^2(\mathbb{R}^3)}(t) + \sum_{|\alpha|\leq 5} \left( \int_0^T \int_{\mathbb{R}^3} \frac{|\overline{\partial}\Gamma^\alpha\psi|^2}{(1+|t-|x||)^{1+\delta}} dxdt\right)^{\frac12} \\
\leq &C \left( \epsilon + \int_0^T \left( \sum_{|\alpha|\leq 5} \|\partial\Gamma^\alpha\phi\|_{L^2(\mathbb{R}^3)}(t)\right) \left( \sum_{|\beta|\leq 2} \|\partial\Gamma^\beta\phi)\|_{L^\infty(\mathbb{R}^3)}(t)\right) dt\right) \\
\leq &C \left( \epsilon + \epsilon^2 \int_0^T \frac{dt}{1+t}\right) \leq C\epsilon \log(2+T).
\end{aligned}$$

This improves over (12.2) and we are done. □

## 13. Another decay estimate

By now we know that decay estimates are very important for understanding long time behaviour of solutions to nonlinear equations. Before we proceed further, we prove an additional decay estimate. This estimate is most useful when the energy grows but one still wants to obtain sharp pointwise decay estimate. For a simple application, see Corollary 13.2. This will also play an important role when we discuss the nonlinear stability of Minkowski space.

**Proposition 13.1.** *Consider the equation*

$$(m^{\alpha\beta} + H^{\alpha\beta})\partial^2_{\alpha\beta}\phi = F$$

*in $(3+1)$-dimensions, where $m$ is the Minkowski metric and $H^{\alpha\beta}$ is a symmetric 2-tensor such that*

$$\|H\|_{L^\infty(D_t)} < \frac14, \quad |H^{uu}| + |H^{uv}| + |H^{uA}| \leq \frac{1+|t-|x||}{10(1+t+|x|)},$$

*where $D_t := \{\frac{t}{2} \leq |x| \leq \frac{3t}{2}\}$. Then, we have the following decay estimates:*

$$\begin{aligned}
&\|(1+t)\partial\phi\|_{L^\infty(\mathbb{R}^3)}(t) \\
\leq &C \sup_{\tau\in[0,t]} \sum_{|\alpha|\leq 1} \|\Gamma^\alpha\phi\|_{L^\infty(\mathbb{R}^3)}(\tau) + C\int_0^t \left( (1+\tau)\|F\|_{L^\infty(D_\tau)} + \sum_{|\alpha|\leq 2} (1+\tau)^{-1} \|\Gamma^\alpha\phi\|_{L^\infty(D_\tau)}\right) d\tau.
\end{aligned}$$

*Proof.* Since we allow $\sum_{|\alpha|=1} \|\Gamma^\alpha\phi\|_{L^\infty(D_\tau)}$ on the right hand side, we can control $\overline{\partial}\phi$ on the left hand side by (9.1) in the proof of Corollary 9.8. It therefore suffices to control $\partial_u\phi$. Moreover, it suffices to control $\partial_u\phi$ in the region $\frac{t}{2} \leq |x| \leq \frac{3t}{2}$. We now use the equation to get

$$|4\partial_v\partial_u(r\phi) + rH^{uu}\partial^2_{uu}\phi| \leq C(r(1+|H|)|\overline{\partial}\overline{\partial}\phi| + r|H^{uv}||\overline{\partial}\partial\phi| + r|H^{uA}||\overline{\partial}\partial\phi| + r|F|).$$

This implies

$$|(4\partial_v + H^{uu}\partial_u)\partial_u(r\phi)| \leq C(\sum_{|\alpha|\leq 2}\frac{|\Gamma^\alpha\phi|}{1+t} + |H^{uu}||\partial_u\phi| + r|F|) \leq C(\sum_{|\alpha|\leq 2}\frac{|\Gamma^\alpha\phi|}{1+t} + (1+t)|F|).$$

Given a point such that $\frac{t}{2} \leq |x| \leq \frac{3t}{2}$, we integrate the estimate above along the the integral curves of $4\partial_v + H^{uu}\partial_u$ towards the past until it hits the boundary of the set $\{\frac{t}{2} \leq |x| \leq \frac{3t}{2}\}$. This gives the desired bound. □

One immediate consequence of the estimate above is the following corollary:

**Corollary 13.2.** *Let $F$ be a smooth function supported in $B(0, t+1)$ for every $t \geq 0$ such that*

$$\sum_{|\alpha|\leq 5}(1+T)^{-\frac{1}{10}}\int_0^T \|\Gamma^\alpha F\|_{L^2(\mathbb{R}^3)}(t)dt + \sup_{t,x}(1+t)^2\log^2(2+t)|F(t,x)| \leq C'.$$

*Suppose $\phi$ is a solution to $\Box\phi = F$ with compactly supported initial data, then*

$$(1+t)\|\partial\phi\|_{L^\infty(\mathbb{R}^3)}(t) \leq C,$$

*for some $C > 0$ depending on initial data and $C'$.*

*Proof.* **Exercise.** □