Lectures on Hyperbolic Equations

Qian Wang

Mathematical Institute, University of Oxford

March 12, 2014

Contents

Part 1. Lectures on conservation laws

References

- L. C. Evans, Partial Differential Equations, 2010.
- L. Hörmander, *Lectures on Nonlinear Hyperbolic Differential Equations*, 1997.
- J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, 1994.
- C. D. Sogge, *Lectures on Nonlinear Wave Equations*, 1995.

- In this part we consider the mathematics of conservation laws.
- Conservation laws typically assert that the rate of change within a region is governed by a flux function controlling the rate of loss/increase through the boundary of the region.
- Let

$$u = u(x,t) = (u_1(x,t), \cdots, u_n(x,t)), \quad x \in \mathbb{R}^n,\ t \geq 0$$

be a vector function whose components are conserved in some physical system under investigation. Let $f : \mathbb{R}^m \to \mathbb{R}^{m \times n}$ be the flux function. Then the conservation law states

$$\frac{d}{dt}\int_{\Omega} u dx = -\int_{\partial\Omega} f(u)\nu dS$$

for any smooth bounded domain $\Omega \subset \mathbb{R}^n$, where ν denotes the outward unit normal to $\partial\Omega$.

- By the divergence theorem we have

$$\int_\Omega u_t dx = -\int_\Omega \operatorname{div} f(u) dx.$$

- Since Ω is arbitrary, we have

$$u_t + \operatorname{div} f(u) = 0 \quad \text{on } \mathbb{R}^n \times (0, \infty) \tag{1}$$

- This covers many equations from applications, including the Euler's equations for compressible gas flow.
- In this course we only consider the scalar case of (1) in one dimension, i.e. u is a scalar function of single variables, together with the initial condition $u(x, 0) = u_0(x)$, $x \in \mathbb{R}$.

1. The method of characteristics

We develop the method of characteristics to solve the nonlinear first order PDE

$$F(x, u, Du) = 0 \quad \text{in } U, \qquad u = g \quad \text{on } \Gamma, \tag{2}$$

where $U \subset \mathbb{R}^n$ is an open set, $x \in U$, $\Gamma \subset \partial U$, $g : \Gamma \to \mathbb{R}$ and $F : \overline{U} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ are given smooth functions. Writing

$$F = F(x, z, \mathbf{p}) = F(x_1, \cdots, x_n, z, p_1, \cdots, p_n),$$

we use the notation

$$D_x F = (F_{x_1}, \cdots, F_{x_n}), \quad D_z F = F_z, \quad D_{\mathbf{p}} F = (F_{p_1}, \cdots, F_{p_n}).$$

The basic idea of the method is as follows:

- Given $x \in U$, find a curve within U connecting x with a point $x_0 \in \Gamma$.
- Determine u along this curve.
- This usually requires the knowledge of Du along this curve.
- Let $x(s)$ be such a curve and set

  $$z(s) = u(x(s)) \quad \text{and} \quad \mathbf{p}(s) = Du(x(s)).$$

  Then $x(s), z(s), \mathbf{p}(s)$ are determined by solving systems of ODEs.

So, the key point is to derive the ODEs governing $x(s), z(s), \mathbf{p}(s)$.

To derive these equations, first

$$\frac{dz}{ds} = \sum_{j=1}^{n} u_{x_j}(x(s))\frac{dx_j}{ds}, \qquad \frac{dp_i}{ds} = \sum_{j=1}^{n} u_{x_i x_j}(x(s))\frac{dx_j}{ds}.$$

In order to eliminate the second derivative $u_{x_i x_j}$, we differentiating the PDE in (2) with respect to x_j to get

$$F_{x_j} + F_z u_{x_j} + \sum_{i=1}^{n} F_{p_i} u_{x_i x_j} = 0.$$

Restricting this equation to the curve $x(s)$, we obtain

$$F_{x_j}(x, z, \mathbf{p}) + F_z(x, z, \mathbf{p})p_j + \sum_{i=1}^{n} F_{p_i}(x, z, \mathbf{p})u_{x_i x_j}(x(s)) = 0.$$

Thus, if we set

$$\frac{dx_i}{ds} = F_{p_i}(x, z, \mathbf{p}),$$

then

$$\frac{dp_i}{ds} = -F_{x_i}(x, z, \mathbf{p}) - F_z(x, z, \mathbf{p})p_i, \quad \frac{dz}{ds} = \sum_{i=1}^{n} p_i F_{p_i}(x, z, \mathbf{p}).$$

We therefore obtain the system of ODEs

$$\begin{cases} \frac{dx}{ds} &= D_{\mathbf{p}} F(x, z, \mathbf{p}), \\ \frac{dz}{ds} &= \mathbf{p} \cdot D_{\mathbf{p}} F(x, z, \mathbf{p}), \\ \frac{d\mathbf{p}}{ds} &= -D_x F(x, z, \mathbf{p}) - D_z F(x, z, \mathbf{p})\mathbf{p}. \end{cases} \tag{3}$$

which is called the characteristic ODEs for (2)

- We still need to determine appropriate initial conditions for the characteristic ODEs (3) using $u = g$ on Γ.
- We use local parametrizations of Γ. Let Γ be locally parametrized by

$$x_i = x_i(\theta_1, \cdots, \theta_{n-1}), \quad i = 1, \cdots, n$$

with parameters $\theta_1, \cdots, \theta_{n-1}$. We will write $x = x(\theta)$ for short.
- Let $x^0 := x(\theta^0)$ be a point on Γ. For the ODEs in (3) it is natural to set $x(0) = x^0$ and $z(0) = z^0 := g(x^0)$. We need to determine $\mathbf{p}(0) = \mathbf{p}^0 := (p_1^0, \cdots, p_n^0)$.
- By the PDE in (2) we have $F(x^0, z^0, \mathbf{p}^0) = 0$.

- Using $u = g$ on Γ, we have $u(x(\theta)) = \tilde{g}(\theta) := g(x(\theta))$. Differentiating with respect to θ_j gives

$$\sum_{i=1}^{n} u_{x_i}(x(\theta)) \frac{\partial x_i}{\partial \theta_j} = \tilde{g}_{\theta_j}(\theta), \quad j = 1, \cdots, n-1.$$

By setting $\theta = \theta^0$ we obtain n equations on $\mathbf{p}^0$:

$$\begin{aligned} &\sum_{i=1}^{n} p_i^0 \frac{\partial x_i}{\partial \theta_j}(\theta^0) = \tilde{g}_{\theta_j}(\theta^0), \quad j = 1, \cdots, n-1, \\ &F(x^0, z^0, \mathbf{p}^0) = 0. \end{aligned} \tag{4}$$

In many situations, $\mathbf{p}^0$ can be obtained by solving (4).

Example 1

Consider the problem

$$uu_x + u_y = 2, \qquad u(x,x) = x.$$

Here $F = F(x, y, z, p_1, p_2) = zp_1 + p_2 - 2$. Since $F_x = F_y = 0$, $F_z = p_1$, $F_{p_1} = z$, and $F_{p_2} = 1$, it follows from the characteristic ODEs (3) that

$$\frac{dx}{ds} = z, \quad \frac{dy}{ds} = 1, \quad \frac{dz}{ds} = p_1 z + p_2.$$

Recall that $z = u(x,y)$, $p_1 = u_x(x,y)$ and $p_2 = u_y(x,y)$, we have

$$\frac{dz}{ds} = 2.$$

To include the boundary condition $u(x,x)=x$, we fix any τ, let $(x(s),y(s))$ be the characteristic curve with

$$(x(0),y(0))=(\tau,\tau).$$

Then $z(0)=\tau$ and thus

$$\begin{cases} \frac{dx}{ds}=z, & x(0)=\tau, \\ \frac{dy}{ds}=1, & y(0)=\tau, \\ \frac{dz}{ds}=2, & z(0)=\tau. \end{cases}$$

Solving these equations give

$$y(s)=s+\tau, \qquad z(s)=2s+\tau, \qquad x(s)=s^2+\tau s+\tau.$$

Now for any (x, y) we determine s and τ such that $(x, y) = (x(s), y(s))$. It yields

$$s = \frac{y - x}{1 - y} \quad \text{and} \quad \tau = \frac{x - y^2}{1 - y}.$$

Therefore

$$u(x, y) = u(x(s), y(s)) = z(s) = 2s + \tau = \frac{2y - y^2 - x}{1 - y}.$$

This solution makes sense only if $y \neq 1$. ∎

When the PDE in (2) has special structures, the characteristic ODEs can be significantly simplified.

- Consider the first order linear PDE

$$\mathbf{b}(x) \cdot Du(x) + c(x)u(x) = 0.$$

Here $F(x, z, \mathbf{p}) = \mathbf{b}(x) \cdot \mathbf{p} + c(x)z$. Since $D_{\mathbf{p}}F = \mathbf{b}(x)$, we have

$$\frac{dx}{ds} = \mathbf{b}(x), \qquad \frac{dz}{ds} = \mathbf{b}(x) \cdot \mathbf{p}(s).$$

Since $\mathbf{p}(s) = Du(x(s)) = -c(x(s))u(x(s)) = -c(x(s))z(s)$, we obtain the simplified characteristic ODEs

$$\frac{dx}{ds} = \mathbf{b}(x), \qquad \frac{dz}{ds} = -c(x)z.$$

The equations on $\mathbf{p}$ are not needed. ∎

- Consider the scalar Hamilton-Jacobi equation

$$u_t + f(u_x) = 0,$$

where $f \in C^1(\mathbb{R})$. Here $F = F(t, x, z, q, p) = q + f(p)$ with $p = u_x$ and $q = u_t$. Consequently

$$F_q = 1, \quad F_p = f'(p), \quad F_t = F_x = F_z = 0.$$

Therefore, it follows from the characteristic ODEs (3) that

$$\begin{aligned} &\frac{dt}{ds} = 1, \qquad \frac{dx}{ds} = f'(p), \qquad \frac{dz}{ds} = q + pf'(p), \\ &\frac{dq}{ds} = 0, \qquad \frac{dp}{ds} = 0. \end{aligned}$$

Thus we may take $s = t$. Since $q = u_t = -f(u_x) = -f(p)$, we obtain the simplified characteristic ODEs

$$\begin{cases} \frac{dx}{dt} = f'(p), \\ \frac{dz}{dt} = pf'(p) - f(p), \\ \frac{dp}{dt} = 0. \end{cases}$$

These equations imply that

- p are constants along characteristics by the last equation .
- Characteristics are straight lines with velocity $f'(p)$ by the first equation.
- By the second equation, u can be obtained along characteristic lines.

We will use these facts to discuss Hamilton-Jacobi equation later. ■

- Consider the initial value problem of the scalar conservation law

$$
\begin{aligned}
&u_t + f(u)_x = 0, \qquad && (x,t) \in \mathbb{R} \times (0,\infty), \\
&u(x,0) = u_0(x), \qquad && x \in \mathbb{R},
\end{aligned} \tag{5}
$$

where f is a C^1 function. The equation can be write as $u_t + f'(u)u_x = 0$. Here $F = F(t,x,u,q,p) = q + f'(u)p$ with $q = u_t$ and $p = u_x$. Since

$$
F_t = F_x = 0, \quad F_q = 1, \quad F_p = f'(u), \quad qu + p = 0,
$$

from the characteristic ODEs (3) we have

$$
\frac{dt}{ds} = 1, \quad \frac{dx}{ds} = f'(u), \quad \frac{du}{ds} = q + pf'(u) = 0.
$$

We can take $s = t$. Thus for (5) the characteristic ODEs become

$$\begin{cases} \frac{dx}{dt} = f'(u), \\ \frac{du}{dt} = 0. \end{cases} \tag{6}$$

From these equation we can conclude

- u are constants along characteristics.
- Characteristics are straight lines with velocity $f'(u)$.

We will use these facts to show the following result.

Lemma 2

The problem (5) cannot have a C^1 solution defined for all $t > 0$ if there exist $x_1 < x_2$ such that $f'(u_0(x_2)) < f'(u_0(x_1))$.

Proof.

- Assume (5) has a C^1 solution defined for all $t > 0$.
- Then u are constants along characteristics and characteristics are straight lines. For characteristic line crossing x-axis at x, its velocity is $f'(u_0(x))$.
- Let l_1, l_2 be the two characteristics lines starting from $(x_1, 0)$ and $x_2, 0)$. Their velocities are $f'(u_0(x_1))$ and $f'(u_0(x_2))$ respectively.

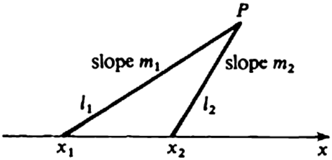


Figure: The plots of l_1 and l_2 whose slopes are $m_1 = 1/f'(u_0(x_1))$ and $m_2 = 1/f'(u_0(x_2))$ respectively,

- Since $f'(u_0(x_2)) < f'(u_0(x_1))$, these two lines must cross at some point P in $t > 0$.
- Along l_i we have $u(x_i, t) = u_0(x_i)$, $i = 1, 2$. Thus u must be discontinuous at P. Contradiction! ■

Conclusion:

- In general C^1 solutions of (5) can exits for only a finite time no matter how smooth u_0 is.
- In order to allow (5) to admit solutions defined for all $t > 0$, the notion of solution should be generalized to include solutions with "discontinuities".

2. Weak solutions and Rankine-Hugoniot condition

Consider again the initial value problem (5), i.e.

$$u_t + f(u)_x = 0, \qquad u(x,0) = u_0(x). \tag{7}$$

To motivate the notion of weak solution, assume u is a C^1 solution of (7). Multiplying (7) by any test function $\varphi \in C_0^\infty(\mathbb{R} \times [0,\infty))$, integrating over $\mathbb{R} \times (0,\infty)$, and using integration by parts, it gives

$$\begin{aligned} 0 &= \int_0^\infty \int_{-\infty}^\infty (u_t + f(u)_x)\varphi dxdt \\ &= \int_0^\infty \int_{-\infty}^\infty (u\varphi_t + f(u)\varphi_x)dxdt + \int_{-\infty}^\infty u_0(x)\varphi(x,0)dx. \end{aligned}$$

Since the last equation makes sense provided that u and u_0 are merely bounded and measurable, it leads to the following definition.

Definition 3

Let $u_0 \in L^\infty(\mathbb{R})$. A function $u \in L^\infty(\mathbb{R} \times (0, \infty))$ is called a weak solution of (7) if

$$\int_0^\infty \int_{-\infty}^\infty (u\varphi_t + f(u)\varphi_x) dxdt + \int_{-\infty}^\infty u_0(x)\varphi(x, 0)dx = 0$$

for all $\varphi \in C_0^\infty(\mathbb{R} \times [0, \infty))$.

Remarks.

(i) If $u \in C^1(\mathbb{R} \times [0, \infty))$ is a classical solution of (7), then u is automatically a weak solution.

(ii) If u is a weak solution of (7) and if u is C^1 in a domain $\Omega \subset \mathbb{R} \times (0, \infty)$, then $u_t + f(u)_x = 0$ in Ω. In fact, for any $\varphi \in C_0^1(\Omega)$ we have by integration by parts that

$$0 = \int_0^\infty \int_{-\infty}^\infty (u\varphi_t + f(u)\varphi_x) dxdt = \int_0^\infty \int_{-\infty}^\infty (u_t + f(u)_x)\varphi dxdt.$$

Since φ is arbitrary, it follows $u_t + f(u)_x = 0$ in Ω.

(iii) If $u_0 \in C(\mathbb{R})$ and $u \in C^1(\mathbb{R} \times [0, \infty))$ is a weak solution of (7), then u is a classical solution. In fact, $u_t + f(u)_x = 0$ in $\mathbb{R} \times (0, \infty)$ by (ii). Thus, by the definition of weak solution and integration by parts, we have

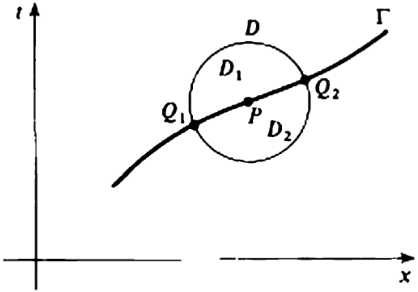
$$0 = \int_{-\infty}^\infty (u(x, 0) - u_0(x))\varphi(x, 0) dx, \quad \forall \varphi \in C_0^1(\mathbb{R} \times [0, \infty)).$$

Therefore $u(x,0) = u_0(x)$ for $x \in \mathbb{R}$. ■

The notion of weak solution places restrictions on the curve of discontinuity.

- Let Γ be a smooth curve across which u has a jump discontinuity, and u is smooth away from Γ.
- Let $P \in \Gamma$ and let D be a small ball in $t > 0$ centered at P. Assume that the part of Γ in D is given by $x = x(t)$, $a \leq t \leq b$.
- Γ splits D into two parts: the left part D_1 and the right part D_2. Let

$$u_l := \lim_{\varepsilon \searrow 0} u(x(t) - \varepsilon, t), \qquad u_r := \lim_{\varepsilon \searrow 0} u(x(t) + \varepsilon, t).$$



- For any $\varphi \in C_0^1(D)$, we have

$$0 = \iint\limits_{D} (u\varphi_t + f(u)\varphi_x)\,dxdt = \left(\iint\limits_{D_l} + \iint\limits_{D_r} \right) (u\varphi_t + f(u)\varphi_x)\,dxdt.$$

Since u is C^1 in D_1 and D_2, we have $u_t + f(u)_x = 0$ in D_1 and D_2. Therefore it follows from the divergence theorem that

$$\begin{aligned}\iint\limits_{D_1} (u\varphi_t + f(u)\varphi_x)dxdt &= \iint\limits_{D_1} ((u\varphi)_t + (f(u)\varphi)_x)dxdt \\ &= \int_{\partial D_1} \varphi(-udx + f(u)dt) \\ &= \int_{\Gamma} \varphi(-u_l dx + f(u_l)dt).\end{aligned}$$

Similarly,

$$\iint\limits_{D_2} (u\varphi_t + f(u)\varphi_x)dxdt = -\int_{\Gamma} \varphi(-u_r dx + f(u_r)dt).$$

Therefore

$$0 = \int_\Gamma \varphi(-[u]dx + [f(u)]dt),$$

where $[u] = u_l - u_r$ and $[f(u)] = f(u_l) - f(u_r)$ denote the jumps across Γ. Let $s := \frac{dx}{dt}$ denote the speed of the curve of discontinuities. Then

$$0 = \int_a^b \varphi(-s[u] + [f(u)])dt.$$

By the arbitrariness of φ, we can conclude that

$$s[u] = [f(u)] \tag{8}$$

at each point on Γ, which is called the *Rankine-Hugoniot condition*.

Proposition 4

If u is a weak solution of (7), then on the curves of discontinuity there must hold the Rankine-Hugoniot condition (8).

We give an example to indicate how to produce weak solutions by the method of characteristics and the Rankine-Hugoniot condition .

Example 5

Consider the initial value problem of Burgers equation

$$u_t + (u^2/2)_x = 0, \quad u(x,0) = u_0(x) := \begin{cases} 1, & x < 0, \\ 1 - x, & 0 \le x \le 1, \\ 0, & x > 1. \end{cases}$$

- We first use the method of characteristics to find the solution defined for a finite time.
- We know that all characteristics are straight lines and u are constants along characteristics lines.
- Since the flux is $f(u) = u^2/2$, the characteristic line crossing x-axis at x_0 is given by

$$x(t) = x_0 + tu_0(x_0), \qquad x_0 \in \mathbb{R}.$$

and on this line

$$u = u_0(x_0).$$

Since all characteristics starting at $(x_0, 0)$ with $0 \leq x_0 \leq 1$ cross at $(1, 1)$, $u(x, t)$ can not be smooth for $t \geq 1$.

- By the knowledge of characteristics, $u(x,t)$ for $t<1$ can be determined as follows:
  - $u(x,t)=1$ for $x<t$ and $u(x,t)=0$ for $x>1$.
  - For (x,t) satisfying $0<t\leq x\leq 1$, the characteristic through it intersects x-axis at $(x_0,0)$ with $x_0=(x-t)/(1-t)$. So

$$u(x,t)=u_0(x_0)=1-x_0=1-\frac{x-t}{1-t}=\frac{1-x}{1-t}.$$

- Therefore, for $t<1$ we have

$$u(x,t)=\begin{cases}1, & x<t,\\ (1-x)/(1-t), & t\leq x\leq 1,\\ 0, & x>1.\end{cases}$$

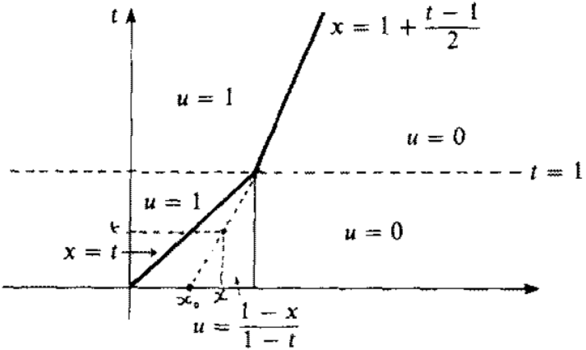
- Next we use the Rankine-Hugoniot condition to define $u(x,t)$ for $t \geq 1$.
  - By the knowledge of characteristics, a curve of discontinuities starting at the point $(1,1)$ is expected with $u = 1$ on the left and $u = 0$ on the right.
  - By the Rankine-Hugoniot condition, the speed of the curve of discontinuities is

$$s(t) = \frac{u_l^2/2 - u_r^2/2}{u_l - u_r} = \frac{1}{2}(u_l + u_r) = \frac{1}{2}.$$

  So the curve is given by $x(t) = 1 + (t-1)/2$, $t \geq 1$. Hence, for $t \geq 1$ we have

$$u(x,t) = \begin{cases} 1, & x < 1 + (t-1)/2, \\ 0, & x > 1 + (t-1)/2. \end{cases}$$

The solution u is depicted in the following figure.



- By definition it is easy to check that the above u is a weak solution. ∎

Example 6 (Nonuniqueness of weak solutions)

Consider the initial value problem of Burgers equation

$$u_t + (u^2/2)_x = 0, \qquad u(x,0) = \begin{cases} 0, & x < 0, \\ 1, & x > 0. \end{cases}$$

The method of characteristics determines the solution everywhere in $t > 0$ except in the sector $0 < x < t$. By defining u in $0 < x < t$ carefully, we obtain two functions

$$u_1(x,t) = \begin{cases} 0, & x < t/2, \\ 1, & x > t/2, \end{cases} \quad \text{and} \quad u_2(x,t) = \begin{cases} 0, & x < 0, \\ x/t, & 0 < x < t, \\ 1, & x > t; \end{cases}$$

both turn out to be weak solutions. ■

3. Entropy conditions

- Example shows that weak solutions of conservation laws are not necessarily unique.
- Criteria should be developed to pick out the "physically relevant" solution.
- Such a criterion is called an entropy condition.
- We motivate the entropy condition for the scalar conservation laws

$$u_t + f(u)_x = 0, \qquad u(x,0) = u_0(x), \tag{9}$$

where $u_0 \in C^1$ and f is C^2 with $f'' > 0$. Assume that (9) has a smooth solution u (thus $u_0' \geq 0$ by Lemma 2).

- Recall that all characteristics of (9) are straight lines given by

$$(x_0 + f'(u_0(x_0))t, t), \qquad x_0 \in \mathbb{R}.$$

- For any (x, t) with $t > 0$ let x_0 be the crossing point of x-axis and the characteristic through (x, t). Since $u(x, t) = u_0(x_0)$ along the characteristic, we have

$$x = x_0 + t\, f'(u(x, t)), \quad \text{i.e. } x_0 = x - t\, f'(u(x, t)).$$

  So u satisfies the equation $u = u_0(x - t\, f'(u))$.
- Taking derivative with respect to x gives

$$u_x(x, t) = \frac{u_0'(x - t\, f'(u))}{1 + u_0'(x - t\, f'(u))f''(u)t}.$$

- If $u_0'(x-tf'(u))=0$, then $u_x(x,t)=0$; If $u_0'(x-tf'(u))>0$, then

$$u_x(x,t)\le\frac{u_0'(x-t\,f'(u))}{u_0'(x-t\,f'(u))f''(u)t}=\frac{1}{f''(u)t}\le\frac{E}{t},$$

where $E=1/\inf\{f''(u):|u|\le\|u_0\|_\infty\}$, here we used $|u(x,t)|\le|u_0\|_\infty$.

- Consequently, we have for any $t>0$, $x\in\mathbb{R}$ and $a>0$ that

$$\frac{u(x+a,t)-u(x,t)}{a}\le\frac{E}{t}.$$

- This last inequality requires no smoothness of u and thus can be used as a criterion to pick out the "right" weak solution.

Definition 7 (Oleinik)

A weak solution u of the scalar conservation laws is said to satisfy the Oleinik entropy condition if there is a constant E such that

$$\frac{u(x+a,t)-u(x,t)}{a} \leq \frac{E}{t}$$

for all $t > 0$ and $x, a \in \mathbb{R}$ with $a > 0$.

We derive another entropy condition due to Lax which is easier to extend for systems of conservation laws.

- Recall that the characteristics are given by

$$(x_0 + f'(u_0(x_0))t, t), \qquad x_0 \in \mathbb{R}.$$

- Assume that, at some point on a curve C of discontinuities, u has distinct left and right limits u_l and u_r and that a characteristic from left and a characteristic from the right hit C at this point. Then

$$f'(u_l) > s > f'(u_r), \tag{10}$$

  where s denote the speed of the discontinuous curve at that point. We call (10) the Lax entropy condition.

Remark. In case $f'' > 0$, Lax entropy condition can be deduced from Oleinik entropy condition:

- Indeed, by Oleinik entropy condition we always have $u_l \geq u_r$ and thus $u_l > u_r$ on the curve of discontinuities.

- Since $f'' > 0$, f' is strictly increasing and thus $f'(u_l) > f'(u_r)$.
- By Rankine-Hugoniot condition, the speed of discontinuous curve is

$$s = \frac{f(u_l) - f(u_r)}{u_l - u_r} = f'(\xi)$$

for some $\xi \in (u_r, u_l)$. Consequently $f'(u_l) > s > f'(u_r)$ which is the Lax entropy condition.

Definition 8

A curve of discontinuity for u is called a shock curve provided both the Rankine-Hugoniot condition and the entropy condition hold.

Question: *Is it possible to show existence and uniqueness of weak solutions of conservation laws satisfying suitable entropy condition?* We will focus on scalar conservation laws with strictly convex flux.

4. Uniqueness of entropy solutions

We will prove the following uniqueness result.

Theorem 9

Consider the initial value problem of the scalar conservation laws

$$\begin{cases} u_t + f(u)_x = 0, & x \in \mathbb{R}, t > 0, \\ u(x,0) = u_0(x), & x \in \mathbb{R}, \end{cases}$$

where f is a C^2 convex function. If $u, v \in L^\infty(\mathbb{R} \times (0,\infty))$ are two weak solutions satisfying the Oleinik entropy condition, then

$$u = v \quad \text{in } \mathbb{R} \times (0,\infty)$$

except a set of measure zero.

Proof. Since $u, v \in L^\infty(\mathbb{R} \times (0,\infty))$, it suffices to show that

$$\int_0^\infty \int_{-\infty}^\infty (u-v)\varphi dxdt = 0, \quad \forall \varphi \in C_0^1(\mathbb{R} \times (0,\infty)). \tag{11}$$

By the definition of weak solution, for any $\psi \in C_0^1(\mathbb{R} \times [0,\infty))$ we have

$$\int_0^\infty \int_{-\infty}^\infty (u\psi_t + f(u)\psi_x)\, dxdt + \int_{-\infty}^\infty u_0(x)\psi(x,0)dx = 0,$$
$$\int_0^\infty \int_{-\infty}^\infty (v\psi_t + f(v)\psi_x)\, dxdt + \int_{-\infty}^\infty u_0(x)\psi(x,0)dx = 0.$$

Therefore

$$0 = \int_0^\infty \int_{-\infty}^\infty \{(u-v)\psi_t + (f(u)-f(v))\psi_x\}\, dxdt.$$

By writing

$$f(u) - f(v) = \int_0^1 \frac{d}{d\tau}[f(\tau u + (1-\tau)v)]d\tau = b(u-v),$$

where

$$b(x,t) := \int_0^1 f'(\tau u(x,t) + (1-\tau)v(x,t))d\tau,$$

then it follows

$$0 = \int_0^\infty \int_{-\infty}^\infty (u-v)\left(\psi_t + b\psi_x\right) dxdt \tag{12}$$

for all $\psi \in C_0^1(\mathbb{R} \times [0,\infty))$.

- If we could solve the linear transport equation

$$\psi_t + b\psi_x = \varphi \tag{13}$$

  for any $\varphi \in C_0^1(\mathbb{R} \times (0, \infty))$ to obtain $\psi \in C_0^1(\mathbb{R} \times [0, \infty))$, then we would obtain (11) from (12).
- Unfortunately, (13) may not have a C_0^1 solution ψ because b is not continuous in general.
- To get around this difficulty, we need to use the mollification technique.
- We take a mollifier, i.e. a function $\omega \in C_0^\infty(\mathbb{R}^2)$ with

$$\omega \geq 0, \quad \iint_{\mathbb{R}^2} \omega(x, t)dxdt = 1, \quad \text{supp}(\omega) \subset \{x^2 + t^2 \leq 1\}.$$

- For any $\varepsilon > 0$ set $\omega_\varepsilon(x,t) = \varepsilon^{-2}\omega(x/\varepsilon, t/\varepsilon)$.
- To regularize u and v, we set $u(x,t) = v(x,t) = 0$ for $t < 0$ and define

$$u_\varepsilon = u * \omega_\varepsilon, \qquad v_\varepsilon = v * \omega_\varepsilon$$

where $*$ denotes the convolution, i.e.

$$u * \omega_\epsilon(x,t) = \iint_{\mathbb{R}^2} u(y,s)\omega_\varepsilon(x-y, t-s)dydt.$$

It is well known that both u_ε and v_ε are smooth functions and

$$|u_\varepsilon| \leq M \quad \text{and} \quad |v_\varepsilon| \leq M, \quad \text{in } \mathbb{R} \times [0,\infty), \qquad (14)$$

where $M > 0$ is a constant such that $|u|, |v| \leq M$.

- We use the Oleinik entropy condition to show for $\alpha > 0$ that

$$\partial_x u_\varepsilon \le E/\alpha \quad \text{and} \quad \partial_x v_\varepsilon \le E/\alpha, \qquad \forall t \ge \alpha. \tag{15}$$

Let $h(x,t) := u(x,t) - Ex/\alpha$. Then for $a \ge 0$ and $t \ge \alpha$

$$h(x+a,t)-h(x,t) = u(x+a,t)-u(x,t)-\frac{Ea}{\alpha} \le \frac{Ea}{t}-\frac{Ea}{\alpha} \le 0.$$

Thus $x \to (h*\omega_\varepsilon)(x,t)$ is decreasing for each $t \ge \alpha$. Since

$$(h*\omega_\varepsilon)(x,t) = u_\varepsilon(x,t) - \frac{Ex}{\alpha} + \frac{E}{\alpha}\iint_{\mathbb{R}^2} y\omega_\varepsilon(y,s)dyds,$$

we obtain

$$0 \ge \partial_x(h*\omega_\varepsilon) = \partial_x u_\varepsilon - E/\alpha, \quad \forall t \ge \alpha.$$

- Next define

$$b_\varepsilon := \int_0^1 f'(\tau u_\varepsilon + (1-\tau)v_\varepsilon)d\tau.$$

Because of (14) and $f \in C^2$, we have $b_\varepsilon \in C^1$ and there is a constant M_1 independent of ε such that

$$|b_\varepsilon(x,t)| \le M_1, \quad (x,t) \in \mathbb{R} \times [0,\infty). \tag{16}$$

- Moreover, for any $\alpha > 0$ there holds

$$\partial_x b_\varepsilon \le C_0 E/\alpha, \qquad \forall t \ge \alpha, \tag{17}$$

where $C_0 := \max\{f''(\xi) : |\xi| \le M\}$. In fact,

$$\partial_x b_\varepsilon = \int_0^1 f''(\tau u_\varepsilon + (1-\tau)v_\varepsilon)\,(\tau\partial_x u_\varepsilon + (1-\tau)\partial_x v_\varepsilon)\,d\tau.$$

Since $f'' \geq 0$, we may use (15) and (14) to derive for $t \geq \alpha$ that

$$\partial_x b_\varepsilon \leq \frac{E}{\alpha} \int_0^1 f''(\tau u_\varepsilon + (1-\tau)v_\varepsilon) d\tau \leq \frac{C_0 E}{\alpha}.$$

- We next prove that $b_\varepsilon \to b$ locally in L^1 as $\varepsilon \to 0$. To see this, using $f \in C^2$ we can write

$$\begin{aligned} & b_\varepsilon(x,t) - b(x,t) \\ & = \int_0^1 \left(f'(\tau u_\varepsilon + (1-\tau)v_\varepsilon) - f'(\tau u + (1-\tau)v) \right) d\tau \\ & = \int_0^1 f''(\xi) \left(\tau(u_\varepsilon - u) + (1-\tau)(v_\varepsilon - v) \right) d\tau, \end{aligned}$$

where ξ is between $\tau u_\varepsilon + (1-\tau)v_\varepsilon$ and $\tau u + (1-\tau)v$.

By (14) we have $|\xi| \leq M$. Therefore

$$|b_\varepsilon(x,t) - b(x,t)| \leq \frac{1}{2} C_0 \left(|u_\varepsilon - u| + |v_\varepsilon - v| \right).$$

Thus for any compact set $K \subset \mathbb{R} \times [0, \infty)$ we have

$$\iint_K |b_\varepsilon - b| dxdt \leq \frac{1}{2} C_0 \iint_K \left(|u_\varepsilon - u| + |v_\varepsilon - v| \right) dxdt$$
$$\to 0 \quad \text{as } \varepsilon \to 0.$$

- For any fixed $\varphi \in C_0^1(\mathbb{R} \times (0, \infty))$, we consider the problem

$$\psi_t^\varepsilon + b_\varepsilon \psi_x^\varepsilon = \varphi, \quad \psi^\varepsilon(x, T) = 0, \tag{18}$$

where $T > 0$ is chosen such that $\varphi = 0$ for $t \geq T$.

By the method of characteristics, the solution of (18) is given by

$$\psi^{\varepsilon}(x,t) = \int_T^t \varphi(x_\varepsilon(s;x,t),s)ds, \qquad (19)$$

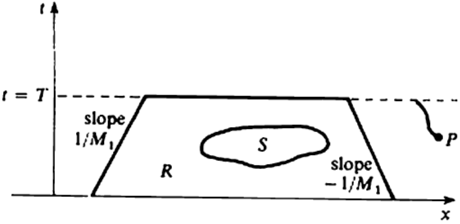
where $x_\varepsilon(s) := x_\varepsilon(s;x,t)$ is defined by

$$\frac{dx_\varepsilon}{ds} = b_\varepsilon(x_\varepsilon,s), \quad x_\varepsilon(t) = x.$$

Since $b_\varepsilon \in C^1$ satisfies (16), x_ε exists for all s and is C^1 with respect to s, x and t. Thus $\psi^\varepsilon \in C^1(\mathbb{R}\times[0,\infty))$.

- We show that $\psi^\varepsilon \in C_0^1(\mathbb{R}\times[0,\infty))$ and $\mathrm{supp}(\psi^\varepsilon)$ are contained in a compact region independent of ε.

To see this, let $S := \text{supp}(\varphi)$. By the choice of T, S is a compact set contained in $\{(x,t) : 0 < t \leq T\}$. In view of (19), $\psi^\varepsilon(x,t) = 0$ for $t \geq T$.



Next let R be the region bounded by the lines $t = 0$, $t = T$ and two lines with slopes $1/M_1$ and $-1/M_1$ such that $S \subset R$. For any $(x,t) \notin R$ with $t < T$, from (16) it follows that

$$x_\varepsilon(s; x, t) \notin R, \qquad \forall t \leq s \leq T$$

Since

$$\frac{d}{ds}\psi^\varepsilon(x_\varepsilon(s;x,t),s) = \psi_s^\varepsilon + \psi_x^\varepsilon \frac{\partial x_\varepsilon}{\partial s} = \psi_s^\varepsilon + b_\varepsilon \psi_x^\varepsilon$$
$$= \varphi(x_\varepsilon(s;x,t),s) = 0$$

for $t \le s \le T$, we have

$$\psi^\varepsilon(x,t) = \psi^\varepsilon(x_\varepsilon(t;x,t),t) = \psi^\varepsilon(x_\varepsilon(T;x,t),T) = 0.$$

Therefore $\mathrm{supp}(\psi^\varepsilon) \subset R$.

- By using (12) with $\psi = \psi^\varepsilon$ and (18) we have

$$0 = \int_0^\infty \int_{-\infty}^\infty (u-v)\left\{\psi_t^\varepsilon + b_\varepsilon \psi_x^\varepsilon + (b - b_\varepsilon)\psi_x^\varepsilon\right\} dxdt.$$

In view of (18) it follows

$$\int_0^\infty \int_{-\infty}^\infty (u - v)\varphi dxdt = \int_0^\infty \int_{-\infty}^\infty (u - v)(b_\varepsilon - b)\psi_x^\varepsilon dxdt. \tag{20}$$

To prove (11), it suffices to show that the right hand side of (20) goes to 0 as $\varepsilon \to 0$.

- We need to estimate ψ_x^ε. We first show that for any $\alpha > 0$ there exists C_α such that

$$|\psi_x^\varepsilon| \leq C_\alpha, \qquad \forall t \geq \alpha. \tag{21}$$

Since $\psi^\varepsilon = 0$ for $t \geq T$, it suffices to show (21) for $\alpha \leq t < T$.

By using (19) we obtain

$$\psi_x^\varepsilon(x,t) = \int_T^t \varphi_x(x_\varepsilon(s,x,t),s)\frac{\partial x_\varepsilon}{\partial x}(s;x,t)ds. \tag{22}$$

Recall that $x_\varepsilon(t;x;t) = x$, we have $\frac{\partial x_\varepsilon}{\partial x}(t;x,t) = 1$. Let

$$a_\varepsilon(s) := \frac{\partial x_\varepsilon}{\partial x}(s;x,t).$$

Then $a_\varepsilon(t) = 1$ and

$$\begin{aligned}\frac{\partial a_\varepsilon}{\partial s} = \frac{\partial}{\partial s}\frac{\partial x_\varepsilon}{\partial x} = \frac{\partial}{\partial x}\frac{\partial x_\varepsilon}{\partial s} = \frac{\partial}{\partial x}b_\varepsilon(x_\varepsilon(s;x,t),s)\\ = \partial_x b_\varepsilon \frac{\partial x_\varepsilon}{\partial x} = (\partial_x b_\varepsilon)a_\varepsilon\end{aligned}$$

Therefore

$$a_\varepsilon(s) = \exp\left(\int_t^s \partial_x b_\varepsilon(x_\varepsilon(\tau; x, t), \tau) d\tau\right).$$

In view of (17), it follows $a_\varepsilon(s) \leq e^{C_0 ET/\alpha}$ for $\alpha \leq t \leq s \leq T$. Thus we have from (22) that

$$|\psi_x^\varepsilon(x, t)| \leq \int_t^T |\varphi_x| a_\varepsilon(s) ds \leq C_\alpha, \qquad \forall \alpha \leq t \leq T.$$

- We next derive the total variation estimate on ψ^ε: For each $t > 0$ let

$$TV_t(\psi^\varepsilon) := \int_{-\infty}^{\infty} |\psi_x^\varepsilon(x, t)|\, dx$$

denote the total variation of the function $\psi^\varepsilon(\cdot, t)$.

Since the supports of ψ^ε) are contained in a compact region independent of ε, it follows from (21) that for any $\alpha > 0$ there is a constant $\hat{C}_\alpha$ independent of ε such that

$$TV_t(\psi^\varepsilon) \leq \hat{C}_\alpha, \qquad \forall t \geq \alpha.$$

We claim that

$$\exists \beta > 0 \text{ such that } TV_t(\psi^\varepsilon) \leq \hat{C}_\beta \text{ for all } 0 < t \leq \beta. \tag{23}$$

To see this, by using $\varphi \in C_0^1(\mathbb{R} \times (0, \infty))$ we may take $\beta > 0$ such that $\varphi = 0$ for $0 \leq t \leq \beta$. It then follows from (18) that

$$\psi_t^\varepsilon + b_\varepsilon \psi_x^\varepsilon = 0 \quad \text{ for } t \leq \beta. \tag{24}$$

Fix $0 \leq t \leq \beta$, let $x_0 < x_1 < \cdots < x_N$ be any partition of $\mathbb{R}$, and set $y_i = x_\varepsilon(\beta; x_i, t)$ for $i = 0, \cdots, N$. Then $y_0 < y_1 < \cdots < y_N$. Since (24) implies that ψ^ε is constant along the characteristic curves $s \to x_\varepsilon(s; x_i, t)$ for $0 \leq s \leq \beta$, we have

$$\psi^\varepsilon(x_i, t) = \psi^\varepsilon(y_i, \beta), \quad i = 0, 1, \cdots, N.$$

Therefore

$$\begin{aligned}\sum_{i=0}^{N-1} |\psi^\varepsilon(x_{i+1}, t) - \psi^\varepsilon(x_i, t)| &\leq \sum_{i=0}^{N-1} |\psi^\varepsilon(y_{i+1}, \beta) - \psi^\varepsilon(y_i, \beta)| \\ &\leq TV_\beta(\psi^\varepsilon).\end{aligned}$$

Taking the supremum over all such partitions gives $TV_t(\psi^\varepsilon) \leq TV_\beta(\psi^\varepsilon) \leq \hat{C}_\beta$.

- Finally we complete the proof by estimating

$$\left|\int_0^\infty \int_{-\infty}^\infty (u-v)(b_\varepsilon - b)\psi_x^\varepsilon dxdt\right| \le I_1 + I_2,$$

where

$$I_1 = \int_0^\alpha \int_{-\infty}^\infty |u-v||b_\varepsilon - b||\psi_x^\varepsilon| dxdt,$$
$$I_2 = \int_\alpha^\infty \int_{-\infty}^\infty |u-v||b_\varepsilon - b||\psi_x^\varepsilon| dxdt.$$

By using (16) and (23) we obtain for $0 < \alpha \le \beta$ that

$$I_1 \le 2M \cdot 2M_1 \int_0^\alpha TV_t(\psi^\varepsilon)dt \le 4MM_1\alpha\hat{C}_\beta.$$

Thus, for any $\eta > 0$ we can take $0 < \alpha \leq \beta$ such that

$$I_1 \leq 4MM_1\alpha\hat{C}_\beta < \eta/2.$$

For this α, recall that the supports of ψ^ε are contained in a compact region independent of ε, we may use (21) and the local convergence of b_ε to b in L^1 to obtain

$$I_2 \leq \eta/2 \quad \text{for sufficiently small } \varepsilon > 0.$$

Consequently, for small $\varepsilon > 0$ there holds

$$\left| \int_0^\infty \int_{-\infty}^\infty (u - v)(b_\varepsilon - b)\psi_x^\varepsilon dxdt \right| \leq \eta.$$

Since $\eta > 0$ is arbitrary, we can conclude the proof. ■

5. Riemann problems

Before giving the general existence result, we consider the scalar conservation law with simple initial values:

$$u_t + f(u)_x = 0, \quad u(x,0) = u_0(x) = \begin{cases} u_l, & x < 0, \\ u_r, & x > 0, \end{cases} \tag{25}$$

where u_l and u_r are constants. This problem is called Riemann problem. We will determine the unique entropy solution explicitly when $f'' > c_0 > 0$.

- Observing that if $u(x,t)$ is a solution of (25), then, for any $\lambda > 0$, $u_\lambda(x,t) = u(\lambda x, \lambda t)$ is also a solution. It is natural to determine the solution of the form $u(x,t) = v(x/t)$.

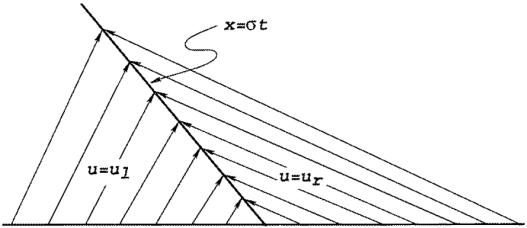
We need to consider two cases: $u_l > u_r$ and $u_l < u_r$.

- Case 1. $u_l > u_r$.
  - Since $f'' > 0$, we have $f'(u_l) > f'(u_r)$. Thus any characteristic line starting from the negative x-axis intersects characteristic lines starting from the positive x-axis.
  - Assume that the curve of discontinuities is $s(t)$. We expect that $s(0) = 0$ and $s'(t) = \sigma$ by Rankine-Hugoniot condition, where
  $$f'(u_r) < \sigma := \frac{f(u_l) - f(u_r)}{u_l - u_r} < f'(u_l).$$
  So $s(t) = \sigma t$.

- Therefore we may define

$$u(x,t) = \begin{cases} u_l, & x < \sigma t, \\ u_r, & x > \sigma t. \end{cases} \tag{26}$$

It is easy to check u is a weak solution. Since $u_l > u_r$, u thus satisfies the Oleinik entropy condition. So, by Theorem 9, u is the unique entropy solution which is called a shock wave.



Shock wave solving Riemann's problem for $u_l > u_r$

- Case 2. $u_l < u_r$.
  - In this case $f'(u_l) < f'(u_r)$. By the method of characteristics, $u = u_l$ for $x < f'(u_l)t$ and $u = u_r$ for $x > f'(u_r)t$, but u is undetermined in the region $f'(u_l)t < x < f'(u_r)t$.
  - In the region $f'(u_l)t < x < f'(u_r)t$, we expect u to be smooth with $u(x,t) = v(x/t)$. Then by $u_t + f(u)_x = 0$ we have

    $$v'(x/t)\,(f'(v(x/t)) - x/t) = 0.$$

    Assuming v' never vanishes, we find $f'(v(x/t)) = x/t$.
  - Since $f'' > c_0 > 0$, $G := (f')^{-1} : \mathbb{R} \to \mathbb{R}$ exists and

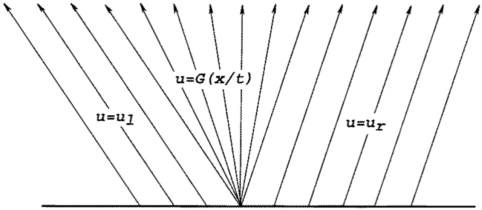
    $$|G(x) - G(y)| \leq |x - y|/c_0$$

    for $x, y \in \mathbb{R}$ (see Lemma 14).
  - Therefore $v(x/t) = G(x/t)$ for $f'(u_l)t < x < f'(u_r)t$.

- Thus we can define

$$u(x,t) = \begin{cases} u_l, & x < f'(u_l)t, \\ G(x/t), & f'(u_l)t < x < f'(u_r)t, \\ u_r, & x > f'(u_r)t. \end{cases} \tag{27}$$

Then u is continuous in $\mathbb{R} \times (0, \infty)$ and $u_t + f(u)_x = 0$ in each of its region of definition. It is easy to check that u is a weak solution.



Rarefaction wave solving Riemann's problem for $u_l < u_r$

- The Oleinik entropy condition can be directly checked case by case; for instance, if $f'(u_l)t < x < x + a < f'(u_r)t$, then

  $$u(x+a,t) - u(x,t) = (f')^{-1}((x+a)/t) - (f')^{-1}(x/t) \leq a/(c_0 t).$$

  So, by Theorem 9, u is the unique entropy solution which is called a rarefaction wave.

Summarizing the above discussion we obtain

Theorem 10

Consider the Riemann problem (25), where $f'' \geq c_0 > 0$.

(i) *If $u_l > u_r$, the unique entropy solution is given by the shock wave (26).*

(ii) *If $u_l < u_r$, the unique entropy solution is given by the rarefaction wave (27).*

6. Existence of entropy solutions

Consider the initial value problem of the scalar conservation laws

$$\begin{cases} u_t + f(u)_x = 0, & (x,t) \in \mathbb{R} \times (0,\infty), \\ u(x,0) = u_0(x), & x \in \mathbb{R}. \end{cases} \tag{28}$$

We will prove the following existence result.

Theorem 11

Let $u_0 \in L^\infty(\mathbb{R})$ and $f \in C^2(\mathbb{R})$ with $f''(\xi) \geq c_0 > 0$ on $\mathbb{R}$. Then (28) has a unique weak solution $u \in L^\infty(\mathbb{R} \times [0,\infty))$ satisfying the Oleinik entropy condition. Moreover

$$\|u(x,t)\|_{L^\infty(\mathbb{R}\times(0,\infty))} \leq \|u_0\|_\infty.$$

- Theorem 11 has several different proofs. We present the one based on the theory of Hamilton-Jacobi equations.
- To motivate it, let $h(x) := \int_0^x u_0(y)dy$ and consider the initial value problem of Hamilton-Jacobi equation

$$\begin{cases} w_t + f(w_x) = 0, & (x,t) \in \mathbb{R} \times (0,\infty), \\ w(x,0) = h(x), & x \in \mathbb{R}. \end{cases} \tag{29}$$

If (29) has smooth solution, we set $u = w_x$. Then $u(x,0) = w_x(x,0) = u_0(x)$. Differentiating the equation in (29) gives

$$u_t = w_{xt} = (w_t)_x = -f(w_x)_x = -f(u)_x.$$

Thus $u = w_x$ is a solution of (28).

- Unfortunately the solution of (29) is not necessarily smooth in general.
- It is necessary to introduce the notion of weak solution of (29).

Definition 12

Consider the problem (29), where h is Lipschitz continuous. A Lipschitz continuous function $w : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ is called a weak solution if

(i) $w(x, 0) = h(x)$ for all $x \in \mathbb{R}$;

(ii) $w_t(x, t) + f(w_x(x, t)) = 0$ for a.e. $(x, t) \in \mathbb{R} \times (0, \infty)$.

- When $f \in C^2$ with $f'' \geq c_0 > 0$, we will show that the solution of (29) is given by the Hopf-Lax formula.
- To motivate the formula, assuming (29) has a C^1 solution. Along a characteristic curve $x(t)$ we set $z(t) := w(x(t), t)$ and $p(t) := w_x(x(t), t)$. Then there hold

$$\begin{cases} \frac{dx}{dt} &= f'(p), \\ \frac{dz}{dt} &= pf'(p) - f(p), \\ \frac{dp}{dt} &= 0. \end{cases} \tag{30}$$

Thus along characteristics p are constants. So, characteristics are straight lines with velocity $f'(p)$. To understand the second equation in (30), we introduce the Legendre-Fenchel conjugate

$$f^*(q) = \sup_{p\in\mathbb{R}} \{pq - f(p)\}, \quad q \in \mathbb{R}.$$

Since f is uniformly convex, the maximum is achieved at p satisfying $q = f'(p)$. Thus

$$f^*(q) = pf'(p) - f(p) \qquad \text{with } f'(p) = q.$$

So $\frac{dz}{dt} = f^*(q)$ with $q = f'(p)$. Fix any $(\bar{x}, \bar{t})$ with $\bar{t} > 0$. For a characteristic line through $(\bar{x}, \bar{t})$ that crosses x-axis at $\bar{y}$, its velocity is $(\bar{x} - \bar{y})/\bar{t}$. Thus, along this characteristic,

$$\frac{dz}{dt} = f^*(\frac{\bar{x} - \bar{y}}{\bar{t}}), \qquad z(0) = h(\bar{y}).$$

Therefore

$$w(\bar{x}, \bar{t}) = z(\bar{t}) = h(\bar{y}) + \bar{t} f^*(\frac{\bar{x} - \bar{y}}{\bar{t}}) \tag{31}$$

This formula is problematic since it involves the unknown $\bar{y}$.

- On the othe hand, by the convexity of f we have for any p

$$-w_t = f(w_x) \geq f(p) + f'(p)(w_x - p).$$

So

$$w_t + f'(p) w_x \leq p f'(p) - f(p) = f^*(f'(p)).$$

Consider the straight line $(x(t), t)$ through $(\bar{x}, \bar{t})$ with velocity $f'(p)$, let y be the intersection point with x-axis. Then

$$f'(p) = (\bar{x} - y)/\bar{t}$$

and

$$\frac{d}{dt}w(x(t),t) \leq f^*(f'(p)) = f^*(\frac{\bar{x}-y}{\bar{t}}).$$

Therefore

$$w(\bar{x},\bar{t}) \leq h(y) + \bar{t}f^*(\frac{\bar{x}-y}{\bar{t}}). \tag{32}$$

Since $f'' \geq c_0 > 0$, f' is strictly increasing with $f'(-\infty) = -\infty$ and $f'(+\infty) = +\infty$. Thus (32) holds for all $y \in \mathbb{R}$ since we can take y to be any number by adjusting p. Since (31) implies that the equality is achieved at some $\bar{y}$, we expect

$$w(x,t) := \inf_{y\in\mathbb{R}} \left\{ h(y) + t\,f^*(\frac{x-y}{t}) \right\} \tag{33}$$

which is called the Hopf-Lax formula.

- The above argument is not rigorous since it requires $w \in C^1$.
- Our goal is to show that (33) gives a weak solution of (29).

We first give some properties on f^*.

Lemma 13

Let f be a C^1 convex function on $\mathbb{R}$. Then the following hold:

(i) *f^* is convex;*

(ii) *For any $A > 0$ we have*

$$\sup_{q \in \mathbb{R}} \{A|q| - f^*(q)\} \leq \sup \{f(x) : |x| \leq A\};$$

(iii) *For any $x \in \mathbb{R}$ we have $\sup_{q \in \mathbb{R}} \{qx - f^*(q)\} = f(x)$.*

Proof.

(i) f^* is convex because f^* is the supremum of linear functions.

(ii) By the definition of f^* we have

$$f^*(q) = \sup_{y\in\mathbb{R}} \{qy - f(y)\} \geq q\frac{Aq}{|q|} - f(\frac{Aq}{|q|}) = A|q| - f(Aq/|q|).$$

Therefore

$$\sup_{q\in\mathbb{R}} \{A|q| - f^*(q)\} \leq \sup_{q\in\mathbb{R}} \{f(Aq/|q|)\} = \sup \{f(x) : |x| \leq A\}.$$

(iii) Since the definition of f^* implies $f^*(q) \geq qx - f(x)$ for all $q \in \mathbb{R}$, we have

$$\sup_{q\in\mathbb{R}} \{qx - f^*(q)\} \leq f(x).$$

To show the reverse inequality, we note that

$$qx - f^*(q) = qx - \sup_{y\in\mathbb{R}} \{qy - f(y)\} = \inf_{y\in\mathbb{R}} \{q(x-y) + f(y)\}$$

Thus

$$\begin{aligned}\sup_{q\in\mathbb{R}} \{qx - f^*(q)\} &= \sup_q \inf_y \{q(x-y) + f(y)\} \\ &\geq \inf_y \left\{f'(x)(x-y) + f(y)\right\}\end{aligned}$$

Since f is convex, we have $f(y) \geq f(x) + f'(x)(y-x)$ and thus

$$f(y) + f'(x)(x-y) \geq f(x), \quad \forall y.$$

So $\sup_{q\in\mathbb{R}}\{qx - f^*(q)\} \geq f(x)$. The proof is complete. ■

Lemma 14

Let $f \in C^2$ be such that $f'' \geq c_0$ for some constant $c_0 > 0$. Then

(i) *$f^* \in C^2$ is strictly convex and $(f^*)' = (f')^{-1}$, where $(f')^{-1}$ denotes the inverse function of f';*

(ii) *$(f^*)'$ is Lipschitz continuous, i.e. for any $p, q \in \mathbb{R}$ there holds*

$$|(f^*)'(p) - (f^*)'(q)| \leq \frac{|p-q|}{c_0}$$

Proof. By the condition on f, f' is strictly increasing with $f'(-\infty) = -\infty$ and $f'(+\infty) = +\infty$, and thus $g := (f')^{-1} : \mathbb{R} \to \mathbb{R}$ exists as a C^1 function with $g'(x) = 1/f''(g(x)) > 0$.

(i) For any $q \in \mathbb{R}$, there always holds $f^*(q) = qx - f(x)$, where x is determined by $q = f'(x)$, i.e. $x = (f')^{-1}(q) = g(q)$. Thus

$$f^*(q) = qg(q) - f(g(q)), \qquad \forall q.$$

This implies that $f^* \in C^1$ and

$$\begin{aligned}(f^*)'(q) &= g(q) + qg'(q) - f'(g(q))g'(q) \\ &= g(q) + qg'(q) - qg'(q) = g(q).\end{aligned}$$

Consequently $(f^*)' = g$ and $f^* \in C^2$ with $(f^*)'' = g' > 0$.

(ii) For any $p, q \in \mathbb{R}$ let $x = (f^*)'(p)$ and $y = (f^*)'(q)$. Then

$$p = f'(x) \quad \text{and} \quad q = f'(y).$$

Since $f'' \geq c_0$, we have

$$f(y) - f(x) - f'(x)(y-x) \geq \frac{1}{2}c_0(y-x)^2,$$
$$f(x) - f(y) - f'(y)(x-y) \geq \frac{1}{2}c_0(x-y)^2.$$

Adding these two inequalities gives

$$c_0(x-y)^2 \leq (f'(x) - f'(y))(x-y) \leq |f'(x) - f'(y)||x-y|$$

This implies that $c_0|x-y| \leq |f'(x) - f'(y)|$, i.e.

$$c_0|(f^*)'(p) - (f^*)'(q)| \leq |p-q|.$$

This completes the proof. ■

Lemma 15

The function w defined by the Hopf-Lax formula (33) is Lipschitz continuous on $\mathbb{R} \times [0, \infty)$ and $w(x, 0) = h(x)$ for $x \in \mathbb{R}$.

Proof. We use

$$Lip(F) := \sup \{|F(x) - F(y)|/|x - y| : x, y \in \mathbb{R} \text{ and } x \neq y\}$$

to denote the Lipschitz constant of a Lipschitz function F.

- We first show that, for each $t > 0$, $w(\cdot, t)$ is Lipschitz with

$$Lip(w(\cdot, t)) \leq Lip(h).$$

To see this, let $x_1, x_2 \in \mathbb{R}$. We may take $y_1 \in \mathbb{R}$ such that

$$w(x_1, t) = h(y_1) + t\, f^*(\frac{x_1 - y_1}{t}).$$

Then

$$
\begin{aligned}
& w(x_2, t) - w(x_1, t) \\
& = \inf \left\{ h(y) + tf^*(\frac{x_2 - y}{t}) \right\} - h(y_1) - tf^*(\frac{x_1 - y_1}{t}) \\
& \leq h(x_2 - x_1 + y_1) - h(y_1) \leq Lip(h)|x_2 - x_1|.
\end{aligned}
$$

Interchanging the role of x_1 and x_2 we then obtain

$$
|w(x_1, t) - w(x_2, t)| \leq Lip(h)|x_1 - x_2|. \qquad (34)
$$

- We next show that there is a constant $C_0 > 0$ such that

$$
|w(x, t) - h(x)| \leq C_0 t, \quad \forall x \in \mathbb{R} \text{ and } t > 0.
$$

Indeed, we first have

$$w(x,t) \leq h(x) + t\,f^*(0).$$

Moreover, by using $h(y) \geq h(x) - Lip(h)|x-y|$ we have

$$\begin{aligned}
w(x,t) &= \inf_{y\in\mathbb{R}} \left\{ h(y) + t\,f^*(\frac{x-y}{t}) \right\} \\
&\geq h(x) - \sup_{y\in\mathbb{R}} \left\{ Lip(h)|x-y| - t\,f^*(\frac{x-y}{t}) \right\} \\
&= h(x) - t \sup_{z\in\mathbb{R}} \{ Lip(h)|z| - f^*(z) \} \\
&\geq h(x) - C_1 t,
\end{aligned}$$

where $C_1 := \sup_{|y|\leq Lip(h)} f(y)$ by Lemma 13 (ii).

- We further show that there is a constant C_2 such that

$$|w(x,t_1) - w(x,t_2)| \le C_2(t_2 - t_1) \qquad (35)$$

for all $x \in \mathbb{R}$ and $0 < t_1 < t_2$. Indeed, letting $y \in \mathbb{R}$ be such that

$$w(x,t_1) = h(y) + t_1 f^*\left((x-y)/t_1\right),$$

we may use the definition of $w(x,t_2)$ to obtain

$$w(x,t_2) \le h(y) + t_2 f^*\left((x-y)/t_2\right).$$

By writing

$$\frac{x-y}{t_2} = \frac{t_1}{t_2}\frac{x-y}{t_1} + \left(1 - \frac{t_1}{t_2}\right) \cdot 0$$

and using the convexity of f^* we have

$$\begin{aligned} w(x,t_2) &\le h(y) + t_2\left\{\frac{t_1}{t_2}f^*(\frac{x-y}{t_1}) + \left(1-\frac{t_1}{t_2}\right)f^*(0)\right\} \\ &= h(y) + t_1 f^*(\frac{x-y}{t_1}) + (t_2-t_1)f^*(0) \\ &= w(x,t_1) + (t_2-t_1)f^*(0). \end{aligned}$$

Therefore

$$w(x,t_2) - w(x,t_1) \le (t_2-t_1)f^*(0), \quad 0 < t_1 < t_2. \tag{36}$$

On the other hand, we may take $z \in \mathbb{R}$ such that

$$w(x,t_2) = h(z) + t_2 f^*((x-z)/t_2).$$

Let $y = \frac{t_1}{t_2}x + (1 - \frac{t_1}{t_2})z$. Since $\frac{x-z}{t_2} = \frac{y-z}{t_1} = \frac{x-y}{t_2-t_1}$, we have

$$\begin{aligned} w(x, t_2) &= h(z) + t_1 f^*(\frac{y-z}{t_1}) + t_2 f^*(\frac{x-z}{t_2}) - t_1 f^*(\frac{y-z}{t_1}) \\ &\geq w(y, t_1) + (t_2 - t_1) f^*(\frac{x-y}{t_2 - t_1}). \end{aligned}$$

Using (34) we have

$$w(y, t_1) \geq w(x, t_1) - Lip(h)|y - x|.$$

Therefore

$$w(x, t_2) \geq w(x, t_1) - Lip(h)|x - y| + (t_2 - t_1) f^*(\frac{x-y}{t_2 - t_1}).$$

Consequently

$$w(x,t_2) \geq w(x,t_1) - (t_2 - t_1) \sup_{\eta \in \mathbb{R}} \{Lip(h)|\eta| - f^*(\eta)\}$$

So, by Lemma 13 (ii), we have

$$w(x,t_2) - w(x,t_1) \geq -C_1(t_2 - t_1), \quad 0 < t_1 < t_2.$$

Combining this with (36) we obtain (35).

- Finally, by writing

$$|w(x_1,t_1) - w(x_2,t_2)| \leq |w(x_1,t_1) - w(x_2,t_1)| + |w(x_2,t_1) - w(x_2,t_2)|,$$

we may use (34) and (35) to complete the proof. ∎

Theorem 16

The function w defined by the Hopf-Lax formula (33) is Lipschitz continuous, is differentiable a.e. on $\mathbb{R} \times (0, \infty)$ and is a weak solution of (29).

Proof. By Lemma 15, w is Lipschitz on $\mathbb{R} \times [0, \infty)$ with $w(\cdot, 0) = h$. So w is differentiable a.e. in $\mathbb{R} \times (0, \infty)$ by Rademacher's Theorem. It remains only to show that

$$w_t(x, t) + f(w_x(x, t)) = 0$$

for any $(x, t) \in \mathbb{R} \times (0, \infty)$ at which w is differentiable.

- We first choose $z \in \mathbb{R}$ such that

$$w(x, t) = h(z) + t\, f^*((x - z)/t).$$

Fix any $0 < \varepsilon < t$ and set $y = (1 - \frac{\varepsilon}{t})x + \frac{\varepsilon}{t}z$. Then

$$w(y, t - \varepsilon) \leq h(z) + (t - \varepsilon) f^*(\frac{y - z}{t - \varepsilon}).$$

Since $\frac{x-z}{t} = \frac{y-z}{t-\varepsilon}$, we have

$$\begin{aligned} w(x, t) - w(y, t - \varepsilon) &\geq t\, f^*(\frac{x - z}{t}) - (t - \varepsilon) f^*(\frac{x - z}{t}) \\ &= \varepsilon f^*(\frac{x - z}{t}). \end{aligned}$$

Therefore

$$\frac{w(x, t) - w(x + \frac{\varepsilon}{t}(z - x), t - \varepsilon)}{\varepsilon} \geq f^*(\frac{x - z}{t}).$$

- Letting $\varepsilon \searrow 0$ gives

$$\frac{x-z}{t} w_x(x,t) + w_t(x,t) \geq f^*(\frac{x-z}{t}).$$

Consequently, by the definition of f^*,

$$\begin{aligned} & w_t(x,t) + f(w_x(x,t)) \\ & \geq f(w_x(x,t)) + f^*(\frac{x-z}{t}) - \frac{x-z}{t} w_x(x,t) \geq 0. \end{aligned}$$

- On the other hand, fix any $q \in \mathbb{R}$ and $\varepsilon > 0$. Then

$$w(x+\varepsilon q, t+\varepsilon) = \inf_{y\in\mathbb{R}} \left\{ h(y) + (t+\varepsilon) f^*(\frac{x+\varepsilon q - y}{t+\varepsilon}) \right\}.$$

- Since $\frac{x+\varepsilon q-y}{t+\varepsilon} = \frac{\varepsilon}{t+\varepsilon}q + \frac{t}{t+\varepsilon}\frac{x-y}{t}$, we may use the convexity of f^* to derive

$$(t+\varepsilon)f^*(\frac{x+\varepsilon q-y}{t+\varepsilon}) \leq \varepsilon\, f^*(q) + t\, f^*(\frac{x-y}{t}).$$

Therefore

$$\begin{aligned} w(x+\varepsilon q, t+\varepsilon) &\leq \varepsilon f^*(q) + \inf_{y\in\mathbb{R}} \left\{ h(y) + tf^*(\frac{x-y}{t}) \right\} \\ &= \varepsilon f^*(q) + w(x,t). \end{aligned}$$

So

$$\frac{w(x+\varepsilon q, t+\varepsilon) - w(x,t)}{\varepsilon} \leq f^*(q).$$

Letting $\varepsilon \searrow 0$ gives

$$qw_x(x,t) + w_t(x,t) \le f^*(q), \quad \forall q \in \mathbb{R}.$$

Therefore, by Lemma 13 (iii),

$$-w_t(x,t) \ge \sup_{q\in\mathbb{R}} \{qw_x(x,t) - f^*(q)\} = f(w_x(x,t)),$$

i.e. $w_t(x,t) + f(w_x(x,t)) \le 0$. The proof is thus complete. ■

We are ready to complete the proof of Theorem 11. To this end, let $h(x) = \int_0^x u_0(y)dy$ and define $w(x,t)$ by the Hopf-Lax formula

$$w(x,t) = \inf_{y\in\mathbb{R}} \left\{ h(y) + t\, f^*(\frac{x-y}{t}) \right\}.$$

By Theorem 16, w is Lipschitz, is differentiable for a.e. (x,t), and

$$\begin{aligned} w_t + f(w_x) &= 0 \quad \text{a.e. in } \mathbb{R}\times(0,\infty), \\ w(x,0) &= h(x), \quad x\in\mathbb{R}. \end{aligned}$$

Lemma 17

Let $u := w_x$. Then u is a weak solution of (28).

Proof. Recall that $Lip(w) \le Lip(h) = \|u_0\|_\infty$, $u \in L^\infty(\mathbb{R}\times(0,\infty))$ with

$$\|u\|_\infty \le Lip(w) \le \|u_0\|_\infty.$$

Next for any $\varphi \in C_0^1(\mathbb{R}\times[0,\infty))$ we have

$$0 = \int_0^\infty \int_{-\infty}^\infty (w_t + f(w_x))\varphi_x dxdt. \tag{37}$$

Since w is Lipschitz, $x \to w(x,t)$ is absolute continuous for each $t \geq 0$ and $t \to w(x,t)$ is absolute continuous for each $x \in \mathbb{R}$. So, integration by parts can be used to obtain

$$\begin{aligned}
&\int_0^\infty \int_{-\infty}^\infty w_t \varphi_x dxdt \\
&= -\int_0^\infty \int_{-\infty}^\infty w \varphi_{xt} dxdt - \int_{-\infty}^\infty w(x,0)\varphi_x(x,0)dx \\
&= \int_0^\infty \int_{-\infty}^\infty w_x \varphi_t dxdt + \int_{-\infty}^\infty w_x(x,0)\varphi(x,0)dx.
\end{aligned}$$

Since $w_x(x,0) = u_0(x)$ for a.e. x, we have

$$\int_0^\infty \int_{-\infty}^\infty w_t \varphi_x dxdt = \int_0^\infty \int_{-\infty}^\infty w_x \varphi_t dxdt + \int_{-\infty}^\infty u_0(x)\varphi(x,0)dx.$$

Combining this with (37) gives

$$0 = \int_0^\infty \int_{-\infty}^\infty (w_x \varphi_t + f(w_x)\varphi_x) dxdt + \int_{-\infty}^\infty u_0(x)\varphi(x,0)dx.$$

Thus $u = w_x$ is a weak solution of (28).

- To complete the proof of Theorem 11, it remains only to show that there is a function $\tilde{u}$ with $u = \tilde{u}$ a.e. in $\mathbb{R} \times (0, \infty)$ such that $\tilde{u}$ satisfies the Oleinik entropy condition.
- To this end, we will use, for each (x, t) with $t > 0$, the minimizer of the function

$$\mathcal{F}_{x,t}(y) := h(y) + tf^*(\frac{x-y}{t}) \quad \text{over } \mathbb{R}.$$

The following lemma shows that for each fixed $t > 0$, if $x_1 < x_2$ then the minimizer of $\mathcal{F}_{x_1,t}(y)$ is always on the left of the minimizer of $\mathcal{F}_{x_2,t}(y)$.

Lemma 18

Assume that $f \in C^2$ satisfies $f'' \geq c_0 > 0$. Fix $t > 0$ and $x_1 < x_2$. If $y_1 \in \mathbb{R}$ is such that

$$\min_{y \in \mathbb{R}} \left\{ h(y) + t\, f^*(\frac{x_1 - y}{t}) \right\} = h(y_1) + t\, f^*(\frac{x_1 - y_1}{t}),$$

then

$$h(y_1) + t\, f^*(\frac{x_2 - y_1}{t}) < h(y) + t\, f^*(\frac{x_2 - y}{t}), \quad \forall y < y_1.$$

Proof. Let $\tau = \frac{y_1 - y}{x_2 - x_1 + y_1 - y}$. Then $0 < \tau < 1$ and

$$x_2 - y_1 = \tau(x_1 - y_1) + (1-\tau)(x_2 - y),$$
$$x_1 - y = (1-\tau)(x_1 - y_1) + \tau(x_2 - y).$$

By the strict convexity of f^*, see Lemma 14 (i), we have

$$f^*(\frac{x_2 - y_1}{t}) < \tau f^*(\frac{x_1 - y_1}{t}) + (1-\tau) f^*(\frac{x_2 - y}{t}),$$
$$f^*(\frac{x_1 - y}{t}) < (1-\tau) f^*(\frac{x_1 - y_1}{t}) + \tau f^*(\frac{x_2 - y}{t}).$$

Adding these two inequalities gives

$$f^*(\frac{x_2 - y_1}{t}) + f^*(\frac{x_1 - y}{t}) < f^*(\frac{x_1 - y_1}{t}) + f^*(\frac{x_2 - y}{t}).$$

Therefore

$$\begin{aligned}
& t\,f^*\left(\frac{x_2 - y_1}{t}\right) + t\,f^*\left(\frac{x_1 - y}{t}\right) + h(y_1) + h(y) \\
& < t\,f^*\left(\frac{x_1 - y_1}{t}\right) + t\,f^*\left(\frac{x_2 - y}{t}\right) + h(y_1) + h(y) \\
& \leq t\,f^*\left(\frac{x_1 - y}{t}\right) + h(y) + t\,f^*\left(\frac{x_2 - y}{t}\right) + h(y);
\end{aligned}$$

for the last inequality we used the fact that y_1 is a minimizer. This implies the conclusion. ■

Now we are able to give the construction of $\tilde{u}$ which is stated in the following result.

Lemma 19

There exists a function $y(x,t)$ defined on $\mathbb{R} \times (0,\infty)$ such that

(i) *for each $t > 0$, $x \to y(x,t)$ is nondecreasing;*

(ii) *for each (x,t) with $t > 0$, $y(x,t)$ is a minimizer of the function*

$$\mathcal{F}_{x,t}(y) := h(y) + tf^*(\frac{x-y}{t}).$$

(iii) *if we set $\tilde{u}(x,t) = (f^*)'(\frac{x-y(x,t)}{t})$, then, for each $t > 0$,*

$$u(x,t) = \tilde{u}(x,t) \qquad \text{for a.e. } x.$$

In particular, $u = \tilde{u}$ for a.e. $(x,t) \in \mathbb{R} \times (0,\infty)$.

Proof.

- Fix $t > 0$. For each $x \in \mathbb{R}$ let $y(x,t)$ be the smallest of those points y giving the minimum of $\mathcal{F}_{x,t}(y)$.
- It follows from Lemma 18 that $x \to y(x,t)$ is nondecreasing and thus $y(\cdot,t)$ is continuous for all but at most countably many x.
- At a point x of continuity of $y(\cdot,t)$, $y(x,t)$ is the unique minimizer of $\mathcal{F}_{x,t}(y)$ over $\mathbb{R}$.
- From Theorem 16 it follows for each fixed $t > 0$ that

$$x \to w(x,t) := \min_{y\in\mathbb{R}} \left\{ h(y) + tf^*(\frac{x-y}{t}) \right\}$$
$$= h(y(x,t)) + tf^*(\frac{x-y(x,t)}{t})$$

is differentiable a.e.

- Since $x \to y(x,t)$ is monotone, it is differentiable a.e. as well. Thus, for a.e. x, $f^*(\frac{x-y(x,t)}{t})$ is differentiable and therefore $x \to h(y(x,t))$ is differentiable as well.
- Consequently for a.e. x

$$u(x,t) = \frac{\partial}{\partial x}\left(h(y(x,t)) + tf^*(\frac{x-y(x,t)}{t})\right)$$
$$= \frac{\partial}{\partial x}(h(y(x,t))) + (f^*)'(\frac{x-y(x,t)}{t})(1-y_x(x,t)).$$

- Since $y(x,t)$ is a minimizer of $\mathcal{F}_{x,t}(y)$ over $\mathbb{R}$, x must be a minimizer of

$$z \to \mathcal{F}_{x,t}(y(z,t)) = h(y(z,t)) + tf^*(\frac{x-y(z,t)}{t}).$$

- Consequently $0 = \frac{\partial}{\partial z}\Big|_{z=x} (\mathcal{F}_{x,t}(y(z,t)))$, i.e.

$$0 = \frac{\partial}{\partial x}(h(y(x,t))) - (f^*)'(\frac{x - y(x,t)}{t})y_x(x,t)$$

We therefore obtain $u(x,t) = (f^*)'(\frac{x-y(x,t)}{t})$ a.e. ■

Theorem 20

Let $f \in C^2$ satisfy $f'' \geq c_0 > 0$, let $u_0 \in L^\infty(\mathbb{R})$ and let $h(x) := \int_0^x u_0(y)dy$. Then the function

$$\tilde{u}(x,t) = (f^*)'(\frac{x - y(x,t)}{t}) \tag{38}$$

defined in Lemma 19 is a weak solution of (28) satisfying the Oleinik entropy condition.

Proof. By condition and Lemma 14, $(f^*)'$ is increasing. Thus, by Lemma 19, we have for any $t > 0$ and $x, a \in \mathbb{R}$ with $a > 0$ that

$$\tilde{u}(x,t) = (f^*)'(\frac{x - y(x,t)}{t}) \geq (f^*)'(\frac{x - y(x+a,t)}{t}).$$

By Lemma 14 (ii), we have

$$\begin{aligned} \tilde{u}(x,t) &\geq (f^*)'(\frac{x + a - y(x+a,t)}{t}) - a/(c_0 t) \\ &= \tilde{u}(x+a,t) - a/(c_0 t). \end{aligned}$$

The proof is complete. ■

Remark. The formula (38) is called the Lax-Oleinik formula. Recall that $(f^*)' = (f')^{-1}$, we have $\tilde{u}(x,t) = (f')^{-1}((x - y(x,t))/t)$.

7. Long time behavior

We prove a uniform decay estimate for the entropy solution of the scalar conservation law

$$u_t + f(u)_x = 0, \qquad u(x,0) = u_0(x) \tag{39}$$

with uniformly convex flux $f(u)$.

Theorem 21

Let $u_0 \in L^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$ and $f \in C^2$ with $f'' \geq c_0 > 0$. Then the entropy solution of (39) satisfies the estimate

$$|u(x,t)| \leq C/t^{1/2},$$

where C is a constant depending only on c_0 and $\|u_0\|_{L^1}$.

Proof. We use the Lax-Oleinik formula

$$u(x,t) = (f^*)'(\frac{x - y(x,t)}{t}).$$

In order to use the Lipschitz continuity of $(f^*)'$, we take $\sigma \in \mathbb{R}$ such that

$$(f^*)'(\sigma) = 0,$$

i.e. $(f')^{-1}(\sigma) = 0$; we can take $\sigma = f'(0)$. Then

$$\begin{aligned} |u(x,t)| &= \left|(f^*)'(\frac{x - y(x,t)}{t}) - (f^*)'(\sigma)\right| \\ &\le \frac{1}{c_0}\left|\frac{x - y(x,t)}{t} - \sigma\right|. \end{aligned} \tag{40}$$

To estimate the right hand side, by the definition of $y(x,t)$ we have

$$
\begin{aligned}
h(y(x,t)) + tf^*(\frac{x-y(x,t)}{t}) &= \min_{y\in\mathbb{R}} \left\{ h(y) + tf^*(\frac{x-y}{t}) \right\} \\
&\le h(x-\sigma t) + tf^*(\sigma)
\end{aligned}
$$

where $h(x) = \int_0^x u_0(\eta)d\eta$. Since $f'' \ge c_0 > 0$, we have

$$
\begin{aligned}
f^*(\frac{x-y(x,t)}{t}) \ge f^*(\sigma) &+ (f^*)'(\sigma)\left(\frac{x-y(x,t)}{t} - \sigma\right) \\
&+ \frac{1}{2}c_0\left(\frac{x-y(x,t)}{t} - \sigma\right)^2.
\end{aligned}
$$

Combining these last two inequalities gives

$$\frac{1}{2}tc_0\left(\frac{x-y(x,t)}{t}-\sigma\right)^2 \le h(x-\sigma t)-h(y(x,t)).$$

Recall the definition of h and $u_0 \in L^1(\mathbb{R})$, we have $|h(x)| \le \|u_0\|_{L^1}$ for all $x \in \mathbb{R}$. Therefore

$$\frac{1}{2}tc_0\left(\frac{x-y(x,t)}{t}-\sigma\right)^2 \le 2\|u_0\|_{L^1},$$

i.e.

$$\left|\frac{x-y(x,t)}{t}-\sigma\right| \le \sqrt{\frac{4\|u_0\|_{L^1}}{c_0 t}}.$$

Combining this with (40) gives the desired estimate. ■

Part 2. Lectures on wave equations

1. Solutions of linear wave equations

We consider the Cauchy problem of linear wave equation

$$\begin{cases} u_{tt} - \triangle u = f(x,t), & x \in \mathbb{R}^n,\ t > 0, \\ u(x,0) = g(x), \quad u_t(x,0) = h(x), & x \in \mathbb{R}, \end{cases} \tag{41}$$

where $\triangle = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ denotes the Laplacian operator on $\mathbb{R}^n$.

- A function $u \in C^2(\mathbb{R}^n \times [0,\infty))$ satisfying (41) is called a classical solution of (41).
- We prove the uniqueness result by deriving energy estimate and establish the existence result of classical solutions by deriving the solution formulae.

1.1. Uniquessness

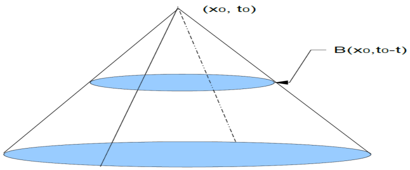
- We show that the Cauchy problem (41) has at most one classical solution.
- We establish uniqueness result by proving a general result, the so-called finite speed propagation property.
- Consider the homogeneous wave equation

$$\Box u := \partial_t^2 u - \triangle u = 0 \quad \text{in } \mathbb{R}^n \times [0, \infty). \tag{42}$$

For any fixed $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$, we introduce

$$C_{x_0,t_0} := \{(x, t) : 0 \le t \le t_0 \text{ and } |x - x_0| \le t_0 - t\}$$

which is called the backward light cone with vertex (x_0, t_0).



The following result says that any "disturbance" originating outside $B_{t_0}(x_0) := \{x \in \mathbb{R}^n : |x - x_0| \leq t_0\}$ at $t = 0$ has no effect on the solution within C_{x_0,t_0}.

Theorem 22 (finite speed of propagation)

Let u be a C^2 solution of (42) in C_{x_0,t_0}. If $u(x,0) \equiv u_t(x,0) \equiv 0$ for $x \in B_{t_0}(x_0)$, then $u \equiv 0$ in C_{x_0,t_0}.

Proof. Consider for $0 \leq t \leq t_0$ the function

$$E(t) := \int_{B_{t_0 - t}(x_0)} \left(|u_t(x,t)|^2 + |\nabla u(x,t)|^2 \right) dx$$
$$= \int_0^{t_0 - t} \int_{\partial B_\tau(x_0)} \left(|u_t(x,t)|^2 + |\nabla u(x,t)|^2 \right) d\sigma(x) d\tau.$$

We have

$$\frac{d}{dt} E(t) = 2 \int_{B_{t_0 - t}(x_0)} \left(u_t(x,t) u_{tt}(x,t) + \nabla u(x,t) \cdot \nabla u_t(x,t) \right) dx$$
$$- \int_{\partial B_{t_0 - t}(x_0)} \left(|u_t(x,t)|^2 + |\nabla u(x,t)|^2 \right) d\sigma(x).$$

Since $\nabla u \cdot \nabla u_t = \mathrm{div}(u_t \nabla u) - u_t \triangle u$, we have

$$\frac{d}{dt}E(t) = 2\int_{B_{t_0-t}(x_0)} u_t \Box u dx + 2\int_{B_{t_0-t}(x_0)} \mathrm{div}(u_t \nabla u) dx - \int_{\partial B_{t_0-t}(x_0)} \left(|u_t|^2 + |\nabla u|^2\right) d\sigma.$$

Using $\Box u = 0$ and the divergence theorem we have

$$\frac{d}{dt}E(t) = 2\int_{\partial B_{t_0-t}(x_0)} u_t \nabla u \cdot \nu d\sigma - \int_{\partial B_{t_0-t}(x_0)} \left(|u_t|^2 + |\nabla u|^2\right) d\sigma,$$

where ν denotes the outward unit normal to $\partial B_{t_0-t}(x_0)$. We have

$$2|u_t \nabla u \cdot \nu| \leq 2|u_t||\nabla u| \leq |u_t|^2 + |\nabla u|^2.$$

Consequently $\frac{d}{dt}E(t) \leq 0$ which implies that

$$E(t) \leq E(0), \quad 0 \leq t \leq t_0.$$

Since $u(\cdot,0) \equiv u_t(\cdot,0) \equiv 0$ on $B_{t_0}(x_0)$, we have $E(0) = 0$. Thus $E(t) \equiv 0$ for $0 \leq t \leq t_0$. Therefore

$$u_t = \nabla u = 0 \quad \text{in } C_{x_0,t_0}.$$

So $u = \textit{constant}$ in C_{x_0,t_0}. Since $u(x,0) = 0$ for $x \in B_{t_0}(x_0)$, we must have $u \equiv 0$ in C_{t_0,x_0}. ∎

Corollary 23

The Cauchy problem (41) of linear wave equation has at most one classical solution.

Proof. Assume that u_1 and u_2 are two classical solutions of (41). Then $u := u_1 - u_2 \in C^2(\mathbb{R}^n \times [0, \infty))$ satisfies

$$\begin{cases} \Box u = u_{tt} - \triangle u = 0 \quad \text{in } \mathbb{R}^n \times (0, \infty), \\ u(x, 0) = 0, \quad u_t(x, 0) = 0, \quad x \in \mathbb{R}^n. \end{cases}$$

Applying Theorem 22 to u, we conclude $u = 0$ in $\mathbb{R}^n \times [0, \infty)$. ■

1.2. Existence

The existence of (41) can be established by solving the following two problems:

$$\begin{cases} \Box u := u_{tt} - \triangle u = 0 \quad \text{in } \mathbb{R}^n \times (0, \infty), \\ u(x, 0) = g(x), \quad u_t(x, 0) = h(x), \quad x \in \mathbb{R}^n \end{cases} \tag{43}$$

and

$$\begin{cases} \Box u := u_{tt} - \triangle u = f(x,t) & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = 0, \quad u_t(x,0) = 0, & x \in \mathbb{R}^n. \end{cases} \tag{44}$$

- If v is the solution of (43) and w is the solution of (44), then $u := v + w$ is the solution of (41).
- We will solve (43) by deriving the explicit solution formula.
- We then solve (44) by reducing it to a problem like (43) using the Duhamel principle.

We now derive the solution formula of (43) when $n = 1, 2, 3$.

Case $n = 1$: Consider the Cauchy problem of 1D homogeneous wave equation

$$\begin{cases} u_{tt} - u_{xx} = 0 & \text{in } \mathbb{R} \times (0, \infty), \\ u(x,0) = g(x), \quad u_t(x,0) = h(x), & x \in \mathbb{R}, \end{cases} \tag{45}$$

where $g \in C^2(\mathbb{R})$ and $h \in C^1(\mathbb{R})$.

- Observing that $u_{tt} - u_{xx} = (\partial_t - \partial_x)(\partial_t + \partial_x)u$. We introduce $v = u_t + u_x$. Then $v_t - v_x = 0$ in $\mathbb{R} \times (0, \infty)$. By the method of Characteristics, we have

$$v(x,t) = v_0(x+t),$$

where $v_0(x) := v(x,0)$.

- So $u_t + u_x = v_0(x+t)$. Let $u_0(x) := u(x,0)$. Then, by the method of characteristics again, it follows

$$\begin{aligned} u(x,t) &= u_0(x-t) + \int_0^t v_0(x-t+2s)ds \\ &= u_0(x-t) + \frac{1}{2}\int_{x-t}^{x+t} v_0(\xi)d\xi. \end{aligned}$$

- The initial conditions give $u_0(x) = g(x)$ and $v_0(x) = h(x) + g'(x)$. Therefore

$$\begin{aligned} u(x,t) &= g(x-t) + \frac{1}{2}\int_{x-t}^{x+t} \left(g'(\xi) + h(\xi)\right) d\xi \\ &= \frac{1}{2}\left(g(x+t) + g(x-t)\right) + \frac{1}{2}\int_{x-t}^{x+t} h(\xi)d\xi. \end{aligned}$$

We therefore obtain the following result.

Theorem 24

Assume that $g \in C^2(\mathbb{R})$ and $h \in C^1(\mathbb{R})$. Then the d'Alembert formula

$$u(x,t) = \frac{1}{2}\left(g(x+t) + g(x-t)\right) + \frac{1}{2}\int_{x-t}^{x+t} h(\xi)d\xi$$

gives the unique classical solution of (45)

We next consider the Cauchy problem (41) in high dimensions.

- The general idea is to reduce the high dimensional problems to one-dimensional problem so that the d'Alembert formula can be used.

- This can be achieved by considering the spherical mean.
- Given $x \in \mathbb{R}^n$ and $r > 0$, we use $B_r(x)$ and $\partial B_r(x)$ to denote the ball of radius r with center x and its boundary respectively. Let ω_n denote the surface area of unit sphere, then

$$|\partial B_r(x)| = \omega_n r^{n-1} \quad \text{and} \quad |B_r(x)| = \frac{1}{n}\omega_n r^n.$$

- Let $u \in C^2(\mathbb{R}^n \times [0, \infty))$ be a solution of (41). For a fixed $x \in \mathbb{R}^n$, define

$$U(r, t; x) := \frac{1}{|\partial B_r(x)|} \int_{\partial B_r(x)} u(y, t) d\sigma(y), \quad r > 0$$

which is called the mean value of u over the sphere $\partial B_r(x)$ at time t.

- Notice that
$$\lim_{r\to 0} U(r,t;x) = u(x,t).$$
If we can find a formula for $U(r,t;x)$ for $r>0$, then we can obtain $u(x,t)$ by taking $r\to 0$.
- Write $U(r,t;x)$ as
$$U(r,t;x) = \frac{1}{\omega_n}\int_{|\xi|=1} u(x+r\xi,t)d\sigma(\xi).$$
Then
$$\begin{aligned}\partial_r U(r,t;x) &= \frac{1}{\omega_n}\int_{|\xi|=1} \nabla u(x+r\xi,t)\cdot\xi d\sigma(\xi)\\ &= \frac{1}{\omega_n r^{n-1}}\int_{\partial B_r(x)} \nabla u(y,t)\cdot\frac{y-x}{r}d\sigma(y).\end{aligned}$$

Since $(y-x)/r$ is the outward unit normal to $\partial B_r(x)$ at y, we may use the divergence theorem to derive

$$\partial_r U(r,t;x) = \frac{1}{\omega_n r^{n-1}} \int_{B_r(x)} \triangle u(y,t)dy.$$

- Using polar coordinates, we have

$$\partial_r U(r,t;x) = \frac{1}{\omega_n r^{n-1}} \int_0^r \int_{\partial B_\tau(x)} \triangle u(y,t)d\sigma(y)d\tau.$$

Consequently

$$\begin{aligned} &\partial_r^2 U(r,t;x) \\ &= \frac{1}{\omega_n r^{n-1}} \int_{\partial B_r(x)} \triangle u(y,t)d\sigma(y) - \frac{n-1}{\omega_n r^n} \int_{B_r(x)} \triangle u(y,t)dy. \end{aligned}$$

.

- By using $u_{tt} - \triangle u = 0$, we have

$$\partial_r^2 U(r,t;x) = \frac{1}{\omega_n r^{n-1}} \int_{\partial B_r(x)} u_{tt}(y,t)d\sigma(y) - \frac{n-1}{r}\partial_r U(r,t;x)$$
$$= \partial_t^2 U(r,t;x) - \frac{n-1}{r}\partial_r U(r,t;x).$$

- By the above expressions, we have

$$\begin{aligned} \lim_{r\to 0} U(r,t;x) &= u(x,t), \\ \lim_{r\to 0} U_r(r,t;x) &= 0, \\ \lim_{r\to 0} U_{rr}(r,t;x) &= \frac{1}{n}\triangle u(x,t). \end{aligned} \tag{46}$$

- Moreover, if u is a C^2 solution of (43), then, for fixed $x \in \mathbb{R}^n$, $U(r, t; x)$ as a function of (r, t) is in $C^2([0, \infty) \times [0, \infty))$ and satisfies the Euler-Poisson-Darboux equation

$$\begin{cases} U_{tt} - U_{rr} - \frac{n-1}{r} U_r = 0 & \text{for } r > 0,\ t > 0, \\ U = G, \quad U_t = H & \text{for } t = 0, \end{cases} \tag{47}$$

where

$$G(r; x) := \frac{1}{|\partial B_r(x)|} \int_{\partial B_r(x)} g(y) d\sigma(y),$$

$$H(r; x) := \frac{1}{|\partial B_r(x)|} \int_{\partial B_r(x)} h(y) d\sigma(y).$$

We hope to transform (47) into the usual 1D wave equation. This can be done easily when $n = 3$. So we consider this case first.

Case $n = 3$. We consider the Cauchy problem (43) of 3D wave equation. The Euler-Poisson-Darboux equation becomes

$$U_{tt} - U_{rr} - \frac{2}{r} U_r = 0.$$

Thus $\partial_r^2(rU) = \partial_t^2(rU)$. Let $\widetilde{U} = rU$, $\widetilde{G} = rG$ and $\widetilde{H} = rH$. Then

$$\begin{cases} \widetilde{U}_{tt} - \widetilde{U}_{rr} = 0 & \text{for } r > 0,\ t > 0, \\ \widetilde{U} = \widetilde{G}, \quad \widetilde{U}_t = \widetilde{H} & \text{at } t = 0 \text{ and } r > 0. \end{cases}$$

Moreover, in view of (46), we have

$$\widetilde{U} = 0,\ \widetilde{U}_r = u(x,t),\ \widetilde{U}_{rr} = 0 \quad \text{when } r = 0.$$

Thus, we may extend $\widetilde{U}$ to $\mathbb{R} \times [0, \infty)$ by odd reflection, i.e. we set

$$\overline{U}(r,t) = \begin{cases} \widetilde{U}(r,t;x), & r \geq 0,\ t \geq 0, \\ -\widetilde{U}(-r,t;x), & r < 0,\ t \geq 0. \end{cases}$$

Then $\overline{U} \in C^2(\mathbb{R} \times [0, \infty))$ and

$$\begin{cases} \overline{U}_{tt} - \overline{U}_{rr} = 0, & -\infty < r < \infty,\ t > 0, \\ \overline{U}(r,0) = \overline{G}(r), \quad \overline{U}_r(r,0) = \overline{H}(r), & -\infty < r < \infty, \end{cases}$$

where

$$\overline{G}(r) = \begin{cases} \widetilde{G}(r;x), & r \geq 0, \\ -\widetilde{G}(-r;x), & r < 0, \end{cases} \qquad \overline{H}(r) = \begin{cases} \widetilde{H}(r;x), & r \geq 0, \\ -\widetilde{H}(-r;x), & r < 0. \end{cases}$$

By the d'Alembert formula,

$$\overline{U}(r,t) = \frac{1}{2}\left(\overline{G}(r+t) + \overline{G}(r-t)\right) + \frac{1}{2}\int_{r-t}^{r+t} \overline{H}(s)ds.$$

Thus

$$\widetilde{U}(r,t;x) = \begin{cases} \frac{1}{2}\left(\widetilde{G}(r+t) + \widetilde{G}(r-t)\right) + \frac{1}{2}\int_{r-t}^{r+t} \widetilde{H}(s)ds, & r > t > 0, \\ \frac{1}{2}\left(\widetilde{G}(r+t) - \widetilde{G}(t-r)\right) + \frac{1}{2}\int_{t-r}^{t+r} \widetilde{H}(s)ds, & 0 \leq r \leq t. \end{cases}$$

Consequently, for $t > 0$ we have

$$u(x,t) = \lim_{r\to 0} \frac{1}{r}\widetilde{U}(r,t;x) = \widetilde{G}'(t) + \widetilde{H}(t).$$

Using the definition of $\widetilde{G}$ and $\widetilde{H}$, and the fact $|\partial B_r(x)| = 4\pi t^2$ in $\mathbb{R}^3$ we obtain

Theorem 25 (Kirchoff formula)

Let $g \in C^3(\mathbb{R}^3)$ and $h \in C^2(\mathbb{R}^3)$. Then

$$u(x,t) = \frac{\partial}{\partial t}\left(\frac{1}{4\pi t}\int_{|y-x|=t} g(y)d\sigma(y)\right) + \frac{1}{4\pi t}\int_{|y-x|=t} h(y)d\sigma(y)$$
$$= \frac{1}{4\pi t^2}\int_{|y-x|=t} \left(g(y) + \nabla g(y)\cdot(y-x) + th(y)\right) d\sigma(y)$$

gives the unique solution $u \in C^2(\mathbb{R}^3 \times [0,\infty))$ of the Cauchy problem (43) for 3D wave equation.

Case $n = 2$:

- The procedure for $n = 3$ does not work for 2D wave equations.
- We use the Hadamard's method of descent to derive the solution formula for 2D wave equation from the Kirchoff formula for 3D wave equation.
- Write $x = (x_1, x_2)$ and $\bar{x} = (x, x_3)$ and consider the Cauchy problem of the 3D wave equation

$$\begin{cases} U_{tt} - \triangle U - U_{x_3x_3} = 0 \quad \text{in } \mathbb{R}^3 \times (0, \infty), \\ U(\bar{x}, 0) = g(x), \quad U_t(\bar{x}, 0) = h(x), \quad \bar{x} \in \mathbb{R}^3, \end{cases}$$

where $\triangle$ denotes 2D Laplacian, i.e. $\triangle U = U_{x_1x_1} + U_{x_2x_2}$.

- By the Kirchoff formula,

$$U(x, x_3, t) = U(\bar{x}, t) = \frac{\partial}{\partial t}\left(\frac{1}{4\pi t}\int_{|\bar{y}-\bar{x}|=t} g(y)d\sigma(\bar{y})\right) + \frac{1}{4\pi t}\int_{|\bar{y}-\bar{x}|=t} h(y)d\sigma(\bar{y})$$

  where $y = (y_1, y_2)$ and $\bar{y} = (y, y_3)$. Since g and h do not depend on y_3, U is independent of x_3 and hence it is a solution of the Cauchy problem (43) of 2D wave equation.
- We simplify U by rewriting the two integrals over the sphere $|\bar{y} - \bar{x}| = t$.

- The sphere $|\bar{y}-\bar{x}|=t$ is a union of the two hemispheres

$$y_3=\phi_\pm(y):=x_3\pm\sqrt{t^2-|y-x|^2},$$

where $|y-x|\le t$. On both hemispheres, we have

$$d\sigma(\bar{y})=\sqrt{1+|\nabla\phi_\pm(y)|^2}dy=\frac{t}{\sqrt{t^2-|y-x|^2}}dy.$$

Therefore

$$U(x,t)=\frac{\partial}{\partial t}\left(\frac{1}{2\pi}\int_{|y-x|<t}\frac{g(y)}{\sqrt{t^2-|y-x|^2}}dy\right)+\frac{1}{2\pi}\int_{|y-x|<t}\frac{h(y)}{\sqrt{t^2-|y-x|^2}}dy.$$

This immediately gives the following result.

Theorem 26 (Poisson formula)

Let $g \in C^3(\mathbb{R}^2)$ and $h \in C^2(\mathbb{R}^2)$. Then

$$u(x,t) = \partial_t \left(\frac{t}{2\pi} \int_{|y|<1} \frac{g(x+ty)}{\sqrt{1-|y|^2}} dy \right) + \frac{t}{2\pi} \int_{|y|<1} \frac{h(x+ty)}{\sqrt{1-|y|^2}} dy$$
$$= \frac{1}{2\pi} \int_{|y-x|<t} \frac{g(y) + th(y) + \nabla g(y) \cdot (y-x)}{\sqrt{t^2 - |y-x|^2}} dy$$

gives the unique solution in $C^2(\mathbb{R}^2 \times [0,\infty))$ of the Cauchy problem (43) for 2D wave equation.

The procedures for $n = 2, 3$ can be extended to derive solution formulae of the Cauchy problems (43) for higher dimensional wave equations.

Since the procedure is lengthy and boring, we state the results without proofs.

Theorem 27

If $g \in C^{[n/2]+2}(\mathbb{R}^n)$ and $h \in C^{[n/2]+1}(\mathbb{R}^n)$, then (43) has a unique solution $u \in C^2([0,\infty) \times \mathbb{R}^n)$, where $[n/2]$ denotes the greatest integer not greater than $n/2$.

Moreover, if $n \geq 3$ is odd, then, with $\gamma_n = 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (n-2)$,

$$u(x,t) = \frac{1}{\gamma_n} \frac{\partial}{\partial t} \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{\frac{n-3}{2}} \left(\frac{t^{n-2}}{|\partial B_t(x)|} \int_{\partial B_t(x)} g d\sigma \right) + \frac{1}{\gamma_n} \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{\frac{n-3}{2}} \left(\frac{t^{n-2}}{|\partial B_t(x)|} \int_{\partial B_t(x)} h d\sigma \right)$$

while, if $n \geq 2$ is even, then, with $\gamma_n = 2 \cdot 4 \cdot \ldots \cdot (n-2) \cdot n$,

$$u(x,t) = \frac{1}{\gamma_n} \frac{\partial}{\partial t} \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{\frac{n-2}{2}} \left(\frac{t^n}{|B_t(x)|} \int_{B_t(x)} \frac{g(y)}{\sqrt{t^2 - |y-x|^2}} dy \right)$$
$$+ \frac{1}{\gamma_n} \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{\frac{n-2}{2}} \left(\frac{t^n}{|B_t(x)|} \int_{B_t(x)} \frac{h(y)}{\sqrt{t^2 - |y-x|^2}} d\sigma \right).$$

Remark.

- Given $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$. Theorem 22 shows that $u(x_0, t_0)$ is completely determined by the values of f and g in the ball $|x - x_0| \leq t_0$.
- When $n \geq 3$ is odd, by the solution formula this result can be strengthened: $u(t_0, x_0)$ depends only on the values of f and g (and derivatives) on the sphere $|x - x_0| = t_0$. This is called the Huygens' principle.

Duhamel Principle

We now consider the inhomogeneous problem (44), i.e.

$$\begin{cases} u_{tt} - \triangle u = f(x,t) & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = 0, \quad u_t(x,0) = 0, & x \in \mathbb{R}, \end{cases} \tag{48}$$

where $f \in C^{[n/2]+1}(\mathbb{R}^n \times [0,\infty))$. We use the Duhamel principle, i.e. for any $s \geq 0$, we first consider the homogeneous problem

$$\begin{cases} w_{tt} - \triangle w = 0 & \text{in } \mathbb{R}^n \times (s,\infty), \\ w = 0, \quad w_t = f(\cdot, s), & \text{when } t = s \end{cases} \tag{49}$$

which has a unique solution, denoted as $w(x,t;s)$; we then define

$$u(x,t) = \int_0^t w(x,t;s)ds. \tag{50}$$

The following result shows that u is the solution of (48).

Theorem 28

Let $f \in C^{[n/2]+1}(\mathbb{R}^n \times [0,\infty))$. Then the u defined by (50) is the unique solution of (48) in $C^2(\mathbb{R}^n \times [0,\infty))$.

Proof. Clearly $u(x,0) = 0$ and

$$u_t(x,t) = w(x,t;t) + \int_0^t w_t(x,t;s)ds = \int_0^t w_t(x,t;s)ds.$$

So $u(x,0) = 0$. Moreover

$$u_{tt}(x,t) = w_t(x,t;t) + \int_0^t w_{tt}(x,t;s)ds = f(x,t) + \int_0^t \triangle w(x,t;s)ds$$

$$= f(x,t) + \triangle \int_0^t w(x,t;s)ds = f(x,t) + \triangle u(x,t). \qquad \blacksquare$$

We conclude this section by giving the explicit solution formulae of (48) for $n = 1, 2, 3$.

- When $n = 1$, by the d'Alembert formula the solution of (49) is given by

$$w(x, t; s) = \frac{1}{2} \int_{x-(t-s)}^{x+(t-s)} f(y, s) dy.$$

Therefore the solution of (48) for $n = 1$ is given by

$$\begin{aligned} u(x, t) &= \frac{1}{2} \int_0^t \int_{x-(t-s)}^{x+(t-s)} f(y, s) dy ds \\ &= \frac{1}{2} \int_0^t \int_{x-\tau}^{x+\tau} f(y, t-\tau) dy d\tau. \end{aligned}$$

- When $n = 3$, by the Kirchoff formula the solution of (49) is

$$w(x,t;s) = \frac{1}{4\pi(t-s)} \int_{|y-x|=t-s} f(y;s)\,d\sigma(y).$$

Therefore, the solution of (48) is

$$\begin{aligned} u(x,t) &= \frac{1}{4\pi} \int_0^t \int_{|y-x|=t-s} \frac{f(y,s)}{t-s}\,d\sigma(y)\,ds \\ &= \frac{1}{4\pi} \int_0^t \int_{|y-x|=\tau} \frac{f(y,t-\tau)}{\tau}\,d\sigma(y)\,d\tau \\ &= \frac{1}{4\pi} \int_{|y-x|\le t} \frac{f(y,t-|y-x|)}{|y-x|}\,dy \end{aligned}$$

which is called the retarded potential because $u(x,t)$ depends on the values of f at the earlier times $t' = t - |y-x|$.

- When $n = 2$, by Poisson formula the solution of (49) is given by

$$w(x,t;s) = \frac{1}{2\pi} \int_{|y-x|<t-s} \frac{f(y,s)}{\sqrt{(t-s)^2 - |y-x|^2}} dy.$$

Therefore the solution of (48) is given by

$$\begin{aligned} u(x,t) &= \frac{1}{2\pi} \int_0^t \int_{|y-x|<t-s} \frac{f(y,s)}{\sqrt{(t-s)^2 - |y-x|^2}} dyds \\ &= \frac{1}{2\pi} \int_0^t \int_{|y-x|<\tau} \frac{f(y,t-\tau)}{\sqrt{\tau^2 - |y-x|^2}} dyd\tau. \end{aligned}$$

2. Local existence of semi-linear wave equations

- We will consider the Cauchy problem of semi-linear wave equation

$$\begin{cases} \Box u := u_{tt} - \triangle u = F(u, \partial u), & \text{in } \mathbb{R}^n \times (0, T], \\ u(x,0) = g(x), \quad u_t(x,0) = h(x), & x \in \mathbb{R}^n \end{cases} \tag{51}$$

where $\partial u = (\partial_t u, \nabla u)$ and $F \in C^\infty$ satisfies $F(0,0) = 0$.

- Under certain conditions on g and h, we will establish a local existence result, i.e. there is a small $T > 0$ such that (51) has a unique solution in $\mathbb{R}^n \times [0, T]$.
- The proof is based on the Picard iteration which defines a sequence $\{u_m\}$; the solution of (51) is obtained by the limit of this sequence.

- The sequence $\{u_m\}$ is defined by solving the Cauchy problem of linear wave equation

$$\begin{cases} \Box u_m = F(u_{m-1}, \partial u_{m-1}), & \text{in } \mathbb{R}^n \times (0, T], \\ u_m(x, 0) = g(x), \quad \partial_t u_m(x, 0) = h(x), & x \in \mathbb{R}^n \end{cases} \tag{52}$$

for $m = 0, 1, \cdots$, where we set $u_{-1} = 0$.
- So it is necessary to understand the Cauchy problems of linear wave equations deeper.
- We need some knowledge on Sobolev spaces.

2.1. The Sobolev spaces H^s

For any fixed $s \in \mathbb{R}$, $H^s := H^s(\mathbb{R}^n)$ denotes the completion of $C_0^\infty(\mathbb{R}^n)$ with respect to the norm

$$\|f\|_{H^s} := \left(\int_{\mathbb{R}^n} (1 + |\xi|^2)^s |\hat{f}(\xi)|^2 d\xi \right)^{1/2},$$

where $\hat{f}(\xi) := \int_{\mathbb{R}^n} e^{-i\langle x,\xi\rangle} f(x)dx$ is the Fourier transform of f.

- H^s is a Hilbert space and $H^0 = L^2$.
- If $s \geq 0$ is an integer, then $\|f\|_{H^s} \approx \sum_{|\alpha|\leq s} \|\partial^\alpha f\|_{L^2}$.
- $H^{s_2} \subset H^{s_1}$ for any $-\infty < s_1 \leq s_2 < \infty$.
- H^{-s} is the dual space of H^s for any $s \in \mathbb{R}$.
- If $s > k + n/2$ for some integer $k \geq 0$, then $H^s \hookrightarrow C^k(\mathbb{R}^n)$ compactly and there is a constant C_s such that

$$\sum_{|\alpha|\leq k} \|\partial^\alpha f\|_{L^\infty} \leq C_s \|f\|_{H^s}, \quad \forall f \in H^s.$$

- Given integer $k \geq 0$, $C^k([0,T], H^s)$ consists of functions $f(x,t)$ such that $t \to \|\partial_t^j f(\cdot,t)\|_{H^s}$ is continuous on $[0,T]$ for $j = 0, \cdots, k$. It is a Banach space under the norm

$$\sum_{j=0}^{k} \max_{0\leq t\leq T} \|\partial_t^j f(\cdot,t)\|_{H^s}.$$

- $L^1([0,T], H^s)$ consists of functions $f(x,t)$ such that

$$\int_0^T \|f(\cdot,t)\|_{H^s} dt < \infty.$$

2.1. Solutions of linear wave equations

Let $\Box = \partial_t^2 - \triangle$ denote the d'Alembertian. We first establish the following energy estimate.

Lemma 29

For any $u \in C^2(\mathbb{R}^n \times [0, T])$ *there holds*

$$\|\partial u(\cdot, t)\|_{L^2} \le \|\partial u(\cdot, 0)\|_{L^2} + \int_0^t \|\Box u(\cdot, \tau)\|_{L^2} d\tau, \quad 0 \le t \le T.$$

Proof. Fix $T_0 > T$ and consider the energy

$$E(t) := \int_{|x| \le T_0 - t} \left(|u_t(x,t)|^2 + |\nabla u(x,t)|^2\right) dx.$$

From the proof of Theorem 22 we have

$$\frac{d}{dt} E(t) \le 2 \int_{|x| \le T_0 - t} u_t(x,t) \Box u(x,t) dx.$$

By the Cauchy-Schwartz inequality we can obtain

$$\begin{aligned}\frac{d}{dt}E(t) &\le 2\left(\int_{|x|\le T_0-t}|u_t(x,t)|^2dx\right)^{1/2}\left(\int_{|x|\le T_0-t}|\Box u(x,t)|^2dx\right)^{1/2}\\ &= 2E(t)^{1/2}\|\Box u(\cdot,t)\|_{L^2(B_{T_0-t}(0))}.\end{aligned}$$

Therefore $\frac{d}{dt}E(t)^{1/2}\le\|\Box u(\cdot,t)\|_{L^2(B_{T_0-t}(0)}$. Consequently

$$\begin{aligned}\|\partial u(\cdot,t)\|_{L^2(B_{T_0-t}(0))} = E(t)^{1/2} &\le E(0)^{1/2}+\int_0^t\|\Box u(\cdot,\tau)\|_{L^2(B_{T_0-t}(0))}d\tau\\ &\le \|\partial u(\cdot,0)\|_{L^2}+\int_0^t\|\Box u(\cdot,\tau)\|_{L^2}d\tau.\end{aligned}$$

Letting $T_0\to\infty$ gives the desired inequality. ■

The energy estimate in Lemma 29 can be extended as follows.

Theorem 30

Let $u \in C^\infty(\mathbb{R}^n \times [0, T])$. Then, for any $s \in \mathbb{R}$, there is a constant C depending on T such that

$$\sum_{|\alpha|\leq 1} \|\partial^\alpha u(\cdot, t)\|_{H^s} \leq C \left(\sum_{|\alpha|\leq 1} \|\partial^\alpha u(\cdot, 0)\|_{H^s} + \int_0^t \|\Box u(\cdot, \tau)\|_{H^s} d\tau \right)$$

for $0 \leq t \leq T$.

Proof. Consider only $s \in \mathbb{Z}$. We may assume that the right hand side is finite. There are three cases to be considered.

Case 1: $s = 0$. We need to establish

$$\sum_{|\alpha|\leq 1} \|\partial^\alpha u(\cdot, t)\|_{L^2} \lesssim \sum_{|\alpha|\leq 1} \|\partial^\alpha u(\cdot, 0)\|_{L^2} + \int_0^t \|\Box u(\cdot, \tau)\|_{L^2} d\tau. \tag{53}$$

To see this, we first use Lemma 29 to obtain

$$\|\partial u(\cdot, t)\|_{L^2} \lesssim \|\partial u(\cdot, 0)\|_{L^2} + \int_0^t \|\Box u(\cdot, \tau)\|_{L^2} d\tau. \tag{54}$$

By the fundamental theorem of Calculus we can write

$$u(x, t) = u(x, 0) + \int_0^t u_t(x, \tau) d\tau.$$

Thus it follows from the Minkowski inequality that

$$\|u(\cdot,t)\|_{L^2} \leq \|u(\cdot,0)\|_{L^2} + \int_0^t \|u_t(\cdot,\tau)\|_{L^2}\,d\tau.$$

Adding this inequality to (54) gives

$$\sum_{|\alpha|\leq 1} \|\partial^\alpha u(\cdot,t)\|_{L^2} \lesssim \sum_{|\alpha|\leq 1} \|\partial^\alpha u(\cdot,0)\|_{L^2} + \int_0^t \|\Box u(\cdot,\tau)\|_{L^2}\,d\tau$$
$$+ \int_0^t \sum_{|\alpha|\leq 1} \|\partial^\alpha u(\cdot,\tau)\|_{L^2}\,d\tau.$$

An application of the Gronwall inequality then gives (53).

Case 2: $s \in \mathbb{N}$**.** Let β be any multi-index with $|\beta| \le s$. We apply (53) to $\partial_x^\beta u$ to obtain

$$\begin{aligned}\sum_{|\alpha|\le 1} \|\partial_x^\beta \partial^\alpha u(\cdot,t)\|_{L^2} &\lesssim \sum_{|\alpha|\le 1} \|\partial_x^\beta \partial^\alpha u(\cdot,t)\|_{L^2} + \int_0^t \|\Box \partial_x^\beta u(\cdot,\tau)\|_{L^2} d\tau \\ &\lesssim \sum_{|\alpha|\le 1} \|\partial_x^\beta \partial^\alpha u(\cdot,0)\|_{L^2} + \int_0^t \|\partial_x^\beta \Box u(\cdot,\tau)\|_{L^2} d\tau.\end{aligned}$$

Summing over all β with $|\beta| \le s$ we obtain

$$\sum_{|\alpha|\le 1} \|\partial^\alpha u(\cdot,t)\|_{H^s} \lesssim \sum_{|\alpha|\le 1} \|\partial^\alpha u(\cdot,0)\|_{H^s} + \int_0^t \|\Box u(\cdot,\tau)\|_{H^s} d\tau.$$

Case 3: $s \in -\mathbb{N}$. We consider

$$v(\cdot,t) := (I-\triangle)^s u(\cdot,t).$$

Since $-s \in \mathbb{N}$, we can apply the estimate established in Case 2 to v to derive that

$$\sum_{|\alpha|\le 1} \|\partial^\alpha v(\cdot,t)\|_{H^{-s}} \lesssim \sum_{|\alpha|\le 1} \|\partial^\alpha v(\cdot,0)\|_{H^{-s}} + \int_0^t \|\Box v(\cdot,\tau)\|_{H^{-s}} d\tau.$$

Since $\Box$ and $(I-\triangle)^s$ commute, we have

$$\Box v(\cdot,\tau) = (I-\triangle)^s \Box u(\cdot,\tau).$$

Therefore

$$\|\Box v(\cdot,\tau)\|_{H^{-s}} = \|\Box u(\cdot,\tau)\|_{H^s}.$$

Consequently

$$\sum_{|\alpha|\le 1} \|\partial^\alpha v(\cdot,t)\|_{H^{-s}} \lesssim \sum_{|\alpha|\le 1} \|\partial^\alpha v(\cdot,0)\|_{H^{-s}} + \int_0^t \|\Box u(\cdot,\tau)\|_{H^s} d\tau.$$

Since $\|\partial^\alpha v(\cdot,t)\|_{H^{-s}} = \|\partial^\alpha u(\cdot,t)\|_{H^s}$, the proof is complete. ■

We now prove the following existence and uniqueness result for the Cauchy problem of linear wave equation

$$\begin{cases} \Box u = f(x,t), & \text{in } \mathbb{R}^n \times (0,T], \\ u(x,0) = g(x), & \partial_t u(x,0) = h(x), \quad x \in \mathbb{R}^n \end{cases} \tag{55}$$

Theorem 31

If $g, h \in C^\infty(\mathbb{R}^n)$ *and* $f \in C^\infty(\mathbb{R}^n \times [0,T])$, *then (55) has a unique solution* $u \in C^\infty(\mathbb{R}^n \times [0,T])$. *If in addition there is* $s \in \mathbb{R}$ *such that*

$$g \in H^{s+1}(\mathbb{R}^n), \quad h \in H^s(\mathbb{R}^n) \quad \textit{and} \quad f \in L^1([0,T], H^s(\mathbb{R}^n)),$$

then

$$u \in C([0,T], H^{s+1}) \cap C^1([0,T], H^s)$$

and, for $0 \le t \le T$ *there holds the estimate*

$$\sum_{|\alpha|\le 1} \|\partial^\alpha u(\cdot,t)\|_{H^s} \lesssim \|g\|_{H^{s+1}} + \|h\|_{H^s} + \int_0^t \|f(\cdot,\tau)\|_{H^s} d\tau.$$

Proof. The existence and uniqueness follow from the previous chapter. The remaining part is a consequence of Theorem 30. ∎

2.2. Semi-linear wave equations

We next consider the semi-linear wave equation (51), i.e.

$$\begin{aligned} \Box u &= F(u, \partial u) \qquad \text{in } \mathbb{R}^n \times (0, T], \\ u(\cdot, 0) &= g, \quad u_t(\cdot, 0) = h, \end{aligned} \tag{56}$$

where $F \in C^\infty$ satisfies $F(0,0) = 0$.

- For this equation, there holds the finite propagation speed property, i.e. if $u \in C^2(\mathbb{R}^n \times [0, T])$ is a solution with $u(x,0) = u_t(x,0) = 0$ for $|x - x_0| \le t_0$, then $u \equiv 0$ in the backward light cone $\mathcal{C}_{x_0,t_0}$. (see Exercise)

Theorem 32

If $g, h \in C_0^\infty(\mathbb{R}^n)$, then there is a $T > 0$ such that (56) has a unique solution $u \in C_0^\infty(\mathbb{R}^n \times [0, T])$.

Proof. 1. We first prove uniqueness. Let u and $\tilde{u}$ be two solutions. Then $v := u - \tilde{u}$ satisfies

$$v_{tt} - \triangle v = R, \quad v(0, \cdot) = 0, \quad v_t(0, \cdot) = 0,$$

where $R := F(u, \partial u) - F(\tilde{u}, \partial \tilde{u})$. It is clear that

$$|R| \le C(|v| + |\partial v|).$$

In view of Theorem 30, we have

$$\sum_{|\alpha| \le 1} \|\partial^\alpha v(\cdot, t)\|_{L^2} \lesssim \int_0^t \|R(\cdot, \tau)\|_{L^2} d\tau \lesssim \int_0^t \sum_{|\alpha| \le 1} \|\partial^\alpha v(\cdot, \tau)\|_{L^2} d\tau.$$

By Gronwall inequality, $\sum_{|\alpha|\le 1} \|\partial^\alpha v\|_{L^2} = 0$. Thus $0 = v = u - \tilde{u}$.

2. Next we prove existence. We first fix an integer $s \ge n + 2$.

- We use the Picard iteration. Let $u_{-1} = 0$ and define u_m, $m \ge 0$, successively by

$$\begin{aligned} \Box u_m &= F(u_{m-1}, \partial u_{m-1}) \quad \text{in } \mathbb{R}^n \times (0, \infty), \\ u_m(\cdot, 0) &= g, \quad \partial_t u_m(\cdot, 0) = h. \end{aligned} \tag{57}$$

  By Theorem 31, all u_m are in $C^\infty(\mathbb{R}^n \times [0, \infty))$.

- For any index γ satisfying $|\gamma| \le s$ we have

$$\Box \partial^\gamma u_m = \partial^\gamma [F(u_{m-1}, \partial u_{m-1})].$$

- Therefore, it follows from Theorem 30 that

$$\sum_{|\beta|\leq 1} \|\partial^\beta \partial^\gamma u_m(\cdot, t)\|_{L^2}$$

$$\leq C_0 \left(\sum_{|\beta|\leq 1} \|\partial^\beta \partial^\gamma u_m(\cdot, 0)\|_{L^2} + \int_0^t \|\partial^\gamma [F(u_{m-1}, \partial u_{m-1})]\|_{L^2} d\tau \right)$$

for all γ with $|\gamma| \leq s$. Summing over all such γ gives

$$\sum_{|\alpha|\leq s+1} \|\partial^\alpha u_m(\cdot, t)\|_{L^2}$$

$$\leq C_0 \left(\sum_{|\alpha|\leq s+1} \|\partial^\alpha u_m(\cdot, 0)\|_{L^2} + \int_0^t \sum_{|\alpha|\leq s} \|\partial^\alpha [F(u_{m-1}, \partial u_{m-1})]\|_{L^2} d\tau \right)$$

- Let

$$A_m(t) := \sum_{|\alpha|\le s+1} \|\partial^\alpha u_m(\cdot, t)\|_{L^2}.$$

Then

$$A_m(t) \le C_0\Big(A_m(0) + \int_0^t \sum_{|\alpha|\le s} \|\partial^\alpha [F(u_{m-1}, \partial u_{m-1})]\|_{L^2} d\tau\Big).$$

By using (57) it is easy to show that

$$A_m(0) \le A_0, \quad m = 0, 1, \cdots$$

for some number A_0 independent of m; in fact we can take A_0 to be a multiple of $\|g\|_{H^{s+1}} + \|h\|_{H^s}$.

- Consequently

$$A_m(t) \le C_0\Big(A_0 + \int_0^t \sum_{|\alpha|\le s} \|\partial^\alpha[F(u_{m-1}, \partial u_{m-1})]\|_{L^2} d\tau\Big). \tag{58}$$

Step 1. We show that there is $0 < T \le 1$ independent of m such that

$$A_m(t) \le 2C_0A_0, \quad \forall 0 \le t \le T \text{ and } m = 0, 1, \cdots . \tag{59}$$

- We prove (59) by induction on m. Since $F(0,0) = 0$ and $u_{-1} = 0$, we can obtain (59) with $m = 0$ from (58). Next we assume that (59) is true for $m = k$ and show that it is also true for $m = k+1$. During the argument we will indicate the choice of T.

In view of (58), we have

$$A_{k+1}(t) \leq C_0\Big(A_0 + \int_0^t \sum_{|\alpha|\leq s} \|\partial^\alpha[F(u_k, \partial u_k)]\|_{L^2} d\tau\Big). \quad (60)$$

Observing that $\partial^\alpha[F(u_k, \partial u_k)]$ is the sum of the terms

$$a(u_k, \partial u_k)\partial^{\beta_1} u_k \cdots \partial^{\beta_l} u_k \partial^{\gamma_1}\partial u_k \cdots \partial^{\gamma_m}\partial u_k$$

where $|\beta_1| + \cdots + |\beta_l| + |\gamma_1| + \cdots + |\gamma_m| = |\alpha|$. Therefore $|\beta_j| \leq |\alpha|/2$ and $|\gamma_j| \leq |\alpha|/2$ except one of the multi-indices.

So $\partial^\alpha[F(u_k, \partial u_k)]$ is the sum of finitely many terms, each is a product of derivatives of u_k in which at most one factor where u_k is differentiated more than $|\alpha|/2 + 1 \leq s/2 + 1$ times.

For $\partial^\gamma u_k$ with $|\gamma| \leq s/2+1$, by Sobolev embedding we have for $r > n/2+1+s/2$ that

$$\sum_{|\gamma|\leq s/2+1} |\partial^\gamma u_k(x,t)| \leq C \sum_{|\gamma|\leq r} \|\partial^\gamma u_k(\cdot,t)\|_{L^2}.$$

Since $s \geq n+2$, we have $s+1 > n/2+1+s/2$ and thus by induction hypothesis

$$\sum_{|\gamma|\leq s/2+1} |\partial^\gamma u_k(x,t)| \leq C \sum_{|\gamma|\leq s+1} \|\partial^\gamma u_k(\cdot,t)\|_{L^2} \leq CA_k(t) \leq 2CC_0A_0. \tag{61}$$

Therefore

$$|\partial^\alpha[F(u_k, \partial u_k)]| \leq C_{A_0} \sum_{|\beta| \leq s+1} |\partial^\beta u_k|, \quad \forall |\alpha| \leq s.$$

Consequently, by the induction hypothesis, we have

$$\sum_{|\alpha| \leq s} \|\partial^\alpha[F(u_k, \partial u_k)]\|_{L^2} \leq C_{A_0} A_k(t) \leq C_{A_0}. \tag{62}$$

In view of (60), we obtain

$$A_{k+1}(t) \leq C_0 \left(A_0 + C_{A_0} t\right) \leq C_0 \left(A_0 + C_{A_0} T\right), \quad 0 \leq t \leq T.$$

So, by taking $0 < T \leq 1$ so small that $C_{A_0} T \leq A_0$, we obtain $A_{k+1}(t) \leq 2C_0 A_0$ for $0 \leq t \leq T$. This completes the proof of (59).

Step 2. Next we show that $\{u_m\}$ is convergent under the norm

$$|||u||| := \max_{0 \le t \le T} \sum_{|\alpha| \le s+1} \|\partial^\alpha u(\cdot, t)\|_{L^2}.$$

To this end, consider

$$E_m(t) := \sum_{|\alpha| \le s+1} \|\partial^\alpha (u_{m+1} - u_m)(\cdot, t)\|_{L^2}.$$

By the definition of $\{u_m\}$, we have

$$\begin{aligned} &\Box (u_{m+1} - u_m) = R_m \quad \text{in } \mathbb{R}^n \times (0, T], \\ &(u_{m+1} - u_m)|_{t=0} = 0, \quad \partial_t (u_{m+1} - u_m)|_{t=0} = 0, \end{aligned}$$

where

$$R_m := F(u_m, \partial u_m) - F(u_{m-1}, \partial u_{m-1}).$$

By the same argument for deriving (58), we obtain

$$E_m(t) \le C_0 \int_0^t \sum_{|\alpha| \le s} \|\partial^\alpha R_m(\cdot, \tau)\|_{L^2} d\tau.$$

By (59) and the similar argument for deriving (62) we have

$$\sum_{|\alpha| \le s} \|\partial^\alpha R_m(\cdot, t)\|_{L^2} \le C E_{m-1}(t).$$

Thus

$$E_m(t) \le C \int_0^t E_{m-1}(\tau) d\tau, \quad m = 1, 2, \cdots .$$

Consequently

$$E_m(t) \leq \frac{(Ct)^m}{m!} \sup_{0\leq t\leq T} E_0(t), \quad m = 0, 1, \cdots .$$

So $\sum_m E_m(t) \leq C_0$. Therefore $\{u_m\}$ converges to some function u under the norm $\|\!| \cdot \|\!|$. By Sobolev embedding, we can conclude $u_m \to u$ in $C^{s+[(1-n)/2]}(\mathbb{R}^n \times [0, T])$ and hence in $C^2(\mathbb{R}^n \times [0, T])$ since $s \geq n+2$. By taking $m \to \infty$ in (57) we obtain that u is a solution of (56).

Step 3. The T obtained in Step 1 depends on s. If we can show (59), i.e.

$$\sum_{|\alpha|\leq s+1} \|\partial^\alpha u_m(\cdot, t)\|_{L^2} \leq A_s, \quad 0 \leq t \leq T$$

for all $m = 0, 1, \cdots$ with $T > 0$ independent of s, then we can conclude that $u \in C^\infty(\mathbb{R}^n \times [0, T])$.

- We now fix $s_0 \geq n + 3$ and let $T > 0$ be such that

$$\max_{0 \leq t \leq T} \sum_{|\alpha| \leq s_0 + 1} \|\partial^\alpha u_m(\cdot, t)\|_{L^2} \leq C_0 < \infty, \quad m = 0, 1, \cdots$$

  and show that for all $s \geq s_0$ there holds

$$\max_{0 \leq t \leq T} \sum_{|\alpha| \leq s+1} \|\partial^\alpha u_m(t, \cdot)\|_{L^2} \leq C_s < \infty, \quad \forall m. \tag{63}$$

- We show (63) by induction on s. Assume that (63) is true for some $s \geq s_0$, we show it is also true with s replaced by $s + 1$.

By the induction hypothesis and Sobolev embedding,

$$\max_{(x,t)\in\mathbb{R}^n\times[0,T]} \sum_{|\alpha|\le s+1-[(n+2)/2]} |\partial^\alpha u_m(x,t)| \le A_s < \infty, \quad \forall m.$$

Since $s \ge n+3$, we have $[(s+4)/2] \le s+1-[(n+2)/2]$. So

$$\max_{(x,t)\in\mathbb{R}^n\times[0,T]} \sum_{|\alpha|\le (s+4)/2} |\partial^\alpha u_m(x,t)| \le A_s, \quad \forall m.$$

This is exactly (61) with s replaced by $s+2$. Same argument there can be used to derive that

$$\max_{0\le t\le T} \sum_{|\alpha|\le s+2} \|\partial^\alpha u_m(\cdot,t)\|_{L^2} \le C_{s+1} < \infty, \quad \forall m.$$

We complete the induction argument and obtain a C^∞ solution. ■

- The interval of existence for semi-linear wave equation could be very small.
- The following theorem gives a criterion on extending solutions which is important in establishing global existence results.

Theorem 33 (Continuation principle)

Assume that u be the solution of the Cauchy problem (56) with $g, h \in C_0^\infty(\mathbb{R}^n)$. Let

$$T_* := \sup \{ T > 0 : u \text{ satisfies (56) on } [0, T] \} .$$

If $T_* < \infty$, then

$$\sum_{|\alpha| \le (n+6)/2} |\partial^\alpha u(t,x)| \notin L^\infty(\mathbb{R}^n \times [0, T_*]). \quad (64)$$

Proof. Assume that (64) does not hold, then

$$\sup_{[0,T_*)\times\mathbb{R}^n} \sum_{|\alpha|\le(n+6)/2} |\partial^\alpha u(t,x)| \le C < \infty.$$

Applying the argument in deriving (59) we have

$$\sup_{\mathbb{R}^n\times[0,T_*)} \sum_{|\alpha|\le s_0+1} \|\partial^\alpha u(\cdot,t)\|_{L^2} \le C_0 < \infty$$

where $s_0 = n+3$. By the argument in Step 3 of the proof of Theorem 32 we obtain for all $s \ge s_0$ that

$$\sup_{[0,T_*)\times\mathbb{R}^n} \sum_{|\alpha|\le s+1} \|\partial^\alpha u(t,\cdot)\|_{L^2} \le C_s < \infty.$$

So u can be extend to $u \in C^\infty([0, T_*] \times \mathbb{R}^n)$.

Since $g, h \in C_0^\infty(\mathbb{R}^n)$, by the finite speed of propagation we can find a number R (possibly depending on T_*) such that $u(x,t) = 0$ for all $|x| \geq R$ and $0 \leq t < T_*$. Consequently

$$u(x, T_*) = \partial_t u(x, T_*) = 0 \quad \text{when } |x| \geq R.$$

Thus, $u(x, T_*)$ and $\partial_t u(x, T_*)$ are in $C_0^\infty(\mathbb{R}^n)$, and can be used as initial data at $t = T_*$ to extend u beyond T_* by theorem 32. This contradicts the definition of T_*. ■

3. Invariant vector fields in Minkowski space

First are some conventions. We will set

$$\mathbb{R}^{1+n} := \{(t, x) : t \in \mathbb{R} \text{ and } x \in \mathbb{R}^n\},$$

where t denotes the time and $x := (x^1, \cdots, x^n)$ the space variable. We sometimes write $t = x^0$ and use

$$\partial_0 = \frac{\partial}{\partial t} \quad \text{and} \quad \partial_j := \frac{\partial}{\partial x^j} \text{ for } j = 1, \cdots, n.$$

For any multi-index $\alpha = (\alpha_0, \cdots, \alpha_n)$ and any function $u(t, x)$ we write

$$|\alpha| := \alpha_0 + \alpha_1 + \cdots + \alpha_n \quad \text{and} \quad \partial^\alpha u := \partial_0^{\alpha_0} \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} u.$$

Given any function $u(t, x)$, we use

$$|\partial_x u|^2 := \sum_{j=1}^{n} |\partial_j u|^2 \quad \text{and} \quad |\partial u|^2 := |\partial_0 u|^2 + |\partial_x u|^2.$$

We will use Einstein summation convention: *any term in which an index appears twice stands for the sum of all such terms as the index assumes all of a preassigned range of values.*

- A Greek letter is used for index taking values $0, \cdots, n$.
- A Latin letter is used for index taking values $1, \cdots, n$.

For instance

$$b^\mu \partial_\mu u = \sum_{\mu=0}^{n} b^\mu \partial_\mu u \quad \text{and} \quad b^j \partial_j u = \sum_{j=1}^{n} b^j \partial_j u.$$

3.1. Vector fields and tensor fields

- We use $x = (x^0, x^1, \cdots, x^n)$ to denote the natural coordinates in $\mathbb{R}^{1+n}$, where $x^0 = t$ denotes time variable.
- A vector field X in $\mathbb{R}^{1+n}$ is a first order differential operator of the form

$$X = \sum_{i=0}^{n} X^\mu \frac{\partial}{\partial x^\mu} = X^\mu \partial_\mu,$$

  where X^μ are smooth functions. We will identify X with (X^μ).
- The collection of all vector fields on $\mathbb{R}^{1+n}$ is called the tangent space of $\mathbb{R}^{1+n}$ and is denoted by $T\mathbb{R}^{1+n}$.

- For any two vector fields $X = X^\mu \partial_\mu$ and $Y = Y^\mu \partial_\mu$, one can define the Lie bracket

$$[X, Y] := XY - YX.$$

Then

$$\begin{aligned}[X, Y] &= (X^\mu \partial_\mu)(Y^\nu \partial_\nu) - (Y^\nu \partial_\nu)(X^\mu \partial_\mu) \\ &= X^\mu Y^\nu \partial_\mu \partial_\nu + X^\mu (\partial_\mu Y^\nu) \partial_\nu - Y^\nu X^\mu \partial_\nu \partial_\mu - Y^\nu (\partial_\nu X^\mu) \partial_\mu \\ &= (X^\mu \partial_\mu Y^\nu - Y^\mu \partial_\mu X^\nu) \partial_\nu = (X(Y^\mu) - Y(X^\mu)) \partial_\mu.\end{aligned}$$

So $[X, Y]$ is also a vector field.

- A linear mapping $\eta : T\mathbb{R}^{1+n} \to \mathbb{R}$ is called a 1-form if

$$\eta(fX) = f\eta(X), \quad \forall f \in C^\infty(\mathbb{R}^{1+n}), X \in T\mathbb{R}^{1+n}.$$

For each $\mu = 0, 1, \cdots, n$, we can define the 1-form dx^μ by

$$dx^\mu(X) = X^\mu, \quad \forall X = X^\mu \partial_\mu \in T\mathbb{R}^{1+n}.$$

Then for any 1-form η we have

$$\eta(X) = X^\mu \eta(\partial_\mu) = \eta_\mu dx^\mu(X), \quad \text{where } \eta_\mu := \eta(\partial_\mu).$$

Thus any 1-form in $\mathbb{R}^{1+n}$ can be written as $\eta = \eta_\mu dx^\mu$ with smooth functions η_μ. We will identify η with (η_μ).

- A bilinear mapping $T : T\mathbb{R}^{1+n} \times T\mathbb{R}^{1+n} \to \mathbb{R}$ is called a (covariant) 2-tensor field if for any $f \in C^\infty(\mathbb{R}^{1+n})$ and $X, Y \in T\mathbb{R}^{1+n}$ there holds

$$T(fX, Y) = T(X, fY) = fT(X, Y).$$

It is called symmetric if $T(X,Y) = T(Y,X)$ for all vector fields X and Y.

- Let
$$(\mathbf{m}_{\mu\nu}) = \text{diag}(-1, 1, \cdots, 1)$$
be the $(1+n)\times(1+n)$ diagonal matrix. We define $\mathbf{m} : T\mathbb{R}^{1+n} \times T\mathbb{R}^{1+n} \to \mathbb{R}$ by
$$\mathbf{m}(X,Y) := \mathbf{m}_{\mu\nu} X^\mu Y^\nu$$
for all $X = X^\mu \partial_\mu$ and $Y = Y^\mu \partial_\mu$ in $T\mathbb{R}^{1+n}$. It is easy to check $\mathbf{m}$ is a symmetric 2-tensor field on $\mathbb{R}^{1+n}$. We call $\mathbf{m}$ the Minkowski metric on $\mathbb{R}^{1+n}$. Clearly
$$\mathbf{m}(X,X) = -\left(X^0\right)^2 + \left(X^1\right)^2 + \cdots + (X^n)^2 .$$

- A vector field X in $(\mathbb{R}^{1+n}, \mathbf{m})$ is called space-like, time-like, or null if

$$\mathbf{m}(X,X) > 0, \quad \mathbf{m}(X,X) < 0, \quad \text{or} \quad \mathbf{m}(X,X) = 0$$

respectively. Consider the three vector fields $X_1 = 2\partial_0 - \partial_1$, $X_2 = \partial_0 - \partial_1$ and $X_3 = \partial_0 - 2\partial_1$. Then X_1 is time-like, X_2 is null, and X_3 is space-like.

- In $(\mathbb{R}^{1+n}, \mathbf{m})$ we define the d'Alembertian

$$\Box = \mathbf{m}^{\mu\nu}\partial_\mu\partial_\nu, \quad \text{where } (\mathbf{m}^{\mu\nu}) := (\mathbf{m}_{\mu\nu})^{-1}.$$

In terms of the coordinates $(t, x^1, \cdots, x^n)$, $\Box = -\partial_t^2 + \triangle$, where $\triangle = \partial_1^2 + \cdots + \partial_n^2$.

3.2. Energy-momentum tensor

- In order to derive the general energy estimates related to $\Box u = 0$, we introduce the so called energy-momentum tensor.
- To see how to write down this tensor, we consider a vector field $X = X^\mu \partial_\mu$ with constant X^μ. Then for any smooth function u we have

$$\begin{aligned}(Xu)\Box u &= X^\rho \partial_\rho u \, \mathbf{m}^{\mu\nu} \partial_\mu \partial_\nu u \\ &= \partial_\mu \left(X^\rho \mathbf{m}^{\mu\nu} \partial_\nu u \partial_\rho u \right) - X^\rho \mathbf{m}^{\mu\nu} \partial_\mu \partial_\rho u \partial_\nu u.\end{aligned}$$

Using the symmetry of $(\mathbf{m}^{\mu\nu})$ we can obtain

$$X^\rho \mathbf{m}^{\mu\nu} \partial_\mu \partial_\rho u \partial_\nu u = \partial_\rho \left(\frac{1}{2} X^\rho \mathbf{m}^{\mu\nu} \partial_\mu u \partial_\nu u \right).$$

Therefore $(Xu)\Box u = \partial_\nu \left(Q[u]_\mu^\nu X^\mu \right)$, where

$$Q[u]_\mu^\nu = \mathbf{m}^{\nu\rho}\partial_\rho u \partial_\mu u - \frac{1}{2}\delta_\mu^\nu \left(\mathbf{m}^{\rho\sigma}\partial_\rho u \partial_\sigma u \right)$$

in which δ_μ^ν denotes the Kronecker symbol, i.e. $\delta_\mu^\nu = 1$ when $\mu = \nu$ and 0 otherwise.

- This motivates to introduce the symmetric 2-tensor

$$Q[u]_{\mu\nu} := \mathbf{m}_{\mu\rho} Q[u]_\nu^\rho = \partial_\mu u \partial_\nu u - \frac{1}{2}\mathbf{m}_{\mu\nu} \left(\mathbf{m}^{\rho\sigma}\partial_\rho u \partial_\sigma u \right)$$

which is called the energy-momentum tensor associated to $\Box u = 0$. Then for any vector fields X and Y we have

$$Q[u](X, Y) = (Xu)(Yu) - \frac{1}{2}\mathbf{m}(X, Y)\mathbf{m}(\partial u, \partial u)$$

- For a 1-form η in $(\mathbb{R}^{1+n}, \mathbf{m})$, its divergence is a function defined by

$$\text{div}\eta := \mathbf{m}^{\mu\nu}\partial_\mu \eta_\nu.$$

For a symmetric 2-tensor field T in $(\mathbb{R}^{1+n}, \mathbf{m})$, its divergence is a 1-form defined by

$$(\text{div}\, T)_\rho := \mathbf{m}^{\mu\nu}\partial_\mu T_{\nu\rho}.$$

- The divergence of the energy-momentum tensor is

$$\begin{aligned}
(\text{div} Q[u])_\rho &= \mathbf{m}^{\mu\nu}\partial_\mu Q[u]_{\nu\rho} \\
&= \mathbf{m}^{\mu\nu}\partial_\mu \left(\partial_\nu u \partial_\rho u - \frac{1}{2}\mathbf{m}_{\nu\rho}\left(\mathbf{m}^{\sigma\eta}\partial_\sigma u \partial_\eta u\right)\right) \\
&= \mathbf{m}^{\mu\nu}\partial_\mu\partial_\nu u \partial_\rho u = (\Box u)\partial_\rho u.
\end{aligned}$$

- Let X be a vector field. Using $Q[u]$ we can introduce the 1-form

$$P_\mu := Q[u]_{\mu\nu} X^\nu.$$

Then its divergence is

$$\begin{aligned}
\mathrm{div} P = \mathbf{m}^{\mu\nu}\partial_\mu P_\nu &= \mathbf{m}^{\mu\nu}\partial_\mu \left(\mathbf{Q}[u]_{\nu\rho} X^\rho\right) \\
&= \mathbf{m}^{\mu\nu}\partial_\mu Q[u]_{\nu\rho} X^\rho + \mathbf{m}^{\mu\nu} Q[u]_{\nu\rho}\partial_\mu X^\rho \\
&= (\mathrm{div} Q[u])_\rho X^\rho + \mathbf{m}^{\mu\nu} Q[u]_{\nu\rho}\partial_\mu X^\rho \\
&= \Box u\, \partial_\rho u\, X^\rho + \mathbf{m}^{\mu\nu} Q[u]_{\nu\rho}\mathbf{m}^{\rho\eta}\partial_\mu X_\eta \\
&= (\Box u) X u + \frac{1}{2} Q[u]^{\mu\rho}\left(\partial_\mu X_\rho + \partial_\rho X_\mu\right).
\end{aligned}$$

where $Q[u]^{\mu\nu} := \mathbf{m}^{\mu\rho}\mathbf{m}^{\sigma\nu} Q[u]_{\rho\sigma}$ and $X_\eta := \mathbf{m}_{\rho\eta} X^\rho$.

- For a vector field X, we define

$$^{(X)}\pi_{\mu\nu} := \partial_\mu X_\nu + \partial_\nu X_\mu$$

which is called the deformation tensor of X with respect to $\mathbf{m}$. Then we have

$$\text{div} P = \partial_\mu(\mathbf{m}^{\mu\nu} P_\nu) = (\Box u) X u + \frac{1}{2} Q[u]^{\mu\nu} \, {}^{(X)}\pi_{\mu\nu}. \qquad (65)$$

- Assume that u vanishes for large $|x|$ at each t. Then for any $t_0 < t_1$, we integrate $\text{div} P$ over $[t_0, t_1] \times \mathbb{R}^n$ and note that ∂_t is the upward unit normal to each slice $\{t\} \times \mathbb{R}^n$, we obtain

$$\iint\limits_{[t_0,t_1]\times\mathbb{R}^n} \text{div} P dx dt = \int\limits_{\{t=t_1\}} Q[u](X, \partial_t) dx - \int\limits_{\{t=t_0\}} Q[u](X, \partial_t) dx.$$

This together with (65) then implies

Theorem 34

Let $u \in C^2(\mathbb{R}^{1+n})$ that vanishes for large $|x|$ at each t. Then for any vector field X and $t_0 < t_1$ there holds

$$\int_{\{t=t_1\}} Q[u](X, \partial_t)dx = \int_{\{t=t_0\}} Q[u](X, \partial_t)dx + \iint_{[t_0,t_1]\times\mathbb{R}^n} (\Box u)Xudxdt + \frac{1}{2}\iint_{[t_0,t_1]\times\mathbb{R}^n} Q[u]^{\mu\nu}\, {}^{(X)}\pi_{\mu\nu}dxdt. \qquad (66)$$

- By choosing X suitably, many useful energy estimates can be derived from Theorem 34.

- For instance, we may take $X = \partial_t$ in Theorem 34. Notice that ${}^{(\partial_t)}\pi = 0$ and

$$Q[u](\partial_t, \partial_t) = \frac{1}{2}\left(|\partial_t u|^2 + |\nabla u|^2\right),$$

we obtain for $E(t) = \frac{1}{2}\int_{\{t\}\times\mathbb{R}^n}(|\partial_t u|^2 + |\nabla u|^2)dx$ the identity

$$E(t) = E(t_0) + \int_{t_0}^{t}\int_{\mathbb{R}^n} \Box u\, \partial_t u dx dt', \quad \forall t \geq t_0.$$

This implies that

$$\frac{d}{dt}E(t) = \int_{\{t\}\times\mathbb{R}^n} \Box u \partial_t u dx \leq \sqrt{2}\|\Box u(\cdot, t)\|_{L^2(\mathbb{R}^n)} E(t)^{1/2}.$$

Therefore

$$\frac{d}{dt}E(t)^{1/2} \le \frac{1}{\sqrt{2}}\|\Box u(\cdot, t)\|_{L^2(\mathbb{R}^n)}.$$

Consequently we obtain the energy estimate

$$E(t)^{1/2} \le E(t_0)^{1/2} + \frac{1}{\sqrt{2}}\int_{t_0}^{t} \|\Box u(\cdot, t')\|_{L^2(\mathbb{R}^n)} dt', \quad \forall t \ge t_0.$$

3.3. Killing vector fields

The identity (66) can be significantly simplified if ${}^{(X)}\pi = 0$. A vector field $X = X^\mu \partial_\mu$ in $(\mathbb{R}^{1+n}, \mathbf{m})$ is called a *Killing vector field* if ${}^{(X)}\pi = 0$, i.e.

$$\partial_\mu X_\nu + \partial_\nu X_\mu = 0 \qquad \text{in } \mathbb{R}^{1+n}.$$

Corollary 35

Let $u \in C^2(\mathbb{R}^{1+n})$ that vanishes for large $|x|$ at each t. Then for any Killing vector field X and $t_0 < t_1$ there holds

$$\int_{\{t=t_1\}} Q[u](X, \partial_t)dx = \int_{\{t=t_0\}} Q[u](X, \partial_t)dx + \iint_{[t_0,t_1]\times\mathbb{R}^n} (\Box u)Xudxdt.$$

- We can determine all Killing vector fields in $(\mathbb{R}^{1+n}, \mathbf{m})$. Write $\pi_{\mu\nu} = {}^{(X)}\pi_{\mu\nu}$, Then

$$\begin{aligned}
\partial_\rho \pi_{\mu\nu} &= \partial_\rho\partial_\mu X_\nu + \partial_\rho\partial_\nu X_\mu, \\
\partial_\mu \pi_{\nu\rho} &= \partial_\mu\partial_\nu X_\rho + \partial_\mu\partial_\rho X_\nu, \\
\partial_\nu \pi_{\rho\mu} &= \partial_\nu\partial_\rho X_\mu + \partial_\nu\partial_\mu X_\rho.
\end{aligned}$$

- Therefore

$$\partial_\mu \pi_{\nu\rho} + \partial_\nu \pi_{\rho\mu} - \partial_\rho \pi_{\mu\nu} = 2\partial_\mu \partial_\nu X_\rho.$$

If X is a Killing vector field, then ${}^{(X)}\pi = 0$ and hence

$$\partial_\mu \partial_\nu X_\rho = 0 \quad \text{for all } \mu, \nu, \rho.$$

Thus each X_ρ is an affine function, i.e. there are constants $a_{\rho\nu}$ and b_ρ such that

$$X_\rho = a_{\rho\nu} x^\nu + b_\rho.$$

Using ${}^{(X)}\pi = 0$ again we have

$$0 = \partial_\mu X_\nu + \partial_\nu X_\mu = a_{\nu\mu} + a_{\mu\nu}.$$

- Therefore $a_{\mu\nu} = -a_{\nu\mu}$ and thus

$$X = X^{\mu}\partial_{\mu} = \mathbf{m}^{\mu\nu}X_{\nu}\partial_{\mu} = \mathbf{m}^{\mu\nu}\left(a_{\nu\rho}x^{\rho} + b_{\nu}\right)\partial_{\mu}$$

$$= \sum_{\nu=0}^{n}\left(\sum_{\rho<\nu} + \sum_{\rho>\nu}\right) a_{\nu\rho}x^{\rho}\mathbf{m}^{\mu\nu}\partial_{\mu} + \mathbf{m}^{\mu\nu}b_{\nu}\partial_{\mu}$$

$$= \sum_{\nu=0}^{n}\sum_{\rho<\nu} a_{\nu\rho}x^{\rho}\mathbf{m}^{\mu\nu}\partial_{\mu} + \sum_{\rho=0}^{n}\sum_{\nu<\rho} a_{\nu\rho}x^{\rho}\mathbf{m}^{\mu\nu}\partial_{\mu} + \mathbf{m}^{\mu\nu}b_{\nu}\partial_{\mu}$$

$$= \sum_{\nu=0}^{n}\sum_{\rho<\nu} \left(a_{\nu\rho}x^{\rho}\mathbf{m}^{\mu\nu}\partial_{\mu} + a_{\rho\nu}x^{\nu}\mathbf{m}^{\mu\rho}\partial_{\mu}\right) + \mathbf{m}^{\mu\nu}b_{\nu}\partial_{\mu}$$

$$= \sum_{\nu=0}^{n}\sum_{\rho<\nu} a_{\nu\rho}\left(x^{\rho}\mathbf{m}^{\mu\nu}\partial_{\mu} - x^{\nu}\mathbf{m}^{\mu\rho}\partial_{\mu}\right) + \mathbf{m}^{\mu\nu}b_{\nu}\partial_{\mu}.$$

Thus we obtain the following result on Killing vector fields.

Proposition 36

Any Killing vector field in $(\mathbb{R}^{1+n}, \mathbf{m})$ *can be written as a linear combination of the vector fields* ∂_μ, $0 \le \mu \le n$ *and*

$$\Omega_{\mu\nu} = (\mathbf{m}^{\rho\mu} x^\nu - \mathbf{m}^{\rho\nu} x^\mu)\, \partial_\rho, \quad 0 \le \mu < \nu \le n.$$

- Since $(\mathbf{m}^{\mu\nu}) = \mathrm{diag}(-1, 1, \cdots, 1)$, the vector fields $\{\Omega_{\mu\nu}\}$ consist of the following elements

$$\Omega_{0i} = x^i \partial_t + t\partial_i, \quad 1 \le i \le n,$$
$$\Omega_{ij} = x^j \partial_i - x^i \partial_j, \quad 1 \le i < j \le n.$$

3.4. Conformal Killing vector fields

- When ${}^{(X)}\pi_{\mu\nu} = f\mathbf{m}_{\mu\nu}$ for some function f, the identity (66) can still be modified into a useful identity. To see this, we use (65) to obtain

$$\begin{aligned}\mathrm{div}P = \partial_\mu(\mathbf{m}^{\mu\nu}P_\nu) &= (\Box u)Xu + \frac{1}{2}f\mathbf{m}^{\mu\nu}Q[u]_{\mu\nu}\\ &= (\Box u)Xu + \frac{1-n}{4}f\mathbf{m}^{\mu\nu}\partial_\mu u\partial_\nu u.\end{aligned}$$

We can write

$$\begin{aligned}&f\mathbf{m}^{\mu\nu}\partial_\mu u\partial_\nu u = \mathbf{m}^{\mu\nu}\partial_\mu(fu\partial_\nu u) - \mathbf{m}^{\mu\nu}u\partial_\mu f\partial_\nu u - fu\Box u\\ &= \mathbf{m}^{\mu\nu}\partial_\mu(fu\partial_\nu u) - \mathbf{m}^{\mu\nu}\partial_\nu\left(\frac{1}{2}u^2\partial_\mu f\right) + \frac{1}{2}u^2\Box f - fu\Box u\\ &= \mathbf{m}^{\mu\nu}\partial_\mu\left(fu\partial_\nu u - \frac{1}{2}u^2\partial_\nu f\right) + \frac{1}{2}u^2\Box f - fu\Box u\end{aligned}$$

Consequently

$$\partial_\mu(\mathbf{m}^{\mu\nu}P_\nu) = (\Box u)Xu + \frac{1-n}{4}\mathbf{m}^{\mu\nu}\partial_\mu\left(fu\partial_\nu u - \frac{1}{2}u^2\partial_\nu f\right) + \frac{1-n}{8}u^2\Box f - \frac{1-n}{4}fu\Box u$$

Therefore, by introducing

$$\widetilde{P}_\mu := P_\mu + \frac{n-1}{4}fu\partial_\mu u - \frac{n-1}{8}u^2\partial_\mu f,$$

we obtain

$$\mathrm{div}\widetilde{P} = \partial_\mu(\mathbf{m}^{\mu\nu}\widetilde{P}_\nu) = \Box u\left(Xu + \frac{n-1}{4}fu\right) - \frac{n-1}{8}u^2\Box f.$$

By integrating over $[t_0, t_1] \times \mathbb{R}^n$ as before, we obtain

Theorem 37

If X is a vector field in $(\mathbb{R}^{1+n}, \mathbf{m})$ with ${}^{(X)}\pi = f\mathbf{m}$, then for any smooth function u vanishing for large $|x|$ there holds

$$\int_{t=t_1} \widetilde{Q}(X, \partial_t)dx = \int_{t=t_0} \widetilde{Q}(X, \partial_t)dx - \frac{n-1}{8} \iint_{[t_0,t_1]\times\mathbb{R}^n} u^2 \Box f dxdt + \iint_{[t_0,t_1]\times\mathbb{R}^n} \left(Xu + \frac{n-1}{4} fu \right) \Box u dxdt,$$

where $t_0 \leq t_1$ and

$$\widetilde{Q}(X, \partial_t) := Q[u](X, \partial_t) + \frac{n-1}{4} \left(fu\partial_t u - \frac{1}{2} u^2 \partial_t f \right).$$

- A vector field $X = X^\mu \partial_\mu$ in $(\mathbb{R}^{1+n}, \mathbf{m})$ is called conformal Killing if there is a function f such that ${}^{(X)}\pi = f\mathbf{m}$, i.e. $\partial_\mu X_\nu + \partial_\nu X_\mu = f\mathbf{m}_{\mu\nu}$.
- Any Killing vector field is conformal Killing. However, there are vector fields which are conformal Killing but not Killing.

  (i) Consider the vector field

$$L_0 = \sum_{\mu=0}^{n} x^\mu \partial_\mu = x^\mu \partial_\mu.$$

  we have $(L_0)^\mu = x^\mu$ and so $(L_0)_\mu = \mathbf{m}_{\mu\nu} x^\nu$. Consequently

$$\begin{aligned} {}^{(L_0)}\pi_{\mu\nu} &= \partial_\mu (L_0)_\nu + \partial_\nu (L_0)_\mu = \partial_\mu(\mathbf{m}_{\nu\eta} x^\eta) + \partial_\nu(\mathbf{m}_{\mu\eta} x^\eta) \\ &= \mathbf{m}_{\nu\eta}\delta_\mu^\eta + \mathbf{m}_{\mu\eta}\delta_\nu^\eta = 2\mathbf{m}_{\mu\nu}. \end{aligned}$$

  Therefore L_0 is conformal Killing and ${}^{(L_0)}\pi = 2\mathbf{m}$.

(ii) For each fixed $\mu = 0, 1, \cdots, n$ consider the vector field

$$K_\mu := 2\mathbf{m}_{\mu\nu} x^\nu x^\rho \partial_\rho - \mathbf{m}_{\eta\nu} x^\eta x^\nu \partial_\mu.$$

We have $(K_\mu)^\rho = 2\mathbf{m}_{\mu\nu} x^\nu x^\rho - \mathbf{m}_{\eta\nu} x^\eta x^\nu \delta_\mu^\rho$. Therefore

$$(K_\mu)_\rho = \mathbf{m}_{\rho\eta}(K_\mu)^\eta = 2\mathbf{m}_{\rho\eta}\mathbf{m}_{\mu\nu} x^\nu x^\eta - \mathbf{m}_{\rho\mu}\mathbf{m}_{\nu\eta} x^\nu x^\eta.$$

By direct calculation we obtain

$$^{(K_\mu)}\pi_{\rho\eta} = \partial_\rho (K_\mu)_\eta + \partial_\eta (K_\mu)_\rho = 4\mathbf{m}_{\mu\nu} x^\nu \mathbf{m}_{\rho\eta}.$$

Thus each K_μ is conformal Killing and $^{(K_\mu)}\pi = 4\mathbf{m}_{\mu\nu} x^\nu \mathbf{m}$. The vector field K_0 is due to Morawetz (1961).

All these conformal Killing vector fields can be found by looking at $X = X^\mu \partial_\mu$ with X^μ being quadratic.

- We can determine all conformal Killing vector fields in $(\mathbb{R}^{1+n}, \mathbf{m})$ when $n \geq 2$.

Proposition 38

Any conformal Killing vector field in $(\mathbb{R}^{1+n}, \mathbf{m})$ can be written as a linear combination of the vector fields

$$\begin{aligned}
&\partial_\mu, \qquad\qquad 0 \leq \mu \leq n,\\
&\Omega_{\mu\nu} = (\mathbf{m}^{\rho\mu} x^\nu - \mathbf{m}^{\rho\nu} x^\mu)\partial_\rho, \quad 0 \leq \mu < \nu \leq n,\\
&L_0 = \sum_{\mu=0}^{n} x^\mu \partial_\mu,\\
&K_\mu = \mathbf{m}_{\mu\nu} x^\nu x^\rho \partial_\rho - \mathbf{m}_{\rho\nu} x^\rho x^\nu \partial_\mu, \quad \mu = 0, 1, \cdots, n.
\end{aligned}$$

Proof. Let X be conformal Killing, i.e. there is f such that

$$^{(X)}\pi_{\mu\nu} := \partial_\mu X_\nu + \partial_\nu X_\mu = f\mathbf{m}_{\mu\nu}. \tag{67}$$

We first show that f is an affine function. Recall that

$$2\partial_\mu\partial_\nu X_\rho = \partial_\mu\pi_{\nu\rho} + \partial_\nu\pi_{\rho\mu} - \partial_\rho\pi_{\mu\nu}.$$

Therefore

$$2\partial_\mu\partial_\nu X_\rho = \mathbf{m}_{\nu\rho}\partial_\mu f + \mathbf{m}_{\rho\mu}\partial_\nu f - \mathbf{m}_{\mu\nu}\partial_\rho f.$$

This gives

$$2\Box X_\rho = 2\mathbf{m}^{\mu\nu}\partial_\mu\partial_\nu X_\rho = (1-n)\partial_\rho f. \tag{68}$$

In view of (67), we have

$$(n+1)f = 2\mathbf{m}^{\mu\nu}\partial_\mu X_\nu$$

This together with (68) gives

$$(n+1)\Box f = 2\mathbf{m}^{\mu\nu}\partial_\mu \Box X_\nu = (1-n)\mathbf{m}^{\mu\nu}\partial_\mu\partial_\nu f = (1-n)\Box f.$$

So $\Box f = 0$. By using again (68) and (67) we have

$$\begin{aligned}(1-n)\partial_\mu\partial_\nu f &= \frac{1-n}{2}\left(\partial_\mu\partial_\nu f + \partial_\nu\partial_\mu f\right) = \partial_\mu\Box X_\nu + \partial_\nu \Box X_\mu \\ &= \Box\left(\partial_\mu X_\nu + \partial_\nu X_\mu\right) = \mathbf{m}_{\mu\nu}\Box f = 0.\end{aligned}$$

Since $n \geq 2$, we have $\partial_\mu\partial_\nu f = 0$. Thus f is an affine function, i.e. there are constants a_μ and b such that $f = a_\mu x^\mu + b$.

Consequently

$$^{(X)}\pi = (a_\mu x^\mu + b)\mathbf{m}.$$

Recall that $^{(L_0)}\pi = 2\mathbf{m}$ and $^{(K_\mu)}\pi = 4\mathbf{m}_{\mu\nu}x^\nu\mathbf{m}$. Therefore, by introducing the vector field

$$\widetilde{X} := X - \frac{1}{2}bL_0 - \frac{1}{4}\mathbf{m}^{\mu\nu}a_\nu K_\mu,$$

we obtain

$$^{(\widetilde{X})}\pi = {}^{(X)}\pi - \frac{1}{2}b\,{}^{(L_0)}\pi - \frac{1}{4}\mathbf{m}^{\mu\nu}a_\nu\,{}^{(K_\mu)}\pi = 0.$$

Thus $\widetilde{X}$ is Killing. We may apply Proposition 36 to conclude that $\widetilde{X}$ is a linear combination of ∂_μ and $\Omega_{\mu\nu}$. The proof is complete. ■

4. Klainerman-Sobolev inequality

We turn to global existence of Cauchy problems for nonlinear wave equations

$$\Box u = F(u, \partial u).$$

This requires good decay estimates on $|u(t,x)|$ for large t. Recall the classical Sobolev inequality

$$|f(x)| \le C \sum_{|\alpha| \le (n+2)/2} \|\partial^\alpha f\|_{L^2}, \quad \forall x \in \mathbb{R}^n$$

which is very useful. However, it is not enough for the purpose. To derive good decay estimates for large t, one should replace ∂f by Xf with suitable vector fields X that exploits the structure of Minkowski space. This leads to Klainerman inequality of Sobolev type.

The formulation of Klainerman inequality involves only the constant vector fields

$$\partial_\mu, \quad 0 \le \mu \le n$$

and the homogeneous vector fields

$$\begin{aligned} L_0 &= x^\rho \partial_\rho, \\ \Omega_{\mu\nu} &= (\mathbf{m}^{\rho\mu} x^\nu - \mathbf{m}^{\rho\nu} x^\mu)\, \partial_\rho, \quad 0 \le \mu < \nu \le n. \end{aligned}$$

There are $m+1$ such vector fields, where $m = \frac{(n+1)(n+2)}{2}$. We will use Γ to denote any such vector field, i.e. $\Gamma = (\Gamma_0, \cdots, \Gamma_m)$ and for any multi-index $\alpha = (\alpha_0, \cdots, \alpha_m)$ we adopt the convention $\Gamma^\alpha = \Gamma_0^{\alpha_0} \cdots \Gamma_m^{\alpha_m}$.

It is now ready to state the Klainerman inequality of Sobolev type, which will be used in the proof of global existence.

Theorem 39 (Klainerman)

Let $u \in C^\infty([0,\infty) \times \mathbb{R}^n)$ vanish when $|x|$ is large. Then

$$(1+t+|x|)^{n-1}(1+|t-|x||)|u(t,x)|^2 \le C \sum_{|\alpha| \le \frac{n+2}{2}} \|\Gamma^\alpha u(t,\cdot)\|_{L^2}^2$$

for $t > 0$ and $x \in \mathbb{R}^n$, where C depends only on n.

We skip the proof of Theorem 39 since the argument is rather lengthy. Before using this result, deeper understanding on the vector fields Γ is necessary.

Lemma 40 (Commutator relations)

Among the vector fields ∂_μ, $\Omega_{\mu\nu}$ and L_0 we have the commutator relations:

$$\begin{aligned}
[\partial_\mu, \partial_\nu] &= 0, \\
[\partial_\mu, L_0] &= \partial_\mu, \\
[\partial_\rho, \Omega_{\mu\nu}] &= \left(\mathbf{m}^{\sigma\mu}\delta^\nu_\rho - \mathbf{m}^{\sigma\nu}\delta^\mu_\rho\right)\partial_\sigma, \\
[\Omega_{\mu\nu}, \Omega_{\rho\sigma}] &= \mathbf{m}^{\sigma\mu}\Omega_{\rho\nu} - \mathbf{m}^{\rho\mu}\Omega_{\sigma\nu} + \mathbf{m}^{\rho\nu}\Omega_{\sigma\mu} - \mathbf{m}^{\sigma\nu}\Omega_{\rho\mu}, \\
[\Omega_{\mu\nu}, L_0] &= 0.
\end{aligned}$$

Therefore, the commutator between ∂_μ and any other vector field is a linear combination of $\{\partial_\nu\}$, and the commutator of any two homogeneous vector fields is a linear combination of homogeneous vector fields.

Proof. These identity can be checked by direct calculation. As an example, we derive the formula for $[\Omega_{\mu\nu}, \Omega_{\rho\sigma}]$. Recall that

$$\Omega_{\mu\nu} = (\mathbf{m}^{\eta\mu} x^{\nu} - \mathbf{m}^{\eta\nu} x^{\mu})\, \partial_{\eta}.$$

Therefore

$$\begin{aligned}
[\Omega_{\mu\nu}, \Omega_{\rho\sigma}] &= \Omega_{\mu\nu} (\mathbf{m}^{\eta\rho} x^{\sigma} - \mathbf{m}^{\eta\sigma} x^{\rho})\, \partial_{\eta} - \Omega_{\rho\sigma} (\mathbf{m}^{\eta\mu} x^{\nu} - \mathbf{m}^{\eta\nu} x^{\mu})\, \partial_{\eta} \\
&= (\mathbf{m}^{\gamma\mu} x^{\nu} - \mathbf{m}^{\gamma\nu} x^{\mu}) \left(\mathbf{m}^{\eta\rho} \delta^{\sigma}_{\gamma} - \mathbf{m}^{\eta\sigma} \delta^{\rho}_{\gamma}\right) \partial_{\eta} \\
&\quad - (\mathbf{m}^{\gamma\rho} x^{\sigma} - \mathbf{m}^{\gamma\sigma} x^{\rho}) \left(\mathbf{m}^{\eta\mu} \delta^{\nu}_{\gamma} - \mathbf{m}^{\eta\nu} \delta^{\mu}_{\gamma}\right) \partial_{\eta} \\
&= \mathbf{m}^{\sigma\mu} (\mathbf{m}^{\eta\rho} x^{\nu} - \mathbf{m}^{\eta\nu} x^{\rho})\, \partial_{\eta} - \mathbf{m}^{\rho\mu} (\mathbf{m}^{\eta\sigma} x^{\nu} - \mathbf{m}^{\eta\nu} x^{\sigma})\, \partial_{\eta} \\
&\quad + \mathbf{m}^{\rho\nu} (\mathbf{m}^{\eta\sigma} x^{\mu} - \mathbf{m}^{\eta\mu} x^{\sigma})\, \partial_{\eta} - \mathbf{m}^{\sigma\nu} (\mathbf{m}^{\eta\rho} x^{\mu} - \mathbf{m}^{\eta\mu} x^{\rho})\, \partial_{\eta} \\
&= \mathbf{m}^{\sigma\mu} \Omega_{\rho\nu} - \mathbf{m}^{\rho\mu} \Omega_{\sigma\nu} + \mathbf{m}^{\rho\nu} \Omega_{\sigma\mu} - \mathbf{m}^{\sigma\nu} \Omega_{\rho\mu}.
\end{aligned}$$

This shows the result. ■

Lemma 41

For any $0 \leq \mu, \nu \leq n$ *there hold*

$$[\Box, \partial_\mu] = 0, \quad [\Box, \Omega_{\mu\nu}] = 0, \quad [\Box, L_0] = 2\Box$$

Consequently, for any multiple-index α *there exist constants* $c_{\alpha\beta}$ *such that*

$$\Box \Gamma^\alpha = \sum_{|\beta| \leq |\alpha|} c_{\alpha\beta} \Gamma^\beta \Box. \tag{69}$$

Proof. Direct calculation. ■

5. Global Existence in higher dimensions

We consider in $\mathbb{R}^{1+n}$ the global existence of the Cauchy problem

$$\begin{aligned} &\Box u = F(\partial u) \\ &u|_{t=0} = \varepsilon f, \qquad \partial_t u|_{t=0} = \varepsilon g, \end{aligned} \tag{70}$$

where $n \geq 4$, $\varepsilon \geq 0$ is a number, and $F : \mathbb{R}^{1+n} \to \mathbb{R}$ is a given C^∞ function which vanishes to the second order at the origin:

$$F(0) = 0, \quad \mathbf{D}F(0) = 0. \tag{71}$$

The main result is as follows.

Theorem 42

Let $n \geq 4$ and let $f, g \in C_c^\infty(\mathbb{R}^n)$. If F is a C^∞ function satisfying (71), then there exists $\varepsilon_0 > 0$ such that (70) has a unique solution $u \in C^\infty([0, \infty) \times \mathbb{R}^n)$ for any $0 < \varepsilon \leq \varepsilon_0$.

Proof. Let

$$T_* := \sup\{T > 0 : (70) \text{ has a solution } u \in C^\infty([0, T] \times \mathbb{R}^n)\}.$$

Then $T_* > 0$ by Theorem 33. We only need to show that $T_* = \infty$. Assume that $T_* < \infty$, then Theorem 33 implies

$$\sum_{|\alpha| \leq (n+6)/2} |\partial^\alpha u(t, x)| \notin L^\infty([0, T_*) \times \mathbb{R}^n).$$

We will derive a contradiction by showing that there is $\varepsilon_0 > 0$ such that for all $0 < \varepsilon \le \varepsilon_0$ there holds

$$\sup_{(t,x)\in[0,T_*)\times\mathbb{R}^n} \sum_{|\alpha|\le(n+6)/2} |\partial^\alpha u(t,x)| < \infty. \tag{72}$$

Step 1. We derive (72) by showing that there exist $A > 0$ and $\varepsilon_0 > 0$ such that

$$A(t) := \sum_{|\alpha|\le n+4} \|\partial\Gamma^\alpha u(t,\cdot)\|_{L^2} \le A\varepsilon, \quad 0 \le t < T_* \tag{73}$$

for $0 < \varepsilon \le \varepsilon_0$, where the sum involves all invariant vector fields ∂_μ, L_0 and $\Omega_{\mu\nu}$.

In fact, by Klainerman inequality in Theorem 39 we have for any multi-index β that

$$|\partial\Gamma^\beta u(t,x)| \le C(1+t)^{-\frac{n-1}{2}} \sum_{|\alpha|\le(n+2)/2} \|\Gamma^\alpha\partial\Gamma^\beta u(t,\cdot)\|_{L^2}.$$

Since $[\Gamma, \partial]$ is either 0 or $\pm\partial$, see Lemma 40, using (73) we obtain for $|\beta| \le (n+6)/2$ that

$$\begin{aligned} |\partial\Gamma^\beta u(t,x)| &\le C(1+t)^{-\frac{n-1}{2}} \sum_{|\alpha|\le n+4} \|\partial\Gamma^\alpha u(t,\cdot)\|_{L^2} \\ &= C(1+t)^{-\frac{n-1}{2}} A(t) \\ &\le CA\varepsilon(1+t)^{-\frac{n-1}{2}}. \end{aligned} \tag{74}$$

To estimate $|\Gamma^\beta u(t,x)|$, we need further property of u. Since $f, g \in C_0^\infty(\mathbb{R}^n)$, we can choose $R > 0$ such that

$$f(x) = g(x) = 0 \quad \text{for } |x| \geq R.$$

By the finite speed of propagation,

$$u(t,x) = 0, \quad \text{if } 0 \leq t < T_* \text{ and } |x| \geq R + t.$$

To show (72), it suffices to show that

$$\sup_{0 \leq t < T_*, |x| \leq R+t} |\Gamma^\alpha u(t,x)| < \infty, \quad \forall |\alpha| \leq (n+6)/2.$$

For any (t,x) satisfying $0 \le t < T_*$ and $|x| < R+t$, write $x = |x|\omega$ with $|\omega| = 1$. Then

$$\begin{aligned}
&\Gamma^\alpha u(t,x) = \Gamma^\alpha u(t,|x|\omega) - \Gamma^\alpha u(t,(R+t)\omega) \\
&= \int_0^1 \partial_j \Gamma^\alpha u(t,(s|x|+(1-s)(R+t))\omega) ds \; (|x|-R-t)\omega^j.
\end{aligned}$$

In view of (74), we obtain for all $|\alpha| \le (n+6)/2$ that

$$\begin{aligned}
|\Gamma^\alpha u(t,x)| &\le CA\varepsilon(1+t)^{-\frac{n-1}{2}}(R+t-|x|) \\
&\le CA\varepsilon(1+t)^{-\frac{n-3}{2}}.
\end{aligned}$$

Step 2. We prove (73).

- Since $u \in C^\infty([0, T_*) \times \mathbb{R}^n)$ and $u(t,x) = 0$ for $|x| \geq R + t$, we have $A(t) \in C([0, T_*))$.
- Using initial data we can find a large number A such that

$$A(0) \leq \frac{1}{4} A\varepsilon. \tag{75}$$

  By the continuity of $A(t)$, there is $0 < T < T_*$ such that $A(t) \leq A\varepsilon$ for $0 \leq t \leq T$.
- Let

$$T_0 = \sup\{T \in [0, T_*) : A(t) \leq A\varepsilon, \forall 0 \leq t \leq T\}.$$

  Then $T_0 > 0$. It suffices to show $T_0 = T_*$.

We show $T_0 = T_*$ be a contradiction argument. If $T_0 < T_*$, then $A(t) \leq A\varepsilon$ for $0 \leq t \leq T_0$. We will prove that for small $\varepsilon > 0$ there holds

$$A(t) \leq \frac{1}{2}A\varepsilon \quad \text{for } 0 \leq t \leq T_0.$$

By the continuity of $A(t)$, there is $\delta > 0$ such that

$$A(t) \leq A\varepsilon \quad \text{for } 0 \leq t \leq T_0 + \delta$$

which contradicts the definition of T_0.

Step 3. It remains only to prove that there is $\varepsilon_0 > 0$ such that

$$A(t) \leq A\varepsilon \text{ for } 0 \leq t \leq T_0 \Longrightarrow A(t) \leq \frac{1}{2}A\varepsilon \text{ for } 0 \leq t \leq T_0$$

for $0 < \varepsilon \leq \varepsilon_0$.

By Klainerman inequality and $A(t) \leq A\varepsilon$ for $0 \leq t \leq T_0$, we have for $|\beta| \leq (n+6)/2$ that

$$|\partial \Gamma^\beta u(t,x)| \leq CA\varepsilon(1+t)^{-\frac{n-1}{2}}, \quad \forall (t,x) \in [0, T_0] \times \mathbb{R}^n. \quad (76)$$

To estimate $\|\partial \Gamma^\alpha u(t,\cdot)\|_{L^2}$ for $|\alpha| \leq n+4$, we use the energy estimate to obtain

$$\|\partial \Gamma^\alpha u(t,\cdot)\|_{L^2} \leq \|\partial \Gamma^\alpha u(0,\cdot)\|_{L^2} + C\int_0^t \|\Box \Gamma^\alpha u(\tau,\cdot)\|_{L^2} d\tau. \quad (77)$$

We write

$$\Box \Gamma^\alpha u = [\Box, \Gamma^\alpha] u + \Gamma^\alpha(F(\partial u))$$

and estimate $\|\Gamma^\alpha(F(\partial u))(\tau,\cdot)\|_{L^2}$ and $\|[\Box, \Gamma^\alpha] u(\tau,\cdot)\|_{L^2}$.

Since $F(0) = \mathbf{D}F(0) = 0$, we can write

$$F(\partial u) = \sum_{j,k=1}^{n} F_{jk}(\partial u)\partial_j u \partial_k u,$$

where F_{jk} are smooth functions. Using this it is easy to see that $\Gamma^\alpha(F(\partial u))$ is a linear combination of following terms

$$F_{\alpha_1 \cdots \alpha_m}(\partial u) \cdot \Gamma^{\alpha_1}\partial u \cdot \Gamma^{\alpha_2}\partial u \cdot \cdots \cdot \Gamma^{\alpha_m}\partial u$$

where $m \geq 2$, $F_{\alpha_1 \cdots \alpha_m}$ are smooth functions and $|\alpha_1| + \cdots + |\alpha_m| = |\alpha|$ with at most one α_i satisfying $|\alpha_i| > |\alpha|/2$ and at least one α_i satisfying $|\alpha_i| \leq |\alpha|/2$.

- In view of (76), by taking ε_0 such that $A\varepsilon_0 \leq 1$, we obtain $\|F_{\alpha_1 \cdots \alpha_m}(\partial u)\|_{L^\infty} \leq C$ for $0 < \varepsilon \leq \varepsilon_0$ with a constant C independent of A and ε.

- Since $|\alpha|/2 \leq (n+4)/2$, using (76) all terms $\Gamma^{\alpha_j}\partial u$, except the one with largest $|\alpha_j|$, can be estimated as

$$\|\Gamma^{\alpha_j}\partial u(t,x)\|_{L^\infty([0,T_0]\times\mathbb{R}^n)} \leq CA\varepsilon(1+t)^{-\frac{n-1}{2}}$$

Therefore

$$\begin{aligned}\|\Gamma^\alpha(F(\partial u))(t,\cdot)\|_{L^2} &\leq CA\varepsilon(1+t)^{-\frac{n-1}{2}}\sum_{|\beta|\leq|\alpha|}\|\Gamma^\beta\partial u(t,\cdot)\|_{L^2}\\ &\leq CA\varepsilon(1+t)^{-\frac{n-1}{2}}A(t). \qquad (78)\end{aligned}$$

Recall that $[\Box,\Gamma]$ is either 0 or $2\Box$. Thus

$$|[\Box,\Gamma^\alpha]u| \lesssim \sum_{|\beta|\leq|\alpha|}|\Gamma^\beta\Box u| \lesssim \sum_{|\beta|\leq|\alpha|}|\Gamma^\beta(F(\partial u))|.$$

Therefore

$$\begin{aligned}\|[\Box, \Gamma^\alpha]u(t,\cdot)\|_{L^2} &\le C \sum_{|\beta|\le|\alpha|} \|\Gamma^\beta(F(\partial u))(t,\cdot)\|_{L^2} \\ &\le CA\varepsilon(1+t)^{-\frac{n-1}{2}}A(t). \qquad (79)\end{aligned}$$

Consequently, it follows from (77), (78) and (79) that

$$\|\partial\Gamma^\alpha u(t,\cdot)\|_{L^2} \le \|\partial\Gamma^\alpha u(0,\cdot)\|_{L^2} + CA\varepsilon \int_0^t \frac{A(\tau)}{(1+\tau)^{\frac{n-1}{2}}} d\tau$$

Summing over all α with $|\alpha| \le n+4$ we obtain

$$A(t) \le A(0) + CA\varepsilon \int_0^t \frac{A(\tau)}{(1+\tau)^{\frac{n-1}{2}}} d\tau \le \frac{1}{4}A\varepsilon + CA\varepsilon \int_0^t \frac{A(\tau)}{(1+\tau)^{\frac{n-1}{2}}} d\tau.$$

By Gronwall inequality,

$$A(t) \le \frac{1}{4} A\varepsilon \exp\left(CA\varepsilon \int_0^t \frac{d\tau}{(1+\tau)^{(n-1)/2}} \right), \quad 0 \le t \le T_0.$$

For $n \ge 4$, $\int_0^\infty \frac{d\tau}{(1+\tau)^{(n-1)/2}} = \frac{2}{n+2} < \infty$. (This is the reason we need $n \ge 4$ for global existence). We now choose $\varepsilon_0 > 0$ so that

$$\exp\left(\frac{2}{n+2} CA\varepsilon_0 \right) \le 2.$$

Thus $A(t) \le A\varepsilon/2$ for $0 \le t \le T_0$ and $0 < \varepsilon \le \varepsilon_0$. The proof is complete. ■

Remark. The proof does not provide global existence result when $n \leq 3$ in general. However, the argument can guarantee existence on some interval $[0, T_\varepsilon]$, where T_ε can be estimated as

$$T_\varepsilon \geq \begin{cases} e^{c/\varepsilon}, & n = 3, \\ c/\varepsilon^2, & n = 2, \\ c/\varepsilon, & n = 1. \end{cases} \tag{80}$$

In fact, let $A(t)$ be defined as before, the key point is to show that, for any $T < T_\varepsilon$,

$$A(t) \leq A\varepsilon \text{ for } 0 \leq t \leq T \Longrightarrow A(t) \leq \frac{1}{2} A\varepsilon \text{ for } 0 \leq t \leq T$$

The same argument as above gives

$$A(t) \leq \frac{1}{4} A\varepsilon \exp\left(CA\varepsilon \int_0^t \frac{d\tau}{(1+\tau)^{(n-1)/2}} \right), \quad 0 \leq t \leq T.$$

Thus we can improve the estimate to $A(t) \leq \frac{1}{2} A\varepsilon$ for $0 \leq t \leq T$ if T_ε satisfies

$$\exp\left(CA\varepsilon \int_0^{T_\varepsilon} \frac{d\tau}{(1+\tau)^{(n-1)/2}} \right) \leq 2$$

When $n \leq 3$, the maximal T_ε with this property satisfies (80).

Remark. For $n = 2$ or $n = 3$, the above argument can guarantee global existence when F satisfies stronger condition

$$F(0) = 0, \quad \mathbf{D}F(0) = 0, \quad \cdots, \quad \mathbf{D}^k F(0) = 0, \tag{81}$$

where $k = 5 - n$. Indeed, this condition guarantees that $F(\partial u)$ is a linear combination of the terms

$$F_{j_1 \cdots j_{k+1}}(\partial u)\partial_{j_1} u \cdots \partial_{j_{k+1}} u.$$

Thus $\Gamma^\alpha(F(\partial u))$ is a linear combination of the terms

$$f_{i_1 \cdots i_r}(\partial u)\Gamma^{\alpha_{i_1}}\partial u \cdot \ldots \cdot \Gamma^{\alpha_{i_r}}\partial u,$$

where $r \geq k + 1$, $|\alpha_1| + \cdots + |\alpha_r| = |\alpha|$ and $f_{i_1 \cdots i_r}$ are smooth functions; there are at most one α_i satisfying $\alpha_i > |\alpha|/2$ and at least k of α_i satisfying $|\alpha_i| \leq |\alpha|/2$.

We thus can obtain

$$\|\Gamma^\alpha(F(\partial u))(t,\cdot)\|_{L^2} \leq CA\varepsilon(1+t)^{-\frac{(n-1)k}{2}}A(t),$$
$$\|[\Box,\Gamma^\alpha]u(t,\cdot)\|_{L^2} \leq CA\varepsilon(1+t)^{-\frac{(n-1)k}{2}}A(t).$$

Therefore

$$A(t) \leq \frac{1}{4}A\varepsilon \exp\left(CA\varepsilon\int_0^t \frac{d\tau}{(1+\tau)^{((n-1)k)/2}}\right).$$

Since $k = 5-n$, $\int_0^\infty \frac{d\tau}{(1+\tau)^{((n-1)k)/2}}$ converges for $n=2$ or $n=3$.

The condition (81) is indeed too restrictive. In next lecture we relax it to include quadratic terms when $n=3$ using the so-called null condition introduced by Klainerman.