MATH 8110 - THEORY OF PARTIAL DIFFERENTIAL EQUATIONS

MARCELO M. DISCONZI[*]

Contents

[*]Vanderbilt University, Nashville, TN, USA. `marcelo.disconzi@vanderbilt.edu`.

1. Abbreviations

- ODE = ordinary differential equation
- PDE = partial differential equation
- LHS = left hand side
- RHS = right hand side
- w.r.t = with respect to
- $\Rightarrow$ = implies
- $\square$ = end of proof
- LHS := RHS means that the LHS is defined by the RHS.
- nd (e.g. 1d, 2d, ...) = n dimensional
- DO = differential operator
- iff = if and only if
- A $\subset$ B = A is a subset of B
- A $\subset\subset$ B = A is compactly contained in B
- Ω = domain in $\mathbb{R}^n$, unless stated otherwise

2. Notation

Unless stated otherwise, the following notation will be used throughout.

2.1. **Basic notation.** We denote by $\{x^i\}_{i=1}^n$ rectangular coordinates in $\mathbb{R}^n$. For problems involving a time variable t, we denote $(t,x) \in \mathbb{R} \times \mathbb{R}^n$, and let $\{x^\mu\}_{\mu=0}^n$ be rectangular coordinates in $\mathbb{R} \times \mathbb{R}^n$, where $x^0 := t$. Latin indices range from 1 to n and Greek indices from 0 to n. Naturally, when we say coordinates in $\mathbb{R}^n$ it could be in a subset $\mathbb{R}^n$ etc. Sometimes we write $\mathbb{R}^{1+n}$ to emphasize to spacetime structure $\mathbb{R} \times \mathbb{R}^n$, but for many general discussions we simply write $\mathbb{R}^n$.

Repeated indices, with one index up and one down, are summed over their range. E.g.:

$$u^i \sigma_i = \sum_{i=1}^{n} u^i \sigma_i$$

2.2. **Multi-index notation.** We denote:

$$\partial_\mu = \frac{\partial}{\partial x^\mu}, \quad \partial_i = \frac{\partial}{\partial x^i}$$

Definition 2.1. A vector of the form $\alpha = (\alpha_0, ..., \alpha_n)$, where each α_μ, $\mu = 0, ..., n$ is a non-negative integer, is called a **multi-index** of order $|\alpha| := \alpha_0 + ... + \alpha_n$.

We similarly write $\alpha = (\alpha_1, ..., \alpha_n)$ when t is not present, and also sometimes write $\vec{\alpha} = (\alpha_1, ..., \alpha_n)$, calling $\vec{\alpha}$ a spatial multi-index.

Given a multi-index α, denote

$$D^\alpha u := \frac{\partial^{|\alpha|} u}{\partial (x^0)^{\alpha_0} \partial (x^1)^{\alpha_1} ... \partial (x^n)^{\alpha_n}},$$

where $u = u(t, x^1, ..., x^n)$. If k is a non-negative integer,

$$D^k u = \{ D^\alpha u \mid |\alpha| = k \},$$
$$|D^k u| := \sqrt{\sum_{|\alpha|=k} |D^\alpha u|^2}$$

We can identify Du with the gradient of u and $D^2 u$ with the Hessian of u. These definitions have a natural interpretation when $u = u(x^1, ..., x^n)$ (so the multi-indices are spatial).

Let α, β be multi-indices. Define

$$\alpha! := \alpha_0! \alpha_1! ... \alpha_n!,$$
$$x^\alpha := (x^0)^{\alpha_0} (x^1)^{\alpha_1} ... (x^n)^{\alpha_n},$$
$$\alpha \leq \beta \text{ means } \alpha_\mu \leq \beta_\mu, \text{ where } \mu = 0, ..., n,$$
$$\binom{\alpha}{\beta} := \frac{\alpha!}{\beta! (\alpha - \beta)}, \text{ where } \beta \leq \alpha,$$
$$\binom{|\alpha|}{\alpha} := \frac{|\alpha|!}{\alpha!}$$

Then we have:

Multi-nomial theorem:

$$(x^0 + x^1 + ... + x^n)^k = \sum_{|\alpha|=k} \binom{|\alpha|}{\alpha!} x^\alpha$$

Product rule:

$$D^\alpha (uv) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^\beta u D^{\alpha - \beta} v$$

Taylor's formula:

$$u(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} D^\alpha u(0) x^\alpha + \mathcal{O}(|x|^{k+1})$$

where we recall the **big-oh notation**:

$$f = \mathcal{O}(g) \text{ as } x \to x_0$$

if there exists a constant $C > 0$ such that

$$|f(x)| \leq C |g(x)|$$

for all x sufficiently close to x_0. Many times x_0 is clear from the context and we write simply $f = \mathcal{O}(g)$.

Remark 2.2. Many times we will provide a definition, introduce a concept, etc., that has a natural generalization to a context studied later on. In these cases, such natural generalizations will be taken for granted.

3. Introduction

PDEs are essentially a generalization of ODEs for functions of several variables.

Definition 3.1. Let Ω be an open set in $\mathbb{R}^n$. We denote by $C^\infty(\Omega, \mathbb{R}^m)$ the set of all infinitely many times differentiable (i.e., smooth) maps $u : \Omega \to \mathbb{R}^m$. We put $C^\infty := C^\infty(\Omega, \mathbb{R})$ (although we can abuse notation and write $C^\infty(\Omega)$ for $C^\infty(\Omega, \mathbb{R}^m)$ if $\mathbb{R}^m$ is clear from the context). We also extend the notation to $C^\infty(\Omega, \mathbb{C}^m)$ etc.

Definition 3.2. Let $\Omega \subset \mathbb{R}^n$ be an open set. A **differential operator** $\mathcal{P}$ on Ω is a map $\mathcal{P} : \mathcal{U} \to C^\infty(\Omega)$, where $\mathcal{U} \subset C^\infty(\Omega)$, of the form

$$(\mathcal{P}u)(x) = P(D^k u(x), D^{k-1}u(x), ..., Du(x), u(x), x),$$

where $x \in \Omega, u \in \mathcal{U}$, and P is a function

$$P : \mathbb{R}^{n^k} \times \mathbb{R}^{n^{k-1}} \times ... \times \mathbb{R}^n \times \mathbb{R} \times \Omega \to \mathbb{R}.$$

The number k above is called the order of the operator. We often identify $\mathcal{P}$ with P and say “the differential operator P.”

Remark 3.3.

- In the above definition, it is implicitly assumed that the first entry in P is not trivial, so that the order of P is well defined. Otherwise, we could take, say, the first order operator $\mathcal{P}u = \partial_x u$ and think of it as the second order operator $Pu = 0 \cdot \partial_x^2 u + \partial_x u$, etc.
- P might in fact be defined only on a subset of $C^\infty(\Omega)$, e.g., $Pu = \frac{1}{\partial_x u}$ is not defined on constants. Situations like this will typically be clear from the context.
- We can generalize the above to $C^\infty(\Omega, \mathbb{R}^m)$.
- Differential operators will naturally extend to more general function spaces we will introduce later on.

Example 3.4. Take $\Omega = \mathbb{R}^2$. Then

$$Pu = \partial_x^2 u + \partial_y^2 u + u^2$$

is a second-order DO. To identify the function P, denote coordinates in $\mathbb{R}^{2^2} \times \mathbb{R}^2 \times \mathbb{R} \times \Omega$ by $z = (p_{xx}, p_{xy}, p_{yx}, p_{yy}, p_x, p_y, p, x, y)$, so $P(z) = p_{xx} + p_{yy} + p^2$.

Observe that the definition of a DO takes all entries into account, ignoring, e.g., $\partial_{xy} u = \partial_{yx} u$ etc.

Definition 3.5. Let P be a DO of order k. P is called

- **linear**, if it has the form

$$(Pu)(x) = \sum_{|\alpha| \le k} a_\alpha(x) D^\alpha u(x),$$

  for some functions a_α.

- **semi-linear**, if it has the form
$$(Pu)(x) = \sum_{|\alpha|=k} a_\alpha(x) D^\alpha(x) + a_0(D^{k-1}u(x), ..., Du(x), u(x), x),$$
for some a_α.
- **quasi-linear** if it has the form
$$(Pu)(x) = \sum_{|\alpha|=k} a_\alpha(D^{k-1}u(x), ..., Du(x), x) D^\alpha u(x) + a_0(D^{k-1}u(x), ..., Du(x), u(x), x).$$
- P is **fully nonlinear** if it depends nonlinearly on derivatives of order k.

Examples will be given below.

Definition 3.6. Let $\Omega \subset \mathbb{R}^n$ be an open set and P a DO of order k in Ω. An equation of the form
$$Pu = 0$$
for an unknown function u is called a $k^{\mathbf{th}}$ **order PDE** in Ω. The PDE is quasi-linear, etc., according to the character of P. In the linear case, we also consider the situation where a function $f : \Omega \to \mathbb{R}$ is given, and call the PDE $Pu = f$ inhomogeneous and $Pu = 0$ homogeneous. A **solution** to a PDE is a function $u : \Omega \to \mathbb{R}$ that satisfies the equation $Pu = 0$.

Similarly, we can define systems of PDEs (we abuse terminology and sometimes call a system of PDEs "a" PDE).

Given a PDE, we are typically interested in questions of the form

- Does a solution exist?
- If solutions exist, are they unique?
- If solutions exist, what are their properties?

4. Examples of PDEs

Laplace's equation:
$$\Delta u = 0,$$
where $\Delta := \frac{\partial^2}{\partial(x^1)^2} + \frac{\partial^2}{\partial(x^2)^2} + ... + \frac{\partial^2}{\partial(x^n)^2}$ is the **Laplacian** operator.

Helmholtz's equation:
$$\Delta u + \lambda u = 0$$

Linear transport equation:
$$\partial_t u + b^i \partial_i u = 0$$

Heat or diffusion equation:
$$\partial_t u - \Delta u = 0$$

Schrödinger's equation:
$$i\partial_t u + \Delta u = 0$$

Wave equation:
$$\Box u = 0$$
where $\Box = -\partial_t^2 + \Delta$ is the D'Alembertian or wave operator.

Eikonal equation:
$$|Du| = 1$$

Minimal surface equation:

$$\operatorname{div}\left(\frac{Du}{(1+|Du|^2)^{\frac{1}{2}}}\right) = 0$$

Burgers' equation:

$$\partial_t u + u\partial_x u = 0$$

Maxwell's equations:

$$\partial_t E - \operatorname{curl} B = 0$$
$$\partial_t B + \operatorname{curl} E = 0$$
$$\operatorname{div} B = \operatorname{div} E = 0$$

Euler's equations for incompressible fluids:

$$\partial_t u + (u \cdot \nabla)u = -\nabla p,$$
$$\operatorname{div} u = 0,$$

$u \cdot \nabla = u^i \frac{\partial}{\partial x^i}$.

Navier-Stokes equations for incompressible fluids:

$$\partial_t u + (u \cdot \nabla)u = -\nabla p + \Delta u,$$
$$\operatorname{div} u = 0$$

Euler's equations for compressible fluids:

$$\partial_t \rho + \operatorname{div}(\rho u) = 0$$
$$\partial_t u + (u \cdot \nabla)u = -\frac{1}{\rho}\nabla p$$
$$p = p(\rho)$$

Vacuum Einstein's equations:

$$\mathrm{Ric}_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} + \Lambda g_{\alpha\beta} = 0,$$

where g is a Lorentzian metric, Ric is the Ricci curvature of g, R the scalar curvature of g, and Λ is a constant (the cosmological constant).

Matter Einstein's equations:

$$\mathrm{Ric}_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} + \Lambda g_{\alpha\beta} = T_{\alpha\beta}$$

where $T_{\alpha\beta}$ is the energy-momentum tensor containing information about matter fields (e.g., electromagnetic fields).

5. Laplace's and Poisson's equation

We are going to study **Laplace's equation**

$$\Delta u = 0,$$

and its non-homogeneous version, **Poisson's equation**

$$\Delta u = f$$

Definition 5.1. A function satisfying Laplace's equation is called a **harmonic function.**

5.1. **Fundamental solution.** To solve $\Delta u = f$, we first consider $\Delta u = 0$ and try the Ansatz

$$u(x) = v(r),$$

$r = |x|$. Computing,

$$\partial_i u(x) = v'(r)\frac{x^i}{r},$$

$$\partial_i^2 u(x) = v''(r)\frac{(x^i)^2}{r^2} + v'(r)\left(\frac{1}{r} - \frac{(x^i)^2}{r^3}\right),$$

and summing:

$$\Delta u = v'' + \frac{n-1}{r}v'(r) \ (n = \text{dimension})$$

Thus $\Delta u = 0$ gives a ODE for v with solution

$$v(r) = \begin{cases} A\ln r + B, & n = 2, \\ \frac{A}{r^{n-2}} + B, & n \geq 3, \end{cases}$$

A, B constants. This motivates the definition:

Definition 5.2. The function

$$\Gamma(x) := \begin{cases} \frac{1}{2\pi}\ln|x|, & n = 2, \\ \frac{1}{n(2-n)\omega_n}\frac{1}{|x|^{n-2}}, & n \geq 3, \end{cases}$$

is called the **fundamental solution to Laplace's equation.** Above, ω_n = volume of $B_1(0)$ in $\mathbb{R}^n$.

Note that

$$|D\Gamma(x)| \leq \frac{C}{|x|^{n-1}}, \ \ |D^2\Gamma(x)| \leq \frac{C}{|x|^n},$$

where above and throughout we adopt the following:

Notation 5.3. We use $C > 0$ to denote a generic constant (depending on fixed data in a given problem) that can vary line-by-line. Unless stated otherwise, $\Omega \subset \mathbb{R}^n$ is an open set.

Definition 5.4. Let $\Omega \subset \mathbb{R}^n$ be an open set. We denote by $C^k(\Omega)$ the space of k-times continuously differentiable functions in Ω and by $C_c^k(\Omega)$ the space of those $u \in C^k(\Omega)$ with compact support.

Theorem 5.5. *Let* $f \in C_c^2(\mathbb{R}^n)$. *Set*

$$u(x) = \int_{\mathbb{R}^n} \Gamma(x-y)f(y)dy$$

Then:

(i) $u \in C^2(\mathbb{R}^n)$,
(ii) $\Delta u = f$ *in* $\mathbb{R}^n$.

Proof. Note that u is well defined ($\Gamma dy \sim r^{2-n}r^{n-1}dr$). Changing variables:

$$u(x) = \int_{\mathbb{R}^n} \Gamma(y)f(x-y)dy,$$

so

$$\frac{u(x+he_i) - u(x)}{h} = \int_{\mathbb{R}^n} \Gamma(y)\left(\frac{f(x+he_i-y) - f(x-y)}{h}\right)dy$$
$$\to \int \Gamma(y)\frac{\partial f}{\partial x^i}(x-y)dy \text{ as } h \to 0$$

since the difference quotient converges uniformly to $\partial_i f$. Similarly, we obtain D^2u, whose continuity follows from that of $D^2 f$.

Fix a small $\epsilon > 0$ and write

$$\begin{aligned}\Delta u &= \int_{B_\epsilon(0)} \Gamma(y)\Delta_x f(x-y)dy + \int_{\mathbb{R}^n\setminus B_\epsilon(0)} \Gamma(y)\Delta_x f(x-y)dy \\ &=: I_1 + I_2.\end{aligned}$$

$$|I_1| \le C\|D^2 f\|_{L^\infty(\mathbb{R}^n)} \int_{B_\epsilon(0)} |\Gamma(y)|dy \le C \begin{cases} \epsilon^2|\ln\epsilon| \\ \epsilon^2 \end{cases} \underset{\epsilon\to 0}{\to} 0$$

Note that $\Delta_x f(x-y) = \Delta_y f(x-y)$. Thus, integrating by parts:

$$\begin{aligned}I_2 &= -\int_{\mathbb{R}^n\setminus B_\epsilon(0)} \nabla\Gamma(y)\cdot\nabla_y f(x-y)dy + \int_{\partial B_\epsilon(0)} \Gamma(y)\frac{\partial f}{\partial \nu}(x-y)d\sigma(y) \\ &= I_{21} + I_{22}.\end{aligned}$$

$I_{22} \to 0$ as $\epsilon \to 0$, $\Gamma(y) \sim \log\epsilon$, ϵ^{2-n}, $d\sigma \sim \epsilon^{n-1}$. By parts again:

$$I_{21} = \int_{\mathbb{R}^n\setminus B_\epsilon(0)} \Delta\Gamma(y) f(x-y)dy - \int_{\partial B_\epsilon(0)} \frac{\partial\Gamma(y)}{\partial\nu} f(x-y)d\sigma(y).$$

Note that $\Delta\Gamma(y) = 0$ away from $y = 0$.

$$\frac{\partial\Gamma}{\partial\nu}(y) = \nabla\Gamma(y)\cdot\nu = \frac{1}{n\omega_n}\frac{y}{|y|^n}\cdot\nu = \frac{-1}{n\omega_n|y|^{n-1}},$$

where $\nu = -\frac{y}{|y|}$.

$$\begin{aligned}I_{21} &= \int_{\partial B_\epsilon(0)} \frac{1}{n\omega_n\epsilon^{n-1}} f(x-y)d\sigma(y) = \frac{1}{\mathrm{vol}(B_\epsilon(x))}\int_{\partial B_\epsilon(x)} f(y)d\sigma(y) \\ &\to f(x) \text{ as } \epsilon \to 0\end{aligned}$$

□

Remark 5.6. Note that u above is not unique (need boundary conditions at ∞).

5.2. **Properties of harmonic functions.**

Theorem 5.7. ***(Mean-value formulas).*** *If $u \in C^2(\Omega)$ is harmonic, then*

$$u(x) = \frac{1}{vol(\partial B_r(x))}\int_{\partial B_r(x)} u d\sigma = \frac{1}{vol(B_r(x))}\int_{B_r(x)} u dy$$

for any ball $B_r(x) \subset\subset \Omega$.

Proof. Changing variables $z = \frac{(y-x)}{r}$

$$\frac{1}{n\omega_n r^{n-1}}\int_{\partial B_r(x)} u(y)d\sigma(y) = \frac{1}{n\omega_n}\int_{\partial B_r(0)} u(x+rz)d\sigma(z) =: f(r)$$

$$f'(r) = \int_{\partial B_r(0)} \nabla u(x+rz)\cdot z d\sigma(z) = \frac{1}{n\omega_n r^{n-1}}\int_{\partial B_1(x)} \nabla u(y)\cdot\frac{y-x}{r}d\sigma(y)$$

Note that $\nu = \frac{y-x}{r}$.

$$= \frac{1}{n\omega_n r^{n-1}}\int_{\partial B_1(x)} \frac{\partial u}{\partial\nu}d\sigma = \frac{1}{n\omega_n r^{n-1}}\int_{B_r(x)} \Delta u dy = 0$$

$$\implies f(r) = f(0) = \lim_{r\to 0} \frac{1}{n\omega_n r^{n-1}} \int_{\partial B_r(x)} u(y)d\sigma(y) = u(x).$$

For the other equality:

$$\int_{B_r(x)} udy = \int_0^r \int_{\partial B_\tau(x)} ud\sigma d\tau = \omega_n r^n u(x).$$

Note that $\int_{\partial B_r(x)} ud\sigma = u(x)n\omega_n\tau^{n-1}$.

□

Theorem 5.8. ***(Converse of the mean value).*** *If $u \in C^2(\Omega)$ satisfies*

$$u(x) = \frac{1}{n\omega_n r^{n-1}} \int_{B_r(x)} ud\sigma$$

for $B_r(x) \subset\subset \Omega$, then u is harmonic.

Proof. If $\Delta u(x) \neq 0$, then $\Delta u > 0$ in some $B_r(x)$ contradicting $f'(r) = 0$.

□

Recall that a **mollifier** can be constructed as

$$\varphi(x) = \begin{cases} Ae^{-\frac{1}{1-|x|^2}}, & |x| < 1, \\ 0, & |x| \geq 1 \end{cases}$$

where A is a constant such that $\int_{\mathbb{R}^n} \varphi = 1$.
Then, $\text{supp}(\varphi) \subset \overline{B_1(0)}$ and $\varphi \in C^\infty(\mathbb{R}^n)$.
For $\epsilon > 0$, we put

$$\varphi_\epsilon(x) := \frac{1}{\epsilon^n}\varphi(\frac{x}{\epsilon})$$

so $\varphi_\epsilon \in C^\infty(\mathbb{R}^n)$, $\text{supp}(\varphi_\epsilon) \subset \overline{B_\epsilon(0)}$, and $\int_{\mathbb{R}} \varphi_\epsilon = 1$. If $u \in L^1_{\text{loc}}(\Omega)$, the function

$$u_\epsilon := \varphi_\epsilon * u$$

(the regularization of u) is defined in

$$\Omega_\epsilon := \{x \in \Omega |\ \text{dist}(x, \partial\Omega) > \epsilon\}$$

and $u_\epsilon \in C^\infty(\Omega_\epsilon)$. Moreover, $u_\epsilon \to u$ a.e., $u_\epsilon \to u$ in $C^0_{\text{loc}}(\Omega)$, $L^\infty_{\text{loc}}(\Omega)$ if $u \in C^0(\Omega)$, $L^\infty_{\text{loc}}(\Omega), 1 \leq p < \infty$.

Theorem 5.9. *Harmonic functions are C^∞.*

Proof. Let u be harmonic in Ω and put $u_\epsilon := u * \varphi_\epsilon$. Then, by the mean value property:

$$\begin{aligned} u_\epsilon(x) &= \int_\Omega \varphi_\epsilon(x-y)u(y)dy = \frac{1}{\epsilon^n}\int_{B_\epsilon(x)} \varphi\left(\frac{x-y}{\epsilon}\right)u(y) \\ &= \frac{1}{\epsilon^n}\int_0^\epsilon \varphi\left(\frac{r}{\epsilon}\right)\int_{\partial B_r(x)} ud\sigma dr = \frac{u(x)}{\epsilon^n}\int_0^\epsilon \varphi\left(\frac{r}{\epsilon}\right) n\omega_n r^{n-1}dr \\ &= u(x)\int_{B_\epsilon(0)} \varphi_\epsilon dy = u(x). \end{aligned}$$

Note that $n\omega_n r^{n-1} = \int_{\partial B_r(0)} d\sigma$.

□

Remark 5.10. One can show that harmonic functions are, in fact, analytic.

Using the mean value formula, we can derive decay estimates (as $|x| \to \infty$) for harmonic functions, leading to:

Theorem 5.11. ***(Liouville's Theorem).*** *If $u : \mathbb{R}^n \to \mathbb{R}$ is harmonic and bounded, then it is constant.*

Proof. Exercise. □

Theorem 5.12. ***(Maximum principle).*** *Suppose that $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ is harmonic in Ω, where Ω is bounded. Then $\max_{\overline{\Omega}} u = \max_{\partial\Omega} u$. Moreover, if Ω is connected and $u(x_0) = \max_{\overline{\Omega}} u$ for some $x_0 \in \Omega$, then u is constant.*

Proof. The first claim is implied by the second, which we prove. Say $u(x_0) = M = \max_{\overline{\Omega}} u$ and let r be such that $0 < r < \text{dist}(x_0, \partial\Omega)$.

By mean value,

$$M = u(x_0) = \frac{1}{\omega_n r^n} \int_{B_r(x)} u dy \leq M,$$

so $u = M$ in $B_r(x_0)$. Thus, $\{x \in \Omega | u(x) = M\}$ is open and closed in Ω. □

Remark 5.13. We need the boundedness assumption, e.g., $u(x, y) = y$ in $\mathbb{R}^2_+$.

Remark 5.14. Changing $u \mapsto -u$, we also get the minimum principle.

Corollary 5.15. *There exists at most one $C^2(\Omega) \cap C^0(\overline{\Omega})$ solution to*

$$\Delta u = f \text{ in } \Omega,$$
$$u = g \text{ on } \partial\Omega$$

with $f \in C^0(\Omega), g \in C^0(\partial\Omega)$.

Exercise: Look up Harnack inequality for Δ.

5.3. **Green's function.** The Green function is the analogue of the fundamental solution in the case of a **boundary-value problem**, i.e., a PDE plus a boundary condition.

In this section, we assume Ω to be bounded and with a C^1-boundary. We are interested in

$$\Delta u = f \text{ in } \Omega,$$
$$u = g \text{ on } \partial\Omega$$

where f and g are given.

Given a $C^2(\overline{\Omega})$ function u, we have (Green's identity):

$$\int_{\Omega \setminus B_\epsilon(x)} (u(y) \Delta y \Gamma(y - x) - \Gamma(y - x) \Delta u(y)) dy$$

$$= \int_{\partial(\Omega \setminus B_\epsilon(x))} (u(y) \frac{\partial \Gamma}{\partial v_y}(y - x) - \Gamma(y - x) \frac{\partial u}{\partial v_a}(y)) d\sigma(y)$$

Observe the following facts:

- $\Delta_y \Gamma(y - x) = 0$ in $\Omega \setminus B_\epsilon(x)$.
- $\int_{\partial B_\epsilon(x)} \Gamma(y - x) \frac{\partial u}{\partial v} d\sigma(y) \to 0$ as $\epsilon \to 0$.
- $\int_{\partial B_\epsilon(x)} u(y) \frac{\partial \Phi}{\partial v_y}(y - x) d\sigma(y) \to -u(x)$ as $\epsilon \to 0$.

Thus,

$$u(x) = \int_{\Omega} \Gamma(y-x)\Delta u(y)dy + \int_{\partial\Omega} u(y)\frac{\partial \Gamma}{\partial \nu_y}(y-x)d\sigma(y) - \int_{\partial\Omega} \Gamma(y-x)\frac{\partial u}{\partial \nu}(y)d\sigma(y).$$

Replacing $\Delta u = f$, $u = g$ on $\partial\Omega$, we get a formula for u except or the term in $\frac{\partial u}{\partial \nu}$. To eliminate this term, suppose that for each x, r^x solves

$$\Delta r^x = 0 \text{ in } \Omega,$$
$$r^x = \Gamma(y-x) \text{ on } \partial\Omega.$$

Then, the function

$$G(x,y) = \Gamma(x-y) - r^x(y)$$

satisfies $\Delta_y G = 0$ for $x \neq y$ and $G = 0$ on $\partial\Omega$.

Repeating the above with $\Gamma \mapsto G$:

$$u(x) = \int_{\Omega} G(x,y)\Delta u(y)dy + \int_{\partial\Omega} u(y)\frac{\partial G(x,y)}{\partial \nu_y}d\sigma(y). \tag{5.1}$$

Definition 5.16. The function G above is called the Green's function for the domain Ω.

From the above we have

Theorem 5.17. ***Representation Formula.*** *If $u \in C^2(\overline{\Omega})$ solves*

$$\Delta u = f \text{ in } \Omega,$$
$$u = g \text{ on } \partial\Omega.$$

where Ω is bounded, $f \in C^0(\Omega)$, and $g \in C^0(\partial\Omega)$, then u is given by (5.1).

Proposition 5.18.

$$G(x,y) = G(y,x), x \neq y$$

Proof. Set $(x \neq y)$

$$f(z) := G(x,z), \quad g(z) := G(y,z)$$

The goal is to show $f(y) = g(x)$. We have $\Delta f(z) = 0$ for $z \neq x$, $\Delta g(z) = 0$ for $z \neq y$, and $f = 0 = g$ on $\partial\Omega$. Let $U := \Omega \setminus (B_\epsilon(x) \cup B_\epsilon(y))$. Then, by Green's identity:

$$\int_{\partial B_\epsilon(x)} (\frac{\partial f}{\partial \nu}g - \frac{\partial g}{\partial \nu}f)d\sigma = \int_{\partial B_\epsilon} (\frac{\partial g}{\partial \nu}f - \frac{\partial f}{\partial \nu}g)d\sigma. \tag{5.2}$$

Since g is smooth near x:

$$\int_{\partial B_\epsilon(x)} \frac{\partial g}{\partial \nu}fd\sigma \to 0 \text{ as } \epsilon \to 0$$

Also

$$\int_{\partial B_\epsilon(x)} \frac{\partial f}{\partial \nu}gd\sigma = \int_{\partial B_\epsilon(x)} \frac{\partial \Gamma(z-x)}{\partial \nu_z}g(z)d\sigma(z) + \int_{\partial B_\epsilon(x)} \frac{\partial r^x(z)}{\partial \nu}g(z)d\sigma(z) \to g(x) \text{ as } \epsilon \to 0.$$

So, the LHS of (5.2) $\to g(x)$. Similarly, the RHS of (5.2) $\to f(y)$. □

Remark 5.19. In the language of distributions, Γ solves

$$\Delta\Gamma = \delta_x \text{ in } \mathbb{R}^n$$

and G solves

$$\Delta G = \delta_x \text{ in } \Omega,$$
$$G = 0 \text{ on } \partial\Omega.$$

5.4. **Explicit formulas.** We have not shown how to find the function r^x for the construction of G (we shall see later). However, for some special domains, it is possible to construct G "directly."

The case $\mathbb{R}^n_+$. In this case, one can verify

$$G(x,y) = \Gamma(y-x) - \Gamma(y-\tilde{x}),$$

where $\tilde{x}$ is the reflection of x, i.e.,

$$\tilde{x} = (x^1, x^2, ..., x^{n-1}, -x^n).$$

In particular,

$$u(x) = \frac{2x^n}{n\omega_n} \int_{\partial\mathbb{R}^n_+} \frac{g(y)}{|x-y|^n} dy$$

solves

$$\Delta u = 0 \text{ in } \mathbb{R}^n_+,$$
$$u = g \text{ on } \partial\mathbb{R}^n_+,$$

for $g \in C^0(\mathbb{R}^{n-1}) \cap L^\infty(\mathbb{R}^{n-1})$. The function

$$\kappa(x,y) := \frac{2x^n}{n\omega_n} \frac{1}{|x-y|^n}$$

is called **Poisson's kernel for** $\mathbb{R}^n_+$.
(Here, Ω is not bounded as in our assumptions above, but one can check that this G still works.)

The case $B_r(0)$**.** In this case,

$$G(x,y) = \Gamma(y-x) - \Gamma(|x|(y-\tilde{x})),$$

where $\tilde{x} = \frac{x}{|x|^2}$ is the inversion through $\partial B_r(0)$. In particular,

$$u(x) = \frac{r^2 - |x|^2}{n\omega_n r} \int_{B_r(0)} \frac{1}{|x-y|^n} dy$$

solves

$$\Delta u = 0 \text{ in } B_r(0),$$
$$u = g \text{ on } \partial B_r(0),$$

if $g \in C^0(\partial B_r(0))$. The function

$$\kappa(x,y) := \frac{r^2 - |x|^2}{n\omega_n r} \frac{1}{|x-y|^n}$$

is called Poisson's kernel for $B_r(0)$.

6. The heat equation

We now study the heat equation

$$\partial_t u - \Delta u = 0$$

and its non-homogenous version

$$\partial_t u - \Delta u = f.$$

6.1. **Fundamental solution.** The heat equation has the scaling invariance $u(t,x) \mapsto u(\lambda^2 t, \lambda x)$, i.e., if u solves the (homogenous) heat equation, so does $v(t,x) = u(\lambda^2 t, \lambda x)$. Thus, the ratios $\frac{|x|^2}{t}$ plays a role in the heat equation and suggests the Ansatz

$$u(t,x) = \frac{1}{t^\alpha} v\left(\frac{x}{t^\beta}\right).$$

solutions of this form satisfy $u(t,x) = \lambda^\alpha u(\lambda t, \lambda^\beta x), \lambda > 0$. Plugging our Ansatz into the equation

$$t^{-(\alpha+2\beta)} \Delta v(y) + \beta t^{-(\alpha+1)} y \cdot \nabla v(y) + \alpha t^{-(\alpha+1)} v(y) = 0,$$

where $y = \frac{x}{t^\beta}$. Setting $\beta = \frac{1}{2}$ gives

$$\Delta v + \frac{1}{2} y \cdot \nabla v(y) + \alpha v = 0.$$

Assuming now v to be radial, $v(y) = \tilde{v}(r)$,

$$\tilde{v}'' + \frac{n-1}{2}\tilde{v}' + \frac{1}{2} r\tilde{v}' + \alpha\tilde{v} = 0$$

Setting $\alpha = \frac{n}{2}$:

$$(r^{n-1}\tilde{v}')' + \frac{1}{2}(r^n \tilde{v})' = 0.$$

We can now solve this ODE and find, assuming $\tilde{v}, \tilde{v}' \to 0$ as $r \to \infty$,

$$\tilde{v}(r) = Ae^{-\frac{r^2}{4}}.$$

Reverting back to u, we are led to:

Definition 6.1. The function

$$\Gamma(t,x) := \begin{cases} \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}}, & t > 0, x \in \mathbb{R}^n, \\ 0, & t < 0, x \in \mathbb{R}^n, \end{cases}$$

is called the **fundamental solution to the heat equation.**

One readily verifies that

$$\int_{\mathbb{R}^n} \Gamma(t,x) dx = 1, \quad t > 0 \tag{6.1}$$

6.2. **The initial-value problem.** We are interested in the **initial-value problem**, a.k.a. **Cauchy problem,** for the heat equation:

$$\partial_t u - \Delta u = 0 \text{ in } (0,\infty) \times \mathbb{R}^n,$$
$$u = g \text{ on } \{t = 0\} \times \mathbb{R}^n,$$

where g is given (the initial data).

Theorem 6.2. *Let $g \in C^0(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. Set*

$$u(t,x) := \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y) dy, \quad t > 0, x \in \mathbb{R}^n.$$

Then:

(i) $u \in C^\infty((0,\infty) \times \mathbb{R}^n)$
(ii) $\partial_t u - \Delta u = 0$ *in* $(0,\infty) \times \mathbb{R}^n$
(iii) $u = g$ *on* $\{t = 0\} \times \mathbb{R}^n$ *in the sense that* $\lim_{\substack{(t,x) \to (0,x_0) \\ t>0}} u(t,x) = g(x_0)$ *for each* $x_0 \in \mathbb{R}^n$.

Proof. For each $\epsilon > 0$, $\frac{1}{t^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}}$ is C^∞ with uniformly bounded derivatives in $[\epsilon, \infty) \times \mathbb{R}^n$ and the derivatives of the fundamental solution are integrable, so $u \in C^\infty((0,\infty) \times \mathbb{R}^n)$. Also:

$$\partial_t u(t,x) - \Delta u(t,x) = \int_{\mathbb{R}^n} (\partial_t \Gamma(t, x-y) - \Delta \Gamma(t, x-y)) g(y) dy = 0.$$

To show (iii), let $\epsilon > 0$ and $\delta > 0$ be such that

$$|g(y) - g(x_0)| < \epsilon \text{ if } |y - x_0| < \delta.$$

$$\begin{aligned}|u(t,x) - g(x_0)| &= \left| \int_{\mathbb{R}^n} \Gamma(t, x-y)(g(y) - g(x_0)) dy \right| \text{ (using (6.1))} \\ &\leq \int_{B_\delta(x_0)} \Gamma(t, x-y)|g(y) - g(x_0)| + \int_{\mathbb{R}^n \setminus B_\delta(x_0)} \Gamma(t, x-y)|g(y) - g(x_0)| dy \\ &=: I_1 + I_2.\end{aligned}$$

For $|x - x_0| < \frac{\delta}{2}$, we have

$$I_1 \leq \epsilon \int_{B_\delta(x_0)} \Gamma(t, x-y) dy \leq \epsilon.$$

For I_2 we have $|x - x_0| < \frac{\delta}{2}$ and $|y - x_0| \geq \delta$ so

$$\begin{aligned}|y - x_0| &\leq |y - x| + |x - x_0| \leq |y - x| + \frac{\delta}{2} \\ &\leq |y - x| + \frac{|y - x_0|}{2} \implies |y - x_0| \leq 2|y - x|.\end{aligned}$$

$$\begin{aligned}I_1 &\leq 2\|g\|_{L^\infty(\mathbb{R}^n)} \frac{C}{t^{\frac{n}{2}}} \int_{\mathbb{R}^n \setminus B_\delta(x_0)} e^{-\frac{|x-y|^2}{4t}} dy \\ &\leq \frac{C}{t^{\frac{n}{2}}} \int_{\mathbb{R}^n \setminus B_\delta(x_0)} e^{\frac{|y-x_0|^2}{16t}} dy = \frac{C}{t^{\frac{n}{2}}} \int_\delta^\infty e^{-\frac{r^2}{t}} r^{n-1} dr.\end{aligned}$$

The RHS$\to 0$ as $t \to 0^+$. Thus, choose t small such that RHS $< \epsilon$. □

Remark 6.3.

- $u \in C^\infty$: the heat equation regularizes the data.
- If $g \geq 0, \neq 0$, localized then $u(t,x) > 0$ for all $x, t > 0$: $\infty-$speed of propagation

6.3. **The non-homogenous problem.** We now consider

$$\begin{aligned}\partial_t u - \Delta u &= f \text{ in } (0,\infty) \times \mathbb{R}^n, \\ u &= 0 \text{ on } \{t = 0\} \times \mathbb{R}^n,\end{aligned}$$

where f is given. We can reduce this to a solution of the homogenous problem. Let us denote by $C^2_1(I \times \mathbb{R}^n)$, $I =$ interval, the space of functions that C^1 is the t variable and C^2 is the x-variable.

Theorem 6.4. *Let $f \in C^2_1([0,\infty,] \times \mathbb{R}^n)$ have compact support. For each $s \geq 0$, let u_s be the solution to the Cauchy problem given by the previous theorem*

$$\begin{aligned}\partial_t u_s - \Delta u_s &= 0 \text{ in } (s,\infty) \times \mathbb{R}^n, \\ u_s &= f(s, \cdot) \text{ in } \{t = 0\} \times \mathbb{R}^n.\end{aligned}$$

Set $u(t,x) := \int_0^t u_s(t,x) ds$. Then

(i) $u \in C^2_1((0,\infty) \times \mathbb{R}^n)$

(ii) $\partial_t u - \Delta u = f$ *in* $(0,\infty) \times \mathbb{R}^n$.

(iii) $\lim_{\substack{(t,x) \to (0,x_0) \\ t > 0}} u(t,x) = 0$ *for each* $x_0 \in \mathbb{R}^n$.

Proof. We can write

$$u(t,x) = \int_0^t \frac{1}{(4\pi(t-s))^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4(t-s)}} f(s,y)dyds.$$
$$= \int_0^t \int_{\mathbb{R}^n} \Gamma(s,y) f(t-s,x-y)dyds$$

Since f has compact support and $\Gamma(s,y)$ is smooth for s near $s=t$, $t>0$, we can differentiate:

$$\partial_t u(t,x) = \int_0^t \int_{\mathbb{R}^n} \Gamma(s,y)\partial_t f(t-s,x-y)dyds + \int_{\mathbb{R}^n} \Gamma(t,y) f(0,x-y)dy.$$

Similarly

$$D_x^2 u(t,x) = \int_0^t \int_{\mathbb{R}^n} \Gamma(s,y) D_x^2 f(t-s,x-y)dyds$$

we see that both $\partial_t u, D_x^2 u$ are continuous.

Compute

$$\partial_u(t,x) - \Delta u(t,x) = \int_0^t \int_{\mathbb{R}^n} \Gamma(s,y)(\partial_t f(t-s,x-y) - \Delta_x f(t-s,x-y))dyds + \int_{\mathbb{R}^n} \Gamma(t,y) f(0,x-y)dy$$

$$= (\int_\epsilon^t \int_{\mathbb{R}^n} + \int_0^\epsilon \int_{\mathbb{R}^n}) \Big(\Gamma(s,y)(-\partial_s - \Delta_y) f(t-s,x-y) \Big) dyds + \int_{\mathbb{R}^n} \Gamma(t,y) f(0,x-y)dy =: I_1 + I_2 + I_3.$$

$$|I_2| \le \left(\|\partial_t f\|_{L^\infty} + \|D^2 f\|_{L^\infty} \right) \int_0^\epsilon \int_{\mathbb{R}^n} \Gamma(s,y)dyds \le C\epsilon.$$

By parts

$$\begin{aligned} I_1 &= \int_\epsilon^t \int_{\mathbb{R}^n} \Gamma(s,y)(-\partial_s - \Delta_y) f(t-s,x-y)dyds \\ &= \int_\epsilon^t \int_{\mathbb{R}^n} \underbrace{(\partial_s \Gamma(s,y) - \Delta\Gamma(s,y))}_{=0} f(t-s,x-y)dyds \\ &- \underbrace{\int_{\mathbb{R}^n} \Gamma(t,y) f(0,x-y)dy}_{=I_3} \\ &+ \int_{\mathbb{R}^n} \Gamma(\epsilon,y) f(t-\epsilon,x-y)dy, \text{ so} \end{aligned}$$

$$\begin{aligned} I_1 + I_3 &= \int_{\mathbb{R}^n} \Gamma(\epsilon,y) f(t-\epsilon,x-y)dy \\ &= \int_{\mathbb{R}^n} \Gamma(\epsilon,x-y) f(t-\epsilon,y)dy. \end{aligned}$$

This is the same integral as in the previous theorem with $t \mapsto \epsilon, g \mapsto f(t-\epsilon,\cdot)$, so $\epsilon \to 0^+$ gives $f(t,x)$.

□

Remark 6.5. The strategy of solving a non-homogeneous problem by reducing it to a homogeneous one applies to many other PDEs and is known as **Duhamel's principle**.

Remark 6.6. There is no uniqueness in a strict sense, in fact

$$\begin{aligned} \partial_t u - \Delta u &= 0 \text{ in } (0,T) \times \mathbb{R}^n, \\ u &= 0 \text{ on } \{t=0\} \times \mathbb{R}^n, \end{aligned}$$

has ∞-many solutions. Uniqueness does hold, however, if one imposes suitable growth $u \sim e^{|x|^2}$.

The heat equation admits further properties that are reminiscent of Laplace's equation, such as a mean value property and a maximum principle.

7. The wave equation

We study the Cauchy problem for the wave equation:

$$\begin{aligned} \Box u = -\partial_t^2 u + \Delta u = 0 \text{ in } (0,\infty) \times \mathbb{R}^n \\ u = g \text{ on } \{t = 0\} \times \mathbb{R}^n \\ \partial_t u = h \text{ on } \{t = 0\} \times \mathbb{R}^n \end{aligned}$$

7.1. **1d: Dirichlet's formula.** In 1d,

$$-\partial_t^2 u + \partial_x^2 u = 0.$$

Set $\alpha := x + t$, $\beta := x - t$ and put

$$u(t, x) = v(\alpha, \beta).$$

Then,

$$\begin{aligned} \partial_t^2 u = \partial_\alpha^2 v - 2\partial_\alpha \partial_\beta v + \partial_\beta^2 v \\ \partial_x^2 u = \partial_\alpha^2 v - 2\partial_\alpha \partial_\beta v + \partial_\beta^2 v \end{aligned}$$

So, $\Box u = -4\partial_\alpha \partial_\beta v = 0$. Thus, $\partial_\alpha v$ is a function only of α, $\partial_\alpha v(\alpha, \beta) = f(\alpha)$, so

$$v(\alpha, \beta) = F(\alpha) + G(\beta)$$

for some F, G.

We see that if u is a C^2 solution, then there exist F, G such that

$$u(t, x) = F(x + t) + G(x - t)$$

Reciprocally, given C^2 F, G the above is a solution of the wave equation. G and F are called **forward** and **backward** waves (G moves the graph to the right, F to the left).

Observe that

$$\begin{aligned} u(0, x) = F(x) + G(x) = g(x) \\ \partial_t u(0, x) = F'(x) - G'(x) = h(x) \\ \implies F(x) - G(x) = \int_0^x h(y)dy + C. \end{aligned}$$

Solving for F and G:

$$\begin{aligned} F(x) = \frac{1}{2}g(x) + \frac{1}{2}\int_0^x h(y)dy + \frac{C}{2} \\ G(x) = \frac{1}{2}g(x) - \frac{1}{2}\int_0^x h(y)dy - \frac{C}{2}. \end{aligned}$$

Since $u(t, x) = F(x + t) + G(x - t)$:

$$u(t, x) = \frac{g(x + t) + g(x - t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} h(y)dy,$$

which is known as **D'Alembert's formula.**

Theorem 7.1. *Let $g \in C^2(\mathbb{R}), h \in C^1(\mathbb{R})$. Then, there exists a unique $u \in C^2([0,\infty) \times \mathbb{R})$ that solves the Cauchy problem for the 1d wave equation with data g, h.*

Proof. Define u by D'Alembert's formula. We easily verify the properties stated in the theorem. □

Definition 7.2. The lines $x+t=$ constant and $x-t=$ constant are called the **characteristics** (or **characteristic curves**) of the 1d wave equation.

7.2. **Domains of dependence of regions of influence.** Suppose $h=0$ and $g(x)=0$ for $x\notin[a,b]$. Since $g(x+t)$ and $g(x-t)$ are constant along the lines $x+t=$ constant and $x-t=$ constant, respectively, we see that $u(t,x)\neq 0$ only possibly for (t,x) that lie in the region determined by the region lying between the characteristics emanating from a and b as indicated in the figure:

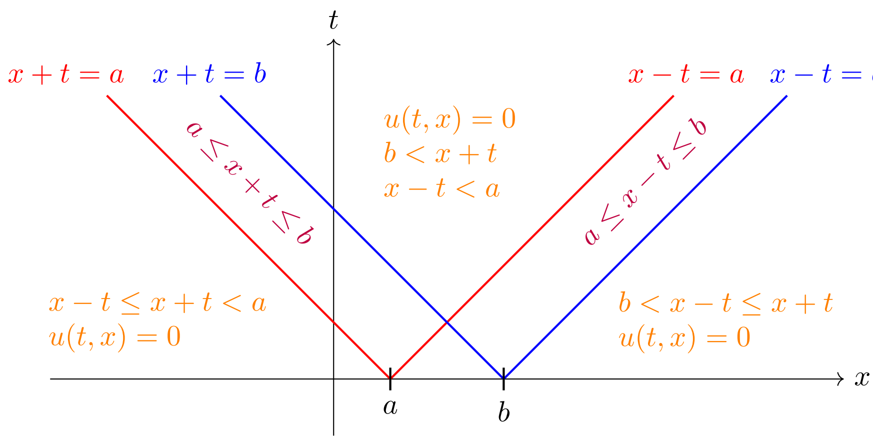


FIGURE 1. Domain of Dependence and the Characteristic Curves

Notation 7.3. Although we ordered the coordinates as (t,x), we will often draw the (t,x) plane with the x-axis on the horizontal.

Suppose now that $g=0$ and that $h(x)=0$ for $x\notin[a,b]$. Then $\int_{x-t}^{x+t}h(y)dy=0$ whenever we have $[x-t,x+t]\cap[a,b]=\emptyset$, i.e., if $x+t<a$ or $x-t>b$. Therefore, $u(t,x)\neq 0$ possibly only in the region $\{x+t\geq a\}\cap\{x-t\leq b\}$, as depicted in the figure:

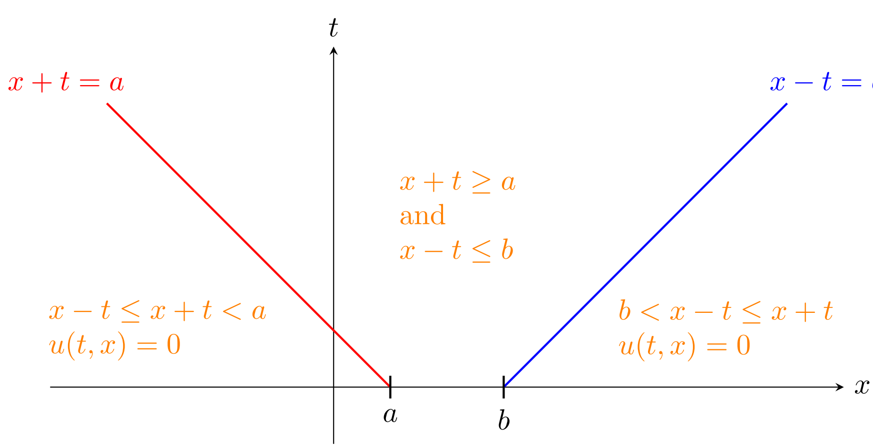


FIGURE 2. Domain of Influence

For general g and h, we can therefore precisely track how the values of $u(t,x)$ are influenced by the values of the initial conditions. It follows that the values of the data on an interval $[a,b]$ can only affect the values of $u(t,x)$ for $(t,x) \in \{x+t \geq a\} \cap \{x-t \leq b\}$. This reflects the fact that waves travel at a finite speed. The region $\{x+t \geq a\} \cap \{x-t \leq b\}$ is called **domain of influence** of $[a,b]$.

Consider now a point (t_0,x_0) and $u(t_0,x_0)$. Let D be the triangle with vertex (t_0,x_0) determined by $x+t = x_0+t_0$, $x-t = x_0+t_0$, and $t=0$:

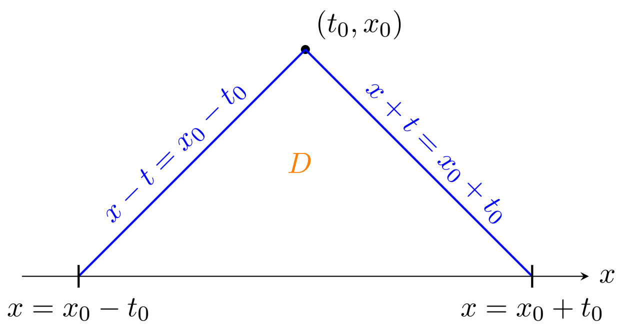


Figure 3. Domain of Dependence

Then,

$$u(t_0,x_0) = \frac{g(x_0+t_0)+g(x_0-t_0)}{2} + \frac{1}{2}\int_{x_0-t_0}^{x_0+t_0} h(y)dy$$

and we see that $u(t_0,x_0)$ is completely determined by the values of the initial data on the interval $[x_0-t_0, x_0+t_0]$. The region D is called the (past) **domain of dependence of** (t_0,x_0)**.**

(Characteristics, domains of dependence/influence are important concepts that will be generalized). Contrast the domain of dependence property with the heat equation.

8. The wave equation in $\mathbb{R}^n$

Here we will study the **Cauchy problem** for the wave equation in $\mathbb{R}^n$, i.e.,

$$\begin{cases} \Box u = 0 \text{ in } [0,\infty) \times \mathbb{R}^n, \\ u = u_0 \text{ on } \{t=0\} \times \mathbb{R}^n, \\ \partial_t u = u_1 \text{ on } \{t=0\} \times \mathbb{R}^n, \end{cases}$$

where $\Box := -\partial_t^2 + \Delta$ is called the D'Alembertian (or the wave operator) and $u_0, u_1 : \mathbb{R}^n \to \mathbb{R}$ are given.

The initial conditions can also be stated as:

$$u(0,x) = u_0(x), \;\; \partial_t u(0,x) = u_1(x), \;\; x \in \mathbb{R}^n$$

Definition 8.1. The sets

$$\mathcal{G}_{t_0,x_0} := \{(t,x) \in (-\infty,+\infty) \times \mathbb{R}^n \;\big|\; |x-x_0| \leq |t-t_0|\},$$
$$\mathcal{G}^+_{t_0,x_0} := \{(t,x) \in (-\infty,+\infty) \times \mathbb{R}^n \;\big|\; |x-x_0| \leq t-t_0\},$$
$$\mathcal{G}^-_{t_0,x_0} := \{(t,x) \in (-\infty,+\infty) \times \mathbb{R}^n \;\big|\; |x-x_0| \leq t_0-t\},$$

are called, respectively, the **light-cone**, **future light-cone**, and **past light-cone with vertex at** (t_0, x_0). The sets

$$\mathcal{K}_{t_0,x_0} := \mathcal{G}_{t_0,x_0} \cap \{t \geq 0\},$$
$$\mathcal{K}^+_{t_0,x_0} := \mathcal{G}^+_{t_0,x_0} \cap \{t \geq 0\},$$
$$\mathcal{K}^-_{t_0,x_0} := \mathcal{G}^-_{t_0,x_0} \cap \{t \geq 0\},$$

are called, respectively, the **light-cone**, **future light-cone**, and **past light-cone for positive time with vertex at** (t_0, x_0). We often omit "for positive time" and refer to the sets $\mathcal{K}$ as light-cones. We also refer to a part of a cone, e.g, for $0 \leq t \leq T$, as the **truncated** (future, past) light-cone.

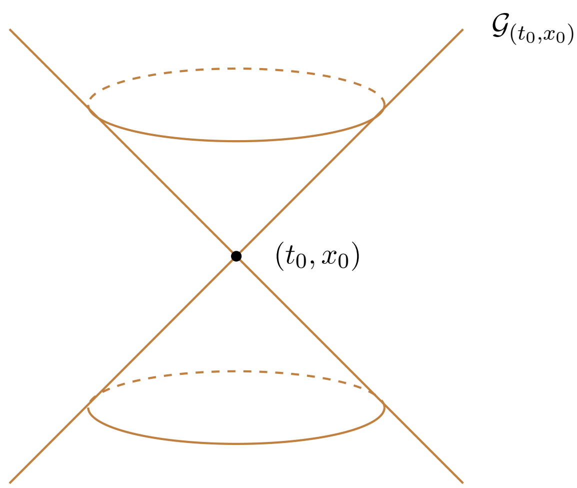


FIGURE 4. Light Cone

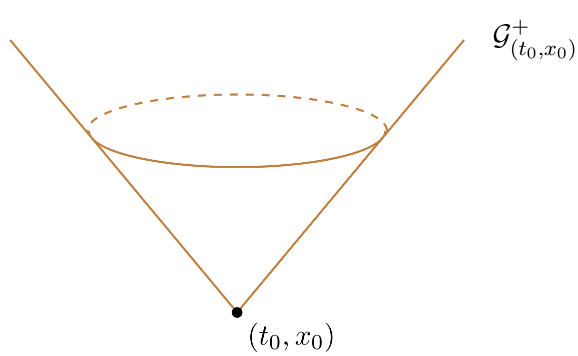


FIGURE 5. Future Light Cone

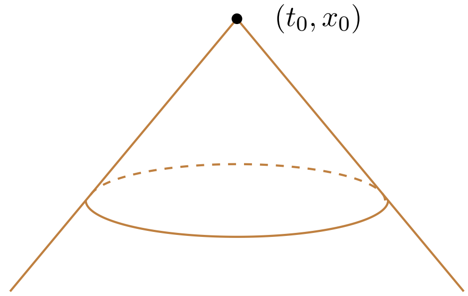


FIGURE 6. Past Light Cone

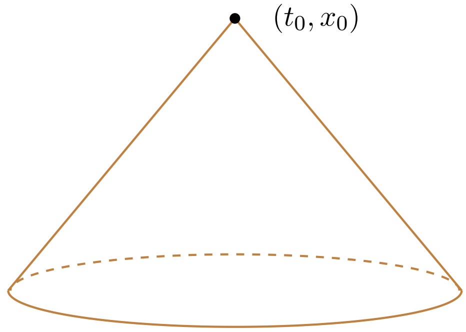


FIGURE 7. Truncated Past Light Cone

Lemma 8.2. *(differentiation of moving regions). Let $\Omega(\tau) \subset \mathbb{R}^n$ be a family of bounded domains with smooth boundary depending smoothly on the parameter τ. Let v be the velocity of the moving boundary $\partial\Omega(\tau)$ and ν the unit outer normal to $\partial\Omega(\tau)$. If $f = f(\tau, x)$ is smooth, then*

$$\frac{d}{d\tau}\int_{\Omega(\tau)} f dx = \int_{\Omega(\tau)} \partial_\tau f dx + \int_{\partial\Omega(\tau)} f v \cdot \nu d\sigma$$

Proof. Change of variables.

□

Theorem 8.3. *(finite propagation speed). Let $u \in C^2([0,\infty) \times \mathbb{R}^n)$ be a solution to the Cauchy problem for the wave equation. If $u_0 = u_1 = 0$ on $\{t = 0\} \times B_{t_0}(x_0)$, then $u = 0$ within $\mathcal{K}^-_{(t_0,x_0)}$. (Thus, the solution at (t_0, x_0) depends on the data on $B_{t_0}(x_0)$ and the cone $\mathcal{K}^-_{(t_0,x_0)}$ is also called a **domain of dependence**).*

Proof. Define the "energy",

$$E(t) = \frac{1}{2}\int_{B_{t_0-t}(x_0)} ((\partial_t u)^2 + |\nabla u|^2) dx, 0 \le t \le t_0.$$

Then,

$$\frac{dE}{dt} = \int_{B_{t_o-t}(x_0)} (\partial_t u \partial_t^2 u + \nabla u \cdot \nabla \partial_t u) dx + \frac{1}{2}\int_{\partial B_{t_o-t}(x_0)} ((\partial_t u)^2 + |\nabla u|^2) v \cdot \nu d\sigma$$

The points on the boundary move inward orthogonally to the spheres $\partial B_{t_0-t}(x_0)$ and with a linear speed in t, thus $v = -\nu$.

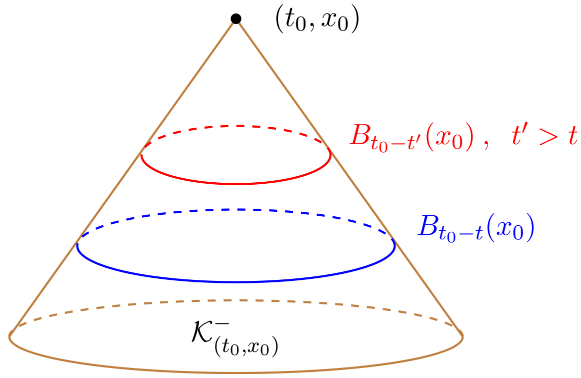


FIGURE 8. Boundaries on the Light Cone

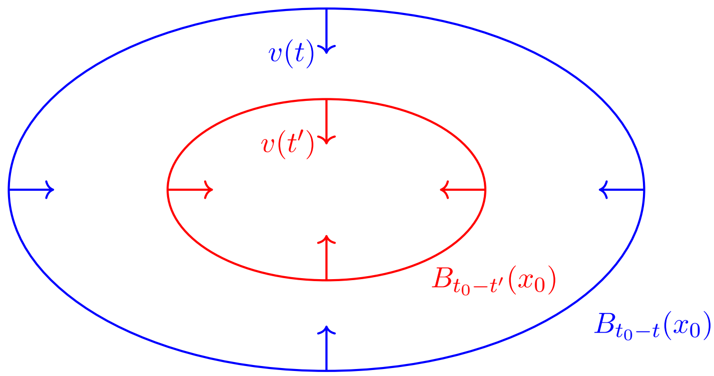


FIGURE 9. Moving Boundaries

Integrating by parts:

$$\int_{B_{t_o-t}(x_0)} \nabla u \cdot \nabla \partial_t u dx = -\int_{B_{t_o-t}(x_0)} \Delta u \partial_t u dx + \int_{\partial B_{t_o-t}(x_0)} \frac{\partial u}{\partial v} \partial_t u d\sigma.$$

Thus,

$$\begin{aligned}\frac{dE}{dt} &= \int_{B_{t_o-t}(x_0)} \underbrace{(\partial_t^2 u - \Delta u)}_{=0} \partial_t u x + \int_{\partial B_{t_o-t}(x_0)} \frac{\partial u}{\partial v} \partial_t u d\sigma - \frac{1}{2}\int_{\partial B_{t_0-t}(x_0)} ((\partial_t u)^2 + |\nabla u|^2) d\sigma \\ &= \int_{\partial B_{t_o-t}(x_0)} (\frac{\partial u}{\partial v}\partial_t u - \frac{1}{2}(\partial_t u)^2 - \frac{1}{2}|\nabla u|^2) d\sigma \\ &\leq \int_{\partial B_{t_o-t}(x_0)} (|\nabla u||\partial_t u| - \frac{1}{2}(\partial_t u)^2 - \frac{1}{2}|\nabla u|^2) d\sigma = 0,\end{aligned}$$

which implies the result.

□

The name, energy, above comes from the fact that $E(t) := \frac{1}{2}\int_{\mathbb{R}^n} ((\partial_t u)^2 + |\nabla u|^2) dx$ indeed represents the total energy of the system at time t, $\frac{1}{2}(\partial_t u)^2$ corresponding to the (local) kinetic energy and $\frac{1}{2}|\nabla u|^2$ the (local) potential energy. Restoring all units we can check that E in fact has units of energy.

Alternatively, one could imagine "discovering" the energy as follows. Multiply the wave equation $-\partial_t^2 u + \Delta u = 0$ by $\partial_t u$ and integrate over $\mathbb{R}^n$:

$$\int_{\mathbb{R}^n} (-\partial_t u \partial_t^2 u + \partial_t u \Delta u) dx = 0$$

Integrating the last term by parts, $\int_{\mathbb{R}^n} \partial_t u \Delta u dx = -\int_{\mathbb{R}^n} \partial_t \nabla u \cdot \nabla u dx$, where we assume that u decays fast enough as $|x| \to \infty$ so there are no boundary terms. Thus,

$$0 = \int_{\mathbb{R}^n} (\partial_t u \partial_t^2 u + \partial_t u \nabla u \cdot \nabla u) dx = \frac{1}{2} \partial_t \int_{\mathbb{R}^n} ((\partial_t u)^2 + |\nabla u|^2) dx,$$

i.e.,

$$E(t) = \frac{1}{2} \int_{\mathbb{R}^n} ((\partial_t u)^2 + |\nabla u|^2) dx$$

is conserved.

Remark 8.4. Inspired by the above, it is customary to call "energy" any quantity that is quadratic on derivatives of the solution, integrated over a region, even if they do not have a direct physical meaning. Such energies are typically obtained by multiplying the equation by a suitable term and integrating by parts, as above, and they play a key role in the study of certain PDEs.

Notation 8.5. Henceforth, we assume that $n \geq 2$. Set

$$\begin{aligned} U(t, x; r) &:= \frac{1}{\text{vol}(\partial B_r(x))} \int_{\partial B_r(x)} u(t, y) d\sigma(y), \\ U_0(x; r) &:= \frac{1}{\text{vol}(\partial B_r(x))} \int_{\partial B_r(x)} u_0(t, y) d\sigma(y), \\ U_1(x; r) &:= \frac{1}{\text{vol}(\partial B_r(x))} \int_{\partial B_r(x)} u_1(t, y) d\sigma(y), \end{aligned}$$

which are spherical averages over $\partial B_r(x)$.

Proposition 8.6. *(Euler-Poisson-Darboux equation). Let $u \in C^m([0, \infty) \times \mathbb{R}^n)$, $m \geq 2$, be a solution to the Cauchy problem for the wave equation. For fixed $x \in \mathbb{R}^n$, consider $U = U(t, x; r)$ as a function of t and r. Then, $U \in C^m([0, \infty) \times [0, \infty))$ and U satisfies the* ***Euler-Poisson-Darboux equation:***

$$\begin{cases} \partial_t^2 U - \partial_r^2 U - \dfrac{n-1}{r} \partial_r U = 0 \text{ in } (0, \infty) \times (0, \infty), \\ \qquad\qquad\qquad U = U_0 \text{ on } \{t = 0\} \times (0, \infty), \\ \qquad\qquad\quad \partial_t U = U_1 \text{ on } \{t = 0\} \times (0, \infty). \end{cases}$$

Proof. Differentiability with respect to t is immediate, as is the differentiability w.r.t. r for $r > 0$.

Arguing as in the proof of the mean value formula for Laplace's equation:

$$\partial_r U(t, x; r) = \frac{r}{n} \frac{1}{\text{vol}(B_r(x))} \int_{B_r(x)} \Delta u(t, y) dy.$$

This implies $\lim_{r \to 0^+} \partial_r U(t, x; r) = 0$. Next,

$$\begin{aligned} \partial_r^2 U(t, x; r) = &\frac{1}{n} \frac{1}{\text{vol}(B_r(x))} \int_{B_r(x)} \Delta u(t, y) dy \\ &+ \frac{r}{n} \partial_r \left(\frac{1}{\text{vol}(B_r(x))} \right) \int_{B_r(x)} \Delta u(t, y) dy + \frac{r}{n} \frac{1}{\text{vol}(B_r(x))} \partial_r \left(\int_{B_r(x)} \Delta u(t, y) dy \right). \end{aligned}$$

But $\partial_r\left(\int_{B_r(x)} \Delta u(t,y)dy \right) = \int_{\partial B_r(x)} \Delta u(t,y) d\sigma(y)$, and recall that $\text{vol}(B_r(x)) = \omega_n r^n$, so

$$\frac{r}{n}\frac{1}{\text{vol}(B_r(x))} = \frac{1}{n\omega_n r^{n-1}} = \frac{1}{\text{vol}(\partial B_r(x))},$$

$$\frac{r}{n}\partial_r\left(\frac{1}{\text{vol}(B_r(x))}\right) = \frac{r}{n}\partial_r\left(\frac{1}{\omega_n r^n}\right) = -\frac{1}{\omega_n r^n} = -\frac{1}{\text{vol}(B_r(x))},$$

so

$$\partial_r^2 U(t,x;r) = \left(\frac{1}{n} - 1\right)\frac{1}{\text{vol}(B_r(x))}\int_{B_r(x)} \Delta u(t,y)dy + \frac{1}{\text{vol}(\partial B_r(x))}\int_{\partial B_r(x)} \Delta u(t,y) d\sigma(y).$$

This implies that $\lim_{r\to 0^+} \partial_r^2 U(t,x;r) = \frac{1}{n}\Delta u(t,x)$.

Proceeding this way we compute all derivatives of U w.r.t. r and conclude that $U \in C^m([0,\infty)\times[0,\infty))$. Returning to the expression for $\partial_r U$:

$$\partial_r U = \frac{r}{n}\frac{1}{\text{vol}(B_r(x))}\int_{B_r(x)} \Delta u = \frac{r}{n}\frac{1}{\text{vol}(B_r(x))}\int_{B_r(x)} \partial_t^2 u.$$

Thus,

$$\begin{aligned}
\partial_r(r^{n-1}\partial_r U) &= \partial_r\left(\frac{r^n}{n\text{vol}(B_r(x))}\int_{B_r(x)} \partial_t^2 u\right) = \partial_r\left(\frac{1}{n\omega_n}\int_{B_r(x)} \partial_t^2 u\right) \\
&= \frac{1}{n\omega_n}\int_{\partial B_r(x)} \partial_t^2 u = \frac{r^{n-1}}{\text{vol}(\partial B_r(x))}\int_{\partial B_r(x)} \partial^2 tu \\
&= r^{n-1}\partial_t^2\left(\frac{1}{\text{vol}(\partial B_r(x))}\int_{\partial B_r(x)} u\right) = r^{n-1}\partial_t^2 U.
\end{aligned}$$

On the other hand:

$$\underbrace{\partial_r(r^{n-1}\partial_r U)}_{=r^{n-1}\partial_t^2 U} = (n-1)r^{n-2}\partial_r U + r^{n-1}\partial_r^2 U,$$

which gives the result.

□

8.1. **Reflection method.** We will use the function $U(t,x;r)$ to reduce the higher dimensional wave equation to the 1d wave equation, for which D'Alembert's formula is available, in the variables t and r. However, $U(t,x;r)$ is defined only for $r \geq 0$, whereas D'Alembert's formula is for $-\infty < r < \infty$. Thus, we first consider:

$$\begin{cases} u_{tt} - u_{xx} = 0 \text{ in } (0,\infty)\times(0,\infty), \\ u = u_0 \text{ on } \{t=0\}\times(0,\infty), \\ \partial_t u = u_1 \text{ on } \{t=0\}\times(0,\infty), \\ u = 0 \text{ on } (0,\infty)\times\{x=0\}, \end{cases}$$

where $u_0(0) = u_1(0) = 0$. Consider odd extensions, where $t \geq 0$:

$$\tilde{u}_0 = \begin{cases} u_0(x), \ x \geq 0, \\ -u_0(-x), \ x \leq 0, \end{cases} \qquad \tilde{u}_1(x) = \begin{cases} u_1(x), \ x \geq 0, \\ -u_1(x), \ x \leq 0. \end{cases}$$

A solution to the problem on $(0,\infty,)\times(0,\infty)$ is obtained by solving

$$\begin{cases} \tilde{u}_{tt} - \tilde{u}_{xx} = 0 \text{ in } (0,\infty)\times\mathbb{R}, \\ \tilde{u} = \tilde{u}_0 \text{ on } \{t=0\}\times\mathbb{R}, \\ \partial_t \tilde{u} = \tilde{u}_1 \text{ on } \{t=0\}\times\mathbb{R}, \end{cases}$$

and restricting to $(0, \infty,) \times (0, \infty)$ where $\tilde{u} = u$. (D'Alembert's formula implies that $\tilde{u}$ will be odd, thus satisfying $\tilde{u}(t, 0) = 0$, if $\tilde{u}_0$ and $\tilde{u}_1$ are odd, i.e., if u_0 and u_1 vanish at $x = 0$.)

D'Alembert's formula gives

$$\tilde{u}(t, x) = \frac{1}{2}(\tilde{u}_0(x + t) + \tilde{u}_0(x - t)) + \frac{1}{2}\int_{x-t}^{x+t} \tilde{u}_1(y)dy.$$

Consider now $t \geq 0$ and $x \geq 0$, so that $\tilde{u}(t, x) = u(t, x)$. Then, $x+t \geq 0$ so that $\tilde{u}_0(x+t) = u_0(x+t)$. If $x \geq t$, then the integration variable y satisfies $y \geq 0$, since $y \in [x - t, x + t]$. In this case $\tilde{u}_1(y) = u_1(y)$. Thus,

$$u(t, x) = \frac{1}{2}(u_0(x + t) + u_0(x - t)) + \frac{1}{2}\int_{x-t}^{x+t} u_1(y)dy$$

for $x \geq t$. If $0 \leq x \leq t$, then $\tilde{u}_0(x - t) = -\tilde{u}_0(-(x - t))$ and

$$\begin{aligned}\int_{x-t}^{x+t} \tilde{u}_1(y)dy &= \int_{x-t}^{0} \tilde{u}_1(y)dy + \int_{0}^{x+t} \tilde{u}_1(y)dy = -\int_{x-t}^{0} u_1(-y)dy + \int_{0}^{x+t} u_1(y)dy \\ &= \int_{-x+t}^{0} u_1(y)dy + \int_{0}^{x+t} u_1(y)dy = \int_{-x+t}^{x+t} u_1(y)dy.\end{aligned}$$

Thus,

$$u(t, x) = \frac{1}{2}(u_0(x + t) - u_0(t - x)) + \frac{1}{2}\int_{-x+t}^{x+t} u_1(y)dy$$

for $0 \leq x \leq t$.

Summarizing:

$$u(t, x) = \begin{cases} \frac{1}{2}(u_0(x + t) + u_0(x - t)) + \frac{1}{2}\int_{x-t}^{x+t} u_1(y)dy, & x \geq t \geq 0, \\ \frac{1}{2}(u_0(x + t) - u_0(t - x)) + \frac{1}{2}\int_{-x+t}^{x+t} u_1(y)dy, & 0 \leq x \leq t. \end{cases}$$

Note that u is not C^2 except if $u''(0) = 0$. Note also that $u(t, 0) = 0$.

This solution can be understood as follows: for $x \geq t \geq 0$, finite propagation speed implies that the solution "does not see" the boundary. For $0 \leq x \leq t$, the waves traveling to the left are reflections on the boundary where $u = 0$.

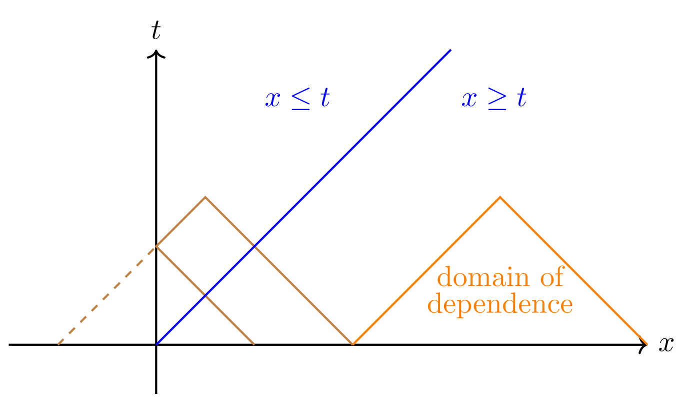


FIGURE 10. Reflection Across the t-axis

8.2. **Solution for $n = 3$: Kirchhoff's formula.** Set $\tilde{U} = rU$, $\tilde{U}_0 = rU_0$, $\tilde{U}_1 = rU_1$, where $\tilde{U}, \tilde{U}_0, \tilde{U}_1$ are as above. Then,

$$\begin{aligned}
\partial_t^2 \tilde{U} &= r\partial_t^2 U = r\left(\partial_r^2 U + \frac{3-1}{r}\partial_r U\right) \\
&= r\partial_r^2 U + 2\partial_r U \\
&= \partial_r^2(rU) = \partial_r^2 \tilde{U},
\end{aligned}$$

so $\tilde{U}$ solves the 1d wave equation on $(0,\infty,)\times(0,\infty)$ with initial conditions $\tilde{U}(0,r) = \tilde{U}_0(r)$, $\partial_t \tilde{U}(0,r) = \tilde{U}_1(r)$.

By the reflection method discussed above, we have

$$\tilde{U}(t,x;r) = \frac{1}{2}(\tilde{U}_0(r+t) - \tilde{U}_0(t-r)) + \frac{1}{2}\int_{-r+t}^{r+t} \tilde{U}_1(y)dy$$

for $0 \le r \le t$, where we used the notation $\tilde{U}_0(r+t)$ and $\tilde{U}_1(y)$ for $\tilde{U}_0(x;r+t), \tilde{U}_1(x;y)$.

From the definition of $\tilde{U}$ and U and the above formula:

$$\begin{aligned}
u(t,x) &= \lim_{r\to 0^+} \frac{1}{\mathrm{vol}(\partial B_r(x))}\int_{\partial B_r(x)} u(t,y)d\sigma(y) \\
&= \lim_{r\to 0^+} U(t,x;r) \\
&= \lim_{r\to 0^+} \frac{\tilde{U}(t,x;r)}{r} \\
&= \lim_{r\to 0^+} \frac{1}{2}\frac{\tilde{U}_0(t+r) - \tilde{U}_0(t-r)}{r} + \lim_{r\to 0^+} \frac{1}{2r}\int_{t-r}^{t+r} \tilde{U}_1(y)dy.
\end{aligned}$$

Note that

$$\lim_{r\to 0^+} \frac{\tilde{U}_0(t+r) - \tilde{U}_0(t-r)}{2r} = \lim_{r\to 0^+} \frac{\tilde{U}_0(t+2r) - \tilde{U}_0(t)}{2r} = \tilde{U}_0'(t)$$

and

$$\lim_{r\to 0^+} \frac{1}{2r}\int_{t-r}^{t+r} \tilde{U}_1(y)dy = \tilde{U}_1(t)$$

(this equality is simply $\lim_{r\to 0^+} \frac{1}{\mathrm{vol}(B_r(x))}\int_{B_r(x)} f(y)dy = f(x)$ for $n = 1$). So,

$$u(t,x) = \tilde{U}_0'(t) + \tilde{U}_1(t)$$

Invoking the definition of $\tilde{U}_0$ and $\tilde{U}_1$:

$$u(t,x) = \frac{\partial}{\partial t}\left(\frac{t}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} u_0(y)d\sigma(y)\right) + \frac{t}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} u_1(y)d\sigma(y).$$

Making the change of variables $z = \frac{y-x}{t}$ (recall that we are treating the $n = 3$ case, so in the calculations that follow $n = 3$, but we write n for the sake of a cleaner notation):

$$\begin{aligned}
\frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} u_0(y)d\sigma(y) &= \frac{1}{n\omega_n t^{n-1}}\int_{\partial B_t(x)} u_0(y)d\sigma(y) \\
&= \frac{1}{n\omega_n t^{n-1}}\int_{\partial B_1(0)} u_0(x+tz)t^{n-1}d\sigma(z) \\
&= \frac{1}{n\omega_n}\int_{\partial B_1(0)} u_0(x+tz)d\sigma(z).
\end{aligned}$$

Then,

$$\frac{\partial}{\partial t}\left(\frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} u_0(y)d\sigma(y)\right) = \frac{1}{n\omega_n}\frac{\partial}{\partial t}\left(\int_{\partial B_1(0)} u_0(x+tz)d\sigma(z)\right)$$
$$= \frac{1}{n\omega_n}\int_{\partial B_1(0)} \nabla u_0(x+tz)\cdot z d\sigma(z).$$

Changing variables back to y, i.e., $y = x + tz$ and recalling that $d\sigma(y) = t^{n-1}d\sigma(z)$:

$$\frac{\partial}{\partial t}\left(\frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} u_0(y)d\sigma(y)\right) = \frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} \nabla u_0(y)\cdot(\frac{y-x}{t})d\sigma(y).$$

Using this in the above expression for $u(t,x)$:

$$u(t,x) = \frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} \big(u_0(y) + tu_1(y)\big)d\sigma(y)$$
$$+ \frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} \nabla u_0(y)\cdot(y-x)d\sigma(y),$$

which is known as **Kirchhoff's formula**.

Theorem 8.7. *Let $u_0 \in C^3(\mathbb{R}^3)$ and $u_1 \in C^2(\mathbb{R}^3)$. Then, there exists a unique $u \in C^2([0,\infty)\times\mathbb{R}^3)$ that is a solution to the Cauchy problem for the wave equation in the three spatial dimensions. Moreover, u is given by Kirchhoff's formula.*

Proof. Define u by Kirchhoff's formula. By construction it is a solution with the stated regularity. Uniqueness follow from the finite speed propagation property.

□

8.3. **Solution for $n = 2$: Poisson's formula.** We now consider $u \in C^2([0,\infty)\times\mathbb{R}^2)$ a solution to the wave equation for $n = 2$. Then

$$v(t,x^1,x^2,x^3) := u(t,x^1,x^2)$$

is a solution for the wave equation in $n = 3$ dimensions with data $v_0(x^1,x^2,x^3) := u(x^1,x^2)$ and $v_1(x^1,x^2,x^3) := u_1(x^1,x^2)$. Let us write $x = (x^1,x^2)$ and $\bar{x}(x^1,x^2,0)$. Thus, from the $n = 3$ case:

$$u(t,x) = v(t,\bar{x}) = \frac{\partial}{\partial t}\left(\frac{t}{\mathrm{vol}(\partial\bar{B}_t(\bar{x}))}\int_{\partial\bar{B}_t(\bar{x})} v_0 d\bar{\sigma}\right) + \frac{t}{\mathrm{vol}(\partial\bar{B}_t(\bar{x}))}\int_{\partial\bar{B}_t(\bar{x})} v_1 d\bar{\sigma},$$

where $\bar{B}_t(\bar{x})$ = ball in $\mathbb{R}^3$ with center $\bar{x}$ and radius t, $d\bar{\sigma}$ = volume element on $\partial\bar{B}_t(\bar{x})$. We now rewrite this formula with integrals involving only variables in $\mathbb{R}^2$.

The integral over $\partial\bar{B}_t(\bar{x})$ can be written as

$$\int_{\partial\bar{B}_t(\bar{x})} = \int_{\partial\bar{B}_t^+(\bar{x})} + \int_{\partial\bar{B}_t^-(\bar{x})},$$

where $\partial\bar{B}_t^+(\bar{x})$ and $\partial\bar{B}_t^-(\bar{x})$ are, respectively, the upper and lower hemispheres of $\partial\bar{B}_t(\bar{x})$.

The upper cap $\partial\bar{B}_t^+(\bar{x})$ is parametrized by

$$f(y) = \sqrt{t^2 - |y-x|^2}, \quad y = (y^1,y^2) \in B_t(x), \quad x = (x^1,x^2),$$

where $B_t(x)$ is the ball of radius t and center x in $\mathbb{R}^2$. Recalling the formula for integrals along a surface given by a graph:

$$\frac{1}{\mathrm{vol}(\partial\bar{B}_t(\bar{x}))}\int_{\partial\bar{B}_t^+(\bar{x})} v_0 d\bar{\sigma} = \frac{1}{4\pi t^2}\int_{B_t(x)} u_0(y)\sqrt{1+|\nabla f(y)|^2}dy,$$

where we used that $v_0(x^1, x^2, x^3) = u_0(x^1, x^2)$. This last fact also implies that

$$\int_{\partial \bar{B}_t^+(\bar{x})} v_0 d\bar{\sigma} = \int_{\partial \bar{B}_t^-(\bar{x})} v_0 d\bar{\sigma},$$

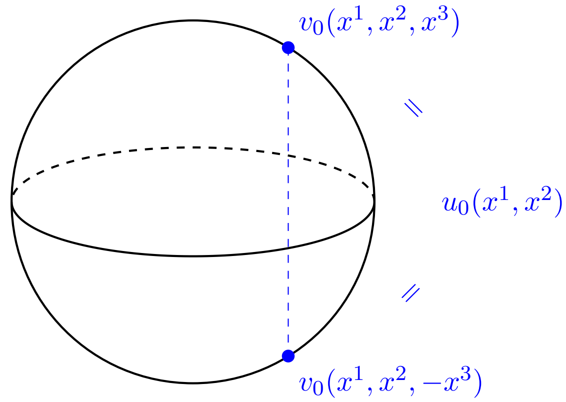


FIGURE 11. 3D Ball

Thus,

$$\begin{aligned}\frac{1}{\operatorname{vol}(\partial \bar{B}_t(\bar{x}))} \int_{\partial \bar{B}_t(\bar{x})} v_0 \, d\bar{\sigma} &= \frac{2}{4\pi t^2} \int_{B_t(x)} u_0(y) \sqrt{1 + |\nabla f(y)|^2} \, dy \\ &= \frac{1}{2\pi t} \int_{B_t(x)} \frac{u_0(y)}{\sqrt{t^2 - |y-x|^2}} \, dy.\end{aligned}$$

In the last step we used

$$1 + |\nabla f(y)|^2 = 1 + \frac{|y-x|^2}{t^2 - |y-x|^2} = \frac{t^2}{t^2 - |y-x|^2}$$

Similarly,

$$\frac{t}{\operatorname{vol}(\partial \bar{B}_t(\bar{x}))} \int_{\partial \bar{B}_t(\bar{x})} v_1 d\bar{\sigma} = \frac{1}{2\pi} \int_{B_t(x)} \frac{u_1(y)}{\sqrt{t^2 - |y-x|^2}} dy.$$

Hence,

$$\begin{aligned}u(t,x) &= \frac{\partial}{\partial t}\left(\frac{1}{2\pi} \int_{B_t(x)} \frac{u_0(y)}{\sqrt{t^2 - |y-x|^2}} dy \right) + \frac{1}{2\pi} \int_{B_t(x)} \frac{u_1(y)}{\sqrt{t^2 - |y-x|^2}} dy \\ &= \frac{1}{2} \frac{\partial}{\partial t}\left(\frac{t^2}{\operatorname{vol}(B_t(x))} \int_{B_t(x)} \frac{u_0(y)}{\sqrt{t^2 - |y-x|^2}} dy \right) + \frac{1}{2} \frac{t^2}{\operatorname{vol}(B_t(x))} \int_{B_t(x)} \frac{u_1(y)}{\sqrt{t^2 - |y-x|^2}} dy.\end{aligned}$$

Changing variables $\frac{y-x}{t} = z$ in the first integral (so $dy = t^2 dz$)

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{t^2}{\operatorname{vol}(B_t(x))} \int_{B_t(x)} \frac{u_0(y)}{\sqrt{t^2 - |y-x|^2}} dy \right) &= \frac{\partial}{\partial t}\left(\frac{t}{\operatorname{vol}(B_1(0))} \int_{B_1(0)} \frac{u_0(x+tz)}{\sqrt{1-|z|^2}} dz \right) \\ &= \frac{1}{\operatorname{vol}(B_1(0))} \int_{B_1(0)} \frac{u_0(x+tz)}{\sqrt{1-|z|^2}} dz + \frac{t}{\operatorname{vol}(B_1(0)} \int_{B_1(0)} \frac{\nabla u_0(x+tz) \cdot z}{\sqrt{1-|z|^2}} dz \\ &= \frac{t}{\operatorname{vol}(B_t(x))} \int_{B_t(x)} \frac{u_0(y)}{\sqrt{t^2 - |y-x|^2}} dy + \frac{t}{\operatorname{vol}(B_t(x))} \int_{B_t(x)} \frac{\nabla u_0(y) \cdot (y-x)}{\sqrt{t^2 - |y-x|^2}} dy,\end{aligned}$$

where in the last step we changed variables back to y. Hence,

$$u(t,x) = \frac{1}{2}\frac{1}{\mathrm{vol}(B_t(x))}\int_{B_t(x)}\left(\frac{tu_0(y)+t^2u_1(y)}{\sqrt{t^2-|y-x|^2}}\right)dy + \frac{1}{2}\frac{1}{\mathrm{vol}(B_t(x))}\int_{B_t(x)}\frac{t\nabla u_0(y)(y-x)}{\sqrt{t^2-|y-x|^2}}dy$$

which is known as Poisson's formula.

Theorem 8.8. *Let $u_0 \in C^3(\mathbb{R}^2)$ and $u_1 \in C^2(\mathbb{R}^2)$. Then, there exists a unique $u \in C^2([0,\infty)\times\mathbb{R}^2)$ that is a solution to the Cauchy problem for the wave equation in two spatial dimensions. Moreover, u is given by Poisson's formula.*

Proof. Define u by Poisson's formula. By construction, it is a solution with the stated regularity. Uniqueness follows from the finite-speed propagation property.

□

8.4. **Solution for arbitrary $n \geq 2$.** The above procedure can be generalized for any $n \geq 2$: for n odd, we show that the suitably radial averages of u satisfies a 1d wave equation for $r > 0$ and invoke the reflection principle; for n even, we view u as a solution in $n+1$ dimensions, apply the result for n odd, and then reduce back to n dimensions. The finite formulas are

n **odd:**

$$u(t,x) = \frac{1}{\beta_n}\frac{\partial}{\partial t}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-3}{2}}\left(\frac{t^{n-2}}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)}u_0 d\sigma\right) + \frac{1}{\beta_n}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-3}{2}}\left(\frac{t^{n-2}}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)}u_1 d\sigma\right)$$

where

$$\beta_n := 1\cdot 3\cdot 5\cdots(n-2),$$

n **even**:

$$u(t,x) = \frac{1}{r^n}\frac{\partial}{\partial t}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-2}{2}}\left(\frac{t^n}{\mathrm{vol}(B_t(x))}\int_{B_t(x)}\frac{u_0(y)}{\sqrt{t^2-|y-x|^2}}dy\right) + \frac{1}{r^n}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-2}{2}}\left(\frac{t^n}{\mathrm{vol}(B_t(x))}\int_{B_t(x)}\frac{u_1(y)}{\sqrt{t^2-|y-x|^2}}dy\right),$$

where

$$r_n := 2\cdot 4\cdots(n-2)\cdot n.$$

Remark 8.9. The method of using the solution in $n+1$ to obtain a solution in n dimensions for n even is known as **method of descent.**

Remark 8.10. We already know that solutions to the wave equation at (t_0, x_0) depend only on the data on $B_{t_0}(x_0)$. For $n \geq 3$ odd, the above shows that the solution depends only on the data on the boundary $\partial B_{t_0}(x_0)$. This fact is known as the **strong Huygens' principle.**

8.5. **The inhomogeneous wave equation.** We now consider

$$\begin{cases} \Box u = f \text{ in } (0,\infty)\times\mathbb{R}^n, \\ \quad u = u_0 \text{ on } \{t=0\}\times\mathbb{R}^n, \\ \partial_t u = u_1 \text{ on } \{t=0\}\times\mathbb{R}^n \end{cases}$$

where $f : [0,\infty) \to \mathbb{R}^n$, $u_0, u_1 : \mathbb{R}^n \to \mathbb{R}$ are given. f is called a source and this is known as the inhomogeneous Cauchy problem for the wave equation. Since we already know how to solve the problem when $f = 0$, by linearity it suffices to consider

$$\begin{cases} \Box u = f \text{ in } (0,\infty)\times\mathbb{R}^n, \\ \quad u = 0 \text{ on } \{t=0\}\times\mathbb{R}^n, \\ \partial_t u = 0 \text{ on } \{t=0\}\times\mathbb{R}^n, \end{cases}$$

Let $u_s(t,x)$ be the solution of

$$\begin{cases} \Box u_s = 0 \text{ in } (s,\infty)\times\mathbb{R}^n, \\ \quad u_s = 0 \text{ on } \{t=s\}\times\mathbb{R}^n, \\ \partial_t u_s = f \text{ on } \{t=s\}\times\mathbb{R}^n, \end{cases}$$

This problem is simply the Cauchy problem with data on $t = s$ instead of $t = 0$, so the previous solutions apply.

For $t \geq 0$, define:

$$u(t,x) := \int_0^t u_s(t,s)ds.$$

Note that $u(0,x) = 0$. We have

$$\partial_t u(t,x) = u_s(t,x)\big|_{s=t} + \int_0^t \partial_t u_s(t,x)ds.$$

Since $u_s(t,x) = 0$ for $t = s$, the first term vanishes, so

$$\partial_t u(t,x) = \int_0^t \partial_t u_s(t,x)ds.$$

Then, $\partial_t u(0,x) = 0$. Taking another derivative:

$$\partial_t^2 u(t,x) = \partial_t u_s(t,x)\big|_{s=t} + \int_0^t \partial_t^2 u_s(t,x)ds.$$

Since $\partial_t u_s(t,x)\big|_{s=t} = f(s,x)\big|_{s=t} = f(t,x)$ and $\partial_t^2 u_s = \Delta u_s$:

$$\begin{aligned} \partial_t^2 u(t,x) &= f(t,x) + \int_0^t \Delta u_s(t,x)ds \\ &= f(t,x) + \Delta \int_0^t u_s(t,x)ds \\ &= f(t,x) + \Delta u(t,x), \text{ i.e.,} \end{aligned}$$

$$\partial_t^2 u - \Delta u = f.$$

Therefore, we conclude that u satisfies the inhomogeneous wave equation with zero initial conditions. We summarize this in the following theorem:

Theorem 8.11. *Let $f \in C^{[\frac{n}{2}]+1}([0,\infty) \times \mathbb{R}^n)$, where $[\frac{n}{2}]$ is the integer part of $\frac{n}{2}$. Let u_s be the unique solution to:*

$$\begin{cases} \Box u_s = 0 \text{ in } (s,\infty) \times \mathbb{R}^n, \\ u_s = 0 \text{ on } \{t = s\} \times \mathbb{R}^n, \\ \partial_t u_s = f \text{ on } \{t = s\} \times \mathbb{R}^n, \end{cases}$$

and define u by

$$u(t,x) = \int_0^t u_s(t,x)ds.$$

Then, $u \in C^2([0,\infty) \times \mathbb{R}^n)$ and is a solution to the Cauchy problem for the wave equation with source f and zero initial conditions.

Remark 8.12. The procedure of solving the inhomogeneous equation by solving a homogeneous one with initial condition f, as seen in the case of the heat equation, is known as the Duhamel principle.

9. Sobolev spaces

We will now introduce and study properties of certain function spaces, called Sobolev spaces, that are very useful for the study of PDEs.

Unless stated otherwise, *in this section Ω denotes a domain $\Omega \subset \mathbb{R}^n$.*

9.1. Weak derivatives.

Definition 9.1. Let $u, v \in L^1_{\text{loc}}(\Omega)$ and α be a multi-index. We say that v is a **α-weak partial derivative** v**,** and write $D^\alpha u = v$, if

$$\int_\Omega u D^\alpha \varphi dx = (-1)^{|\alpha|} \int_\Omega v\varphi dx$$

for every $\varphi \in C_c^\infty(\Omega)$.

The intuition is that the integration by parts formula holds. Functions in $C_c^\infty(\Omega)$ are often called **test functions**. The space of functions whose all derivatives exist up to order k is denoted $W^k(\Omega)$.

Example 9.2. If $v(x) = \begin{cases} 1, & -1 < x < 0, \\ x+1, & 0 \le x < 1, \end{cases}$ then $v'(x) = \begin{cases} 0, & -1 < x < 0, \\ 1, & 0 < x < 1, \end{cases}$ is a weak-derivative.

$$\int_{-1}^1 v(x)\varphi'(x)dx = \int_{-1}^0 \varphi'(x)dx + \int_0^1 (x+1)\varphi'(x)dx$$
$$= \varphi(0) - \int_0^1 \varphi(x)dx - \varphi(0) = -\int_{-1}^1 v'(x)\varphi(x)dx.$$

If, however, $v(x) = \begin{cases} 1, & -1 < x < 0, \\ x+2, & 0 \le x < 1, \end{cases}$ then, as we shall see, weak derivatives do not exist.

Lemma 9.3. *Weak derivatives, if they exist, are unique.*

Proof. If v, w are weak derivatives of u, then

$$\int_\Omega u D^\alpha \varphi = (-1)^{|\alpha|} \int_\Omega v\varphi = (-1)^{|\alpha|} \int_\Omega w\varphi, \quad \int_\Omega \varphi(v - w) = 0$$

for all test functions $\varphi, v = w$ a.e..

□

Thus, *weak and classical derivatives agree* when the latter exists.

Lemma 9.4. *Let $u \in L^1_{loc}(\Omega)$ and suppose that $D^\alpha u$ exists, where α is a multi-index. Then, if $\epsilon <$dist$(x,\partial\Omega)$,*

$$(D^\alpha u_\epsilon)(x) = (D^\alpha u)_\epsilon(x),$$

where $(\cdot)_\epsilon$ is the regularization.

Proof.

$$\begin{aligned} D^\alpha u_\epsilon(x) &= D^\alpha\left[\frac{1}{\epsilon_n}\int_\Omega \varphi\left(\frac{x-y}{\epsilon}\right)u(y)dy\right] = \frac{1}{\epsilon^n}\int_\Omega D^\alpha_x\left(\varphi\left(\frac{x-y}{\epsilon}\right)\right)u(y) \\ &= (-1)^{|\alpha|}\frac{1}{\epsilon^n}\int_\Omega D^\alpha_y\left(\varphi\left(\frac{x-y}{\epsilon}\right)\right)u(y)dy = \frac{1}{\epsilon}\int_\Omega \varphi\left(\frac{x-y}{\epsilon}\right)D^\alpha u(y)dy \\ &= (D^\alpha u)_\epsilon(x) \end{aligned}$$

□

Theorem 9.5. *Let $u, v \in L^1_{loc}(\Omega)$. Then, $D^\alpha u = v$ iff there exists a sequence of $C^\infty(\Omega)$ functions $u_k \underset{L^1_{loc}(\Omega)}{\to} u$ such that $D^\alpha u_k \underset{L^1_{loc}(\Omega)}{\to} v$.*

Proof. If $D^\alpha u = v$, we take a sequence of regularizations suitably multiplied by cut-off functions $(u_\epsilon \in C^\infty(\Omega_\epsilon))$.

If $u_k \to u$ and $D^\alpha u_k \to v$ in $L^1_{\text{loc}}(\Omega)$, let $\varphi \in C^\infty_c(\Omega)$.

$$\int_\kappa v\varphi = \int_\Omega D^\alpha u_k \varphi = (-1)^{|\alpha|}\int_\Omega u_k D^\alpha\varphi = (-1)^{|\alpha|}\int_\kappa u_k D^\alpha \varphi,$$

where κ is a compact set such that $\text{supp}(\varphi) \subset \kappa \subset \Omega$. Passing to the limit, we have the result.

□

Using this approximation property and the definition, we can prove a series of basic properties.

Theorem 9.6. *We have*

(i) $D(uv) = uDv + vDu$, *if* $u,\ v,\ uv,\ uDv + vDu \in W^1(\Omega)$.

(ii) $D^\alpha D^\beta u = D^\beta D^\alpha u = D^{\alpha+\beta}u$ *if* $u \in W^k(\Omega)$, $|\alpha| + |\beta| \le k$ *(More generally, if any two of these weak derivatives exist, then they all exist and coincide.)*

(iii) If $u, v \in W^k(\Omega)$, *so do linear combinations*

(iv) $u \in W^k(\Omega) \implies u \in W^k(\Omega')$, $\Omega' \subset \Omega$, Ω' *open.*

(v) $\varphi u \in W^k(\Omega)$ *if* $u \in W^k(\Omega)$, $\varphi \in C^\infty_c(\Omega)$, *and the product rule holds.*

(vi) If $\psi : \Omega \to \Omega'$ *is a* C^1 *diffeomorphism,* $u \in W^1(\Omega)$, *then* $v := u \circ \psi^{-1} \in W^1(\Omega')$ *and* $\partial_i u(x) = \frac{\partial y^j}{\partial x^i}\partial_j v(y)$.

Proof. (ii), (iii), (iv) are trivial. (v) by induction on k. (i) and (vi) by approximation by smooth functions. □

Theorem 9.7. *(Chain rule). If $\psi \in C^1(\mathbb{R})$, $\psi' \in L^\infty(\mathbb{R})$, $u \in W^1(\Omega)$, then $\psi \circ u \in W^1(\Omega)$ and $D(\psi \circ u) = \psi'(u)Du$.*

Proof. Let $u_k,\ Du_k \to u,\ Du$ in $L^1_{\text{loc}}(\Omega)$. Fix $\kappa \subset\subset \Omega$. Then,

$$\int_\kappa |\psi(u_k) - \psi(u)| \le \|\psi'\|_{L^\infty(\mathbb{R})}\int_\kappa |u_k - u| \to 0$$

$$\int_\kappa |\psi'(u_k)Du_k - \psi'(u)Du| \le \|\psi'\|_{L_\infty(\mathbb{R})}\underbrace{\int_\kappa |Du_k - Du|}_{\to 0} + \int_\kappa |\psi'(u_k) - \psi'(u)||Du|$$

Up to a subsequence, $u_k \to u$ a.e. in κ, so $\psi'(u_k) \to \psi(u)$ a.e. in κ since ψ is C^1. So $\int_\kappa |\psi'(u_k) - \psi'(u)||Du| \to 0$ by dominated convergence. Thus, $\psi(u_k) \to \psi(u)$ and $D(\psi \circ u_k) = \psi'(u_k)Du_k \to \psi'(u)Du$ in $L^1_{\text{loc}}(\Omega)$, so $D(\psi \circ u) = \psi'(u)Du$ by our characterization (which works for finitely many derivatives).

□

Recall that $u^+ = \max\{u, 0\},\ u^- = -\min\{u, 0\},\ \text{so } u = u^+ - u^-,\ |u| = u^+ + u^-.$

Proposition 9.8. *If $u \in W^1(\Omega)$ then $u^+, u^-, |u| \in W^1(\Omega)$, and*

$$Du^+ = \begin{cases} Du \text{ if } u > 0 \\ 0 \text{ if } u \leq 0 \end{cases}, Du^- = \begin{cases} 0 \text{ if } u \geq 0, \\ -Du \text{ if } u < 0, \end{cases}$$

$$D|u| = \begin{cases} Du \text{ if } u > 0, \\ 0 \text{ if } u = 0, \\ -Du \text{ if } u < 0. \end{cases}$$

Proof. Fix $\epsilon > 0$ and set

$$\psi_\epsilon(u) = \begin{cases} (u^2 + \epsilon^2)^{\frac{1}{2}} - \epsilon, & u > 0, \\ 0, & u \leq 0 \end{cases}$$

so $\psi_\epsilon \in C^1(\mathbb{R}),\ \psi'(\mathbb{R}) \in L^\infty(\mathbb{R})$. By the previous theorem,

$$D(\psi_\epsilon \circ u) = \psi'_\epsilon(u)Du = \begin{cases} \frac{u}{(u^2+\epsilon^2)^{\frac{1}{2}}} Du, & u > 0, \\ 0, & u \leq 0. \end{cases}$$

For $\psi \in C_c^\infty(\Omega)$,

$$\begin{array}{ccc} \displaystyle\int_\Omega \psi_\epsilon(u) D\psi & = & \displaystyle -\int_{\{u>0\}} \frac{u}{(u^2+\epsilon^2)^{\frac{1}{2}}} Du\psi \\ \Big\downarrow \epsilon \to 0 & & \Big\downarrow \epsilon \to 0 \\ \displaystyle\int_\Omega u + D\psi & & \displaystyle -\int_{\{u>0\}} Du\psi = -\int_\Omega \chi_{\{u>0\}} Du\psi. \end{array}$$

The remaining results follow from $u^- = (-u)^+,\ |u| = u^+ + u^-$.

□

Corollary 9.9. *Let $u \in W^1(\Omega)$. Then $Du = 0$ a.e. on any set where u is constant.*

Proof. $Du = Du^+ - Du^-$.

□

The converse is also true: $Du = 0$, then $0 = (Du)_\epsilon = Du_\epsilon \implies u_\epsilon = \text{constant} = c_\epsilon$. $u_\epsilon \to u$ is $L^1_{\text{loc}}(\Omega)$; this convergence can only happen if the numerical sequence c_ϵ converges. So $u = \text{constant}$ a.e..

We also have:

Lemma 9.10. *Let ψ be continuous and have piecewise continuous first derivatives with $\psi' \in L^\infty(\mathbb{R})$. If $u \in W^1(\Omega)$, then $f \circ u \in W^1(\Omega)$ and*

$$D(y, u) = \begin{cases} \psi'(u)Du, & u \notin L, \\ 0, & u \in L, \end{cases}$$

where L is the set of corner points of ψ.

Proof. By induction and translations we can reduce it to the case of one corner at the origin. If $\psi_1, \psi_2 \in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and $\psi(u) = \psi_1(u)$, $\psi(u) = \psi_2(u)$ for $u \geq 0$, $u \leq 0$, respectively, then $\psi(u) = \psi_1(u^+) - \psi_2(u^-)$ and the result follows.

□

Thus, if u is weakly differentiable, so is u^2, $f(u)$, etc.

Notation 9.11. From now on, D will always denote weak derivatives unless explicitly said otherwise.

9.2. **Sobolev spaces and their basic properties.**

Definition 9.12. Let $1 \leq p \leq \infty$ and k be an integer. We define the **Sobolev space**

$$W^{k,p}(\Omega) := \big\{ u : \Omega \to \mathbb{R} \big| D^\alpha u \in L^p(\Omega), |\alpha| \leq k \big\},$$

where D^α are weak derivatives of u. We endow $W^{k,p}(\Omega)$ with the norm

$$\begin{aligned} \|u\|_{W^{k,p}(\Omega)} = \|u\|_{k,p} &:= \left(\sum_{|\alpha| \leq k} \int_\Omega |D^\alpha u|^p dx \right)^{\frac{1}{p}} \\ &= \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|^p_{L^p(\Omega)} \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty, \\ \|u\|_{W^{k,\infty}} = \|u\|_{k,\infty} &:= \sum_{|\alpha| \leq k} \operatorname*{ess\,sup}_\Omega |D^\alpha| \\ &= \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)}. \end{aligned}$$

If $p = 2$, we denote $W^{k,2}(\Omega) = H^k(\Omega)$ with the **inner product**

$$(u, v)_k := \sum_{|\alpha| \leq k} \int_\Omega D^\alpha u D^\alpha v dx,$$

where again D^α are weak derivatives. The corresponding norm is denoted $\| \cdot \|_k$. We also define

$$W_0^{k,p}(\Omega) := \text{ closure of } C_c^\infty(\Omega) \text{ in } W^{k,p}(\Omega).$$

We also define $W^{k,p}_{\text{loc}}(\Omega)$ in the usual way.

Remark 9.13.

- $W^{0,p}(\Omega) = L^p(\Omega)$
- $W_0^{0,p}(\Omega) = L^p(\Omega)$ if $1 \leq p \leq \infty$ since $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for $1 \leq p \leq \infty$.
- We have the embeddings
$$W_0^{k,p}(\Omega) \hookrightarrow W^{k,p}(\Omega) \hookrightarrow L^p(\Omega)$$
- $u \in W_0^{k,p}(\Omega)$ iff there exists a sequence $C_c^\infty(\Omega) \ni u_k \to u$ in $W^{k,p}(\Omega)$. Thus, we think of $W_0^{k,p}(\Omega)$ as the set of $u \in W^{k,p}(\Omega)$ such that "$D^\alpha u = 0$" on $\partial\Omega$ for $|\alpha| \leq k - 1$ (this will be given a precise interpretation later on).
- We have obvious generalizations to $\mathbb{C}$, vector-valued functions, etc.
- D^α is a bounded operator from $W^{k,p}(\Omega)$ to $W^{k-|\alpha|,p}(\Omega)$.

Example 9.14. Let $\Omega = B_1(0)$ and set

$$u(x) = \frac{1}{|x|^l}, \quad x \neq 0.$$

Then, $|Du(x)| = \frac{l}{|x|^{l+1}}$. Let $\varphi \in C_c^\infty(\Omega)$.

$$\int_{\Omega\setminus B_\epsilon(0)} u\partial_i\varphi = -\int_{\Omega\setminus B_\epsilon(0)} \partial_i u\varphi + \int_{\partial B_\epsilon(0)} u\varphi v^i.$$

Assume $l < n-1$. Then,

$$\left|\int_{\partial B_\epsilon(0)} u\varphi v^i\right| \le \|\varphi\|_{L^\infty(\Omega)} \int_{\partial B_\epsilon(0)} \frac{1}{\epsilon^l} \underbrace{d\sigma}_{=\epsilon^{n-1}d\omega} \to 0, \quad \epsilon \to 0.$$

If $l < n-1$, then $Du \in L^1(\Omega)$ so

$$\int_\Omega u\partial_i\varphi = -\int_\Omega \partial_i u\varphi$$

Also, $\frac{l}{|x|^{l+1}} \in L^p(\Omega)$ iff $(l+1)p < n$. Thus, $u \in W^{1,p}(\Omega)$ iff $l < \frac{n-p}{p}$.

Example 9.15. Let $\{a_k\} \subset \Omega = B_1(0)$ be a dense countable subset. Put $u(x) = \sum_{k=1}^\infty \frac{1}{2^k}|x-a_k|^{-l}$. For $l < \frac{n-p}{p}$, $n \in W^{1,p}(\Omega)$. Note that u is unbounded on each open set for $0 < l < \frac{n-p}{p}$.

Theorem 9.16. *$W^{k,p}(\Omega)$ is a Banach space. $H^k(\Omega)$ is a Hilbert space.*

Proof. Obviously $\|\cdot\|_{k,p}$ is indeed a norm and $\|u\|_{k,p} = 0$ iff $u = 0$ a.e..

Let $\{u_k\}$ be a Cauchy sequence in $W^{k,p}(\Omega)$. Then $\{D^\alpha u_k\}$ is Cauchy in $L^p(\Omega)$ for each $|\alpha| \le k$, so there exist functions u_α such that $D^\alpha u_k \to u_\alpha$ in $L^p(\Omega)$. Let $\varphi \in C_c^\infty(\Omega)$,

$$\int_\Omega uD^\alpha\varphi = \lim\int_\Omega u_k D^\alpha\varphi = (-1)^{|\alpha|}\lim\int_\Omega D^\alpha u_k\varphi = (-1)^{|\alpha|}\int_\Omega u_\alpha\varphi,$$

so $u_\alpha = D^\alpha u$ and $u \in W^{k,p}(\Omega)$.

□

Theorem 9.17. *$W^{k,p}(\Omega)$ is separable if $1 \le p < \infty$, and is uniformly convex and reflexive if $1 < p < \infty$.*

Proof. Let $\mu(k,n)$ be the number of multi-indices α such that $|\alpha| \le k$, and for each α let Ω_α be a copy of Ω, so the $\mu(k,n)$ domains Ω_α are disjoint. Set

$$\Omega_{(k)} := \bigcup_{|\alpha|\le k}\Omega_\alpha.$$

Given $u \in W^{k,p}(\Omega)$, let v be the function on $\Omega_{(k)}$ that coincides with $D^\alpha u$ in Ω_α. The map $\Gamma : W^{k,p}(\Omega) \to L^p(\Omega_{(k)}) u \mapsto v$ is an isometry. Because $W^{k,p}(\Omega)$ is complete, the image X of Γ is a closed subspace of $L^p(\Omega_{(k)})$. The result follows from $W^{k,p}(\Omega) = \Gamma^{-1}(X)$.

□

9.3. **Approximation by smooth functions.** Given a subset $A \subset \mathbb{R}^n$ and a collection $\mathcal{O}$ of open sets covering A, $A \subset \bigcup_{U\in\mathcal{O}} U$, recall that a (smooth) partition of unity of A subordinate to $\mathcal{O}$ is a collection Ψ of $C_c^\infty(\mathbb{R}^n)$ functions ψ such that

(i) $0 \le \psi \le 1$;
(ii) If $\kappa \subset\subset A$, all but finitely many ψ vanish identically on κ;
(iii) For every $\psi \in \Psi$ there exists a $U \in \mathcal{O}$ such that $\text{supp}(\psi) \subset U$;
(iv) For every $x \in A$, $\sum_{\psi\in\Psi}\psi(x) = 1$.

It is a standard theorem in topology that smooth partitions of unit exist.

Proposition 9.18. *If $u \in W^{k,p}(\Omega)$, $1 \le p < \infty$, then $u_\epsilon \to u$ in $W^{k,p}_{loc}(\Omega)$ as $\epsilon \to 0$, where u_ϵ is the regularization of u.*

Proof. This follows immediately from the properties of u_ϵ. ☐

The previous proposition is a local approximation by smooth functions. The next theorem improves this to a global approximation.

Theorem 9.19. *$C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$, $1 \le p < \infty$.*

Proof. For $j = 1, 2, \dots$ set

$$\Omega_j := \left\{ x \in \Omega \;\middle|\; |x| < j \text{ and } \operatorname{div}(x, \partial\Omega) > \frac{1}{j} \right\}$$

and $\Omega_{-1} = \Omega_0 = \emptyset$. Set

$$U_j := \Omega_{j+1} \cap (\mathbb{R}^n \setminus \bar{\Omega}_{j-1})$$

Then $\mathcal{O} = \{U_j\}$ covers Ω. Let Ψ be a corresponding partition of unity and let ψ_j be the sum of the (finitely many) $\psi \in \Psi$ whose supports are in U_j. Then $\psi_j \in C_c^\infty(U_j)$ and

$$\sum_{j=1}^{\infty} \psi_j(x) = 1.$$

If $0 < \epsilon < \frac{1}{(j+1)(j+2)}$, then $(\psi_j u)_\epsilon$ has support in

$$V_j := \Omega_{j+2} \cap (\mathbb{R}^n \setminus \Omega_{j-2}) \subset\subset \Omega.$$

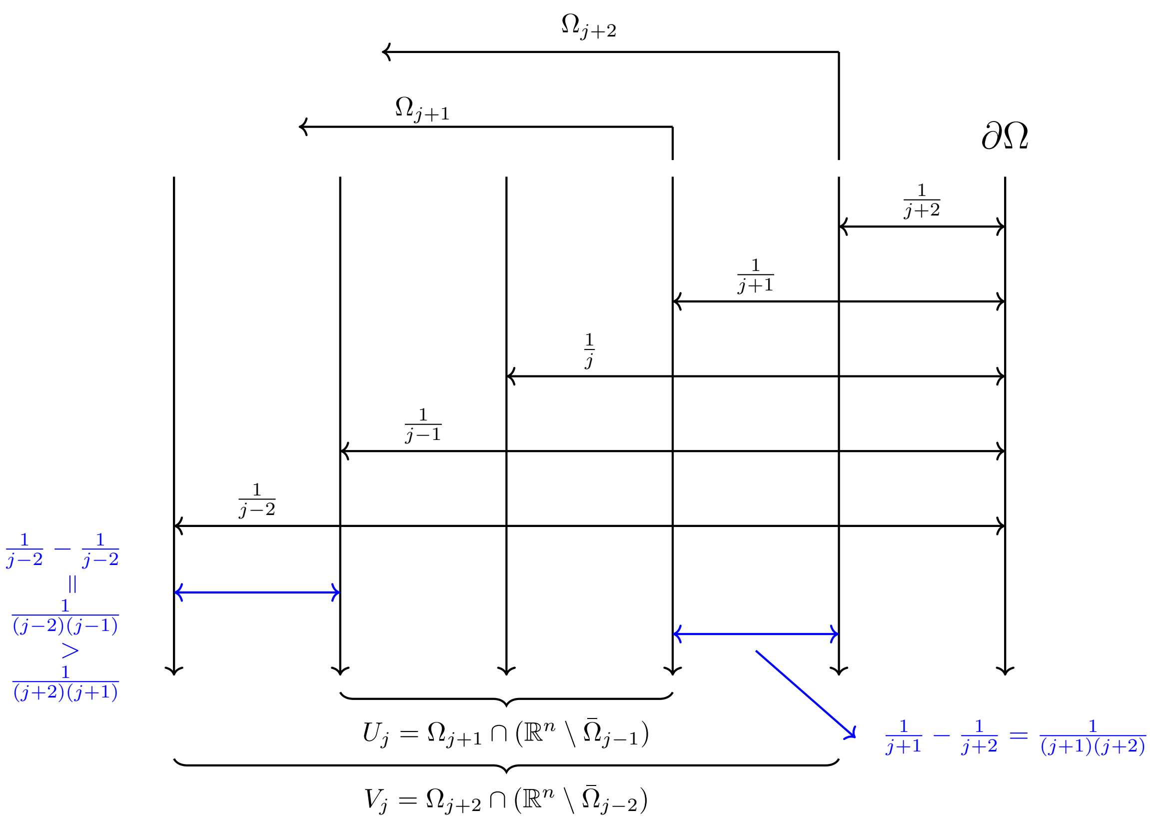


FIGURE 12. Visual Demonstration of Ω

And since $\psi_j u \in W^{k,p}(\Omega)$, we can choose (by previous local proposition) ϵ_j such that

$$\|(\psi_j u)_{\epsilon_j} - \psi_j u\|_{k,p} = \|(\psi_j u)_{\epsilon_j} - \psi_j u\|_{W^{k,p}(v_j)} < \frac{\epsilon}{2^j}$$

where $\epsilon > 0$ is given. Set

$$\psi = \sum_{j=1}^{\infty} (\psi_j u)_{\epsilon_j}$$

Then $\psi \in C^\infty(\Omega)$ since for any $\Omega' \subset\subset \Omega$ only finitely many terms in the sum are non-zero. Now, for $x \in \Omega_j$

$$u(x) = \sum_{i=1}^{j+2} \psi_i(x) u(x), \quad \psi(x) = \sum_{i=1}^{j+2} (\psi_i u)_{\epsilon_i}(x)$$

Therefore,

$$\|u - \psi\|_{W^{k,p}(\Omega_j)} \leq \sum_{i=1}^{j+2} \|(\psi_i u)_{\epsilon_i} - \psi_i u\|_{k,p} < \epsilon,$$

and the result follows from the monotone convergence theorem. (Apply it to $f_j := \chi_{\Omega_j} \sum_{|\alpha| \leq k} |D^\alpha(u - \psi)|^p$ to get

$$\|u - \psi\|_{W^{k,p}(\Omega_j)} \to \|u - \psi\|_{W^{k,p}(\Omega)}$$

as $j \to \infty$.

□

One often sees Sobolev spaces defined as

$$\tilde{W}^{k,p}(\Omega) := \text{ completion of } C^k(\Omega) \text{ with respect to the } \|\cdot\|_{k,p} \text{ norm.}$$

We observe that $\tilde{W}^{k,p}(\Omega) \subset W^{k,p}(\Omega)$. For the set $X := \{u \in C^k(\Omega) \mid \|u\|_{k,p} < \infty\}$, it is contained in $W^{k,p}(\Omega)$. Because $W^{k,p}(\Omega)$ is complete, the identity map on X extends to an isometry between $\tilde{W}^{k,p}(\Omega)$ and the closure of X in $W^{k,p}(\Omega)$. We identify $\tilde{W}^{k,p}(\Omega)$ with this closure.

In view of the previous theorem, any element in $W^{k,p}(\Omega)$ is a limit point of a sequence of smooth functions, i.e., any $u \in W^{k,p}(\Omega)$ belongs to the closure of $C^\infty(\Omega)$ w.r.t. the $\|\cdot\|_{k,p}$ norm. Thus,

$$W^{k,p}(\Omega) \subset W^{k,p}(\tilde{\Omega}), \quad 1 \leq p < \infty$$

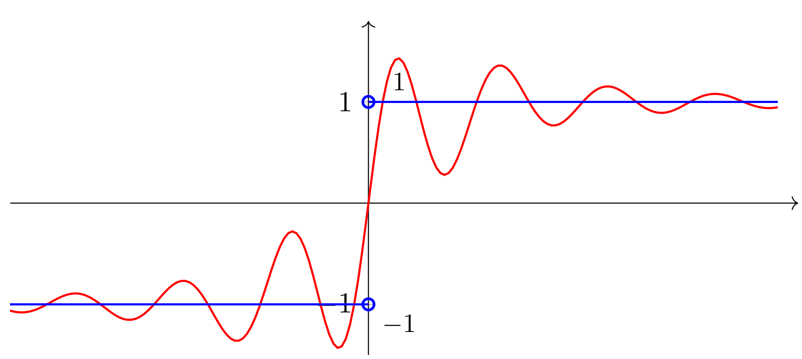
Hence

$$W^{k,p}(\Omega) = W^{k,p}(\tilde{\Omega})$$

This equivalence was first proven by Meyers and Serrin in '64.

This theorem *cannot* be extended to $p = \infty$:

Example 9.20. $\Omega = (-1, 1)$, $u(x) = |x|$. $u'(x) = \frac{x}{|x|}$ for $x \neq 0$. If $0 < \epsilon < \frac{1}{2}$, there does not exist a $\psi \in C^1(\Omega)$ such that $\|\psi' - u'\|_{L^\infty(\Omega)} < \epsilon$.



The density $C^\infty(\Omega)\cap W^{k,p}(\Omega)\to W^{k,p}(\Omega)$ makes *no assumption* on $\partial\Omega$. On the other hand, the derivatives of the smooth functions approximating $u\in Wk,p(\Omega)$ can become unbounded near $\partial\Omega$. We thus ask if it is possible to show that $C^\infty(\bar\Omega)\cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$ or, more generally, if $C^m(\bar\Omega)\cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$. Without further assumptions on $\partial\Omega$, the answer is no.

Example 9.21. $\Omega=\{(x,y)\in\mathbb{R}^2\big|0<|x|<1,0<y<1\}$ (strictly speaking this is not a domain, but the argument can be adapted to a domain).

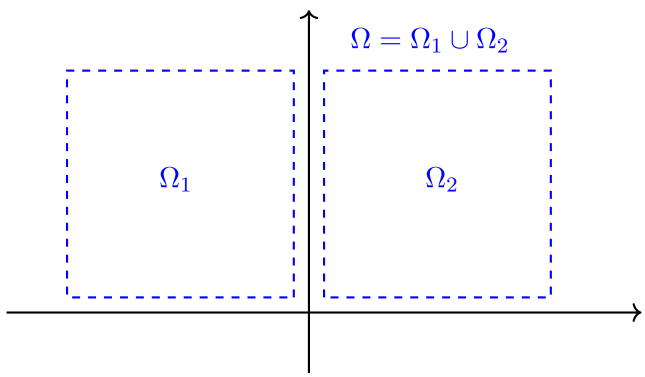


FIGURE 13. Graph of Ω

Put $u(x,y)=\begin{cases}1, & x>0,\\ 0, & x<0.\end{cases}$ Then, $u\in W^{1,p}(\Omega)$. Suppose that there exist $\psi\in C^1(\bar\Omega)$ such that $\|u-\psi\|_{1,p}<\epsilon$. Let $L=\bar\Omega_1$, $R=\bar\Omega_2$, so $\bar\Omega=L\cup R$. We have $\|\psi\|_{L^1(L)}\le\|\psi\|_{L^p(L)}<\epsilon$, $\|1-\psi\|_{L^1(R)}\le\|1-\psi\|_{L^p(R)}<\epsilon$. Set $\varphi=\int_0^1\psi(x,y)dy$. Since

$$\begin{aligned}\int_{-1}^0\varphi(x)dx&=\int_{-1}^0\int_0^1\psi(x,y)dydx\le\|\psi\|_{L^1(L)}<\epsilon,\\ \int_0^1\varphi(x)dx&=\int_0^1\int_0^1\psi(x,y)dydx=\int_0^1\int_0^1(\psi-1)dydx+1\\ &\ge-\int_0^1\int_0^1|\psi-1|dydx+1=-\|\psi-1\|_{L^1(R)}+1>-\epsilon+1\end{aligned}$$

We conclude that there must exist $-1\le a<0$ and $0<b\le1$ such that $\varphi(a)<\epsilon,\varphi(b)>1-\epsilon$. Thus

$$\begin{aligned}1-\epsilon-\epsilon<\varphi(b)-\varphi(a)&=\int_a^b\varphi'(x)dx\le\int_{\bar\Omega}|\partial_x\psi(x,y)|dxdy\\ &\le2^{\frac1q}\|\partial_x(\psi)\|_{L^p(\bar\Omega)}=2^{\frac1q}\|\partial_x\psi-\partial_xu\|_{L^p(\Omega)}\\ &\le2^{\frac1q}\|\psi-u\|_{W^{1,p}(\Omega)}<2^{\frac1q}\epsilon,\end{aligned}$$

where we used $\frac1q+\frac1p=1$ and $Du=0$. Thus $1<(2+2^{\frac1q})\epsilon$, which cannot be true for small $\epsilon>0$.

The problem above is caused by the fact that Ω is on both sides of part of its boundary. The following condition prevents this.

Definition 9.22. A domain Ω satisfies the **segment condition** if for every $x\in\partial\Omega$ there exists a neighborhood U_x and a nonzero vector y_x such that if $z\in\bar\Omega\cap U_x$, then $z+ty_x\in\Omega,0<t<1$.

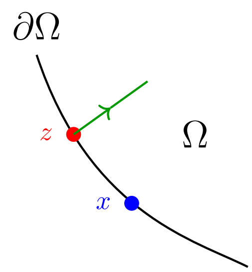


FIGURE 14. Segment Condition I

Look at points z on $\partial\Omega$ and move a bit inside the domain along the segment $z + ty_x$, we stay inside the domain, and this is uniform on each U_x.

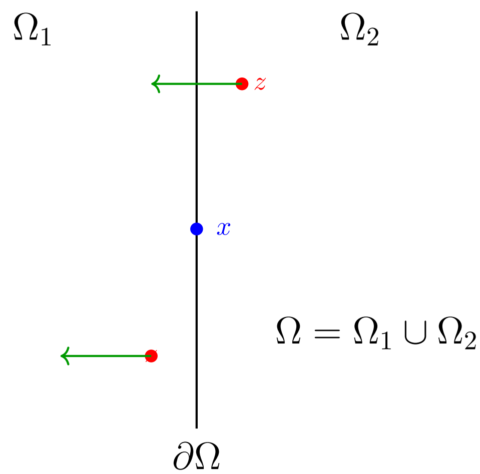


FIGURE 15. Segment Condition II

No matter how small we fix U_x and y_x, for some z close to $\partial\Omega$ the line crosses $\partial\Omega$. (y_x has to be uniform on U_x).

If the segment condition is satisfied, then $\partial\Omega$ must be $n-1$ dimensional and Ω cannot lie on both sides of $\partial\Omega$. (for $\partial\Omega \neq \emptyset$).

Theorem 9.23. *If Ω satisfies the segment condition, then the set of restrictions to Ω of functions in $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\Omega)$, $1 \leq p < \infty$. In particular, $C^\infty(\bar{\Omega})$ is dense in $W^{k,p}(\Omega)$, $1 \leq p < \infty$.*

Proof. We begin with some reductions.

Let $\psi \in C_c^\infty(\mathbb{R}^n)$ satisfy: $\psi(x) = 1$ if $|x| \leq 1$, $\psi(x) = 0$ if $|x| \geq 2$, $|D^\alpha \psi(x)| \leq C$ for $|\alpha| \leq k$. Set $\psi_\epsilon(x) := \psi(\epsilon x)$. Then, $\psi_\epsilon(x) = 1$ for $|x| \leq \frac{1}{\epsilon}, \psi_\epsilon(x) = 0$ for $|x| \geq \frac{2}{\epsilon}$, $|D^\alpha \psi_\epsilon(x)| \leq C\epsilon^{|\alpha|} \leq C$ for $0 < \epsilon \leq 1$. If $u \in W^{k,p}(\Omega)$, then $u_\epsilon := \psi_\epsilon u \in W^{k,p}(\Omega)$, has bounded support, and

$$|D^\alpha u_\epsilon| \leq C \sum_{\beta \leq \alpha} D^\beta u D^{\alpha-\beta} \psi_\epsilon \leq C \sum_{\beta \leq \alpha} D^\beta u.$$

Set $\Omega_\epsilon := \{x \in \Omega \mid |x| > \frac{1}{\epsilon}\}$,

$$\|u - u_\epsilon\|_{k,p} = \|u \underbrace{(1-\psi_\epsilon)}_{=0, |x| \leq \frac{1}{\epsilon}}\|_{k,p} = \|u - u_\epsilon\|_{W^{k,p}(\Omega_\epsilon)}$$

$$\leq C\|u\|_{W^{k,p}(\Omega_\epsilon)} + C\|u_\epsilon\|_{W^{k,p}(\Omega_\epsilon)} \leq C\|u\|_{W^{k,p}(\Omega_\epsilon)} \to 0$$

as $\epsilon \to 0$. Thus, u can be approximated by functions with compact support (but not smooth at this point).

We can thus assume that $\kappa := \{x \in \Omega \mid u(x) \neq 0\}$ has a bounded support. Set

$$F := \bar{\kappa} \setminus \bigcup_{x \in \partial\Omega} U_x,$$

where U_x are the sets in the definition of the segment condition. F is compact and contained in Ω. So, we can find U_0 such that $F \subset\subset U_0 \subset\subset \Omega$ and among the U_x's finitely many $U_1, ..., U_l$ such that $\bar{\kappa} \subset U_0 \cup ... \cup U_l$. We can find further open sets v_j, $j = 0, ..., l$, such that $v_j \subset\subset U_j$, $\bar{\kappa} \subset v_0 \cup ... \cup v_l$.

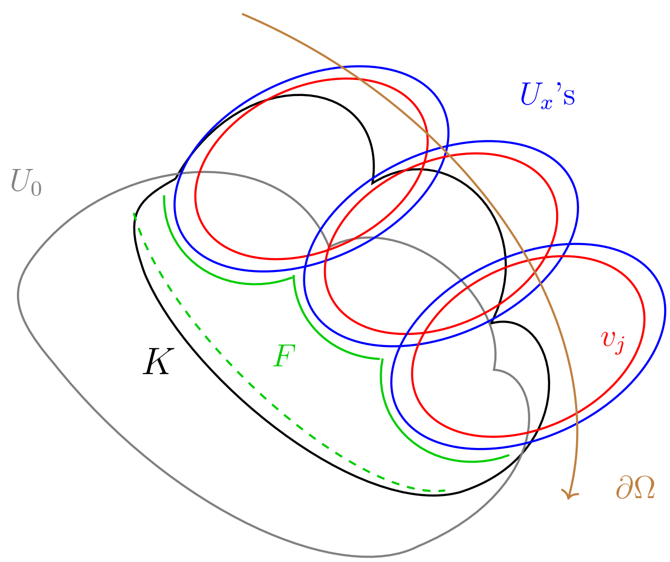


Figure 16. Compact $F \subset\subset U_0 \subset\subset \Omega$

Let Ψ be a partition of unity subordinate to the v_j's. Let ψ_j be the sum of the (finitely many) ψ's whose supports are in v_j. Let $u_j := \psi_j u$. If for each $j = 0, .., l$ we can find $\psi_j \in C_c^\infty(\mathbb{R}^n)$ such that $\|u_j - \psi_j\|_{k,p} < \frac{\epsilon}{l+1}$, then we are done since, with $\psi := \sum_{j=0}^{l} \psi_j$,

$$\|u - \psi\|_{k,p} = \Big\| \sum_{\psi \in \Psi} \psi u - \sum_{j=0}^{l} \psi_j \Big\|_{k,p} = \Big\| \sum_{j=0}^{l} (u_j - \psi j) \Big\|_{k,p}$$

$$\leq \sum_{j=0}^{l} \|u_j - \psi_j\|_{k,p} < \epsilon.$$

For $j = 0$, we find ψ_0 by direct regularization (see above Prop. prior to the theorem on density of $C^\infty(\Omega)$ functions).

Fix a $j \in \{1, ..., l\}$. Since u_j has support in $v_j \cap \bar{\kappa}$, we can extend it to be $\equiv 0$ outside Ω. In particular, $u_j \in W^{k,p}(\mathbb{R}^n \setminus \Gamma)$, where

$$\Gamma := \bar{v}_j \cap \partial\Omega$$

Let y be the non-zero vector corresponding to U_j in the segment condition. Pick t such that

$$0 < t < \min\left\{1, \frac{\operatorname{dist}(v_j, \mathbb{R}^n \setminus U_j)}{|y|}\right\}$$

and set $\Gamma_t := \{x - ty \mid x \in \Gamma\}$.

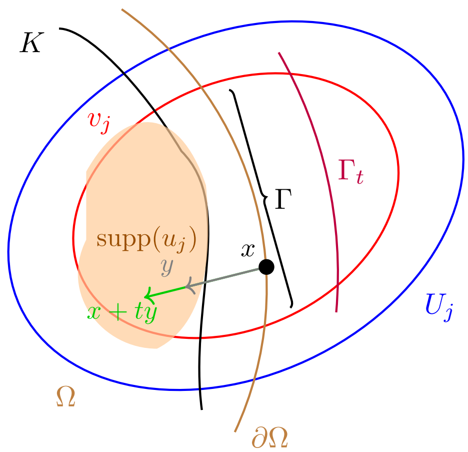


FIGURE 17. Segment Condition III

The segment condition gives that $\Gamma_t \cap \bar{\Omega} = \emptyset$, and our choice of t also guarantees that $\Gamma_t \subset U_j$. Since $u_j \in W^{k,p}(\mathbb{R}^n \setminus \Gamma)$, the function

$$u_{j,t}(x) := u_j(x + ty)$$

belongs to $W^{k,p}(\mathbb{R}^n \setminus \Gamma_t)$. Translations are continuous in L^p, so $D^\alpha u_{j,t} \to D^\alpha u_j$ in $L^p(\Omega)$ as $t \to 0$. Hence $u_{j,t} \to u_j$ in $W^{k,p}(\Omega)$ as $t \to 0$. It thus suffices to find $\psi_j \in C_c^\infty(\mathbb{R}^n)$ such that $\|u_{j,t} - \psi_j\|_{k,p}$ is small. Since

$$\Omega \cap U_j \subset\subset \mathbb{R}^n \setminus \Gamma_t,$$

we can take a regularization of $u_{j,t}$ as in the proposition mentioned above.

□

The key idea of the proof is that we want to regularize u to get a C^∞ function. As we want to include boundary points, the usual regularization would involve averaging u outside Ω. However, we do not need to make the regularization at $x \in \partial\Omega$ with an average centered at x; we can average about another point in the interior. This is what the translations of the boundary does. This is a useful idea when dealing with boundaries.

Corollary 9.24.

$$W_0^{k,p}(\mathbb{R}^n) = W^{k,p}(\mathbb{R}^n)$$

Remark 9.25. The corollary is in general not true for Ω. Also, since Lipchitz boundaries (properly defined for non-compact $\partial\Omega$) satisfy the segment condition, we obtain the density result when $\partial\Omega$ is Lipschitz.

We will now use the density of smooth functions to study coordinate transformations.

Theorem 9.26. *Let $\Omega, \mathcal{D}$ be domains in $\mathbb{R}^n$. Suppose that there exists a one-to-one and onto $\Psi : \Omega \to \mathcal{D}$ such that $\Psi^j \in C^k(\Omega)$, $(\Psi^{-1})^j \in C^k(\mathcal{D})$ have bounded derivatives, $j = 1, ..., n$, and $\frac{1}{\mathcal{C}} \leq |det\ D\Psi| + |det\ D\Psi^{-1}| \leq \mathcal{C}$, for some $\mathcal{C} > 1, k \geq 1$. Given $u \in W^{k,p}(\mathcal{D})$, define $\tilde{\Psi}(u) : \Omega \to \mathbb{R}$ by $\tilde{\Psi}(u)(x) = u(\Psi(x))$. Then, $\tilde{\Psi}$ transforms $W^{k,p}(\Omega)$ boundedly onto $W^{k,p}(\mathcal{D})$ and has bounded inverse, $1 \leq p < \infty$.*

Proof. The map is well-defined for a.e. defined functions since $k \geq 1$. Let $\{u_j\} \subset C^\infty(\Omega)$ converge to u in $W^{k,p}(\mathcal{D})$. Let $|\alpha| \leq k$ be a multi-index. Successive applications of the chain rule and product rule give that $(y = \Psi(x))$

$$\begin{aligned} D^\alpha \tilde{\Psi}(u_j)(x) &= \sum_{p \leq \alpha} p_{\alpha\beta}(x) D_y^\beta u_j(y) \\ &= \sum_{p \leq \alpha} p_{\alpha\beta}(x) \tilde{\Psi}(D^\beta u_j)(x), \end{aligned}$$

where $p_{\alpha\beta}$ is a polynomial of degree $\leq |\beta|$ in derivatives of Ψ^j, $j = 1, ..., n$ of order $\leq |\alpha|$. Let $\psi \in C_c^\infty(\Omega)$.

$$\begin{aligned} (-1)^{|\alpha|} \int_\Omega \tilde{\Psi}(u_j)(x) D^\alpha \psi(x) dx \quad &= \quad \sum_{\beta \leq \alpha} \int_\Omega p_{\alpha\beta}(x) \tilde{\Psi}(D^\beta u_j)(x) \psi(x) dx \\ =(-1)^{|\alpha|} \int_{\mathcal{D}} \underbrace{\tilde{\Psi}(u_j)(\Psi^{-1}(y))}_{u_j(y)} D^\alpha \psi(\Psi^{-1}(y)) |\det\, D\Psi^{-1}(y)| dy \quad & \quad \sum_{\beta \leq \alpha} \int_{\mathcal{D}} p_{\alpha\beta}(\Psi^{-1}(y)) \underbrace{\tilde{\Psi}(D^\beta u_j)(\Psi^{-1}(y))}_{D^\beta u_j(y)} \psi(\Psi^{-1}(y)) |\det\, D\Psi^{-1}(y)| dy \end{aligned}$$

Since $D^\beta u_j \to u$, we can replace u_j by u above and change variables back to get

$$(-1)^{|\alpha|} \int_\Omega \tilde{\Psi}(u)(x) D^\alpha \psi(x) dx = \sum_{\beta_p \leq \alpha} \int_\Omega p_{\alpha\beta}(x) \tilde{\Psi}(D^\beta u)(x) \psi(x) dx,$$

so $\tilde{\Psi}(u)$ is $W^k(\Omega)$ and

$$D^\alpha \tilde{\Psi}(u)(x) = \sum_{\beta \leq \alpha} p_{\alpha\beta}(x) \tilde{\Psi}(D^\beta u)(x).$$

Then,

$$\begin{aligned} \int_\Omega |D^\alpha \tilde{\Psi}(u)(x)|^p dx &\leq \mathcal{C} \max_{|\beta| \leq |\alpha|} \sup_{x \in \Omega} |p_{\alpha\beta}(x)|^p \int_\Omega \underbrace{|\tilde{\Psi}(D^\beta u)(x)|^p}_{|(D^\beta u)(\Psi(x))|^p} dx \\ &\leq \mathcal{C} \max_{|\beta| \leq |\alpha|} \int_{\mathcal{D}} |D^\beta u(y)|^p |\det\, D\Psi^{-1}(y)| dy \\ &\leq \mathcal{C} \|u\|_{W^{k,p}(\mathcal{D})} \implies \|\tilde{\Psi}(u)\|_{W^{k,p}(\Omega)} \leq \mathcal{C} \|u\|_{W^{k,p}(\mathcal{D})}. \end{aligned}$$

Repeating the argument for Ψ^{-1} gives the result. □

9.4. **Extensions.** Given $u \in W^{k,p}(\Omega)$, can we extend outside Ω? In other words, does there exist $\tilde{u} \in W^{k,p}(\mathbb{R}^n)$ such that $\tilde{u} = u$ in Ω? We begin by making this notion more precise:

Definition 9.27. Let $\Omega \subset \mathbb{R}^n$ be a domain, $k \geq 0$ an integer and $1 \leq p < \infty$. A linear map $E : W^{k,p}(\Omega) \to W^{k,p}(\mathbb{R}^n)$ is called a **(k,p)-extension** (or simply extension if k, p are understood), in Ω if there exist a constant $\kappa = \kappa(k,p)$ such that

(i) $Eu(x) = u(x)$ a.e. in Ω
(ii) $\|Eu\|_{W^{k,p}(\mathbb{R}^n)} \leq \kappa \|u\|_{W^{k,p}(\Omega)}$

for all $u \in W^{k,p}(\Omega)$. E is called a **strong k-extension** (or strong extension if k is understood) in Ω if it is a linear operator mapping a.e. defined functions in Ω to a.e. defined functions in $\mathbb{R}^n$, and such that, for every $1 \leq p < \infty$ and every $0 \leq m \leq k$, the restriction of E to $W^{m,p}(\Omega)$ is a (k,p)-extension. E is called a **total extension** in Ω if it is a strong extension for every k (necessarily extends from $C^k(\bar{\Omega})$ to $C^k(\mathbb{R}^n)$).

Lemma 9.28. *Let $\Omega = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x^n > 0\}$. Then, there exists a strong k-extension operator in Ω. Moreover, for every multi-index α, $|\alpha| \leq k$, there exists a strong $(k - |\alpha|)$-extension operator E_α in Ω such that*

$$D^\alpha E u = E_\alpha D^\alpha u.$$

Proof. Set

$$Eu(x) = \begin{cases} u(x), & x^n > 0, \\ \sum_{j=1}^{k+1} \lambda_j u(x^1, ..., x^{n-1}, -jx_n), & x^n < 0, \end{cases}$$

$$E_\alpha u(x) = \begin{cases} u(x), & x^n > 0, \\ \sum_{j=1}^{k+1} (-j)^{\alpha_n} \lambda_j u(x^1, ..., x^{n-1}, -jx_n), & x^n < 0, \end{cases}$$

where $\lambda_1, ..., \lambda_{k+1}$ are the unique solution to

$$\sum_{j=1}^{k+1} (-j)^l \lambda_j = 1, \quad l = 0, ..., k$$

If $u \in C^k(\bar{\mathbb{R}}^n_+)$, then $Eu \in C^k(\mathbb{R}^n)$ and

$$D^\alpha Eu = E_\alpha D^\alpha u, \quad |\alpha| \leq k.$$

(E.g., $k = 1, \lambda_1 + \lambda_2 = 1, -\lambda_1 - 2\lambda_2 = 1, \lambda_1 = 3, \lambda_2 = -2, Eu(X) = 3u(x^1, ..., -x^n) - 2u(x^1, ..., -2x^n)$ $\implies Eu|_{\{x^n=0\}} = u,\ \partial_i Eu|_{\{x^n=0\}} = \partial_i u,\ i < n,\ \partial_n Eu|_{\{x^n=0\}} = \partial_u u|_{\{x^n=0\}}$).

Then,

$$\int_{\mathbb{R}^n} |D^\alpha Eu|^p = \int_{\mathbb{R}^n_+} |D^\alpha u|^p + \int_{\mathbb{R}^n_-} \left| \sum_{j=1}^{k+1} (-j)^{\alpha_n} \lambda_j D^\alpha u(x^1, ..., x^{n-1}, -jx^n) \right|^p$$
$$\leq \mathcal{C} \int_{\mathbb{R}^n} |D^\alpha u|^p$$

By density of $C^\infty(\bar{\mathbb{R}}^n_+)$ in $W^{k,p}(\mathbb{R}^n_+)$, we obtain that the inequality is valid for $u \in W^{k,p}(\mathbb{R}^n)$, $0 \leq m \leq k$, so E is a strong k-extension. A similar argument, noting that $D^\beta E_\alpha = E_{\alpha+\beta} u$, shows that E_α is a strong $(k - |\alpha|)$-extension.

□

To obtain the result for general Ω, we need conditions on $\partial\Omega$.

Definition 9.29. A domain $\Omega \subset \mathbb{R}^n$ satisfies the **strong local Lipschitz condition** if there exist $\delta > 0$, $M > 0$, a locally finite open cover $\{U_j\}$ of $\partial\Omega$, for each j a real valued function f_j of $n-1$ variables, such that

(i) for some finite R, every collection of $R+1$ open sets U_j has empty intersection;
(ii) for every pair
$$x, y \in \Omega_\delta := \{z \in \Omega \mid \mathrm{dist}(z, \partial\Omega) < \delta\}$$
such that $|x - y| < \delta$, there exists j such that
$$x, y \in V_j := \{z \in U_j \mid \mathrm{dist}(z, \partial U_j) > \delta\}$$
(iii) each f_j satisfies a Lipschitz condition with constant M
(iv) For some Cartesian coordinate system
$$(\theta_{j,1}, ..., \theta_{j,n}) \text{ in } U_j, \quad \Omega \cap U_j = \{\theta_{j,n} < f_j(\theta_{j,1}, ..., \theta_{j,n-1})\}.$$

Remark 9.30. For bounded Ω, these conditions reduce to the requirement that Ω has a locally Lipschitz boundary.

Definition 9.31. A domain $\Omega \subset \mathbb{R}^n$ satisfies the **uniform C^k condition** if there exists a locally finite open cover $\{U_j\}$ of $\partial\Omega$, a sequence $\{\psi_j\}$ of C^k functions taking U_j onto $B_1(0)$, with C^k inverses $\psi_j^{-1} = \phi_j$ such that

(i) for some finite R, every collection of $R+1$ open sets U_j has empty intersection;

(ii) for some $\delta > 0$,

$$\Omega_\delta := \{z \in \Omega \mid \operatorname{dist}(z, \partial\Omega) < \delta\} \subset \bigcup_{j=1}^{\infty} \phi_j(B_{\frac{1}{2}}(0))$$

(iii) for each j,

$$\psi_j(U_j \cap \Omega) = B_1(0) \cap \{z^n > 0\}$$

(iv) there exists a constant M such that

$$|D^\alpha \phi_j^i(x)| \leq M \quad \forall x \in U_j$$

$$|D^\alpha \psi_j^i(y)| \leq M \quad \forall y \in B_1(0)$$

and every $|\alpha| \leq k$.

Theorem 9.32. *Let Ω satisfy the uniform C^k condition and $\partial\Omega$ be bounded. Then, there exists a strong k-extension operator E in Ω. Moreover, if α and β are multi-indices with $|\beta| \leq |\alpha| \leq k$, then there exists a linear operator $E_{\alpha\beta}$ continuous from $W^{l,p}(\Omega)$ into $W^{l,p}(\mathbb{R}^n)$, $1 \leq l \leq k - |\alpha|$, $1 \leq p < \infty$, such that*

$$D^\alpha(Eu)(x) = \sum_{|\beta| \leq |\alpha|} E_{\alpha\beta}(D^\beta u)(x)$$

for all $u \in W^{|\alpha|,p}(\Omega)$.

Theorem 9.33. *Let Ω satisfy the strong Lipschitz condition. Then, there exists a total extension operator in Ω.*

The proofs can be found in Stein, E. "Singular integrals and differentiability properties of functions." For the first theorem, the basic idea is to use our theorem on coordinate transformations to reduce the problem to $\mathbb{R}^n_+$, for which we already have the result. For the second theorem, the key idea is to use that the distance function to $\partial\Omega$ is Lipschitz, and then work with suitable approximations of the distance function.

Remark 9.34.

- The assumption of $\partial\Omega$ bounded in the first theorem is not essential and can be removed in most reasonable cases.
- We have the implications:
  uniform C^k regularity ($k \geq 2$) $\implies$ strong local Lipschitz $\implies$ segment condition.

9.5. **The Sobolev embedding theorem.** Our goal in this section is to answer the following question: if $u \in W^{k,p}(\Omega)$, does u belong (in a non-trivial way) to some other function space?

In order to answer this question, it is helpful to establish some conventions about $C^k(\bar{\Omega})$. So far, we have only considered $C^k(\bar{\Omega})$ when Ω is bounded, in which case $C^k(\bar{\Omega})$ is a Banach space. (An exception was the density of $C^\infty(\bar{\Omega})$ in $W^{k,p}(\Omega)$ when Ω satisfies the segment condition. But in that case the statement was that restrictions from $C_c^\infty(\mathbb{R}^n)$ to Ω are dense in $W^{k,p}(\Omega)$, in which case we can entirely avoid talking about $C^\infty(\bar{\Omega})$.) For our purposes, we are interested in considering other function spaces that are Banach spaces. However, the set of functions continuously differentiable up to order k on $\bar{\Omega}$ might not be a Banach space if Ω is not bounded. Thus, we henceforth adopt the following.

Definition 9.35. Let $\Omega \subset \mathbb{R}^n$ be a domain and $k \geq 0$ an integer. We define $\mathbf{C^k_B(\Omega)}$ as the space of all $u \in C^k(\Omega)$ such that $D^\alpha u$ is bounded, $0 \leq |\alpha| \leq k$. $C^k_B(\Omega)$ is a Banach space with norm

$$\|u\|_{C^k_B(\Omega)} = \sum_{|\alpha| \leq k} \sup_{x \in \Omega} |D^\alpha u(x)|.$$

If $u \in C^0(\Omega)$ is bounded and uniformly continuous, it admits a unique, bounded continuous extension to $\bar{\Omega}$. We define $\mathbf{C^k(\bar{\Omega})}$ to be the space of those $u \in C^k(\Omega)$ such that $D^\alpha u$ is bounded and uniformly continuous on Ω (so in particular $D^\alpha u$ extends to $\bar{\Omega}$), $0 \leq |\alpha| \leq k$. $C^k(\bar{\Omega})$ is a Banach space with the norm

$$\|u\|_{C^k(\bar{\Omega})} = \sum_{|\alpha| \leq k} \sup_{x \in \bar{\Omega}} |D^\alpha u(x)|.$$

Let $0 < \gamma \leq 1$. We define the **Hölder space** $\mathbf{C^{k,\gamma}(\bar{\Omega})}$ as the subspace of $C^k(\bar{\Omega})$ of those u such that $D^\alpha u$ satisfies a **Hölder condition with exponent** γ, i.e., those u for which there exist a constant $M > 0$ such that

$$|D^\alpha u(x) - D^\alpha u(y)| \leq M|x-y|^\gamma,$$

$0 \leq |\alpha| \leq k$. $C^{k,\gamma}(\bar{\Omega})$ is a Banach space with the norm

$$\|u\|_{C^{k,\gamma}(\bar{\Omega})} := \|u\|_{C^k(\bar{\Omega})} + \max_{0 \leq |\alpha| \leq k} \sup_{\substack{x,y, \in \Omega \\ x \neq y}} \frac{|D^\alpha u(x) - D^\alpha u(y)|}{|x-y|^\gamma}.$$

The quantity $[u]_{\gamma,\Omega} = \sup_{\substack{x,y, \in \Omega \\ x \neq y}} \frac{|u(x)-u(y)|}{|x-y|^\gamma}$ is called the γ **Hölder semi-norm**.

Remark 9.36.

- The notation $C^k(\bar{\Omega})$ can be confusing if $\bar{\Omega}$ is not bounded, as we can have $u \in C^k(\Omega')$, $\bar{\Omega} \subset\subset \Omega'$ (so u is k-times continuously differentiable up to $\partial\Omega$) but $u \notin C^k(\bar{\Omega})$ because the derivatives are not bounded. In other words, by $C^k(\bar{\Omega})$ we always mean the Banach space.
- It is common to write $\|\cdot\|_{C^k_B(\Omega)} = \|\cdot\|_{C^k(\Omega)}$.
- $C^{0,1}(\bar{\Omega})$ is the space of Lipschitz functions on $\bar{\Omega}$.
- If $0 < \delta < \gamma \leq 1$, it is not true that $C^{k,\gamma}(\bar{\Omega}) \subset C^{k,\delta}(\bar{\Omega})$ and, more generally, if $k+\gamma < m+\delta$, it is not true that $C^{m+\delta}(\bar{\Omega}) \subset\subset C^{k+\gamma}(\bar{\Omega})$. For example, take $\Omega = \{(x,y) \in \mathbb{R}^2 \mid y < |x|^{\frac{1}{2}},\ x^2+y^2 < 1\}$.

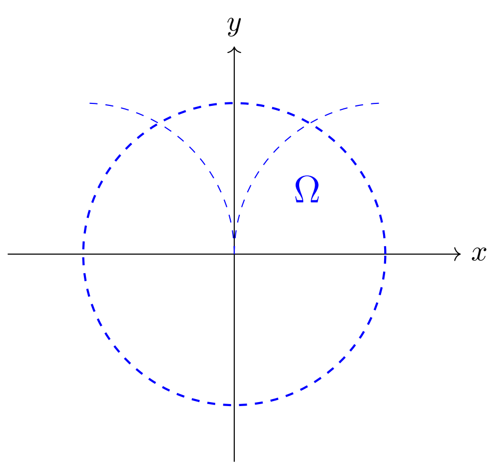


FIGURE 18. $\Omega = \{(x,y) \in \mathbb{R}^2 \mid y < |x|^{\frac{1}{2}},\ x^2+y^2 < 1\}$

Pick $1 < \beta < 2$. Set

$$u(x,y) = \begin{cases} \operatorname{sign} x\, y^\beta, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

Then, $u \in C^1(\bar{\Omega})$, but if $\frac{\beta}{2} < \gamma \leq 1$ then $u \notin C^{0,\gamma}(\bar{\Omega})$, so $C^1(\bar{\Omega}) \not\subset C^{0,\alpha}(\bar{\Omega})$.

- We also have $C^{k,1}(\bar{\Omega}) \not\subset C^{k+1}(\bar{\Omega})$ (since Lipschitz functions need not to be differentiable everywhere e.g., $|x|$) and $C^{k+1}(\bar{\Omega}) \not\subset C^{k,1}(\bar{\Omega})$ (above example).
- For nice domains, however, the above expected inclusions hold.
- But it does hold that $C^{k,\alpha}(\bar{\Omega}) \subset C^{k,\beta}(\bar{\Omega}), \beta < \alpha$
- Hölder spaces capture functions like $x^{\frac{1}{2}}$ (which $\in C^{0,\frac{1}{2}}(\mathbb{R})$): not differentiable, but better than just continuous.
- The following is also used in the literature to define the Hölder norm:

$$\|u\|_{\tilde{C}^{k,\alpha}(\bar{\Omega})} = \|u\|_{C^k(\bar{\Omega})} + \max_{|\alpha|=k} \sup_{\substack{x,y,\in\Omega \\ x\neq y}} \frac{|D^\alpha u(x) - D^\alpha u(y)|}{|x-y|^\gamma}.$$

  The $\tilde{C}^{k,\alpha}(\bar{\Omega})$ and $C^{k,\alpha}(\bar{\Omega})$ norms are not equivalent for arbitrary domains. However, since $\|\cdot\|_{\tilde{C}^{k,\alpha}(\bar{\Omega})} \leq \|\cdot\|_{C^{k,\alpha}(\bar{\Omega})}$, the embeddings into $C^{k,\alpha}(\bar{\Omega})$ that we establish below will automatically give embeddings into $\tilde{C}^{k,\alpha}(\bar{\Omega})$.

We now introduce the types of domains we will consider.

Definition 9.37. Let y be a non-zero vector in $\mathbb{R}^n$. For each $x \neq 0$, let $\sphericalangle(x,y)$ be the angle between the position vector x and y. Given $\varrho > 0$ and $0 < \theta \leq \pi$, the set

$$\mathcal{C} = \mathcal{C}_{y,\varrho,\theta} := \left\{ x \in \mathbb{R}^n \mid x = 0 \text{ or } 0 < |x| \leq \varrho, \ \sphericalangle(x,y) \leq \frac{\theta}{2} \right\}.$$

is called a finite cone of height ϱ, axis direction y, and (aperture) angle θ with vertex at the origin. The set $z + \mathcal{C}$ is a cone with the same properties but vertex at z.

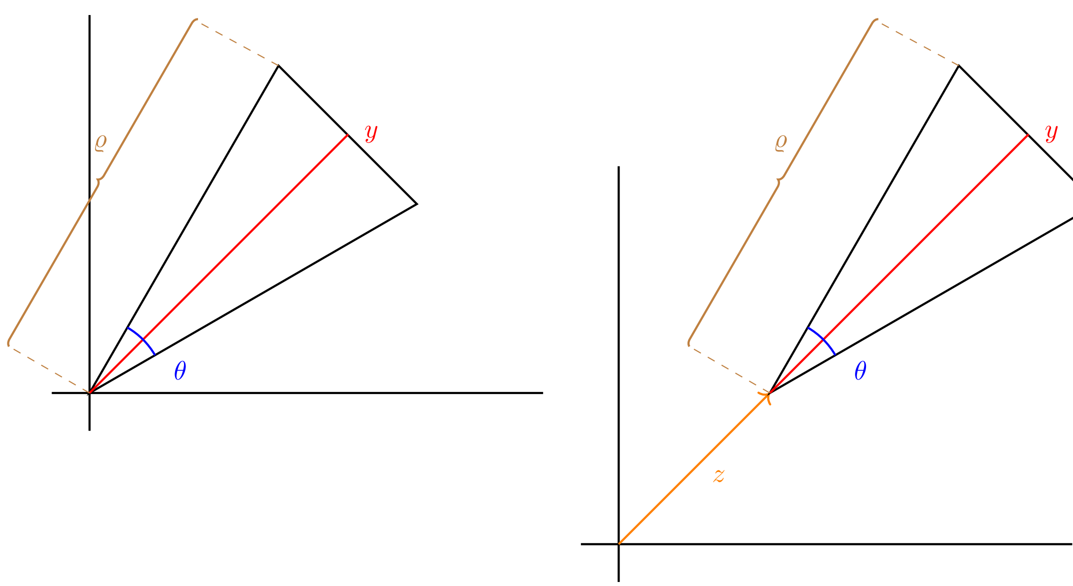


Figure 19. Finite Cones $\mathcal{C}$ and $z + \mathcal{C}$ of Heights ρ

Definition 9.38. A domain in $\mathbb{R}^n$ satisfies the **cone condition** if there exists a finite cone $\mathcal{C}$ such that each $x \in \Omega$ is the vertex of a cone $\mathcal{C}_x$ contained in Ω and congruent to $\mathcal{C}$ (i.e., each $\mathcal{C}_x$ is obtained from $\mathcal{C}$ by rigid motions).

The intuition is that, up to rigid motions, we can fit the cone $\mathcal{C}$ inside Ω, with a vertex on $\partial\Omega$. Thus, the domain can have corners with uniform aperture but cannot become arbitrarily this in one direction. (see examples below).

Definition 9.39. A domain $\Omega \subset \mathbb{R}^n$ satisfies the **uniform cone condition** if there exists a locally finite open cover $\{U_j\}$ of $\partial\Omega$ and a corresponding sequence $\{\mathcal{C}_j\}$ of finite cones, each congruent to some fixed finite cone $\mathcal{C}$ (i.e., each $\mathcal{C}_j$ is obtained from $\mathcal{C}$ by rigid motions), such that

(i) There exists a constant $M < \infty$ such that every U_j has $\operatorname{diam}(U_j) < M$;
(ii) There exists a $0 < \delta < \infty$ such that $\Omega_\delta \subset \bigcup_{j=1}^{\infty} U_j$, where $\Omega_\delta := \{x \in \Omega \mid \operatorname{dist}(x, \partial\Omega) < \delta\}$;
(iii) For every j,

$$Q_j := \bigcup_{x \in U_j \cap \Omega} (x + \mathcal{C}_j) \subset \Omega;$$

(iv) For some finite R, every collection of $R+1$ of the sets Q_j has empty intersection.

We have the following implications: uniform C^k regularity ($k \geq 2$) $\implies$ strong local Lipschitz $\implies$ uniform cone condition $\implies$ segment condition.

Note that all these four conditions imply that $\partial\Omega$ is $(n-1)$ dimensional and that Ω lies on one side of its boundary (since the segment condition implies so). But this is not the case of the cone condition since. In particular, the cone condition does not imply the segment condition. However, uniform cone condition $\implies$ cone condition

Example 9.40. $\Omega = \{(x, y) \in \mathbb{R}^2 \mid 0 < |x| < 1,\ 0 < y < 1\}$ (not a domain, but we can modify it). However, a *bounded* domain satisfying the cone condition can be decomposed into a finite union of subdomains each of which satisfies the strong local Lipschitz condition (and hence the segment condition).

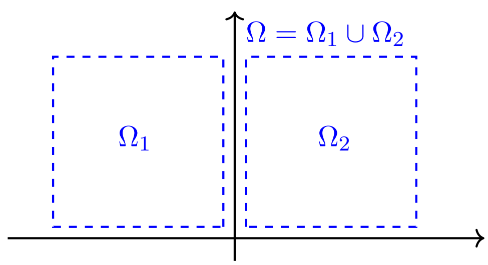


FIGURE 20. Condition Check Example 1

Cone condition? Yes. Uniform cone condition? No. Segment condition? No.

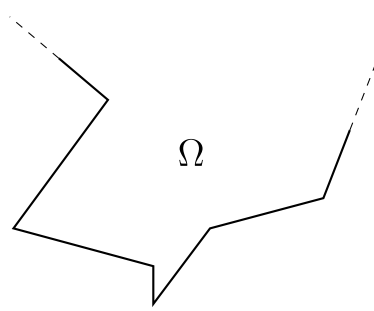


FIGURE 21. Condition Check Example 2

Cone condition? Yes. Uniform cone condition? Yes. Segment condition? Yes.

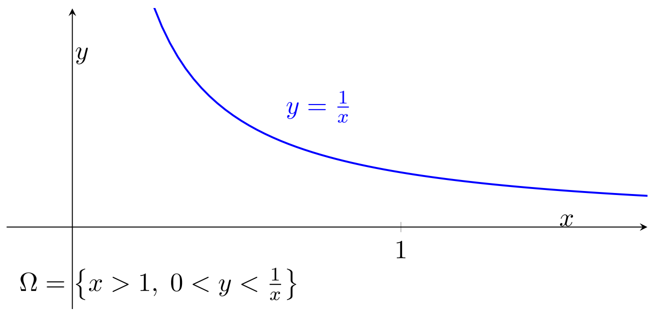


FIGURE 22. Condition Check Example 3

Cone condition? No. Uniform cone condition? No. Segment condition? Yes.

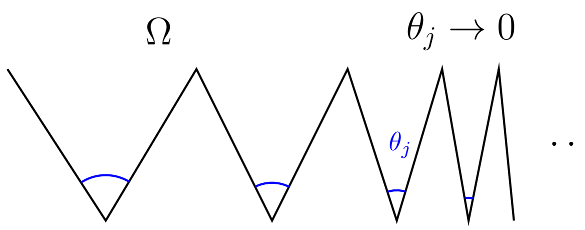


FIGURE 23. Condition Check Example 4

Cone condition? No. Uniform cone condition? No. Segment condition? Yes.

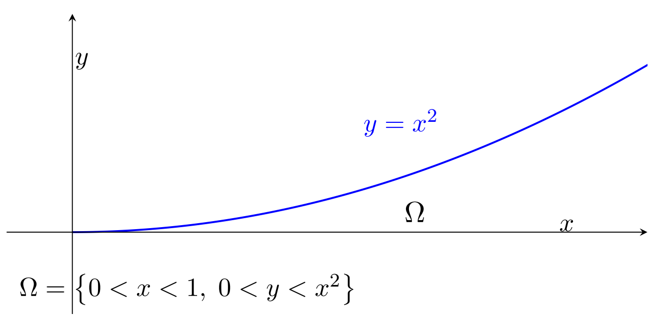


FIGURE 24. Condition Check Example 5

Cone condition? No. Uniform cone condition? No. Segment condition? Yes.

We need one more definition before we can answer the question posed above.

Definition 9.41. Let X, Y be normed spaces. We say that X **is (continuously) embedded in** Y, and write $X \hookrightarrow Y$, if X is a vector subspace of Y and the identity map on X is continuous, i.e., there exists a constant $M > 0$ such that $\|x\|_Y \le M\|x\|_X$ for all $x \in X$.

Theorem 9.42. ***(Sobolev's embedding theorem).*** *Let Ω be a domain in $\mathbb{R}^n$. Let $j \ge 0$ and $k \ge 1$ be integers. Let $1 \le p < \infty$.*

Case I. *Suppose that Ω satisfies the cone condition.*
Case I.A. If either $kp > n$ or $k = n$ and $p = 1$:

$$W^{k+j,p}(\Omega) \hookrightarrow C^j_B(\Omega).$$

Moreover,

$$W^{k+j,p}(\Omega) \hookrightarrow W^{j,q}(\Omega), \;\; p \le q < \infty$$

In particular, $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega), \;\; p \le q < \infty$.

Case I.B. If $kp = n$*:*

$$W^{k+j,p}(\Omega) \hookrightarrow W^{j,q}(\Omega), \;\; p \le q < \infty.$$

In particular, $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega), \;\; p \le q < \infty$.

Case I.C. If $kp < n$*:*

$$W^{k+j,p}(\Omega) \hookrightarrow W^{j,q}(\Omega), \;\; p \le q \le p^* := \frac{np}{n-kp}$$

(i.e., $\frac{1}{p} - \frac{1}{p^*} = \frac{k}{n}$*)*
In particular, $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega), \;\; p \le q \le p^* := \frac{np}{n-kp}$.

The constants in these embeddings depend only on n, k, p, q, j *and the dimensions of the cone* $\mathcal{C}$ *in the cone condition.*

Case II. *Suppose that* Ω *satisfies the strong local Lipschitz condition. Then, the target space in the first embedding above can be replaced with* $C^j(\bar{\Omega})$*. If, moreover,* $kp > n > (k-1)p$*, then*

$$W^{k,p}(\Omega) \hookrightarrow C^{j,p}(\bar{\Omega}), \;\; 0 < p \le k - \frac{n}{p}$$

and if $kp > n = (k-1)p$*, then*

$$W^{k,p}(\Omega) \hookrightarrow C^{j,p}(\bar{\Omega}), \;\; 0 < p < 1.$$

Also, if $n = k-1$ *and* $p = 1$*, then this last embedding holds for* $r = 1$ *as well. The constants in these embeddings depend only on* n, k, p, j *and the data on the strong local Lipschitz condition.*

Case III. *All the above embeddings hold for arbitrary* Ω *if the* W *space being embedded is replaced by* W_0*.*

Note that it suffices to prove the above embeddings for $j = 0$*. The other cases follow by applying the* $j = 0$ *case to* $D^\alpha u$*,* $|\alpha| \le j$*. We begin estimating* u *by some suitable weighted averages.*

Lemma 9.43. *Let* $\Omega \subset \mathbb{R}^n$ *be a domain satisfying the cone condition and* $k \ge 1$ *be an integer. Then, there exists a constant* K *depending only on* n, k*, and the parameters* ϱ *and* θ *of the cone condition such that*

$$\begin{aligned}|u(x)| \le K \sum_{|\alpha| \le k-1} r^{|\alpha|-n} \int_{\mathcal{C}_{x,r}} |D^\alpha u(y)| dy \\ + K \sum_{|\alpha| = \kappa} \int_{\mathcal{C}_{x,r}} |D^\alpha u(y)| |x-y|^{k-n} dy\end{aligned}$$

for all $u \in C^\infty(\Omega)$*, every* $x \in \Omega$*, and every* $0 < r \le \varrho$*, where* $\mathcal{C}_{x,r} := \big\{ y \in \mathcal{C}_x \;\big|\; |x-y| \le r \big\}$*,* $\mathcal{C}_x =$ *cone with vertex at* x *as in the cone condition.*

Proof. Set $f(t) = u\big(tx + (1-t)y\big)$ for $x \in \Omega, \; y \in \mathcal{C}_{x,r}$. Recall Taylor's formula:

$$f(1) = \sum_{j=0}^{k-1} \frac{1}{j!} f^{(j)}(0) + \frac{1}{(k-1)!} \int_0^1 (1-t)^{k-1} f^{(k)}(t) dt$$

But

$$f^{(j)}(t) = \sum_{|\alpha|=j} \frac{j!}{\alpha!} D^\alpha u\big(tx + (1-t)y\big)(x-y)^\alpha,$$

so $t = 1$ gives

$$|u(x)| \le \sum_{|\alpha|\le k-1} \frac{1}{\alpha!}|D^\alpha u(y)||x-y|^{|\alpha|} + \sum_{|\alpha|=k} \frac{k}{\alpha!}|x-y|^k \int_0^1 (1-t)^{k-1}|D^\alpha u\big(tx+(1-t)y\big)|dt.$$

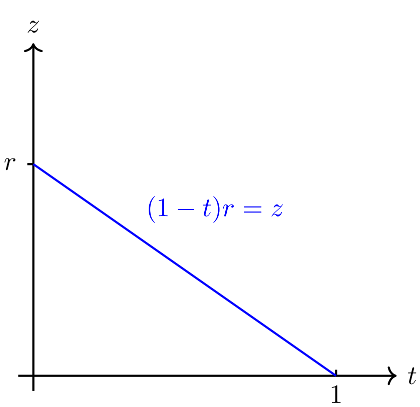
Integrate over $\mathcal{C}_{x,r}$ w.r.t y and using $|x-y|\le r$

$$ar^n|u(x)| \le \sum_{|\alpha|\le k-1} \frac{r^{|\alpha|}}{\alpha!}\int_{\mathcal{C}_{x,r}} |D^\alpha u(y)|dy + \sum_{|\alpha|=k} \frac{k}{\alpha!}\int_{\mathcal{C}_{x,r}} |x-y|^k \int_0^1 (1-t)^{k-1}|D^\alpha u\big(tx+(1-t)y\big)|dtdy.$$

Here, a is the constant in $\text{vol}(\mathcal{C}_x) = a\varrho^n$, so $\text{vol}(\mathcal{C}_{x,r}) = ar^n$. Next,

$$\begin{aligned}
&\int_{\mathcal{C}_{x,r}} |x-y|^k \int_0^1 (1-t)^{k-1}|D^\alpha u\big(tx+(1-t)y\big)|dtdy \\
&= \int_0^1 (1-t)^{k-1}\int_{\mathcal{C}_{x,r}} |D^\alpha u\big(tx+(1-t)y\big)||x-y|^k dydt \\
&= \int_0^1 (1-t)^{-n-1}\int_{\mathcal{C}_{x,(1-t)r}} |z-x|^k|D^\alpha u(z)|dzdt
\end{aligned}$$

where $z = tx+(1-t)y \implies dz = (1-t)^n dy$ and $z-x = (1-t)(y-x)$.



$$0\le t\le 1,\ \ 0\le z\le (1-t)r$$
$$0\le t\le 1-\tfrac{z}{r},\ \ 0\le z\le r$$

FIGURE 25. $z = (1-t)r$

General: $0\le t\le 1-\frac{|z-x|}{r}$
$z\in\mathcal{C}_{x,r}$

$$= \int_{\mathcal{C}_{x,r}} |z-x|^k|D^\alpha u(z)|\underbrace{\int_0^{1-\frac{|z-x|}{r}} (1-t)^{-n-1}dt}_{=\frac{r^n|z-x|^{-n}}{n}-\frac{1}{n}}dz$$

$$\le \frac{r^n}{n}\int_{\mathcal{C}_{x,r}} |z-x|^{k-n}|D^\alpha u(z)|dz.$$

Divide by r^n to get the result.

□

Proof. **of case I.A.** We will show that (recall that it suffices to show the $j = 0$ case):

$$|u(x)| \leq C\|u\|_{k,p}.$$

Take $u \in W^{k,p}(\Omega) \cap C^\infty(\Omega)$. The above inequality is a direct consequence of the previous Lemma if $k = n$, $p = 1$. For $kp > n$, if $p = 1$, then $k > n$ and again the inequality follows from the previous Lemma. Consider thus $kp > n, p > 1$. Apply Hölder's inequality with $\frac{1}{p} + \frac{1}{q} = 1$ and $r = \varrho$ in the Lemma

$$\begin{aligned}\int_{\mathcal{C}_{x,\varrho}} |D^\alpha u(y)|dy &\leq \mathrm{vol}(\mathcal{C}_{x,\varrho})^{\frac{1}{q}}\|D^\alpha u\|_{L^p(\mathcal{C}_{x,\varrho})} \\ &= a^{\frac{1}{q}}\varrho^{\frac{n}{q}}\|D^\alpha u\|_{L^p(\mathcal{C}_{x,\varrho})}\end{aligned}$$

$$\int_{\mathcal{C}_{x,\varrho}} |D^\alpha u(y)\||x-y|^{k-n}dy \leq \left(\int_{\mathcal{C}_{x,\varrho}} \left(|x-y|^{k-n}\right)^q dy\right)^{\frac{1}{q}}\|D^\alpha u\|_{L^p(\mathcal{C}_{x,\varrho})}$$

The integral is finite if $k \geq n$. If $k < n$, then, as $kp > n$, $(k-n)q = (k-n)\frac{p}{p-1} = (\underbrace{kp}_{>n} - np)\frac{1}{p-1} > -n$, so (since we can assume $\varrho < 1$)

$$|x-y|^{(k-n)q} = \frac{1}{|x-y|^{(n-k)q}} \leq \frac{1}{|x-y|^{n-\epsilon}} \iff |x-y|^{n-\epsilon} \leq |x-y|^{(n-k)q}$$

$$\iff (n-k)q \leq n-\epsilon \iff -n+\epsilon \leq -(n-k)q = (k-n)q,$$

i.e., $-n < (k-n)q$. Thus,

$$\int_{\mathcal{C}_{x,\varrho}} \left(|x-y|^{k-n}\right)^q dy \leq C\int_{B_\varrho(0)} r^{-n+\epsilon}r^{n-1}dr < \infty.$$

Thus,

$$\begin{aligned}|u(x)| &\leq \kappa\sum_{|\alpha|\leq k-1}\varrho^{|\alpha|-n}\int_{\mathcal{C}_{x,\varrho}}|D^\alpha u(y)|dy + \kappa\sum_{|\alpha|=k}\int_{\mathcal{C}_{x,\varrho}}|D^\alpha u(y)||x-y|^{k-n}dy \\ &\leq C\sum_{|\alpha|\leq k-1}\varrho^{|\alpha|-\frac{n}{p}}\|D^\alpha u\|_{L^p(\mathcal{C}_{x,\varrho})} + C\sum_{|\alpha|=k}\|D^\alpha u\|_{L^p(\mathcal{C}_{x,\varrho})} \\ &\leq C\|u\|_{k,p} \quad \text{since } \mathcal{C}_{x,\varrho} \subset \Omega.\end{aligned}$$

If $u \in W^{k,p}(\Omega)$, then there is a $u_j \to u$ in $W^{k,p}(\Omega)$, $u_j \in W^{k,p}(\Omega) \cap C^\infty(\Omega)$. By the above,

$$|u_j(x) - u_l(x)| \leq C\|u_j - u_l\|_{k,p}$$

so $u_j \to u$ in $C^0_B(\Omega)$ and thus u $C^0_B(\Omega)$. The second embedding follows from interpolation: $u \in L^p(\Omega)$, now we know that $u \in L^\infty(\Omega)$, so $u \in L^q(\Omega)$, $\frac{1}{q} = \frac{\theta}{p} + \frac{1-\theta}{\infty} = \frac{\theta}{p}, 0 < \theta < 1$, i.e., $u \in L^q(\Omega)$, $p \leq q < \infty$ and $\|u\|_{L^q(\Omega)} \leq \|u\|^\theta_{L^p(\Omega)}\|u\|^{1-\theta}_{L^\infty(\Omega)} \leq C\|u\|_{k,p}$.

□

To continue and prove cases I.B. and I.C., let us introduce χ_r to be the characteristic function of $B_r(0)$, and $G_k(x) = |x|^{k-n}$. Note that

$$\chi_r G(x) = \begin{cases} |x|^{k-n}, & |x| < r, \\ 0, & |x| \geq r, \end{cases}$$

If $k \leq n$ and $0 < r \leq 1$,

$$\chi_r(x) \leq \chi_r G_k(x) \leq G_k(x)$$

Lemma 9.44. *Let* $p \geq 1$ *and* $k \geq 1$ *be an integer. There exists a constant* $\kappa > 0$ *such that* $\chi_r G_k * |u| \in L^p(\mathbb{R}^n)$ *and*

$$\|\chi_r G_k * |u|\|_{L^p(\mathbb{R}^n)} \leq \kappa r^k \|u\|_{L^p(\mathbb{R}^n)}$$

for every $r > 0$ *and every* $u \in L^p(\mathbb{R}^n)$. *In particular*

$$\|\chi_1 * |u|\|_{L^p(\mathbb{R}^n)} \leq \|\chi_1 G_k * |u|\|_{L^p(\mathbb{R}^n)} \leq \kappa \|u\|_{L^p(\mathbb{R}^n)}$$

Proof. Write

$$\chi_r G_k * |u|(x) = \int_{\mathbb{R}^n} |u(y)|\; \chi_r(x-y)|x-y|^{k-n} dy$$

$$= \int_{B_r(x)} |u(y)||x-y|^{k-n} dy = \int_{B_r(x)} |u(y)||x-y|^{-m}|x-y|^{m+k-n} dy.$$

If $p > 1$, use Hölder, $\frac{1}{p} + \frac{1}{q} = 1$,

$$\leq \left(\int_{B_r(x)} |u(y)|^p |x-y|^{-mp} dy \right)^{\frac{1}{p}} \left(\int_{B_r(x)} |x-y|^{(m+k-n)q} dy \right)^{\frac{1}{q}}$$

Since $dy \sim t^{n-1} dt$, the second term is finite if $(m+k-n)q + n - 1 > -1 \iff (m+k)p - n > 0$, in which case it gives

$$C\big(r^{(m+k-n)q+n}\big)^{\frac{1}{q}} = Cr^{m+k-\frac{n}{p}}.$$

If $p = 1$, then $|x-y|^{m+k-n} \leq r^{m+k-n}$ if $m+k-n \geq 0$, thus

$$\int_{B_r(x)} |u(y)||x-y|^{-m}|x-y|^{m+k-n} dy \leq r^{m+k-n} \int_{B_r(x)} |u(y)||x-y|^{-m} dy.$$

Thus in either case we have

$$\chi_r G_k * |u|(x) = \int_{B_r(x)} |u(y)||x-y|^{k-n} dy \leq Cr^{m+k-\frac{n}{p}} \left(\int_{B_r(x)} |u(y)|^p |x-y|^{-mp} dy \right)^{\frac{1}{p}}$$

for m in the above range. Then,

$$\int_{\mathbb{R}^n} |\chi_r G_k * |u|(x)|^p dx \leq Cr^{(m+k)p-n} \int_{\mathbb{R}^n} \int_{B_r(x)} |u(y)|^p |x-y|^{-mp} dy dx$$

$$= Cr^{(m+k)p-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(y)|^p \chi_{B_r(x)}(y) |x-y|^{-mp} dy dx$$

$$= Cr^{(m+k)p-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(y)|^p \chi_r(x-y) |x-y|^{-mp} dy dx$$

Observe that

$$|u|^p * \big(\chi_r |x|^{-mp}\big) = \int_{\mathbb{R}^n} |u(y)|^p \chi_r(x-y)|x-y|^{-mp} dy$$

$$\||u|^p * (\chi_r |x|^{-mp})\|_{L^1(\mathbb{R}^n)} = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(y)|^p \chi_r(x-y)|x-y|^{-mp} dy dx$$

Applying Young's inequality

$$\|f * g\|_{L^{p_3}(\mathbb{R}^n)} \leq \|f\|_{L^{p_1}(\mathbb{R}^n)}\; \|g\|_{L^{p_2}(\mathbb{R}^n)},$$

$\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p_3} + 1$ with $p_1 = p_2 = p_3 = 1, f = |u|^p, g = \chi_r |x|^{-mp}$, we find

$$\||u|^p * (\chi_r |x|^{-mp})\|_{L^1(\mathbb{R}^n)} \leq \||u|^p\|_{L^1(\mathbb{R}^n)} \|\chi_r |x|^{-mp}\|_{L^1(\mathbb{R}^n)}$$

$$= \|u\|^p_{L^p(\mathbb{R}^n)} \|\chi_r |x|^{-mp}\|_{L^1(\mathbb{R}^n)}$$

$$\|\chi_r |x|^{-mp}\|_{L^1(\mathbb{R}^n)} = \int_{B_r(0)} |x|^{-mp} dx \leq Cr^{n-mp}$$

provided $-mp+n-1>-1$, i.e., $n>mp$. Hence,

$$\begin{aligned}\|\chi_r G_k * |u|\|^p_{L^p(\mathbb{R}^n)} &\le C r^{(m+k)p-n}\|u\|^p_{L^p(\mathbb{R}^n)} r^{n-mp}\\ &= C r^{kp}\|u\|^p_{L^p(\mathbb{R}^n)}\end{aligned}$$

Thus, we get the result provided we can find m satisfying all the above conditions, i.e.,

$$(m+k)p-n>0 \text{ and } n>mp,$$

i.e.,

$$\frac{n}{p}-k<m<\frac{n}{p}.$$

Since by assumption $k\ge 1$, this is possible.

□

We need the following.

Lemma 9.45. *Let $p>1$, $kp<n$, $p^*=\frac{np}{n-kp}$. There exists a constant $\kappa>0$ such that*

$$\|\chi_1 * |u|\|_{L^{p^*}(\mathbb{R}^n)} \le \|\chi_1 G_k * |u|\|_{L^{p^*}(\mathbb{R}^n)} \le \|G_k * |u|\|_{L^p(\mathbb{R}^n)} \le \kappa\|u\|_{L^p(\mathbb{R}^n)}$$

for all $u\in L^p(\mathbb{R}^n)$.

Proof. The proof is like in the previous Lemma, by a careful (albeit more complicated) analysis of convolutions. □

Proof. **of case I.C.,** $p>1$. Recall: $j=0$, $kp<n$, $p\le q\le p^*:=\frac{np}{n-kp}$, and we want to show $W^{k,p}(\Omega)\hookrightarrow L^q(\Omega)$.

Let $u\in C^\infty(\Omega)\cap W^{k,p}(\Omega)$, extend u to $\mathbb{R}^n$ by making it identically zero in $\mathbb{R}^n\setminus\Omega$. Recall the estimate

$$\begin{aligned}|u(x)| \le \kappa &\sum_{|\alpha|\le k-1} r^{|\alpha|-n}\int_{\mathcal{C}_{x,r}}|D^\alpha u(y)|dy\\ &+\kappa\sum_{|\alpha|=k}\int_{\mathcal{C}_{x,r}}|D^\alpha u(y)||x-y|^{k-n}dy\end{aligned}$$

Taking $r=\varrho$, and writing ($\varrho\le 1$)

$$\begin{aligned}\int_{\mathcal{C}_{x,\varrho}}|D^\alpha u(y)||x-y|^{m-n}dy &\le \int_{B_1(x)}|D^\alpha u(y)||x-y|^{m-n}dy\\ &=\int_{\mathbb{R}^n}|D^\alpha u(y)|\chi_{B_1(x)}(y)|x-y|^{m-n}dy\\ &=\int_{\mathbb{R}^n}|D^\alpha u(y)|\chi_1(x-y)|x-y|^{m-n}dy\\ &=(\chi_1 G_m)*|D^\alpha u|(x)\end{aligned}$$

for $m=n$ or $m=k$. Then,

$$|u(x)|\le C\sum_{|\alpha|\le k-1}\chi_1 * |D^\alpha u|(x)+\sum_{|\alpha|=k}(\chi_1 G_k)*|D^\alpha u|u(x).$$

Then,

$$\begin{aligned}\|u\|_{L^p(\Omega)} &\leq C \sum_{|\alpha|\leq k-1} \|\chi_1 * |D^\alpha u|\|_{L^p(\Omega)} + C \sum_{|\alpha|=k} \|(\chi_1 G_k) * |D^\alpha u|\|_{L^p(\Omega)} \\ &\leq C \sum_{|\alpha|\leq k-1} \|\chi_1 * |D^\alpha u|\|_{L^p(\mathbb{R}^n)} + C \sum_{|\alpha|=k} \|(\chi_1 G_k) * |D^\alpha u|\|_{L^p(\mathbb{R}^n)}\end{aligned}$$

By the next-to-previous Lemma above:

$$\begin{aligned}&\leq C \sum_{|\alpha|\leq k-1} \|D^\alpha u\|_{L^p(\mathbb{R}^n)} + C \sum_{|\alpha|=k} \|D^\alpha u\|_{L^p(\mathbb{R}^n)} \\ &= C \sum_{|\alpha|\leq k-1} \|D^\alpha u\|_{L^p(\Omega)} + C \sum_{|\alpha|=k} \|D^\alpha u\|_{L^p(\Omega)} \\ &\leq C\|u\|_{W^{k,p}(\Omega)}.\end{aligned}$$

Similarly, using the last Lemma:

$$\|u\|_{L^{p^*}(\Omega)} \leq C\|u\|_{W^{k,p}(\Omega)}$$

(Here is where we use $p > 1$, since the last Lemma requires $p > 1$).

Let $\frac{1}{q} = \frac{\theta}{p} + \frac{(1-\theta)}{p^*}$, $0 \leq \theta \leq 1$, so $p \leq q \leq p^*$. Interpolating

$$\begin{aligned}\|u\|_{L^q(\Omega)} &\leq \|u\|^\theta_{L^p(\Omega)} \, \|u\|^{1-\theta}_{L^{p^*}(\Omega)} \\ &\leq C\|u\|_{W^{k,p}(\Omega)}.\end{aligned}$$

The immediate inequality extends to an arbitrary $u \in W^{k,p}(\Omega)$ via approximation by smooth functions.

□

Proof. **of remaining cases I.** To complete the proof of case I, we need to establish case I.C. for $p = 1$ and case I.B. The proofs are somewhat lengthy, but again involve ideas like averaging, interpolation, and exploiting the cone condition. We will omit their proofs for the sake of brevity.

□

Proof. **of case II.** Recall that it suffices to prove the case $j = 0$, so we need to show

$$W^{k,p}(\Omega) \hookrightarrow C^{0,\gamma}(\bar{\Omega}),$$

where $kp > n$ and

(a) $0 < \gamma \leq k - \frac{n}{p}$ for $n > (k-1)p$,
(b) $0 < \gamma < 1$ for $n = (k-1)p$, $p > 1$,
(c) $0 < \gamma \leq 1$ for $n = (k-1)p$, $p = 1$.

In particular, we have the embedding in $C^0(\bar{\Omega})$. Since the strong local Lipschitz condition implies the segment condition, by part I, we know that

$$\sup_{x\in\Omega} |u(x)| \leq C\|u\|_{k,p}.$$

So it remains to show

$$\sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|u(x) - u(y)|}{|x-y|^\gamma} \leq C\|u\|_{k,p}$$

with γ as above. We claim that we can reduce the problem to proving this inequality for the case $k = 1$. For, if $k > 1$, then case I.C. with $j = 1$ gives

(ã) $W^{(k-1)+1,p}(\Omega) \hookrightarrow W^{1,p^*}(\Omega)$, $p^* = \frac{np}{n-(k-1)p}$, for $(k-1)p < n$.

Case I.B. gives

($\tilde{b}$) $W^{(k-1)+1,p}(\Omega) \hookrightarrow W^{1,q}(\Omega)$, $p \leq q < \infty$, for $(k-1)p = n$
Case I.A. with $p = 1$ gives
($\tilde{c}$) $W^{(k-1)+1,1} \hookrightarrow W^{1,\infty}(\Omega)$, for $(k-1) = n$.

Cases ($\tilde{a}$), ($\tilde{b}$), and ($\tilde{c}$) correspond to the relation between k, p and n in (a), (b), (c) above. Thus, we see that all cases are covered by the following statement: If $u \in W^{1,p}(\Omega), n \leq p < \infty$, then

$$\sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|u(x)-u(y)|}{|x-y|^\gamma} \leq C\|u\|_{1,p}, 0 < \gamma \leq 1 - \frac{n}{p}.$$

Assume first that $u \in C^\infty(\Omega)$ and Ω is a cube with unit edges. For $0 < t < 1$, let Q_t be a closed cube with edges having length t and faces parallel to Ω.

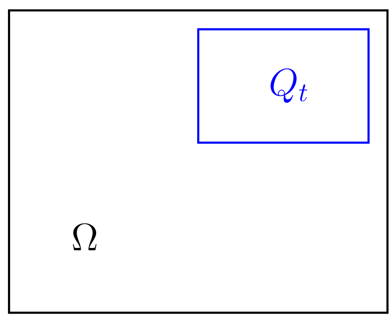


FIGURE 26. $Q_t \subset \Omega$

If $x, y \in \Omega$, $|x-y| = \sigma < 1$, then there exists Q_σ such that $x, y \in Q_\sigma$. For $z \in Q_\sigma$:

$$u(z) - u(x) = \int_0^1 \frac{d}{dt}\Big(u\big(x + t(z-x)\big)\Big)dt.$$

$$|u(x) - u(z)| \leq \int_0^1 |\nabla u\big(x + t(z-x)\big)||z-x|dt$$

Since the diagonal of a n-dim cube of edge σ has length $\sigma\sqrt{n}$:

$$|u(x) - u(z)| \leq \sigma\sqrt{n}\int_0^1 |\nabla u\big(x + t(z-x)\big)|dt$$

$$\begin{aligned} |u(x) - \frac{1}{\sigma^n}\int_{Q_\sigma} u(z)dz| &\leq \frac{1}{\sigma^n}\int_{Q_\sigma} |u(x) - u(z)|dz \\ &\leq \frac{\sqrt{n}}{\sigma^{n-1}}\int_{Q_\sigma}\int_0^1 |\nabla u\big(x + t(z-x)\big)|dtdz \\ &= \frac{\sqrt{n}}{\sigma^{n-1}}\int_0^1\int_{Q_\sigma} |\nabla u\big(x + t(z-x)\big)|dzdt. \end{aligned}$$

Set $\theta = x + t(z-x)$, so

$$= \frac{\sqrt{n}}{\sigma^{n-1}}\int_0^1\int_{Q_{t\sigma}} |\nabla u(\theta)|t^{-n}d\theta dt = \frac{\sqrt{n}}{\sigma^{n-1}}\int_0^1 \Big(\int_{Q_{t\sigma}} |\nabla u(\theta)|d\theta\Big)t^{-n}dt$$

Applying the Hölder to the $Q_{t\sigma}$ integral, $\frac{1}{p}+\frac{1}{q}=1$,

$$\leq \frac{\sqrt{n}}{\sigma^{n-1}}\int_0^1 \|\nabla u\|_{L^p(Q_{t\sigma})} \underbrace{(\mathrm{vol}(Q_{t\sigma}))^{\frac{1}{q}}}_{=(t\sigma)^{\frac{n}{q}}=(t\sigma)^{n-\frac{n}{p}}} t^{-n}dt$$

$$\leq \sqrt{n}\sigma^{1-\frac{n}{p}}\|\nabla u\|_{L^p(\Omega)} \underbrace{\int_0^1 t^{-\frac{n}{p}}dt}_{<\infty \text{ since } p>n}.$$

The same inequality holds with y instead of x, so

$$\begin{aligned}|u(x)-u(y)| &\leq \left|u(x)-\frac{1}{\sigma^n}\int_{Q_\sigma}u(z)dt\right| + \left|u(y)-\frac{1}{\sigma^n}\int_{Q_\sigma}u(z)dz\right| \\ &\leq C\sigma^{1-\frac{n}{p}}\|\nabla u\|_{L^p(\Omega)} \leq C|x-y|^{1-\frac{n}{p}}\|\nabla u\|_{L^p(\Omega)},\end{aligned}$$

thus

$$\frac{|u(x)-u(y)|}{|x-y|^r} \leq C|x-y|^{1-\frac{n}{p}-r}\|u\|_{1,p},$$

and the desired inequality holds when Ω is a cube since $0<r\leq 1-\frac{n}{p}$. Any parallelepiped can be transformed into a cube, implying that the inequality holds for parallelepipeds as well.

We now consider a general Ω satisfying the strong local Lipschitz condition. Let $\delta, M, \Omega_\delta, U_j$, and V_j be as in the definition of such domains. We recall here the definition for convenience: $\Omega\subset\mathbb{R}^n$ satisfies the strong local Lipschitz condition if there exist $\delta>0$, $M>0$, a locally finite open cover $\{U_j\}$ of $\partial\Omega$, for each j a real valued function f_j of $n-1$ variables, such that

(i) for some finite R, every collection of $R+1$ open sets U_j has empty intersection;

(ii) for every pair

$$x,y\in\Omega_\delta := \big\{z\in\Omega \mid \mathrm{dist}(z,\partial\Omega)<\delta\big\}$$

such that $|x-y|<\delta$ there exists j such that

$$x,y\in V_j := \big\{z\in U_j \mid \mathrm{dist}(z,\partial U_j)>\delta\big\}$$

(iii) each f_j satisfies a Lipschitz condition with constant M

(iv) For some Cartesian coordinate system

$$(\theta_{j,1},...,\theta_{j,n}) \text{ in } U_j,\ \Omega\cap U_j = \big\{\theta_{j,n}<f_j(\theta_{j,1},...,\theta_{j,n})\big\}.$$

There exists a parallelepiped P, whose dimensions depend only on δ and M, with the following properties:

- For each j, there exists a parallelepiped P_j congruent to P and having one vertex at the origin such that for every $x\in V_j\cap\Omega$, we have $x+P_j\subset\Omega$.
- There exist constants δ_0 and δ_1, $\delta_0\leq\delta$, such that if $x,y\in V_j\cap\Omega$ and $|x-y|<\delta_0$, then there exists a $z\in(x+P_j)\cap(y+P_j)$ satisfying $|x-z|+|y-z|\leq\delta_1|x-y|$.

Let $x,y\in\Omega$. We consider the following possibilities:

- $|x-y|<\delta_0\leq\delta$ and $x,y\in\Omega_\delta$. Then,

$$|u(x)-u(y)|\leq|u(x)-u(z)|+|u(z)-u(y)|,$$

where $z\in(x+P_j)\cap(y+P_j)$. We can apply the previous inequality in $x+P_j$ and $y+P_j$ so

$$\begin{aligned}|u(x)-u(y)| &\leq C|x-z|^r\|u\|_{1,p}+C|y-z|^r\|u\|_{1,p} \\ &\leq |x-y|^r\|u\|_{1,p}\end{aligned}$$

since $|x-z|,|y-z|\leq\delta_1|x-y|$.

- $|x-y| < \delta_0, \quad x \in \Omega_\delta,\ y \in \Omega \setminus \Omega_\delta$. Then $x \in V_j$ for some j and we can again apply the inequality in $x + P_j$, $y + P_j$.
- $|x-y| < \delta_0, \quad x, y \in \Omega \setminus \Omega_\delta$. Then we apply the inequality to $x + \tilde{P}, y + \tilde{P}$ where $\tilde{P}$ is any parallelepiped congruent to P with a vertex at the origin.
- $|x-y| \geq \delta_0$. Then,

$$|u(x) - u(y)| \leq |u(x)| + |u(y)| \leq C\|u\|_{1,p} \leq C\delta_0^{-r}|x-y|^r\|u\|_{1,p}.$$

An approximation by smooth functions produces the result for a general $u \in W^{1,p}(\Omega)$.

□

The following pictures illustrate the above ideas.

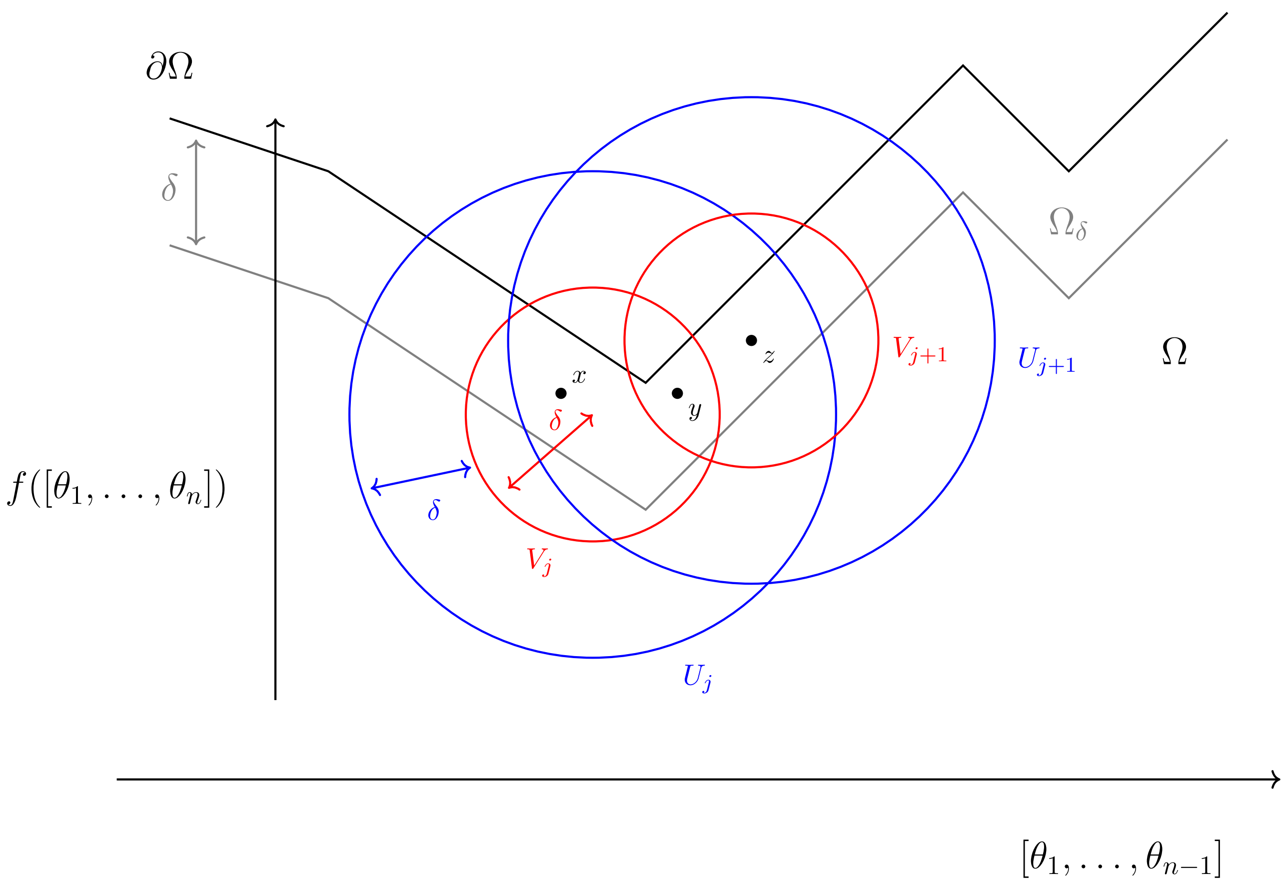


Figure 27. Strong Local Lipschitz Condition

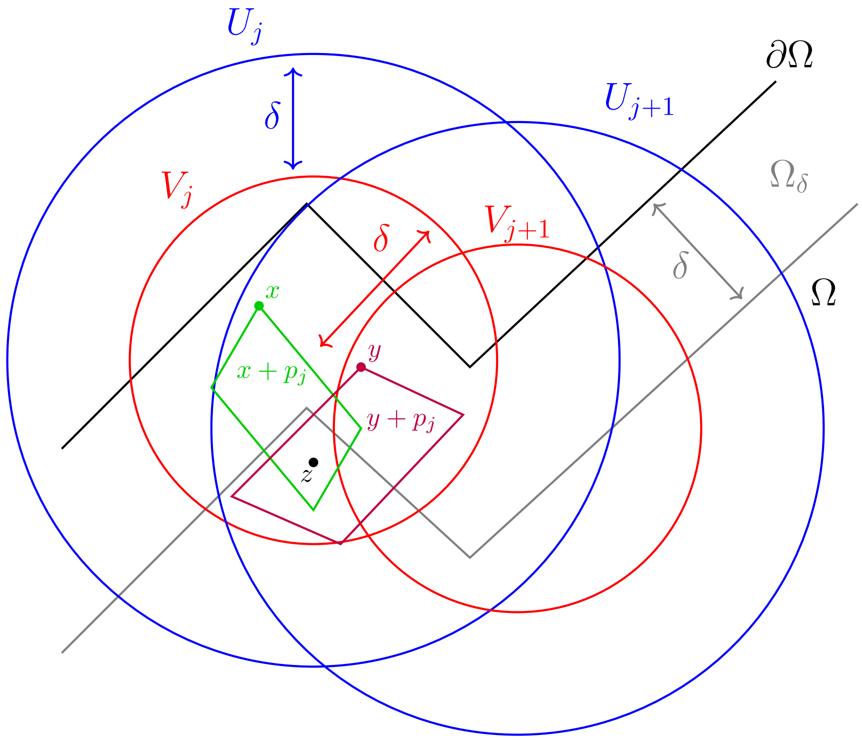


FIGURE 28. Proof of Case II:

Proof. **of case III.** The operator that maps $u : \Omega \to \mathbb{R}$ to $\tilde{u} : \mathbb{R}^n \to \mathbb{R}$ by extending u to be zero outside Ω is an isometry of $W_0^{k,p}(\Omega)$ into $W^{k,p}(\mathbb{R}^n)$. We can then apply cases I and II to $W^{k,p}(\mathbb{R}^n)$. □

9.6. **Sobolev's inequality.** We have the embedding $W^{k,p}(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n)$ for some values of q depending on the cases $kp >, =, < n$. We will now refine this embedding:

Theorem 9.46. ***(Sobolev's inequality).*** *Let $k \geq 1$ be an integer and p satisfy $kp < n$. Then there exists a constant $\kappa > 0$ such that*

$$\|u\|_{L^q(\mathbb{R}^n)} \leq \kappa \sum_{|\alpha|=k} \|D^\alpha u\|_{L^p(\mathbb{R}^n)}$$

for every $u \in C_c^\infty(\mathbb{R}^n)$ if and only if

$$q = p^* = \frac{np}{n - kp}.$$

Proof. We start proving the if-part. It suffices to prove the case $k = 1$, since higher k cases can be obtained by induction. Moreover, if suffices to prove the following inequality for $\varrho = 1$:

$$\int_{\mathbb{R}^n} |u|^{\frac{n}{n-1}} dx \leq \kappa \left(\sum_{j=1}^n \int_{\mathbb{R}^n} |\partial_j u| dx \right)^{\frac{n}{n-1}}.$$

For, if $1 < p < n$, $p^* = \frac{np}{n-p}$, then we apply the above to $|u|^r$, $r = (n-1)\frac{p^*}{n}$, to get

$$\begin{aligned}
\int_{\mathbb{R}^n} (|u|^r)^{\frac{n}{n-1}} dx = \int_{\mathbb{R}^n} |u|^{p^*} dx &\leq C\bigg(\sum_{j=1}^n \int_{\mathbb{R}^n} \big|\partial_j |u|^r \big| dx \bigg)^{\frac{n}{n-1}} \\
&\leq C\bigg(\sum_{j=1}^n \int_{\mathbb{R}^n} |u|^{r-1} |\partial_j u| dx \bigg)^{\frac{n}{n-1}}, \text{ apply Hölder} \\
&\leq C\bigg\{ \sum_{j=1}^n \bigg[\int_{\mathbb{R}^n} \big(|u|^{r-1}\big)^{\frac{p}{p-1}} dx \bigg]^{\frac{p-1}{p}} \bigg[\int_{\mathbb{R}^n} |\partial_j u|^p dx \bigg]^{\frac{1}{p}} \bigg\}^{\frac{n}{n-1}} \\
&= C\bigg(\int_{\mathbb{R}^n} |u|^{p^*} dx \bigg)^{\frac{n(p-1)}{p(n-1)}} \bigg(\sum_{j=1}^n \bigg[\int_{\mathbb{R}^n} |\partial_j u|^p dx \bigg]^{\frac{1}{p}} \bigg)^{\frac{n}{n-1}}
\end{aligned}$$

after using $(r-1)\frac{p}{p-1} = ((n-1)\frac{p^*}{n} - 1)\frac{p}{p-1} = (\frac{n-1}{n}\frac{np}{n-p} - 1)\frac{p}{p-1} = \frac{((n-1)p-n+p)}{n-p}\frac{p}{p-1} = \frac{np}{n-p}p^*$. Also, $r\frac{n}{n-1} = (n-1)\frac{p^*}{n} \cdot \frac{n}{n-1} = p^*$,

$$\|u\|_{L^{p^*}(\mathbb{R}^n)}^{p^*} \leq C\big(\|u\|_{L^{p^*}}\big)^{\frac{p^* n(p-1)}{p(n-1)}} \bigg(\sum_{j=1}^n \bigg[\int_{\mathbb{R}^n} |\partial_j u|^p dx \bigg]^{\frac{1}{p}} \bigg)^{\frac{n}{n-1}},$$

which gives the result since

$$1 - \frac{n(p-1)}{p(n-1)} = \frac{(n-1)p - n(p-1)}{p(n-1)} = \frac{n-p}{p(n-1)},$$

$$p^* \frac{(n-1)}{n} \frac{(n-p)}{p(n-1)} = \frac{np}{n-p}\frac{n-p}{np} = 1.$$

So let's prove the inequality. Since

$$u(x) = \int_{-\infty}^{x^i} \partial_i u(x^1, ..., x^{i-1}, t^i, x^{i+1}, ..., x^n) dt^i,$$

we have

$$|u(x)| \leq \int_{-\infty}^{\infty} \big|\nabla u(x^1, ..., x^{i-1}, t^i, x^{i+1}, ..., x^n)\big| dt^i.$$

Thus

$$|u(x)|^{\frac{n}{n-1}} \leq \prod_{i=1}^n \bigg(\int_{-\infty}^{\infty} \big|\nabla u(x^1, ..., x^{i-1}, t^i, x^{i+1}, ..., x^n)\big| dt^i \bigg)^{\frac{1}{n-1}}.$$

Integrating in x^1 and omitting the argument $(x^1, ..., x^{i-1}, t^i, x^{i+1}, ..., x^n)$

$$\begin{aligned}
\int_{-\infty}^{+\infty} |u(x)|^{\frac{n}{n-1}} dx^1 &\leq \int_{-\infty}^{+\infty} \prod_{i=1}^n \bigg(\int_{-\infty}^{+\infty} |\nabla u| dt^i \bigg)^{\frac{1}{n-1}} dx^1 \\
&= \bigg(\int_{-\infty}^{\infty} |\nabla u| dt^1 \bigg)^{\frac{1}{n-1}} \int_{-\infty}^{\infty} \prod_{i=2}^n \bigg(\int_{-\infty}^{\infty} |\nabla u| dt^i \bigg)^{\frac{1}{n-1}} dx^1
\end{aligned}$$

Apply Hólder's inequality

$$\int |u_1 ... u_l| dx \leq \prod_{i=1}^l \|u_i\|_{L^{p_i}}, \quad \frac{1}{p_1} + ... + \frac{1}{p_l} = 1$$

in the x^1 variable to the $n-1$ functions

$$\left(\int_{-\infty}^{\infty} |\nabla u| dt^1\right) \quad i = 2, ..., n \text{ with } p_i = n-1,$$

$$\leq \left(\int_{-\infty}^{\infty} |\nabla u| dt^1\right)^{\frac{1}{n-1}} \prod_{i=2}^{n} \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla u| dt^i dx^1\right)^{\frac{1}{n-1}},$$

Integrate w.r.t. x^2:

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u|^{\frac{n}{n-1}} dx^1 dx^2 \leq & \int_{-\infty}^{\infty} \Bigg\{ \left(\int_{-\infty}^{\infty} |\nabla u| dt^1\right)^{\frac{1}{n-1}} \\ & \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla u| dt^2 dx^1\right)^{\frac{1}{n-1}} \prod_{i=3}^{n} \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla u| dt^i dx^1\right)^{\frac{1}{n-1}} \Bigg\} dx^2 \\ = & \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla u| dt^2 dx^1\right)^{\frac{1}{n-1}} \int_{-\infty}^{\infty} \Bigg\{ \left(\int_{-\infty}^{\infty} |\nabla u| dt^1\right)^{\frac{1}{n-1}} \\ & \prod_{i=3}^{n} \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla u| dt^i dx^1\right)^{\frac{1}{n-1}} \Bigg\} dx^2.\end{aligned}$$

Apply again Hölder's inequality to the $n-1$ functions on the right in the x^2 variable:

$$\begin{aligned}\leq & \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla u| dt^2 dx^1\right)^{\frac{1}{n-1}} \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla u| dt^1 dx^2\right)^{\frac{1}{n-1}} \\ & \prod_{i=3}^{n} \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla u| dt^i dx^1 dx^2\right)^{\frac{1}{n-1}}.\end{aligned}$$

If $u = 2$, we are done. Otherwise, we continue to get

$$\int_{\mathbb{R}^n} |u|^{\frac{n}{n-1}} dx \leq \left(\int_{\mathbb{R}^n} |\nabla u| dx\right)^{\frac{n}{n-1}}.$$

Now we prove the only if part. If

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C \sum_{|\alpha|=k} \|D^\alpha u\|_{L^p(\mathbb{R}^n)}$$

holds for all $u \in C_c^\infty(\mathbb{R}^n)$, then it holds for $u_t(x) = u(tx)$, $t > 0$. Changing variables we find

$$\|u_t\|_{L^q(\mathbb{R}^n)} \leq C t^{k - \frac{n}{p} + \frac{n}{q}} \sum_{|\alpha|=k} \|D^\alpha u_t\|_{L^p(\mathbb{R}^n)}.$$

Since we must have $k - \frac{n}{p} + \frac{n}{q} = 0$, we find $q = p^*$.

□

Remark 9.47. If Ω is bounded, we can get

$$\|u\|_{L^p(\Omega)} \leq C \|\nabla u\|_{L^p(\Omega)},$$

which is known as **Poincaré's inequality**, $u \in C_c^\infty(\Omega)$. Thus, $\|\nabla u\|_{L^p(\Omega)}$ is a norm on $W_0^{1,p}(\Omega)$ equivalent to $\|\cdot\|_{1,p}$. There are other related inequalities which are also known as Poincaré's inequality.

9.7. Compact embeddings.

Definition 9.48. Let X and Y be Banach spaces. We say that X **is compactly embedded into** Y if X is continuously embedded into Y and X and each bounded sequence in X is pre-compact in Y (so bounded sequences in X have convergent subsequences in Y).

We recall the following theorem from analysis:

Theorem 9.49. ***Pre-compactness in*** $L^p(\Omega)$. *A bounded subset $\mathcal{B} \subset L^p(\Omega), 1 \leq p < \infty$, is pre-compact if and only if for every $\epsilon > 0$ there exists a $\delta > 0$ and a subset $\kappa \subset\subset \Omega$ such that*

$$\int_\Omega |\tilde{u}(x+h) - \tilde{u}(x)|^p dx < \epsilon^p$$

and

$$\int_{\Omega \setminus \bar{\kappa}} |u(x)|^p dx < \epsilon^p$$

for every $u \in \mathcal{B}$ and every $h \in \mathbb{R}^n$ with $|h| < \delta$, where

$$\tilde{u}(x) = \begin{cases} u(x), & x \in \Omega, \\ 0, & x \notin \Omega. \end{cases}$$

Theorem 9.50. ***Reillich-Kondrachov theorem.*** *Let Ω be a domain in $\mathbb{R}^n$ and $\Omega_0 \subset \Omega$ a bounded subdomain. Let $k \geq 1$ and $j \geq 0$ be integers, and $1 \leq p < \infty$. The embeddings below are compact under the stated hypotheses:*

Case I. *Ω satisfies the cone condition, $kp \leq n$,*

$$W^{k+j,p}(\Omega) \hookrightarrow W^{j,q}(\Omega_0), \;\; kp < n, \;\; 1 \leq q < p^* = \frac{np}{n-kp}$$

or $kp = n, \;\; 1 \leq q < \infty$.

Case II. *Ω satisfies the cone condition, $kp > n$,*

$$W^{k+j,p}(\Omega) \hookrightarrow C^j_B(\Omega_0)$$
$$W^{k+j,p}(\Omega) \hookrightarrow W^{j,q}(\Omega_0), \;\; 1 \leq q < \infty.$$

Case III. *Ω satisfies the strong local Lipschitz condition,*

$$W^{k+j,p}(\Omega) \hookrightarrow C^j(\bar{\Omega}_0), kp > n$$
$$W^{k+j,p}(\Omega) \hookrightarrow C^{j,r}(\bar{\Omega}_0), kp > n \geq (k-1)p \text{ and } 0 < r < k - \frac{n}{p}.$$

Case IV. *Ω is an arbitrary domain. Then all of the above embeddings hold with $W^{k+j,p}(\Omega)$ replaced by $W^{k+j,p}_0(\Omega)$. In particular, if Ω is bounded, we can take $\Omega_0 = \Omega$ above.*

Proof. As for the previous embeddings, it suffices to prove the case $j = 0$. We can also assume Ω_0 to satisfy the cone condition. For, if $\mathcal{C}$ is a cone for the cone condition of Ω, let Ω' be the union of all cones congruent to $\mathcal{C}$ that are contained in Ω and have non-empty intersection with Ω_0. Then we have $\Omega_0 \subset \Omega' \subset \Omega$. Ω' is bounded and satisfies the cone condition. If $W^{k,p}(\Omega) \hookrightarrow X(\tilde{\Omega})$ is compact, so is $W^{k,p}(\Omega) \hookrightarrow X(\Omega_0)$ by taking the restriction. We will also use that the composition of a continuous embedding with a compact one is compact.

Proof of case I. Consider the case when

$$kp < n \text{ and } 1 \leq q < p^* = \frac{np}{n-kp}.$$

We will reduce the proof to $q = 1$ by the following claim:

Claim 9.51. *If $W^{k,p}(\Omega) \hookrightarrow L^{q^*}(\Omega_0)$ and $W^{k,p}(\Omega) \hookrightarrow L^{q_1}(\Omega_0)$ compactly, then $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega_0)$ compactly for $q_1 \le q < q^*$, where $q^* < \infty$.*

To prove the claim, interpolate

$$\|u\|_{L^q(\Omega_0)} \le \|u\|^{\theta}_{L^{q_1}(\Omega)} \|u\|^{1-\theta}_{L^{q^*}(\Omega_0)} \le C\|u\|^{\theta}_{L^{q_1}(\Omega_0)} \|u\|^{1-\theta}_{W^{k,p}(\Omega)}$$

If $\{u_j\}$ is bounded in $W^{k,p}(\Omega)$, then it has a subsequence converging in $L^{q_1}(\Omega_0)$ by assumption, so it is Cauchy in $L^q(\Omega_0)$.

In our case, we have the embedding

$$W^{k,p}(\Omega) \hookrightarrow L^{p^*}(\Omega)$$

by Sobolev embedding (case I.C.), so it suffices to get compactness $W^{k,p}(\Omega) \hookrightarrow L^1(\Omega_0)$.

Let $\mathcal{B}$ be a set of bounded functions in $W^{k,p}(\Omega)$. We will use the L^p-precompactness theorem above, so let

$$\tilde{u}(x) = \begin{cases} u(x), & x \in \Omega, \\ 0, & x \notin \Omega. \end{cases}$$

and $\epsilon > 0$ be given.

Let

$$\Omega_j = \big\{x \in \Omega \mid \operatorname{dist}(x, \partial\Omega) > \frac{2}{j}\big\},$$

$j = 1, 2, ...$ We have:

$$\begin{aligned} \int_{\Omega_0 \setminus \Omega_j} |u(x)| dx &\le \Big(\int_{\Omega_0 \setminus \Omega_j} |u(x)|^{p^*} dx \Big)^{\frac{1}{p^*}} \operatorname{vol}(\Omega_0 \setminus \Omega_j)^{1-\frac{1}{p^*}} \\ &\le C\|u\|_{W^{k,p}(\Omega)} \operatorname{vol}(\Omega_0 \setminus \Omega_j)^{1-\frac{1}{p^*}} \end{aligned}$$

Since $p^* > 1$, we have that for every $u \in \mathcal{B}$

$$\int_{\Omega_0 \setminus \Omega_j} |u(x)| dx < \epsilon$$

if j is large enough. Similarly, if j is large enough

$$\int_{\Omega_0 \setminus \Omega_j} \big|\tilde{u}(x+h) - \tilde{u}(x)\big| dx < \frac{\epsilon}{2}.$$

Now, let $|u| < \frac{1}{j}$. Then $x \in \Omega_j \implies x + th \in \Omega_{2j}, \;\; 0 \le t \le 1.$

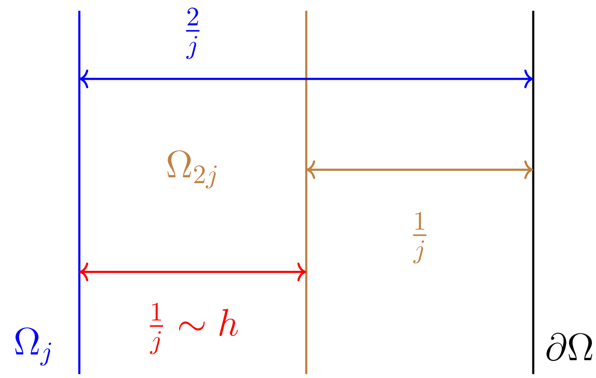


FIGURE 29. Ω_j, Ω_{2j}, ∂_Ω

Then,

$$\begin{aligned}\int_{\Omega_j} \big|u(x+h)-u(x)\big|dx &\leq \int_{\Omega_j}\int_0^1 \left|\frac{d}{dt}u(x+th)\right|dtdx \\ &\leq |h|\int_{\Omega_{2j}} |\nabla u(x)|dx \\ &\leq C|h|\|u\|_{W^{k,p}(\Omega)}.\end{aligned}$$

(prove the above inequality for C^∞ first).

So $\int_{\Omega_j}|u(x+h)-u(x)|dx < \frac{\epsilon}{2}$ if $|h|$ is sufficiently small, and we get $\int_{\Omega_0}|\tilde{u}(x+h)-\tilde{u}(x)|dx < \epsilon$. Thus $\mathcal{B}$ is pre-compact in $L^1(\Omega_0)$.

The case $kp = n$ is proven with similar ideas. For example, if then we have, with $1 \leq r < p$,

$$W^{k,p}(\Omega) \hookrightarrow W^{k,p}(\Omega_0) \hookrightarrow W^{k,r}(\Omega_0) \hookrightarrow L^q(\Omega_0),$$

and the latter is compact as showed above (recall that Ω_0 can be assumed to satisfy the cone condition).

Proof of case III. Consider Ω satisfying the strong local Lipschitz condition and

$$kp > n \geq (k-1)p \text{ and } 0 < \gamma < k - \frac{n}{p}.$$

Let β be such that $\gamma < \beta < k - \frac{n}{p}$. We have

$$W^{k,p}(\Omega) \underset{\text{Sobolev, case II}}{\hookrightarrow} C^{0,\beta}(\bar{\Omega}) \underset{\text{by restriction}}{\hookrightarrow} C^{0,\beta}(\bar{\Omega}_0) \underset{\text{compact by } \bar{\Omega}_0 \text{ bounded and Arzelá-Ascoli}}{\hookrightarrow} C^{0,\gamma}(\bar{\Omega}_0)$$

If $kp > n$, let $l \geq 0$ be an integer such that

$$(k-l)p > n \geq (k-l-1)p.$$

Then

$$W^{k,p}(\Omega) \hookrightarrow W^{k-l,p}(\Omega) \hookrightarrow C^{0,\alpha}(\bar{\Omega}_0) \hookrightarrow C^0(\bar{\Omega}_0)$$

where the last two embeddings are similar to above with the last one compact.

Proof of case II. Since Ω_0 can be assumed to satisfy the cone condition and is bounded, we can write

$$\Omega_0 = \bigcup_{j=1}^{M} \Omega_j$$

where each Ω_j satisfies the strong local Lipschitz condition (this is a property of the cone condition that we will not prove). Then

$$W^{k,p}(\Omega) \hookrightarrow W^{k,p}(\Omega_j) \hookrightarrow C^0(\bar{\Omega}_j)$$

where the last embedding is compact as above. If $\{u_i\}$ is bounded in $W^{k,p}(\Omega)$, we can then select a subsequence whose restriction to Ω_j converges in $C^0(\bar{C}_l)$ for each j. But then it converges in $C^0_B(\Omega_0)$. The other embedding follows from Hölder's inequality since Ω_0 is bounded, so $C^0(\Omega_0) \hookrightarrow L^q(\Omega_0)$ for any q.

Proof of case IV. This follows from the embedding $W_0^{k+j,p}(\Omega) \hookrightarrow W^{k+j,p}(\mathbb{R}^n)$ obtained from extending functions to be zero outside Ω.

□

9.8. **Traces.** In order to treat boundary value problems, we need to be able to talk about the restriction of Sobolev to $\partial\Omega$.

Theorem 9.52. ***(trace theorem)***. *Assume that Ω is bounded and $\partial\Omega$ is C^1. There exists a bounded linear operator $T : W^{1,p}(\Omega) \to L^p(\partial\Omega), 1 \leq p < \infty$, such that $Tu = u|_{\partial\Omega}$ if $u \in W^{1,p}(\Omega) \cap C^0(\bar{\Omega})$. (Tu is called the **trace** of u on $\partial\Omega$).*

Proof. Suppose first that $u \in C^1(\bar{\Omega})$ and that $\partial\Omega$ is flat near a point $z \in \partial\Omega$ and such that

$$B_r(z) \cap \{x^n \geq 0\} \subset \bar{\Omega}, \;\; B_r(z) \cap \{x^n < 0\} \subset \mathbb{R}^n \setminus \bar{\Omega}$$

for some $r > 0$. Let $\psi \in C_c^\infty(B_r(z))$ be such that $\psi \geq 0$ and $\psi = 1$ on $B_{\frac{r}{2}}(z)$. Let $\Gamma := \partial\Omega \cap B_{\frac{r}{2}}(z)$ and $y = (x^1, ..., x^{n-1})$. Then,

$$\begin{aligned}
\int_\Gamma |u|^p dy &\leq \int_{\{x^n=0\}} \psi |u|^p dx \\
&= -\int_{B_r(z)\cap\{x^n\geq 0\}} \partial_u(\psi|u|^p)dx \\
&= -\int_{B_r(z)\cap\{x^n\geq 0\}} |u|^p \partial_u \psi dx - \int_{B_r(z)\cap\{x^n\geq 0\}} \psi p |u|^{p-1} (\text{sign } u)\partial_u u dx. \\
&\leq C\int_{B_r(z)\cap\{x^n\geq 0\}} |u|^p dx + C\int_{B_r(z)\cap\{x^n\geq 0\}} |u|^{p-1}|\nabla u| dx \\
&\leq C\|u\|_{1,p}^p
\end{aligned}$$

since $|u|^{p-1}|\nabla u| \leq C(|u|^{(p-1)q} + |\nabla u|^p) = C(|u|^p + |\nabla u|^p)$.

Since $\partial\Omega$ is compact and each $z \in \partial\Omega$ has a neighborhood where $\partial\Omega$ can be flattened, we get the inequality for $C^1(\bar{\Omega})$ functions. Using an approximation we get the inequality for $u \in W^{1,p}(\Omega)$.

If $u \in W^{1,p}(\Omega) \cap C^0(\bar{\Omega})$, observe first that Tu does not depend on which sequence we use to approximate u. If we take the sequence we constructed in our poof of the density of $C^\infty(\bar{\Omega})$, that sequence converges uniformly to a limit $u \in C^0(\bar{\Omega})$, so $Tu = u|_{\partial\Omega}$.

□

In the case of $W_0^{1,p}(\Omega)$, we have:

Theorem 9.53. *Let Ω be bounded and $\partial\Omega$ be C^1. Let $u \in W^{1,p}(\Omega)$. Then $u \in W_0^{1,p}(\Omega)$ if and only if $Tu = 0$.*

Proof. We will omit the proof, but this should not be surprising since $u \in W_0^{1,p}(\Omega)$ is the limit of functions compactly supported in Ω.

□

It is possible to define Sobolev spaces $W^{s,p}(\Omega)$ with s not an integer. In fact, we will do this for $p=2$. In this case the trace theorem can be strengthened to

$$T : W^{1,p}(\Omega) \to W^{1-\frac{1}{p},p}(\partial\Omega).$$

If $\partial\Omega$ is sufficiently regular, we get similar results for $W^{s,p}(\Omega),\quad s>1$. In particular, for $p=2$:

$$T : H^s(\Omega) \to H^{s-\frac{1}{2}}(\partial\Omega).$$

9.9. **Sobolev spaces of fractional and negative order.** It is possible to generalize the definition of Sobolev spaces for $W^{s,p}(\Omega)$ with $s\in\mathbb{R}$. Here, we will do this in the case $p=2$ and $\Omega=\mathbb{R}^n$. First, we recall some basic facts about the Fourier transform.

Facts about the Fourier transform.

- A function $u \in C^\infty(\mathbb{R}^n,\mathbb{C})$ is called a **Schwartz function** if for every pair of multi-indices α and β there exists a constant $\kappa_{\alpha,\beta}$ such that for all $x\in\mathbb{R}^n$

$$|x^\alpha D^\beta u(x)| \le \kappa_{\alpha,\beta}$$

  The space of Schwartz functions is denoted by $\mathcal{S}=\mathcal{S}(\mathbb{R}^n)$.
- If $u\in\mathcal{S}$, the **Fourier transform** of u is defined by

$$\hat{u}(\xi) \equiv \mathcal{F}(u)(\xi) := \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(x)dx.$$

  The **inverse Fourier transform** of $\hat{u}\in\mathcal{S}$ is

$$u(x) \equiv \mathcal{F}^{-1}(\hat{u})(\xi) := \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}\hat{u}(\xi)d\xi.$$

  $\mathcal{F}$ and $\mathcal{F}^{-1}$ are continuous maps (with respect to the Schwartz topology) $\mathcal{S}\to\mathcal{S}$ that are in fact inverses of each other.
  (The topology on $\mathcal{S}$ is given by the metric

$$d(u,v)=\sum_{j=1}^\infty 2^{-j}\frac{p_j(u-v)}{1+p_j(u-v)}$$

  where $\{p_j\}$ is the (countable) set of all semi-norms

$$p_{\alpha,\beta}(u)=\sup_{x\in\mathbb{R}^n}|x^\alpha D^\beta u(x)|.\ )$$

- Parseval's formula

$$\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\hat{u}(\xi)\bar{\hat{v}}(\xi)d\xi = \int_{\mathbb{R}^n} u(x)\bar{v}(x)dx.$$

  Observe that

$$\widehat{D^\alpha u} = i^{|\alpha|}\xi^\alpha\hat{u}$$

  This will motivate the definition of $H^s(\mathbb{R}^n)$ for $s\in\mathbb{R}$.

Definition 9.54. A continuous linear form on $\mathcal{S}$ (a.k.a. continuous functional, i.e. a linear map $f:\mathcal{S}\to\mathbb{C}$ that is continuous, $f(u_j)\to f(u)$ if $u_j\to u$ in $\mathcal{S}$) is called a **tempered distribution**. The space of tempered distributions is denoted $\mathcal{S}'$.

Definition 9.55. The **Fourier transform of** $f\in\mathcal{S}'$ is defined by

$$\hat{f}(u)=f(\hat{u}),\quad u\in\mathcal{S}.$$

Definition 9.56. Let $s \in \mathbb{R}$. We define the **Sobolev space** $H^s = H^s(\mathbb{R}^n)$ as the space of $u \in \mathcal{S}'$ such that $\hat{u}$ is a measurable function with the property that $\hat{u}(\xi)(1+|\xi|^2)^{\frac{s}{2}}$ is square-integrable. We sometimes write $H^{(s)}$ if we want to stress that s can be any real number. A norm in H^s is given by

$$\|u\|_s \equiv \|u\|_{(s)} := \left(\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\hat{u}(\xi)|^2 \big(1+|\xi|^2\big)^s d\xi \right)^{\frac{1}{2}}$$

and an inner product by

$$(u,v)_s \equiv (u,v)_{(s)} := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{u}(\xi)\bar{\hat{v}}(\xi)\big(1+|\xi|^2\big)^s d\xi.$$

To see the definition, notice that if $k \geq 0$ is an integer, then $\widehat{D^k u} = i^k |\xi|^k \hat{u}$, so

$$\|Du\|_{L^2} = \|\, |\xi|^k \hat{u} \,\|_{L^2} = \left(\int_{\mathbb{R}^n} (|\xi|^2)^k \, |\hat{u}(\xi)|^2 d\xi \right)^{\frac{1}{2}}.$$

More generally, we have

$$\frac{1}{C}(1+|\xi|^2)^k \leq \sum_{|\alpha|\leq k} \xi^{2\alpha} \leq C(1+|\xi|^2)^k.$$

So the norms $\|u\|_k$ and $\|u\|_{(k)}$ are equivalent. One can also show that if $u \in H^{(s)}$ then u is $|\alpha|$-times weakly differentiable, $|\alpha| \leq s$. These observations show that $H^{(s)}$ agrees with the previous definition of H^s when s is a non-negative integer.

All the basic properties H^k, $k \geq 0$ integer, remain valid for $H^{(s)}$, including

- Density of $C_c^\infty(\mathbb{R}^n)$. We also have density of $\mathcal{S}$.
- $D^\alpha : H^{(s)} \to H^{(s-|\alpha|)}$ is bounded.
- $(u, D^\alpha v)_0 = (-1)^{|\alpha|}(D^\alpha u, v), \;\; u, v \in H^{(s)}, \;\; |\alpha| \leq s.$
- The Sobolev embedding theorems and compact embeddings hold for $H^{(s)}$, where s replace k in the statements and the inequalities among s, n, and p are interpreted accordingly (including $p^* = \frac{np}{n-sp}$).

Definition 9.57. Let $u \in H^{(s)}$. We define $(1-\Delta)^t u$ as the tempered distribution whose Fourier transform is $(1+|\xi|^2)^t \hat{u}(\xi)$, i.e.,

$$(\widehat{(1-\Delta)^t u})(\xi) = \big(1+|\xi|^2\big)^t(\xi).$$

We obtain that

$$\|(1-\Delta)^{\frac{t}{2}} u\|_{(s-t)} = \|u\|_{(s)},$$

so $\big(1-\Delta\big)^{\frac{t}{2}}$ is a bounded linear map from $H^{(s)}$ to $H^{(s-t)}$. It has a bounded inverse given by $(1-\Delta)^{-\frac{t}{2}}$. Thus, we see that

$$H^s(\mathbb{R}^n) = (1-\Delta)^{-\frac{s}{2}}(L^2(\mathbb{R}^n)).$$

We also have

$$\big((1-\Delta)^{\frac{t}{2}} u, v\big)_0 = \big(u, (1-\Delta)^{\frac{t}{2}} v\big)_0$$

for $u, v \in H^{(s)}, \;\; s \geq 0, \; t \leq s$.

9.10. **Duality.** We will now investigate the dual space of $W^{k,p}(\Omega)$.

Notation 9.58. Given $1 \leq p < \infty$, we set

$$p' = \begin{cases} \infty, & p = 1, \\ \frac{p}{p-1}, & 1 < p < \infty, \\ 1, & p = \infty. \end{cases}$$

We recall the following construction used in the proof that $W^{k,p}(\Omega)$ is reflexive for $1 \leq p < \infty$:

Let $\mu(k,n)$ be the number of multi-indices α such that $|\alpha| \leq k$, and for each α let Ω_α be a copy of Ω, so the $\mu(k,n)$ domains Ω_α are disjoint. Set

$$\Omega_{(k)} := \bigcup_{|\alpha| \leq k} \Omega_\alpha.$$

Given $v : \Omega_{(k)} \to \mathbb{R}$, we write v_α for $v|_{\Omega_\alpha}$, so we can identify v with a vector (v_α), $v_\alpha : \Omega_\alpha \to \mathbb{R}$. Given $u \in W^{k,p}(\Omega)$, let v be the function $\Omega_{(k)}$ that coincides with $D^\alpha u$ in Ω_α. The map $\Gamma : W^{k,p}(\Omega) \to L^p(\Omega_{(k)})$ $u \mapsto v$ is an isometry. Because $W^{k,p}(\Omega)$ is complete, the image X of Γ is a closed subspace of $L^p(\Omega_{(k)})$ and we have $W^{k,p}(\Omega) = \Gamma^{-1}(X)$. We will use these constructions below.

The dual space of $W^{k,p}(\Omega)$ is defined in the usual way as dual of a Banach space:

Definition 9.59. The **dual space of** $W^{k,p}(\Omega)$, denoted $(W^{k,p}(\Omega))'$, is defined as the space of continuous linear forms on $W^{k,p}(\Omega)$.

Our goal is to characterize the $(W^{k,p}(\Omega))'$. We will use certain dualities realized by the L^2 inner product so it will be convenient to denote

$$\langle u, v \rangle = \int_\Omega u(x)v(x)dx$$

provided the RHS makes sense.

Lemma 9.60. *To every $f \in (L^p(\Omega_{(k)}))'$, $1 \leq p < \infty$, there corresponds a unique $v \in L^{p'}(\Omega_{(k)})$, such that*

$$f(u) = \int_{\Omega_{(k)}} v(x)u(x)dx = \sum_{|\alpha| \leq k} \int_{\Omega_\alpha} v_\alpha(x)u_\alpha(x)dx = \sum_{|\alpha| \leq k} \langle v_\alpha, u_\alpha \rangle$$

for all $u \in L^p(\Omega_{(k)})$. So $(L^p(\Omega_{(k)}))' = L^{p'}(\Omega_{(k)})$.

Proof. This is simply the Riesz representation theorem applied to $L^p(\Omega_{(k)})$ under our notational conventions.

□

Theorem 9.61. ***(Riesz representation for Sobolev spaces).*** *Let $1 \leq p < \infty$, $k \geq 1$ an integer. For every $f \in (W^{k,p}(\Omega))'$, there exists a $v \in L^{p'}(\Omega_{(k)})$ such that*

$$f(u) = \sum_{|\alpha| \leq k} \langle v, D^\alpha u \rangle \tag{9.1}$$

for all $u \in W^{k,p}(\Omega)$. Furthermore

$$\|f\|_{(W^{k,p}(\Omega))'} = \inf_{\mathcal{U}} \|v\|_{L^{p'}(\Omega_{(k)})} = \min_{\mathcal{U}} \|v\|_{L^{p'}(\Omega_{(k)})} \tag{9.2}$$

where $\mathcal{U}$ is the set of all $v \in L^{p'}(\Omega_{(k)})$ for which (9.1) holds for all $u \in W^{k,p}(\Omega)$, and the last equality in (9.2) indicates that the infimum is attained. If $1 < p < \infty$, then the v satisfying (9.1) and (9.2) is unique.

Proof. Define, for elements in X:

$$f^*(\Gamma u) = f(u).$$

so $f^* \in X'$ since Γ is an isometric isomorphism. Then,

$$\|f^*\|_{X'} = \sup_{\|\Gamma u\|_X \leq 1} f^*(\Gamma u) = \sup_{\|u\|_{W^{k,p}(\Omega)} \leq 1} f(u) = \|f\|_{(W^{k,p}(\Omega))'}.$$

By Hahn-Banach, there exists a (norm-preserving) extension $\hat{f}$ of f to $L^p(\Omega_{(k)})$. By the previous Lemma, there exists a unique $v \in L^{p'}(\Omega_{(k)})$ such that

$$\hat{f}(u) = \sum_{|\alpha| \leq k} \langle v_\alpha, u_\alpha \rangle, \quad u \in L^p(\Omega_{(k)})$$

Thus, for $u \in W^{k,p}(\Omega)$

$$\begin{aligned} f(u) = f^*(\Gamma u) = \hat{f}(\Gamma u) &= \sum_{0 \leq |\alpha| \leq k} \langle v_\alpha, (\Gamma u)_\alpha \rangle \\ &= \sum_{|\alpha| \leq k} \langle v_\alpha, D^\alpha u \rangle. \end{aligned}$$

This proves (9.1). (Observe that uniqueness of v is guaranteed for $\hat{f}$, i.e., such that $\hat{f}(u) = \sum_{|\alpha| \leq k} \langle v_\alpha, u_\alpha \rangle$ for all $u \in L^p(\Omega_{(k)})$, but not necessarily for f, i.e., not for all $u \approx \Gamma u \in X$.)

As seen, $\|f\|_{\left(W^{k,p}(\Omega)\right)'} = \|f^*\|_{X'}$, but the later equals

$$\|\hat{f}\|_{\left(L^p(\Omega_{(k)})\right)'} = \|v\|_{L^{p'}(\Omega_{(k)})}.$$

Now we have to show that

$$\|f\|_{\left(W^{k,p}(\Omega)\right)'} = \inf_{\mathcal{U}} \|w\|_{L^{p'}(\Omega_{(k)})} = \min_{\mathcal{U}} \|w\|_{L^{p'}(\Omega_{(k)})}.$$

We have already identified a $v \in L^{p'}(\Omega_{(k)})$ such that $\|f\|_{\left(W^{k,p}(\Omega)\right)'} = \|v|\|_{L^{p'}(\Omega_{(k)})}$. Thus, it suffices to show that if $w \in L^{p'}(\Omega_{(k)})$ is such that

$$f(u) = \sum_{|\alpha| \leq k} \langle w_\alpha, D^\alpha u \rangle$$

for all $u \in W^{k,p}(\Omega)$, then $\|w\|_{L^{p'}(\Omega_{(k)})} \geq \|v\|_{L^{p'}(\Omega_{(k)})}$. But such a w agrees with f^* on X, so it will be an extension of f^* to $L^p(\Omega_{(k)})$, and thus it must have norm at least equal to $\|f^*\|_{X'} = \|v\|_{L^{p'}(\Omega_{(k)})}$.

It remains to show uniqueness when $1 < p < \infty$. Suppose the conclusion holds for v_1 and v_2 attaining the minimum, so $\|v_1\|_{L^{p'}(\Omega_{(k)})} = \|f\|_{(W^{k,p}(\Omega))'} = \|v_2\|_{L^{p'}(\Omega_{(k)})} = 1$, where we can assume $= 1$ upon redefining f as $f/\|f\|_{(W^{k,p}(\Omega))'}$, and for all $u \in W^{k,p}(\Omega)$,

$$f(u) = \sum_{|\alpha| \leq k} \langle v_1, D^\alpha u \rangle = \sum_{|\alpha| \leq k} \langle v_2, D^\alpha u \rangle$$

First we claim that there exists a unique $x \in X$ such that

$$f^*(x) = \|x\|_{L^p(\Omega_{(k)})} = 1.$$

Since $\|f\|_{(W^{k,p}(\Omega_{(k)}))'} = \|f^*\|_{X'} = 1$, there exists $\{x_i\} \subset X$ such that $\|x_i\|_{L^p_{(\Omega_{(k)})}} = 1$ and $|f^*(x_i)| \to 1$. Modifying $\{x_i\}$ if needed (multiply by -1) we can assume that $f^*(x_i) \to 1$.

Because $L^p(\Omega_{(k)})$ is uniformly convex for $1 < p < \infty$, given $0 < \epsilon \le 2$, there exists a $\delta > 0$ such that (since $\|x_i\| = 1$) if $\|x_i - x_j\|_{L^p(\Omega_{(k)})} \ge \epsilon$, then $\frac{\|x_i + x_j\|}{2}_{L^p(\Omega_{(k)})} \le 1 - \delta$, thus if $\frac{\|x_i+x_j\|}{2}_{L^p(\Omega_{(k)})} > 1 - \delta$ we must have $\|x_i - x_j\|_{L^p(\Omega_{(k)})} < \epsilon$.

For large i we have $f^*(x_i) > 1 - \delta$ thus for large i, j we also have $f^*(\frac{x_i+x_j}{2}) > 1 - \delta$. Then, because f^* is continuous with norm 1:

$$1 - \delta < f^*\left(\frac{x_i + x_j}{2}\right) \le \frac{\|x_i + x_j\|}{2}_{L^p(\Omega_{(k)})}.$$

Hence $\|x_i - x_j\|_{L^p(\Omega_{(k)})} < \epsilon$, and $\{x_i\}$ is Cauchy and $x_i \to x$ in $L^p(\Omega_{(k)})$. Since X is closed, $x \in X$. Clearly $\|x\|_{L^p(\Omega_{(k)})} = 1$ and $f^*(x) = 1$. To obtain uniqueness, suppose that there are two such x's, x_1 and x_2. Then we apply the above argument to the sequence $\{x_1, x_2, x_1, x_2, ...\}$ which must converge.

Since by assumption v_1 and v_2 are two representatives of f^*, we have

$$f^*(x) = 1 = \sum_{|\alpha| \le k} \langle (v_1)_\alpha, x_\alpha \rangle = \sum_{|\alpha| \le k} \langle (v_2)_\alpha, x_\alpha \rangle.$$

Next, consider the claim: given $w \in L^p(\Omega_{(k)})$ with $\|w\|_{L^p(\Omega_{(k)})} = 1$, there exists at most one $l \in (L^p(\Omega_{(k)}))'$ such that $\|l\|_{(L^p(\Omega_{(k)}))'} = 1$ and $l(w) = 1$.

Let $\tilde{v}_1$ and $\tilde{v}_2$ be the extensions of v_1 and v_2, considered linear functionals on X, to $L^p(\Omega_{(k)})$ given by Hahn-Banach. Thus $\|\tilde{v}_1\|_{L^{p'}(\Omega_{(k)})} = 1 = \|\tilde{v}_2\|_{L^{p'}(\Omega_{(k)})}$ (observe that though $\tilde{v}_1 = f^* = \tilde{v}_2$ on X, we cannot claim $\tilde{v}_1 = \tilde{v}_2$ because the Hahn-Banach extension might not be unique), and by foregoing we have $\tilde{v}_1(x) = 1 = \tilde{v}_2(x)$. Thus, $\tilde{v}_1 = \tilde{v}_2$ by the claim.

It remains to prove the claim. Suppose there are two such l's, l_1 and l_2, $l_1 \neq l_2$. Then $l_1(u) \neq l_2(u)$ for some $u \in L^p(\Omega_{(k)})$. We can assume that

$$l_1(u) - l_2(u) = 2$$

upon replacing u by multiple of itself, and that

$$l_1(u) = 1 \text{ and } l_2(u) = -1$$

by replacing u with its sum with a suitable multiple of w. Thus

$$l_1(w + tu) = 1 + t,$$
$$l_2(w - tu) = 1 + t, \quad t > 0.$$

Since $\|l_1\|_{L^p(\Omega_{(k)})'} = 1 = \|l_2\|_{L^p((\Omega_{(k)}))'}$, we have

$$1 + t = l_1(w + tu) \le \|w + tu\|_{L^p(\Omega_{(k)})}$$
$$1 + t = l_2(w - tu) \le \|w - tu\|_{L^p(\Omega_{(k)})}.$$

Recall the L^p-parallelogram inequalities

$$\left\|\frac{a+b}{2}\right\|_{L^p}^p + \left\|\frac{a-b}{2}\right\|_{L^p}^p \ge \frac{1}{2}\|a\|_{L^p}^p + \frac{1}{2}\|b\|_{L^p}^p, \quad 1 < p \le 2,$$
$$\left\|\frac{a+b}{2}\right\|_{L^p}^{p'} + \left\|\frac{a-b}{2}\right\|_{L^p}^{p'} \ge \left(\frac{1}{2}\|a\|_{L^p}^p + \frac{1}{2}\|b\|_{L^p}^p\right)^{p'-1}, \quad 2 \le p < \infty.$$

If $1 < p \le 2$, we get

$$1 + t^p\|u\|_{L^p(\Omega_{(k)})}^p = \left\|\frac{(w + tu) + (w - tu)}{2}\right\|_{L^p(\Omega_{(k)})}^p + \left\|\frac{(w + tu) - (w - tu)}{2}\right\|_{L^p(\Omega_{(k)})}^p$$
$$\ge \frac{1}{2}\|w + tu\|_{L^p(\Omega_{(k)})}^p + \frac{1}{2}\|w - tu\|_{L^p(\Omega_{(k)})}^p \ge (1 + t)^p,$$

which cannot hold for all $t > 0$. If $2 \le p < \infty$, we apply the second inequality

$$1 + t^{p'} \|u\|^{p'}_{L^p(\Omega_{(k)})} = \left\| \frac{(w+tu)+(w-tu)}{2} \right\|^{p'}_{L^p(\Omega_{(k)})} + \left\| \frac{(w+tu)-(w-tu)}{2} \right\|^{p'}_{L^p(\Omega_{(k)})}$$
$$\ge \left(\frac{1}{2} \|w+tu\|^p_{L^p(\Omega_{(k)})} + \frac{1}{2} \|w-tu\|^p_{L^p(\Omega_{(k)})} \right)^{p'-1} \ge (1+t)^{p'}$$

which again is an impossibility. □

Definition 9.62. Consider $C^\infty_c(\Omega)$. For each compact subset $\kappa \subset \Omega$, let $\mathcal{D}_\kappa(\Omega)$ be the set of all $u \in C^\infty_c(\Omega)$ such that $\operatorname{supp}(u) \subset \kappa$. Define a family of semi-norms by

$$p_{\kappa,m}(u) = \sup_{\substack{|\alpha| \le m \\ x \in \kappa}} |D^\alpha u(x)|, \quad m \text{ integer},$$

$\mathcal{D}_\kappa(\Omega)$ is then a locally convex topological vector space. The strict inductive limit of $\mathcal{D}_\kappa(\Omega)$, when κ varies over all compact subsets of Ω is a locally convex topological vector space. We denote $C^\infty_c(\Omega)$ with this topology by $\mathcal{D}(\Omega)$.

A consequence of the definition is that a sequence $\{u_j\}$ converges to u in $\mathcal{D}(\Omega)$ if and only if (i) there exists a compact set $\kappa \subset \Omega$ such that $\operatorname{supp}(u_j) \subset \kappa$ for all j, (ii) for any multi-index α, the sequence $\{D^\alpha u_j\}$ converges uniformly to $D^\alpha u$ in κ.

Definition 9.63. A continuous linear form on $\mathcal{D}(\Omega)$ is called a **distribution.** The space of distributions on Ω is denoted $\mathcal{D}'(\Omega)$.

Thus, $f \in \mathcal{D}'(\Omega)$ if and only if it is a linear map $f : \mathcal{D}(\Omega) \to \mathbb{R}$ such that $f(u_j) \to f(u)$ when $u_j \to u$ in $\mathcal{D}(\Omega)$.

Definition 9.64. Let $f \in \mathcal{D}'(\Omega)$ and α be a multi-index. The α-derivative of f is the distribution $D^\alpha f$ defined by

$$D^\alpha f(u) = (-1)^{|\alpha|} f(D^\alpha u)$$

for every $u \in \mathcal{D}(\Omega)$. Observe that distributions are infinitely many times differentiable.

The motivation for this definition is clear: if ψ is smooth, it defines a distribution by

$$f_\psi(u) = \int_\Omega \psi(x) u(x) dx,$$

and $D^\alpha \psi$ by

$$f_{D^\alpha \psi}(u) = \int_\Omega D^\alpha \psi(x) u(x) dx.$$

Then, integrating by parts:

$$f_{D^\alpha \psi}(u) = (-1)^{|\alpha|} \int_\Omega \psi(x) D^\alpha u(x) dx = (-1)^{|\alpha|} f_\psi(D^\alpha u).$$

Let $f \in (W^{k,p}(\Omega))'$. By the above theorem, $f(u) = \sum_{|\alpha| \le k} \langle v_\alpha, D^\alpha u \rangle$ for some $v \in L^{p'}(\Omega_{(k)})$. Since $C^\infty_c(\Omega) \subset W^{k,p}(\Omega)$, $f|_{C^\infty_c(\Omega)}$ is well defined. To see that $f|_{C^\infty_c(\Omega)}$ defines a distribution, let $u_j \to u$ in $\mathcal{D}(\Omega)$. Then

$$f(u_j - u) = \sum_{|\alpha| \le k} \int_{\Omega_\alpha} v_\alpha(x) D^\alpha (u_j - u) dx$$
$$\le \sum_{|\alpha| \le k} \|v_\alpha\|_{L^{p'}(\Omega_\alpha)} \|D^\alpha (u_j - u)\|_{L^p(\Omega)}$$
$$= \sum_{|\alpha| \le k} \|v_\alpha\|_{L^{p'}(\Omega_\alpha)} \|D^\alpha (u_j - u)\|_{L^p(\kappa)}$$

for some compact κ; but $\|D^\alpha(u_j - u)\|_{L^p(\kappa)} \to 0$ as $j \to \infty$ since $D^\alpha u_j \to D^\alpha u$ uniformly on κ. Therefore, elements of $(W^{k,p}(\Omega))'$ can be viewed as extensions to $W^{k,p}(\Omega)$ of distributions. Similarly, $(W_0^{k,p}(\Omega))'$ can be viewed as a space of extensions to $W_0^{k,p}(\Omega)$ of continuous linear forms on $\mathcal{D}(\Omega)$. In the case of $W_0^{k,p}(\Omega)$, this in fact characterizes the dual:

Theorem 9.65. *Let $k \geq 1$ be an integer and $1 \leq p < \infty$. $(W^{k,p}(\Omega))'$ is isometrically isomorphic to a Banach space X consisting of those distributions $f \in \mathcal{D}'(\Omega)$ that have the form*

$$f = \sum_{|\alpha| \leq k} (-1)^{|\alpha|} D^\alpha f_{v_\alpha}, \quad f_{v_\alpha}(u) = \langle v_\alpha, u\rangle, \quad u \in \mathcal{D}(\Omega) \tag{9.3}$$

for some $v \in L^{p'}(\Omega_{(k)})$ and having norm

$$\|f\|_X = \inf_{\mathcal{U}} \|v\|_{L^{p'}(\Omega_{(k)})} = \min_{\mathcal{U}} \|v\|_{L^{p'}(\Omega_{(k)})}$$

where $\mathcal{U}$ is the set of all $v \in L^{p'}(\Omega_{(k)})$ for which f is given by (9.3).

Proof. First observe that the set of $f \in \mathcal{D}'(\Omega)$ of the form (9.3) with norm $\|f\|_X = \inf_{\mathcal{U}} \|v\|_{L^{p'}(\Omega_{(k)})}$ is a normed space.

Let $f \in \mathcal{D}'(\Omega)$ have the form (9.3). Let us show that it has a unique continuous extension to $W_0^{k,p}(\Omega)$. Let $u_j \to u$ in $W_0^{k,p}(\Omega)$, $\{u_j\} \subset C_c^\infty(\Omega)$. Then,

$$\begin{aligned} |f(u_j) - f(u_i)| &\leq \sum_{|\alpha| \leq k} |\langle f_{v_\alpha}, D^\alpha(u_j - u_i)\rangle| \\ &\leq \sum_{|\alpha| \leq k} \|D^\alpha(u_j - u_i)\|_{L^p(\Omega)} \|v_\alpha\|_{L^{p'}(\Omega)} \to 0 \text{ as } i, j \to 0. \end{aligned}$$

Thus $\{f(u_i)\}$ is Cauchy in $\mathbb{R}$ and converges. If we take another sequence, the limit is the same so we can define

$$\tilde{f}(u) = \lim_{j \to \infty} f(u_j)$$

for $u \in W_0^{k,p}(\Omega)$. Observe that f is linear, thus it defines an element of $(W_0^{k,p}(\Omega))'$.

Let $f \in (W_0^{k,p}(\Omega))'$. Then $f : W_0^{k,p}(\Omega) \to \mathbb{R}$ has a norm preserving extension $f^* : W^{k,p}(\Omega) \to \mathbb{R}$ and thus f^* has the form (9.3) by a theorem above and its norm is as stated. This applies in particular to the extension $\tilde{f}$ so the infimum is realized. Thus we have a norm preserving map $f \in X \mapsto \tilde{f} \in (W_0^{k,p}(\Omega))'$.

Reciprocally, if $f \in (W_0^{k,p}(\Omega))'$, then as just seen it is given by

$$f(u) = f^*(u) = \sum_{|\alpha| \leq k} \langle v_\alpha, D^\alpha u\rangle, \quad u \in W_0^{k,p}(\Omega)$$

with norm as indicated. As seen above, f restrictied to $\mathcal{D}(\Omega)$ gives rise to a distribution, i.e., we have a norm preserving map $f \in (W_0^{k,p}(\Omega))' \mapsto f_{\mathcal{D}(\Omega)} \in X \subset \mathcal{D}'(\Omega)$.

Consequently, X is complete since $(W_0^{k,p}(\Omega))'$ is.

□

Observe that the above argument does not hold, in general, for $W^{k,p}(\Omega)$: to uniquely extend $f \in X$ to an element of $(W_0^{k,p}(\Omega))'$ we used that any $u \in W_0^{k,p}(\Omega)$ is a limit of elements in $C_c^\infty(\Omega)$, where f is initially defined, but elements in $W^{k,p}(\Omega)$ cannot in general by approximated by $C_c^\infty(\Omega)$. In other words, when $W_0^{k,p}(\Omega)$ is a proper subspace of $W^{k,p}(\Omega)$, $f : W^{k,p}(\Omega) \to \mathbb{R}$ is not determined by its restriction to $C_c^\infty(\Omega)$. Thus, $f \in X$ extends to $W^{k,p}(\Omega)$ (by extending to $W_0^{k,p}(\Omega)$ uniquely and then to $W^{k,p}(\Omega)$ by Hahn-Banach) but this extension in general is not unique.

Definition 9.66. The Banach space X in the previous theorem, identified with $(W_0^{k,p}(\Omega))'$, is denoted $W^{-k,p'}(\Omega)$.

Observe that $W^{-k,p'}(\Omega)$ is separable and reflexive for $1 < p < \infty$.

We will now give another characterization of $W^{-k,p'}(\Omega)$ for $1 < p < \infty$.

Definition 9.67. Let $1 < p < \infty$. Any element $v \in L^{p'}(\Omega)$ determines a functional on $W_0^{k,p}(\Omega)$ by $f_v(u) = \langle v, u\rangle$, since

$$|f_v(u)| \leq \|v\|_{L^{p'}(\Omega)}\|u\|_{L^p(\Omega)} \leq \|v\|_{L^{p'}(\Omega)}\|u\|_{W^{k,p}(\Omega)}.$$

We define the $(-k, p')$**-norm of** $v \in L^{p'}(\Omega)$ as the norm of f_v, i.e.,

$$\|v\|_{-k,p'} \equiv \|v\|_{W^{-k,p'}(\Omega)} := \|f_v\|_{(W_0^{k,p}(\Omega))'} = \sup_{\substack{u\in W^{k,p}(\Omega)\\ \|u\|_{k,p}\leq 1}} |f_v(u)|$$

$$= \sup_{\substack{u\in W^{k,p}(\Omega)\\ \|u\|_{k,p}\leq 1}} |\langle v, u\rangle|.$$

Observe that $\|v\|_{-k,p'} \leq \|v\|_{L^{p'}(\Omega)}$ and

$$|\langle v, u\rangle| = \|u\|_{k,p}\left|\langle \frac{u}{\|u\|_{k,p}}, v\rangle\right| \leq \|u\|_{k,p}\|v\|_{-k,p'}$$

for all $u \in W^{k,p}(\Omega)$ and $v \in L^{p'}(\Omega)$, which is known as the **generalized Hölder's inequality.**

Theorem 9.68. *Let $1 < p < \infty$ and $k \geq 1$ be an integer. Then*

$$W^{-k,p'}(\Omega) = \textit{completion of } L^{p'}(\Omega) \textit{ w.r.t. } \|\cdot\|_{-k.p'}.$$

Proof. First we show that

$$X = \left\{f_v \mid v \in L^{p'}(\Omega)\right\}$$

where f_v is as above, is dense in $(W_0^{k,p}(\Omega))'$. If that is not the case, then there exists a $F \in (W_0^{k,p}(\Omega))' \setminus \bar{X}$. By one of the corollaries of Hahn-Banach, there exists $l \in (W_0^{k,p}(\Omega))''$ such that $l(F) \neq 0$ and $l\,|_{\bar{X}} = 0$. By reflexivity, there exists a $F_l \in W_0^{k,p}(\Omega)$ such that $l(\tilde{F}) = \tilde{F}(F_l)$ for every $\tilde{F} \in (W_0^{k,p}(\Omega))'$. Then, for $\tilde{F} \in X$ (so $\tilde{F} = f_v$), $\tilde{F}(F_l) = f_v(F_l) = \langle v, F_l\rangle$, but $\tilde{F}(F_l) = l(\tilde{F}) = l(f_v) = 0$, so $\langle v, F_l\rangle = 0$ for all $v \in L^{p'}(\Omega)$. Hence, $F_l = 0$ and thus $l = 0$ by $l(\tilde{F}) = \tilde{F}(F_l)$, a contradiction.

Let Y be the completion of $L^{p'}(\Omega)$ w.r.t. $\|\cdot\|_{-k,p'}$. Define $T : Y \to (W_0^{k,p}(\Omega))'$ by

$$T(y) = \lim_{j\to\infty} f_{v_j}$$

where $v_j \to y$ in Y and $\lim_{j\to\infty} f_{v_j}$ is the limit in $(W_0^{k,p}(\Omega))'$. Then

(i) T is well defined. If $\lim_{j\to\infty} v_j = y = \lim_{j\to\infty} w_j$, the limits in Y, then, since $\|v\|_{-k,p'} = \|f_v\|_{(W_0^{k,p}(\Omega))'}$,

$$\|f_{v_j} - f_{w_j}\|_{(W_0^{k,p}(\Omega))'} = \|v_j - w_j\|_{-k,p} \to 0,$$

so $T(y) = \lim_{j\to\infty} f_{v_j} = \lim_{j\to\infty} w_j$, limits in $(W_0^{k,p}(\Omega))'$.

(ii) T is linear.

(iii) T is one-to-one. If $T(y) = 0$, then

$$0 = \lim_{j\to\infty} \|f_{v_j}\|_{(W_0^{k,p}(\Omega))'} = \lim_{j\to\infty} \|v_j\|_{-k,p'}$$

so $y = \lim_{j\to\infty} v_j = 0$ in Y.

(iv) T is onto. Let $f \in (W_0^{k,p}(\Omega))'$. By the above density, $f = \lim_{j\to\infty} f_{v_j}$ in $(W_0^{k,p}(\Omega))'$ for some $\{v_j\} \subset X$. But

$$\|v_j - v_i\|_{-k,p'} = \|f_{v_j} - f_{v_i}\|_{(W_0^{k,p}(\Omega))'}$$

thus $\{v_j\}$ is Cauchy in Y, $v_j \to y$ and

$$f = \lim_{j\to\infty} f_{v_j} = T(y).$$

If $T(y) = f$, then we have

$$T(y) = \lim_{j\to\infty} f_{v_j}, \text{ limit in } (W_0^{k,p}(\Omega))'$$

for some $v_j \to y$ in Y. Because the limit $\lim_{j\to\infty} f_{v_j}$ is in $(W_0^{k,p}(\Omega))'$,

$$\begin{aligned}\|T(y)\|_{(W_0^{k,p}(\Omega))'} &= \lim_{j\to\infty} \|f_{v_j}\|_{(W_0^{k,p}(\Omega))'} = \lim_{j\to\infty} \|v_j\|_{-k,p'} \\ &= \|y\|_{-k,p}\end{aligned}$$

since $v_j \to y$ in Y. Thus, Y is an isometric isomorphism.

□

We observe the following consequences of the above theorem and its proof:

- Since any $f \in (W_0^{k,p}(\Omega))'$ is of the form

$$f = \lim_{j\to\infty} f_{v_j}, \quad v_j \to v \in Y,$$

we can extend the notation $\langle , \rangle$ to mean

$$\langle v, u\rangle = f_v(u) = \lim_{j\to\infty} f_{v_j}(u) = \lim_{j\to\infty} \langle v_j, u\rangle$$

for all $v \in Y$ and $u \in W_0^{k,p}(\Omega)$. Thus, any linear functional f on $W_0^{k,p}(\Omega)$ can be represented as

$$f(u) = \langle v, u\rangle$$

for some $v \in W^{-k,p}(\Omega)$, i.e., some v with finite $\|\cdot\|_{-k,p'}$ norm.

- We can extend the generalized Hölder inequality:

$$|\langle v, u\rangle| \le \|v\|_{-k,p'}\|u\|_{k,p},$$

$v \in W^{-k,p'}(\Omega),\ u \in W^{k,p}(\Omega)$.

Remark 9.69. The same argument as above shows that $(W_0^{k,p}(\Omega))'$, $1 < p < \infty$, can be identified with the completion of $L^{p'}(\Omega)$ with respect to the norm

$$|||v|||_{-k,p'} = \sup_{\substack{u\in W^{k,p}(\Omega) \\ \|u\|_{k,p}\le 1}} |\langle v, u\rangle|,$$

$v \in L^{p'}(\Omega)$, and that the above representation and generalized Hölder's inequality also hold. Sometimes the dual of $W^{k,p}(\Omega)$ is also denoted $W^{-k,p'}(\Omega)$ with the $\|\cdot\|_{-k,p'}$, norm denoted by $|||\cdot|||_{-k,p'}$.

Remark 9.70. We can verify that for $p = 2$ and $\Omega = \mathbb{R}^2$, the above construction of H^{-s} agrees with $H^{(-s)}$.

Definition 9.71. We define the pairing between $L^p\big([0,T],H^{(s)}\big)$ and $L^{p'}\big([0,T],H^{(-s)}\big)$, as

$$\langle u,v\rangle = \int_0^T ((1-\Delta)^{\frac{s}{2}}u,(1-\Delta)^{-\frac{s}{2}}v)_0 dt,$$

for $u\in L^p([0,T],H^{(s)})$, $v\in L^p([0,T],H^{(-s)})$.

(Note that here $\langle\,,\rangle$ is denoting something different than above.)

The definition makes sense because $((1-\Delta)^{\frac{s}{2}}u,(1-\Delta)^{-\frac{s}{2}}v)_0$ is integrable in t:

$$|((1-\Delta)^{\frac{s}{2}}u,(1-\Delta)^{-\frac{s}{2}}v)_0| \le \|u(t)\|_{H^{(s)}}\|v(t)\|_{H^{(-s)}},$$

so

$$|\langle u,v\rangle| \le \|u\|_{L^p\big([0,T],H^{(s)}\big)}\|v\|_{L^{p'}\big([0,T],H^{(-s)}\big)}$$

Notation 9.72. We abbreviate

$$L^p\big([0,T],H^{(s)}\big) = L^p_tH^{(s)}_x = L^p_tH^s_x$$

Theorem 9.73. *Given* $f\in\big(L^1_tH^{(s)}_x\big)'$, *there exists a* $v\in L^\infty_tH^{(-s)}_x$ *such that*

$$f(u)=\langle u,v\rangle$$

for all $u\in L^t_1H^{(s)}_x$.

Proof. This follows from the duality between $H^{(s)}$ and $H^{(-s)}$ and the Riesz representation theorem. We leave it as an exercise. □

9.11. **Some miscellaneous inequalities.** We collect here some inequalities that will be use later on, mostly in the study of non-linear problems. Their proofs can be found in most books on the topic.

- Let $u_1,...,u_l\in H^l(\mathbb{R}^n)\cap L^\infty(\mathbb{R}^n)$, $\alpha_1,...,\alpha_l$ be multi-indices with $\sum_{i=1}^l|\alpha_l|=k$. Then

$$\|D^{\alpha_1}u_1\ldots D^{\alpha_l}u_l\|_{L^2(\mathbb{R}^n)} \le C\sum_{i=1}^l\|D^ku_i\|_{L^2(\mathbb{R}^n)}\prod_{j\ne i}\|u_j\|_{L^\infty(\mathbb{R}^n)}$$

  (This inequality also holds in $C^\infty_c(\Omega)$ by considering zero extensions)
- Let $F\in C^\infty(\mathbb{R}^{n+1})$ be such that $F(x,0)=0$ for every $x\in\mathbb{R}^n$. Assume that for each non-negative integer j and multi-index α there exists a continuous increasing function $F_{\alpha,j}$ such that

$$|D^\alpha_xD^\alpha_yF(x,y)|\le F_{\alpha,j}(|y|)$$

  for all $(x,y)\in\mathbb{R}^{n+1}$. Then, if $u\in H^k(\mathbb{R}^n)\cap L^\infty(\mathbb{R}^n)$, $F(\cdot,u)\in W^k(\mathbb{R}^n)$ (i.e., it is k-times weakly differentiable) and its derivatives are given by the chain rule. Moreover, there exists a continuous increasing function C such that

$$\|F(\cdot,u)\|_{H^k(\mathbb{R}^n)}\le C\big(\|u\|_{L^\infty(\mathbb{R}^n)}\big)\|u\|_{H^k(\mathbb{R}^n)}.$$

- If $u\in W^{k,p}(\Omega)$, $v\in W^{k,p}(\Omega)$, $k>\frac{n}{p}$, then $uv\in W^{k,p}(\Omega)$ and

$$\|uv\|_{W^{k,p}(\Omega)}\le C\|u\|_{W^{k,p}(\Omega)}\|v\|_{W^{k,p}(\Omega)}.$$

  These assertions hold for Ω satisfying the cone condition.
- Interpolation:

$$\|u\|_{H^{(s_2)}(\mathbb{R}^n)}\le\|u\|^{\frac{s_3-s_2}{s_3-s_1}}_{H^{(s_1)}(\mathbb{R}^n)}\|u\|^{\frac{s_2-s_1}{s_3-s_1}}_{H^{(s_3)}(\mathbb{R}^n)},$$

  $s_1<s_2<s_3$.

For Ω satisfying the cone condition, we also have, $0 \leq m \leq k$,

$$\|u\|_{W^{k,p}(\Omega)} \leq C\big(\epsilon\|u\|_{W^{k,p}(\Omega)} + \epsilon^{-\frac{m}{k-m}}\|u\|_{L^p(\Omega)}\big)$$
$$|u|_{W^{k,p}(\Omega)} \leq C\big(\epsilon|u|_{W^{k,p}(\Omega)} + \epsilon^{-\frac{m}{k-m}}\|u\|_{L^p(\Omega)}\big)$$

where $|u|_{W^{k,p}(\Omega)} := \big(\sum_{|\alpha|=k}|D^\alpha u|^p\big)^{\frac{1}{p}}$

$$\|u\|_{W^{k,p}(\Omega)} \leq C\|u\|^{\frac{m}{k}}_{W^{k,p}(\Omega)}\|u\|^{\frac{(k-m)}{k}}_{W^{m,p}(\Omega)}$$

Moreover, let q satisfy: $p \leq q < \infty$ if $kp > n$; $p \leq q \leq \infty$ if $kp = n$; $p \leq q \leq p^* = \frac{kp}{n-kp}$ if $kp < n$. Then

$$\|u\|_{L^q(\Omega)} \leq C\|u\|^{\theta}_{W^{k,p}(\Omega)}\|u\|^{1-\theta}_{L^p(\Omega)},$$

$\theta = \frac{n}{kp} - \frac{n}{kq}$.

10. Necessary and sufficient condition for existence of solutions to linear PDEs

In this section we will assume Ω to be a bounded domain with smooth boundary. We will also consider only operations of second order. More general domains and operators also be considered by adapting the proofs below.

Consider the boundary-value problem

$$(BVP)\begin{cases} Lu = f \text{ in } \Omega \\ Nu = 0 \text{ on } \partial\Omega \end{cases}$$

where throughout we assume

$$Lu = a^{ij}\partial_i\partial_j u + b^i\partial_i u + cu,$$
$$Nu = \alpha\frac{\partial u}{\partial \nu} + \beta\cdot D_\tau u + \gamma u,$$

where D_τ is the tangential gradient and $a^{ij}, b^j, c, \alpha, \beta, \gamma$, are smooth in $\bar{\Omega}$ and $a^{ij} = a^{ji}$.

Consider the following formal computation.

$$\begin{aligned}
\int_\Omega vLu &= \int_\Omega v(a^{ij}\partial_i\partial_j u + b^i\partial_i u + cu) \\
&= -\int_\Omega a^{ij}\partial_i v\partial_j u - \int_\Omega v\partial_i a^{ij}\partial_j u - \int_\Omega b^i\partial_i vu - \int_\Omega \partial_i b^i vu \\
&\quad + \int_\Omega cuv + \int_{\partial\Omega} a^{ij}v\partial_j u\nu_i + \int_{\partial\Omega} b^i vu\nu_i \\
&= \int_\Omega a^{ij}\partial_i\partial_j vu + \int_\Omega \partial_j a^{ij}\partial_i vu + \int_\Omega \partial_i a^{ij}\partial_j vu + \int_\Omega \partial_i\partial_j a^{ij}vu \\
&\quad - \int_\Omega b^i\partial_i vu - \int_\Omega \partial_i b^i vu + \int_\Omega cuv + \int_{\partial\Omega} a^{ij}v\partial_j u\nu_i + \int_{\partial\Omega} b^i vu\nu_i \\
&\quad - \int_{\partial\Omega} a^{ij}\partial_i vu\nu_j - \int_{\partial\Omega} \partial_i a^{ij}vu\nu_j \\
&= \int_\Omega u\big[a^{ij}\partial_i\partial_j v + (2\partial_j a^{ij} - b^i)\partial_i v + (c - \partial_i b^i + \partial_i\partial_j a^{ij})v\big] \\
&\quad + \int_{\partial\Omega} (a^{ij}v\partial_j u + b^i vu - a^{ij}\partial_j vu - \partial_j a^{ij}vu)\nu_j.
\end{aligned}$$

This motivates the following.

Definition 10.1. The **formal adjoint** L^* of L is the operator

$$L^*u = a^{ij}\partial_i\partial_j u + (2\partial_j a^{ij}\partial v - b^i)\partial_i v + (c - \partial_i b^i + \partial_i\partial_j a^{ij})u$$

Definition 10.2. Let $C^\infty_N(\bar\Omega)$ be the space of $u \in C^\infty(\bar\Omega)$ satisfying $Nu = 0$ on $\partial\Omega$. A function $v \in C^1(\bar\Omega)$ is said to satisfy the **adjoint boundary condition** $N^*v = 0$ if

$$(Lu, v)_0 = (u, L^*v)$$

for all $u \in C^\infty_N(\bar\Omega)$. The space of all $v \in C^\infty(\bar\Omega)$ satisfying $N^*v = 0$ will be denoted $C^\infty_{N^*}(\Omega)$.

Example 10.3. If $Nu = u$ (Dirichlet boundary condition), then $N^*v = v$ since in this case the boundary term becomes

$$\int_{\partial\Omega}(a^{ij}v\partial_j u + b^i vu - a^{ij}\partial_j vu - \partial_j a^{ij}vu)\nu_i = \int_{\partial\Omega} a^{ij}\partial_j u\nu_i v$$

thus one needs $v = 0$ on $\partial\Omega$, provided $a^{ij}\partial_j u\nu_i \neq 0$ (say, if L is a an elliptic operator, to be defined later).

Definition 10.4. Let k be an integer. We say that $u \in H^k(\Omega)$ is a **weak solution** to BVP if

$$(u, L^*v)_0 = (f, v)_0$$

for all $v \in C^\infty_{N^*}(\Omega)$.

The idea is that if u is a smooth solution, then integration by parts gives the above equality for $v \in C^\infty_{N^*}(\bar\Omega)$.

In the next theorem and what follows, $H^{-s}(\Omega)$ is the dual space of $H^s(\Omega)$ (and not of $H^s_0(\Omega)$).

Theorem 10.5. *Let $s, t \geq 0$ be integers. Then, there exists a weak solution $u \in H^s(\Omega)$ to BVP for each $f \in H^t(\Omega)$ if and only if there exists a constant $\kappa > 0$ such that*

$$\|v\|_{-t} \leq \kappa\|L^*v\|_{-s}$$

for all $v \in C^\infty_{N^*}(\Omega)$ (recall $\|\cdot\|_s = \|\cdot\|_{H^s(\Omega)}$).

Remark 10.6. Observe that the theorem makes no statement about uniqueness.

Proof. We first need some auxiliary constructions. Let ($k \geq 0$, integer)

$$T : \underbrace{H^{-k}(\Omega)}_{\substack{\text{completion of } L^2(\Omega) \\ \text{w.r.t. } \|\cdot\|_{-k}}} \to (H^k(\Omega))'$$

be the isometric isomorphism that identifies these two spaces. Let

$$R : (H^k(\Omega))' \to H^k(\Omega)$$

be the isometric isomorphism given by the Riesz representation theorem. Set

$$(u, v)_{-k} := (R \circ T(u), R \circ T(v))_k.$$

This defines an inner product on $H^{-k}(\Omega)$. We have

$$\begin{aligned}(u, u)_{-k} &= (R \circ T(u), R \circ T(u))_k = \|R(T(u))\|^2_k \\ &= \|T(u)\|^2_{(H^k(\Omega))'} = (\sup_{\substack{v \in H^k(\Omega) \\ \|v\|_k = 1}} |(u, v)|)^2 = \|u\|^2_{-k},\end{aligned}$$

so $(\cdot,\cdot)_{-k}$ generates the $H^{-k}(\Omega)$ topology. We already know that an element in $(H^k(\Omega))'$ is represented by an element in $H^{-k}(\Omega)$. Conversely, a $g \in (H^k(\Omega))'$ is uniquely represented by a $v \in H^k(\Omega)$ via

$$g(u) = (v,u)_0, u \in H^{-k}(\Omega).$$

The argument is similar to what we did for the first identification, showing that the functions of the form $g_v(u) = (v,u), v \in H^k(\Omega)$ form a dense set. Moreover,

$$\|g_v\|_{(H^{-k}(\Omega))'} = \sup_{\substack{u\in H^{-k}(\Omega)\\ \|u\|_{-k}\leq 1}} |(v,u)_0|$$

But we can also write $(v,u)_0 = f_u(v)$, $f_u \in (H^k(\Omega))'$. By one of the corollaries of Hahn-Banach, we can choose a u' such that $f_{u'}(v) = \|v\|_k$ and $\|f_{u'}\|_{(H^k(\Omega))'} = 1$ so

$$\|g_v\|_{(H^{-k}(\Omega))'} = \sup_{\substack{u\in H^{-k}(\Omega)\\ \|u\|_{-k}\leq 1}} |(v,u)_0| \geq |f_{u'}(v)| = \|v\|_k,$$

thus $\|g_v\|_{(H^{-k}(\Omega))'} = \|v\|_k$ by the generalized Hölder inequality.

Therefore, the construction $T : H^{-k}(\Omega) \to (H^k(\Omega))'$ carries over to $T : H^k(\Omega) \to (H^{-k}(\Omega))'$, where we slightly abuse notation by calling the later map T as well (so we have $T : H^{-k}(\Omega) \to (H^k(\Omega))'$ for $k \in \mathbb{Z}$).

We can now prove the theorem.

Assume the estimate. Set

$$X = L^* C^\infty_{N^*}(\bar{\Omega}) \subset H^{-s}(\Omega).$$

For $f \in H^t(\Omega)$, set $F : X \to \mathbb{R}$ by

$$F(L^*v) = (f,v)_0.$$

This is well defined because the estimate gives that $v = 0$ if $\|L^*v\|_{-s} = 0$. By the generalized Hölder inequality,

$$|F(L^*v)| \leq \|f\|_t\|v\|_{-t} \leq \kappa\|f\|\|L^*v\|_{-s},$$

thus F is a bounded linear functional on X. By Hahn-Banach, F extends to $\tilde{F} : H^{-s}(\Omega) \to \mathbb{R}$, i.e., $\tilde{F} \in (H^{-s}(\Omega))'$. By the foregoing discussion, the linear functional $\tilde{F}$ can be represented by a $u \in H^s(\Omega)$ via

$$\tilde{F}(w) = (u,w)_0$$

for all $w \in H^{-s}(\Omega)$. In particular, if $w \in X$,

$$(u,w)_0 = (u, L^*v)_0 = \tilde{F}(L^*v) = F(L^*v) = (f,v)_0,$$

i.e., $(u, L^*v)_0 = (f,v)_0$ for all $v \in C^\infty_{N^*}(\bar{\Omega})$, showing that u is a weak solution.

We can now prove the converse. Assume that for each $f \in H^t(\Omega)$ there exists a weak solution $u = u_f \in H^s(\Omega)$. The map

$$g_f(v) = (f,v)_0,\ v \in H^{-t}(\Omega),$$

defines a bounded linear function on $H^{-t}(\Omega)$, i.e., $g_f \in (H^{-t}(\Omega))'$. Since $H^{-t}(\Omega)$ is a Hilbert space as seen above, by the Riesz representation theorem there exists a $R(g_f) \in H^{-t}(\Omega)$ such that

$$g_f(v) = (R(g_f), v)_{-t}$$

On the other hand, $R(g_f)$ is the image by $T^{-1} \circ R^{-1}$ of an element in H^t, and this element must be f since the maps involved are all isometric isomorphisms:

$$(H^{-t}(\Omega))' \xrightarrow{R} H^{-t}(\Omega) \xrightarrow{T} (H^t(\Omega))' \xrightarrow{R} H^t(\Omega),$$

$$(T^{-1} \circ R^{-1}(f), v)_{-t} = g_f(v) = (f,v)_0$$

This holds for every $v \in H^{-t}(\Omega)$. In particular, for $v \in C^\infty_{N^*}(\bar\Omega)$ we have

$$(T^{-1} \circ R^{-1}(f), v)_{-t} = (f, v)_0 = (u_f, L^* v)_0,$$
$$|(T^{-1} \circ R^{-1}(f), v)_{-t}| \leq \|u_f\|_s \|L^* v\|_{-s} \leq C_f \|L^* v\|_{-s}.$$

Since the maps T and R are isomorphisms, any $w \in H^{-t}(\Omega)$ is of the form $T^{-1} \circ R^{-1}(f)$ for some f. Thus, we define, for each $v \in C^\infty_{N^*}(\bar\Omega)$, the functional

$$J_v(w) = \Bigg(\underbrace{T^{-1} \circ R^{-1}(f)}_{w}, \frac{v}{\|L^* v\|_{-s}}\Bigg)_{-t}.$$

Hence, we have a family $\{J_v\}_{v \in C^\infty_{N^*}(\bar\Omega)} \subset (H^{-t}(\Omega))'$ with the property that for each $w \in H^t(\Omega)$

$$|J_v(w)| \leq C_w,$$

i.e., the family is pointwise bounded. Therefore, by the Banach-Steinhaus theorem, the family $\{J_v\}_{v \in C^\infty_{N^*}(\bar\Omega)}$ is uniformly bounded, i.e., $\|J_v\|_{(H^{-t}(\Omega))'} \leq C$ for all $v \in C^\infty_{N^*}(\bar\Omega)$. Thus, for any $w \in H^{-t}(\Omega)$

$$\left| \left(\frac{w}{\|w\|_{-t}}, \frac{v}{\|L^* v\|_{-s}} \right)_{-t} \right| \leq C.$$

Choosing $w = v$ gives

$$\frac{\|v\|^2_{-t}}{\|v\|_{-t}\|L^* v\|_{-s}} \leq C,$$

hence the result. □

10.1. **Egorov's counterexample: a PDE that is not locally solvable at the origin.** For this section, we will further restrict the notion of weak solution. We continue to assume L and Ω to have the same form as in the previous section.

Definition 10.7. We say that $u \in H^s(\Omega)$ is a weak solution to $Lu = f$ in Ω if $(u, L^* v)_0 = (f, v)_0$ for all $v \in C^\infty_c(\Omega)$, $s, t, \in \mathbb{Z}$.

Definition 10.8. Let $\Omega \subset \mathbb{R}^n$ contain the origin. We say that L is **locally solvable at the origin** if given $f \in C^\infty_c(\Omega)$, there exists a $\tilde\Omega \subset \Omega$, $\tilde\Omega \ni 0$, and a $u \in H^{-s}(\Omega), s \in \mathbb{N}$, such that $Lu = f$ holds weakly in $\tilde\Omega$.

We will henceforth consider the operator

$$Lu = \partial_t^2 u - a^2(t)\partial_x^2 u + b(t)\partial_x u,$$

$(t, x) \in \mathbb{R}^2$, $a, b \in C^\infty(\mathbb{R})$. We will present an example, due to Egorov, of a choice of, a, b such that L is not locally solvable at the origin. The construction is a bit tedious, involving some cumbersome computations, and we will give only the main steps.

Lemma 10.9. *If $Lu = f$ always has a (weak) solution in $\Omega \subset \mathbb{R}^2$ for any given $f \in C^\infty_c(\Omega)$ then there exists a $C > 0$ and $N \in \mathbb{N}$ such that*

$$\|v\|_0 \leq C\|L^* v\|_N$$

for all $v \in C^\infty_c(\Omega)$.

Proof. Using that now we established $H^{-k}(\Omega) \simeq (H^k(\Omega))'$, the necessary condition for existence can be extended for $s, t \in \mathbb{Z}$. Thus there exist $s, t \in \mathbb{Z}$ and $C > 0$ such that:

$$\|v\|_s \leq C\|L^* v\|_t \tag{10.1}$$

If $s \geq 0$, then $\|v\|_0 \leq \|v\|_s$ and we choose $N \geq t$. Otherwise, note first that we can assume $t \geq s$ since if $t < s$ then we can choose $\tilde{t} \geq s$ and work with $\tilde{t}$ instead (so $\|L^*v\|_t \leq \|L^*v\|_{\tilde{t}}$). Since $D_x^\alpha v \in C_c^\infty(\Omega)$ if $v \in C_c^\infty(\Omega)$, we can apply the inequality (10.1) to $D_x^\alpha v$,

$$\|D_x^\alpha v\|_s \leq C\|L^* D_x^\alpha v\|_t = C\|D_x^\alpha L^* v\|_t \leq C\|L^*v\|_{t+|\alpha|}$$

where we used that $L^*v = \partial_t^2 v - a(t)\partial_x^2 v - b(t)\partial_x v$. This latter expression also gives

$$\begin{aligned}\|\partial_t^2 v\|_{s-1} &\leq C \underbrace{\|L^*v\|_{s-1}}_{\substack{\leq \|L^*v\|_{t+1} \\ \text{by } s \leq t}} + \underbrace{\|\partial_x^2 v\|_{s-1} + \|\partial_x v\|_{s-1}}_{\substack{\leq \|L^*v\|_{t+1} \\ \text{by the above with } D_x^\alpha = \partial_x^2 \\ \text{and } s-1 \text{ in place of } s}} \\ &\leq C\|L^*v\|_{t+1}\end{aligned}$$

Then

$$\begin{aligned}\|v\|_{s+1} &\leq \underbrace{\|v\|_s}_{\leq \|L^*v\|_t} + \underbrace{\|\partial_t^2 v\|_{s-1} + \|\partial_x^2 v\|_{s-1}}_{\leq C\|L^*v\|_{t+1}} \\ &\leq C\|L^*v\|_{t+1}.\end{aligned}$$

Iterating this argument

$$\|v\|_{s+l} \leq C\|L^*v\|_{t+l}$$

and we choose l such that $s + l = 0$ and $N = t + l$. □

Theorem 10.10. *There exist $a, b \in C^\infty(\mathbb{R})$ such that given $f \in C_c^\infty(\Omega)$, Ω containing the origin in $\mathbb{R}^2, Lu = f$ has no weak solution $u \in H^{-s}(\Omega), s \in \mathbb{N}$.*

Remark 10.11. This is not yet saying that L is not locally solvable.

Proof. Set

$$a(t) := \begin{cases} e^{-t^2 - \sin^{-2}\left(\frac{1}{t^2}\right)}, & t > 0, \\ 0, & t \leq 0, \end{cases}$$

$$b(t) := \begin{cases} -2a(t)\xi'(t) - a'(t), & t > 0 \\ 0, & t \leq 0, \end{cases}$$

where $\xi(t) - sin^{-4}(\frac{1}{t}) - \ln(t)$. One can verify that these functions are smooth.

Notice that a oscillates very fast on the intervals $I_\mu := \left(\frac{1}{\pi(\mu+1)}, \frac{1}{\pi\mu}\right)$. We will use this to violate the inequality in the previous Lemma by constructing a sequence of functions $v_{\mu\lambda} \in C_c^\infty(I_\mu x(-\frac{2}{\lambda}, \frac{2}{\lambda}))$ which makes the RHS smaller than the LHS for large μ, λ.

We first search for approximate solutions $L^*v_{\mu v} \approx 0$. The change of variables $\bar{t} = t, \bar{x} = x - \int_0^t a(t')dt'$ is smooth near the origin and setting $\bar{v}(\bar{t}, \bar{x}) = v(t, x)$,

$$\begin{aligned}L^*\bar{v} &= \frac{\partial \bar{v}}{\partial \bar{t}^2} - 2a\frac{\partial^2 \bar{v}}{\partial \bar{x}\partial \bar{t}} - (b + a')\frac{\partial \bar{v}}{\partial \bar{x}} \\ &= \frac{\partial^2 \bar{v}}{\partial \bar{t}^2} - 2a\frac{\partial^2 \bar{v}}{\partial \bar{x}\partial \bar{t}} - 2\xi' a\frac{\partial \bar{v}}{\partial \bar{x}}\end{aligned}$$

We now drop the $^-$'s. Set

$$v_{\mu\lambda}(t, x) = \sum_{i=0}^{l} \frac{1}{\lambda^i} z_i(t) w_i(\lambda x)$$

for $z_i \in C_c^\infty(I_\mu), w_i \in C_C^\infty(-1,1)$, so $v_{\mu\lambda} \in C_c^\infty(J_{\mu\lambda})$, $J_{\mu\lambda} := I_\mu \times (-\frac{1}{\lambda}, \frac{1}{\lambda})$. Calculate

$$\begin{aligned} L^* v_{\mu\lambda} &= -\lambda 2a w_0'(z_0' + \xi' z_0) \\ &\quad + z_0'' w_0 - 2a w_1'(z_1' + \xi' z_1) \\ &\quad + \frac{1}{\lambda}(z_1'' w_1 - 2a w_2'(z_2' + \xi' z_2)) \\ &\quad + \dots \\ &\quad + \frac{1}{\lambda^l} z_l'' w_l. \end{aligned}$$

Take $w_l \in C_c^\infty(-1,1)$ and set

$$w_i(\tilde{x}) = (\frac{d}{d\tilde{x}})^{l-i} w_l(\tilde{x}), \quad \tilde{x} = \lambda x,$$

so $w_{i+1}' = w_i$. Set

$$\begin{aligned} z_0 &= e^{-\xi} \\ z_{i+1}(t) &= e^{-\xi(t)} \int_0^t \frac{z_i''(t')}{2a(t')} e^{\xi(t')} dt'. \end{aligned}$$

then $z_i \in C^\infty(I_\mu)\big(e^{-\sin^{-4}(\frac{1}{t})}$ dominates $(\frac{z_i''}{2a})e^{\xi}$ because the latter is $o(t^{-\alpha} e^{\beta \sin^{-2}(\frac{2}{t})})$, $\alpha, \beta > 0\big)$. Then, with these choices $v_{\mu\lambda} \in C_c^\infty(J_{\mu\lambda})$. Moreover,

$$\|L^* v_{\mu\lambda}\|_{H^N(J_{\mu\lambda})} \leq \underbrace{C_{\mu l N}}_{\substack{\text{does not} \\ \text{depend on } \lambda}} \lambda^{-N},$$

and

$$\|v_{\mu\lambda}\|^2_{L^\alpha(J_{\mu\lambda})} \geq \int_{J_{\mu\lambda}} e^{-2\xi(t)} w_0^2(\lambda x)\, dt\, dx - \underbrace{D_{\mu l}}_{\substack{\text{does not depend} \\ \text{on } \lambda \text{ or } N}} \lambda^{-1}$$

Compute

$$\int_{-\frac{1}{\lambda}}^{\frac{1}{\lambda}} w_0^2(\lambda x)\, dx \underset{\lambda x = \tilde{x}}{=} \int_{-1}^{1} w_0^2(\tilde{x}) \lambda^{-1}\, d\tilde{x} = \lambda^{-1} \underbrace{C_l}_{\substack{\text{depends on } l \\ \text{because } w_0 \text{ does}}}$$

Thus

$$\|v_{\mu\lambda}\|^2_{L^2(J_{\mu\lambda})} \geq \underbrace{E_{\mu l}}_{\substack{\text{does not depend} \\ \text{on } \lambda \text{ or } N}} \lambda^{-1}$$

Thus, for large μ and large λ (choosing λ after fixing μ large) we violate the inequality of the previous lemma. □

Corollary 10.12. *There exists $f \in C_c^\infty(\mathbb{R}^2)$ such that $Lu = f$ has no weak solution $u \in H^{-s}(\Omega)$, $s \in \mathbb{N}$, for any Ω containing the origin.*

Proof. For a fixed Ω containing the origin, set

$$X_s(\Omega) := \big\{ f \in C_c^\infty(\mathbb{R}^2) \mid Lu = f \text{ has a weak solution } u \in H^{-s}(\Omega) \text{ such that } \|u\|_{-s} \leq |s| + 1 \big\}$$

$$X(\Omega) := \big\{ f \in C_c^\infty(\mathbb{R}^2) \mid Lu = f \text{ has a weak solution } u \in H^{-s}(\Omega) \text{ for some } s \in \mathbb{N} \big\}.$$

If $f \in X(\Omega)$, then $f \in X_s(\Omega)$ for some s. For, if $u \in H^{-s_0}(\Omega)$ is a weak solution, then since $\|u\|_{-s} \leq \|u\|_{-s_0}, -s \leq -s_0$, we have $u \in H^{-s}(\Omega)$, $-s \leq -s_0$. If there is no s such that $\|u\|_s \leq |s|+1$, then we have $\|u\|_{-s} > |s|+1$ for all $-s \leq -s_0$, so taking $-s$ to be sufficiently negative we can make $\|u\|_{-s_0}$ arbitrarily large. Thus

$$X(\Omega) = \bigcup_{s=-\infty}^{\infty} X_s(\Omega).$$

$X(\Omega)$ is a Fréchet space with topology given by the collection of semi-norms $\|\cdot\|_s, s \in \mathbb{N}$. Thus, its topology is generated by open sets

$$\mathcal{U}_{s_1,\ldots,s_l} = \big\{x \in X(\Omega) \mid \|x\|_{s_1} < \epsilon, \ldots, \|x\|_{s_l} < \epsilon\big\}.$$

Hence $x_i \to x$ in $X(\Omega)$ iff $\|x_i - x\|_s \to 0$ for each $s \in \mathbb{N}$.

Let $\{f_j\} \subset X_s(\Omega)$ converge to f. For each j there exists f_j such that $Lu_j = f_j$ weakly and $\|u_j\|_{-s} \leq |s|+1$, so the sequence $\{u_j\}$ is bounded in $H^{-s}(\Omega)$ and thus has a convergent subsequence, still denoted $\{u_j\}$, converging to a limit $u \in H^{-s}(\Omega)$. But $(f_j, v)_0 \to (f, v)_0$ for all $v \in C_c^\infty(\Omega)$. We also have $(u_j, L^*v)_0 \to (u, L^*v)_0$ for all $v \in C_c^\infty(\Omega)$, since $H^{-s}(\Omega) \hookrightarrow H^{-s-1}(\Omega)$ compactly, $-s-1 < -s$, $u_j \to u$ in $H^{-s-1}(\Omega)$, which gives the claim. Therefore,

$$Lu = f \text{ weakly},$$

and we conclude that $X_s(\Omega)$ is closed.

Let $f \in X_s(\Omega)$. By the above, there exists a $\hat{f} \in C_c^\infty(\Omega)$ such that $Lu = \hat{f}$ has no weak solution $u \in H^{-s}(\Omega)$ for any $s \in \mathbb{N}$. If $f + t\hat{f} \in X_s(\Omega)$, then there exists a $w \in H^{-s}(\Omega)$ such that $Lw = f + t\hat{f}$ weakly. And since $X_s(\Omega)$, there exists a $z \in H^{-s}(\Omega)$ such that $Lz = f$ weakly. But then

$$\frac{1}{t}L(w-z) = \frac{1}{t}(f + t\hat{f} - f) = \hat{f} \text{ weakly},$$

contradicts the properties of $\hat{f}$. Since $f + t\hat{f} \to f$ as $t \to 0^+$, f cannot be an interior point. Thus, $X_s(\Omega)$ has an empty interior.

Therefore, $X(\Omega)$ is of first category. Now, set

$$Y_i := \big\{f \in C_c^\infty(\mathbb{R}) \mid Lu = f \text{ has a weak solution } u \in H^{-s}(\Omega_i) \text{ for some } s \in \mathbb{N}\big\},$$

where $\{\Omega_i\}_{i=1}^\infty$ is a collection of nested open sets with smooth boundary that form a countable basis for open sets containing the origin.

By the above each Y_i is of first category. By the Baire category theorem, $C_c^\infty(\mathbb{R}^n) \neq \bigcup_{i=1}^\infty Y_i$, thus there exists an $f \in C_c^\infty(\mathbb{R}^n) \setminus \bigcup_{i=1}^\infty Y_i$, and such f has the desired property.

□

11. Linear elliptic PDEs

In this section we consider the following boundary value problem for a unknown function u

$$\text{(BVP)} \begin{cases} Lu = f \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases}$$

where Ω is a bounded domain in $\mathbb{R}^n$ and L is given by

$$Lu = -\partial_i(a^{ij}\partial_j u) + b^i \partial_i u + cu$$

and $a^{ij} = aji$.

Remark 11.1. If the coefficients a^{ij} are sufficiently regular we can write

$$Lu = -a^{ij}\partial_i\partial_j u + \underbrace{(b^i - \partial_j a^{ij})}_{=\tilde{b}^i}\partial_i u + cu$$

the operator L is said to be in **divergence form** if written as in (BVP) and in non-divergence form if written as in this last expression. The negative sign in $-\partial_i(a^{ij}\partial_j u)$ is for convenience (we will consider integration by parts).

Definition 11.2. The operator L is **(uniformly) elliptic** in Ω if there exists a constant $\Lambda > 0$ such that

$$a^{ij}(x)\xi_i\xi_j \geq \Lambda|\xi|^2$$

a.e. in Ω for all $\xi \in \mathbb{R}^n$.

An obvious example is the case $a^{ij} = \delta^{ij}, b^i = 0, c = 0$, so $L = -\Delta$. The motivation for this definition is as follows. The definition says that the matrix (a^{ij}) is positive definite, with smallest eigenvalue $\geq \Lambda$. This implies that given $x_0 \in \Omega$, there exists a coordinate transformation $y^j = \psi^j(x^1, ..., x^n)$ in the neighborhood of x_0 such that Lu reads

$$(\tilde{L}\tilde{u})(y) = -\frac{\partial}{\partial y^i}\big(\tilde{a}^{ij}(y)\frac{\partial}{\partial y^j}\tilde{u}(y)\big) + \tilde{b}^i\frac{\partial}{\partial \tilde{y}^i}\tilde{u}(y) + \tilde{c}(y)\tilde{u}(y),$$

where $\tilde{a}^{ij}(y_0) = \delta^{ij}, y_0 = \psi(x_0)$. So an elliptic operator is locally comparable to the Laplacian. We will see that *many of the basic properties of Laplace's equation remain valid for elliptic operators.* Note that if a^{ij} are the components of a (inverse) Riemannian metric, the above change of coordinates is realized by normal coordinates.

Definition 11.3. Let L be an elliptic operator and assume that $a^{ij}, b^i, c \in L^\infty(\Omega)$. The **bilinear form**

$$B : H_0'(\Omega) \times H_0'(\Omega) \to \mathbb{R}$$

associated with L is

$$B(u, v) := \int_\Omega (a^{ij}\partial_i u\partial_j v + b^i v\partial_i u + cvu)dx.$$

We say that $u \in H_0'(\Omega)$ is a **weak solution to (BVP)** if

$$B(u, v) = (f, v)_0$$

for all $v \in H_0'(\Omega)$.

The idea of weak solutions is that if u is u satisfies (BVP) pointwise, multiplying $Lu = f$ by $v \in H_0'(\Omega)$ and integrating by parts we get $B(u, v) = (f, v)_0$.

Remark 11.4. Because a weak solution u is in $H_0'(\Omega)$, it has zero trace on $\partial\Omega$ (when the trace is well-defined). Whenever talking about $u|_{\partial\Omega}$ it will always be meant in the trace sense.

Definition 11.5. We say that u is a **strong solution** to $Lu = f$ in Ω if u is twice weakly differentiable and satisfies $Lu = f$ a.e. in Ω. u is a strong solution to (BVP) if it is a strong solution to $Lu = f$ such that $u|_{\partial\Omega} = 0$.

Observe that if u is a weak solution that is sufficiently regular, then we can integrate by parts and obtain that it is a strong solution.

11.1. **Existence of weak solution.** *Strategy to solve (BVP):*

- Find (unique) weak solutions: easier because weak solutions are more general.
- Prove regularity: show that the weak solution is in fact sufficiently differentiable, so it is a strong solution.

We will need the following theorems from functional analysis.

Theorem 11.6. ***(Lax-Milgram theorem).*** *Let H be a real Hilbert space and let $B : H \times H \to \mathbb{R}$ be a bilinear form that is bounded, i.e.,*

$$|B(x, y)| \leq h\|x\|\|y\|$$

for some constant $h > 0$ and all $x, y \in H$, and coercive, i.e.,

$$l\|x\|^2 \leq B(x, y)$$

for some constant $l > 0$ and all $x, y \in H$. Let $f : H \to \mathbb{R}$ be a bounded linear functional. Then, there exists a unique $z \in H$ such that

$$B(z, x) = \langle f, x\rangle$$

for all $x \in H$, where $\langle\, , \rangle$ is the pairing between H and H'.

Definition 11.7. Let X and Y be real Banach spaces. A bounded linear map $\kappa : X \to Y$ is called **compact** if given a bounded sequence $\{x\}_{j=1}^{\infty} \subset X$, the subsequence $\{\kappa x\}_{j=1}^{\infty} \subset Y$ is pre-compact in Y, i.e., $\{\kappa x\}_{j=1}^{\infty}$ has a convergent subsequence.

Theorem 11.8. *Let H be a real Hilbert space. If $\kappa : H \to H$ is compact, so is its adjoint $\kappa^* : H \to H$.*

Theorem 11.9. ***(Fredholm alternative).*** *Let H be a real Hilbert space and $\kappa : H \to H$ a compact operator. Then*

(i) ker$(I - \kappa)$ is finite dimensional.
(ii) range$(I - \kappa)$ is closed.
(iii) range$(I - \kappa)$ = ker$(I - \kappa^*)^{\perp}$.
(iv) ker$(I - \kappa) = \{0\}$ if and only if range$(I - \kappa) = H$.
(v) dim ker$(I - \kappa)$ = dim$(I - \kappa^*)$.

(I is the identity operator.)

Theorem 11.10. *There exists constants $k, l > 0$ and $m \geq 0$ such that*

$$|B(u, v)| \leq k\|u\|_{H_0'(\Omega)}\|v\|_{H_0'(\Omega)}$$

and

$$l\|u\|^2_{H_0'(\Omega)} \leq B(u, u) + m\|u\|^2_{L^2(\Omega)}$$

for all $u, v \in H_0'(\Omega)$.

Proof.

$$\begin{aligned} |B(u, v)| \leq \sum_{i,j=1}^{n} \|a^{ij}\|_{L^\infty(\Omega)} \int_\Omega |\nabla u||\nabla v| \\ + \sum_{i=1}^{\infty} \|b^i\|_{L^\infty(\Omega)} \int_\Omega |\nabla u||v| + \|c\|_{L^\infty(\Omega)} \int_\Omega |u||v| \\ \leq C\|u\|_{H_0'(\Omega)}\|v\|_{H_0'(\Omega)} \end{aligned}$$

The ellipticity of L gives

$$\Lambda \int_\Omega |\nabla u|^2 \leq \int_\Omega a^{ij}\partial_i u \partial_j u = B(u,u) - \int_\Omega (b^i \partial_i u u + cu^2)$$
$$\leq B(u,u) + \underbrace{\sum_{i=1}^{\infty} \|b^i\|_{L^\infty(\Omega)} \int_\Omega \underbrace{|\nabla u||u|}_{\leq \epsilon|\nabla u|^2 + \frac{1}{4\epsilon}u^2}}_{\leq \frac{\Lambda}{2}\int_\Omega |\nabla u|^2 + C\int_\Omega u^2 \text{ if } \epsilon \text{ is sufficiently small.}} + \|c\|_{L^\infty(\Omega)} \int_\Omega u^2$$
$$\implies \frac{\Lambda}{2}\int_\Omega |\nabla u|^2 \leq B(u,u) + C\int_\Omega u^2.$$

Because $u \in H_0'(\Omega)$, we have (Poincaré inequality)

$$\|u\|_{L^2(\Omega)} \leq C\|\nabla u\|_{L^2(\Omega)},$$

which gives the result. □

Theorem 11.11. *There exists a $m \geq 0$ such that, for each $\mu \geq m$, and each $f \in L^2(\Omega)$, there exists a unique weak solution $u \in H_0'(\Omega)$ to*

$$\begin{cases} Lu + \mu u = f \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega. \end{cases}$$

Proof. Take $m \geq 0$ from the previous theorem. The bilinear form

$$B_\mu(u,v) := B(u,v) + \mu(u,v)$$

corresponds to the operator $L_\mu u := Lu + \mu u$. B_μ then satisfies the assumptions of Lax-Milgram since $\mu(u,u) \geq l(u,u)$. Given $f \in L^2(\Omega)$, set

$$\langle f, v\rangle := (f,v)_0,$$

$v \in H_0'(\Omega)$, which is a bounded linear functional, so by Lax-Milgram there exists a unique $u \in H_0'(\Omega)$ such that

$$B_\mu(u,v) = (f,v)_0$$

for all $v \in H_0'(\Omega)$.

□

Definition 11.12. The **formal adjoint** to L is the operator L^* defined by

$$L^* v = -\partial_i(a^{ij}\partial_j v) - b^i\partial_i v + (c - \partial_i b^i)v$$

if $b \in C'(\bar{\Omega})$. The **adjoint bilinear form** $B^* : H_0'(\Omega) \times H_0'(\Omega) \to \mathbb{R}$ is defined by

$$B^*(v,u) = B(u,v).$$

We say that $v \in H_0'(\Omega)$ is a **weak solution to the adjoint problem**

$$\begin{cases} L^* v = f \text{ in } \Omega, \\ v = 0 \text{ on } \partial\Omega, \end{cases}$$

if $B^*(v,u) = (f,u)_0$ for all $u \in H_0'(\Omega)$.

The above definition is again inspired by integration by parts:

$$\begin{aligned}\int_\Omega vLu &= -\int_\Omega \partial_i(a^{ij}\partial_j u)v + \int_\Omega b^i\partial_i uv + \int_\Omega cuv \\ &= \int_\Omega a^{ij}\partial_j u\partial_i v + \int_\Omega b^i\partial_i uv + \int_\Omega cuv = B(u,v) \\ &= -\int_\Omega \partial_j(a^{ij}\partial_i v)u - \int_\Omega (b^i\partial_i vu + \partial_i b^i uv) + \int_\Omega cuv \\ &= \int_\Omega uL^*v.\end{aligned}$$

On the other hand, the bilinear form $B^*(v,u)$ such that

$$\int_\Omega L^*vu = B^*(v,u),$$

so $B^*(v,u) = B(u,v)$.

The next theorem characterizes the solvability of (BVP).

Theorem 11.13. ***(Fredholm alternative for elliptic operators).*** *Exactly one of the following statements holds:*

(i) For each $f \in L^2(\Omega)$ there exists a unique weak solution u to

$$\textit{(BVP)}\begin{cases} Lu = f \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases}$$

or else

(ii) There exists a weak solution $u \neq 0$ to

$$\textit{(BVP - H)}\begin{cases} Lu = 0 \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases}$$

Furthermore, if (ii) holds, then the dimension of the subspace $N \subset H_0'(\Omega)$ of weak solutions to (BVP-H) is finite and equals the dimension of the subspace $N^* \subset H_0'(\Omega)$ of weak solutions to

$$\textit{(BVP - H)}^*\begin{cases} L^*v = 0 \text{ in } \Omega, \\ v = 0 \text{ on } \partial\Omega, \end{cases}$$

Finally, (BVP) admits a weak solution iff $(f,v)_0 = 0$ for all $v \in N^*$.

Remark 11.14. The theorem says, loosely speaking, that we can solve (BVP) iff f is orthogonal to the kernel of the adjoint operator. Compare with the similar statement in linear algebra.

Proof. Let $m \geq 0$ be the constant from the previous theorem and set

$$L_m u := Lu + mu,$$

with corresponding bilinear form

$$B_m(u,v) := B(u,v) + m(u,v)_0.$$

Then, given $g \in L^2(\Omega)$ there exists a unique $u \in H_0'(\Omega)$ such that $B_m(u,v) = (g,v)_0$ for all $v \in H_0'(\Omega)$. This defines a linear map $L_m^{-1} : L^2(\Omega) \to H_0'(\Omega), g \mapsto u$.

$$u \in H_0'(\Omega) \text{ is a weak solution to (BVP) iff } B(u,v) = (f,v)_0$$

which is equivalent to

$$B_m(u,v) = B(u,v) + m(u,v)_0 = (f + mu, v)_0$$

for all $v \in H_0'(\Omega)$. But this means that $u = L_m^{-1}(g)$ with $g = f + mu$, i.e.,

$$u = L_m^{-1}(f + mu).$$

We can write this equation as

$$u - \underbrace{mL_m^{-1}u}_{=:\kappa u} = \underbrace{L_m^{-1}f}_{=:h}.$$

Let us show that κ defines a compact operator $\kappa : L^2(\Omega) \to L^2(\Omega)$. We have, for any $g \in L^2(\Omega)$,

$$m\|u\|^2_{H_0'(\Omega)} \le B_m(u,u) \underset{\text{taking } v=u}{=} (g,u)_0 \le \|g\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)} \le \|g\|_{L^2(\Omega)}\|u\|_{H_0'(\Omega)}$$
$$\implies m\|u\|_{H_0'(\Omega)} \le \|g\|_{L^2(\Omega)}.$$

But $u = L_m^{-1}g = \frac{1}{m}\kappa g$ (if $m = 0$ then $L_m = L$ and there is nothing left to prove). Thus, $\kappa : L^2(\Omega) \to H_0'(\Omega)$ is bounded. Since $H_0'(\Omega) \hookrightarrow L^2(\Omega)$ compactly, $\kappa : L^2(\Omega) \to L^2(\Omega)$ is compact. Thus, by the Fredholm alternative, $\ker(I - \kappa) = \{0\}$ if and only if $\text{range}(I - \kappa) = H$. Thus, either $u - \kappa u = h$ has a unique solution $u \in L^2(\Omega)$ for each $h \in L^2(\Omega)$ or else $u - \kappa u = 0$ has a non-zero solution $u \in L^2(\Omega)$. In the former case, taking $h = L_m^{-1}f$, we have $h \in H_0'(\Omega)$ thus by the foregoing u is in fact in $H_0'(\Omega)$ and is a weak solution to (BVP). In the latter case, necessarily $m > 0$, and the dimension of the space N of solution to

$$u - \kappa u = 0$$

is finite and equals the dimension of space N^* of solution, to

$$v - \kappa^* v = 0.$$

Let us show that $u - \kappa u = 0$ iff u is a weak solution to (BVP-H) and $v - \kappa^* v = 0$ iff v is a weak solution to (BVP-H)*.

Observe that:

$$u - \kappa u = 0 \Leftrightarrow \frac{1}{m}u = L_m^{-1}u \overset{\substack{L_m^{-1}g=u \text{ means} \\ B_m(u,v)=(g,v)_0 \text{ for all} \\ v\in H_0'(\Omega) \text{ with } g=\frac{1}{m}u}}{\Longleftrightarrow} B_m\left(\frac{1}{m}u, v\right) = (u,v)_0 \Leftrightarrow$$
$$B(u,v) + m(u,v)_0 = (mu,v)_0 \Leftrightarrow B(u,v) = 0$$

for all $v \in H_0'(\Omega)$, i.e., $u \in H_0'(\Omega)$ is a weak solution to (BVP-H). ($u \in H_0'(\Omega)$ because $u = mL_m^{-1}u$, and $L_m^{-1} : L^2(\Omega) \to H_0'(\Omega) \subset L^2(\Omega)$).

For (BVP-H)*, notice that the formal adjoint L_m^* is $L_m^* u = L^* u + mu$, with corresponding bilinear form

$$B_m^*(v,u) = B_m(u,v).$$

Thus we also obtain a bounded operator

$$m(L_m^*)^{-1} = \tilde{\kappa} : L^2(\Omega) \to H_0'(\Omega)$$

that is compact into $L^2(\Omega)$. Then, for any $u, v \in L^2(\Omega), \kappa u, \tilde{\kappa} v \in H_0'(\Omega)$ and

$$B_m(\kappa u, \tilde{\kappa} v) = mB_m(L_m^{-1}u, \tilde{\kappa} v) = m(u, \tilde{\kappa} v)_0$$
$$\to L_m^{-1}u = w \Leftrightarrow B_m(w,z) = (u,z)_0 \quad \forall z \in H_0'(\Omega)$$

$$B_m(\kappa u, \tilde{\kappa} v) = B_m^*(\tilde{\kappa} v, \kappa u) = mB_m^*((L_M^*)^{-1}v, \kappa u) = m(v, \kappa u)_0$$
$$\to (L_m^*)^{-1}v = z \Leftrightarrow B_m^*(z,w) = (v,w)_0 \quad \forall w \in H_0'(\Omega)$$

Thus

$$(u, \tilde{\kappa} v)_0 = (v, \kappa u)_0 = (\kappa^* v, u)_0$$

for all $u, v \in L^2(\Omega)$. Thus $\kappa^* = m(L_m^*)^{-1}$ and the same argument used for $u - \kappa u = 0$ shows that $v - \kappa v^* = 0$ iff v is a weak solution to (BVP-H)*.

For the last statement of the theorem, note that since $u - \kappa u = h$ has a solution iff $h \underset{L^2}{\perp} N^*$, with $h = L_m^{-1} f$ we have

$$(h, v)_0 = \frac{1}{m}(\kappa f, v)_0 = \frac{1}{m}(f, \kappa^* v) = \frac{1}{m}(f, v),$$

$v \in N^*$.

□

Theorem 11.15. *There exists at most a countable set $\Sigma \subset \mathbb{R}$ such that*

$$\begin{cases} Lu = \lambda u + f \ \text{ in } \Omega \\ \ \ u = 0 \ \text{ on } \partial\Omega \end{cases}$$

has a unique weak solution for each $f \in L^2(\Omega)$ if and only if $\lambda \notin \Sigma$. If Σ is infinite, then $\Sigma = \{\lambda_j\}_{j=1}^{\infty}$, and λ_j is a non-decreasing sequence with $\lambda_j \to \infty$ as $j \to \infty$.

Proof. Considering again, L_m^{-1}, this follows from properties of eigenvalues of compact operators.

□

Definition 11.16. Σ is called the spectrum of L and the vales $\lambda \in \Sigma$ the eigenvalues of L.

Corollary 11.17. *If $\lambda \notin \Sigma$, then there exists a constant $C > 0$ such that*

$$\|u\|_{L^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}$$

for $u \in H_0'(\Omega)$ the unique weak solution to

$$\begin{cases} Lu = \lambda u + f \ \text{ in } \Omega \\ \ \ u = 0 \ \text{ on } \partial\Omega \end{cases}$$

Proof. If not, there exist $\{f\}_{j=1}^{\infty} \subset L^2(\Omega)$, $\{u_j\}_{j=1}^{\infty} \subset H_0'(\Omega)$ weak solutions such that $\|u_j\|_{L^2(\Omega)} > j\|f\|_{L^2(\Omega)}$. Replacing u_j, f_j by $\frac{u_j}{\|u_j\|_{L^2(\Omega)}}, \frac{f_j}{\|u_j\|_{L^2}}$, we can assume that $\|u\|_{L^2(\Omega)} = 1$, then $f_j \to 0$ in $L^2(\Omega)$. Since

$$m\|u\|^2_{H_0'(\Omega)} \leq \underbrace{B(u.u)}_{(u,f)_0} + l\|u\|^2_{L^2(\Omega)},$$

$\|u\|_{H_0'(\Omega)} \leq C$ so $u_j \to u$ weakly in $H_0'(\Omega)$ and $u_j \to u$ in $L^2(\Omega)$. It follows that u is a weak solution to

$$\begin{cases} Lu = \lambda u \text{ in } \Omega \\ \ \ u = 0 \text{ on } \partial\Omega \end{cases}$$

Since $\lambda \notin \Sigma$, $u = 0$, but $\|u\|_{L^2(\Omega)} = 1$. □

Remark 11.18. The constant above $\to \infty$ as $\lambda \to \Sigma$.

11.2. **Elliptic regularity.**

Theorem 11.19. ***(interior regularity).*** *Let $a^{ij} \in C^1(\Omega)$ $(b^i, c \in L^\infty(\Omega))$, $f \in L^2(\Omega)$ and $u \in H_0'(\Omega)$ be a weak solution to BVP. Then $u \in H^2_{loc}(\Omega), Lu = f$ a.e. in Ω, and for each $\tilde{\Omega} \subset\subset \Omega$, there exists a constant $C = C(\tilde{\Omega}, a^{ij}, b^i, c)$ such that*

$$\|u\|_{H^2(\tilde{\Omega})} \leq C\big(\|f\|_{L^2} + \|u\|_{L^2}\big)$$

Proof. Let Ω' be such that

$$\tilde{\Omega} \subset\subset \Omega' \subset\subset \Omega$$

and $0 \le \psi \le 1$ be a $C_c^\infty(\Omega')$ such that $\psi = 1$ in $\tilde{\Omega}$. Since $B(u,v) = (f,v)_0$ for all $v \in H_0'(\Omega)$, we have

$$\int_\Omega a^{ij}\partial_i u \partial_j v = \int_\Omega \tilde{f} v, \tag{11.1}$$

$\tilde{f} = f - b^i\partial_i u - cu$. Let

$$D_l^h u(x) := \frac{u(x+he_l) - u(x)}{h}$$

be the difference quotient of u, assume $|h|$ small and set

$$v := -D_l^{-h}(\psi^2 D_l^h u).$$

Note that $v \in H_0'(\Omega)$. We also note the formulas

$$\int_\Omega z D_l^{-h} w = -\int_\Omega w D_l^h z$$

(integration by parts), z compactly supported, h small, and

$$D_l^h(zw) = z^h D_l^h w + w D_l^h z$$

(product rule) where $z^h(x) := z(x+he_l)$. Then

$$\begin{aligned}
LHS(11.1) &= -\int_\Omega a^{ij}\partial_i u \partial_j (D_l^{-h}(\psi^2 D_l^h h)) \\
&= \int_\Omega D_l^h(a^{ij}\partial_j u)\partial_j(\psi^2 D_l^h h) \\
&= \int_\Omega \left[(a^{ij})^h D_l^h \partial_j u \partial_j(\psi^2 D_l^h u) + D_l^h a^{ij}\partial_i u \partial_j(\psi^2 D_l^h u)\right] \\
&= \int_\Omega (a^i j)^h D_l^h \partial_j u D_l^h \partial_j u \psi^2 \\
&+ \int_\Omega \left[2\psi\partial_j u (a^{ij})^h D_l^h \partial_j u D_l^h u + D_l^h a^{ij}\partial_i u D_l^h \partial_j u \psi^2 \right. \\
&\quad \left. + 2\psi\partial_j\psi D_l^h a^{ij}\partial_i u D_l^h u\right] \;=\; I_1 + I_2.
\end{aligned}$$

By ellipticity (with $\xi = \psi D_l^h \nabla u$)

$$\Lambda \int_\Omega \psi^2 \left|D_l^h \nabla u\right|^2 \le I_1.$$

For I_2,

$$\begin{aligned}
|I_2| &\le C\int_{\Omega'} \Big(\underbrace{\psi|D_l^h\nabla u||D_l^h u|}_{\le \epsilon\psi^2|D_l^h\nabla u|^2 + \frac{C}{\epsilon}|D_l^h u|^2} + \psi^2 \underbrace{|\nabla u||D_l^h \nabla u|\Big)}_{\le \frac{C}{\epsilon}|\nabla u|^2 + |D_l^h\nabla u|^2} \\
&\le \epsilon \int_\Omega \psi^2 |D_l^h \nabla u|^2 + \frac{C}{\epsilon}\int_{\Omega'} \left(|D_l^h u|^2 + |\nabla u|^2\right).
\end{aligned}$$

We will show in a lemma below that

$$\int_{\Omega'} |D_l^h u|^2 \le C\int_\Omega |\nabla u|^2, \text{ thus}$$

$$|I_2| \le \epsilon \int_\Omega \psi^2 |D_l^h \nabla u|^2 + C\int_\Omega |\nabla u|^2,$$

so

$$LHS(11.1) = I_1 + I_2 \geq \frac{\Lambda}{2}\int_\Omega \psi^2 |D_l^h \nabla u|^2 - C\int_\Omega |\nabla u|^2$$

Next,

$$\begin{aligned} RHS(11.1) &= \int_\Omega \tilde{f} v \leq C\int_\Omega \big(|f| + |\nabla u| + |u|\big)|v| \\ &\leq C\int_\Omega \big(f^2 + u^2 + |\nabla u|^2\big) + \epsilon\int_\Omega v^2. \end{aligned}$$

By the Lemma below

$$\begin{aligned} \int_\Omega v^2 &\leq C\int_\Omega |\nabla(\psi^2 D_l^h u)|^2 \\ &\leq C\int_{\Omega'} |D_l^h u|^2 + C\int_{\Omega'} \psi^2 |D_l^h \nabla u|^2, \text{ so} \\ RHS(11.1) &\leq \epsilon\int_\Omega \psi^2 |D_l^h \nabla u|^2 + C\int_\Omega \big(f^2 + u^2 + |\nabla u|^2\big). \end{aligned}$$

Therefore

$$\int_{\tilde{\Omega}} |D_l^h \nabla u|^2 \leq \int_\Omega \psi^2 |D_l^h \nabla u|^2 \leq C\int_\Omega \big(f^2 + u^2 + |\nabla u|^2\big).$$

By the lemma below $\nabla u \in H'_{\text{loc}}(\Omega)$ so $u \in H^2_{\text{loc}}(\Omega)$ and

$$\|u\|_{H^2(\tilde{\Omega})} \leq C\big(\|u\|_{H'(\Omega)} + \|f\|_{L^2(\Omega)}\big).$$

Next, observe that if we repeat the argument in a set $\tilde{\Omega} \subset \Omega' \subset \Omega$, we in fact have

$$\|u\|_{H^2(\tilde{\Omega})} \leq C\big(\|u\|_{H'(\Omega')} + \|f\|_{L^2(\Omega)}\big).$$

In (11.1) choose $v = \psi^2 u$, $0 \leq \psi \leq 1$, $\psi = 1$ in Ω', $\psi C_c^\infty(\Omega)$. Then

$$\int_\Omega \psi^2 |\nabla u|^2 \leq C\int_\Omega \big(u^2 + f^2\big).$$

□

Lemma 11.20. *Let $u \in W^{1,p}(\Omega), 1 \leq p < \infty$. For each $\tilde{\Omega} \subset\subset \Omega$,*

$$\|D^h u\|_{L^p(\tilde{\Omega})} \leq C\|\nabla u\|_{L^p(\Omega)},$$

$D^h u = (D_1^h u, ..., D_n^h u)$, $0 < |h| < \frac{1}{2} dist(\tilde{\Omega}, \partial\Omega)$.

If $1 < p < \infty$ and $u \in L^p(\tilde{\Omega})$ satisfies

$$\|D^h u\|_{L^p(\tilde{\Omega})} \leq C,$$

$0 < |h| < \frac{1}{2} dist(\tilde{\Omega}, \partial\Omega)$, *then* $u \in W^{1,p}(\tilde{\Omega})$ *and* $\|\nabla u\|_{L^p(\tilde{\Omega})} \leq C$.

Proof. Assume first u smooth. Then

$$u(x + he_l) - u(x) = \int_0^1 \frac{d}{dt} u(x + the_l)dt = \int_0^1 \nabla u(x + the_l) \cdot he_l dt$$

$$\left|\frac{u(x + he_l) - u(x)}{h}\right| \leq \int_0^1 \big|\nabla u(x + te_l)\big| dt$$

Set

$$\int_{\tilde{\Omega}} |D^h u|^p dx$$

Then, using Jensen's inequality,

$$\begin{aligned}\int_{\tilde{\Omega}} |D^h u|^p dx &\leq C \sum_{l=1}^n \int_{\tilde{\Omega}} \int_0^1 |\nabla u(x + te_l)|^p dt dx \\ &\leq C \int_0^1 \int_{\tilde{\Omega}} |\nabla u(x + te_l)|^p dx dt \\ &\leq C \|\nabla u\|^p_{L^p(\Omega)}.\end{aligned}$$

Assume now $\|D^h u\|_{L^p(\tilde{\Omega})} \leq C, 1 < p < \infty$. Then

$$\sup_h \|D^{-h} u\|_{L^p(\tilde{\Omega})} < \infty,$$

so there exists a $v_l \in L^p(\tilde{\Omega})$ such that $D_l^{-h} u \to v_l$ weakly in $L^p(\tilde{\Omega})$ as $h \to 0$, so

$$\begin{aligned}\int_{\tilde{\Omega}} u \partial_l \psi = \int_{\Omega} u \partial_l \psi &= \lim_{h \to 0} \int_{\Omega} u D_l^h \psi = - \lim_{h \to 0} \int_{\tilde{\Omega}} D_l^{-h} u \psi \\ &= - \int_{\tilde{\Omega}} v_l \psi = - \int_{\Omega} v_l \psi,\end{aligned}$$

so $v_l = \partial_l u$ in the weak sense, $\nabla u \in L^p(\Omega)$.

□

Theorem 11.21. ***(higher elliptic regularity).*** *Let the coefficients $a^{ij}, b^i, c \in C^{k+1}(\Omega)$ and $f \in H^k(\Omega), k \geq 0$ an integer. If $u \in H_0'(\Omega)$ is a weak solution to (BVP), then $u \in H^{k+2}_{loc}(\Omega)$ and for each $\tilde{\Omega} \subset\subset \Omega$*

$$\|u\|_{H^{k+2}(\tilde{\Omega})} \leq C\big(\|f\|_{H^k(\Omega)} + \|u\|_{L^2(\Omega)}\big).$$

In particular, if $a^{ij}, b^i, c, f \in C^\infty(\bar{\Omega})$, then $u \in C^\infty(\Omega)$.

Proof. This is proven by induction in k, with $k = 0$ done above.

□

Theorem 11.22. ***(boundary regularity).*** *Assume that $a^{ij} \in C^1(\bar{\Omega}), b^i, c \in L^\infty(\Omega), f \in L^2(\Omega)$, and $\partial\Omega$ is C^2. Let $u \in H_0'(\Omega)$ be a weak solution to (BVP). Then $u \in H^2(\Omega), Lu = f$ a.e. in Ω, and there exists a constant $C = C(\Omega, a^{ij}, b^i, c)$ such that*

$$\|u\|_{H^2(\Omega)} \leq C\big(\|f\|_{L^2} + \|u\|_{L^2}\big).$$

Proof. Consider first the case

$$\Omega = B_1(0) \cap \mathbb{R}^n_+.$$

Set $\tilde{\Omega} := B_{\frac{1}{2}}(0) \cap \mathbb{R}^n_+$ and let $\psi \in C_c^\infty(B_1(0))$ be such that $0 \leq \psi \leq 1$, $\psi = 1$ in $B_{\frac{1}{2}}(0)$. Note that $\psi = 1$ in $\tilde{\Omega}$.

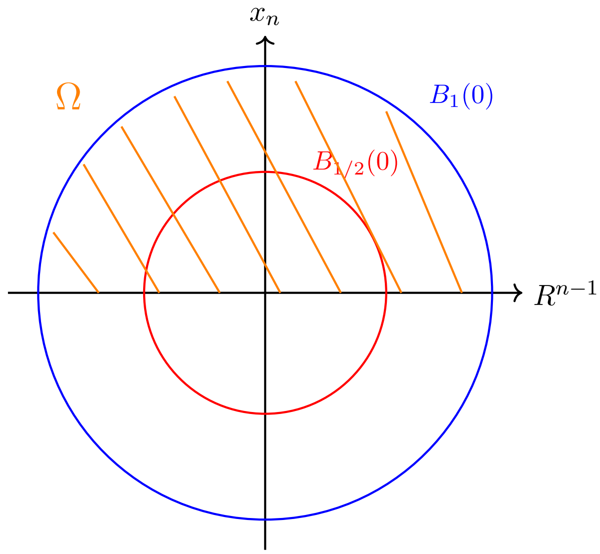


FIGURE 30. $\Omega = B_1(0) \cap \mathbb{R}^n_+$

We have

$$\int_\Omega a^{ij} \partial_i u \partial_j v = \int_\Omega \tilde{f} v$$

for all $v \in H_0'(\Omega)$, where $\tilde{f} := f - b^i \partial_i u - cu$.

For h small and $l \in \{1, ..., x^{n-1}\}$, put

$$v := -D_l^{-h}(\psi^2 D_l^h u).$$

We have

$$\begin{aligned} v(x) &= -D_l^{-h}\left(\psi^2 \frac{u(x+he_k) - u(x)}{h}\right) \\ &= -\frac{1}{h^2}\left[\psi^2(x - he_l)\big(u(x) - u(x - he_l)\big) - \psi^2(x)\big(u(x + he_l) - u(x)\big)\right] \end{aligned}$$

The RHS has a weak derivative for $x \in \Omega$ (recall that $1 \le l \le n-1$) and $u = 0$ on $\{x^n = 0\}$ (in the trace sense), thus $v \in H_0'(\Omega)$. We thus repeat the proof given in the first regularity to conclude

$$\partial_l u \in H'(\tilde{\Omega}), l = 1, ..., n-1,$$

and

$$\|\partial_l \nabla u\|_{L^2(\tilde{\Omega})} \le C\big(\|f\|_{L^2(\Omega)} + \|u\|_{H'(\Omega)}\big), l = 1, ..., n-1.$$

Since we already know $u \in H^2_{\text{loc}}(\Omega)$, we have $Lu = f$ a.e. in Ω, thus, since $a^{ij} \in C^1(\bar{\Omega})$,

$$-a^{ij}\partial_i \partial_j u + (b^i - \partial_j a^{ij})\partial_i u + cu = f$$

so

$$a^{nn}\partial_n^2 u = -\sum_{\substack{i,j=1 \\ i+j<2u}}^{n} a^{ij}\partial_i \partial_j u + \big(b^i - \partial_j a^{ij}\big)\partial_i u + cu - f.$$

By the above estimate for $\partial_l \nabla u$, $l = 1, ..., n-1$, we have $\partial_n^2 u \in L^2(\tilde{\Omega})$ so $u \in H^2(\tilde{\Omega})$, ($a^{nn} \ge C > 0$ by ellipticity, take $\xi = e_n$). As before we also obtain

$$\|u\|_{H^2(\tilde{\Omega})} \le C\big(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}\big).$$

Using the compactness of $\bar{\Omega}$ and local flattenings of the boundary, we obtain the general result, including the desired estimate in $H^2(\Omega)$.

□

Remark 11.23. Observe how we made specific use of the structure of the equation to estimate $\partial_n^2 u$.

Theorem 11.24. ***(higher boundary regularity)****. Assume that $a^{ij}, b^i, c \in C^{k+1}(\bar{\Omega}), f \in H^k(\Omega)$, and $\partial\Omega$ is C^{k+2}, $k \geq 0$ an integer. Let $u \in H_0'(\Omega)$ be a weak solution to (BVP). Then, $u \in H^{k+2}(\Omega)$, and we have the estimate*

$$\|u\|_{H^{k+2}(\Omega)} \leq C(\|f\|_{H^k(\Omega)} + \|u\|_{L^2(\Omega)}).$$

In particular, $u \in C^\infty(\bar{\Omega})$ if all the data is C^∞ (up to the boundary).

Proof. Again by induction.

□

Remark 11.25.

- The regularity theorems say, roughly, that u gains two derivatives in relation to f so u is "as regular as possible".
- If u is the unique weak solution, then an argument similar to the one used to prove $\|u\|_{L^2}(\Omega) \leq C\|f\|_{L^2(\Omega)}$ gives

$$\|u\|_{H^{k+2}(\Omega)} \leq C\|f\|_{H^k(\Omega)}$$

11.3. **Maximum principles.** We will now assume an elliptic operator of the form

$$Lu = a^{ij}\partial_i\partial_j u + b^i\partial_i u + cu$$

in a domain Ω, with $a^{ij}, b^i, c \in L^\infty(\Omega)$ and $a^{ij} \in C^0(\Omega)$. Ellipticity is as before,

$$a^{ij}\,\xi_i\,\xi_j \geq \Lambda|\xi|^2.$$

The basic intuition of maximum principles is that if $x_0 \in \Omega$ is a maximum for u over $\bar{\Omega}$ and $c = 0$, then $Lu(x_0) = a^{ij}\partial_i\partial_j u(x_0) \leq 0$ since $(\partial_i\partial_j u(x_0))$ is non-positive and $a^{ij}(x_0)$ is positive-definite by ellipticity. Thus, if $Lu > 0$, u cannot have an interior maximum.

Throughout, replacing u by $-u$ we obtain statements for minimum.

Theorem 11.26. ***(weak maximum principle).*** *Let $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$ satisfy*

$$Lu \geq 0 (\leq 0)$$

in a bounded domain Ω, and suppose that $c = 0$. Then the maximum (minimum) of u is achieved on $\partial\Omega$.

Proof. Since $a'' > 0$, we can choose $\alpha > 0$ large so that

$$Le^{\alpha x'} = (\alpha^2 a'' + \alpha b')e^{\alpha x'} > 0.$$

For any $\epsilon > 0$,

$$L\big(u + \epsilon e^{\alpha x'}\big) > 0,$$

thus $u + \epsilon e^{\alpha x'}$ achieves is maximum on $\partial\Omega$ by the above argument. So does u by taking $\epsilon \to 0$.

□

Corollary 11.27. *Consider the same assumptions as above but instead suppose $c \leq 0$. Then*

$$\sup_\Omega u \leq \sup_{\partial\Omega} u^+ (\inf_\Omega u \geq \inf_{\partial\Omega} u^-)$$

$u^+ = \max\{u^+, 0\}$*, $u^- = \min\{u, 0\}$. In particular,* $\sup_\Omega |u| = \sup_{\partial\Omega} |u|$ *if $Lu = 0$.*

Proof. Set

$$\Omega^+ = \{x \in \Omega \mid u(x) > 0\}.$$

If $Lu \geq 0$ then

$$L_+ u = a^{ij}\partial_i\partial_j u + b^i \partial_i u \geq -cu \geq 0 \text{ in } \Omega^+$$

since $-cu \geq 0$ in Ω^+. Then the maximum of u in $\bar{\Omega}^+$ is achieved on $\partial\Omega^+$. Since $u \leq 0$ in $\Omega \setminus \Omega^+$, and $u(x_0) = 0$ if $x_0 \in \partial\Omega^+ \cap \Omega$, we must have $\sup_\Omega u \leq \sup_{\partial\Omega} u^+$.

□

Corollary 11.28. *Under the same assumptions of the previous corollary, if $Lu \geq Lv$ in Ω and $u \leq v$ on $\partial\Omega$, then $u \leq v$ in Ω. In particular, $u = v$ if $Lu = Lv$ in Ω and $u = v$ on Ω.*

Remark 11.29. The assumption $c \leq 0$ cannot be relaxed, as there exist positive eigenvalues to the problem $\Delta u + \lambda u = 0$ in Ω, $u = 0$ on $\partial\Omega$.

Definition 11.30. A domain $\Omega \subset \mathbb{R}^n$ satisfies the **interior sphere condition** at $x_0 \in \partial\Omega$ if there exists a $x \in \Omega$ and a $r > 0$ such that $B_r(x) \subset \Omega$ and $x_0 \in \partial B_r(x)$.

Lemma 11.31. *Suppose that L satisfies $c = 0$, and $Lu \geq 0$ in Ω, where $u \in C^2(\Omega)$. Let $x_0 \in \partial\Omega$ and suppose that*

(i) u is continuous at x_0;

(ii) $u(x_0) > u(x)$ for all $x \in \Omega$;

(iii) Ω satisfies an interior sphere condition at x_0. Then, the outer normal derivative of u at x_0, if it exists, satisfies the strict inequality

$$\frac{\partial u}{\partial v}(x_0) > 0.$$

If $c \leq 0$, the conclusion holds provided that $u(x_0) \geq 0$. If $u(x_0) = 0$, the conclusion holds irrespectively of the sign of c.

Proof. Let $B_R(y) \subset \Omega$ be such that $x_0 \in \partial B_R(y)$, by the interior sphere condition. Fix $0 < \rho < R$, set

$$v(x) = e^{-\alpha r^2} - e^{-\alpha R^2},$$

$r = |x - y| > \rho$, $\alpha > 0$ to be chosen. For $c \leq 0$ we have

$$\begin{aligned} Lv(x) &= \big(a^{ij}\partial_i\partial_j + b^i\partial_i + c\big)v(x) \\ &= e^{-\alpha r^2}\big[4\alpha^2 a^{ij}(x_i - y_i)(x_j - y_j) - 2\alpha(a^{ii} + b^i(x_i - y_i)\big] + cv \\ &\geq e^{-\alpha r^2}\big[4\alpha^2\Lambda r^2 - 2\alpha(a^{ii} + |b|r) + c\big] \quad \big(|b| = |(b^1, ..., b^n)|\big) \\ &\geq e^{-\alpha r^2}\big[4\alpha^2\Lambda\rho^2 - 2\alpha(a^{ii} + |b|r) + c\big] \end{aligned}$$

since $v \leq e^{-\alpha r^2}$ and $c \leq 0$. Thus, taking α large enough,

$$Lv \geq 0 \text{ in } \Omega' = B_R(y) - B_\rho(y).$$

Since $u - u(x_0) < 0$ on $\partial B_\rho(y)$, there exists a $\epsilon > 0$ such that $u - u(x_0) + \epsilon v \leq 0$ on $\partial B_\rho(y)$. We also have $u - u(x_0) + \epsilon v \leq 0$ on $\partial B_R(y)$ since $v = 0$ there. Moreover,

$$\begin{aligned} L\big(u - u(x_0) + \epsilon v\big) &= \underbrace{Lu}_{\geq 0} - cu(x_0) + \epsilon \underbrace{Lv}_{\geq 0} \\ &\geq -cu(x_0) \geq 0 \text{ in } \Omega'. \end{aligned}$$

where the last inequality is valid for $c = 0$ or $c \leq 0$ with $u(x_0) \geq 0$. Since $u \in C^2(\Omega)$ and u is continuous at x_0, we have $u \in C^2(\Omega') \cap C^0(\bar{\Omega}')$. Thus, by a corollary of the weak maximum principle, ($Lw \geq Lz$ in $\Omega, w \leq z$ on $\partial\Omega \implies w \leq z$ in Ω) we have $u - u(x_0) + \epsilon v \leq 0$ in Ω'. Because the function $u - u(x_0) + \epsilon v$ vanishes at x_0, we conclude that its normal derivative at x_0 cannot be negative, so

$$\frac{\partial u}{\partial v}(x_0) \geq -\epsilon \frac{\partial v}{\partial v}(x_0).$$

But $-\frac{\partial v}{\partial v}(x_0) = -\frac{1}{R}\nabla v(x_0) \cdot x_0 = -v'(R) = 2\alpha R e^{-\alpha R^2} > 0$. For c arbitrary, if $u(x_0) = 0$ the above argument works with L replaced by $L - c^+$.

□

Theorem 11.32. ***(strong maximum principle).*** *Suppose that $u \in C^2(\Omega)$ satisfies $Lu \geq 0$ (≤ 0) in Ω and $c = 0$. If u achieves a maximum (minimum) in Ω, then u is constant. If $c \leq 0$, then u cannot achieve a non-negative maximum (non-positive minimum) in Ω unless it is constant.*

Proof. Suppose u achieves a maximum M in Ω. If u is not constant, $\Omega^- := \{x \in \Omega \mid u(x) < M\}$ is not empty, neither is $\partial\Omega^- \cap \Omega$. Thus, there exists a x_0 such that $\operatorname{dist}(x_0, \partial\Omega^-) < \operatorname{dist}(x_0, \partial\Omega)$. Let $B_r(x_0)$ be the largest ball centered at x_0 such that $B_r(x_0) \subset \Omega^-$. Then $u(y) = M$ for some $y \in \partial B_r(x_0)$ and $u < M$ in $B_r(x_0)$. By the above Lemma $\frac{\partial u}{\partial v}(y) > 0$, but $\nabla u(y) = 0$ since y is an interior maximum.

□

Remark 11.33. The proof also gives that if $c(x) < 0$ at some $x \in \Omega$ then the constant in theorem must be zero and if u vanishes at the interior maximum (minimum) then $u = 0$ regardless of the sign of c.

Theorem 11.34. *Let $Lu \geq f(= f)$ in a bounded domain Ω, $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$, and assume that $c \leq 0$. Then, there exists a constant $C > 0$ depending only on the diameter of Ω and on $\beta = \frac{\|(b^1,\ldots,b^n)\|}{\Lambda} L^\infty(\Omega)$, such that*

$$\sup_{\Omega} u\big(|u|\big) \leq \sup_{\partial\Omega} u^+\big(|u|\big) + c \sup_{\Omega} \frac{|f^-|}{\Lambda}\left(\frac{|f|}{\Lambda}\right).$$

($f^- = \inf\{f, 0\}$, $u^+ = \sup\{u, 0\}$).

Proof. Let Ω lie in the slab $0 < x^1 < d$, and set

$$L_0 = a^{ij}\partial_i\partial_j + b^i\partial_i.$$

Then, if $\alpha \geq \beta + 1$,

$$\begin{aligned}
L_0 e^{\alpha x'} &= \big(\alpha^2 a'' + \alpha b'\big)e^{\alpha x'} \geq \big(\alpha^2\Lambda + \alpha b'\big)e^{\alpha x'} \\
&\geq \big(\alpha^2\Lambda = \alpha|b'|\big)e^{\alpha x'} \geq \big(\alpha^2\Lambda - \alpha\|(b^1, \ldots, b^n)\|_{L^\infty(\Omega)}\big)c^{\alpha x'} \\
&= \big(\alpha^2\Lambda - \alpha\Lambda\beta\big)e^{\alpha x'} \geq \Lambda.
\end{aligned}$$

Set

$$v = \sup_{\partial\Omega} u^+ + (e^{\alpha d} - e^{\alpha x'}) \sup_{\Omega} \frac{|f^-|}{\Lambda} \geq 0$$

Then

$$\begin{aligned}
Lv = L_0 v + cv &= \underbrace{-(L_0 e^{\alpha x'})}_{\leq -\Lambda} \sup_{\Omega} \frac{|f^-|}{\Lambda} + \underbrace{cv}_{\leq 0} \\
&\leq -\sup_{\Omega} |f^-|, \text{ thus} \\
L(v - u) = Lv - Lu &\leq -\sup_{\Omega}|f^-| - f \leq 0.
\end{aligned}$$

We also have $v - u \geq 0$ on $\partial\Omega$. Thus, by one of the corollaries of the weak maximum principle ($Lw \geq Lz$ in Ω, $w \leq z$ on $\partial\Omega$, then $w \leq z$ in Ω). $u \leq v$ in Ω.

Thus,

$$u \le \sup_{\partial\Omega} u^+ + (e^{\alpha d} - e^{\alpha x'}) \sup_{\Omega} \frac{|f^-|}{\Lambda}$$
$$\le \sup_{\partial\Omega} u^+ + (e^{\alpha d} - 1) \sup_{\Omega} \frac{|f^-|}{\Lambda}.$$

□

Corollary 11.35. *Let $Lu = f$ in a bounded domain Ω, $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$. Let C be the constant of the previous theorem. Suppose that*

$$A = 1 - c \sup_{\Omega} \frac{c^+}{\Lambda} > 0.$$

Then

$$\sup_{\Omega} |u| \le \frac{1}{A}\left(\sup_{\partial\Omega} |u| + C \sup_{\Omega} \frac{|f|}{\Lambda} \right).$$

Proof. Write $Lu = (L_0 + c)u = f$ as

$$(L_0 + c^-)u = f - c^+ u =: \tilde{f}.$$

From the theorem,

$$\sup_{\Omega} |u| \le \sup_{\partial\Omega} |u| + C \sup_{\Omega} \frac{|\tilde{f}|}{\Lambda}$$
$$\le \sup_{\partial\Omega} |u| + C\left(\sup_{\Omega} \frac{|f|}{\Lambda} + \sup_{\Omega} |u| \sup_{\Omega} \frac{|u^+|}{\Lambda} \right),$$

so

$$\left(1 - C \sup_{\Omega} \frac{|c^+|}{\Lambda}\right) \sup |u| \le \sup_{\partial\Omega} |u| + C \sup_{\Omega} \frac{|f|}{\Lambda}.$$

□

Remark 11.36. Since we can take $C = e^{\alpha d - 1} \to 0$ as $d \to 0$, the above corollary implies uniqueness, hence solvability, of the Dirichlet problem on any sufficiently thin bounded domain.

12. Nonlinear elliptic equations

We will investigate the solvability of equations of the form

$$Lu + f(\cdot, u) = 0, \tag{12.1}$$

where L is an elliptic operator. When f is non-linear in u, this is a semi-linear elliptic equation. Our goal is to illustrate some techniques, thus we will consider special cases of (12.1) (but the ideas can be adapted to more general settings). We will also briefly consider more general nonlinear equations. The arguments we will employ in this section are soft, thus, it is instructive to consider a more general setting. Therefore, in this section we will take (M, g) to be a closed Riemannian manifold and let Δg be the Laplacian w.r.t. g, which in local coordinates reads $\Delta g = \frac{1}{\sqrt{\det g}} \partial_i \big(\sqrt{\det g}\, g^{ij} \partial_j\big)$. Students not familiar with geometry can take $M = \pi^n$ (n-dimensional torus) and $\Delta g = \Delta$. Here are the facts that need to be known for our analysis:

- We can define the spaces $W^{k,p}(M)$ and the embedding theorems go through.
- The Fredholm alternative remains valid.
- Elliptic regularity remains valid. In fact, we will use that elliptic regularity holds in $W^{k,p}(M)$, $1 < p < \infty$. Thus, $f \in W^{k,p}(M)$, $Lu = f \implies u \in W^{k+2,p}(M)$ if the coefficients of L are sufficient regular.

- The strong maximum principle remains valid.
- As a consequence of the above, $W^{k,p}(M) = L\big(W^{k+2,p}(M)\big) \bigoplus \ker L^*, 1 < p < \infty$, and when L is invertible $L^{-1} : W^{k,p}(M) \to W^{k+2,p}(M)$ is an isomorphism.

12.1. **The method of sub- and super-solutions.**

Theorem 12.1. *Consider in (M, g) the equation*

$$\Delta_g u + f(\cdot, u) = 0,$$

where $f \in C^\infty(M \times \mathbb{R})$. Suppose that there exist functions $\psi_-, \psi_+ \in C^2(M)$ such that $\psi_- \leq \psi_+$ and

$$\Delta_g \psi_- + f(\cdot, \psi_-) \geq 0,$$
$$\Delta_g \psi_+ + f(\cdot, \psi_+) \leq 0.$$

then there exists a smooth solution u to $\Delta_g u + f(\cdot, u) = 0$.

Remark 12.2. ψ_- and ψ_+ are called, respectively, sub- and super-solutions to the equation.

Proof. Let A be a constant such that

$$-A \leq \psi_- \leq \psi_+ \leq A$$

and choose a constant $c > 0$ large enough such that

$$F(x, t) = ct + f(x, t)$$

is increasing in $t \in [-A, A]$ for each $x \in M$. Set

$$Lu := -\Delta_g u + cu.$$

By the maximum principle, $\ker(L) = \{0\}$. Notice that if $u = \text{const}$ solves $Lu = 0$, then $u = 0$ (since $c > 0$). But if $-\Delta_g u + cu \leq 0$ then u cannot have a non-negative maximum unless it is constant, so $u \leq 0$; and if $-\Delta_g u + cu \geq 0$ then u cannot have a non-positive minimum, so $u \geq 0$. (Note that here we are applying the strong maximum principle to the equation multiplied by -1; max and min always achieved by compactness of M.)

(Alternatively, we can see uniqueness by:

$$-\Delta_g u + cu = 0 \underset{\text{by parts}}{\Longrightarrow} \int_M \big(|\nabla_g u|^2 + cu^2\big) d\,\text{vol}_g = 0 \implies u = 0.)$$

We also note that L is a positive operator, i.e., $Lu_1 \geq Lu_2 \implies u_1 \geq u_2$, since the maximum principle gives $L(u_1 - u_2) \geq 0 \implies u_1 - u_2$ cannot have a non-positive minimum, so $u_1 - u_2 \geq 0$.

Thus we have an isomorphism

$$L^{-1} : W^{k,p}(\Omega) \to W^{k+2,p}(\Omega), 1 < p < \infty.$$

Let $v \in C^2(M)$. Then $v \in W^{2,p}(M)$ for any p since M is compact. Thus $L^{-1}v \in W^{4,p}(M)$ for any $1 < p < \infty$. Taking p large enough so that $4 - \frac{n}{p} > 2$, by the Sobolev embedding $W^{4,p}(M) \hookrightarrow C^{2,\alpha}(M)$, we have that $L^{-1}v \in C^2(M)$.

Define inductively

$$\underline{\psi}_0 = \psi_-, \ \ \underline{\psi}_l = L^{-1}(F(\cdot, \ \underline{\psi}_{l-1}))$$
$$\bar{\psi}_0 = \psi_+, \ \ \bar{\psi}_l = L^{-1}(F(\cdot, \ \bar{\psi}_{l-1}))$$

Observe that

$$L\underline{\psi}_1 = F(\cdot, \psi_-) = c\psi_- + f(\cdot, \psi_-) \geq c\psi_- - \Delta_g \psi_- = L\psi_-$$

and

$$L\bar{\psi}_1 = F(\cdot, \psi_+) = c\psi_+ + f(\cdot, \psi_+) \le c\psi_+ - \Delta_g \psi_+ = L\psi_+,$$

thus

$$L\psi_- \le L\underline{\psi}_1 = F(\cdot, \psi_-) \underset{\substack{\text{since } F \text{ is increasing} \\ \text{in its second argument}}}{\le} F(\cdot, \psi_+) = L\bar{\psi}_1 \le L\psi_+$$

Hence $\psi_- \le \underline{\psi}_1 \le \bar{\psi}_1 \le \psi_+$ by the positivity of L. Repeating the argument,

$$\psi_- \le \underline{\psi}_{l-1} \le \underline{\psi}_l \le \bar{\psi}_l \le \bar{\psi}_{l-1} \le \psi_+ \text{ for every } l.$$

Thus, we have monotone bounded sequences $\{\underline{\psi}_l\}, \{\bar{\psi}_l\}$, thus they converge pointwise to limits $\underline{u}$ and $\bar{u}$, $\underline{u} \le \bar{u}$.

Since $|\underline{\psi}_l| \le C$, we have $\|F(\cdot, \underline{\psi}_l(\cdot))\|_{L^p(M)} \le C$ for any p . Because $L\underline{\psi}_l = F(\cdot, \underline{\psi}_{l-1})$, elliptic regularity gives

$$\|\underline{\psi}_l\|_{W^{2,p}(M)} \le C\big(\|F(\ , \underline{\psi}_{l-1})\|_{L^p(M)} + \|\underline{\psi}_l\|_{L^p(M)}\big) \le C.$$

By the compact embedding $W^{2,p}(M) \hookrightarrow C^{1,\alpha}(\Omega), 2 > \frac{n}{p}$, we have that $\{\underline{\psi}_l\} \subset C^{1,\alpha}(M)$ and converges, up to a subsequence in $C^{1,\alpha}(M)$. But $v \in C^{1,\alpha}(M)$ implies that for any p, we find $\|v\|_{W^{1,p}(M)} \le C\|v\|_{C^{1,\alpha}(M)}$. Hence

$$\|\underline{\psi}_l\|_{W^{3,p}(M)} \le C\big(\|F(\cdot, \underline{\psi}_{l-1})\|_{W^{1,p}(M)} + \|\underline{\psi}_l\|_{L^p(M)}\big) \le C$$

and by the compact embedding $W^{3,p}(M) \hookrightarrow C^{2,\alpha}(M)$ we obtain convergence in $C^2(M)$. We can thus pass to the limit in the equation to obtain

$$L\underline{\psi}_l = F(\cdot, \underline{\psi}_{l-1}) \implies L\underline{u} = F(\cdot, \underline{u}).$$

Similarly $L\bar{u} = F(\cdot, \bar{u})$. But

$$Lu = F(\cdot, u) \Leftrightarrow -\Delta_g u + cu = cu + f(\cdot, u) \Leftrightarrow \Delta_g u + f(\cdot, u) = 0.$$

Applying elliptic regularity inductively as above to $\Delta_g u + f(\cdot, u) = 0$ we conclude $u \in C^\infty(M)$. □

Remark 12.3. There might be many sub- and super-solutions. E.g., $\Delta_g u = f(\cdot, u) + \cos u, |f| \le 1$. Then $u(x) = 2m\pi$ and $u(x) = (2m-1)\pi$ and all super- and sub-solutions, respectively, so we find at least one solution u on each interval $(2m-1)\pi \le u \le 2m\pi$.

We will now give a proof of the easy case of the uniformization theorem.

Theorem 12.4. *Let (M, g) be a closed two-dimensional Riemannian manifold with the Euler characteristic $\chi(M) < 0$. Let $\tilde{\kappa} \le 0$ be a smooth function in M that is not identically zero. Then, there exists a metric $\tilde{g}$ conformed to g such that $\kappa(\tilde{g}) = \tilde{\kappa}$, where κ is the Gauss curvature. In particular, we can take $\tilde{\kappa} = -1$ and obtain the uniformization theorem in the negative case.*

Proof. Write $\tilde{g} = e^{2u} g$. Then the scalar curvature of $\tilde{g}$ and g are related by

$$\tilde{R} = e^{-2u}(R - 2\Delta_g u).$$

Since scalar = 2 Gauss, the Gauss curvatures are related by

$$\Delta u - \kappa + \tilde{\kappa} e^{2u} = 0.$$

Hence, we only need to find u solving this equation.

Let us find a super-solution ψ_+. We claim that we can find $v \in C^\infty(M)$ such that

$$\Delta_g v = \tilde{\kappa}_0 - \tilde{\kappa},$$

where $\tilde{\kappa}_0 = (\operatorname{vol}_g(M))^{-1}\int_M \tilde{\kappa} d\operatorname{vol}_g$. Then $\int_M(\tilde{\kappa}_0 - \kappa)d\operatorname{vol}_g = 0$, i.e., $\tilde{\kappa}_0 - \tilde{\kappa}$ is L^2 orthogonal to constants. But $\ker(\Delta_g) = \mathbb{R}$ since $\Delta_g u = 0 \implies \int_M |\nabla_g u|^2 d\operatorname{vol}_g = 0$. Thus, by Fredholm, we cam solve $\Delta_g v = \tilde{\kappa}_0 - \kappa$. Elliptic regularity gives $v \in C^\infty(M)$.

Set $\psi_+ = av + b,\ a, b \in \mathbb{R}$. Since $\tilde{\kappa}_0 < 0$ by our assumptions, we can choose a such that $a\tilde{\kappa} < \kappa(x)$ for all $x \in M$. Then, take b so large that $e^{2(av+b)} - a > 0$. Then

$$\Delta_g\psi_+ - \kappa + \tilde{\kappa}e^{2\psi_+} = a\Delta_g v - \kappa + \tilde{\kappa}e^{2\psi_+}$$
$$= a(\tilde{\kappa}_0 - \kappa) - \kappa + \tilde{\kappa}e^{2\psi_+} = a\tilde{\kappa}_0 - \kappa + \tilde{\kappa}\big(e^{2(av+b)} - a\big) < 0.$$

Next, we will find a sub-solution ψ_- such that $\psi_- \leq \psi_+$.

Let v solve

$$\Delta_g v = \kappa - \kappa_0,\ \kappa_0 = (\operatorname{vol}_g(M))^{-1}\int_M \kappa d\operatorname{vol}_g,$$

which can be found by the same arguments as above. Put $\psi_- = v - c$, where $c \in \mathbb{R}$ is large enough so that $\psi_- \leq \psi_+$. Then

$$\Delta\psi_- - \kappa + \tilde{\kappa}e^{2\psi_-} = \kappa - \kappa_0 + \kappa + \tilde{\kappa}e^{2v-2c}$$
$$= -\kappa_0 + \tilde{\kappa}e^{2v-2c}$$

Since $\kappa_0 < 0$, by $\chi(M) < 0$ and Gauss-Bonnet, we can choose c large such that RHS > 0.

□

Remark 12.5. It is possible to give a full proof of the uniformization theorem using PDE methods.

12.2. **Implicit function theorem methods.** We recall some notions of functional analysis.

Definition 12.6. Let X, Y be Banach spaces, $U \subset X$ an open set, and $f : U \to Y$. We say that f has a **Gateaux derivative** at $x \in U$ if

$$f'(x, y) := \frac{d}{dt}f(x + ty)\Big|_{t=0}$$

exists for every $y \in X$. We say that f has a **Fréchet derivative** at $x \in U$ if there exists a continuous linear map $Df(x) : X \to Y$ such that

$$f(x + y) = f(x) + Df(x)(y) + \sigma(|y|)$$

for every y such that $x + y \in U$, in which case one sees that $Df(x)$ is in fact defined for every $y \in X$. We say that f is **continuously differentiable** (or C^1) at x if the map

$$x \in U \mapsto Df(x) \in L(X, Y)$$

is continuous.

Theorem 12.7. *Let X, Y be Banach spaces, $U \subset X$ open, $f : U \to Y$. If f has a Gateaux derivative $f'(x, y)$ in U which is linear in y, and if the map $x \in U \mapsto f'(x, \cdot) \in L(X, Y)$ is continuous, then f is Fréchet differentiable in U and $Df(x)(y) = f'(x, y)$.*

Theorem 12.8. ***(implicit function theorem).*** *Let X, Y, Z be Banach spaces. Let $f : X \times Y \to Z$ be continuously differentiable. Suppose that $f(x_0, y_0) = 0$ and that $Df(x_0, y_0)(0, \cdot) : Y \to Z$ is a (Banach space) isomorphism. Then there exists a neighborhood $U \times V \ni (x_0, y_0)$ and a Fréchet differentiable map $g : U \to V$ such that $f(x, g(x)) = 0$ for all $x \in U$.*

Definition 12.9. Let $p : C^\infty(\Omega) \to C^\infty(\Omega)$ be a differential operator. Its **linearization** at $u \in C^\infty(\Omega)$ is the linear operator

$$L_u v = Lv = \frac{d}{dt}P(u + tv)\Big|_{t=0}.$$

This definition extends to P defined on $H^s(\Omega)$ etc.

As an application, let $\Omega \subset \mathbb{R}^3$ be a bounded set with smooth boundary. Let $h : \partial\Omega \to \mathbb{R}$, and consider the problem of extending h to Ω as a perturbation of the identity that is volume-preserving, i.e.,

$$\text{Jac}\,(\text{id} + \nabla u) = 1,$$

where u extends h and $\text{id} + \nabla u$ is the map

$$x \in \Omega \mapsto x + \nabla u(x) \in \mathbb{R}^3.$$

Expanding the Jacobian, we see that Jac $(\text{id} + \nabla u) = 1$ is equivalent to

$$\Delta u + \mathcal{N}(u) = 0,$$

where

$$\mathcal{N}(f) = f_{xx}f_{yy} + f_{xx}f_{zz} + f_{yy}f_{zz} - f_{xy}^2 - f_{xz}^2 - f_{yz}^2 + \det(D^2 f).$$

Thus, given h, we seek to solve

$$\begin{cases} \Delta u + \mathcal{N}(u) &= 0 \text{ in } \Omega, \\ u &= h \text{ on } \partial\Omega, \end{cases} \tag{12.2}$$

which is a fully nonlinear BVP. If $h = 0$, then $u = 0$ is a solution. Thus, we expect that a solution u is small if h is small. But for small u, the equation reads

$$\Delta u + \mathcal{O}(|D^2 u|^2) = 0$$

and since $|D^2u|^2 << |D^2 u|$ for u small, we have a perturbation of $\Delta u = 0$.

Theorem 12.10. *Let $s > \frac{3}{2}$ and $B_\delta^{s+2}(\partial\Omega)$ be the open ball of radius δ in $H^{s+2}(\partial\Omega)$, where Ω is a bounded domain with smooth boundary in $\mathbb{R}^3$. If δ is sufficiently small, there exists a solution u to (12.2).*

Proof. Given $h \in H^{s+2}(\partial\Omega)$ define

$$F : H^{s+2}(\partial\Omega) \times H^{s+\frac{5}{2}}(\Omega) \to H^{s+2}(\partial\Omega) \times H^{s+\frac{1}{2}}(\Omega)$$
$$F(h, u) = (u|_{\partial\Omega} - h,\ \Delta u + \mathcal{N}(u)).$$

This is well defined since $D^2 u \in C^0(\Omega)$ by Sobolev embedding and $u|_{\partial\Omega} \in H^{s+2}(\partial\Omega)$ by the trace theorem. We have $F(0, 0) = 0$. F is C^1 in the neighborhood of the origin and

$$D_2 F(0, 0)(w) = DF(0, 0)(0, w) = (w|_{\partial\Omega}, \Delta w).$$

Given $(g, f) \in H^{s+2}(\partial\Omega) \times H^{s+\frac{1}{2}}(\Omega)$, there exists a unique $w \in H^{s+\frac{5}{2}}(\Omega)$ solving

$$\begin{cases} \Delta w &= f \text{ in } \Omega, \\ w &= g \text{ on } \partial\Omega, \end{cases}$$

and elliptic regularity gives

$$\|w\|_{H^{s+\frac{5}{2}}(\Omega)} \leq C \left(\|f\|_{H^{s+\frac{5}{2}}(\Omega)} + \|g\|_{H^{s+2}(\partial\Omega)} \right),$$

so $D_2 F(0, 0)$ is an isomorphism (we did not quite see how to solve the Dirichlet problem in these fractional spaces, but it follow by similar ideas to what we used; solutions with $u \neq 0$ on $\partial\Omega$ follow by considering a problem for $u - g$ with homogeneous boundary condition).

By the implicit function theorem, there exists a $f = \psi(h)$ solving (12.2) if h is small.

□

The implicit function theorem is generally a good tool to find solutions by perturbations.

12.3. **The continuity method.** The basic idea of the continuity method is the following. Suppose we want to solve $P(u) = 0$. We embed this problem into a one parameter family of problems.

$$P_t(u) = 0, \;\; 0 \leq t \leq 1,$$

where $P_1(u) = P(u)$. We then consider

$$A = \{t \in [0,1] \mid P_t(u) = 0 \text{ has a solution}\}.$$

The goal is to show that $A \neq \emptyset$ and that A is open and closed, so that $A = [0,1]$. The usual strategy is:

- to show $A \neq \emptyset$, choose P_t so that $P_0(u) = 0$ is easy to solve
- to show that A is open, use the implicit function theorem. to show that if $P_{t_0}(u) = 0$ has a solution, so does $P_t(u) = 0$ for all t near t_0.
- to show closedness, use estimates for solutions to show that if $\{t_i\} \subset A$, $t_i \to t$, then there exists a subsequence of $\{u_i\}$, where u_i solves $F_{t_i}(u_i) = 0$, converging in a topology such that $F_{t_i}(u_i) \to F_t(u)$.

We will now illustrate the method with the equation

$$\Delta_g u + f - he^u = 0,$$

where $f, h > 0$ (compare to the equation studied in the uniformization theorem).

Theorem 12.11. *Let (M, g) be a closed Riemannian manifold and $f, h : M \to \mathbb{R}$ be smooth functions satisfying $f, h > 0$. Then, there exists a $u \in C^\infty(M)$ solving*

$$\Delta_g u + f - he^u = 0,$$

Proof. Define

$$F(t, u) = \Delta_g u - hu + t(f - h(e^u - u)).$$

Then $F(1, u) = \Delta_g u + f - he^u$. Set

$$A = \{t \in [0,1] \mid F(t,u) = 0 \text{ has a solution } u \in C^2(\Omega)\}$$

For $t = 0$, $u = 0$ solves $F(0, u) = 0$, so $A \neq \emptyset$.

Suppose that $F(t_0, u_0) = 0$. The linearization of F at (t_*, u_*) is

$$L_{(t_*, u_*)} = \Delta_g v - h(1 - t_* + t_* e^{u_*})v.$$

Let $p > \frac{u}{2}$ so that $W^{2,p}(M) \hookrightarrow C^0(M) \subset L^p(M)$. Thus F defines a map

$$F : \mathbb{R} \times W^{2,p}(M) \to L^p(M)$$

and so does $L_{(t_*, u_*)}$

$$L_{(t_*, u_*)} : W^{2,p}(M) \to L^p(M),$$

which is a bounded linear map between these spaces. Consider

$$\begin{aligned} \|L_{(t_*, u_*)} - L_{(\tilde{t}, \tilde{u})}\| &= \sup_{\|v\|_{W^{k,p}(M)} = 1} \left| \left(L_{(t_*, u_*)} - L_{(\tilde{t}, \tilde{u})}\right) v \right| \\ &= \sup_{\|v\|_{W^{k,p}(M)} = 1} \left| h\left(1 - t_* + t_* e^{u_*}\right) v - h(1 - \tilde{t} + \tilde{t} e^{\tilde{u}}) v \right| \\ &= \left| h(\tilde{t} - t_*) + h(t_* e^{u_*} - \tilde{t} e^{\tilde{u}}) \right|. \end{aligned}$$

Since $W^{k,p}(M) \hookrightarrow C^0(M)$ and M is compact, we can make this expression as small as we want by taking t_* close to $\tilde{t}$ in $\mathbb{R}$ and u_* close to $\tilde{u}$ in $W^{k,p}(M)$. Thus, L depends continuously on (t, u), the

Gateaux derivative equals the linearization which equals the Fréchet derivative, and F is C^1.
For (t_0, u_0) we have, $0 \leq t_0 < 1$, we have

$$L_{(t_0,u_0)}v = Lv = \Delta_g v - h(1 - t_0 + t_0 e^{u_0})v$$

and $1 - t_0 + t_0 e^{u_0} > 0$ and, since $h > 0$, by the maximum principle $L_{(t_0 v_0)}v = 0 \Rightarrow v = 0$. By the Fredholm alternative and the elliptic regularity, L is a Banach space isomorphism $L : W^{k,p}(M) \to L^p(M)$. By the implicit function theorem, there exists a u_t solving

$$F(t, u_t) = 0$$

for t near t_0. Bootstrapping the regularity of u as we did in the sub-/super-solutions theorem, we find $u_t \in C^2(M)$, Thus, A is open.

Suppose now that we here a C^2 solution

$$F(t, u) = \Delta_g u - hu + t(f - h(e^u - u)) = 0$$

At a max of u

$$\begin{aligned} 0 &\geq \Delta u = hu - t(f - h(e^u - u)) \\ &= -tf + h((1-t)u + te^u) \end{aligned}$$

Since $h > 0$ and $e^x \geq 1 + x$, and $t \geq 0$.

$$\begin{aligned} 0 &\geq -tf + h(1-t)u + ht(1+u) \\ &= -tf + hu + \underbrace{ht}_{\geq 0} \\ &\geq -tf + hu \end{aligned}$$

Since $t \leq 1$,

$$u \leq \frac{tf}{h} \leq \frac{f}{h} \leq \sup_M \frac{f}{h} = \mathcal{C} < \infty$$

since $h > 0$ and M is compact. Applying a similar argument to the minimum of u, we conclude

$$\|u\|_{C^0(M)} \leq \mathcal{C}$$

where $\mathcal{C}$ does not depend on t. Using this bound, writing $F(t, u) = 0$ as

$$\Delta u = hu - t(f - h(e^u - u)) = \tilde{f}.$$

$\|\tilde{f}\|_{L^p(M)} \leq \mathcal{C}$, applying elliptic regularity and bootstrapping the regularity of u as before, we get

$$\|u\|_{C^{2,\alpha}(M)} \leq \mathcal{C},$$

where $\mathcal{C}$ does not depend on $t \in [0, 1]$. If $\{t_i\} \subset A$, $t_i \to t$, let $\{u_i\}$ be corresponding solutions to $F(t_i, u_i) = 0$. Then

$$\|u_i\|_{C^{2,\alpha}(M)} \leq \mathcal{C}$$

where $\mathcal{C}$ does not depend on i. Thus, by Arzelá-Ascoli, up to a subsequence, $u_i \to u$ in $C^2(M)$. We can thus pass to the limit in the equation to obtain $F(t, u) = 0$, so A is closed. Thus, $A = [0, 1]$ and we found a C^2 solution. Bootstrapping the regularity of this solution as above, we find $u \in C^\infty(M)$. $\square$

13. Linear hyperbolic equations

We will now study linear hyperbolic equations, which are generalizations of the wave equation in a similar manner as we saw that elliptic equations are generalizations of Laplace's equation.

Lemma 13.1. ***(Gronwall Lemma)***. *Let A, φ, and u be non-negative functions on $[T_0, T] \subset \mathbb{R}$, $u \in L^\infty([T_0,T])$, $\phi \in L^1([T_0,T])$, and A is non-decreasing. Suppose that*

$$u(t) \leq A(t) + \int_{T_0}^t \varphi(\tau) u(\tau) d\tau$$

for all $t \in [T_0, T]$. Then,

$$u(t) \leq A(t) e^{\int_{T_0}^t \varphi(\tau) d\tau}$$

for all $t \in [T_0, T]$.

Proof. It suffices to prove it for $t \geq T$ since A is non-decreasing. So we can assume

$$A = A(T) = \text{constant}.$$

Set

$$F(t) := A + \int_{T_0}^t \varphi(\tau) u(\tau) d\tau.$$

Then:

- F is differentiable a.e.
- $F' = \varphi u$
- F is absolutely continuous.

Then

$$G(t) := F(t) e^{-\int_{T_0}^t \varphi(\tau) d\tau}$$

- is absolutely continuous (since $\int_{T_0}^T \varphi(\tau) d\tau$ is).
- G is differentiable a.e.

We have

$$\begin{aligned} G' &= \underbrace{F'(t)}_{=\phi(t)u(t)} e^{-\int_{T_0}^t \varphi(\tau) d\tau} - F(t)\varphi(t) e^{-\int_{T_0}^t \varphi(\tau) d\tau} \\ &= \underbrace{\varphi(t)}_{\geq 0} \underbrace{(u(t) - F(t))}_{\leq 0} e^{-\int_{T_0}^t \varphi(\tau) d\tau} \leq 0 \end{aligned}$$

Then

$$G(t) \leq G(T_0) = F(T_0) = A \implies F \leq A e^{\int_{T_0}^t \varphi(\tau) d\tau}.$$

Since $u \leq F$, the result follows.

□

Remark 13.2. For simplicity, we will work on the interval $[0, T]$, i.e., $T_0 = 0$. It will be clear the results will hold on $[T_0, T]$, $T_0 \neq 0$.

13.1. **Linear first-order symmetric hyperbolic systems.** Let us consider the initial-value problem or Cauchy problem

$$\begin{cases} A^\mu \partial_\mu u + B \quad u = f \text{ in } [0,T] \times \mathbb{R}^n, \\ \qquad\qquad\quad\; u = u_0 \text{ on } \{t = 0\} \times \mathbb{R}^n, \end{cases} \tag{13.1}$$

where $f : [0,T] \times \mathbb{R}^n \to \mathbb{R}^d$, $A, B : [0,T] \times \mathbb{R}^n \to M_{d \times d} = d \times d$ matrices and $u : [0,T] \times \mathbb{R}^n \to \mathbb{R}^d$ is the unknown.

Definition 13.3. We say that the PDE in (13.1) is a **(linear) first-order symmetric hyperbolic system (FOSH)** if the matrices A^μ are symmetric and $A^0 = A^t$ is uniformly positive definite, i.e.,

$$A^0(x)(\xi,\xi) \geq C|\xi|^2 \quad \text{for all } x \in [0,T]\times\mathbb{R}^n.$$

Notation 13.4. We will often write (t,x) for points in $[0,T]\times\mathbb{R}^n$, i.e., $(t,x)\in[0,T]\times\mathbb{R}^n$. Denote

$$M_T \equiv M := [0,T]\times\mathbb{R}^n.$$
$$\Sigma_t := \big\{(t,x)\in M\big\}.$$

We often write u for u^t = transpose of u if it is clear from the context, e.g., $uAu = u^tAu$, $A = d\times d$ matrix.

Theorem 13.5. ***(Basic Energy Estimate)*** *Assume that u is a smooth solution to the FOSH system (13.1) such that $u(t,\cdot)$ and $\partial_t u(t,\cdot)$ are Schwartz functions with constants that are uniform in t, i.e.,*

$$\left|x^{\vec{\alpha}}\right|\left(\left|D^{\vec{\beta}}u\right| + \left|D^{\vec{\beta}}\partial_t u\right|\right) \leq C_{\vec{\alpha},\vec{\beta}}$$

in M_T (note that f then also satisfies similar bounds). Suppose that A^μ, B and all their derivatives are bounded in M_T. Set

$$E(t) = \frac{1}{2}\int_{\Sigma_t} uA^0u\,dx \equiv \frac{1}{2}\int_{\mathbb{R}^n} uA^0u\,dx.$$

Then, there exists a constant $\mathcal{C}>0$ independent of u such that

$$\sqrt{E(t)} \leq \left(\sqrt{E(0)} + \mathcal{C}\int_0^t \|f(\tau,\cdot)\|_{L^2(\sigma_\tau)}d\tau\right)e^{\mathcal{C}t}.$$

for all $t\in M_T$.

Proof. Compute

$$\begin{aligned}
\partial_t E &= \underbrace{\frac{1}{2}\int_{\Sigma_t}\partial_t uA^0u + \frac{1}{2}\int_{\Sigma_t} uA^0\partial_t u}_{=\int_{\Sigma_t} uA^0\partial_t u \text{ by symmetry of } A^0} + \frac{1}{2}\int_{\Sigma_t} u\partial_t A^0 u \\
&= -\int_{\Sigma_t}(uA^i\partial_i u + uBu - uf) + \frac{1}{2}\int_{\Sigma_t} u\partial_t A^0u, \text{ by symmetry of } A^i: \\
&= -\int_{\Sigma_t}\left(\frac{1}{2}\partial_i(uA^iu) - \frac{1}{2}u\partial_iA^iu\right) + \int_{\Sigma_t}\left(\frac{1}{2}u\partial_tA^0u + uf - uBu\right) \\
&\underset{\text{by parts}}{=} \int_{\Sigma_t}\left(\frac{1}{2}u\partial_\mu A^\mu u - uBu + uf\right).
\end{aligned}$$

We have

$$\int_{\Sigma_t}\left(u\partial_\mu A^\mu u - uBu\right) \leq C\int_{\Sigma_t}|u|^2 \leq C\int_{\Sigma_t} uA^0u = CE(t)$$
$$\int_{\Sigma_t} uf \leq \|u\|_{L^2(\Sigma_t)}\|f\|_{L^2(\Sigma_t)} \leq \sqrt{E(t)}\,\|f\|_{L^2(\Sigma_t)},$$

so

$$\partial_t E \leq CE + C\sqrt{E}\|f\|_{L^2}.$$

Setting $E_\varepsilon(t) = E(t)+\varepsilon, \varepsilon>0$, the same inequality holds for $E_\varepsilon(t)$, so

$$\partial_t E_\varepsilon \leq CE_\varepsilon + C\sqrt{E_\varepsilon}\|f\|_{L^2},$$

so

$$\frac{\partial_t E_\varepsilon}{\sqrt{E_\varepsilon}} = 2\partial_t\sqrt{E_\varepsilon} \leq C\sqrt{E_\varepsilon} + \|f\|_{L^2}.$$

Then

$$\sqrt{E_\varepsilon(t)} \le \sqrt{E_\varepsilon(0)} + C\int_0^t \|f\|_{L^2(\Sigma_\tau)} d\tau + C\int_0^t \sqrt{E_\varepsilon(\tau)} d\tau$$

By Gronwall's Lemma:

$$\sqrt{E_\varepsilon(t)} \le \left(\sqrt{E_\varepsilon(0)} + C\int_0^t \|f\|_{L^2(\Sigma_\tau)}\right) e^{Ct}$$

which gives the result taking $\epsilon \to 0$.

□

Definition 13.6. The **commutator** of two differential operators P and Q is defined as

$$[P,Q] := PQ - QP$$

wherever the RHS is defined.

Remark 13.7. If P and Q have orders k and l, respectively, and are linear, then $[P,Q]$ has order $k+l-1$, since

$$P = \sum_{|\alpha|\le k} a_\alpha D^\alpha,\ Q = \sum_{|\alpha|\le l} b_\alpha D^\alpha$$

$$PQ = \sum_{|\alpha|=k, |\beta|=l} a_\alpha b_\beta D^{\alpha+\beta} + \text{terms where at least one derivative falls on } b_\alpha$$

$$QP = \sum_{|\alpha|=k, |\beta|=l} b_\beta a_\alpha D^{\beta+\alpha} + \text{terms where at least one derivative falls on } a_\alpha$$

Corollary 13.8. *(higher order energy estimates). Under the same assumptions of the theorem,*

$$\partial_t E_k \le CE_k + C\sqrt{E_k}\ \|f\|_{H^k(\Sigma_t)},$$
$$E_k(t) \le C_T\left(\sqrt{E_k(0)} + \int_0^t \|f(\tau,\cdot)\|_{H^k(\Sigma_\tau)} d\tau\right)$$

where C_T = constant depending on T and

$$E_k(t) := \frac{1}{2}\sum_{|\vec{\alpha}|\le k}\int_{\Sigma_t} D^{\vec{\alpha}} u A^0 D^{\vec{\alpha}} u dx.$$

Proof. Write the equation as $Lu = f$ and $E_k = E_k(u)$. Then, applying $D^{\vec{\alpha}}$ to the equation.

$$\underbrace{D^{\vec{\alpha}} Lu}_{=L(D^\alpha \vec{u}) + [D^{\vec{\alpha}}, L]u} = D^{\vec{\alpha}} f \implies LD^{\vec{\alpha}} u = D^{\vec{\alpha}} f + [L, D^{\vec{\alpha}}]u$$

Applying the basic energy estimate to $\vec{D}^\alpha u$ with f replaced by $D^{\vec{\alpha}} f + [L, D^{\vec{\alpha}}]u$ (or, more precisely, applying an intermediate inequality that was obtained in the proof of the basic energy estimate):

$$\partial_t E(D^{\vec{\alpha}} u) \le CE(D^{\vec{\alpha}} u) + C\sqrt{E(D^{\vec{\alpha}} u)}\left(\|D^{\vec{\alpha}} f + [L, D^{\vec{\alpha}}]u\|_{L^\alpha(\Sigma_t)}\right).$$

We have, since $|\vec{\alpha}| \leq k$, $\|D^{\vec{\alpha}} f\|_{L^\alpha(\Sigma_t)} \leq \|f\|_{H^k(\Sigma_t)}$.

$$\begin{aligned}
[D^{\vec{\alpha}}, L]u &= [D^{\vec{\alpha}}, A^\mu \partial_\mu]u + [D^{\vec{\alpha}}, B]u \\
[D^{\vec{\alpha}}, A^\mu \partial_\mu]u &= D^{\vec{\alpha}}(A^\mu \partial_\mu u) - A^\mu \partial_\mu D^{\vec{\alpha}} u \\
&= \sum_{\vec{\beta} \leq \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^\mu \partial_\mu D^{\vec{\beta}} u - A^\mu \partial_\mu D^{\vec{\alpha}} u \\
&= \underbrace{\sum_{\vec{\beta} = \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^\mu \partial_\mu D^{\vec{\beta}} u}_{= A^\mu \partial_\mu D^{\vec{\alpha}} u} + \sum_{\vec{\beta} < \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^\mu \partial_\mu D^{\vec{\beta}} u - A^\mu \partial_\mu D^{\vec{\alpha}} u \\
&= \sum_{\vec{\beta} < \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^\mu \partial_\mu D^{\vec{\beta}} u \\
&= \sum_{\vec{\beta} < \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^0 \partial_0 D^{\vec{\beta}} u + \sum_{\vec{\beta} < \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^i \partial_i D^{\vec{\beta}} u
\end{aligned}$$

The second term gives

$$\left\| \sum_{\vec{\beta} < \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^i \partial_i D^{\vec{\beta}} u \right\|_{L^2(\sum_t)} \leq C\|u\|_{H^k(\sum_t)}$$

For the first term, we use the equation and the fact that A^0 is invertible to write

$$\partial_t u = (A^0)^{-1}(f - A^i \partial_i u)$$

so that

$$\begin{aligned}
\sum_{\vec{\beta} < \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^0 \partial_0 D^{\vec{\beta}} u &= \sum_{\vec{\beta} < \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^0 D^{\vec{\beta}} \partial_0 u \\
&= \sum_{\vec{\beta} < \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^0 D^{\vec{\beta}} \left((A^0)^{-1}(f - A^i \partial_i u)\right)
\end{aligned}$$

So

$$\begin{aligned}
\left\| \sum_{\vec{\beta} < \vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} D^{\vec{\alpha}-\vec{\beta}} A^0 \partial_0 D^{\vec{\beta}} u \right\|_{L^\alpha} &\leq C\|f\|_{H^{h-1}(\sum_t)} + C\|u\|_{H^k(\sum_t)} \\
&\leq C\|f\|_{H^k(\sum_t)} + C\|u\|_{H^k(\sum_t)},
\end{aligned}$$

where we used that $(A^0)^{-1}$ is smooth since A^0 is $[D^{\vec{\alpha}}, \beta]u$ is handled similarly. Thus

$$\begin{aligned}
\partial_t E(D^{\vec{\alpha}} u) &\leq C E(D^{\vec{\alpha}} u) + C\sqrt{E(D^{\vec{\alpha}} u)} \left(\|D^{\vec{\alpha}} f + [L, D^{\vec{\alpha}}]u\|_{L^2(\Sigma_t)} \right). \\
&\leq C E(D^{\vec{\alpha}} u) + C\sqrt{E(D^{\vec{\alpha}} u)}\, \|u\|_{H^k(\Sigma_t)} + C\sqrt{E(D^{\vec{\alpha}} u)}\, \|f\|_{H^k(\Sigma_t)}
\end{aligned}$$

Since A^0 is positive definite and bounded,

$$\frac{1}{C}\|u\|^2_{H^k(\Sigma_t)} \leq E_k(t) \leq C\|u\|^2_{H^k(\Sigma_t)},$$

so, using that $E(D^{\vec{\alpha}} u) \leq E_k$,

$$\partial_t E(D^{\vec{\alpha}} u) \leq C E_k + \sqrt{E_k}\|f\|_{H^k(\Sigma_t)}.$$

Summing over α and using

$$E_k = \sum_{|\alpha|\leq k} E(D^{\vec{\alpha}}u)$$

we have the first inequality. Dividing by $\sqrt{E_k}$ (or $\sqrt{E_k}+\epsilon$, as before) and using Gronwall's inequality gives the result.

□

Next, we extend the result for negative k:

Corollary 13.9. *Under the same assumptions as above,*

$$\|u(t,\cdot)\|_{H^k(\Sigma_t)} \leq C_T \left(\|u(0,\cdot)\|_{H^k(\Sigma_0)} + \int_0^t \|f(\tau,\cdot)\|_{H^k(\Sigma_\tau)} d\tau \right)$$

for any $k \in \mathbb{Z}$.

Proof. We already have the result for $k \geq 0$, so take $k < 0$. Set

$$v := (1-\Delta)^k u.$$

Because $(1-\Delta)^k$ maps $\mathcal{S}$ into itself (it is defined using the Fourier transform that maps $\mathcal{S}$ into itself), v satisfies Schwartz bounds similar to u.

We have

$$\begin{aligned} \|u\|_{H^k(\Sigma_t)} &= \|(1-\Delta)^{\frac{k}{2}} u\|_{L^2(\Sigma_t)} = \|(1-\Delta)^{-\frac{k}{2}} v\|_{L^2(\Sigma_t)} \\ &= \|v\|_{H^{-k}|\Sigma_t|} \leq C\sqrt{E_{-k}(v)} \end{aligned}$$

Since v identically satisfies the equation

$$Lv = F =: Lv,$$

the energy estimates give

$$\|u\|_{H^k(\Sigma_t)} \leq C\sqrt{E_{-k}(v)(0)} + C\int_0^t \|Lv\|_{H^{-k}(\Sigma_\tau)} d\tau$$

Since

$$\sqrt{E_{-k}(v)(0)} \leq \|v\|_{H^{-k}(\Sigma_0)} = \|u\|_{H^k(\Sigma_0)},$$

it remains to estimate the term $\|Lv\|_{H^{-k}(\Sigma_t)}$.

Observe the identity

$$\begin{aligned} (1-\Delta)^{-k}Lv + [L,(1-\Delta)^{-k}]v &= (1-\Delta)^{-k}Lv + L(1-\Delta)^{-k}v - (1-\Delta)^{-k}Lv \\ &= L\underbrace{(1-\Delta)^{-k}v}_{=u} = Lu = f, \end{aligned}$$

so

$$(1-\Delta)^{-k}Lv = f - [L,(1-\Delta)^{-k}]v$$

Apply $(1-\Delta)^{\frac{k}{2}}$ and take the L^2 norm:

$$\begin{aligned} \underbrace{\|(1-\Delta)^{-\frac{k}{2}}Lv\|_{L^2(\Sigma_t)}}_{=\|Lv\|_{H^{-k}(\Sigma_t)}} &\leq \underbrace{\|(1-\Delta)^{\frac{k}{2}}f\|_{L^2(\Sigma_t)}}_{=\|f\|_{H^k(\Sigma_t)}} \\ &\quad + \underbrace{\|(1-\Delta)^{\frac{k}{2}}[L,(1-\Delta)^{-k}]v\|_{L^2(\Sigma_t)}}_{=\|[L,(1-\Delta)^{-k}]v\|_{H^k(\Sigma_t)}} \end{aligned}$$

Expanding $(1-\Delta)^{-k} = \sum_{|\vec{\alpha}|\leq -2k} a_{\vec{\alpha}} D^{\vec{\alpha}}$, then (recall $k<0$)

$$
\begin{aligned}
[L,(1-\Delta)^{-k}]v &= [A^\mu\partial_\mu, \sum_{|\vec{\alpha}|\leq -2k} a_{\vec{\alpha}} D^{\vec{\alpha}}]v \\
&= \sum_{|\vec{\alpha}|\leq -2k} \left(A^\mu \partial_\mu (a_{\vec{\alpha}} D^{\vec{\alpha}} v) - a_{\vec{\alpha}} D^{\vec{\alpha}} (A^\mu \partial_\mu v) \right) \\
&= \sum_{|\vec{\alpha}|\leq -2k} A^\mu \partial_\mu a_{\vec{\alpha}} D^{\vec{\alpha}} v - \sum_{|\vec{\alpha}|\leq -2k} \sum_{\vec{\beta}<\vec{\alpha}} \binom{\vec{\alpha}}{\vec{\beta}} a_{\vec{\alpha}} D^{\vec{\alpha}-\vec{\beta}} A^\mu \partial_\mu D^{\vec{\beta}} v
\end{aligned}
$$

Thus

$$
\begin{aligned}
\left\|[L,(1-\Delta)^{-k}]v\right\|_{H^k(\Sigma_t)} &\leq C\sum_{l=0}^{-2k} \|D^l v\|_{H^k(\Sigma_t)} + C\sum_{l=0}^{-2k-1} \|D^l \partial_t v\|_{H^k(\Sigma_t)} \\
&\leq C\|v\|_{H^{-k}(\Sigma_t)} + C\|\partial_t v\|_{H^{-k-1}(\Sigma_t)}
\end{aligned}
$$

Using that $D: H^s \to H^{s-1}$ is bounded.

Let us estimate $\|\partial_t v\|_{H^{-k-1}(\Sigma_t)}$. Set $\tilde{L} = (A^0)^{-1}L$. Then, arguing similarly to above:

$$
(1-\Delta)^{-k}\tilde{L}v + [\tilde{L},(1-\Delta)^{-k}]v = \tilde{L}\underbrace{(1-\Delta)^{-k}v}_{=u} = (A^0)^{-1}f
$$

But $(1-\Delta)^{-k}\tilde{L}v = (1-\Delta)^{-k}\left(\partial_t v + (A^0)^{-1}A^i\partial_i v\right)$, so

$$
\begin{aligned}
(1-\Delta)^{-k}\partial_t v &= -(1-\Delta)^{-k}\left((A^0)^{-1}A^i\partial_i v\right) \\
&= -[\tilde{L},(1-\Delta)^{-k}]v \\
&= (A^0)^{-1} f
\end{aligned}
$$

Since $\|\partial_t v\|_{H^{-k-1}(\Sigma_t)} = \|(1-\Delta)\|^{\frac{-k-1}{2}}\partial_t v\|_{L^2(\Sigma_t)}$, we apply $(1-\Delta)^{\frac{k-1}{2}}$ and estimate in L^2

$$
\begin{aligned}
\|(1-\Delta)^{\frac{-k-1}{2}}\partial_t v\|_{L^2} \leq\ & C\|(1-\Delta)^{-k} Dv\|_{H^{k-1}(\Sigma_t)} \\
&+ C\|[\tilde{L},(1-\Delta)^{-k}]v\|_{H^{k-1}(\Sigma_t)} \\
&+ C\|f\|_{H^{k-1}(\Sigma_t)}
\end{aligned}
$$

$(1-\Delta)^{-k}Dv$ contains at most $-2k+1$ derivatives and $[\tilde{L},(1-\Delta)^{-k}]$ at most $2k$ derivatives and no time derivative, so the first two terms on the RHS are bounded by

$$
\|D^{-2k+1}v\|_{H^{k-1}(\Sigma_t)} \leq \|v\|_{H^{-k}(\Sigma_t)}.
$$

Thus

$$
\|\partial_t v\|_{H^{-k-1}(\Sigma_t)} \leq C\|v\|_{H^{-k}(\Sigma_t)} + C\|f\|_{H^{k-1}(\Sigma_t)}.
$$

Putting all together and invoking Gronwall's inequality gives the result.

□

Corollary 13.10. *Under the same assumptions as above,*

$$
\|u(t,\cdot)\|_{H^k(\Sigma_t)} \leq C_T\left(\|u(T,\cdot)\|_{H^k(\Sigma_t)} + \int_t^T \|f(\tau,\cdot)\|_{H^k(\Sigma_\tau)} d\tau\right)
$$

for any $k \in \mathbb{Z}$.

Proof. Set

$$
\begin{aligned}
\hat{L}v(t,x) :=& - A^0(T-t,x)\partial_t v(t,x) + A^i(T-t,x)\partial_i v(t,x) \\
&+ B(T-t,x)v^(t,x \\
\hat{u}(t,x) :=& \, u(T-t,x).
\end{aligned}
$$

Then $\partial_t \hat{u}(t,x) = -\partial_t u(T-t,x), \partial_i \hat{u}(t,x) = \partial_i u(T-t,x)$, so that

$$
\begin{aligned}
\hat{L}\hat{u}(t,x) =& - A^0(T-t,x)\partial_t x\hat{u}(t,x) + A^i(T-t,x)\partial_i \hat{u}(t,x) \\
&+ B(T-t,x)\hat{u}(t,x) \\
=& A^0(T-t,x)\partial_t u(T-t,x) + A^i(T-t,x)\partial_i u(T-t,x) \\
&+ B(T-t,x)u(T-t,x) \\
=& (Lu)(T-t,x)
\end{aligned}
$$

The operator $-\hat{L}$ satisfies the same assumptions as L, thus we have an estimate

$$
\|u(t,\cdot)\|_{H^k(\Sigma_t)} \leq C \left(\|v(0,\cdot)\|_{H^k(\Sigma_0)} + \int_0^t \|\hat{L}v(\tau,\cdot)\|_{H^k(\Sigma_\tau)} d\tau \right).
$$

For $v = \hat{u}$

$$
\begin{aligned}
\int_0^t \|\hat{L}\hat{u}(\tau,\cdot)\|_{H^k(\Sigma_\tau)} &= \int_0^t \|Lu(T-\tau,\cdot)\|_{H^k(\Sigma_{T-\tau})} d\tau \\
&= -\int_T^{T-t} \|Lu(s)\|_{H^k(\Sigma_s)} ds \\
&= \int_{T-t}^T \|Lu(s)\|_{H^k(\Sigma_s)} ds,
\end{aligned}
$$

so, since $Lu = f$

$$
\|u(T-t,\cdot)\|_{H^k(\Sigma_{T-t})} \leq C \left(\|u(T,\cdot)\|_{H^k(\Sigma_T)} + \int_{T-t}^T \|f(s)\|_{H^k(\Sigma_s)} ds \right).
$$

Since any $t' \in [0,T]$ is of the form $t' = T - t$ for some $t \in [0,T]$, we have the result. $\square$

Definition 13.11. We denote

$$
\mathcal{C}_{t_0,x_0,c_0} := \{(t,x) \in M_T \big| 0 < t < t_0, x \in B_{t_0 - c_0 t}(x_0)\}
$$

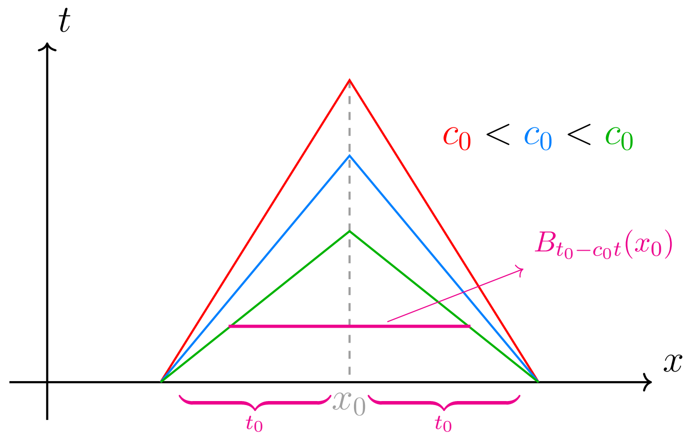


FIGURE 31. $\mathcal{C}_{t_0,x_0,c_0}$

Theorem 13.12. ***(Domain of Dependence and Uniqueness).*** *Let $u \in C^1\big([0,T]\times\mathbb{R}^n\big)$ solution to the FOSH system (13.1), where the A^μ are C^1, bounded, and have bounded derivatives, B is C^0 and bounded, and f is continuous. There exists a $c_0 > 0$, depending on the lower bound of A^0 and the upper bounded of A^i, such that if $u_0 = 0$ on $B_r(x_0)$ and $f = 0$ in $\mathcal{C}_{r,x_0,c_0}$, then $u = 0$ in $\mathcal{C}_{r,x_0,c_0}$, In particular, solutions are unique.*

Proof. Consider, for $a > 0$,

$$\begin{aligned}\partial_\mu(e^{-at}uA^\mu u) &= e^{-at}2uA^\mu\partial_\mu u + e^{-at}u\partial_\mu A^\mu u - ae^{-at}uA^0u\\ &= 2e^{-at}u(f-Bu) + e^{-at}u\partial_\mu A^\mu u - ae^{-at}uA^0u.\end{aligned}$$

Integrate over $\mathcal{C} = \mathcal{C}_{r,x_0,c_0}$, where c_0 will be chosen.

$$\begin{aligned}\int_{\mathcal{C}}\partial_\mu(e^{-at}uA^\mu u) &= \int_{\partial\mathcal{C}} e^{-at}uA^\mu u v_\mu\\ &= \int_{\partial_L\mathcal{C}} e^{-at}uA^\mu u\nu_\mu - \underbrace{\int_{B_r(x_0)} e^{-at}uA^0u}_{=0},\end{aligned}$$

where ν is the unit outer normal to $\mathcal{C}$, $\partial_L\mathcal{C}$ is the lateral boundary so that $\partial\mathcal{C} = \partial_L\mathcal{C}\cup B_r(x_0)$, and we used $\nu = (-1,0,...,0)$ on $B_r(x_0)$. We can make the components ν_i as small as we want by taking c_0 large enough, so that $\nu_0 > 0$ and

$$\begin{aligned}uA^\mu u\nu_\mu &= uA^0u\nu_0 + uA^iu\nu_i\\ &\geq C_1|u|^2 - C_2|u|^2 > C|u|^2\end{aligned}$$

Thus $\int_{\mathcal{C}}\partial_\mu(e^{-at}uA^\mu u) \geq 0$. On the other hand

$$\int_{\mathcal{C}}\left(2e^{-at}u(\overbrace{f}^{=0 \text{ in } \mathcal{C}} - Bu) + e^{-at}u\partial_\mu A^\mu u - ae^{-at}\overbrace{uA^0u}^{\geq C|u|^2}\right)$$

$$\leq \int_{\mathcal{C}} e^{-at}\left(-2uBu + u\partial_\mu A^\mu u - aC|u|^2\right) \leq 0$$

if we choose a large enough. Thus RHS ≤ 0, $LHS \geq 0$, thus, since LHS = RHS, both sides must vanish. Since we can freely choose a (large enough), we must have $u = 0$ in $\mathcal{C}$.

□

Definition 13.13. Let $L = A^\mu \partial_\mu + B$ be a first-order symmetric hyperbolic operator. The **formal adjoint of** L **is**

$$\begin{aligned} L^* u &= -\partial_t(A^0 u) - \partial_i(A^i u) + B^* u \\ &= -A^\mu \partial_\mu u - \partial_\mu A^\mu u + B^* u, \end{aligned}$$

where $B^* =$ transpose of B.

The motivation for the definition comes from integration by parts, e.g., if $\psi \in C_c^\infty(M_T)$,

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^n} \psi A^\mu \partial_\mu u dx dt &= \int_{\mathbb{R}^n} \int_0^T \underbrace{\psi A^0}_{=A^0\psi} \partial_t u dt dx + \int_0^T \int_{\mathbb{R}^n} \underbrace{\psi A^i}_{A^i\psi} \partial_i u dx dt \\ &= -\int_0^T \int_{\mathbb{R}^n} \partial_t (A^0 \psi) u dx dt - \int_0^T \int_{\mathbb{R}^n} \partial_i (A^i \psi) u dx dt, \end{aligned}$$

where there are no boundary terms because $\psi \in C_c^\infty(M_T)$.

Theorem 13.14. *Consider the Cauchy problem for the FOSH system (13.1). Assume that* $u_0 \in C_c^\infty(\mathbb{R}^n), f \in C_c^\infty(\mathbb{R} \times \mathbb{R}^n)$*,* A^μ, B *are* C^∞ *with all derivatives bounded. Moreover, there exists a compact set* $\kappa \subset \mathbb{R}^n$ *such that* $u = 0$ *outside* $[0, T] \times \kappa$*.*

Proof. The uniqueness and the statement about κ follow from the domain of dependence/uniqueness result.

Let $\varphi \in C_c^\infty(\mathbb{R} \times \mathbb{R}^n)$ be such that $\varphi(t, x) = 0$ for $t \geq T$. Since $-L^*$ is also a first-order symmetric operator, we have, by one of the above corollaries,

$$\begin{aligned} \|\varphi\|_{H^{-k}(\Sigma_t)} &\leq C \left(\underbrace{\|\varphi\|_{H^{-k}_{(\Sigma_T)}}}_{=0} + \int_t^T \|L^*\varphi\|_{H^{-k}_{(\Sigma_\tau)}} d\tau \right) \\ &\leq C \int_t^T \|L^*\varphi\|_{H^{-k}_{(\Sigma_\tau)}} d\tau \end{aligned}$$

This implies, in particular, that if $L^*\varphi = 0$ for $t \in [0, T]$ then $\varphi = 0$. Given $\psi \in L^1\big([0, T], H^k(\mathbb{R}^n)\big)$, for φ as above and $k \geq 1$, set

$$\langle \varphi, \psi \rangle := \int_0^T (\varphi, \psi)_0 dt$$

which is well defined by the generalized Cauchy-Szhwartz inequality.

For φ as above, $L^*\varphi \in L^1\big([0, T], H^{-k}(\mathbb{R}^n)\big)$. Let $X \subset L^1\big([0, T], H^{-k}(\mathbb{R}^n)\big)$ be the subspace spanned by $L^*\varphi, \varphi$ as above. Define $F_\psi : X \to \mathbb{R}$ by

$$F_\psi(L^*\varphi) = \langle \psi, \varphi \rangle.$$

Note that F_ψ is well defined ($L^*\varphi = 0 \implies \varphi = 0$ for $0 \leq t \leq T$) and is bounded by the above energy estimate. By Hahn-Banach, F_ψ extends to a bounded linear functional $\tilde{F}_\psi$ on $L^1\big([0, T], H^{-k}(\mathbb{R}^n)\big)$ (with same norm as F_ψ). By one of our duality theorems, there exists a $u \in L^\infty\big([0, T], H^k(\mathbb{R}^n)\big)$ such that

$$\tilde{F}_\psi(v) = \int_0^T (u, v)_0 dt = \langle u, v \rangle$$

for all $v \in L^1\big([0, T], H^{-k}(\mathbb{R}^n)\big)$.

In particular, for elements in X,

$$\tilde{F}_\psi(L^*\varphi) = F_\psi(L^*\varphi) = \int_0^T (u, L^*\varphi)_0 dt$$
$$= \langle \psi, \varphi \rangle = \int_0^T (\psi, \varphi)_0 dt,$$

i.e.,

$$\int_0^T (\psi, \varphi)_0 dt = \int_0^T (u, L^*\varphi)_0 dt.$$

Consider now f as in the theorem, but assume further that $f(t,x) = 0$ for $t \leq 0$. Take $\psi = f$ above and extend u to be identically zero for $t < 0$, so $u \in L^\infty\big((-\infty, T], H^k(\mathbb{R}^n)\big)$. Therefore,

$$\int_{-\infty}^T (f, \varphi)_0 dt = \int_{-\infty}^T (u, L^*\varphi)_0 dt$$

for all $\varphi \in C_c^\infty(\mathbb{R} \times \mathbb{R}^n)$ such that $\varphi(t,x) = 0$ for $t \geq T$. We would like to integrate by parts to obtain $(Lu, \varphi)_0$ and then $Lu = f$, but u is not regular enough in time. Thus, we proceed as follows.

Let $\tilde{u} \in L^2_{\text{loc}}((-\infty, T) \times \mathbb{R}^n)$ such that $\tilde{u}$ is k-times weakly differentiable with respect to x and such that

$$\int_{-\infty}^T \int_{\mathbb{R}^n} u\varphi dx dt = \int_{-\infty}^T \int_{\mathbb{R}^n} \tilde{u}\varphi dx dt$$

for all $\varphi \in C^\infty((-\infty, T) \times \mathbb{R}^n)$ (the existence of such $\tilde{u}$ can be demonstrated; $\tilde{u}$ is "essentially" u). Applying this to φ replaced by $L^*\varphi$, and using the above, we have

$$\int_{-\infty}^T \int_{\mathbb{R}^n} f\varphi dx dt = \int_{-\infty}^T \int_{\mathbb{R}^n} L^*\varphi \tilde{u} dx dt$$

for all $\varphi \in C_c^\infty((-\infty, T) \times \mathbb{R}^n)$. Write

$$\int_{-\infty}^T \int_{\mathbb{R}^n} L^*\varphi \tilde{u} dx dt = \int_{-\infty}^T \int_{\mathbb{R}^n} -\partial_t(A^0\varphi)\tilde{u} dx dt - \int_{-\infty}^T \int_{\mathbb{R}^n} \underbrace{\partial_i(A^i\varphi)}_{\substack{\text{we can integrate by} \\ \text{parts because of the} \\ x\text{-differentiability of } \tilde{u}}} \tilde{u} dx dt + \int_0^T \int_{\mathbb{R}^n} \underbrace{B^*\varphi\tilde{u}}_{=\varphi B\tilde{u}} dx dt$$
$$= -\int_{-\infty}^T \int_{\mathbb{R}^n} \partial_t(A^0\varphi)\tilde{u} dx dt + \int_{-\infty}^T \int_{\mathbb{R}^n} \varphi A^i \partial_i \tilde{u} dx dt + \int_{-\infty}^T \int_{\mathbb{R}^n} \varphi B\tilde{u} dx dt.$$

Therefore

$$\int_{-\infty}^T \int_{\mathbb{R}^n} f\varphi dx dt = -\int_{-\infty}^T \int_{\mathbb{R}^n} \partial_t(A^0\varphi)\tilde{u} dx dt + \int_{-\infty}^T \int_{\mathbb{R}^n} \varphi A^i \partial_i \tilde{u} dx dt + \int_{-\infty}^T \int_{\mathbb{R}^n} \varphi B\tilde{u} dx dt.$$

Write this as

$$\int_{-\infty}^T \int_{\mathbb{R}^n} \varphi(f - A^i\partial_i\tilde{u} - B\tilde{u}) dx dt = -\int_{-\infty}^T \int_{\mathbb{R}^n} \partial_t(A^0\varphi)\tilde{u} dx dt$$

Any $\psi \in C_c^\infty((-\infty, T) \times \mathbb{R}^n)$ can be written as $\psi = A^0\varphi$ for some $\varphi \in C_c^\infty((-\infty, T) \times \mathbb{R}^n)$ by our assumptions on A^0, so the above reads, using that $(A^0)^{-1}$ is symmetric

$$\int_{-\infty}^T \int_{\mathbb{R}^n} \psi[(A^0)^{-1}(f - A^i\partial_i\tilde{u} - B\tilde{u})] dx dt = -\int_{-\infty}^T \int_{\mathbb{R}^n} \partial_t\psi\tilde{u} dx dt$$

for all $\psi \in C_c^\infty((-\infty, T) \times \mathbb{R}^n)$. This shows that $\tilde{u}$ has a ∂_t weak derivative. Moreover, $\partial_t\tilde{u}$ is given by

$$\partial_t\tilde{u} = (A^0)^{-1}(f - A^i\partial_i\tilde{u} - B\tilde{u}).$$

Since u is k-times weakly differentiable in x the RHS admits $k-1$ weak spatial derivatives, so $D^{\vec{\alpha}}\partial_t\tilde{u} = \partial_t D^{\vec{\alpha}}\tilde{u}$ exists, $|\vec{\alpha}| \leq k-1$. We can now iterate this argument: apply the above identity with ψ replaced by $\partial_t D^{\vec{\alpha}}\psi, |\vec{\alpha}| \leq k-2$.

$$\int_{-\infty}^{T}\int_{\mathbb{R}^n} \partial_t D^{\vec{\alpha}}\psi[(A^0)^{-1}(f - A^i\partial_i\tilde{u} - B\tilde{u})]dxdt = -\int_{-\infty}^{T}\int_{\mathbb{R}^n} \partial_t(\partial_t D^{\vec{\alpha}}\psi)\tilde{u}dxdt.$$

Since, by the above, $\partial_t D^{\vec{\alpha}}\partial_i u$ exists, we can integrate by parts on the LHS to conclude that $\partial_t^2 D^{\vec{\alpha}}$ weak derivative of $\tilde{u}$ exists. Proceeding in this way, we conclude that

$$\partial_t^j D^{\vec{\alpha}}\tilde{u}, \ \ j + |\vec{\alpha}| \leq k$$

exists weakly. Since we can take k very large, by Sobolev embedding we conclude that $\tilde{u}$ is, say, C^l for some large l. (We have not said that $\tilde{u} \in H^k(\mathbb{R}^n)$, but to conclude that $\tilde{u}$ is C^l it suffices to apply the Sobolev embedding theorem to $\varphi\tilde{u}$, with test functions φ).

We can now integrate by parts to get $Lu = f$ pointwise in $(-\infty, T] \times \mathbb{R}^n$.

To conclude, finally, that $\tilde{u}$ is C^∞, observe that for a (large) k we obtain a C^l solution and for a different (large) k' we obtain a $C^{l'}$ solution, and in principle these two solutions need not to coincide. However, since we can assume $l, l' \geq 1$, the previous uniqueness result says that both solutions coincide. Thus $\tilde{u}$ is C^l for all l hence smooth.

Observe that since $\tilde{u}$ is C^∞ in $(-\infty, T] \times \mathbb{R}^n$ and vanishes identically for $t < 0$, we in fact have that $\tilde{u} = 0$ on $\{t = 0\} \times \mathbb{R}^n$ (so this $\tilde{u}$ is not yet the solution to the Cauchy problem).

We now remove the assumption that f vanishes for $t \leq 0$. Let $0 \leq \varphi \leq l$ be a smooth function on $\mathbb{R}$ such that $\varphi(t) = 0, t \leq 0, \varphi(t) = 1, t \geq 1$. Set

$$f_\epsilon(t,x) = \varphi\left(\frac{t}{\epsilon}\right) f(t,x).$$

For any $\epsilon > 0$, we have by the above a solution u_ϵ to $Lu_\epsilon = f_\epsilon$ such that $u_\epsilon(t,x) = 0$ for $t \leq 0$. ($\leq$ and not only $<$ by the above). By the energy estimates

$$\|u_\epsilon - u_{\epsilon'}\|_{H^k(\Sigma_t)} \leq C\int_0^t \left|\varphi\left(\frac{t}{\epsilon}\right) - \varphi\left(\frac{t}{\epsilon'}\right)\right| \|f\|_{H^k(\Sigma_\tau)} d\tau.$$

Thus, u_ϵ, for any $t \in [0,T]$, u_ϵ converges to a limit in $H^k(\Sigma_t)$, for any $k, t \in [0,T]$. Solving for $\partial_t u_\epsilon$ in the equation, we get convergence of the time derivatives as long as $t > 0$. Hence, we have a smooth solution to $Lu = f$ in $(0,T) \times \mathbb{R}^n$. Let us show that this solution extends to $t = 0$. We have, for any k:

$$\|u_\epsilon\|_{H^k(\Sigma_t)} \leq C\int_0^t \|f_\epsilon\|_{H^k(\Sigma_\tau)} d\tau \leq C\int_0^t \|f\|_{H^k(\Sigma_\tau)} d\tau.$$

The RHS is independent of ϵ, thus the inequality holds for u too. Hence (taking k large), u and $D^{\vec{\alpha}}u$ converge to zero as $t \to 0^+$. We can solve for $\partial_t u$ in the equation and then get convergence of $\partial_t u$ as $t \to 0^+$ (to whatever it has to converge according to the expression for $\partial_t u$ determined by the equation). Inductively, we get convergence of higher time derivatives.

Thus, we obtain a $C^\infty([0,T] \times \mathbb{R}^n)$ function u solving $Lu = f$ in $[0,T] \times \mathbb{R}^n$ and satisfying $u = 0$ on $\{t = 0\} \times \mathbb{R}^n$. To obtain the correct initial condition, we take $C_c^\infty(\mathbb{R}) \quad 0 \leq \varphi \leq 1, \varphi = 1$ for $-1 \leq t \leq T+1$ and consider the problem for $u - \varphi u_0$.

□

Remark 13.15. Observe how we get the x-differentiability from "general" arguments and the t-differentiability by using the equation to "solve" for ∂_t (even in the case of the weak derivatives). Compare with what we did for elliptic equations to get the $\partial_n^2 u$ regularity on the boundary.

Corollary 13.16. *Consider the Cauchy problem for the FOSH system (13.1) in $\mathbb{R} \times \mathbb{R}^n$. Assume that $u_0 \in C_c^\infty(\mathbb{R}^n)$, that $f \in C^\infty(\mathbb{R} \times \mathbb{R}^n)$, and that A^μ and B are C^∞. Suppose that for each $[T_1, T_2] \subset \mathbb{R}$ there exist a, b such that*

$$\xi A^0 \xi \geq a|\xi|^2, |A^\mu| \leq b.$$

Then, there exists a unique $u \in C^\infty(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^d)$ solving (13.1).

Proof. As before, uniqueness follows from a theorem further above. For existence, we now do not necessarily have f compactly supported and A^μ, B with bounded derivatives, so we will use cut-off functions.

Fix $T > 0$. Let c_0 be as in the domain of dependence theorem. c_0 depends only on the constants a, b on $[0, T]$. Let $r \geq Tc_0 + 1$ and set

$$\mathcal{C}_r := \left\{ (t, x) \in [0, T] \;\middle|\; x \in B_{r - c_0 t}(0) \right\}.$$

Let $\varphi \in C_c^\infty(\mathbb{R} \times \mathbb{R}^n)$, $0 \leq \varphi \leq 1$, $\varphi = 1$ on $\mathcal{C}_{2r+2c_0T}$, and $\psi \in C_c^\infty(\mathbb{R}^n)$ satisfy $\psi = 1$ on $B_r(0)$, $\psi = 0$ on $B_{2r}(0)$.

Set (see picture below)

$$\tilde{A}^0 = \varphi A^0 + (1 - \varphi) A^0(0, 0) \,,$$
$$\tilde{A}^i = \varphi A^i, \quad \tilde{B} = \varphi B,$$
$$\tilde{u}_0 = \psi u_0, \quad \tilde{f} = \varphi \psi f$$

and consider

$$\tilde{A}^\mu \partial_\mu \tilde{u} + \tilde{B} \tilde{u} = \tilde{f} \text{ in } [0, T] \times \mathbb{R}^n,$$
$$\tilde{u} = \tilde{u}_0 \text{ on } \{t = 0\} \times \mathbb{R}^n.$$

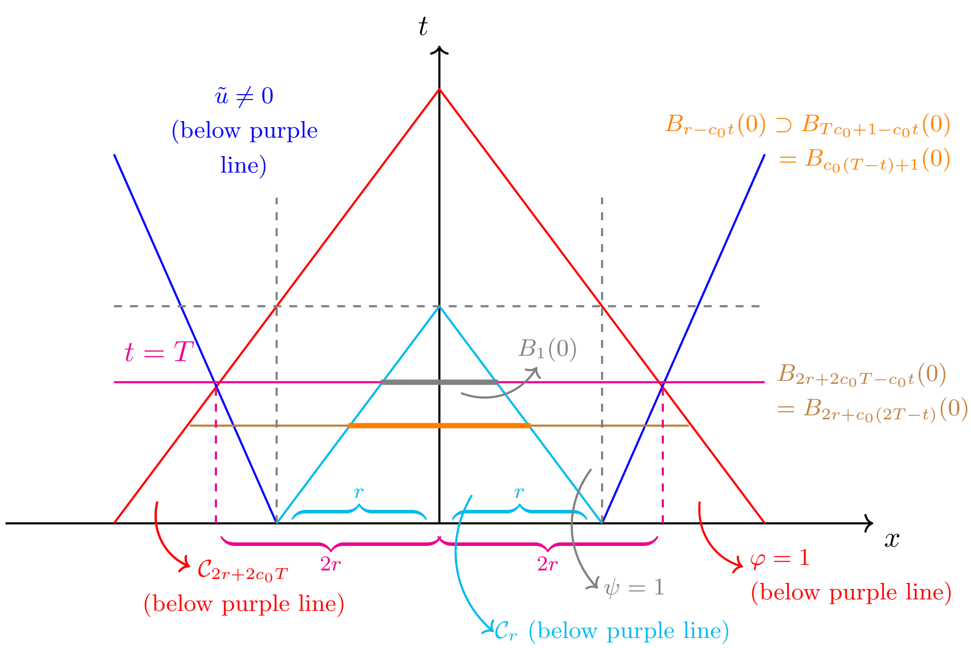


FIGURE 32. Visual Demonstration

Let $\tilde{u}$ be the smooth solution. Since

$$\begin{aligned}\xi\tilde{A}^0\xi &= \varphi\xi A^0\xi + (1-\varphi)\xi A^0(0,0)\xi \\ &\geq \varphi a|\xi|^2 + (1-\varphi)a|\xi|^2 \\ &= a|\xi|^2 \\ |\tilde{A}^i| &\leq |A^i| \leq b,\end{aligned}$$

we can apply the domain of dependence property to $\tilde{u}$ with the same constant c_0. Thus $\tilde{u}$ vanishes outside $B_{2r+c_0t}(0), t \in [0,T]$. In the region where $\tilde{u}$ does not vanish, $\tilde{A}^\mu = A, \tilde{B} = B$, so $\tilde{u}$ solve

$$\begin{aligned}A^\mu\partial_\mu\tilde{u} + B\tilde{u} = \tilde{f}, \\ \tilde{u} = \tilde{u}_0,\end{aligned}$$

in $[0,T]\times\mathbb{R}^n$. In $\mathcal{C}_r$, $\tilde{f} = f$, so $\tilde{u}$ solves the original equation in $\mathcal{C}_r$.

Next, repeat the argument with $\phi > r$, obtaining a solution u' to the original equations in $\mathcal{C}_\phi$ that agrees with $\tilde{u}$ in $\mathcal{C}_r$. In this way, we can repeat the argument and use uniqueness to obtain a solution in $[0,\infty)\times\mathbb{R}^n$.

To obtain the solution for $(-\infty,0]$, we reverse time.

□

Remark 13.17. We will often refer to "reversing time" arguments, so let us write the details here once. Set $\tau := -t$,

$$\begin{aligned}\tilde{A}^\mu(t,x) &:= A^\mu(-t,x), \\ \tilde{B}(t,x) &:= \tilde{B}(t,x) \\ \tilde{f}(t,x) &:= f(t,x)\end{aligned}$$

Consider

$$\begin{aligned}\tilde{L}\tilde{u}(t,x) &= \tilde{A}^0(t,x)\partial_t\tilde{u}(t,x) - \tilde{A}^i(t,x)\partial_i\tilde{u}(t,x) - \tilde{B}(t,x)\tilde{u}(t,x) \\ &= -\tilde{f}(t,x).\end{aligned}$$

For $t \geq 0$, we obtain a solution by the above. Then, setting $u(-t,x) := \tilde{u}(t,x)$, $\partial_t u(-t,x) = -\partial_t\tilde{u}(t,x)$,

$$\begin{aligned}-\tilde{A}(t,x)\partial_t u(-t,x) - \tilde{A}^i(t,x)\partial_i u(-t,x) - \tilde{B}(t,x)u(-t,x) = -\tilde{f}(t,x) \\ A(-t,x)\partial_t u(-t,x) + \tilde{A}^i(-t,x)\partial_i u(-t,x) + B(-t,x)u(-t,x) = f(-t,x)\end{aligned}$$

which is the original equation evaluated at $(-t,x)$.

13.2. **Linear hyperbolic/wave equations.**

Definition 13.18. A **Lorentz matrix** g is a $(n+1)\times(n+1)$ symmetric invertible matrix with the following property. Denoting the components ofl; /;;g by $g_{\mu\nu}$, μ, $\nu = 0,...,n$, $g_{00} < 0$ and $g_{ij}, i,j = 1,...,n$, are the components of a positive definite matrix. We denote the components of g^{-1} by $g^{\mu\nu}$. Observe that g^{-1} is also a Lorentzian matrix, i.e., $g^{00} < 0$ and g^{ij} are the components of a positive definite matrix.

A **Lorentzian metric in $[0,T]\times\mathbb{R}^n$** is a map of g from $[0,T]\times\mathbb{R}^n$ to the set of Lorentzian matrices with **uniform bounds**

$$\begin{aligned}g_{00}(t,x) \leq C < 0, \quad g_{ij}\xi^i\xi^j \geq C|\xi|^2, \\ g^{00}(t,x) \leq C < 0, \quad g^{ij}(t,x)\xi_i\xi_j \geq C|\xi|^2\end{aligned}$$

for all $(t,x) \in [0,T]\times\mathbb{R}^n$ and all $\xi \in \mathbb{R}^n$.

Definition 13.19. A second order linear system of **hyperbolic PDEs**, a.k.a. **a (system of) linear wave equation(s)** in $[0, T] \times \mathbb{R}^n$ is a system of the form

$$g^{\mu\nu}\partial_\mu\partial_\nu u + a^\mu\partial_\mu u + bu = f,$$

where $a^\mu, b : [0, T] \times \mathbb{R}^n \to d \times d$ matrices, $f : [0, T] \times \mathbb{R}^n \to \mathbb{R}^d$, g is a Lorentzian metric, and $u : [0, T] \times \mathbb{R}^n \to \mathbb{R}^d$ is the unknown.

Theorem 13.20. ***(Basic Energy Estimate for Wave Equations).*** *In the above (system) linear wave equation, assume that g, a^μ, b and f are smooth and all their derivatives are bounded. Let u be a smooth solution with the property that for each $T > 0$ there is a compact set $\kappa \subset \mathbb{R}^n$ with $u(t, x) = 0, t \in [0, T], x \notin \kappa$.*

Then

$$\sqrt{E(t)} \leq \left(\sqrt{E(0)} + C\int_0^t \|f(\tau, \cdot)\|_{L^2(\Sigma_\tau)} d\tau\right) e^{ct}$$

for some constant independent of u, where

$$E(t) := \frac{1}{2}\int_{\Sigma_t} (-g^{00}|\partial_t u|^2 + g^{ij}\partial_i u\partial_j u + |u|^2)dx$$

($g^{ij}\partial_i u\partial_j u = g^{ij}\partial_i u^T\partial_j u$, but as before we omit the transpose sign in u^T).

Proof. We have

$$\begin{aligned}
\partial_t E =& \int_{\Sigma_t} \left(-g^{00}\partial_t u\partial_t^2 u + g^{ij}\partial_i u\partial_j\partial_t u\right) \\
&+ \frac{1}{2}\int_{\Sigma_t} \left(-\partial_t g^{00}|\partial_t u|^2 + \partial_t g^{ij}\partial_i u\partial_j u + 2u\partial_t u\right) =: I_1 + I_2.
\end{aligned}$$

$$\begin{aligned}
I_1 =& \int_{\Sigma_t} \left(-g^{00}\partial_t u\partial_t^2 u + g^{ij}\partial_i u \underbrace{\partial_j}_{\text{by parts}} \partial_t u\right) \\
=& -\int_{\Sigma_t} \partial_t u\left(g^{00}\partial_t^2 u + g^{ij}\partial_i\partial_j u\right) - \int_{\Sigma_t} \partial_j g^{ij}\partial_i u\partial_t u \\
=& g^{\mu\nu}\partial_\mu\partial_\nu u - 2g^{0i}\partial_t\partial_i u \\
=& f - a^\mu\partial_\mu u - bu - 2g^{0i}\partial_t\partial_i u \\
=& 2\int_{\Sigma_t} g^{0i}\partial_t u\partial_i\partial_t u - \int_{\Sigma_t} \partial_t u\left(f - a^\mu\partial_\mu u - bu\right) - \int_{\Sigma_t} \partial_j g^{ij}\partial_i u\partial_t u \\
=& \int_{\Sigma_t} g^{0i}\underset{\text{by parts}}{\partial_i \left(|\partial_t u|^2\right)} - \int_{\Sigma_t} \partial_t u\left(f - a^\mu\partial_\mu u - bu\right) - \int_{\Sigma_t} \partial_j g^{ij}\partial_i u\partial_t u \\
=& -\int_{\Sigma_t} \left(\partial_i g^{0i}|\partial_t u|^2 - \partial_t u\left(f - a^\mu\partial_\mu u - bu\right)\right) - \int_{\Sigma_t} \partial_j g^{ij}\partial_i u\partial_t u \\
\leq& CE + C\sqrt{E}\|f\|_{L^2(\Sigma_t)}
\end{aligned}$$

since, under our assumptions

$$\frac{1}{C}E \leq \int_{\Sigma_t} \left(|\partial_t u|^2 + |\nabla u|^2 + |u|^2\right) \leq CE.$$

Using this we also obtain $I_2 \leq E$ so

$$\partial_t E \leq CE + \sqrt{E}\|f\|_{L^2(\Sigma_t)}$$

and the result follows by dividing by $\sqrt{E}$ and using Grönwall.

$\square$

Theorem 13.21. *Let $g_I, I = 1, ..., d$, be smooth Lorentzian metrics in $[0, \infty) \times \mathbb{R}^n$ such that for each $T > 0$, g_I satisfies uniform bounds in the sense of the definition of Lorentzian metrics in $[0, T] \times \mathbb{R}^n$. Let $a_J^{I\mu}, b_J^I, f^I \in C^\infty([0, \infty) \times \mathbb{R}^n), I = 1, .., d$. Let $u_0^I, u^I \in C^\infty(\mathbb{R}^n), I = 1, ..., d$. Consider the Cauchy problem*

$$\begin{aligned} g_I^{\mu\nu}\partial_\mu\partial_\nu u^I + a_J^{I\mu}\partial_\mu u^J + b_J^I u^I = f^I \text{ in } [0,\infty)\times\mathbb{R}^n, \\ u^I(0,\cdot) = u_0^I \text{ on } \{t=0\}\times\mathbb{R}^n, \\ \partial_t u^I(0,\cdot) = u_1^I \text{ on } \{t=0\}\times\mathbb{R}^n, \end{aligned}$$

where $I, J = 1, ..., d$ and there is a sum over J (but not over I). Then, there exists a unique smooth solution $u \in C^\infty([0, \infty) \times \mathbb{R}^n, \mathbb{R}^d)$. If the data has compact support and $f_I(t, x) = 0$ for $t \in [0, T], x \notin \kappa =$ compact, then there exists a compact set $\tilde{\kappa} \subset \mathbb{R}^n$ such that $u(t, x) = 0$ if $t \in [0, T], x \notin \tilde{\kappa}$. Moreover, the following domain of dependence property holds: given $T > 0$, there exists a $c_0 > 0$ such that if u_0^I, u_1^I vanish $B_r(x_0)$ and f^I vanishes in $\mathcal{C}_{r,x_0,c_0}$, then u vanishes in $\mathcal{C}_{r,x_0,c_0}$. This last statement requires only g_I to be C^1 with uniform bounds, $a_J^{I\mu}, b_J^I$ and f to be continuous, and u to be a C^2 solution.

Proof. This can be proven with ideas very similar to what we used for FOSH linear systems. For existence, we derive higher order energy estimates by considering the equation satisfies by $D^{\vec{\alpha}}u$, and then involve functional analytic methods. For uniqueness and the domain of dependence property, we integrate over $\mathcal{C}_{r,x_0,c_0}$, choose c_0 appropriately, and analyze the boundary integrals. We will, however, take a shortcut, as follows.

Set for each $I = 1, ..., d$, and $i, j, k = 1, ..., n$,

$$\begin{aligned} v^I = (v_i^I, ..., v_{n+2}^I) := (\partial_1 u^I, ..., \partial_n u^I, \partial_t u^I, u^I), \\ A_{ij}^{I0} := g_I^{ij}, \\ A_{n+1,n+1}^{I0} = A_{n+2,n+2}^{I0} := 1, \\ A_{i,n+1}^{Ik} = A_{n+1,i}^{Ik} := g_I^{ik}, \\ A_{n+1,n+1}^{Ik} := 2g_I^{0k}. \end{aligned}$$

From this we obtain $(n+2) \times (n+2)$ matrices A^{I0} and A^{Ik}, where the entries that have not been defined above are set to zero.

Set

$$\begin{aligned} h_{n+1}^I := -f^I, \\ d_{J\,n+1,i}^I := -a_J^{I\,i}, \\ d_{J\,n+1,n+1}^I := -a_J^{I\,0}, \\ d_{J\,n+1,n+2}^I := -b_J^I, \\ d_{J\,n+2,n+1}^I := -\delta_J^I, \end{aligned}$$

and the remaining components are set to zero.

Then, if u is a solution to the wave equation, v satisfies the FOSH linear system:

$$A^{I0}\partial_t v^I - A^{Ii}\partial_i v^I + d_J^I v^J = h^I \tag{13.2}$$

(no sum over I). Moreover, $v(0,\cdot)$ satisfies

$$\begin{aligned} v^I(0,\cdot) &= (v_1(0,\cdot),...,v_n(0,\cdot),v_{n+1}(0,\cdot),v_{n+2}(0,\cdot)) \\ &= (\partial_1 u^I(0,\cdot),...,\partial_n u^I(0,\cdot),\partial_t u^I(0,\cdot),u^I(0,\cdot)). \end{aligned}$$

Observe that the initial data has the property that

$$\partial_i v^I_{n+2}(0,\cdot) = v^I_i(0,\cdot). \tag{13.3}$$

From our assumptions, we can apply our results on FOSH systems to obtain a smooth solution v with a domain of dependence. Assume further that the initial data satisfies (13.3). From (13.2), taking the j-component:

$$\begin{aligned} (A^{I0}\partial_t v^I)_j - (A^{Ii}\partial_i v^I)_j + (d^I_J v^J)_j &= \overbrace{(h^I)_j}^{=0} \\ A^{I0}_{jM}(\partial_t v^I)_M - A^{Ii}_{jM}(\partial_i v^I)_M + \overbrace{d^I_{J\ jL}}^{=0}(v^J)_L &= 0 \\ \overbrace{A^{I\,0}_{jk}}^{=g^{jk}_I}(\partial_t v^I)_k + \overbrace{A^{I\,0}_{j\ n+1}}^{=0}(\partial_t v^I)_{n+1} + \overbrace{A^{I\,0}_{j\ n+2}}^{=0}(\partial_t v^I)_{n+2} & \\ \overbrace{-A^{Ii}_{jk}}^{=0}(\partial_i v^I)_k - \overbrace{A^{Ii}_{jn+1}}^{=g^{ij}_I}(\partial_i v^I)_{n+1} - \overbrace{A^{Ii}_{jn+2}}^{=0}(\partial_i v^I)_{n+2} &= 0 \\ g^{jk}_I \partial_t v^I_k &= g^{ij}_I \partial_i v^I_{n+1} \\ \to \overbrace{m_{lj}\, g^{jk}_I}^{=\delta^k_l}\ \partial_t v^I_k &= \overbrace{m_{lj}\, g^{ij}_I}^{=\delta^i_l \partial_i v^I_{n+1}}\ \partial_i v^I_{n+1} \\ \to \partial_t v^I_i &= \partial_i v^I_{n+1}, \end{aligned}$$

where m is the inverse of (g^{ij}_I) (not necessarily $m_{ij} = g_{ij}$).

Similarly, taking the $n+2$ component

$$\begin{aligned} (A^{I0}\partial_t v^I)_{n+2} - (A^{Ii}\partial_i v^I)_{n+2} + (d^I_J v^J)_{n+2} &= \overbrace{(h^I)_{n+2}}^{=0} \\ A^{I0}_{n+2\ M}(\partial_t v^I)_M - A^{Ii}_{n+2\ M}(\partial_i v^I)_M + \overbrace{d^I_{J\ n+2L}}^{=-\delta^J_I,\,L=n+1}(v^J)_L &= 0 \\ \overbrace{A^{I0}_{n+2,k}}^{=0}(\partial_t v^I)_k + \overbrace{A^{I0}_{n+2,n+1}}^{=0}(\partial_t v^I)_{n+1} + \overbrace{A^{I0}_{n+2,n+2}}^{=0}(\partial_t v^I)_{n+2} & \\ -\overbrace{A^{Ii}_{n+2,k}}^{=0}(\partial_i v^I)_k - \overbrace{A^{Ii}_{n+2,n+1}}^{=0}(\partial_i v^I)_{n+1} - \overbrace{A^{Ii}_{n+2,n+2}}^{=0}(\partial_i v^I)_{n+2} &= 0 \\ \implies \partial_t v^I_{n+2} &= v^I_{n+1}. \end{aligned}$$

Set $u^I := v^I_{n+2}$. Then by (13.3) $\partial_i u^I(0,\cdot) = v^I_i(0,\cdot)$.

Then

$$\begin{aligned}
\partial_i u^I(t,\cdot) &= \partial_i u^I(0,\cdot) + \int_0^t \partial_t \underbrace{\partial_i u^I(\tau,\cdot)}_{=v^I_{n+2}(\tau,\cdot)} d\tau \\
&= v^I_i(0,\cdot) + \int_0^t \partial_i \underbrace{\partial_t v^I_{n+2}(\tau,\cdot)}_{=v^I_{n+1}(\tau,\cdot)} d\tau \\
&= v^I_i(0,\cdot) + \int_0^t \underbrace{\partial_i v^I_{n+1}(\tau,\cdot)}_{=\partial_t v^I_i(\tau,\cdot)} d\tau \\
&= v^I_i(t,\cdot)
\end{aligned}$$

Also

$$\partial_t u = \partial_t v_{n+2} = v_{n+1}.$$

Hence, v is of the form stated in terms of u, and u satisfies the original equation.

□

14. Local existence and uniqueness for quasilinear wave equations

Our goal is to study systems like the above linear wave equation where now g (and a, b, etc.) depend on u. For this, we used to make some specific choices about this dependence.

Definition 14.1. We say that a C^k map

$$g : \mathbb{R}^{nd+2d+n+1} \to \text{space of } (n+1)\times(n+1) \text{ Lorentzian matrices}$$

is a $(\mathbf{C^k}, \mathbf{n}, \mathbf{d})$**−admissible metric**, or **admissible metric** for short, if:

- For every multi-index $\alpha = (\alpha_1, ..., \alpha_{nd+2d+n+1})$ such that $|\alpha| \le k$ and every compact interval $I = [T_1, T_2]$ there exists a continuous increasing function
$$h_{I,\alpha} : \mathbb{R} \to \mathbb{R}$$
such that
$$|D^\alpha g_{\mu\nu}(t,x,\xi)| \le h_{I,\alpha}(|\xi|)$$
for all $\mu, \nu = 0, ..., n,\ x \in \mathbb{R}^n,\ t \in I,\ \xi \in \mathbb{R}^{nd+2d}$.
- For every compact interval $I = [T_1, T_2]$, there exist $a_1, a_2, a_3 > 0$ such that for every $(t,x,\xi) \in I \times \mathbb{R}^{nd+2d+n}$
$$g_{00}(t,x,\xi) \le -a_1, \ \ g_{ij}(t,x,\xi)\,\xi^i\,\xi_j \ge a_2|\xi|^2 \text{ for all } \xi \in \mathbb{R}^n,$$
and
$$\sum_{\mu,\nu=0}^n |g_{\mu\nu}(t,x,\xi)| \le a_3.$$

It follows from the above definitions that if g is (C^0, n, d)-admissible then $g(\cdot,\cdot,\xi)$ is a Lorentzian metric.

Definition 14.2. A map $f : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^d$ has **local compact support in x** if for any compact interval $[T_1, T_2]$ there exists a compact set $\kappa \subset \mathbb{R}^n$ such that

$$f(t,x) = 0,\ t \in [T_1,T_2],\ x \notin \kappa.$$

A smooth function $f : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^d$ that has local compact support in x can be regarded as an element of

$$C^m(\mathbb{R}, H^k(\mathbb{R}^n, \mathbb{R}^d))$$

for any $m, k \geq 0$ integers. This is not necessarily the case if f has the property that for any fixed $t, f(t, \cdot)$ has compact support. Consider $\varphi \in C_c^\infty(\mathbb{R}^n), \varphi \notin 0$. Set

$$f(t,x) := \begin{cases} \varphi(x^1 - \frac{1}{t}, x^2, .., x^n) & t > 0, \\ 0 & t \leq 0. \end{cases}$$

Then f is smooth for $t > 0$ and for $t < 0$. For each $(0, x)$, there exists a neighborhood $\mathcal{U} \ni (0, x)$ such that $f = 0$ in $\mathcal{U}$. Therefore, f is smooth. For fixed t, $f(t, \cdot)$ has compact support. For $t \leq 0$,

$$\|f(t,\cdot)\|_{L^2(\mathbb{R}^n)} = 0$$

but for $t > 0$

$$\|f(t,\cdot)\|_{L^2} > 0,$$

so $f \notin C^0(\mathbb{R}, H^0(\mathbb{R}^n, \mathbb{R}))$.

Definition 14.3. A C^k-map

$$f : \mathbb{R}^{nd+2d+n+1} \to \mathbb{R}^d$$

is called a $(\mathbf{C^k}, \mathbf{n}, \mathbf{d})$**-admissible nonlinearity**, or **admissible nonlinearity** for short, if:

- For every multi-index $\alpha = (\alpha_1, ..., \alpha_{nd+2d+n+1})$ such that $|\alpha| \leq k$ and every compact interval $I = [T_1, T_2]$ there exists a continuous increasing function

  $$h_{I,\alpha} : \mathbb{R} \to \mathbb{R}$$

  such that

  $$|D^\alpha f(t, x, \xi)| \leq h_{I,\alpha}(|\xi|)$$

  for all $t \in I,\ x \in \mathbb{R}^n,\ \xi \in \mathbb{R}^{nd+2d}$.
- The function of (t, x) defined by $f(t, x, 0)$ has a local compact support in x.

Let $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ and $u : \Omega \to \mathbb{R}^d$ be differentiable. Let g be an admissible metric and f an admissible nonlinearity. Define $g[n]$ as the Lorentzian metric

$$g[u](t, x,) = g(t, x, u(t, x), \partial_t u(t, x), ..., \partial_n u(t, x))$$

and $f[u]$ as the function

$$f[u](t, x) = f(t,\ x,\ u(t, x),\ \partial_t u(t, x), ...,\ \partial_n u(t, x)).$$

Note that

$$(\overbrace{t, x}^{\in \mathbb{R}\times\mathbb{R}^n}, \underbrace{\overbrace{u(t,x)}^{\in\mathbb{R}^d}, \overbrace{\partial_t u(t,x)}^{\in\mathbb{R}^d}, \overbrace{\partial_1 u(t,x), ..., \partial_n(t,x)}^{\in\mathbb{R}^{nd}}}_{\xi}) \in \mathbb{R}^{nd+2d+n+1}$$

explains the above choices.

Definition 14.4. A $(\mathbf{n}, \mathbf{d})$ **admissible majorizer**, or simply **majorizer**, is a map that associates to each (C^∞, n, d)-admissible metric g, (C^∞, n, d)-admissible nonlinearity f and compact interval

$$I = [T_1, T_2],$$

a continuous function

$$z_I[g, f] : \mathbb{R}^m \to [0, \infty),$$

where m is fixed, with the property that

$$z_{I_1}[g, f] \leq z_{I_2}[g, f]$$

whenever $I_1 \subset I_2$.

Definition 14.5. A $(\mathbf{n}, \mathbf{d})$ **admissible constant**, or simply **admissible constant**, is a map that associates to each (C^∞, n, d)-admissible metric g, (C^∞, n, d)-admissible nonlinearity f and compact interval $I = [T_1, T_2]$, a real number $C_I[g, f] > 0$, with the property that $C_{I_1}[g, f] \leq C_{I_2}[g, f]$ if $I_1 \subset I_2$.

We will often omit the arguments $[g, f]$ and write z_I, C_I.

Notation 14.6. We will henceforth write $g_{\mu\nu}$ for $g_{\mu\nu}[u]$ and $g^{\mu\nu}$ for $g^{\mu\nu}[u]$. Similarly, we will write f for $f[u]$. Or sometimes we write $g_u^{\mu\nu}$, f_u, etc. Also, we can assume, without loss of generality, that $g^{00} = -1$.

We will study the Cauchy problem for the quasilinear wave equation

$$\begin{cases} g^{\mu\nu}\partial_\mu\partial_\nu u = f \text{ in } [0, T] \times \mathbb{R}^n, \\ u(0, \cdot) = u_0 \text{ on } \{t = 0\} \times \mathbb{R}^n, \\ \partial_t u(0, \cdot) = u_1 \text{ on } \{t = 0\} \times \mathbb{R}^n. \end{cases} \tag{14.1}$$

(with $g^{\mu\nu} = g_u^{\mu\nu}, f = f_u$).

Theorem 14.7. ***(uniqueness)***. *Let g be a C^1 admissible metric and f a C^1 admissible nonlinearity. Let u and v be two solutions to (14.1) with $u_0 = v_0, u_1 = v_1$. Then $u = v$.*

Proof. Write

$$g_u^{\mu\nu}\partial_\mu\partial_\nu u = f_u$$
$$g_v^{\mu\nu}\partial_\mu\partial_\nu v = f_v$$

so

$$g_u^{\mu\nu}\partial_\mu\partial_\nu(u - v) = (g_v^{\mu\nu} - g_u^{\mu\nu})\partial_\mu\partial_\nu v + f_u - f_v.$$

By the fundamental theorem of calculus

$$\begin{aligned} f_u - f_v &= \int_0^1 \frac{d}{dt}\big(f(tu + (1-t)v)\big)dt \\ &= \int_0^1 \nabla_\xi f\big(tu + (1-t)v\big) \cdot (\xi_u - \xi_v)dt \\ &= \tilde{f}(u - v) + \hat{f}^\mu\partial_\mu(u - v) \end{aligned}$$

for some continuous functions $\tilde{f}, \hat{f}^\mu$, and similarly

$$g_u^{\mu\nu} - g_v^{\mu\nu} = \tilde{g}^{\mu\nu}(u - v) + \hat{g}^{\mu\nu,\lambda}\partial_\lambda(u - v),$$

for some continuous function $\tilde{g}^{\mu\nu}$, $\hat{g}^{\mu\nu,\lambda}$. Thus, with $w := u - v$,

$$g_u^{\mu\nu}\partial_\mu\partial_\nu w = \tilde{g}^{\mu\nu}\partial_\mu\partial_\nu v w + \hat{g}^{\mu\nu,\lambda}\partial_\mu\partial_\nu v\partial_\lambda w + \tilde{f}w + \hat{f}^\mu\partial_\mu w.$$

This is a linear wave equation for w for which our uniqueness results apply.

□

Notation 14.8. Denote

$$\mathcal{M}_k[u](t) := \|u\|_{H^{k+1}(\Sigma_t)} + \|\partial_t u\|_{H^k(\Sigma_t)}$$

$$\mathcal{N}[u](t) := \sum_{|\vec{\alpha}|+j\leq 2} \|D^{\vec{\alpha}}\partial_t^j u\|_{L^\infty(\Sigma_t)}.$$

Theorem 14.9. *Let g be a C^∞ admissible metric and f a C^∞ admissible nonlinearity. Let $u_0, u_1 \in C_c^\infty(\mathbb{R}^n, \mathbb{R}^d)$, $v \in C^\infty(\mathbb{R}\times\mathbb{R}^n, \mathbb{R}^d)$. Suppose that v has local compact support in x. Set $g_v = g[v], f_v = f[v]$, and let u be the solution to*

$$\begin{aligned} g_v^{\mu\nu}\partial_\mu\partial_\nu u &= f_v \text{ in } \mathbb{R}\times\mathbb{R}^n, \\ u(0,\cdot) &= u_0 \text{ on } \{t=0\}\times\mathbb{R}^n, \\ \partial_t u(0,\cdot) &= u_1 \text{ on } \{t=0\}\times\mathbb{R}^n \end{aligned}$$

(which exists and is of local compact support in x).

Set

$$E_k[v,u](t) := \frac{1}{2}\sum_{|\vec{\alpha}|\le k}\int_{\sum_t}\left(-g_v^{00}|D^{\vec{\alpha}}\partial_t u|^2 + g_v^{ij}D^{\vec{\alpha}}\partial_i u\cdot D^{\vec{\alpha}}\partial_j u + |D^{\vec{\alpha}}u|^2\right)dx.$$

Let $I = [0,T]$.
Then, there exist admissible majorizers and constants $z_{1,I}$, $z_{2,I}$, $c_{1,I}$, $c_{2,I}$, $c_{3,I}$ such that

$$\begin{aligned} &\mathcal{M}_k[u](t) \le c_{1,I}\,\mathcal{M}_k[u](0) + \int_0^t\left[c_{2,I} + z_{1,I}\,(\mathcal{N}[v])\,((1+\mathcal{N}[u])\,\mathcal{M}_k[v] + \mathcal{M}_k[u])\right]d\tau, \\ &\partial_t E_k[v,u] \le c_{3,I} + z_{2,I}(\mathcal{N}[u],\mathcal{N}[v])\left(\mathcal{M}_k^2[v] + E_k[v,u]\right). \end{aligned}$$

If $f(t,x,0) = 0$, $c_{2,I}$ can be omitted. (Under our conventions, $g_v^{00} = -1$, but it is convenient to write it for keeping track of things.)

We will sometimes write $g(t,x,v)$ for $g(t,x,v,\partial v)$ etc. $\partial = (\partial, D), D =$ spatial.

Proof. Note that

$$\frac{1}{C_I}E_k^{\frac{1}{2}}[v,\tilde{v}](t) \le \mathcal{M}_k[\tilde{v}](t) \le C_I E_k^{\frac{1}{2}}[v,\tilde{v}](t)$$

provided $\tilde{v}$ is such that this is defined. To simplify the notation, write $E_k = E_k[u,v], E = E_0$. Compute

$$\begin{aligned} \partial_t E &= \int_{\sum_t}\left(-g_v^{00}\partial_t u\partial_t^2 u + g_v^{ij}\partial_i u\underbrace{\partial_j\partial_t u}_{\text{by parts}}\right) + \frac{1}{2}\int_{\sum_t}\left(-\partial_t g_v^{00}|\partial_t u|^2 + \partial_t g_v^{ij}\partial_i u\partial_j u + 2u\partial_t u\right) \\ &= -\int_{\sum_t}\left(\underbrace{g_v^{00}\partial_t^2 u + g_v^{ij}\partial_i\partial_j u}_{=g_v^{\mu\nu}\partial_\mu\partial_\nu u - 2g_v^{i0}\,\partial_i\partial_t u}\right)\partial_t u - \int_{\sum_t}\partial_i g_v^{ij}\partial_j u\partial_t u + \frac{1}{2}\int_{\sum_t}\left(-\partial_t g_v^{00}|\partial_t u|^2 + \partial_t g_v^{ij}\partial_i u\partial_j u + 2u\partial_t u\right) \\ &= -\int_{\sum_t} f\partial_t u + 2\int_{\sum_t} g_v^{i0}\overbrace{\partial_i\partial_t u\partial_t u}^{=\frac{1}{2}\partial_i(|\partial_t u|^2),\text{ by parts}} + \frac{1}{2}\int_{\sum_t}\left(-\partial_t g_v^{00}|\partial_t u|^2 + \partial_t g_v^{ij}\partial_i u\partial_j u + 2u\partial_t u - 2\partial_i g_v^{ij}\partial_j u\partial_t u\right) \\ &= -\int_{\sum_t} f\partial_t u + \frac{1}{2}\int_{\sum_t}\left(-\partial_t g_v^{00}|\partial_t u|^2 + \partial_t g_v^{ij}\partial_i u\partial_j u + 2u\partial_t u - 2\partial_i g_v^{ij}\partial_j u\partial_t u - 2\partial_i g_v^{i0}|\partial_t u|^2\right) \end{aligned}$$

Since $|D^\alpha g_{\mu\nu}(t,x,\xi)| \le h_{I,\alpha}(|\xi|)$ and g depends on up to ∂v,

$$|\partial g_v^{\mu\nu}| \le z_I(\mathcal{N}_I[v]).$$

Thus

$$\partial_t E \le z_I(\mathcal{N}[v])E + C_I\|f_v\|_{L^2\sum_t}\sqrt{E}$$

Next, since $E_k[v,u] = \sum_{|\vec{\alpha}|\le k}E[v, D^{\vec{\alpha}}u]$, differentiate the equation:

$$g_v^{\mu\nu}\partial_\mu\partial_\nu D^{\vec{\alpha}}u = D^{\vec{\alpha}}f_v + [g_v^{\mu\nu}\partial_\mu\partial_\nu, D^{\vec{\alpha}}]u.$$

Apply the above energy inequality to this to get

$$\partial_t E_k \leq z_I(\mathcal{N}(v))E_k + C_I \|f_v\|_{H^k(\sum_t)}\sqrt{E_k} + C_I \left\| [g_v^{\mu\nu}\partial_\mu\partial_\nu, D^{\vec{\alpha}}]u \right\|_{L^2(\sum_t)} \sqrt{E_k}.$$

To estimate the commutator,

$$g_v^{\mu\nu}\partial_\mu\partial_\nu D^{\vec{\alpha}}u - D^{\vec{\alpha}}(g_v^{\mu\nu}\partial_\mu\partial_\nu u)$$

note that it is a sum of terms of the form (up to constants)

$$D^{\vec{\beta}}\partial_i g_v^{\mu\nu} D^{\vec{r}}\partial_\mu\partial_\nu u$$

where $|\vec{\beta}| + |\vec{r}| = |\vec{\alpha}| - 1$ (the ∂_i is there because at least one derivative falls on $g_v^{\mu\nu}$). Write

$$D^{\vec{\beta}}\partial_i g_v^{\mu\nu} D^{\vec{r}}\partial_\mu\partial_\nu u = D^{\vec{\beta}}\partial_i(g_v^{\mu\nu} - g_0^{\mu\nu})D^{\vec{r}}\partial_\mu\partial_\nu u + D^{\vec{\beta}}\partial_i g_0^{\mu\nu} D^{\vec{r}}\partial_\mu\partial_\nu u,$$

$g_0 = g[v=0]$. Since $\partial_i g_v^{\mu\nu} = 0$ if $\mu = \nu = 0$, in $\partial_\mu\partial_\nu u$ at least one of the derivatives is spatial. Then, using some of our inequalities for Sobolev spaces,

$$\|D^{\alpha_1}u_1, ..., D^{\alpha_l}u_l\|_{L^2(\Omega)} \leq C\sum_{i=1}^{l}\|D^k u_i\|_{L^2(\Omega)}\prod_{j\neq i}\|u_j\|_{L^\infty(\Omega)}$$

$|\alpha_1| + ... + |\alpha_l| = k$, we find ($D =$ spatial derivative)

$$\begin{aligned}\|D^{\vec{\beta}}\overbrace{\partial_i(g_v^{\mu\nu} - g_0^{\mu\nu})}^{u_1} D^{\vec{r}}\overbrace{\partial_\mu\partial_\nu u}^{u_2}\|_{L^2(\sum_t)} &\leq C\|D^{k-1}D(g_v - g_0)\|_{L^2(\sum_t)}\|D\partial u\|_{L^\infty(\sum_t)} \\ &+ C\|D(g_v - g_0)\|_{L^\infty(\sum_t)}\|D^{k-1}D\partial u\|_{L^2(\sum_t)}\end{aligned}$$

Since v has local compact support in x,

$$g(t,x,v) = g(t,x,0) = g_0(t,x)$$

for $x \notin \kappa \subset \sum_t$. ($\kappa$ compact depending on T),

$$\|D(g_v - g_0)\|_{L^\infty(\sum_t)} \leq z_I(\mathcal{N}[v]).$$

Recalling another inequality:

$$\|F(\cdot,u)\|_{H^k(\mathbb{R}^n)} \leq \overbrace{C(\|u\|_{L^\infty(\mathbb{R}^n)})}^{\text{continuous, increasing}} \|u\|_{H^k(\mathbb{R}^n)}$$

for F such that $F(x,0) = 0$

$$|D_x^\alpha D_y^\alpha F(x,y)| \leq \underbrace{F_{\alpha,j}(|y|)}_{\text{continuous, increasing}},$$

we find $\big((g_v - g_0)(t,x,0) = 0\big)$

$$\begin{aligned}\|D^{k-1}D(g_v - g_0)\|_{L^2(\sum_t)} &\leq C\|g_v - g_0\|_{H^k(\sum_t)} \\ &\leq C(\|\partial v\|_{L^\infty})\|\partial v\|_{H^k(\sum_t)} \\ &\leq z_I(\mathcal{N}[v])\mathcal{M}_k[v]\end{aligned}$$

(we get ∂v because $g = g(\partial v)$). Therefore

$$\begin{aligned}\|D^{\vec{\beta}}\partial_i(g_v^{\mu\nu} - g_0^{\mu\nu})D^{\vec{r}}\partial_\mu\partial_\nu u\|_{L^2(\sum_t)} &\leq z_I(\mathcal{N}[v])\mathcal{M}_k[v]\mathcal{N}[u] + z_I(\mathcal{N}[v])\mathcal{M}_k[u] \\ &\leq z_I(\mathcal{N}[v])\big(\mathcal{M}_k[u] + \mathcal{N}[u] + \mathcal{M}_k[v]\big)\end{aligned}$$

Write $f_v = (f_v - f_0) + f_0$, $f_0(t,x) = f(t,x,0)$. Since f_0 has a compact local support in x, $\|f_0\|_{H^k(\sum_t)} \leq C_I$. Applying the above inequality for F to $f - f_0$:

$$\|f - f_0\|_{H^k(\sum_t)} \leq z_I(\mathcal{N}[v])\mathcal{M}_k[v].$$

Putting it all together:

$$\begin{aligned}\partial_t E_k \le & z_I(\mathcal{N}[v])E_k + C_I\|f_v\|_{H^k(\sum_t)}\sqrt{E_k}\\ & + \left\|[g_v^{\mu\nu}\partial_\mu\partial_\nu, D^{\vec{\alpha}}]u\right\|_{L^2(\sum_t)}\sqrt{E_k}\\ \le & z_I(\mathcal{N}[v])E_k + (C_I + z_I(\mathcal{N}[v])\mathcal{M}_k[v])\sqrt{E_k}\\ & + z_I(\mathcal{N}[v])(\mathcal{M}_k[u] + \mathcal{N}[u]\mathcal{M}_k[v])\sqrt{E_k}\\ \le & z_I(\mathcal{N}(v))E_k + C_I\sqrt{E_k}\\ & + z_I(\mathcal{N}[v])\big[\mathcal{M}_k[v] + \mathcal{N}[u]\mathcal{M}_k[v] + \mathcal{M}_k[u]\big]\sqrt{E_k}.\end{aligned}$$

Using $C_I\sqrt{E_k} \le CC_I^2 + CE_k$

$$\begin{aligned}\big(\mathcal{M}_k[v] + \mathcal{N}[u]\mathcal{M}_k[v]\big)\sqrt{E_k} &= \big(1+\mathcal{N}[u]\big)\mathcal{M}_k[v]\sqrt{E_k}\\ &\le C\big(1+\mathcal{N}[u]\big)^2\mathcal{M}_k^2[v] + CE_k\\ \mathcal{M}_k[u]\sqrt{E_k} &\le C_I E_k,\end{aligned}$$

we obtain the second inequality; for the first inequality, divide by $\sqrt{E_k}$, integrate, and use $\sqrt{E_k} \approx M_k$.

□

Lemma 14.10. *Let g be a C^∞ admissible metric and f a C^∞ admissible nonlinearity. Let $u_{0,i}$, $u_{1,i} \in C_c^\infty(\mathbb{R}^n), v_i \in C^\infty(\mathbb{R}\times\mathbb{R}^n, \mathbb{R}^d)$ have local compact support in $x, i = 1, 2$. Set $g_i = g[v_i], f_i = f[v_i]$, and let u_i be the solution to*

$$\begin{aligned} g_i^{\mu\nu}\partial_\mu\partial_\nu u_i &= f_i \text{ in } \mathbb{R}\times\mathbb{R}^n,\\ u_i(0,\cdot) &= u_{0,i} \text{ on } \{t=0\}\times\mathbb{R}^n,\\ \partial_t u_i(0,\cdot) &= u_{1,i} \text{ on } \{t=0\}\times\mathbb{R}^n\end{aligned}$$

Let $I = [0,T],\ v = v_2 - v_1,$ and $u = u_2 - u_1$.

Then, there exist majorizers $z_{1,I}$, $z_{2,I}$ and an admissible constant C_I such that, for $t \in I$,

$$\begin{aligned}\mathcal{M}_0[u](t) \le C_I \exp\left(\int_0^t z_{1,I}(\mathcal{N}[v_2])\,d\tau\right)\cdot\Bigg[&\mathcal{M}_0[u](0)\\ &+ \int_0^t z_{2,I}(\mathcal{N}[u_1], \mathcal{N}[v_1], \mathcal{N}[v_2])\,\mathcal{M}_0[v]\,d\tau\Bigg]\end{aligned}$$

Proof. u satisfies

$$\begin{aligned} g_{v_2}^{\mu\nu}\partial_\mu\partial_\nu u &= g_{v_2}^{\mu\nu}\partial_\mu\partial_\nu u_2 - g_{v_2}^{\mu\nu}\partial_\mu\partial_\nu u_1\\ &= g_{v_2}^{\mu\nu}\partial_\mu\partial_\nu u_2 - g_{v_1}^{\mu\nu}\partial_\mu\partial_\nu u_1 + (g_{v_1}^{\mu\nu} - g_{v_2}^{\mu\nu})\partial_\mu\partial_\nu u_1\\ &= f_{v_2} - f_{v_1} + (g_{v_2}^{\mu\nu} - g_{v_1}^{\mu\nu})\partial_\mu\partial_\nu u_1\end{aligned}$$

Set

$$E = \frac{1}{2}\int_{\sum_t}(-g_{v_2}^{00}|\partial_t u|^2 + g_{v_2}^{ij}\partial_i u\partial_j u + |u|^2)$$

and proceed as above to get

$$\begin{aligned}\partial_t E \le\ & z_I(\mathcal{N}(v_2))E + \|f_{v_2} - f_{v_1}\|_{L^2(\sum_t)}\sqrt{E}\\ & + \|(g_{v_2}^{\mu\nu} - g_{v_1}^{\mu\nu})\partial_\mu\partial_\nu u\|_{L^2(\sum_t)}\sqrt{E}\end{aligned}$$

(recall that in the basic estimate the term in power one in the energy is multiplied by a $\mathcal{N}(\cdot)$ term that comes from differentiating $g^{\mu\nu}_{(\cdot)}$, this is why we have $\mathcal{N}(v_2)$ here.)
The differences can be estimated using the fundamental theorem of calculus as we did for uniqueness; $\partial_\mu\partial_\nu u_1$ gives a $\mathcal{N}(u_1)$ term (recall that $\partial_t^2 u_1$ does not appear).

Thus,

$$\begin{aligned}\left\|(g^{\mu\nu}_{v_2} - g^{\mu\nu}_{v_1})\partial_\mu\partial_\nu u_1\right\|_{L^2(\sum_t)} &\leq z_1(\mathcal{N}(u_1))\|\underbrace{(g^{\mu\nu}_{v_2} - g^{\mu\nu}_{v_1})}_{\sim\int_0^1 \partial g(v,\partial v)}\|_{L^2(\sum_t)} \\ &\leq z_1\big(\mathcal{N}(u_1), \mathcal{N}(v_1), \mathcal{N}(v_2)\big)\left(\|v_2 - v_1\|_{L^2(\sum_t)} + \|\partial v_2 - \partial v_1\|_{L^2(\sum_t)}\right) \\ &\leq z_1\big(\mathcal{N}(u_1), \mathcal{N}(v_1), \mathcal{N}(v_2)\big)\mathcal{M}_0(v),\end{aligned}$$

Similarly

$$\|f_{v_2} - f_{v_1}\|_{L^2(\sum_t)} \leq z_I(\mathcal{N}(v_1), \mathcal{N}(v_2))\mathcal{M}_0(v).$$

Thus

$$\partial_t E \leq z_1(\mathcal{N}(v_2))E + z_I(\mathcal{N}(u_1), \mathcal{N}(v_1), \mathcal{N}(v_2))\mathcal{M}_0(v).$$

Dividing by $\sqrt{E}$, integrating, using $\sqrt{E} \approx \mathcal{M}_0[u]$ and involving Gronwall produces the result.

□

Theorem 14.11. *Let g be a C^∞ admissible metric and f a C^∞ admissible nonlinearity. Let $u_0 \in H^{k+1}(\mathbb{R}^n, \mathbb{R}^d), u_1 \in H^k(\mathbb{R}^n, \mathbb{R}^d)$, where $k > \frac{n}{2} + 1$. Then, there exists a $T > 0$ and a unique $u \in C^2_B([0,T] \times \mathbb{R}^n, \mathbb{R}^d)$ which is a solution to*

$$\begin{aligned} g^{\mu\nu}\partial_\mu\partial_\nu u &= f \text{ in } [0,T] \times \mathbb{R}^n, \\ u(0,\cdot) &= u_0 \text{ on } \{t=0\} \times \mathbb{R}^n, \\ \partial_t u(0,\cdot) &= u_1 \text{ on } \{t=0\} \times \mathbb{R}^n \end{aligned}$$

Moreover, u has regularity

$$\begin{aligned} u &\in C^0([0,T], H^{k+1}(\mathbb{R}^n, \mathbb{R}^d)), \\ \partial_t u &\in C^1([0,T], H^k(\mathbb{R}^n, \mathbb{R}^d)). \end{aligned}$$

Finally, for any $t \in [0,T]$ we have

$$\mathbb{E}_k(t) \leq \big(\mathbb{E}_k(0) + C_I t\big)e^{\int_0^t z_I(\mathcal{N}(u))dt},$$

where the RHS depends only on T and an upper bound for $\|u_0\|_{H^{k+1}(\sum_0)}$ and $\|u_1\|_{H^k(\sum_0)}, \mathbb{E}(t) = E[u,u](t)$.

Proof. The proof will be split into several steps.

Set-up. Let $\{u_{0,i}\}$ and $\{u_{1,i}\}$ be sequences of $C^\infty_c(\mathbb{R}^n, \mathbb{R}^d)$ converging to u_0 and u_1 in H^{k+1} and H^k, respectively. We can assume that

$$\|u_{0,i}\|_{H^{k+1}} + \|u_{1,i}\|_{H^k} \leq C_0 + 1,$$

where

$$C_0 := \|u_0\|_{H^{k+1}} + \|u_1\|_{H^k}.$$

Set

$$v_0(t,x) := u_{0,0}(x),$$

which has local compact support in x. Define v_{i+1} inductively as follows. Given $v_i \in C^\infty(\mathbb{R}\times\mathbb{R}^n, \mathbb{R}^d)$ with local compact support in x, let $v_{i+1} \in C^\infty(\mathbb{R}\times\mathbb{R}^n, \mathbb{R}^d)$ be the unique solution to

$$\begin{aligned} g^{\mu\nu}_{i+1}\partial_\mu\partial_\nu v_{i+1} &= f_{i+1} \text{ in } \mathbb{R}\times\mathbb{R}^n, \\ v_{i+1}(0,\cdot) &= u_{0,i+1} \text{ on } \{t=0\}\times\mathbb{R}^n, \\ \partial_t v_{i+1}(0,\cdot) &= u_{1,i+1} \text{ on } \{t=0\}\times\mathbb{R}^n \end{aligned}$$

where $g_{i+1} = g[v_i], f_{i+1} = f[v_i]$. Note that v_{i+1} has local compact support in x.

<u>Boundedness.</u> Let us assume inductively that

$$\mathcal{M}_k[v_{i-1}] \le \mathcal{C},\ \mathcal{M}_k[v_{i-2}] \le \mathcal{C}$$

for some $\mathcal{C}$ and $0 \le t \le T$. We have

$$\begin{aligned} \mathcal{N}[v_i] &\le C\|D^{\le 2}v_i\|_{L^\infty} + C\|D^{\le 1}\partial_t v_i\|_{L^\infty(\sum_t)} + C\|\partial_t^2 v_i\|_{L^\infty(\sum_t)} \\ &\le C\|v_i\|_{H^{k+1}(\sum_t)} + C\|\partial_t v_i\|_{H^k(\sum_t)} + C\|\partial_t^2 v_i\|_{L^\infty(\sum_t)} \end{aligned}$$

where we used Sobolev's embedding. Using the equation and our assumptions the last term is $\partial_t^2 v_i$ is bounded by

$$\begin{aligned} \|\partial_t^2 v_i\|_{L^\infty(\sum_t)} &\le C\|g_{v_i-1}\|_{L^\infty(\sum_t)}\left(\|D^2 v_i\|_{L^\infty(\sum_t)} + \|D\partial_t v_i\|_{L^\infty(\sum_t)} + \|f_{v_{i-1}}\|_{L^\infty(\sum_t)}\right) \\ &\le z_I(\mathcal{C})(\mathcal{M}_k[v_i] + z_I(\mathcal{C})) \\ &\le z_I(\mathcal{C})(1+\mathcal{M}_k[v_i]) \end{aligned}$$

after using Sobolev's embedding, our assumptions on g, f the miscellaneous inequalities we used before, and the induction hypothesis. Thus

$$\mathcal{N}[v_i] \le z_I(\mathcal{C})(1+\mathcal{M}_k[v_i]).$$

Similarly,

$$\begin{aligned} \mathcal{N}[v_{i-1}] &\le z_I(\mathcal{C})(1+\mathcal{M}_k[v_{i-1}]) \\ &\le z_I(\mathcal{C}) \end{aligned}$$

using again the induction hypothesis. We now use the energy estimate

$$\mathcal{M}_k[u](t) \le C_{1,I}\mathcal{M}_k[u](0) + \int_0^t \Big[C_{2,I} + z_{1,I}(\mathcal{N}[v])\big((1+\mathcal{N}[u])\mathcal{M}_k[v] + \mathcal{M}_k[u]\big)\Big]d\tau,$$

with $u \mapsto v_i, v \mapsto v_{i-1}$ to get

$$\begin{aligned} \mathcal{M}_k[v_i] &\le C_I\mathcal{M}_k[v_i](0) + \int_0^t \left[C_I + \underbrace{z_I(\mathcal{N}[v_{i-1}])}_{\le z_I(\mathcal{C})}\left((1+\underbrace{\mathcal{N}[v_i]}_{\le z_I(\mathcal{C})(1+\mathcal{M}_k[v_i])})\underbrace{\mathcal{M}_k[v_{i-1}]}_{\le\mathcal{C}} + \mathcal{M}_k[v_i]\right)\right]d\tau \\ &\le C_I\mathcal{M}[v_i](0) + z_I(\mathcal{C})\int_0^t (1+\mathcal{M}_k[v_i])d\tau. \end{aligned}$$

By Gronwall

$$\mathcal{M}_k[v_i] \le C_I(\mathcal{M}_k[v_i](0) + tz_I(\mathcal{C}))e^{tz_I(\mathcal{C})}.$$

Choosing $\mathcal{C}$ large enough, depending on C_0, C_I and T small enough, depending on $\mathcal{C}$ we have

$$\mathcal{M}_k[v_i] \le \mathcal{C}.$$

We need to verify the $i = 1, 2$ case, i.e., we need $\mathcal{M}_k[v_0] \le \mathcal{C}, \mathcal{M}_k[v_1] \le \mathcal{C}$. For $i = 1$, we have

$$\mathcal{M}_k[v_0] = \mathcal{M}_k[u_{0,0}] \le C_0 + 1,$$

so choose $\mathcal{C} \geq C_0 + 1$ For $i = 1$, i.e., v_1, we apply the above induction argument. For this, we need $\mathcal{N}[v_{i-1}] = \mathcal{N}[v_0] \leq z_I(\mathcal{C})$, which above we obtained using the induction hypothesis for v_{i-2}, which would give v_{-1} (see the above, where it says "similarly"). But here we have $\mathcal{N}[v_0] \leq z_I(\mathcal{C})$ directly from the fact that v_0 is constant in time and from Sobolev embedding.

Lower norm convergence. From linear theory we know that

$$v_i \in C^0([0,T], H^{k+1}(\mathbb{R}^n)) \cap C^1([0,T], H^k(\mathbb{R}^n)).$$

In the estimate for differences

$$\mathcal{M}_0[u](t) \leq C_I \exp \int_0^t z_{1,I}(\mathcal{N}[v_2]) d\tau \cdot \left[\mathcal{M}_0[u](0) + \int_0^t z_{2,I}(\mathcal{N}[u_1], \mathcal{N}[v_1], \mathcal{N}[v_2]) \mathcal{M}_0[v] d\tau \right],$$

$v = v_2 - v_1$, and $u = u_2 - u_1$, choose $v_2 \mapsto v_i,\ v_1 \mapsto v_{i-1},\ u_2 \mapsto v_{i+1},\ u_1 \mapsto v_i$ to get

$$\mathcal{M}_0[v_{i+1} - v_i] \leq C_I \exp \int_0^t z_I(\mathcal{N}[v_i]) d\tau \ \left[\mathcal{M}_0[v_{i+1} - v_i](0) + \int_0^t z_I(\mathcal{N}[v_i], \mathcal{N}[v_{i-1}]) \mathcal{M}_0[v_i - v_{i-1}] d\tau \right]$$

By the foregoing,

$$\mathcal{N}_I[v_i] \leq z_I(\mathcal{C})(1 + \mathcal{M}_k[v_i]) < z_I(\mathcal{C}),$$

so

$$C_I e^{t z_I(\mathcal{C})} \left[\mathcal{M}[v_{i+1} - v_i](0) + z_I(\mathcal{C}) \int_0^t \mathcal{M}_0[v_i - v_{i-1}] d\tau \right]$$

Put $a_i = \sup_{0 \leq t \leq T} \mathcal{M}_0[v_{i+1} - v_i](t)$, We can assume that T is small enough and the approximating initial data sequence are such that

$$C_I e^{t z_I(\mathcal{C})} t z_I(\mathcal{C}) \leq \frac{1}{2}$$

and

$$C_I e^{t z_I(\mathcal{C})} \mathcal{M}_0[v_{i+1} - v_i](0) \leq 2^{-i}.$$

Then

$$a_i \leq 2^{-i} + \frac{1}{2} a_{i-1}.$$

Then

$$a_2 \leq \frac{1}{4} + \frac{1}{2} a_1,$$
$$a_3 \leq \frac{1}{8} + \frac{1}{2} a_2 \leq \frac{1}{8} + \frac{1}{8} + \frac{1}{4} a_1,$$
$$a_i \leq \frac{i-1}{2^i} + \frac{a_1}{2^{i-1}}.$$

Thus

$$\begin{aligned} \mathcal{M}_0[v_{i+j} - v_i] &\leq \mathcal{M}_0[v_{i+j} - v_{i+j-1}] + \ldots + \mathcal{M}_0[v_{i+1} - v_i] \\ &\leq \frac{i+j-2}{2^{i+j-1}} + \frac{a_1}{2^{i+j-2}} + \ldots + \frac{i-1}{2^i} + \frac{a_1}{2^{i-1}}. \end{aligned}$$

Since $\sum \frac{i}{2^i}$ converges, we conclude that $\{v_i\}$ is Cauchy in $C^0([0,T], H^1(\mathbb{R}^n)) \cap C^1([0,T], L^2(\mathbb{R}^n))$, hence converges.

Higher norm convergence. For any $0 < l < k$, interpolation

$$\|u\|_{H^{(s_2)}(\mathbb{R}^n)} \leq \|u\|_{H^{(s_1)}(\mathbb{R}^n)}^{\frac{s_3 - s_2}{s_3 - s_1}} \|u\|_{H^{(s_3)}(\mathbb{R}^n)}^{\frac{s_2 - s_1}{s_3 - s_1}}, \quad s_1 < s_2 < s_3$$

gives,

$$\begin{aligned}\|v_{i+j}-v_i\|_{H^{l+1}(\sum_t)}+&\|\partial_t v_{i+j}-\partial_t v_i\|_{H^l(\sum_t)}\\ &\leq \underbrace{\|v_{i+j}-v_i\|_{H^l(\sum_t)}^{\frac{k+1-(l+1)}{k+1-1}}}_{\to 0}\underbrace{\|v_{i+j}-v_i\|_{H^{k+1}(\sum_t)}^{\frac{l+1-1}{k+1-1}}}_{\text{bounded by energy estimates}}\\ &+\overbrace{\|\partial_t v_{i+j}-\partial_t v_i\|_{H^0(\sum_t)}^{\frac{k-l}{k-0}}}\overbrace{\|\partial_t v_{i+j}-\partial_t v_i\|_{H^k(\sum_t)}^{\frac{l-0}{k-0}}}\end{aligned}$$

hence $\{v_i\}$ converges in $C^0([0,T],H^{l+1}(\mathbb{R}^n))\cap C^l([0,T],H^2(\mathbb{R}^n))$, $l<k$. Since $k>\frac{n}{2}+1$, we can take l (not necessarily integer) so that $l>\frac{n}{2}+1$. Hence, by Sobolev embedding, the sequence converges in $C^0([0,T],C^2(\mathbb{R}^n))\cap C^1([0,T],C^1(\mathbb{R}^n))$. Using the equation we get that $\partial_t^2 v_i$ converges in $C^0([0,T],C^0(\mathbb{R}^n)$. We conclude that the sequence converges in $C_B^2([0,T]\times\mathbb{R}^n)$, hence we obtain a C^2 solution.

Top-norm boundedness. Denote by u the above solution. We already know that $u\in H^{l+1}(\sum_t),\partial_t u\in H^l(\sum_t), l<k$. For each fixed t, the sequence $\{v_i(t,\cdot)\}$ converges in $H^{l+1}(\sum_t)$ and is bounded in $H^{k+1}(\sum_t)$, hence the limit is in $H^{k+1}(\sum_t)$. Similarly, for $\partial_t u$, thus $u(t,\cdot)\in H^{k+1}(\sum_t),\partial_t u(t,\cdot)\in H^k(\sum_t)$.

It remains to show regularity with t and the energy bound.

Time-continuity: weak. Let us first show that u is weakly continuous with respect to t, i.e., given a bounded linear functional φ on $H^{k+1}(\mathbb{R}^n)$, the map $t\mapsto\varphi(u(t,\cdot))$ is continuous.

φ is represented by an element in $H^{-k-1}(\mathbb{R}^n)$ which we still denote by φ, so

$$\langle\varphi,u\rangle=\int_{\mathbb{R}^n}\varphi u.$$

Let φ_j be a sequence of Schwartz functions converging to φ and v_i as above.

$$\begin{aligned}\langle\varphi,u(t,\cdot)\rangle-&\langle\varphi,v_i(t,\cdot)\rangle\\ &=\langle\varphi,u(t,\cdot)\rangle-\langle\varphi_j,u(t,\cdot)\rangle+\langle\varphi_j,u(t,\cdot)\rangle-\langle\varphi_j,v_i(t,\cdot)\rangle\\ &\quad-\langle\varphi_j,v_i(t,\cdot)\rangle-\langle\varphi,v_i(t,\cdot)\rangle,\end{aligned}$$

for $0<l<k$.

$$\begin{aligned}\left|\langle\varphi,u(t,\cdot)\rangle-\langle\varphi,v_i(t,\cdot)\rangle\right|&\leq\|\varphi_j-\varphi\|_{H^{-k-1}(\mathbb{R}^n)}\,\|u(t_i)\|_{H^{k+1}(\mathbb{R}^n)}\\ &+\|\varphi_j\|_{H^{-l-1}(\mathbb{R}^n)}\,\|u(t,\cdot)-v_i(t,\cdot)\|_{H^{l+1}(\mathbb{R}^n)}+\|\varphi_j-\varphi\|_{H^{-k-1}(\mathbb{R}^n)}\,\|v_i(t,\cdot)\|_{H^{k+1}(\mathbb{R}^n)}\end{aligned}$$

Fix j large enough so that

$$\|\varphi_j-\varphi\|_{H^{-k-1}(\mathbb{R}^n)}\left(\|u(t,\cdot)\|_{H^{k+1}(\mathbb{R}^n)}+\|v_i(t,\cdot)\|_{H^{k+1}}\right)<\epsilon$$

which is possible since $\|u(t,\cdot)\|_{H^{k+1}(\mathbb{R}^n)}+\|v_i(t,\cdot)\|_{H^{k+1}}$ is bounded by the above. Then choose i (depending on j) such that

$$\|\varphi_j\|_{H^{-l-1}(\mathbb{R}^n)}\|u(t,\cdot)-v_i(t,\cdot)\|_{H^{l+1}(\mathbb{R}^n)}<\epsilon$$

which is possible by the convergence for $l<k$. Thus,

$$\left|\langle\varphi,u(t,\cdot)-v_i(t,\cdot)\rangle\right|<2\epsilon,\ \ 0\leq t\leq T,$$

and v_i converges uniformly in t to u with respect to the weak topology. But the v_i's are weakly continuous in t since they belong to $C^0([0,T], H^{k+1}(\mathbb{R}^n))$ thus u is the uniform limit of weak continuous functions, and hence weakly continuous. A similar argument applies to $\partial_t u$.

Time-continuity: strong. Let us show that

$$\lim_{t\to 0^+} \Big(\|u(t,\cdot) - u(0,\cdot)\|_{H^{k+1}(\mathbb{R}^n)} + \|\partial_t u(t,\cdot) - \partial_t u(0,\cdot)\|_{H^k(\mathbb{R}^n)} \Big) = 0$$

i.e., right continuity at $t = 0$. Left-continuity follows by reversing time and continuity at any t_0 by taking $u(t_0), \partial_t u(t_0)$ as initial data for the problem on $[t_0, T]$. We will use the estimate

$$\mathbb{E}_k(t) \leq (\mathbb{E}_k(0) + C_I t) e^{\int_0^t z_I(\mathcal{N}(u))dt}$$

that we will prove later. Set

$$h^{ij}(x) = g^{ij}(0, x, u, \partial u),$$

i.e., $= g^{ij}(0, x, u(0,x), \partial u(0,x))$.

Under our assumptions and the results established so far, the following is an inner product on $H^{k+1}(\mathbb{R}^n) \times H^k(\mathbb{R}^n)$ equivalent to the standard one:

$$((v_1, v_2), (w_1, w_2)) := \frac{1}{2} \sum_{|\vec{\alpha}| \leq k} \int_{\mathbb{R}^n} \Big(h^{ij} \partial_i D^{\vec{\alpha}} v_1 \partial_j D^{\vec{\alpha}} w_1 + D^{\vec{\alpha}} v_1 D^{\vec{\alpha}} w_1 + D^{\vec{\alpha}} v_2 D^{\vec{\alpha}} w_2 \Big)\, dx.$$

Compute

$$\begin{aligned}
\big((u-u_0,\, \partial_t u - u_1),\, (u-u_0,\, \partial_t u - u_1)\big) &= \big((u,\partial_t u) - (u_0,u_1),\, (u,\partial_t u) - (u_0,u_1)\big) \\
&= \big((u,\partial_t u),\, (u,\partial_t u)\big) + \big((u_0,u_1),\, (u_0,u_1)\big) \\
&\quad - 2\big((u,\partial_t u),\, (u_0,u_1)\big)
\end{aligned}$$

Since $g^{00} = -1$, $g(0,x,u,\partial u) = h$, we have

$$\mathbb{E}_k(0) = ((u_0,u_1),(u_0,u_1)).$$

The map $u \mapsto ((u,\partial_t u),(u_0,u_1))$ defines a linear functional, thus by the weak continuity established above we have

$$\lim_{t\to 0^+} \big((u,\partial_t u),(u_0,u_1)\big) = ((u_0,u_1),(u_0,u_1)) = \mathbb{E}_k(0).$$

Thus

$$\begin{aligned}
&\lim_{t\to 0^+} \sup\, ((u-u_0, \partial_t u - u_1),(u-u_0,\partial_t u - u_1)) \\
&\qquad = \lim_{t\to 0^+} \sup\, ((u,\partial_t u),(u,\partial_t u) + ((u_0,u_1),(u_0,u_1)) - 2 \lim_{t\to 0^+} \sup\, ((u,\partial_t u),(u_0,u_1)) \\
&\qquad = \lim_{t\to 0^+} \sup\, ((u,\partial_t u),(u,\partial_t u)) - \mathbb{E}_k(0) \\
&\qquad = \lim_{t\to 0^+} \sup\, \Big[(u,\partial_t u),(u,\partial_t u)) - \mathbb{E}_k(t) \Big] + \lim_{t\to 0^+} \sup\, \mathbb{E}_k(t) - \mathbb{E}_k(0).
\end{aligned}$$

From the energy estimate, $\lim_{t\to 0^+} \sup \mathbb{E}_k(t) \leq \mathbb{E}_k(0)$, so

$$\lim_{t\to 0^+} \sup\, \big((u-u_0,\partial_t u - u_1),(u-u_0,\partial_t u - u_1)\big) \leq \lim_{t\to 0^+} \sup\, \Big[\big((u_1,\partial_t u),(u,\partial_t u)\big) - \mathbb{E}_k(t) \Big].$$

Write

$$\begin{aligned}
\big((u,\partial_t u),(u,\partial_t u)\big)-\mathbb{E}_k(t) &= \frac{1}{2}\sum_{|\vec{\alpha}|\le k}\int_{\sum_t}\Big(|\partial_t D^{\vec{\alpha}}u|^2+h^{ij}\partial_i D^{\vec{\alpha}}u\partial_j D^{\vec{\alpha}}u+|D^{\vec{\alpha}}u|^2\Big)\,dx\\
&-\frac{1}{2}\sum_{|\vec{\alpha}|\le k}\int_{\sum_t}\Big(-g_u^{00}|\partial_t D^{\vec{\alpha}}u|^2+g_u^{ij}\partial_i D^{\vec{\alpha}}u\partial_j D^{\vec{\alpha}}u+|D^{\vec{\alpha}}u|^2\Big)\,dx\\
&=\frac{1}{2}\sum_{|\vec{\alpha}|\le k}\int_{\sum_t}\big(h^{ij}-g_u^{ij}\big)\,D_iD^{\vec{\alpha}}u\cdot D_jD^{\vec{\alpha}}u dx\\
&\le C\|h^{-1}-g^{-1}\|_{L^\infty(\sum_t)}\overbrace{\|u\|_{H^{k+1}(\sum_t)}}^{\le C}
\end{aligned}$$

For $\frac{n}{2}+1<l<k$, by Sobolev embedding, our assumptions, the miscellaneous inequalities as above, the fundamental theorem of calculus, and writing

$$g^{-1}(0,x,\partial u(0,x))-g^{-1}(t,x,\partial u(0,x))+g^{-1}(t,x,\partial u(0,x))-g^{-1}(t,x,\partial u(t,x))$$

we have

$$\|h^{-1}-g^{-1}\|_{L^\infty(\sum_t)}\le z_I(\mathcal{C})\|\partial u_0-\partial u\|_{H^l(\sum_t)}$$

which goes to zero when $t\to 0^+$ since we have strong continuity in the $H^{l+1}\times H^l$-norm. Thus

$$0\le\lim_{t\to 0^+}\sup\big((u-u_0,\partial_t u-u_1),(u-u_0,\partial_t u-u_1)\big)=0$$

which gives the result by the inner-product equivalence.

Energy estimate for u.
It remains to show

$$\mathbb{E}_k(t)\le\big(\mathbb{E}_k(0)+C_It\big)e^{\int_0^t z_I(\mathcal{N}(u))dt}$$

(Note that we have to prove this without using the strong continuity since the proof of the latter relied on this estimate.)
Recall

$$\partial_t E_k[v,u]\le c_{3,I}+z_{2,I}\big(\mathcal{N}[u],\mathcal{N}[v]\big)\Big(\mathcal{M}_k^2[v]+E_k[v,u]\Big).$$

Apply this inequality with $v\mapsto v_i$, $u\mapsto v_{i+1}$, $u_0\mapsto u_{0,i}$, $u_1\mapsto u_{1,i}$

$$E_k(v_i,v_{i+1})\le E_k(v_i,v_{i+1})(0)+\int_0^t\Big[c_I+z_I\big(\mathcal{N}(v_i),\mathcal{N}(v_{i+1})\big)\big(\mathcal{M}_k^2(v_i)+E_k(v_i,v_{i+1})\big)\Big]d\tau$$

We compare

$$\begin{aligned}
\big|E_k(v_i,v_{i+1})-E_k(u,v_{i+1})\big| &= \Big|\frac{1}{2}\sum_{|\vec{\alpha}|\le k}\int_{\sum_t}(-g_{v_i}^{00}|\partial_t D^{\vec{\alpha}}v_{i+1}|^2+g_{v_i}^{jl}\partial_j D^{\vec{\alpha}}v_{i+1}\partial_l D^{\vec{\alpha}}v_{i+1}-|D^{\vec{\alpha}}v_{i+1}|^2)dx\\
&-\frac{1}{2}\sum_{|\vec{\alpha}|\le k}\int_{\sum_t}(-g_u^{00}|\partial_t D^{\vec{\alpha}}v_{i+1}|^2+g_u^{jl}\partial_j D^{\vec{\alpha}}v_{i+1}\partial_l D^{\vec{\alpha}}v_{i+1}-|D^{\vec{\alpha}}v_{i+1}|^2)dx\Big|\\
&\le\|g_{v_i}^{-1}-g_u^{-1}\|_{L^\infty(\sum_t)}C_I(\mathcal{C}).
\end{aligned}$$

Since $v_i\mapsto u$ in $C_B^2([0,T]\times\mathbb{R}^n)$, we have that the RHS $\to 0$, thus

$$\begin{aligned}
\lim_{i\to\infty}\sup E_k(u,v_{i+1}) &= \lim_{i\to\infty}\sup E_k(v_i,v_{i+1})\\
&\le\lim_{i\to\infty}\sup E_k(v_i,v_{i+1})(0)\\
&+\int_0^t\Big[c_I+\lim_{i\to\infty}\sup z_I\big(\mathcal{N}(v_i),\mathcal{N}(v_{i+1})\big)\big(\mathcal{M}_k^2(v_i)+E_k(v_i,v_{i+1})\big)\Big]d\tau
\end{aligned}$$

(we used the reverse Fatou's lemma to move the lim sup inside the integral, which can be invoked by $E_k(v_i, v_{i+1}) \leq C_I(\mathcal{C})$).

By the C^2 convergence, $\mathcal{N}(v_i) \to \mathcal{N}(u)$ and, from how the initial data was chosen, $E_k(v_i, v_{i+1})(0) \to \mathbb{E}_k(0)$. Moreover, $\mathcal{M}_k^2(v_i) \leq C_I E(u, v_i)$ by the properties of u established so far (and consequently of g_u) so

$$\lim_{i\to\infty} \sup \mathcal{M}_k^2(v_i) \leq C_I \lim_{i\to\infty} \sup E(u, v_i).$$

Thus

$$\lim_{i\to\infty} \sup E_k(u, v_{i+1}) \leq \mathbb{E}_k(0) + \int_0^t \left[c_I + z_I(\mathcal{N}(u)) \lim_{i\to\infty} \sup E_k(u, v_i) \right] d\tau$$

Gronwalling:

$$\lim_{i\to\infty} \sup E_k(u, v_{i+1}) \leq (\mathbb{E}_k(0) + tC_I)\, e^{\int_0^t z_I(\mathcal{N}(u)) d\tau}.$$

Since v_i converges in $H^{l+1} \times H^l$ and is bounded in $H^{k+1} \times H^k$, it converges weakly to u in $H^{k+1} \times H^k$ (this remains true for the equivalent inner product since it is bounded there too). So

$$\begin{aligned} \mathbb{E}_k(u) &= \lim_{i\to\infty} \sup \frac{1}{2} \sum_{|\alpha|\leq k} \int_{\sum_t} \left(-g_u^{00} \partial_t D^{\vec{\alpha}} u \partial_t D^{\vec{\alpha}} v_i + g_u^{lj} \partial_l D^{\vec{\alpha}} u \partial_j D^{\vec{\alpha}} u + D^{\vec{\alpha}} u D^{\vec{\alpha}} v_i \right) \\ &\leq \lim_{i\to\infty} \sup \big((u, \partial_t u), (v_i, \partial_t v_i) \big) \\ &\leq \lim_{i\to\infty} \sup \underbrace{\|(u, \partial_t u\|_u}_{=\mathbb{E}_k^{\frac{1}{2}}(u)} \underbrace{\|(v_i, \partial_t v_i)\|_u}_{=E_k^{\frac{1}{2}}(u, v_i)} \end{aligned}$$

where we used the Cauchy-Schwarz inequality for the equivalent inner product (with norm denoted $\|\cdot\|_u$). Hence, dividing,

$$\mathbb{E}_k^{\frac{1}{2}}(u) \leq \lim_{i\to\infty} \sup E_k^{\frac{1}{2}}(u, v_i).$$

which implies the result. □

14.1. **Continuation criterion and smooth solutions.** We are interested in the following questions: if a solution is defined on $[0, T]$, can it be continued past T? If the initial data is C^∞, is the solution?

Theorem 14.12. *Let g be a C^∞ admissible metric and f a C^∞ admissible nonlinearity. Let $u_0 \in H^{k+1}(\mathbb{R}^n, \mathbb{R}^d)$, $u_1 \in H^k(\mathbb{R}^n, \mathbb{R}^d)$, where $k > \frac{n}{2} + 1$. Let $u \in C_B^2([0, T] \times \mathbb{R}^n, \mathbb{R}^d)$ be a solution to*

$$\begin{aligned} g^{\mu\nu} \partial_\mu \partial_\nu u &= f \text{ in } [0, T] \times \mathbb{R}^n, \\ u(0, \cdot) &= u_0 \text{ on } \{t = 0\} \times \mathbb{R}^n, \\ \partial_t u(0, \cdot) &= u_1 \text{ on } \{t = 0\} \times \mathbb{R}^n \end{aligned}$$

$T > 0$. Then u has regularity

$$\begin{aligned} u &\in C^0([0, T], H^{k+1}(\mathbb{R}^n, \mathbb{R}^d)), \\ \partial_t u &\in C^1([0, T], H^k(\mathbb{R}^n, \mathbb{R}^d)). \end{aligned} \tag{14.2}$$

and for any $t \in [0, T]$,

$$\mathbb{E}_k(t) \leq \big(\mathbb{E}_k(0) + C_I t \big) e^{\int_0^t z_I(\mathcal{N}(u)) dt}, \tag{14.3}$$

where $\mathbb{E}_k$ is as in the previous theorem. Let T_k be the supremum of T, for which u is a C^2 solution defined on $[0, T]$ and satisfying (14.2). Then either $T_k = \infty$ or

$$\lim_{t\to T_k^-} \sup_{0\leq\tau\leq t} \mathcal{N}[u](\tau) = \infty.$$

Proof. We know (14.2) and (14.3) to hold on a possibly smaller interval, i.e., the interval where the iteration of the previous theorem converges. Let I be the set of times $T_0 \in [0,T]$ such that (14.2) and (14.3) hold on $[0,T_0]$. We already have that I is not empty. For $T_0 \in I$, we have $(u(T_0,\cdot),\partial_t u(T_0,\cdot)) \in H^{k+1}\times H^k$, so we can take it as initial data and obtain a solution defined on $[T_0,T_0+\epsilon]$ for some $\epsilon>0$ and satisfying (14.2) and (14.3) on $[T_0,T_0+\epsilon]$. Since (14.2) and (14.3) hold on $[0,T_0]$ by the definition of T_0, (14.2) holds on $[0,T_0+\epsilon]$. Moreover

$$\mathbb{E}_k(t) \le \big(\mathbb{E}_k(0)+ct\big)e^{\int_0^t z(\mathcal{N}(u))d\tau},\ \ 0\le t\le T_0,$$

$$\mathbb{E}_k(t) \le \big(\mathbb{E}_k(T_0)+C(t-T_0)\big)e^{\int_{T_0}^t z(\mathcal{N}(u))d\tau},\ \ T_0\le t\le T_0+\epsilon.$$

Applying the first inequality with $t=T_0$ we find that the second inequality gives

$$\begin{aligned}
\mathbb{E}_k(t) &\le \big(\mathbb{E}_k(T_0)+C(t-t_0)\big)e^{\int_{T_0}^t z(\mathcal{N}[u])d\tau}\\
&\le \Bigg[(\mathbb{E}_k(0)+\underbrace{CT_0}_{\le Ct \text{ since } t\in[T_0,T_0+\epsilon]})e^{\int_0^{T_0} z(\mathcal{N}[u])d\tau}+\underbrace{C(t-t_0)}_{\le Ct}\Bigg]e^{\int_{T_0}^t z(\mathcal{N}[u])d\tau}\\
&\le \Big[\big(\mathbb{E}_k(0)+Ct\big)e^{\int_0^{T_0} z(\mathcal{N}[u])d\tau}\Big]\,e^{\int_{T_0}^t z(\mathcal{N}[u])d\tau}\\
&\le \big(\mathbb{E}_k(0)+Ct\big)e^{\int_0^t z(\mathcal{N}[u])d\tau},
\end{aligned}$$

showing that I is open.

Let $T_i\to T_0, T_i\in I$. Since $T_i\in I$, (14.3) holds for each i, hence $\big(u(T_i,\cdot),\partial_t u(T_i,\cdot)\big)\in H^{k+1}\times H^k$ is uniformly bounded independent of i. Because the time of existence depends only on the size of the data (and the structure of the nonlinearities), this uniform bound gives a solution for each data $\big(u(T,\cdot),\partial_t u(T_i,\cdot)\big)$ defined on $[T_i,T_i+\epsilon]$ where $\epsilon>0$ is independent of i. Hence we get a solution satisfying (14.2) past T_0 and by continuity (14.3) holds as well. Hence I is closed, thus (14.2) and (14.3) hold.

For the characterization of T_k, suppose that $T_k<\infty$ but

$$\lim_{t\to T_k^-}\ \sup_{0\le\tau\le t}\mathcal{N}[u](\tau)<\infty.$$

Then $\mathcal{N}[u]$ is bounded on $[0,T_k)$. Then (14.3) implies that $\mathbb{E}_k$ has a uniform bound on $[0,T_k)$. Arguing as above, we can construct a solution past T_k, contradicting the definition of T_k.

□

Corollary 14.13. *Under the same assumptions of the local existence and uniqueness theorem, if* $u_0,u_1\in C_c^\infty(\mathbb{R}^n)$ *then the solution is* C_∞.

Proof. This follows from the fact that the data is in H^k for any k and that T_k in the previous theorem is independent of k.

□

Remark 14.14. If g does not depend on ∂u, we can replace $k>\frac{n}{2}+1$ by $k>\frac{n}{2}$ and $\mathcal{N}[u]$ to involve up to first derivatives of u only. This can be seen by inspection in the local existence and uniqueness proof.

15. The role of the characteristics

We will now discuss the concept of characteristic manifolds, or characteristics for short. These play a role similar to the cones in the standard wave equation. It will be important to distinguish between elements of the tangent and cotangent space at a point, even if we will consider primarily equations defined on $[0,T]\times\mathbb{R}^n$ (the generalization for manifolds will be straightforward though). We begin with several definitions and give a motivation further below.

Definition 15.1. Consider in $X=[0,T]\times\mathbb{R}^n$ a linear scalar differential operator L of order k with principal part

$$P=\sum_{|\alpha|=k} a_\alpha D^\alpha.$$

For each $x\in X$ and each $\xi\in T_x^*X$, we can associate a polynomial of degree k in T_x^*X, called the **characteristic polynomial** of at x, by

$$P(x,\xi)=\sum_{|\alpha|=k} a_\alpha(x)\xi^\alpha,$$

where $\xi^\alpha=\xi_0^{\alpha_0}\dots\xi_n^{\alpha_n}$ and we abuse notation, using P for both the principal part of the operator and its characteristic polynomial. The cone $V_x(P)\subset T_x^*X$ is defined by

$$P(x,\xi)=0,$$

called the **characteristic cone** (at x). (Although the set need not to be a cone in all cases, but see below).

Example 15.2. For the wave equation (more precisely, the wave operator, but we abuse terminology).

$$-u_{tt}+\Delta u=0$$

the characteristic cone at any x is given by (the boundary of) the light-cones $-\xi_0^2+|\vec{\xi}|^2=0$.

Example 15.3. For the transport equation

$$\partial_t u+\vec{b}\cdot\nabla u=0$$

the characteristic cones have the form

$$\xi_0+\vec{b}\cdot\vec{\xi}=0$$

Identifying $\vec{b}$ with a one form:

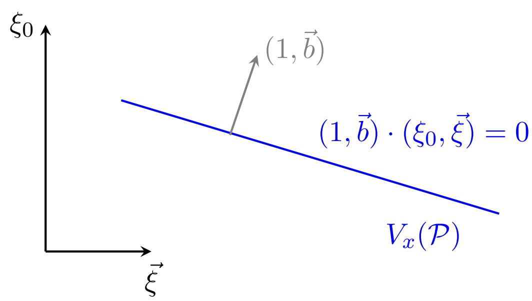


FIGURE 33. $\vec{b}$ and the One Form

Definition 15.4. A regular hypersurface $\sum\subset X$ (i.e., a hypersurface for which tangent vectors are well-defined) is called a **characteristic manifold**, **characteristic surface**, or simply a **characteristics** for P (or for L) if the following holds. $\sum$ can be locally represented as $\{x\in X\mid\phi(x)=0\}$, $d\phi\neq 0$. Setting $\xi=d\phi$, $\xi\in V_x(P)$ for all $x\in\{\phi(x)=0\}$. We also call $\sum$ a **null** hypersurface.

It is convenient to also define characteristics when we have a curve instead of a surface. More precisely, we say that the flow lines of the vector field u are the characteristics for the operator $u^\mu\partial_\mu$ (this is motivated by the characteristic cone $u^\mu\xi_\mu=0$).

Example 15.5. For the wave operator $-u_{tt} + c^2 u_{xx}$, the characteristics are the curves $x \pm ct =$ const. The characteristics cones are $-\xi_0^2 + c^2\xi_1^2 = 0$, $\xi_0 = \pm c|\xi_1|$. Take, e.g., $\phi(t,x) = x \pm ct$. $d\phi = dx \pm cdt = \xi,\ \xi_0 \pm c|\xi_1|$.

In order to generalize the above to vector-valued differential operators (i.e., for systems of PDEs) we need to define the principal part of a mixed order operator (i.e., consider systems of PDEs where the equations might have different orders).

Definition 15.6. Let $L : C^\infty(\Omega, \mathbb{R}^d) \to C^\infty(\Omega, \mathbb{R}^d)$ be a linear differential operator. Assume that it is possible to find non-negative integers m_I, n_J such that

$$(Lu)^J = h_I^J(x, D^{m_I - u_J})u^I + b^J(x, D^{m_K - n_J - 1})u^K,$$

$x \in \Omega$, $I, J, K = 1, ..., d$, sum over I, K, where $h_I^J(x, D^{m_I - n_J})$ is a homogenous linear differential operator of order $m_I - n_J$ (which could be identically zero) and $b^J(x, D^{m_K - n_J - 1})$ is a linear differential operator of order $m_K - n_J - 1$ (which could also be zero). Under these conditions, the **principal part** of L is the operator

$$P = (h_I^J(x, D^{m_I - n_J}))_{I,J=1,\dots,d}$$

and the **characteristic polynomial** is defined as

$$P(x,\xi) = \det(h_I^J(x,\xi)).$$

The definitions of **characteristic cones** and **characteristics** extend to this situation.

Remark 15.7. We also call $h_I^J(x,\xi)$ the characteristic matrix and $\det(h_I^J(x,\xi))$ the characteristic determinant of P. Note that it is a homogenous polynomial (in ξ) of degree $\sum_{I=1}^d m_I - \sum_{J=1}^d n_J$. The indices m_I, n_J are defined up to an overall additive factor.

Example 15.8. Consider the system

$$\begin{aligned} u^\mu \partial_\mu \phi_1 &= \phi_1 + \phi_2 + \phi_3 \\ g^{\mu\nu}\partial_\mu\partial_\nu\phi_2 &= \partial\phi_1 + \partial\phi_2 + \partial\phi_3 \\ g^{\mu\nu}u^\lambda\partial_\mu\partial_\nu\partial_\lambda\phi_3 &= \partial^2\phi_1 + \partial^2\phi_2 + \partial^2\phi_1 \end{aligned}$$

This system has the above structure with the choice

$$\phi_1 : m_1 = 3, n_1 = 2 \qquad \phi_2 : m_2 = 3, n_2 = 1 \qquad \phi_3 : m_3 = 3, n_3 = 0.$$

Then the principal part is given by the LHS,

$$P\begin{pmatrix}\phi_1\\ \phi_2\\ \phi_3\end{pmatrix} = \begin{pmatrix} u^\mu\partial_\mu & 0 & 0 \\ 0 & g^{\mu\nu}\partial_\mu\partial_\nu & 0 \\ 0 & 0 & g^{\mu\nu}u^\lambda\partial_\mu\partial_\nu\partial_\lambda \end{pmatrix}\begin{pmatrix}\phi_1\\ \phi_2\\ \phi_3\end{pmatrix}$$

even though there is a $\partial^2\phi_1, \partial^2\phi_2$ dependence on the RHS. To understand this, recall that we already know the definition of the principal part if all equations have the same order. Transforming the system into this case by taking ∂^2 of the first equation and ∂ of the second, we find

$$\begin{aligned} \partial^3\phi_1 &= \partial^2\phi_1 + \partial^2\phi_2 + \partial^2\phi_3 \\ \partial^3\phi_2 &= \partial^2\phi_1 + \partial^2\phi_2 + \partial^2\phi_3 \\ \partial^3\phi_3 &= \partial^2\phi_1 + \partial^2\phi_2 + \partial^2\phi_1 \end{aligned}$$

and we see that the RHS is indeed lower order.

Definition 15.9. We say that $P(x,\xi)$ is a **hyperbolic polynomial** at x if there exists $\xi \in T_x^*X$ with the following property: given a non-zero $\theta \in T_x^*X$ that is not parallel to ξ, the line $\lambda\xi + \theta$, where $\lambda \in \mathbb{R}$ is a parameter, intersects the cone $V_x(P)$ at k distinct real points. We say that L is a **hyperbolic operator** at x if $P(x,\xi)$ is a hyperbolic polynomial at x. We simply say hyperbolic if these conditions hold at every x or on a domain that is implicitly understood. (Recall that k is the degree of $P(x,\xi)$).

Remark 15.10. The definition of hyperbolicity varies across the literature and sometimes qualifiers such as strong, strict, weak hyperbolic are used to make distinctions among different definitions (see below).

Example 15.11. For the wave equation, any ξ in the interior of the light-cone satisfies this property. For the transport equation, any ξ not parallel to $\xi_0 + \vec{b}\cdot\vec{\xi} = 0$ satisfies the property

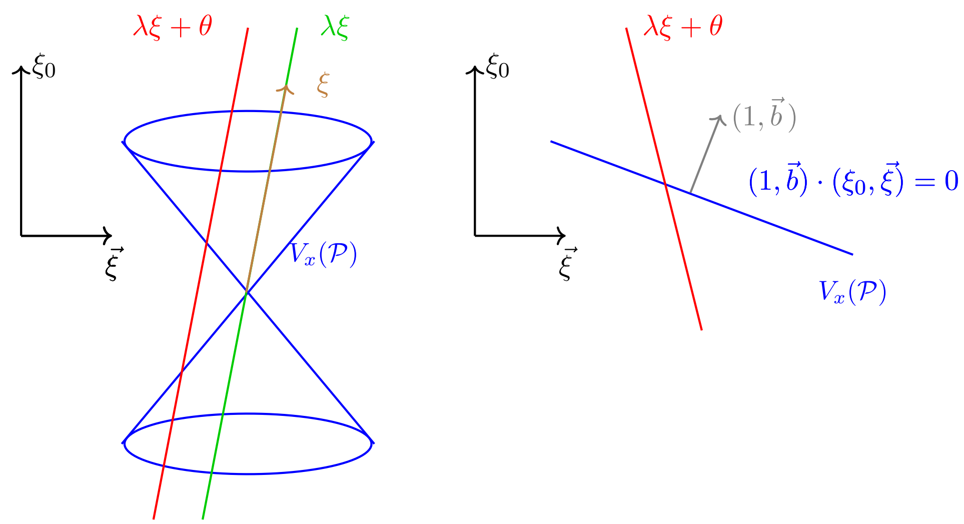


FIGURE 34. The Light Cone and $V_x(\mathcal{P})$

Theorem 15.12. ***(Leray).*** *If $P(x,\xi)$ is hyperbolic at x and dim $X \geq \xi$, then the set of ξ's satisfying the definition forms the interior of two opposite convex half-cones $\Gamma_x^{*,+}(P), \Gamma_x^{*,-}(P)$ with $\Gamma_x^{*,\pm}(P)$ not empty and whose boundaries belong (but need not to coincide) with $V_x(P)$.*

Example 15.13. Let g be a Lorentzian metric and u be a timelike unit vectorfield, i.e., $g_{\alpha\beta}u^\alpha u^\beta = -1$. The operator $g^{\mu\nu}u^\lambda\partial_\mu\partial_\nu\partial_\lambda$ is hyperbolic.

$$P(x,\xi) = g^{\mu\nu}u^\lambda\partial_\mu\partial_\nu\partial_\lambda = 0,$$

thus

$$\begin{cases} g^{\mu\nu}\xi_\mu\xi_\nu = 0 \\ u^\lambda\xi_\lambda = 0 \end{cases}$$

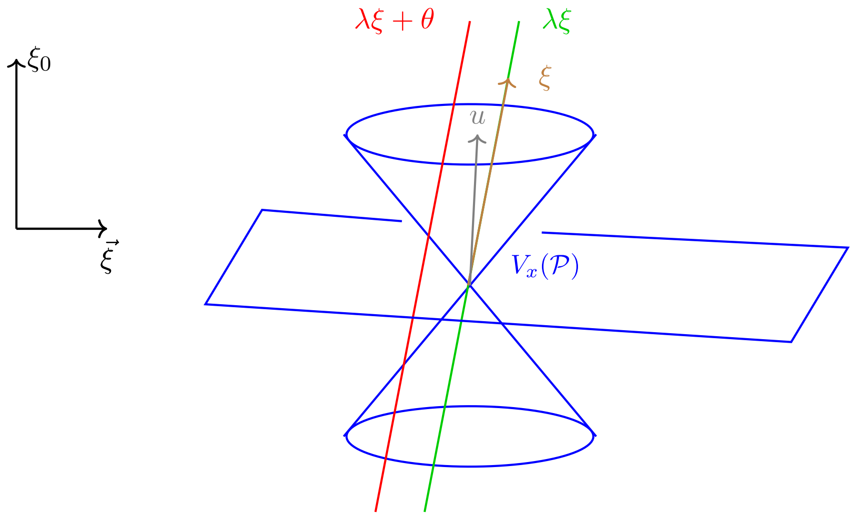


FIGURE 35. Light Cone and Parallel $\lambda\xi + \theta$, $\lambda\xi$

$u^\mu \xi_\mu = g^{\mu\nu} u_\mu \xi_\nu = 0.$

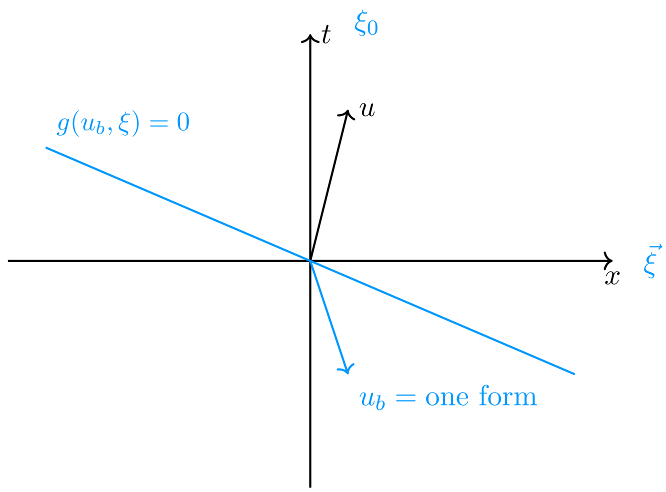


FIGURE 36. Illustration

Definition 15.14. We say that L is **weakly hyperbolic** at x if $P(x,\xi)$ is a product

$$P(x,\xi) = P_1(x,\xi) \ldots P_\mu(x,\xi),$$

where each $P_i(x,\xi)$ is a hyperbolic polynomial, and the intersections

$$\Gamma_x^{*,+}(P) := \bigcap_{i=1}^{\mu} \Gamma_x^{*,+}(P_i)$$

$$\Gamma_x^{*,-}(P) := \bigcap_{i=1}^{\mu} \Gamma_x^{*,-}(P_i),$$

have non-empty interior, where $\Gamma_x^{*,\pm}(P_i)$ are the convex cones associated to $P_i(x,\xi)$. If in addition P is diagonal (i.e., $h_I^J = 0$ for $I \neq J$), and each diagonal entry of P is a hyperbolic operator (at x), then we say that L is a **hyperbolic operator in diagonal form** (at x).

The reason to consider the intersection of the cones can be understood with the following example. Consider

$$-u_{tt} + u_{xx} + u_{yy} = 0,$$
$$-v_{xx} + v_{tt} + v_{yy} = 0,$$

which are two wave equations, with x playing the role of a time variable for the second equation. But

$$-u_{tt} + u_{xx} + u_{yy} = v,$$
$$-v_{xx} + v_{tt} + v_{yy} = u,$$

is not a coupled system of wave equations because they do not share a common direction of evolution, i.e., a common time variable. This is reflected in the corresponding cones having empty intersections

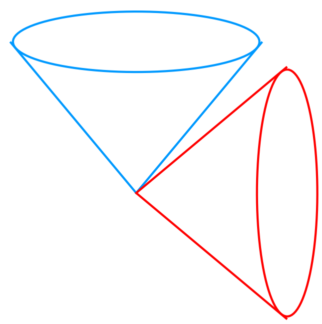


FIGURE 37. Empty Intersection

Consider, instead,

$$-u_{tt} + u_{xx} + u_{yy} = v,$$
$$-2v_{tt} - \frac{4\sqrt{3}}{11}v_{tx} + \frac{32}{121}v_{xx} + v_{yy} = u,$$

This is a system of coupled wave equations with a common evolution direction. The cones are (omitting the y variable) depicted as:

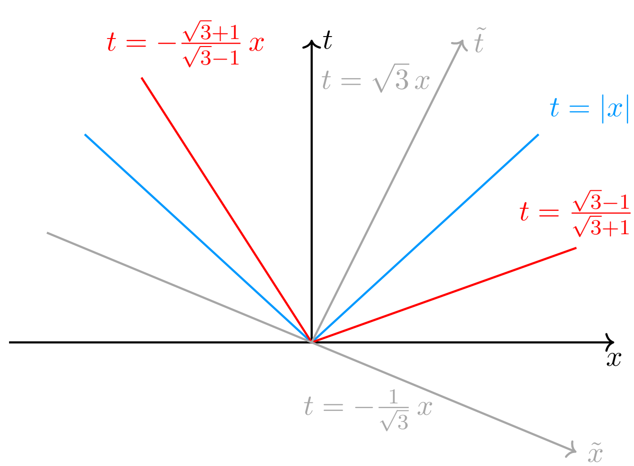


FIGURE 38. Cones with Omitted y Variable

The blue cone corresponds to u and the red cone to v. To see the latter, consider the coordinate system given by $t + \frac{1}{\sqrt{3}}x = 0$, $t - \sqrt{3} = 0$, i.e., consider the variables

$$\tilde{x} = \frac{\sqrt{3}}{2}x - \frac{1}{2}t, \quad \tilde{t} = \frac{1}{2}x + \frac{\sqrt{3}}{2}t$$

with respect to which the red cone is a standard light cone. Set

$$w(\tilde{t}, \tilde{x}) = v(t, x).$$

Then

$$\begin{aligned}
\tilde{x} + \frac{1}{\sqrt{3}}\tilde{t} &= \left(\frac{\sqrt{3}}{2} + \frac{1}{\sqrt{3}}\right)x, \quad x = \frac{2\sqrt{3}}{11}\tilde{x} + \frac{2}{11}\tilde{t} \\
\sqrt{3}\,\tilde{t} - \tilde{x} &= 2t, \quad t = \frac{\sqrt{3}}{2}\tilde{t} - \frac{1}{2}x, \quad \begin{cases} x_{\tilde{x}} = \frac{2\sqrt{3}}{11}, \ x_{\tilde{t}} = \frac{2}{11} \\ t_{\tilde{t}} = \frac{\sqrt{3}}{2}, \ t_{\tilde{x}} = -\frac{1}{2} \end{cases} \\
w_{\tilde{t}} &= v_t\, t_{\tilde{t}} + v_x\, x_{\tilde{t}} = \frac{\sqrt{3}}{2}v_t + \frac{2}{11}v_x \\
w_{\tilde{t}\tilde{t}} &= \frac{\sqrt{3}}{2}(v_{tt}t_{\tilde{t}} + v_{tx}x_{\tilde{t}}) + \frac{2}{11}(v_{xt}t_{\tilde{t}} + v_{xx}x_{\tilde{t}}) \\
&= \frac{9}{4}v_{tt} + \frac{2\sqrt{3}}{11}v_{tx} + \frac{4}{121}v_{xx} \\
w_{\tilde{x}} &= v_t t_{\tilde{x}} + v_x x_{\tilde{x}} = -\frac{1}{2}v_t + \frac{2\sqrt{3}}{11}v_x \\
w_{\tilde{x}\tilde{x}} &= -\frac{1}{2}(v_{tt}t_{\tilde{x}} + v_{tx}x_{\tilde{x}}) + \frac{2\sqrt{3}}{11}(v_{xt}t_{\tilde{x}} + v_{xx}x_{\tilde{x}}) \\
&= \frac{1}{4}v_{tt} - \frac{2\sqrt{3}}{11}v_{tx} + \frac{36}{121}v_{xx},
\end{aligned}$$

thus

$$-w_{\tilde{t}\tilde{t}} + w_{\tilde{x}\tilde{x}} = -2v_{tt} - \frac{4\sqrt{3}}{11}v_{tx} + \frac{32}{121}v_{xx},$$

so the v part is a wave equation with time $\tilde{t}$ and cones given by $\tilde{t} = |\tilde{x}|$, i.e., $\tilde{t} \geq 0$ and $\tilde{t} = \tilde{x}$, $\frac{1}{2}x + \frac{\sqrt{3}}{2}t = \frac{\sqrt{3}}{2}x - \frac{1}{2}t$, $\ t = \frac{\sqrt{3}-1}{\sqrt{3}+1}x$, $\ \tilde{t} = -\tilde{x}$, $\ \frac{1}{2}x + \frac{\sqrt{3}}{2}t = -\frac{\sqrt{3}}{2}x + \frac{1}{2}t$, $\ t = -\frac{1}{\sqrt{3}}x$.
The next definition dualizes the above constructions.

Definition 15.15. We define the **dual convex cone** $\mathcal{C}_x^+(P)$ at T_xX as the set of v's $\in T_xX$ such that $\xi(v) \geq 0$ for every $\xi \in \Gamma_x^{*,+}(P)$. We similarly define $\mathcal{C}_x^-(P)$ and set $\mathcal{C}_x(P) = \mathcal{C}_x^+(P) \cup \mathcal{C}_x^-(P)$. If the cones $\mathcal{C}_x^+(P)$ and $\mathcal{C}_x^-(P)$ can be continuously distinguished with respect to X then X is called **time oriented** and we can define a future and past time direction. A path in X is called **future (past) timelike** if its tangent at each point belongs to $\mathcal{C}_x^+(P)$ ($\mathcal{C}_x^-(P)$), and **future (past) causal** if its tangent at each point belongs or is tangent to $\mathcal{C}_x^+(P)$ ($\mathcal{C}_x^-(P)$). A regular surface $\sum$ (i.e., a surface for which tangent vectors are well defined) in X is called **spacelike** if its tangent vectors at each point are exterior to $\mathcal{C}_x(P)$.

Remark 15.16. Despite the terminology, the above definitions are made in terms of the principal part P of the operator L, without reference to a Lorentzian metric. The terminology captures the close connections among hyperbolic equations and Lorentzian geometry.

Let us now give a motivation for the definition of characteristics. Consider the linear PDE

$$\begin{aligned} Lu &= \sum_{|\alpha|\leq k} a_\alpha D^\alpha u = \sum_{|\alpha|=k} a_\alpha D^\alpha u + \sum_{|\alpha|\leq k-1} a_\alpha D^\alpha u \\ &= Pu + Qu = f, \end{aligned}$$

where the a_α's are $d \times d$ matrices (so we assume the equations to all have the same order for simplicity) and f is given. Consider the Cauchy problem for L with data given on $\{x^0 = 0\}$, i.e.,

$$\begin{aligned} Lu &= f \text{ in } [0,T] \times \mathbb{R}^n, \\ D^\alpha u(0,\cdot) &= u_\alpha \text{ on } \{x^0 = 0\} \times \mathbb{R}^n, |\alpha| \leq k-1. \end{aligned}$$

(At this point we are not assuming L to be hyperbolic or have any structure, so we do not think of x^0 as time.) If the Cauchy problem is uniquely solvable, then in particular all derivatives of u on $\sum = \{x^0 = 0\}$ are uniquely determined in terms of the initial data (and f).

Writing

$$Pu = \sum_{|\alpha|=k} a_\alpha D^\alpha u = a_{\alpha^*} D^{\alpha^*} u + \sum_{\substack{|\alpha|=k \\ \alpha \neq \alpha^*}} a_\alpha D^\alpha u,$$

where $\alpha^* = (k, 0, ..., 0)$ the equation gives

$$a_{\alpha^*} D^{\alpha^*} u\big|_\Sigma = f\big|_\Sigma - \sum_{\substack{|\alpha|=k \\ \alpha \neq \alpha^*}} a_\alpha D^\alpha u\big|_\Sigma - Qu\big|_\Sigma$$

The RHS is entirely determined by the data. Thus, for $D^{\alpha^*} u\big|_\Sigma$ to be (algebraically) determined by the data we need

$$\det(a_{\alpha^*}|_\Sigma) \neq 0$$

Observe that if we define $\phi(x) = x^0$, then $\sum$ is $\{\phi(x) = 0\}$, $\xi = d\phi = (1, 0, ..., 0)$, and $a_{\alpha^*} = a_\alpha \xi^\alpha$, so the condition becomes

$$\det(a_\alpha \xi^\alpha |_\Sigma) \neq 0.$$

Differentiating the equation and arguing as above, we can inductively algebraically find all derivatives of u along $\sum$ in terms of the data. I.e., we can formally solve the Cauchy problem if $\sum$ is non-characteristic.

Consider now the case where we give Cauchy data along a hypersurface $\sum$ given by a level set $\{\phi(x) = 0\}$, with $d\phi \neq 0$. Let us again ask whether all derivatives of u along $\sum$.

Since $d\phi \neq 0$, in the neighborhood of each $x \in \sum$ $\partial_\alpha \phi(x) \neq 0$ for some $\alpha = 0, ..., n$. Let us assume for simplicity that $\partial_0 \phi \neq 0$. Then

$$\tilde{x}^\alpha = \begin{cases} \phi(x), & \alpha = 0, \\ x^\alpha, & \alpha \neq 0 \end{cases}$$

defines a change of variables. Then

$$\frac{\partial u}{\partial x^\alpha} = \frac{\partial u}{\partial \tilde{x}^\beta} \frac{\partial \tilde{x}^\beta}{\partial x^\alpha} = M_\alpha^\beta \frac{\partial u}{\partial \tilde{x}^\beta}, \quad M_\alpha^\beta = \frac{\partial \tilde{x}^\beta}{\partial x^\alpha}.$$

We can write $D = M\tilde{D}$, $\tilde{D} = (\frac{\partial}{\partial \tilde{x}^0}, ..., \frac{\partial}{\partial \tilde{x}^n})$. Inductively we find, for $|\alpha| = k$

$$D^\alpha = (M\tilde{D})^\alpha + R^\alpha$$

where R^α is a differential operator of order $\leq k-1$ and $(M\tilde{D})^\alpha$ is obtained by considering M as a constant matrix, i.e., not applying differentiation to M when we write the several terms in $(M\tilde{D})^\alpha$, e.g.,

$$\frac{\partial}{\partial x^{\alpha_2}}\frac{\partial}{\partial x^{\alpha_1}}u = \frac{\partial}{\partial x^{\alpha_2}}\left(M^\beta_{\alpha_1}\frac{\partial u}{\partial \tilde{x}^\beta}\right) = M^\beta_{\alpha_1}\frac{\partial^2 u}{\partial \tilde{x}^\beta \partial \tilde{x}^\beta}\frac{\partial \tilde{x}^\beta}{\partial x^{\alpha_2}} + \frac{\partial M^\beta_{\alpha_1}}{\partial x^{\alpha_2}}\frac{\partial u}{\partial \tilde{x}^\beta} = M^\beta_{\alpha_1}M^\beta_{\alpha_2}\frac{\partial^2 u}{\partial \tilde{x}^\beta \partial \tilde{x}^\beta} + R_{\alpha_1\alpha_2}$$

Then, in $\tilde{x}$ coordinates P becomes

$$P = \sum_{|\alpha|=k} a_\alpha D^\alpha = \sum_{|\alpha|=k} a_\alpha (M\tilde{D})^\alpha =: \sum_{|\alpha|=k} \tilde{a}_\alpha \tilde{D}^\alpha$$

Since derivatives of order l in x translate to derivatives of order l in $\tilde{x}$ and vice-versa, the data on $\sum$ gives data on $\tilde{\sum} = \{\tilde{x}^0 = 0\}$ for the Cauchy problem in $\tilde{x}$ coordinates, and the inversibility of the coordinate transformation implies that we can determined $D^k u|_\Sigma$ in terms of the data on $\sum$ if and only if we can determine $\tilde{D}^k u|_{\tilde{\sum}}$ in terms of the data on $\tilde{\sum}$. But the latter holds (by the above case $\phi(x) = x^n$) if and only if

$$\det(\tilde{a}_{\alpha^*}) \neq 0,$$

$a^* = (k, 0, ..., 0)$. Now,

$$\tilde{a}_{(k,0,...,0)} = a_\alpha (M\tilde{\xi})^\alpha \text{ with } \tilde{\xi} = (1, 0, ..., 0),$$

i.e.,

$$(M\tilde{\xi})_\beta = M^r_\beta \tilde{\xi}_r = M^0_\beta, \;\; (M\tilde{\xi})^\alpha = (M^0_0)^{\alpha_0}...(M^0_n)^{\alpha_n}$$

(e.g., in 2d, $\alpha = (\alpha_0, \alpha_1)$

$$a_{\alpha_0\alpha_1}\frac{\partial^2 u}{\partial x^{\alpha_0}\partial x^{\alpha_1}} = M^\beta_{\alpha_0}M^r_{\alpha_1}\frac{\partial^2 u}{\partial \tilde{x}^\beta \partial \tilde{x}^r} + ... = M^0_{\alpha_0}M^0_{\alpha_1}\frac{\partial^2 u}{\partial \tilde{x}^0 \partial \tilde{x}^0} + ...$$

But $M^0_{\alpha_0} = M^\beta_{\alpha_0}\tilde{\xi}_\beta$ for $\tilde{\xi} = (1, 0, ..., 0)$.) But

$$M^0_\beta = \frac{\partial \tilde{x}^0}{\partial x^\beta} = \frac{\partial \phi}{\partial x^\beta} = \partial_\beta \phi = (d\phi)_\beta,$$

i.e.,

$$M\tilde{\xi} = d\phi = \xi$$

and the condition to determine $D^k u|_{\tilde{\Sigma}}$ reads

$$\det(\tilde{a}_{\alpha^*}) = \det(a_\alpha (M\xi)^\alpha) = \det(a_\alpha \xi^\alpha) \neq 0,$$

i.e., $\sum$ must be non-characteristic.

From this we conclude the following: Data can be freely specified only on non-characteristic hypersurfaces. If $\sum$ is characteristic, then there must be compatibility relations among the data, also called constraints. E.g., consider

$$-a(t,x)u_{tt} + u_{xx} + \partial_t u = 0 \text{ in } [0,\infty) \times \mathbb{R},$$
$$u(0,\cdot) = g,$$
$$\partial_t u(0,\cdot) = h.$$

Suppose that $a(0,x) = 0$. Then, if u is a solution, restricting to $t = 0$,

$$u_{xx}(0,x) + \partial_t u(0,x) + g_{xx} + h = 0,$$

so g and h cannot be freely specified.

The image starting with a non-characteristic hypersurface and "bending" becomes characteristic.

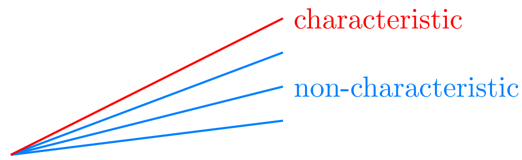


FIGURE 39. Non-characteristic to Characteristic

Since in this case we transition from the ability to freely specify data to data that is constrained, we can imagine that this means that along the characteristic the values of the derivatives of u have already been determined by the values assigned as data "before" we reached the characteristics.

For hyperbolic operators, something like this is true:

For weakly hyperbolic (thus also hyperbolic) operators, under very general conditions we can define the causal part of a point (which is the analogue of the past light cone with vertex at x) and prove a domain-of-dependence property, i.e, that a solution at x depends only on its value on the causal past of x, and the boundary of the causal part of x is a characteristic manifold (exactly like the boundary of the past light cone is characteristic for the wave operator).

It follows that we can roughly say that for hyperbolic operators, information "propagates" along the characteristics. This makes the study of the characteristics important in hyperbolic problems.

One way of understanding the above domain-of-dependence property is as follows. Suppose that we have an operator L with the property that gives Cauchy data on a subset S of non-characteristic hypersurface $\sum$, there exists a unique local solution in a neighborhood of S. Let us not worry about the precise hypotheses regularity, etc., but let us imagine that everything is "sufficiently well behaved" so that what follows makes sense. If the operator is hyperbolic, or the problem satisfies the assumptions of the Cauchy-Kovalevskaya theorem (see below) or of the Holmgren's uniqueness theorem (a type of generalization of the uniqueness part of the Cauchy-Kovalevskaya theorem), then such a situation is true. Consider given initial data, so the solution is uniquely determined in a neighborhood U of S. Let $\tilde{U}$ be the largest neighborhood of S where such solution is uniquely determined by the data. Then $\partial\tilde{U}$ is characteristic. To see this, suppose a portion $V \subset \partial\tilde{U}$ is not characteristic.

The solution $\tilde{U}$ induces data on V. Then, we can solve the problem and obtain a solution in a neighborhood $\tilde{V}$ of V that continues the solution beyond $\tilde{U}$. Since the solution in $\tilde{U}$ is uniquely determined by the data on S, so is the data induced on V, and thus the solution in $\tilde{V}$. But this would contradict the definition of $\tilde{U}$.

The above argument is only heuristic, but gives an idea of the domain-of-dependence property.

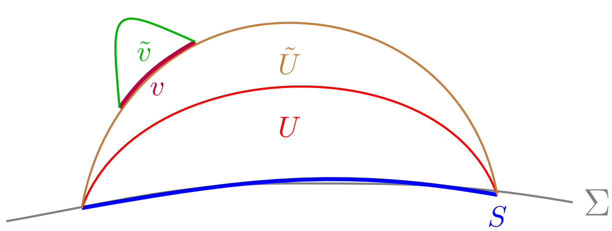


FIGURE 40. Illustration

Definition 15.17. All of the above notions generalize to quasilinear equations. In this case, given a function v, the quasilinear operator

$$(Lu)^J = h_I^J(x, u, ..., D^{m_K - u_J - 1}u, D^{m_I - n_J})u^I + b^J(x, D^{m_K - n_J - 1}u^K)$$

becomes a linear operator if we replace

$$(\tilde{L}u)^J = h_I^J(x, v, ..., D^{m_K - u_J - 1}v, D^{m_I - n_J})u^I + b^J(x, D^{m_K - n_J - 1}v^K)$$

so that all previous notions apply. In this case we talk about L being hyperbolic (at x) for a given v etc. We are particularly interested in the case when v the solution if set (if it exists) or is the initial data, which makes sense since the terms in v involve only up $m_K - n_J - 1$ derivatives of v, although in the mixed-order case more care has to be taken in defining what derivatives must be specified as data, and there might be compatibility conditions even for non-characteristic surfaces (think of previous examples of the system with indices for ϕ_1, ϕ_2, ϕ_3).

Remark 15.18. The previous argument of solving $D^k u|_\Sigma$ in terms of the data if $\sum$ is non-characteristic (which applies to quasilinear equations as well) can be used to produce solutions (in a neighborhood of $\sum$) as follows. Inductively we determine all derivatives $D^l u|_\Sigma$. Then (taking $\sum = \{x^0 = t = 0\}$ for simplicity), we consider the formal expansion.

$$u(t, x) = \sum_{|\alpha|=0}^{\infty} \frac{1}{\alpha!} D^\alpha u(0, x) t^\alpha$$

If everything in the problem is analytic, then we can show that the above series converges and is a solution. This is the content of the **Cauchy-Kovalevskaya theorem.** This procedure of solving for $D^k u(0, x)$ can also be useful for non-analytic data: although the series will not converge in general for non-analytic data, for data in **Gevrey spaces** (which is a generalization of analytic functions) we can still work with the formal power series as a consistent tool to produce solutions. More generally, we can consider the truncated series

$$u(t, x) = \sum_{|\alpha|=0}^{N} \frac{1}{\alpha!} D^\alpha u(0, x) t^\alpha, N >> 1$$

to obtain an approximate solution. Such an idea is often useful when dealing with a technique called the **Nash-Moser iteration**, which can often be applied to PDEs for which energy/a priori estimates are not directly available.

16. Einstein's equations

Einstein's equations are the fundamental equations of general relativity. The basic problem we are interested is the following: find a Lorentzian manifold (M, g) where **Einstein's equations**

$$\text{Ric}(g) - \frac{1}{2} R(g) d = T,$$

or, in coordinates,

$$R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} = \mathcal{T}_{\alpha\beta}$$

are satisfied. Here, $R_{\alpha\beta} = (\text{Ric}(g))_{\alpha\beta}$ is the Ricci curvature of the Lorentzian metric g, R is the scalar curvature of g, and $\mathcal{T}$ is the **energy-momentum tensor** of matter, which contains information about matter and energy interacting with gravity and depends on the problem (e.g., if we are studying a fluid interacting with gravity, $\mathcal{T}$ has a certain form; if we are studying electromagnetic fields interacting with gravity $\mathcal{T}$ has another form; see further below for examples).
When $\mathcal{T} = 0$, we have the **vacuum Einstein equations**

$$R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} = 0.$$

In relativity, a vacuum can be dynamic and quite complex, and we should not think of it as "empty space where nothing happens." We sometimes use the term **matter Einstein equations** or **Einstein equations with matter** to refer to the case $\mathcal{T} \neq 0$. The terminology "matter" is used because in general relativity we call matter anything that is not gravity (so electromagnetic radiation would be called matter).

Although Einstein's equation can be studied in any dimension, we will consider only $\underline{\dim(M) = 4}$, which is the case of most physical interest.

Taking the trace of Einstein's equations,

$$\operatorname{tr}(\mathcal{T}) = g^{\alpha\beta}\mathcal{T}_{\alpha\beta} = -R.$$

We can thus equivalently write

$$R_{\alpha\beta} = \mathcal{T}_{\alpha\beta} - \frac{1}{2}\operatorname{tr}(\mathcal{T})g_{\alpha\beta}.$$

In particular, in the case of vacuum, the Einstein equations can be written as

$$R_{\alpha\beta} = 0.$$

Thus, the vacuum Einstein equations correspond to the geometric problem of finding Ricci-flat Lorentzian four-manifolds.

We will work in local coordinates with $\{x^\alpha\}_{\alpha=0}^3$ an arbitrary coordinate system (when needed, later on, we will specify a specific coordinate system). Since we do not have a canonical system of coordinates, we do not at this point think of x^0 as a time coordinate (although there will be a way of constructing a time coordinate later on).

Most of the features we are interested in are already present in the vacuum case, so we consider this case in detail (later we comment on the case with matter), and by "Einstein" we will mean "vacuum Einstein" when there is no confusion.

Since Ricci involves up to two derivatives of the metric, we see $R_{\alpha\beta}$ as a second-order differential operator acting on (the component) $g_{\alpha\beta}$. Thus, Einstein equations are a system of second-order nonlinear PDEs (in fact, quasilinear, see below) for $g_{\alpha\beta}$.

We are interested in the Cauchy problem for Einstein's equations. For this, we need to understand which hypersurfaces are appropriate to prescribe initial data, i.e., which hypersurfaces are non-characteristic. A direct computation using the definition of Ricci curvature gives

$$\begin{aligned} R_{\alpha\beta} =& -\frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\nu g_{\alpha\beta} - \frac{1}{2}g^{\mu\nu}\partial_\alpha\partial_\beta g_{\mu\nu} \\ &+ \frac{1}{2}g^{\mu\nu}\partial_\alpha\partial_\nu g_{\mu\beta} + \frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\beta g_{\alpha\nu} + \overbrace{H_{\alpha\beta}(\partial g)}^{\text{terms involving up to } \partial g}. \\ =:& \tilde{R}_{\alpha\beta}(\partial^2 g) + H_{\alpha\beta}(\partial g) \end{aligned}$$

To make the notation more clear, first consider α_β at a metric h (recall the remarks on the definition of characteristics for quasilinear problems):

$$\begin{aligned} \tilde{R}_{\alpha\beta}(\partial^2 g) =& -\frac{1}{2}h^{\mu\nu}\partial_\mu\partial_\nu g_{\alpha\beta} - \frac{1}{2}h^{\mu\nu}\partial_\alpha\partial_\beta g_{\mu\nu} \\ &+ \frac{1}{2}h^{\mu\nu}\partial_\alpha\partial_\nu g_{\mu\beta} + \frac{1}{2}h^{\mu\nu}\partial_\mu\partial_\beta g_{\alpha\nu} \end{aligned}$$

This is to viewed as an operator acting on the unknown variable g, which is organized as a vector valued function (the function u taking values in $\mathbb{R}^d$ in the previous notation) which we write as the ten-component vector

$$u = (g_{00}, g_{01}, ..., g_{33}) .$$

(There are 16 components $g_{\alpha\beta}$, but it suffices to consider 10 independent components since $g_{\alpha\beta} = g_{\beta\alpha}$). Thus, the characteristic matrix of $\tilde{R}$ at h acting on u is

$$\begin{aligned}(P(h,\xi)u)_{\alpha\beta} = &- \frac{1}{2}h^{\mu\nu}\xi_\mu\xi_\nu g_{\alpha\beta} - \frac{1}{2}h^{\mu\nu}\xi_\alpha\xi_\beta g_{\mu\nu} \\ &+ \frac{1}{2}h^{\mu\nu}\xi_\alpha\xi_\nu g_{\mu\beta} + \frac{1}{2}h^{\mu\nu}\xi_\mu\xi_\beta g_{\alpha\nu}.\end{aligned}$$

For any h and any ξ, $P(h,\xi)$ always has a kernel, as we can always pick u with entries

$$g_{\alpha\beta} = \xi_\alpha\xi_\beta, \quad \xi \neq 0$$

so

$$(P(h,\xi)u)_{\alpha\beta} = 0,$$

so

$$\det P(h,\xi) = 0.$$

Hence, every hypersurface is characteristic for Einstein's equations. This can be viewed as a consequence of the diffeomorphism invariance of the equations. One way of seeing this is as follows. Suppose that the Cauchy problem can be uniquely solved for data given on a hypersurface $\sum$, with the solution defined on a neighborhood U of $\sum$. So we have a unique g satisfying

$$\mathrm{Ric}(g) = 0$$

in U and taking the correct data (we leave aside for a moment what it means to prescribe data for Einstein's equations, but at least considering them as PDEs for the components $g_{\alpha\beta}$ in local coordinates, it is not too difficult to make sense of it). Take a diffeomorphism $\varphi : U \to U$ such that $\varphi =$ identity on a neighborhood $\tilde{U}$ of $\sum$ with $\tilde{U}$ properly contained in U. Set $h = \varphi^*(g)$. Then

$$\mathrm{Ric}(h) = 0.$$

Moreover, $h = g$ on $\sum$ (and their derivatives also agree on a coordinate chart), but $h \neq g$ in U. Thus, h and g are two different solutions to the Cauchy problem.

Of course, h and g are isometric, but from a "purely PDE" point of view, they are different metrics hence different solutions. This tells us that when we consider solutions to Einstein's equations we will have to do it up to isometries.

16.1. **The constraint equations.** The fact that every hypersurface is characteristics for Einstein's equations implies that initial data cannot be prescribed arbitrarily. To understand the constraints that the initial data has to satisfy, consider a time oriented Lorentzian manifold (M, g) with an embedded smooth spacelike hypersurface $\sum$. We require $\sum$ to be spacelike because it should play the role of the $t = 0$ surface where we prescribe data. Let $\bar{g}$ and $\bar{h}$ the induced metric on $\sum$ and its second fundamental form (M, g). Then $\bar{g}$ and $\bar{h}$ must satisfy the Gauss-Codazzi equations:

$$\bar{R}^\delta_{\alpha\beta\gamma} = \bar{g}^J_\alpha \, \bar{g}^\sigma_\beta \, \bar{g}^\tau_\gamma \, \bar{g}^\delta_\theta \, R^\theta_{J\sigma\tau} - \bar{h}_{\alpha\gamma} \, \bar{h}^\delta_\beta + \bar{h}_{\beta\gamma} \, \bar{h}^\delta_\alpha,$$

$$\bar{\nabla}_\alpha \bar{h}^\alpha_\beta - \bar{\nabla}_\beta \, \bar{h}^\alpha_\alpha = R_{\gamma\delta} N^\delta \bar{g}^\gamma \beta,$$

where $\bar{R}$ is the Riemann curvature of $\bar{g}$, $\bar{\nabla}$ the covariant derivative associated with $\bar{g}$, N is the future-pointing unit normal to $\sum$, and indices are raised and lowered with g.
Note that $\bar{g}_{\alpha\beta} = g_{\alpha\beta} + N_\alpha N_\beta$ since

$$\bar{g}_{\alpha\beta}N^\beta = N_\alpha + N_\alpha \underbrace{N_\beta N^\beta}_{=-1} = 0,$$

so given a vector field X we have

$$X^\alpha = \underbrace{-X^\beta N_\beta N^\alpha}_{\text{projection onto } N} + \underbrace{\bar{g}^\alpha_\beta X^\beta}_{\substack{\text{(projection onto} \\ \text{orthogonal to } N) = \bar{X}}}$$

so that

$$\begin{aligned} g_{\alpha\beta}\bar{X}^\alpha \bar{X}^\beta &= g_{\alpha\beta}\, \bar{g}^\alpha_\gamma\, X^\gamma\, \bar{g}^\beta_\delta\, X^\delta \\ &= g_{\alpha\beta}(g^\alpha_\gamma + N^\alpha N_\gamma)(g^\beta_\delta + N^\beta N_\delta)X^\gamma X^\delta \\ &= (g_{\beta\gamma} + N_\beta N_\gamma)(g^\beta_\delta + N^\beta N_\delta)X^\gamma X^\delta \\ &= (g_{\beta\delta} + N_\gamma N_\delta + N_\delta N_\gamma - N_\gamma N_\delta)X^\gamma X^\delta \\ &= (g_{\beta\delta} + N_\gamma N_\delta)X^\gamma X^\delta \\ &= \bar{g}_{\alpha\beta}\, X^\gamma X^\delta \end{aligned}$$

as should be for the induced metric.

Observe that

$$\bar{R}_{\beta\delta} = \bar{g}^{\alpha\gamma}\bar{R}_{\alpha\beta\gamma\delta},$$
$$\bar{R} = \bar{g}^{\beta\delta}\bar{R}_{\beta\delta},$$

where $\bar{R}_{\alpha\beta}$ and $\bar{R}$ are the Ricci and scalar curvature of $\bar{g}$. Thus, if g satisfies Einstein's equations,

$$\begin{aligned} \bar{R} &= \bar{g}^{\alpha\gamma}\, \bar{g}^{\beta\delta}\, \bar{R}_{\alpha\beta\gamma\delta} \\ &= \bar{g}^{\alpha\gamma}\bar{g}^{\beta\delta}\left(\bar{g}^J_\alpha\, \bar{g}^\sigma_\beta\, \bar{g}^\tau_\gamma\, \bar{g}_{\delta\theta}\, R^\theta_{J\sigma\tau} - \bar{h}_{\alpha\gamma}\, \bar{h}_{\beta\delta} + \bar{h}_{\beta\gamma}\, \bar{h}_{\alpha\delta}\right) \\ &= \underbrace{\bar{g}^{J\tau}\bar{g}^\sigma_\theta R^\theta_{J\sigma\tau}}_{=(g^{J\tau}+N^{J\tau})(g^\sigma_\theta+N^\sigma N_\theta)R^\theta_{J\sigma\tau}} - (\bar{g}^{\alpha\beta}\bar{h}_{\alpha\gamma})^2 + \bar{g}^{\alpha\gamma}\, \bar{g}^{\beta\delta}\, \bar{h}_{\beta\gamma}\, \bar{h}_{\alpha\delta} \\ &= \underbrace{g^{J\tau}g^\sigma_\theta R_{J\delta\tau\theta}}_{R} + N^\sigma N_\theta R^\theta_\sigma + N^J N^\tau g^\sigma_\theta \underbrace{R^\theta_{J\sigma\tau}}_{=-R^\theta_{J\sigma\tau}=+R^\theta_{J\sigma\tau}} + \underbrace{N^J N^\tau N^\sigma N_\theta R^\theta_{J\sigma\tau}}_{=0 \text{ by } R^\theta_{J\sigma\tau}=-R^\theta_{\sigma J\tau}} \\ &= R + 2R_{\alpha\beta}N^\alpha N^\beta = 2(R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta})N^\alpha N^\beta = 0, \end{aligned}$$

where we used $\bar{g}_{\alpha\beta}\bar{g}^\beta_\delta = \bar{g}_{\alpha\delta}$ and similar identities that can be verified directly.

So

$$\bar{R} + (\bar{h}^{\bar{\alpha}}_{\bar{\alpha}})^2 - \bar{h}_{\bar{\alpha}\bar{\beta}}\bar{h}^{\bar{\alpha}\bar{\beta}} = 0,$$

where the barred indices indicated that contraction is with respect to $\bar{g}$. Also:

$$\bar{\nabla}_\alpha \bar{h}^\alpha_\beta - \bar{\nabla}_\beta \bar{h}^\alpha_\alpha = R_{\gamma\delta}N^\delta \bar{g}^\gamma_\beta = 0.$$

But

$$\begin{aligned} g^{\alpha\gamma}\bar{\nabla}_\alpha \bar{h}_{\gamma\beta} - g^{\alpha\gamma}\bar{\nabla}_\beta \bar{h}_{\alpha\gamma} &= (\bar{g}^{\alpha\gamma} - N^\alpha N^\gamma)\bar{\nabla}_\alpha \bar{h}_{\gamma\beta} - (\bar{g}^{\alpha\gamma} - \bar{N}^\alpha \bar{N}^\gamma)\bar{\nabla}_\beta \bar{h}_{\alpha\gamma} \\ &= \bar{\nabla}_{\bar{\alpha}}\bar{h}^{\bar{\alpha}}_\beta - \bar{\nabla}_\beta \bar{h}^{\bar{\alpha}}_{\bar{\alpha}} - N^\alpha N^\gamma \bar{\nabla}_\alpha \bar{h}_{\gamma\beta} + N^\alpha N^\gamma \bar{\nabla}_\beta \bar{h}_{\alpha\gamma}, \end{aligned}$$

compute

$$\begin{aligned}
-N^\alpha N^\gamma \bar{\nabla}_\alpha \bar{h}_{\gamma\beta} + N^\alpha N^\gamma \bar{\nabla}_\beta \bar{h}_{\alpha\gamma} &= N^\alpha N^\gamma (\bar{\nabla}_\beta \bar{h}_{\alpha\gamma} - \bar{\nabla}_\alpha \bar{h}_{\gamma\beta}) \\
&= N^\alpha N^\gamma (\bar{g}^J_\beta \nabla_J \bar{h}_{\alpha\gamma} - \bar{g}^J_\alpha \nabla_J \bar{h}_{\gamma\beta}) \\
&= N^\alpha N^\gamma (\bar{g}^J_\beta \nabla_J \nabla_\alpha N_\gamma - \bar{g}^J_\alpha \nabla_J \nabla \gamma N_\beta) \\
&= 0,
\end{aligned}$$

where we used

$$\bar{\nabla}_\beta \bar{h}_{\alpha\gamma} = \bar{g}^J_\beta \nabla_J \bar{h}_{\alpha\gamma}, \quad \bar{h}_{\alpha\beta} = \nabla_\alpha N_\beta$$

$$N_\gamma N^\gamma = -1 \implies N^\gamma \nabla_\alpha N_\gamma = 0, \quad N^\alpha \bar{g}^J_\alpha = 0.$$

Thus,

$$\bar{\nabla}_{\bar{\alpha}} \bar{h}^{\bar{\alpha}}_\beta - \bar{\nabla}_\beta \bar{h}^{\bar{\alpha}}_{\bar{\alpha}} = 0.$$

We are led to the following.

Definition 16.1. An **initial data set** for the (vacuum) Einstein equations is a triple $(\sum, g_0, h)$ where $\sum$ is a three-dimensional manifold, g_0 is a Riemannian metric $\sum$, and h is a symmetric two tensor on $\sum$, such that the **Einstein constraint equations,**

$$\begin{aligned}
R_{g_0} - |h|^2_{g_0} + (\mathrm{tr}_{g_0} h)^2 = 0, \\
\nabla_{g_0} \mathrm{tr}_{g_0} h - \mathrm{div}_{g_0} h = 0,
\end{aligned}$$

are satisfied. The first equation is called the **Hamiltonian constraint** and the second the **momentum constraint**.

(Above, the notation should be self-explanatory, but $R_{g_0}, tr_{g_0}, |\cdot|_{g_0}, \nabla_{g_0}$, and div_{g_0} are, respectively, the scalar curvature, trace, norm, covariant derivative, and divergence with respect to g_0.)

Definition 16.2. The Cauchy problem for (vacuum) Einstein's equations consists of the following. Given an initial data set $(\sum, g_0, h)$, find a Lorentzian manifold (M, g) where Einstein's equations are satisfied and where $(\sum, g_0)$ embeds isometrically with second fundamental form h.

Several remarks are in order.

We roughly think of $\sum$ as playing the role of the $\{t = 0\}$ surface, g_0 and h as $g|_{t=0}$ and $\partial_t g|_{t=0}$, although even in a heuristic sense this cannot be quite correct. If we consider the equations in local coordinates, we should be prescribing $g_{\alpha\beta}|_{t=0}$ and $\partial_t g_{\alpha\beta}|_{t=0}$ for the independent $g_{\alpha\beta}$, but since g_0 and h are symmetric tensors on a <u>three</u>-dimensional manifold, they only have six independent components, so we can only prescribe (taking coordinates where $x^0 = t = 0$ represents $\sum$) $g_{ij}|_{t=0} = (g_0)_{ij}$, $\partial_t g_{ij}|_{t=0} = h_{ij}$. We will see that the missing components $g_{0\mu}, \partial_t g_{0\mu}$ can be chosen more or less freely. This is what physicists call gauge freedom, and it reflects the fact that Einstein's equations are geometric, i.e., invariant by diffeomorphisms. Thinking of diffeomorphisms as giving locally a change of coordinates, we have four coordinate functions $\{x^\alpha\}^3_{\alpha=0}$ that we can freely reparametrize to fix $g_{0\mu}$ the way we want.

Before investigating the Cauchy problem, we need to make sure initial data exists, i.e., that is is possible to find g_0 and h satisfying the constraints. This turns out to be a problem on its own. When appropriately formulated, the constraint equations become a system of nonlinear elliptic equations for (g_0, h). The solvability of this system depends, among other things, on the topology of $\sum$. Moreover, often we want more than just satisfying the constraints. For example, if $\sum$ is not compact, we would like to find (g_0, h) with prescribed asymptotics.

We will not discuss how to solve the constraints. It suffices to know that there are plenty of situations where this can be done, so statements about initial data sets are not vacuous. The study of the constraint equations has a long history and continues to be an active field of research.

16.2. **The Cauchy problem.** In this section, if regularity is omitted, it means C^∞.

We will use a special system of coordinates to study the Cauchy problem.

Definition 16.3. A coordinate system $\{x^\alpha\}_{\alpha=0}^3$ in a Lorentzian manifold (M, g) is said to form **wave coordinates** if

$$\Box_g x^\alpha = 0, \quad \alpha = 0, ..., 3.$$

$\Box_g$ is the wave operator with respect to x^α, i.e., $\Box_g = \nabla^\mu \nabla_\mu$. $\Box_g x^\alpha$ means $\Box_g$ acting on the scalar function x^α for each α.

Remark 16.4. Notice that wave coordinates depend on the metric.

Remark 16.5. In the Riemannian case, $\nabla_\mu \nabla^\mu$ is the Laplacian of g, so the corresponding coordinates are called harmonic coordinates. We sometimes use the term harmonic even in the Lorentzian set. We have

$$\begin{aligned}
\Box_g x^\alpha &= \nabla_\mu \nabla^\mu x^\alpha \\
&= g^{\mu\nu} \nabla_\mu \nabla_\nu x^\alpha \\
&= g^{\mu\nu} (\partial_\mu \nabla_\nu x^\alpha - \Gamma^\lambda_{\mu\nu} \nabla_\lambda x^\alpha) \\
&= g^{\mu\nu} (\partial_\mu \underbrace{\partial_\nu x^\alpha}_{\delta^\alpha_\nu} - \Gamma^\lambda_{\mu\nu} \underbrace{\partial_\lambda x^\alpha}_{\delta^\alpha_\lambda}) \\
&= -g^{\mu\nu} \Gamma^\alpha_{\mu\nu},
\end{aligned}$$

hence harmonic coordinates can be also characterized by

$$g^{\mu\nu} \Gamma^\alpha_{\mu\nu} = 0.$$

Lemma 16.6. *It is always possible to construct wave coordinates in the neighborhood of a point.*

Proof. Consider coordinates $\{x^\alpha\}_{\alpha=0}^3$ about a point p. We can assume $x^\alpha(p) = 0$ and identify a neighborhood of p where the x^α's are defined with a domain U containing the origins in $\mathbb{R}^4$, with the x^α's identified with the corresponding coordinates in $\mathbb{R}^4$. Denote $t := x^0$.

Consider, for each $\alpha = 0, ..., 3$, the Cauchy problem

$$\begin{aligned}
\Box_g y^i &= 0 \text{ in } \mathbb{R} \times \mathbb{R}^3, \\
y^i(0, x^1, x^2, x^3) &= x^i \text{ on } \{t = 0\} \times \mathbb{R}^3, \\
\partial_t y^i(0, x^1, x^2, x^3) &= 0 \text{ on } \{t = 0\} \times \mathbb{R}^3,
\end{aligned}$$

$$\begin{aligned}
\Box_g y^0 &= 0 \text{ in } \mathbb{R} \times \mathbb{R}^3, \\
y^0(0, x^1, x^2, x^3) &= 0 \text{ on } \{t = 0\} \times \mathbb{R}^3, \\
\partial_t y^0(0, x^1, x^2, x^3) &= 1 \text{ on } \{t = 0\} \times \mathbb{R}^3,
\end{aligned}$$

$\Box_g$ computed with respect to x^α coordinates.

(If the x^α's were wave coordinates, then $y^\alpha = x^\alpha$ would be solutions). This problem admits a smooth solution y^α. In a sufficiently small neighborhood of the origin the y^α form a coordinate system by the implicit function theorem (since they agree with x^α on the initial slice). But since $\Box_g y^\alpha$ is coordinate invariant, it also holds that

$$\Box_g y^\alpha = 0$$

with $\Box_g$ coupled with respect to y-coordinates.

$\Box$

We will now derive a useful identity. Set

$$\Gamma^\alpha = g^{\mu\nu}\Gamma^\alpha_{\mu\nu}$$

Then

$$\begin{aligned} 2\Gamma^\alpha &= g^{\mu\nu} g^{\alpha\lambda}(\partial_\mu g_{\lambda\nu} + \partial_\nu g_{\lambda\mu} - \partial_\lambda g_{\mu\nu}), \\ 2 g_{\sigma\alpha}\Gamma^\alpha &= g^{\mu\nu}(\partial_\mu g_{\alpha\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}) \\ &= 2g^{\mu\nu}\partial_\mu g_{\sigma\nu} - g^{\mu\nu}\partial_\sigma g_{\mu\nu}. \end{aligned}$$

Thus

$$2\partial_\tau(g_{\sigma\alpha}\Gamma^\alpha) = 2g^{\mu\nu}\partial_\mu\partial_\tau g_{\sigma\nu} - g^{\mu\nu}\partial_\tau\partial_\sigma g_{\mu\nu} + \underbrace{H_{\tau\sigma}(\partial g)}_{\text{depending on at most } \partial g}$$

Switching the roles of τ and σ and adding the resulting expressions,

$$2\partial_\tau(g_{\sigma\alpha}\Gamma^\alpha) + 2\partial_\sigma(g_{\tau\alpha}\Gamma^\alpha) = 2g^{\mu\nu}\partial_\mu\partial_\tau g_{\sigma\nu} + 2\partial_\mu\partial_\sigma g_{\tau\nu} - 2g^{\mu\nu}\partial_\tau\partial_\sigma g_{\mu\nu} + H_{\tau\sigma}(\partial g)$$

Recalling that

$$R_{\alpha\beta} = -\frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\nu g_{\alpha\beta} - \frac{1}{2}g^{\mu\nu}\partial_\alpha\partial_\beta g_{\mu\nu} + \frac{1}{2}g^{\mu\nu}\partial_\alpha\partial_\nu g_{\mu\beta} + \frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\beta g_{\alpha\nu} + H_{\alpha\beta}(\partial g).$$

Comparing to the above

$$R_{\alpha\beta} = -\frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\nu g_{\alpha\beta} + \frac{1}{2}\big[\partial_\alpha(g_{\beta\mu}\Gamma^\mu) + \partial_\beta(g_{\alpha\mu}\Gamma^\mu)\big] + H_{\alpha\beta}(\partial g).$$

In wave coordinates, the term in brackets vanishes, so

$$R_{\alpha\beta} = -\frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\nu g_{\alpha\beta} + H_{\alpha\beta}(\partial g).$$

The principal part now is a diagonal matrix with entries $-\frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\nu$; i.e., Einstein's equations in wave coordinates read

$$-\frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\nu g_{\alpha\beta} + H_{\alpha\beta}(\partial g) = 0$$

which is a system of quasilinear wave equations for which we have a local existence and uniqueness theorem. The problem, however, is that wave coordinates depend on the metric, i.e., we need g to construct wave coordinates, but g is what we are trying to solve in the first place. We will see how to overcome this difficulty in the next theorem using a nice trick due to Choquet-Bruhat in the 50's, when she established the first local existence and uniqueness result for Einstein's equations.

Definition 16.7. We call $R^{\text{red}}_{\alpha\beta} := -\frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\nu g_{\alpha\beta}$ the **reduced Ricci tensor** and $R^{\text{red}}_{\alpha\beta} + H_{\alpha\beta}(\partial g) = 0$ the **reduced (vacuum) Einstein equation.**

Theorem 16.8. *Let $\mathcal{I} = (\sum, g_0, h)$ be a smooth initial data set for the vacuum Einstein equations. Then there exists a solution to the Cauchy problem for the vacuum Einstein equations with data $\mathcal{I}$.*

Remark 16.9. We sometimes call a Lorentzian manifold (M, g) that is a solution with data $\mathcal{I}$ a **development** or **Einsteinian development** of $\mathcal{I}$.

Proof. Consider $\mathbb{R} \times \sum$, let $p \in \sum$, $\{x^i\}_{i=1}^3$ be coordinates on an open set U about p in $\sum$ and define coordinates on an open set $\tilde{U}$ about p in $\mathbb{R} \times \sum$, with $\tilde{U} \cap \sum \subset U$, by $\{x^\alpha\}_{\alpha=0}^3$, with $t := x^0$ a coordinate on $\mathbb{R}$. We can identify $\tilde{U}$ with an open set in $\mathbb{R} \times \mathbb{R}^3$ and p with the origin. In order to apply our theorem for smooth solutions to quasilinear wave equations, we need to formulate the problem in $[0, T] \times \mathbb{R}^3$, have compactly supported data, and guarantee that that the principal part is always a metric even when the data vanishes.

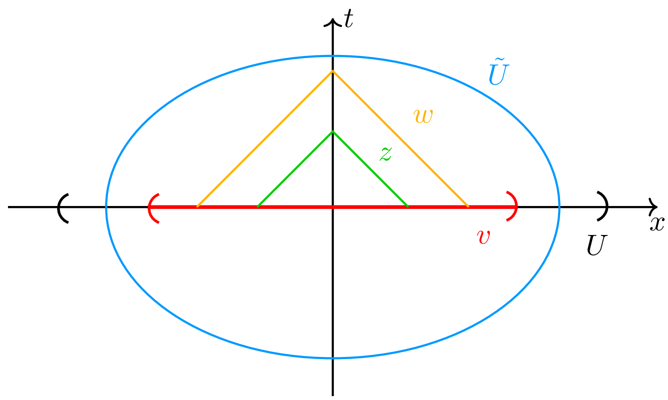


FIGURE 41. $V \subset\subset \tilde{U} \cap \{t = 0\}$

Let $V \subset\subset \tilde{U} \cap \{t = 0\}$, $\varphi \in C_c^\infty(\mathbb{R}^3)$ be such that $0 \leq \varphi \leq 1$, $\varphi = 1$ in V, $\varphi = 0$ outside $\tilde{U} \cap \{t = 0\}$. Consider the following initial data on $\{t = 0\}$:

$$
\begin{aligned}
g_{ij}(0, \cdot) &= \varphi (g_0)_{ij}, \\
g_{00}(0, \cdot) &= -\varphi, \\
g_{0i}(0, \cdot) &= 0, \\
\partial_t g_{ij}(0, \cdot) &= \varphi k_{ij}.
\end{aligned}
$$

To specify $\partial_t g_{i\mu}(0, \cdot)$, recall

$$
\begin{aligned}
g_{\sigma\alpha}\Gamma^\alpha &= g^{\mu\nu}\partial_\mu g_{\sigma\nu} - \frac{1}{2} g^{\mu\nu}\partial_\sigma g_{\mu\nu} \\
&= g^{00}\partial_t g_{\sigma 0} + g^{0i}\partial_t g_{\sigma i} + g^{i\nu}\partial_i g_{\sigma\nu} - \frac{1}{2} g^{00}\partial_\sigma g_{00} - g^{0i}\partial_\sigma g_{0i} - \frac{1}{2} g^{ij}\partial_\sigma g_{ij}.
\end{aligned}
$$

For $g_{00} = -1, g_{0i} = 0$ (so $g^{00} = -1, g^{0i} = 0$) the RHS reads

$$
-\partial_t g_{\sigma 0} + \frac{1}{2}\partial_\sigma g_{00} + D_\sigma
$$

where D_σ (for "data") is $D_\sigma = g^{i\nu}\partial_i g_{\sigma\nu} - \frac{1}{2} g^{ij}\partial_\sigma g_{ij}$. Thus, we can choose $\partial_t g_{0\mu}$ such that $\Gamma^\alpha = 0$ on V. Indeed

$$
\sigma = i: \quad -\partial_t g_{i0} + \frac{1}{2}\underbrace{\partial_i g_{00}}_{=0} + D_i = 0 \implies \partial_t g_{i0} = D_i.
$$

$$
\sigma = 0: \quad -\partial_t g_{00} + \frac{1}{2}\partial_i g_{00} + D_0 = 0 \implies \partial_t g_{00} = 2D_0.
$$

where D_i and D_0 are already known from the previous data choice.

Hence, we put

$$
\partial_t g_{0\mu}(0, \cdot) = \varphi \times (\text{above choice}).
$$

This gives $\Gamma^\alpha = 0$ on V.

Let $\tilde{g} = \tilde{g}(g^{\alpha\beta})$ (which has unknown $U = (g_{00}, ..., g_{33})$ in the previous language) be such that $\tilde{g}(0) =$ Minkowski, $\tilde{g}(g_{\alpha\beta}) = g_{\alpha\beta}$ in a compact neighborhood of $g_{\alpha\beta} = g_{\alpha\beta}(0, \cdot)|_V$, where $g_{\alpha\beta}(0, \cdot)$ is as above. We can also assume that the derivatives of $\tilde{g}$ with respect to its arguments are bounded.

Under these conditions, there exists a unique smooth solution $g_{\alpha\beta}$ to

$$-\frac{1}{2}\tilde{g}^{\mu\nu}\partial_\mu\partial_\nu g_{\alpha\beta} + H_{\alpha\beta}(\partial g) = 0$$

defined on some time interval $[0, T]$. We can take T so small so that g is a Lorentzian metric (since it is one at $t = 0$). By domain of dependence considerations and our choice of $\tilde{g}$, we have $\tilde{g} = g$ in some sufficiently small neighborhood W of p. Therefore, we obtained a solution to the reduced Einstein equation in W taking the correct data on $W \cap \{t = 0\}$. It remains to show that this is a solution to the full (i.e., non-reduced) Einstein equations in W.

We know that $\Gamma^\alpha = 0$ on V. A computation using the constraint equations, which we leave as an exercise, shows that $\partial_t\Gamma^\alpha = 0$ on V. Since g is a metric in W, we can consider its Ricci and scalar curvatures, and the Bianchi identities give

$$\nabla_\alpha(R^\alpha_\beta - \frac{1}{2}Rg^\alpha_\beta) = 0.$$

Using that $R^{\text{red}}_{\alpha\beta} = 0$, another computation that we leave as an exercise gives

$$-\frac{1}{2}g_{\beta\alpha}g^{\mu\nu}\partial_\mu\partial_\nu\Gamma^\alpha + h_\beta = 0$$

or

$$-\frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\nu\Gamma^\alpha + g^{\alpha\beta}h_\beta = 0$$

where h_β is linear in Γ^α, $\alpha = 0, ..., 3$. We thus see that the Γ^α are solutions to a linear system of wave equations in W. After properly setting this system as a system in $[0, T] \times \mathbb{R}^3$ with ideas similar as above, by uniqueness and domain of dependence considerations we conclude that $\Gamma^\alpha = 0$ in a domain $Z \subset W$ with $Z \cap W \subset V$, since $\Gamma^\alpha, \partial_t\Gamma^\alpha = 0$ on V. But if $\Gamma^\alpha = 0$ then Ricci = Ricci$^{\text{red}}$, and our solution to the reduced equations in Z is in fact a solution to the full Einstein equations. (In other words, we showed that the coordinates $\{x^\alpha\}$ we have using are in fact wave coordinates in Z. We did this by showing that if the coordinates are wave coordinates at $t = 0$, which as seen we can arrange to be the case, then they remain wave coordinates for $t > 0$. This procedure is sometimes called "propagation of the gauge").

Before continuing, we make the following observation. By domain of dependence properties and our discussion of characteristics, the solution at a point $r \in Z$ in wave coordinates depends only on the data $J^-(r) \cap \sum$, where $J^-(r)$ is the causal part of r with respect to g expressed in wave coordinates. Using that our solution is a solution to the full Einstein equations in W and that the causal part is invariantly defined, we can show that we have a solution to Einstein's equations in W regardless of the coordinate system and that the solution at a point r depends on the data $J^-(r) \cap \sum$.

Summarizing, given a point $p \in \{t = 0\} \times \sum$, we obtained a solution to Einstein's equations in a neighborhood $Z_p \in \mathbb{R} \times \sum$ of p taking the correct data on $z \cap \sum$. Moreover, the solution at $r \in Z_p$ depends only on $J^-(r) \cap \sum$, where we can assume that $J^-(r) \in Z_p$ for all $r \in Z_p$.

The next step will be to define (M, g) as the union of all such Z_p's. For this, we need to show that solutions agree on intersections. More precisely, let us show the following.

Let $q, r \in \sum$ and let Z_q, Z_r, be corresponding neighborhoods as above. Denote g_q and and g_r the corresponding solutions. Assume that $(Z_q \cap \sum) \cap (Z_r \cap \sum) \neq \emptyset$. Then, for any $w \in (Z_q \cap \sum) \cap (Z_r \cap \sum)$ there must exist neighborhoods U_q and U_r of w in Z_q and Z_r, respectively, and a diffeomorphism $\psi : U_q \to U_r$ such that $g_q = \psi^*(g_r)$.

Take normal coordinates $\{y^i\}$ at w relative to g_0. We can assume the normal coordinates to be defined inside $Z_q \cap \sum$. Construct, as in a previous lemma, wave coordinates $\{x^\alpha\}_{\alpha=0}^3$ in a neighborhood U_q of w in Z_q. Then $(x^{-1})^*(g_q)$ satisfies the reduced Einstein equations in a neighborhood of the origin in $\mathbb{R} \times \mathbb{R}^3$. We carry a similar construction of wave coordinates $\{\tilde{x}^\alpha\}_{\alpha=0}^3$ a neighborhood of U_r of w in Z_r and obtain a solution $(\tilde{x}^{-1})^*(g_r)$ to the reduced Einstein equations in a neighborhood of the origin in $\mathbb{R} \times \mathbb{R}^3$. Let W be the intersection of both neighborhoods of the origin just mentioned. Because $\{y^i\}$ is intrinsically determined by g_0 and Z_q and Z_r induce on $(Z_q \cap \sum) \cap (Z_r \cap \sum)$ the same data, x and $\tilde{x}$ agree on $\{t = 0\} \times V$, where V is some neighborhood of w in $\sum$. Thus, $(x^{-1})^*(g_q)$ and $(\tilde{x}^{-1})^*(g_r)$ are both solutions to the reduced Einstein equations in W with the same data on $\{t = 0\} \cap W$. Therefore we have $(x^{-1})^*(g_q) = (\tilde{x}^{-1})^*(g_r)$ (possibly shrinking W for uniqueness by domain of dependence) thus

$$g_q = (\tilde{x}^{-1} \circ x)^*(g_r).$$

We have thus constructed (M, g), finishing the proof.

□

The above solution is not unique, even in a geometric sense, as we can take a neighborhood M' of $\sum$ in M and consider $(M', g|_{M'})$ and $(M', g|_{M'})$ and (M, g) and $(M', g|_{M'})$ are not isometric in general. To get uniqueness, first we need the following definition.

Definition 16.10. A closed achronal set $\sum \subset M$, M endowed with a metric g, is called a **Cauchy surface** if every inextendible causal curve in M intersects $\sum$, and only once if the curve is timelike. A development (M, g) of $\mathcal{I} = (\sum, g_0, h)$ is called **globally hyperbolic** if $\sum \subset M$ is a Cauchy surface.

Theorem 16.11. *Every initial data set $\mathcal{I} = (\sum, g_0, h)$ admits one and only one maximal globally hyperbolic development.*

Proof. We already know that $\mathcal{I}$ admits one globally hyperbolic development (we have not showed, but we can prove that by taking M to be a sufficiently small neighborhood of $\sum$, $\sum$ will be a Cauchy surface). Let G be the set of all globally hyperbolic developments of $\mathcal{I}$ modulo isometries. We say that $(M, g) \leq (N, h)$ if (M, g) embeds isometrically into (N, h) keeping $\sum$ fixed. This is a partial order in G, so by Zorn's lemma, there exists a maximal element. Uniqueness is obtained because if there exists a (M, g) that does not embeds isometrically into $(M^{\max}, g^{\max})$, we can glue (M, g) and $(M^{\max}, g^{\max})$ to construct a larger solution.

□

Let us finish with some remarks.

Observe that the need for a correct choice of coordinates does not come specifically from the fact that we are dealing with abstract manifolds. We still need to construct wave coordinates in our proof even if $\sum = \mathbb{R}^3$, where we have some a priori canonical coordinates. The wave coordinates depend on the metric, we see that in a sense we constructed the coordinates alongside the solution. In fact, a related approach would be to couple Einstein's equations with the equations determining wave coordinates, i.e., consider the coupled system

$$\text{Ricci}(g) = 0,$$
$$\Box_g x^\alpha = 0,$$

with suitable initial conditions. (Note that this would be the case even in $\mathbb{R}\times\mathbb{R}^3$). Although we have not done so, such a situation, where we need solution-dependent coordinates that are determined alongside the equations of motion themselves, are very common in hyperbolic PDEs.

Let us make a brief comment on the case with matter. Since

$$\nabla^\alpha(R_{\alpha\beta}-\frac{1}{2}Rg_{\alpha\beta})=0,$$

a necessary condition to solve

$$R_{\alpha\beta}-\frac{1}{2}Rg_{\alpha\beta}=\mathcal{T}_{\alpha\beta}$$

is that

$$\nabla_\alpha\mathcal{T}^\alpha_\beta=0.$$

This provides supplementary equations for the matter fields that couple to Einstein's equations. For example, the energy-momentum of a scalar field is

$$\mathcal{T}_{\alpha\beta}=\nabla_\alpha\varphi\nabla_\beta\varphi-\frac{1}{2}\nabla_\mu\varphi\nabla^\mu\varphi g_{\alpha\beta},$$

so that

$$\begin{aligned}\nabla_\alpha\mathcal{T}^\alpha_\beta&=\nabla^\alpha\nabla_\alpha\varphi\nabla_\beta\varphi+\nabla_\alpha\varphi\nabla^\alpha\nabla_\beta\varphi-\nabla_\beta\nabla_\mu\varphi\nabla^\mu\varphi\\&=\nabla^\alpha\nabla_\alpha\varphi\nabla_\beta\varphi\end{aligned}$$

since ∇ is torsion-free. Thus, if $d\phi\neq 0$

$$\nabla^\alpha\nabla_\alpha\varphi=0.$$

Hence, Einstein's equations coupled to a scalar field read

$$\begin{aligned}R_{\alpha\beta}-\frac{1}{2}Rg_{\alpha\beta}=\mathcal{T}_{\alpha\beta}&=\nabla_\alpha\varphi\nabla_\beta\varphi-\frac{1}{2}\nabla_\mu\varphi\nabla^\mu\varphi g_{\alpha\beta},\\\nabla^\mu\nabla_\mu\varphi&=0.\end{aligned}$$

Since $T^\alpha_\alpha=-\nabla_\alpha\varphi\nabla^\alpha\varphi$, we can also write

$$\begin{aligned}R_{\alpha\beta}=\mathcal{T}_{\alpha\beta}-\frac{1}{2}\mathrm{tr}(\mathcal{T})g_{\alpha\beta}&=\nabla_\alpha\varphi\nabla_\beta\varphi,\\\nabla_\alpha\nabla^\alpha\varphi&=0.\end{aligned}$$

All the previous, including the propagation of the gauge, apply if we can solve the reduced coupled system

$$\begin{aligned}R^{\mathrm{red}}_{\alpha\beta}&=\nabla_\alpha\varphi\nabla_\beta\varphi,\\\nabla_\alpha\nabla^\alpha\varphi&=0,\end{aligned}$$

for g and φ. The same remains true if we consider a different matter model, i.e., another energy-momentum tensor $\mathcal{T}$: the argument boils down to solving the reduced coupled system.

17. Elements of the characteristic geometry of quasilinear wave equations

Here we will present some ideas that form the basis of modern techniques for the study of quasilinear wave equations. These techniques are tied to the geometry of the characteristics $g^{\mu\nu}\xi_\mu\xi_\nu=0$, and thus are loosely referred to as the study of the characteristic geometry or, since $g^{\mu\nu}\xi_\mu\xi_\nu=0$ says that ξ null (in the sense of Lorentzian geometry), also as the study of the null geometry. Unless stated otherwise, we will work only in $3+1$ dimensions.

We will consider wave equations of the form

$$g^{\mu\nu}(\varphi)\partial_\mu\partial_\nu\varphi+f(\varphi,\partial\varphi)=0,$$

where it will be convenient to call the solution φ instead of u. Note that we consider metrics that depend on φ but not on its derivatives. We will also often omit the dependence on φ and write $g^{\mu\nu} = g^{\mu\nu}(\varphi)$.

Although in specific problems the form of $f(\varphi, d\varphi)$ is in general relevant, for most part we will be considering properties tied to the principal part, in which case we will consider

$$g^{\mu\nu}(\varphi)\partial_\mu\partial_\nu\varphi = 0.$$

It is also convenient to consider geometric wave equations

$$\Box_{g(\varphi)}\varphi = 0$$

where $\Box_g$ is the wave operator associated with g (which depends on φ)

$$\Box_g = \frac{1}{\sqrt{\det g}}\partial_\alpha(\sqrt{\det g}\, g^{\alpha\beta}\partial_\beta)$$

or equivalently

$$\Box_g = \nabla_\alpha\nabla^\alpha$$

where ∇ is the covariant derivative of g (which depends on φ).

17.1. **The role of decay.** Having settled the question of local existence for quasilinear wave equations, the following question is natural: when is a local-in-time solution a global-in-time solution, i.e., a solution that exists for all time? We saw that a solution can be continued past T if

$$\lim_{t\to T^-}\sup_{0\leq t}\mathcal{N}[\varphi] < \infty,$$

where $\mathcal{N}[\varphi] \sim \partial\varphi$ (recall that here g does not depend on $\partial\varphi$). But an inspection on the proof that something weaker is needed, i.e., it suffices to have

$$\lim_{t\to T^-}\sup_{0\leq t}\int_0^t z(\mathcal{N}[\varphi])d\tau$$

i.e., we want to show that $z(\mathcal{N}[\varphi])d\tau$ is integrable in time on $[0, T]$. In particular, if it is always integrable in time (i.e., integrable on $[0, T]$ for any $T > 0$) then the solution will be global.

For many cases of interest, the structure of the equations is such that the worst terms in $z(\mathcal{N}[\varphi])$ are the ones that contribute linearly. E.g., if $\varphi \sim \frac{1}{(1+t)^2}$, which is integrable, and z is quadratic, then φ^2 is actually better when it comes to time integrability. Also, although $N[\varphi] \sim \varphi + \partial\varphi$, so we actually need integrability of both φ and $\partial\varphi$, experience shows that it is typically $\partial\varphi$ that it is the vilain that is potentially non-integrable (this can be motivated partially from the theory of shocks, where it is shown tat for large classes of equations that form singularities, φ remains bounded whereas $\partial\varphi$ blows up). Finally, working out our energy estimates in more detail for the type of wave equations we are considering in this section shows that indeed solutions can be continued as long as

$$\lim_{t\to T^-}\int_0^t \|\partial\varphi\|_{L^\infty(R^3)}d\tau < \infty$$

which we will take as the basic continuation criterion that motivates the discussion.

This means that solutions can be continued if we can establish uniform pointwise decay (meaning: decay in time) estimates of the form

$$|\partial\varphi(t, x)| \leq \frac{C}{(1+t)^{1+\epsilon}}$$

Is this possible? To investigate this question, let us look at the most basic problem, i.e., the standard linear wave equation in $\mathbb{R}\times\mathbb{R}^3$, $-\partial_t^2\varphi + \Delta\varphi = 0$. (Of course, such solutions are gobal, but let us look at their decay properties as a guide for the nonlinear problem). Kirchhoff's formula gives

$$\begin{aligned}\varphi(t,x) =&\frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)}\big(\varphi_0(y)+t\varphi_1(y)\big)d\sigma(y)\\&+\frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)}\nabla\varphi_0(y)\cdot(y-x)\,d\sigma(y)\end{aligned}$$

where $\varphi(0,\cdot)=\varphi_0$, $\partial_t\varphi(0,\cdot)=\varphi_1$. Let us assume that φ_0 and φ_1 are compactly supported (It is not difficult to see without restrictions on the support of the initial data decay might not hold. E.g., if $\varphi_0=0$ and $\varphi_1=1$, then $\varphi(t,x)=t$ is the solution.) Then, since $\varphi_0=\varphi_1=0$ outside $B_R(0)$ for some $R>0$,

$$\begin{aligned}\varphi(t,x) =&\frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)\cap B_R(0)}\big(\varphi_0(y)+t\varphi_1(y)\big)d\sigma(y)\\&+\frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)\cap B_R(0)}\nabla\varphi_0(y)\cdot(y-x)\,d\sigma(y),\end{aligned}$$

thus

$$\begin{aligned}|\varphi(t,x)|&\leq\frac{C}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)\cap B_R(0)}(1+t+|y-x|)d\sigma(y)\\&\leq\frac{C(1+t)}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)\cap B_R(0)}d\sigma(y),\end{aligned}$$

where we used that since we are integrating on (a portion of) the ball centered at x and of radius t, $|y-x|=t$. The area of $\partial B_t(x)\cap B_R(0)$ is at most $4\pi R^2$ and $\mathrm{vol}(\partial B_t(x))\sim t^2$, so

$$|\varphi(t,x)|\leq\frac{C}{t}.$$

Since $\varphi(t,x)$ has compact support in x for any fixed t, $|\varphi(t,x)|\leq C$ for $0\leq t\leq T$ for some fixed T, so we can also write

$$|\varphi(t,x)|\leq\frac{C}{1+t}.$$

Since $\partial\varphi$ is also a solution to the wave equation, we also get

$$|\partial\varphi(t,x)|\leq\frac{C}{1+t}.$$

Therefore, the estimate we obtained is not good enough to have $\partial\varphi$ integrable, but it is almost good: we got a borderline estimate, and if we can improve the estimate by an $\epsilon>0$, having $(1+t)^{1+\epsilon}$ instead of $1+t$, then we would have integrability.

To investigate this further, first notice the following. If we denote by $\mathcal{S}(\varphi_1)$ ($\mathcal{S}$ for solution) the solution to the Cauchy problem with $\varphi_0=0$, then

$$\varphi=\mathcal{S}(\varphi_1)+\partial\mathcal{S}(\varphi_0)$$

is a solution with data (φ_0,φ_1) since

$$\begin{aligned}-\partial_t^2\varphi+\Delta\varphi&=(-\partial_t^2+\Delta)\mathcal{S}(\varphi_1)+\partial_t(-\partial_t^2+\Delta)\mathcal{S}(\varphi_0)\\&=0+0,\end{aligned}$$

$\partial_t\varphi=\partial_t\mathcal{S}(\varphi_1)+\partial_t^2\mathcal{S}(\varphi_0)=\partial_t\mathcal{S}(\varphi_1)+\Delta\mathcal{S}(\varphi_0)$, but $\mathcal{S}(\varphi_0)$ is the solution that satisfies $\mathcal{S}(\varphi_0)\big|_{t=0}=0$, $\partial_t\mathcal{S}(\varphi_0)\big|_{t=0}=\varphi_0$ and $\mathcal{S}(\varphi_1)$ is the solution that satisfies $\mathcal{S}(\varphi_1)\big|_{t=0}$ and $\partial_t\mathcal{S}(\varphi_1)\big|_{t=0}=\varphi_1$, so

$$\varphi\big|_{t=0}=\varphi_0,\ \ \partial_t\varphi\big|_{t=0}=\varphi_1.$$

Therefore, we can reduce the problem to the case $\varphi_0 = 0$, so

$$\varphi(t,x) = \frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} \varphi_1(y)d\sigma(y).$$

Continuing to assume compactly supported data, since we always have $|\varphi(t,x)| \leq C$ for $0 \leq t \leq T$, T fixed, it suffices to obtain decay for $t \geq T$, so we consider $t \geq 2R$, where supp $(\varphi_1) \subset B_R(0)$. In this situation, we have the alternative formula

$$\varphi(t,x) = \frac{1}{r} f\big(r-t, w, \frac{1}{r}\big)$$

where $f : \mathbb{R} \times S^2 \times [0, \frac{1}{R}] \to \mathbb{R}$ is smooth and vanishes for $|\rho| \geq R$, $f = f(\rho, w, z)$. Here, $r = |x|$ and $x = rw$.

To see this, let us make a change of variables. We can take $x = re_1$

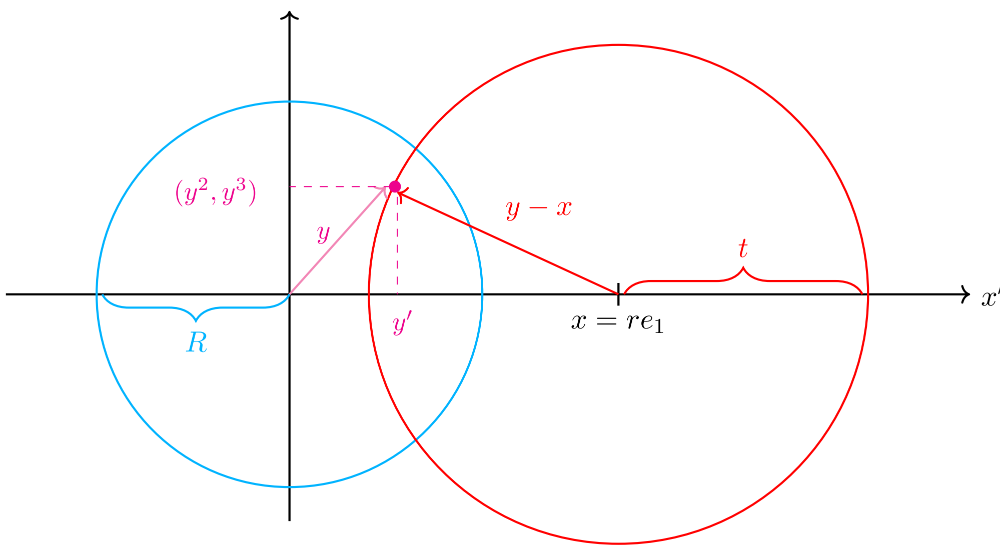


FIGURE 42. Change of Variables: $x = re_1$

$$|y-x| = t, x = re_1, \quad y - x = (y'-r, y^2, y^3)$$
$$t^2 = |y-x|^2 = (r-y)^2 + (y^2)^2 + (y^3)^2 \Rightarrow y' = r - \sqrt{t^2 - ((y^2)^2 + (y^3)^2)}$$

Note that $\sqrt{(y^2)^2 + (y^3)^2} \leq R$, since $y - x \in B_R(0)$ and if $|(y^2, y^3)| > R$ that will not be the case. Therefore, $B_t(x)$ can be parameterized as

$$(y^1, y^2, y^3) = \big(r - \sqrt{t^2 - ((y^2)^2 + (y^3)^2}, \ y^2, \ y^3\big) = \psi(y^2, y^3)$$

Thus, from multivariable calculus

$$d\sigma(y) = \left|\partial_{y^2}\psi \times \partial_{y^3}\psi\right| \, dy^2 dy^3$$
$$= \frac{t}{\sqrt{t^2 - ((y^2)^2 + (y^3)^2)}} \, dy^2 dy^3$$

Then,

$$\varphi(t,x) = \frac{t}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} \varphi_1(y)\, d\sigma(y)$$

$$= \frac{t}{4\pi t^2} \int_{B_R(0)} \varphi_1 \left(r - \sqrt{t^2 - ((y^2)^2 + (y^3)^2)},\ y^2,\ y^3 \right) \frac{t}{\sqrt{t^2 - ((y^2)^2 + (y^3)^2)}}\, dy^2 dy^3$$

$$= \frac{1}{4\pi} \int_{B_R(0)} \varphi_1 \left(r - \sqrt{t^2 - ((y^2)^2 + (y^3)^2)},\ y^2,\ y^3 \right) \frac{1}{\sqrt{t^2 - ((y^2)^2 + (y^3)^2)}}\, dy^2 dy^3$$

Set $\rho = r - t$. Then

$$\begin{aligned}\sqrt{t^2 - ((y^2)^2 + (y^3)^2)} &= \left((\rho - r)^2 - (y^2)^2 + (y^3)^2\right)^{1/2} \\ &= r \left(\left(1 - \frac{\rho}{r}\right)^2 - \frac{(y^2)^2 + (y^3)^2}{r^2} \right)^{1/2}\end{aligned}$$

$$r - \sqrt{t^2 - ((y^2)^2 + (y^3)^2)} = r - r \left(\left(1 - \frac{\rho}{r}\right)^2 - \frac{(y^2)^2 + (y^3)^2}{r^2} \right)^{1/2}$$

$$= r \left[1 - \left(\left(1 - \frac{\rho}{r}\right)^2 - \frac{(y^2)^2 + (y^3)^2}{r^2} \right)^{1/2} \right]$$

$$= r \left[1 - \left(\left(1 - \frac{\rho}{r}\right)^2 - \frac{(y^2)^2 + (y^3)^2}{r^2} \right)^{1/2} \right] \frac{1 + \left(\left(1 - \frac{\rho}{r}\right)^2 - \frac{(y^2)^2 + (y^3)^2}{r^2} \right)^{1/2}}{1 + \left(\left(1 - \frac{\rho}{r}\right)^2 - \frac{(y^2)^2 + (y^3)^2}{r^2} \right)^{1/2}}$$

$$= \frac{r \left[1 - \left(1 - \frac{\rho}{r}\right)^2 + \frac{(y^2)^2 + (y^3)^2}{r^2} \right]}{1 + \left(\left(1 - \frac{\rho}{r}\right)^2 - \frac{(y^2)^2 + (y^3)^2}{r^2} \right)^{1/2}}$$

$$= \frac{2\rho + \frac{(y^2)^2 + (y^3)^2 - \rho^2}{r^2}}{1 + \left(\left(1 - \frac{\rho}{r}\right)^2 - \frac{(y^2)^2 + (y^3)^2}{r^2} \right)^{1/2}}$$

Inspecting the above expressions we see that both are smooth functions of $(\rho,\ \frac{1}{r},\ y^2,\ y^3)$, but y^2 and y^3 are integrated away.

By the strong Huygens' principle, the solution vanishes for $|\rho| = |t - r| \geq R$, and for $t \geq 2R$ it vanishes for $|x| \leq R$, so $r \geq R$ or $\frac{1}{r} \leq \frac{1}{R}$, as claimed. This construction is smooth on its dependence on the base point $x = re_1$, so for a general x we also obtain smooth dependence on w.

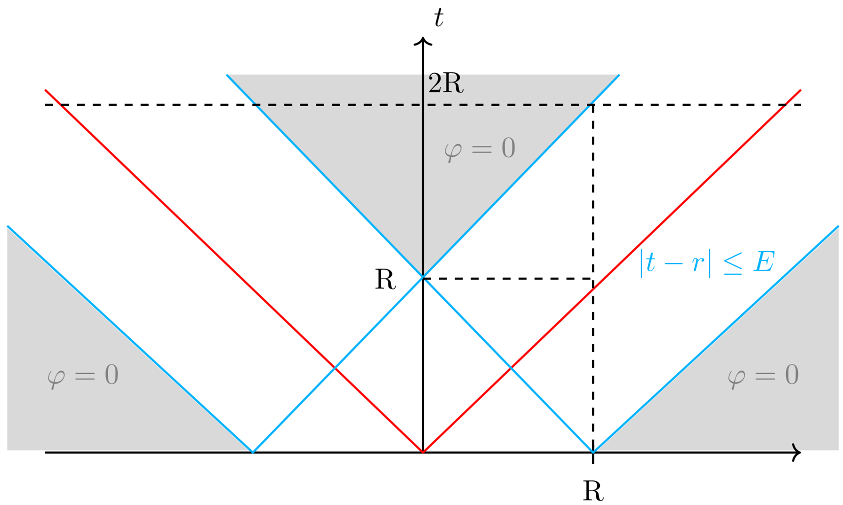


FIGURE 43. Huygens' Principle

Consider the vector fields

$$L = \partial_t + \partial_r, \;\; \underline{L} = \partial_t - \partial_r,$$
$$e_1 = \frac{1}{r}\partial_\theta, \;\; e_2 = \frac{1}{r\sin\theta}\partial_\phi,$$

where $\partial_r = w^i \partial_i = \frac{x^i}{r}\partial_i$ is the radial derivative and $\partial_\theta, \partial_\phi$ are the derivatives in spherical coordinates (with the convention $x^1 = r\sin\theta\cos\phi,\; x^2 = r\sin\theta\sin\phi,\; x^3 = r\cos\theta$). If we introduce the rotation vector fields

$$R_1 = x^2\partial_3 - x^3\partial_2, \; R_2 = x^3\partial_1 - x^1\partial_3, \; R_3 = x_1\partial_2 - x^2\partial_1,$$

then

$$e_1 = -\sin\phi\frac{R_1}{r} + \cos\phi\frac{R_2}{r},$$
$$e_2 = \frac{1}{\sin\theta}\frac{R_3}{r}.$$

Compute

$$L\varphi = (\partial_t + \partial_r)\,\frac{1}{r}\,f(r-t, \omega, \tfrac{1}{r})$$
$$= -\frac{1}{r}f_\rho - \frac{1}{r^2}f + \frac{1}{r}f_\rho - \frac{1}{r^3}f_z = \mathcal{O}\left(\frac{1}{r^2}\right)$$

$$
\begin{aligned}
R_1\varphi =& (x^2\partial_3 - x^3\partial_2)\,\frac{1}{r}f(r-t,\omega,\tfrac{1}{r})\\
=& -\frac{1}{r^2}x^2\partial_3 r + \frac{x^3}{r^2}\partial_2 r + \frac{1}{r}x^2\left(f_\rho\partial_3 r + f_{\omega^i}\partial_3\omega^i - f_z\frac{1}{r^2}\partial_3 r\right)\\
& -\frac{1}{r}x^3\left(f_\rho\partial_2 r + f_{\omega^i}\partial_2\omega^i - f_z\frac{1}{r^2}\partial_2 r\right)\\
=& -\frac{1}{2}x^2\frac{x^3}{r^3} + \frac{x^3}{r^3}x^2 + \frac{1}{r}f_\rho\left(x^2\frac{x^3}{r} - x^3\frac{x^2}{r}\right) - f_z\left(\frac{x^2x^3}{r^4} - \frac{x^2x^3}{r^4}\right)\\
& + \frac{1}{r}f_{\omega^i}\quad \underbrace{\left(x^2\partial_3\omega^i - x^3\partial_2\omega^i\right)}_{=\frac{x^2}{r}(\delta_3^i-\omega^3\omega^i)-\frac{x^3}{r}(\delta_2^i-\omega^2\omega^i)}\\
=& \,\mathcal{O}\left(\frac{1}{r}\right),
\end{aligned}
$$

so $\frac{R_1}{r}\varphi = \mathcal{O}(\frac{1}{r^2})$ and similarly $\frac{R_2}{r}\varphi, \frac{R_3}{r}\varphi = \mathcal{O}(\frac{1}{r^2})$ thus $e_1\varphi, e_2\varphi = \mathcal{O}(\frac{1}{r^2})$.

Finally,

$$
\begin{aligned}
\underline{L}\varphi &= (\partial_t - \partial_r)\frac{1}{r}f(r-t,\omega,\frac{1}{r})\\
&= -\frac{1}{r}f_\rho + \frac{1}{r^2}f - \frac{1}{r}f\rho - \frac{1}{r^3}f_z = \mathcal{O}(\frac{1}{r})
\end{aligned}
$$

Since the solution is not zero only for $|t-r|\le R$, we have $\mathcal{O}(\frac{1}{r}) = \mathcal{O}(\frac{1}{t})$. Thus we have the decay estimates

$$
L\varphi,\ \ e_1\varphi,\ \ e_2\varphi = \mathcal{O}(\frac{1}{t^2}),\ \ \underline{L}\varphi = \mathcal{O}(\frac{1}{t})
$$

$\{L, \underline{L}, e_1, e_2\}$ form a basis of $\mathbb{R}^4$, thus we can reexpress $\partial\varphi$ in terms of $L\varphi,\ \underline{L}\varphi,\ e_1\varphi,\ e_2\varphi$. Except for $\underline{L}\varphi$, all the other derivatives decay better than our previous $\frac{1}{t}$ estimate. Can we get $\underline{L}\varphi = \mathcal{O}(\frac{1}{t^2})$?

The answer is no, because we can show that the $\frac{1}{t}$ decay for solutions to wave equation is sharp (i.e., it cannot be improved for arbitrary compactly supported data). To see this, consider again Kirchhoff's formula with $\varphi_0 = 0$ and $\varphi_1 = 1$ on $B_1(0)$, φ_1 compactly supported and $\varphi_1 \ge 0$. Then

$$
\varphi(t,x) = \frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)} t\varphi_1(y)\,d\sigma(y)
$$

$$
\begin{aligned}
&= \frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)\cap B_1(0)} t\varphi_1(y)\,d\sigma(y) + \underbrace{\frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)\setminus(B_t(x)\cap B_1(0))} t\varphi_1(y)\,d\sigma(y)}_{\text{non-negative}}\\
&\ge \frac{1}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)\cap B_1(0)} t\varphi_1(y)\,d\sigma(y) = \frac{t}{\mathrm{vol}(\partial B_t(x))}\int_{\partial B_t(x)\cap B_1(0)} d\sigma(y)
\end{aligned}
$$

For any $|x| = t$ and $t \le 1$, the area of $\partial B_t(x)\cap B_1(0)$ is $\ge C > 0$.

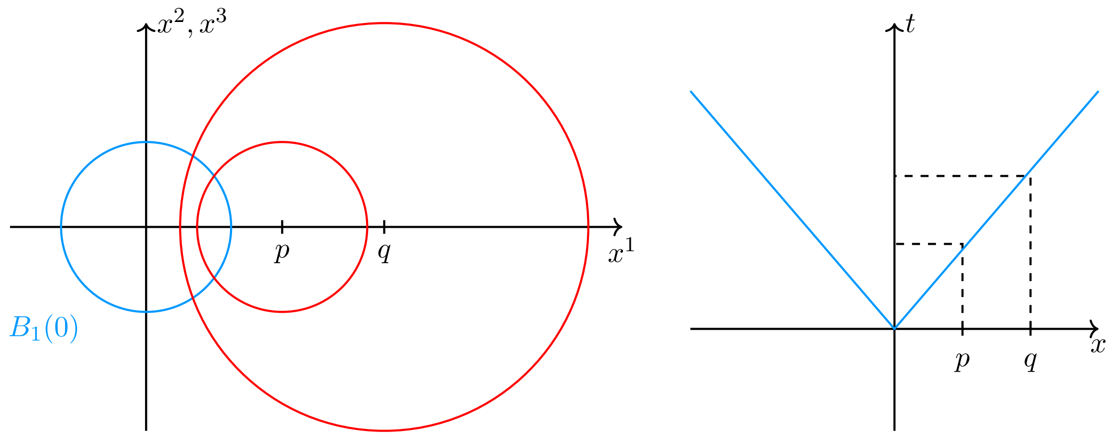


FIGURE 44. $\partial B_t(x) \cap B_1(0)$ and $|x| = t$

Thus, $\varphi(t,x) \geq \frac{C}{t}$.

However, we learned that the obstacle to getting better decay is the derivative in the direction $\underline{\text{L}}$; derivatives in the other direction decay better.

What this means for the quasi-linear problem is that we should not expect to get better decay than $\frac{1}{t}$, as we should not expect the non-linear problems to behave better than the standard linear wave equation. So, in general, we should not expect $\|\partial\varphi\|_{L^\infty}$ to be integrable in time. Does that mean then that we cannot get $\int_0^t \|\partial\varphi\|_{L^\infty} d\tau < \infty$ and hence global solutions? The key here is the word "general". We know that there are examples of solutions that blow up, so we should not expect to get $\int_0^t \|\partial\varphi\|_{L^\infty} d\tau < \infty$. However, we know at least one global solution: the zero solution. The correct question then is whether we can get global solutions for small data solutions. The answer is yes for important class equations. These include quasi-linear wave equations that satisfy a condition called the null-condition. For small data, Einstein's equations also admit global solutions, where small here refers not to initial data that are perturbations of the zero a solution (since for Einstein's equations solutions need to be a metric) but rather to perturbations of the Minkowski space by Christodoulou and Klainerman (the stability of Minkowski space in fact refers to more than the existence of small data global solutions; it also says that such solutions are stable in a precise sense).

Inspired by the linear case, the first thing to do to address global existence is to understand the obstruction to proving better decay that $\frac{1}{t}$. Thus, we try to identify, as in the linear case, the directions along which we have a better decay. Such derivatives, such as the L, e_1, e_2 derivatives in the linear case, are commonly referred to as **"good derivatives,"** whereas the ones where better decay does not hold, such as the $\underline{L}$ derivative in the linear case, are referred to as **"bad derivatives."** Roughly, we expect the good derivatives not to cause too much trouble, and the bad derivatives are the ones that need to be carefully estimated, and where the small data assumption enters crucially.

In order to identify the good vs. bad derivatives, we once again take a clue for the linear problem. Since we are interested in perturbations of the zero solution, it is natural to consider $g^{\mu\nu}(\varphi)\partial_\mu\partial_\nu\varphi + ... = 0$ such that $g^{\mu\nu}(\varphi = 0) = m\mu\nu$, where m is the Minkowski. We expect that atleast for small time, $m\mu\nu\partial_\mu\partial_\nu\varphi$ will be a good approximation for $g^{\mu\nu}\partial_\mu\partial_\nu\varphi$ since the data for φ is small. In other words, if we consider, say,

$$g^{\mu\nu}(\varphi)\partial_\mu\partial_\nu\varphi = 0$$

we can look at its linearization at $\varphi = 0$, which is precisely

$$m\mu\nu\partial_\mu\partial_\nu\varphi = 0.$$

This suggests that if good derivatives exist for the non-linear problem, they should be approximated by $L = \partial_t + \partial_r$, $e_1 = \frac{1}{r}\partial_\theta$, and $e_2 = \frac{1}{r\sin\theta}\partial_\phi$, at least for small time. In such a case, the bas derivative in the non-linear problem should be something close to $\underline{L} = \partial_t - \partial_u$. However, the problem with this reasoning is that we want to prove the global existence, and we do not know that $\varphi = 0$ is a good approximation for large time (in face, we can only make such a statement if we know the solutions to be global). Thus, we try to abstract from the linear problem the geometric features of L, $\underline{L}$, e_1, e_2, that can make sense in the non-linear problem even if we do not know $\varphi \approx 0$. For this, observe the following: L, e_1, e_2 are tangent to the light-cones, whereas $\underline{L}$ is transverse. Moreover, L and $\underline{L}$ are null with respect to the Minkowski metric, whereas e_1, e_2 are spacelike.

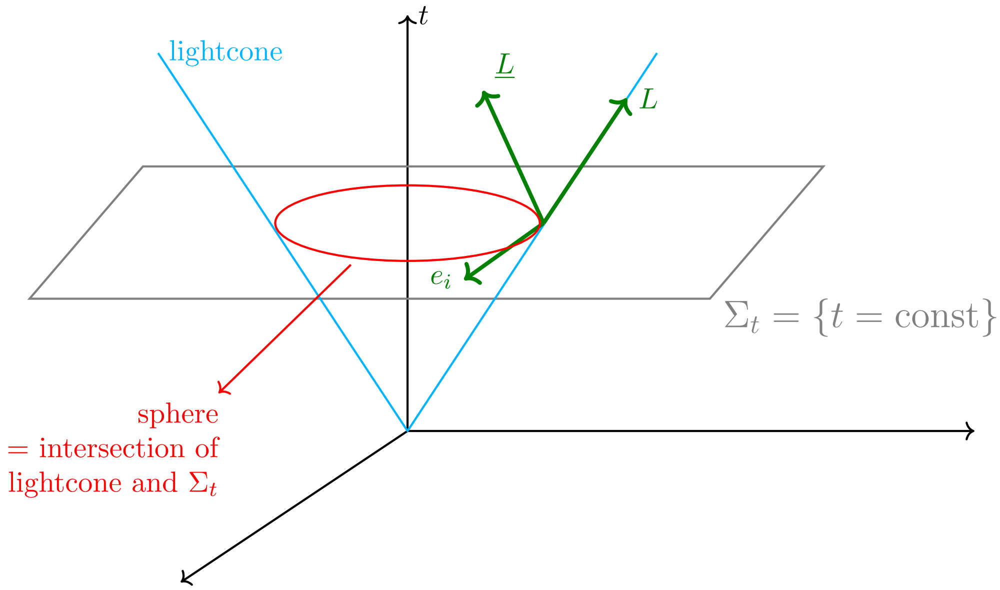


FIGURE 45. L, L, and the Minkowski metric I

$$m_{\mu\nu}L^\mu L^\nu = -1 + \delta_{ij}\frac{x^i}{r}\frac{x^j}{r} = -1 + \frac{x^i x_i}{r^2} = 0$$

$$m_{\mu\nu}\underline{L}^\mu\underline{L}^\nu = -1 + \delta_{ij}\left(\frac{-x^i}{r}\right)\left(\frac{-x^j}{r}\right) = -1 + \frac{x^i x_i}{r^2} = 0$$

To compute further, consider the Minkowski metric in spherical coordinates:

$$m = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2\sin^2\theta \end{pmatrix}$$

So

$$m_{\mu\nu}e_1^\mu e_2^\nu = \left\langle \frac{1}{r}\partial_\theta,\ \frac{1}{r}\partial_\theta \right\rangle = 1,$$

$$m_{\mu\nu}e_2^\mu e_2^\nu = \left\langle \frac{1}{r\sin\theta}\partial_\phi,\ \frac{1}{r\sin\theta}\partial_\theta \right\rangle = 1.$$

Furthermore,

$$m_{\mu\nu} L^\mu \underline{L}^\nu = -1 + \delta_{ij} \frac{x^i}{r} \left(\frac{-x^j}{r} \right) = -2,$$

$$m_{\mu\nu} L^\mu e_i^\nu = \langle -\partial_t + \partial_\sigma,\ e_i \rangle = 0, \quad \text{(this is easy to see using } m \text{ in spherical coordinates)}$$

$$m_{\mu\nu} e_1^\mu e_2^\nu = \left\langle \frac{1}{r}\partial_\phi,\ \frac{1}{r \sin_\theta}\partial_\sigma \right\rangle = 0,$$

So, the vectors are also orthogonal.

Since the lightcones are characteristics for the operator $m^{\mu\nu}\partial_\mu\partial_\nu$, this suggests the following approach to identify good and bad derivatives for the nonlinear problem: consider the characteristic hypersurfaces of $g^{\mu\nu}\partial_\mu\partial_\nu$, which will be null hypersurfaces for the metric g, and construct a frame of vectors $\{L, \underline{L}, e_1, e_2\}$ where L and $\underline{L}$ are null vectors (with respect to g), L is tangent to the characteristics and $\underline{L}$ transverse, e_1 and e_2 are spacelike unit vectors (with respect to g) tangent to $\{t = 0\} \cap \{\text{characteristics}\}$, all the vectors are orthogonal (with respect to g) with exception of $L, \underline{L}$. In particular, they form a basis of $\mathbb{R}^4$.

We then expect that, under reasonable assumptions, we can show that L, e_1, e_2 are good derivatives, i.e., $L\varphi,\ e_1\varphi,\ e_2\varphi$ have better decay than $\frac{1}{t}$.

The crucial observation now is the following: since $g = g(\varphi)$, the characteristics, and thus the vectors $L,\ \underline{L},\ e_1,\ e_2$ will depend on the solution φ. Thus, before trying to prove that we have good derivatives and eventual global existence, we need to understand in detail the properties of these vector fields. This leads to a study of the characteristic geometry of the g.

17.2. **Null Frames.** Although we are primarily interested in the case when $g = g(\varphi)$ comes from a solution φ to a quasilinear problem, what follows applies to a general metric g. Throughout we will consider the spacetime $(\mathbb{R}^4), g$, although our constructions are local and apply equally to a general Lorentzian manifold. We write ∇ and $|\cdot|$ for the covariant derivative and norm of g, writing ∇_g and $|\cdot|$ if we want to emphasize this dependence. Indices will be raised and lowered with g.

Definition 17.1. A **null-frame** is a basis $\{e_\mu\}_{\mu=0}^3$ to the tangent space at x (i.e., a basis of $\mathbb{R}^4$ in our case) vary smoothly with the base point x, such that

$$\begin{aligned} &g(e_1, e_1) = g(e_2, e_2) = 1, \quad g(e_1, e_2) = 0, \\ &g(e_i, e_j) = 0, \quad i = 1, 2, \quad j = 3, 4, \\ &g(e_3, e_3) = g(e_4, e_4) = 0, \quad g(e_3, e_4) = -2\mu,\ \mu > 0. \end{aligned}$$

We often write $e_3 = \underline{L}, e_4 = L$.

Example 17.2. The base example is $\{e_1,\ e_2,\ \underline{L},\ L\}$ constructed above for the Minkowski metric. In that case $\mu = 1$. The choice of μ should be viewed as a normalization since $L,\ \underline{L}$ are null, so we cannot "fix their length" by prescribing $g(L, L),\ g(\underline{L}, \underline{L})$

For another example, consider the Schwarzschild metric

$$g = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2,$$

$M \geq 0$. Then we can take e_1, e_2 as in Minkowski and

$$\underline{L} = \partial_t - \left(1 - \frac{2M}{r}\right)\partial_r, \quad L = \partial_t + \left(1 - \frac{2M}{r}\right)\partial_r.$$

For the Kerr metric,

$$g = -\frac{\Delta - a^2 \sin^2\theta}{\Sigma} dt^2 + \frac{\Sigma}{\Delta} dr^2 - 4aMr\frac{\sin^2\theta}{\Sigma} dt\, d\phi + A\sin^2\theta\, d\phi^2 + \Sigma d\theta^2,$$

where $\Sigma = r^2 + a^2\cos^2\theta,\ \Delta = r^2 + a^2 - 2Mr,\ A = (r^2+a^2)^2 - a^2\Delta\sin^2\theta$. (The Kerr metric reduces to the Schwarzschild one if $a = 0$), we can take

$$\begin{aligned} e_1 &= \frac{1}{\sqrt{\Sigma}}\partial_\theta, \quad e_2 = \frac{1}{\sqrt{\Sigma}\sin\theta}(\partial_\phi + a\sin^2\theta\,\partial_t), \\ \underline{L} &= \partial_t + \frac{a}{r^2+a^2}\partial_\phi - \frac{\Delta}{r^2+a^2}\partial_r \\ L &= \partial_t + \frac{a}{r^2+a^2}\partial_\phi + \frac{\Delta}{r^2+a^2}\partial_r. \end{aligned}$$

(For the Kerr metric, e_1, e_2 are not tangent to spheres. Although this property is desirable (see our motivation above), it is not strictly needed and is not part of the definition of a null frame.)

The dual basis of a null frame is

$$\begin{aligned} e^1 &:= (e_1)^* = e_1, \quad e^2 := (e_2)^* = e_2, \\ e^3 &:= (e_3)^* = -\frac{1}{2\mu}e_4, \quad e^4 := (e_4)^* = -\frac{1}{2\mu}e_3, \end{aligned}$$

where duality is defined in the usual fashion

$$e^\alpha(e_\beta) = \delta^\alpha_\beta,$$

and $(e^\alpha)^*(e_\beta) = g(e_\alpha, e_\beta)$. In fact, if $g_{\alpha\beta}$ are the components of g relative to a null frame and $g^{\alpha\beta}$ its inverse, then

$$e^\alpha = g^{\alpha\beta}e_\beta$$

since

$$\begin{aligned} e^\alpha(e_\beta) &= g(g^{\alpha\beta}e_\gamma, e_\beta) = g^{\alpha\beta}g(e_\gamma, e_\beta) \\ &= g^{\alpha\beta}g_{\gamma\beta} = \delta^\alpha_\beta. \end{aligned}$$

Thus, relative to a null frame

$$g = e^1\otimes e^1 + e^2\otimes e^2 - 2\mu e^3\otimes e^4 - 2\mu e^4\otimes e^3,$$

or in matrix form

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2\mu \\ 0 & 0 & -2\mu & 0 \end{pmatrix}.$$

We can also express g in coordinates but relative to a null frame. More precisely, we have,

$$g_{\alpha\beta} = -\frac{1}{2\mu}L_\alpha \underline{L}_\beta - \frac{1}{2\mu}\underline{L}_\alpha L_\beta + \not{g}_{\alpha\beta}$$

where $\not{g}$ is positive definite on the space orthogonal to the space spanned by L and $\underline{L}$ and it vanishes on span$\{L, \underline{L}\}$. To see this, define

$$\not{g}_{\alpha\beta} = g_{\alpha\beta} + \frac{1}{2\mu}L_\alpha\underline{L}_\beta + \frac{1}{2\mu}\underline{L}_\alpha L_\beta.$$

Then

$$\not{g}_{\alpha\beta}L^\alpha L^\beta = \underbrace{g_{\alpha\beta}L^\alpha L^\beta}_{=0} + \frac{1}{2\mu}\underbrace{L_\alpha L^\alpha}_{=0}\underline{L}_\beta L^\beta + \frac{1}{2\mu}\underline{L}_\alpha L^\alpha \underbrace{L_\beta L^\beta}_{=0} = 0$$

$$\not{g}_{\alpha\beta}\underline{L}^\alpha\underline{L}^\beta = \underbrace{g_{\alpha\beta}\underline{L}^\alpha\underline{L}^\beta}_{=0} + \frac{1}{2\mu}L_\alpha\underline{L}^\alpha\underbrace{\underline{L}_\beta\underline{L}^\beta}_{=0} + \frac{1}{2\mu}\underbrace{\underline{L}_\alpha\underline{L}^\alpha}_{=0}L_\beta\underline{L}^\beta = 0$$

$$\not{g}_{\alpha\beta}\underline{L}^\alpha L^\beta = \underbrace{g_{\alpha\beta}\underline{L}^\alpha L^\beta}_{=-2\mu} + \frac{1}{2\mu}\underbrace{L_\alpha\underline{L}^\alpha}_{=-2\mu}\underbrace{\underline{L}_\beta L^\beta}_{=-2\mu} + \frac{1}{2\mu}\underbrace{\underline{L}_\alpha\underline{L}^\alpha}_{=0}\underbrace{L_\beta L^\beta}_{=0} = 0$$

$$\not{g}_{\alpha\beta}e^\alpha_A e^\beta_B = g_{\alpha\beta}e^\alpha_A e^\beta_B + \underbrace{\frac{1}{2\mu}L_\alpha e^\alpha_A\underline{L}_\beta e^\beta_B + \frac{1}{2\mu}\underline{L}_\alpha e^\alpha_A L_\beta e^\beta_B}_{\text{(all products are zero)}},$$

$A, B = 1, 2$, showing the claim.

At this point, it is convenient to introduce:

Notation 17.3. We use uppercase Latin letters, $A, B, \ldots$ to denote indices 1, 2. The sum convention is adopted for such indices, including when repeated indices are both up or down. The inverse of $\not{g}$ is given in coordinates by

$$(\not{g}^{-1})^{\alpha\beta} = \sum_{A=1,2} e^\alpha_A e^\beta_A = \underbrace{e^\alpha_A e^\beta_A}_{\text{sum over A=1,2.}}$$

(see below for the notation $^{-1}$ in $\not{g}^{-1}$).

Induced

$$\begin{aligned}(\not{g}^{-1})^{\alpha\beta}\not{g}_{\beta\delta} &= e^\alpha_A e^\beta_A(g_{\beta\delta} + \frac{1}{2\mu}L_\beta\underline{L}_\delta + \frac{1}{2\mu}\underline{L}_\beta L_\delta) \\ &= e^\alpha_A e^\beta_A g_{\beta\delta}.\end{aligned}$$

To confirm that $\not{g}^{-1}$ is the inverse, we need to show that this last expression is the identity on the space orthogonal to span$\{L, \underline{L}\}$ and vanishes on span$\{L, \underline{L}\}$. The latter follows from

$$(\not{g}^{-1})^{\alpha\beta}\not{g}_{\alpha\delta}L^\delta = e^\alpha_A e^\beta_A g_{\beta\delta}L^\delta = e^\alpha_A\underbrace{e^\beta_A L_\beta}_{=0} = 0,$$

$$(\not{g}^{-1})^{\alpha\beta}\not{g}_{\alpha\delta}\underline{L}^\delta = e^\alpha_A e^\beta_A g_{\beta\delta}\underline{L}^\delta = e^\alpha_A\underbrace{e^\beta_A\underline{L}_\beta}_{=0} = 0.$$

If $X = X^A e_A = X^A e^\alpha_A\partial_\alpha = X^\alpha\partial_\alpha$, where X^A are the components of X relative to $\{e_A\}^2_{A=1}$ and X^α the components relative to $\{\partial_\alpha\}^3_{\alpha=1}$. Then

$$\begin{aligned}(\not{g}^{-1})^{\alpha\beta}\not{g}_{\alpha\delta}X^\delta &= e^\alpha_A e^\beta_A g_{\beta\delta}X^C e^\delta_C, \\ &= e^\alpha_A X^C\underbrace{g_{\beta\delta}e^\beta_A e^\delta_C}_{=\delta_{AC}} = X^A e^\alpha_A = X^\alpha,\end{aligned}$$

showing the claim.

Remark 17.4. In general, the components of the inverse of $\not{g}$ are not obtained by simply raising an index, unless $\not{g}$ is defined in terms of g and we raise and lower indices with g, which is the case here but might not be true in general (e.g., g is the acoustical metric but we raise/lower indices with the Minkowski metric).

This is because indices are raised and lowered with g, and $g^{\alpha\beta}$ is both the inverse matrix of $g_{\alpha\beta}$ and $g_{\alpha\beta}$ raised indices when indices are raised with g. But here, we have the metric $\not{g}$ but are raising and lowering indices with g. That is why we wrote $(\not{g}^{-1})^{\alpha\beta}$, indicating explicitly that it is the $\alpha\beta$ entry of the inverse matrix of $\not{g}$ (whose entries are $\not{g}_{\alpha\beta}$). In particular, $\not{g}^{\alpha\beta}$ means what it means for any other tensor, namely, $\not{g}_{\alpha\beta}$ raised the indices with the spacetime metric:

$$\not{g}^{\alpha\beta} = g^{\alpha\gamma} g^{\beta\delta} \not{g}_{\gamma\delta}.$$

It is useful to see that $\not{g}^{\alpha\beta}$ is in fact the inverse of $\not{g}_{\alpha\beta}$ only because we raise indices with g itself:

$$\begin{aligned}
\not{g}^{\alpha\beta}\not{g}_{\beta\delta} &= g^{\alpha\sigma} g^{\beta\tau} \not{g}_{\sigma\tau}\not{g}_{\beta\delta} \\
&= g^{\alpha\sigma} g^{\beta\tau} \left(g_{\sigma\tau} + \frac{1}{2\mu} L_\sigma \underline{L}_\tau + \frac{1}{2\mu} \underline{L}_\sigma L_\tau \right) \left(g_{\beta\delta} + \frac{1}{2\mu} L_\beta \underline{L}_\delta + \frac{1}{2\mu} \underline{L}_\beta L_\delta \right) \\
&= \left(g^{\alpha\beta} + \frac{1}{2\mu} L^\alpha \underline{L}^\beta + \frac{1}{2\mu} \underline{L}^\alpha L^\beta \right) \left(g_{\beta\delta} + \frac{1}{2\mu} L_\beta \underline{L}_\delta + \frac{1}{2\mu} \underline{L}_\beta L_\delta \right) \\
&= \Big(\delta^\alpha_\delta + \frac{1}{2\mu} L^\alpha \underline{L}_\delta + \frac{1}{2\mu} \underline{L}^\alpha L_\delta + \frac{1}{2\mu} L^\alpha \underline{L}_\delta + \frac{1}{2\mu} \underline{L}^\alpha L_\delta \\
&\quad + \frac{1}{4\mu^2} L^\alpha \underline{L}_\delta \underbrace{L^\beta \underline{L}_\beta}_{=-2\mu} + \frac{1}{4\mu^2} L^\alpha L_\delta \underbrace{\underline{L}^\beta \underline{L}_\beta}_{=0} + \frac{1}{4\mu^2} \underline{L}^\alpha \underline{L}_\delta \underbrace{L^\beta L_\beta}_{=0} + \frac{1}{4\mu^2} \underline{L}^\alpha L_\delta \underbrace{L^\beta \underline{L}_\beta}_{=-2\mu} \Big) = \delta^\alpha_\delta.
\end{aligned}$$

Next, we claim that

$$\delta^\beta_\alpha + \frac{1}{2\mu} L^\beta \underline{L}_\alpha + \frac{1}{2\mu} \underline{L}^\beta L_\alpha$$

projects a vector onto the space spanned by e_1, e_2.

Induced,

$$\begin{aligned}
X &= X^A e_A + X^{\underline{L}} \underline{L} + X^L L \\
&= X^A e^\alpha_A \partial_\alpha + X^{\underline{L}} \underline{L}^\alpha \partial_\alpha + X^L L^\alpha \partial_\alpha.
\end{aligned}$$

If we denote by $X^{(\alpha)}$ the components of X relative to $\{e_\alpha\}_{\alpha=0}^3$, the above can be written in concise form

$$X = X^{(\alpha)} e^\beta_\alpha \partial_\alpha$$

(recall $e_3 = \underline{L}, e_4 = L$). Then,

$$\left(\delta^\beta_\alpha + \frac{1}{2\mu} L^\beta \underline{L}_\alpha + \frac{1}{2\mu} \underline{L}^\beta L_\alpha \right) X^\alpha$$

$$\begin{aligned}
&= \left(\delta^\beta_\alpha + \frac{1}{2\mu} L^\beta \underline{L}_\alpha + \frac{1}{2\mu} \underline{L}^\beta L_\alpha \right) X^{(\gamma)} e^\alpha_\gamma \\
&= X^{(r)} e^\beta_\gamma + \frac{1}{2\mu} L^\beta X^{(\gamma)} \underbrace{\underline{L}_\alpha e^\beta_\gamma}_{\substack{0 \text{ if } \gamma \neq 4 \\ -2\mu \text{ if } \gamma = 4}} + \frac{1}{2\mu} \underline{L}^\beta X^{(\gamma)} \underbrace{L_\alpha e^\alpha_\gamma}_{\substack{0 \text{ if } \gamma \neq 3 \\ -2\mu \text{ if } \gamma = 3}} \\
&= X^{(\gamma)} e^\beta_\gamma - X^{(4)} L^\beta - X^{(3)} \underline{L}^\beta \\
&= X^A e^\beta_A + X^{(3)} e^\beta_3 + X^{(4)} e^\beta_4 - X^{(4)} L^\beta - X^{(3)} \underline{L}^\beta,
\end{aligned}$$

but $e_3 = \underline{L},\ e_4 = L$, so $e^\beta_3 = \underline{L}^\beta,\ e^\beta_4 = L^\beta$, thus

$$= X^A e^\beta_A,$$

showing the claim.

Notation 17.5. We denote by $\not\Pi$ the projection onto span$\{e_1, e_2\}$. By the above

$$\not\Pi_\alpha^\beta = \delta_\alpha^\beta + \frac{1}{2\mu} L^\beta \underline{L}_\alpha + \frac{1}{2\mu} \underline{L}^\beta L_\alpha$$

Observe that the gradient $\nabla\varphi_1$ expressed relative to a null frame is

$$\nabla\varphi = e_1(\varphi)e_1 + e_2(\varphi)e_2 - \frac{1}{2\mu} e_4(\varphi)e_3 - \frac{1}{2\mu} e_3(\varphi)e_4.$$

Indeed,

$$\begin{aligned}
&\nabla\varphi = (\nabla\varphi)^A e_A + (\nabla\varphi)^{\underline{L}} \underline{L} + (\nabla\varphi)^L L, \quad \text{so,} \\
&(\nabla\varphi)^A = g(\nabla\varphi, e_A), \quad (\nabla\varphi)^{\underline{L}} = -\frac{1}{2\mu} g(\nabla\varphi, L), \\
&(\nabla\varphi)^L = -\frac{1}{2\mu} g(\nabla\varphi, \underline{L}).
\end{aligned}$$

But

$$\begin{aligned}
g(\nabla\varphi, e_\alpha) &= g_{\beta\gamma} \nabla^\beta \varphi e_\alpha^\gamma = g_{\beta\gamma} g^{\beta\delta} \nabla_\delta \varphi e_\alpha^\gamma \\
&= \nabla_\gamma \varphi e_\alpha^\gamma = \partial_\gamma \varphi e_A^\gamma = e_\alpha^\gamma \partial_\varphi \varphi \\
&= e_\alpha(\varphi).
\end{aligned}$$

Similarly, given a vector field X,

$$X = X^A e_A + X^{\underline{L}} \underline{L} + X^L L,$$

$$X^A = g(X, e_A), \quad X^{\underline{L}} = -\frac{1}{2\mu} g(X, L), \quad X^L = -\frac{1}{2\mu} g(X, \underline{L}),$$

So

$$X = g(X, e_A)e_A - \frac{1}{2\mu} g(X, L)\underline{L} - \frac{1}{2\mu} g(X, \underline{L})L.$$

Let us denote the contractions of X with $e_A, \underline{L}$, and L, respectively, by

$$X_A := X_\alpha e_A^\alpha, \quad X_{\underline{L}} := X_\alpha \underline{L}^\alpha, \quad X_L := X_\alpha L^\alpha.$$

Then

$$X = X_A e_A - \frac{1}{2\mu} X_L \underline{L} - \frac{1}{2\mu} X_{\underline{L}} L.$$

Notation 17.6. An **eikonal function** or **optical function** is a solution to the **eikonal equation**

$$g^{\mu\nu} \partial_\mu u \, \partial_\nu u = 0.$$

It follows that the level sets of u are characteristic manifolds for the operator $g^{\mu\nu}\partial_\mu\partial_\nu$. Observe that if u is an optical function, then the vector field

$$L = -\nabla u$$

is null (the negative sign is conventional). Since the gradient is orthogonal to the level surfaces (this result is also true for Lorentzian metrics), L is both orthogonal to the characteristics (below we show it is also tangent).

Example 17.7. For the Minkowski metric, $u = t - r$ and $\underline{u} = t + r$ are eikonal functions. Their level sets correspond to lightcones. Note that $-\nabla u = \partial_t + \partial_r = L, \quad -\nabla \underline{u} = \partial_t - \partial_r$.

Example 17.8. Consider $g = -dt^2 + g_{ij}dx^i dx^j$. Let u be an eikonal function and assume that the intersection of $\{u = \text{const}\}$ with $\{t = \text{const}\}$ is a topological sphere (this can be achieved by choosing initial data for the eikonal equation appropriately).

Set

$$S_{\tau,v} := \big\{(x,t)\,|\,t = \tau,\ u(t,x) = v\big\}.$$

Put

$$L = -\nabla u = \partial_t u\,\partial_t - g^{ij}\partial_i u\,\partial_j.$$

∂_t is orthogonal to $\sum_\tau$ and L orthogonal to $\{u = v\}$, thus $g^{ij}\partial_i u\,\partial_j$ is orthogonal to $S_{\tau,v}$. Another way of seeing this is to notice that $g^{ij}\partial_i u\,\partial_j$ is the gradient of the function $u(\tau,\cdot)$ on $\sum_\tau$, whose level sets are by assumption spheres; these spheres are precisely $S_{\tau,v}$. Write $\tilde{N} = -g^{ij}\partial_i u\,\partial_j$, so

$$\begin{aligned} g(\tilde{N},\tilde{N}) &= g_{\alpha\beta}\tilde{N}^\alpha\tilde{N}^\beta = g_{ij}\tilde{N}^i\tilde{N}^j = g_{ij}g^{ik}\partial_k u\,g^{jl}\partial_l u \\ &= g^{kl}\partial_k u\,\partial_l u = (\partial_t u)^2. \end{aligned}$$

Since u satisfies $g^{\mu\nu}\partial_\mu u\,\partial_\nu u = -(\partial^t u)^2 + g^{ij}\partial_i u\,\partial_j u = 0$.

Set

$$a := \frac{1}{\partial_t u}.$$

Then, $N = a\tilde{N} = \frac{1}{\partial_t u}g^{ij}\partial_j u\,\partial_i$ is a unit normal vector field to the spheres $S_{\tau,v}$. Let $\{e_1, e_2\}$ be an orthonormal frame on $S_{\tau,v}$(with respect to the metric induced on $S_{\tau,v}$) and set

$$\underline{L} = a(\partial_t - N), \quad L = \frac{1}{a}(\partial_t + N).$$

Then $\{e_1,\, e_2,\, \underline{L},\, L\}$ is a null frame normalized by

$$g(\underline{L}, L) = g(\partial_t, \partial_t) - g(N,N) = -2,$$

where we used $g(\partial_t, N) = 0$.

Observe that L is also tangent to $\mathcal{C} = \{u = v\}$. Indeed, the tangent space to $\{u = v\}$ at a point p is the orthogonal complement to a normal to $T_p\mathcal{C}$. Since $L = -\nabla u$, L is orthonormal to $\mathcal{C}$. But $\{e_1, e_2, L\}$ are linearly independent vectors orthonormal to L (L is null thus orthonormal to itself), and since $\dim T_p\mathcal{C} = 3$, $T_p\mathcal{C} = \text{span}\{e_1,\, e_2,\, L\}$. In particular, L is tangent to $\mathcal{C}$.

Since $\underline{L}$ is linearly independent with respect to $\{e_1,\, e_2,\, L\}$, $\underline{L}$ is transverse to $\mathcal{C}$. Thus, we have obtained the analogue of the previous Minkowski space picture.

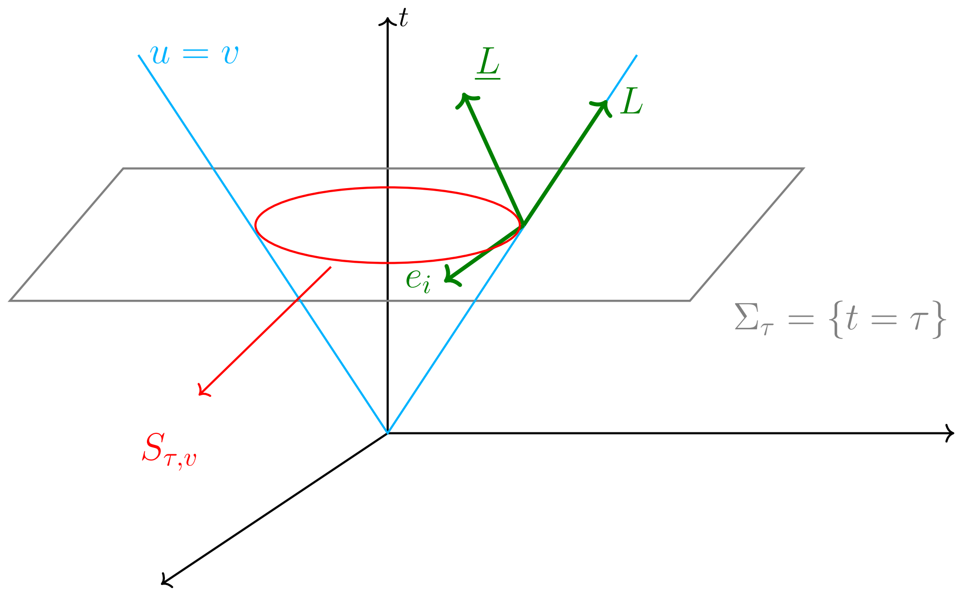


FIGURE 46. L, $\underline{\text{L}}$, and the Minkowski Metric II

17.3. **The null-structure equations.** We saw that we can construct a null frame adapted to the characteristics with the help of an eikonal function. As seen, we hope to prove the decay estimates showing that L, e_A are good derivatives. Because L, e_A are themselves dependent on φ, this turns out to be a difficult coupled problem. With the current techniques known to date, appropriate decay estimates for solutions can only be derived with the help of some complementary estimates for several associated geometric tensors. Such complementary estimates are obtained with the help of a system of elliptic and evolution equations for various geometric quantities associated with the characteristics. Such equations are known as the null-structure equations.

Set-up. Throughout below we consider the situation where the intersections of the level sets of u with $\sum_t$ are topological spheres, like in the example above (although we do not assume, like in that example, that $g_{00} = -1,\ g_{0i} = 0$). Let T be the future-pointing timelike unit normal to $\sum_t$, N the unit outer normal to $S_{t,u}$ inside $\sum_t$, and put $L = T + N$, $\underline{L} = T - N$. Construct a null frame $\{e_1,\ e_2,\ \underline{L},\ L\}$ by considering e_1, e_2 an orthonormal frame on $S_{t,u}$. Note that $g(L, \underline{L}) = -2$.

Let us begin with some basic definitions. Since $\nabla_{e_\alpha} e_\beta$ is a vector field, we can express it relative to a null frame. Thus

$$\nabla_{e_\alpha} e_\beta = \Gamma^\gamma_{\alpha\beta}\, e_\gamma$$

for some (locally defined) functions $\Gamma^\gamma_{\alpha\beta}$ known as the **frame coefficients**. These are exact analogues of the Christoffel symbols when we express ∇X in a coordinate basis. However, their symmetry properties are not the same as in the Christoffel symbols. E.g.,

$$\begin{aligned} g(e_A, e_B) = \delta_{AB} \Rightarrow g(\nabla_{e_C} e_A,\ e_B) + g(e_A,\ \nabla_{e_C} e_B) = 0 \\ \Gamma^D_{C\,A}\, g(e_D,\ e_B) + \Gamma^D_{C\,B}\, g(e_A,\ e_D) = 0 \\ \Gamma^B_{C\,A} + \Gamma^A_{C\,B} = 0. \end{aligned}$$

Recall that $\not\Pi$ is the projection onto span $\{e_1, e_2\}$, which in our case corresponds to projections onto the spheres $S_{t,u}$. We can extend it to projections of arbitrary tensors. For example, if w is a two tensor

$$(\not\Pi \xi)_{\alpha\beta} := \not\Pi^\gamma_\alpha\, \not\Pi^\delta_\beta\, \xi_{\gamma\delta}.$$

We say that a tensor is **tangent** (to $S_{t,u}$) if $\not\Pi\xi = \xi$. For tangent tensors, their $\not g$ norm is defined, e.g.,

$$|\xi|^2_{\not g} = (\not g^{-1})^{\alpha\gamma}(\not g^{-1})^{\beta\delta}\xi_{\alpha\beta}\,\xi_{\gamma\delta}.$$

Note that the above is also given by

$$(\not g^{-1})^{\alpha\gamma}(\not g^{-1})^{\beta\delta}\xi_{\alpha\beta}\,\xi_{\gamma\delta} = e^{\alpha}_{A}\,e^{\gamma}_{A}\,e^{\beta}_{B}\,e^{\delta}_{B}\,\xi_{\alpha\beta}\,\xi_{\gamma\delta} = \xi_{AB}\,\xi_{AB}.$$

If ξ is a $S_{t,u}$-tangent tensor, we can view it as a tensor defined on the entire spacetime by extending it to be zero when contracted against L or $\underline{L}$

We denote by $\not\nabla$ the **projection of the covariant derivative** onto $S_{t,u}$, i.e.,

$$\not\nabla_X\,\xi := \not\prod\nabla_X\xi$$

for any vector field X and tensor ξ. If $\not D$ denotes the connection of the metric $\not g$ on $S_{t,u}$ and both X and ξ are tangent to $S_{t,u}$, then $\not\nabla_X\,\xi = \not D_X\,\xi$.

If ξ is a symmetric $(0,2)$ tensor on $S_{t,u}$, its **trace** relative to $\not g$ is defined as

$$\mathrm{tr}_{\not g}\,\xi = (\not g^{-1})^{\alpha\beta}\xi_{\alpha\beta} = \xi_{AA}.$$

We then define the **trace-free part** of ξ as

$$\hat{\xi} := \xi - \frac{1}{2}\mathrm{tr}_{\not g}\,\xi\,\not g.$$

Notation 17.9. We will abbreviate

$$\nabla_A := \nabla_{e_A}$$

For a $S_{t,n}$-tangent one-form, its **$\not g$-divergence** and **$\not g$-curl** relative to $\{e_A\}_{A=1}^2$ are defined, respectively by

$$\not{\mathrm{div}}\,\xi := \not\nabla_A\,\xi_A,$$
$$\not{\mathrm{curl}}\,\xi := e^{AB}\not\nabla_A\,\xi_B,$$

where ε^{AB} is anti-symmetric on AB with $\varepsilon^{12} = 1$.

Recall that the indices $A, \underline{L}$, and L in the tensors represent contractions with $e_A, \underline{L}$, and L. Such contractions are always taken after the derivatives, e.g.

$$\begin{aligned}\not\nabla_A\,\xi_B &= e^{\alpha}_{A}\,e^{\beta}_{B}\not\nabla_\alpha\,\xi_\beta = e^{\alpha}_{A}\,e^{\beta}_{B}(\not\prod\nabla\,\xi)_{\alpha\beta}\\ &= e^{\alpha}_{A}\,e^{\beta}_{B}\not\prod{}^{\gamma}_{\alpha}\not\prod{}^{\delta}_{\beta}\nabla_\gamma\xi_\delta\end{aligned}$$

If ξ is a $S_{t,u}$-tangent symmetric $(0,2)$ tensor, its **$\not g$-divergence** and **$\not g$-curl** (relative to $\{e_A\}_{A=1}^2$) are defined by

$$\not{\mathrm{div}}\,\xi_A := \not\nabla_B\,\xi_{AB},$$
$$\not{\mathrm{curl}}\,\xi_A := e^{BC}\not\nabla_B\,\xi_{CA}.$$

We denote by $\underline{\prod}$ the projection onto $\sum_t$. IT is not difficult to see that

$$\underline{\prod}{}^{\alpha}_{\beta} = \delta^{\beta}_{\alpha} + T_\alpha T^\beta.$$

We write $\underline{\mathrm{g}}$ for the metric induced on $\sum_t$. For a general tensor ξ, we put

$$\underline{\xi} := \underline{\prod}\,\xi.$$

We denote by $\mathcal{L}$ the Lie derivative. According to our conventions, we have

$$\not{\mathcal{L}}_X \xi := \not\prod \mathcal{L}_X \xi,$$
$$\underline{\mathcal{L}}_X \xi := \underline{\prod} \mathcal{L}_X \xi.$$

Definition 17.10. Let T denote the unit timelike future pointing normal to $\sum_t$ and N the unit outer normal to the spheres $S_{\tau,v} = \sum_\tau \cap \{u = v\}$. Consider a null frame $\{e_1, e_2, \underline{L}, L\}$ as described above (so $L = T + N$, $\underline{L} = T - N$, $g(L, \underline{L}) = -2$) and the above definitions and notation.

We introduce the following quantities, called **connection coefficients:**

Second fundamental form of $\sum_t$:

$$h(X, Y) := -g(\nabla_X T, Y),$$

$X, Y \in T \sum_\tau$.

Second fundamental form of $S_{t,u}$:

$$\theta_{AB} := g(\nabla_A N, e_B).$$

Null second fundamental forms of $S_{t,u}$:

$$\mathcal{X}_{AB} := g(\nabla_A L, e_B), \quad \underline{\mathcal{X}}_{AB} := g(\nabla_A \underline{L}, e_B).$$

$S_{t,n}$**-tangent torsions:**

$$\mathcal{S}_A := \frac{1}{2} g(\nabla_{\underline{L}} L, e_A), \quad \underline{\mathcal{S}}_A := \frac{1}{2} g(\nabla_L \underline{L}, e_A).$$

The null-lapse:

$$b := \frac{1}{|\nabla_{\underline{g}} u|_{\underline{g}}}, \quad \text{g = metric induced on } \sum_t .$$

Proposition 17.11. *The following relations hold: h, θ, X and $\underline{X}$ are symmetric and*

$$h = -\frac{1}{2} \underline{\mathcal{L}}_T\, \underline{g} = -\frac{1}{2} \underline{\mathcal{L}}_T\, g,$$
$$\mathcal{X} = \frac{1}{2} \not{\mathcal{L}}_L \not{g} = \frac{1}{2} \not{\mathcal{L}}_L\, g,$$
$$\underline{\mathcal{X}} = \frac{1}{2} \not{\mathcal{L}}_{\underline{L}} \not{g} = \frac{1}{2} \not{\mathcal{L}}_{\underline{L}}\, g,$$
$$\not\nabla_N N = -\not\nabla \ln b,$$
$$\not\nabla_A N_B = \theta_{AB},$$

Moreover

$$\nabla_A L = \mathcal{X}_{AB}\, e_B - h_{AN} L, \qquad \nabla_A \underline{L} = \mathcal{X}_{AB}\, e_B + h_{AN} \underline{L},$$
$$\nabla_L L = (-h_{NN} + g(\nabla_T T, L))L, \qquad \nabla_L \underline{L} = 2 \underline{\mathcal{S}}_A\, e_A + h_{NN} \underline{L},$$
$$\nabla_L \underline{L} = 2 \mathcal{S}_A e_A + h_{NN} L, \qquad \nabla_{\underline{L}} \underline{L} = -2(\not\nabla_A \ln b) e_A - h_{NN} \underline{L},$$
$$\nabla_L e_A = \not\nabla_L e_A + \underline{\mathcal{S}}_A L, \qquad \nabla_B e_A = \not\nabla_B e_A = \not\nabla e_A + \frac{1}{2} X_{AB} \underline{L} + \frac{1}{2} \underline{X}_{AB} L$$

Finally

$$\mathcal{X}_{AB} = \theta_{AB} - h_{AB}, \quad \underline{\mathcal{X}}_{AB} = -\theta_{AB} - h_{AB},$$
$$\underline{\mathcal{S}}_A = -h_{AN}, \qquad \mathcal{S} = \not\nabla \ln b + h_{AN}.$$

Proof. The first properties (symmetry and identities before "moreover") are standard. The second ones (between "moreover" and "finally") follow from direct computations and the definitions, recalling that for any vector field:

$$X = g(X, e_A)e_A - \frac{1}{2}g(X, L)\underline{L} - \frac{1}{2}g(X, \underline{L})L.$$

For example

$$\nabla_A L = \underbrace{g(\nabla_A L, e_B)}_{=\mathcal{X}_{AB}} e_B - \frac{1}{2} \underbrace{g(\nabla_A L, L)}_{=\frac{1}{2}e_A g(L,L)=0} \underline{L} - \frac{1}{2}g(\nabla_A L, \underline{L})L$$

and

$$\begin{aligned} g(\nabla_A L, \underline{L}) &= g(\nabla_A (T+N), T-N) \\ &= \underbrace{g(\nabla_A T, T)}_{=\frac{1}{2}e_A g(T,T)=0} - \underbrace{g(\nabla_A T, N) + g(\nabla_A N, T)}_{=-g(\nabla_A T,N)+\underbrace{e_A g(N,T)}_{=0}-g(N,\nabla_A T)=2h_{AN}} - \overbrace{g(\nabla_A N, N)}^{=\frac{1}{2}e_A g(N,N)=0} \end{aligned}$$

So $\nabla_A L = \mathcal{X}_{AB}\, e_B - h_{AN}\, L$.

The other relations are proved similarly. The remaining relations (after "finally") follow from the previous ones and the definitions.

□

For the next theorem, we introduce the **volume ratio**:

$$\mathcal{V} := \frac{\sqrt{\det \not{g}}}{\sqrt{\det \not{e}}}$$

where $\not{e}$ is the round metric on $\mathcal{S}^2$. $\mathcal{V}$ is a partial measure of how far $S_{t,u}$ is from being round.

We also define the **mass aspect function** to be

$$\mathcal{M} := \underline{L}\, \mathrm{tr}_{\not{g}}\, \mathcal{X} + \frac{1}{2}\, \mathrm{tr}_{\not{g}}\, \mathcal{X}\, \mathrm{tr}_{\not{g}}\, \underline{\mathcal{X}}.$$

The idea of introducing $\mathcal{M}$ is that we want to control $\underline{L}\, \mathrm{tr}_{\not{g}}\, \mathcal{X}$. However, this quantity does not satisfy a good evolution equation, but $\mathcal{M}$, which can be viewed as modifications of $\underline{L}\, \mathrm{tr}_{\not{g}}\, \mathcal{X}$, does.

The connection coefficients satisfy the following PDEs, known as the **null-structure equations**.

Theorem 17.12. *The connection coefficients satisfy*

$$
\begin{aligned}
L\mathcal{V} &= \mathcal{V}\,\mathrm{tr}_{\slashed{g}}\mathcal{X}\\
Lb &= (-k_{NN} + g(\nabla_T T, L))b\\
L\,\mathrm{tr}_{\slashed{g}}\mathcal{X} + \frac{1}{2}(\mathrm{tr}_{\slashed{g}}\mathcal{X})^2 &= -|\hat{\mathcal{X}}|_{\slashed{g}} - k_{NN}\,\mathrm{tr}_{\slashed{g}} - R_{LL}\\
\slashed{\nabla}_L\hat{\mathcal{X}}_{AB} + (\mathrm{tr}_{\slashed{g}}\mathcal{X})\hat{\mathcal{X}}_{AB} &= -k_{NN}\hat{\mathcal{X}}_{AB} - R_{LALB} + \frac{1}{2}R_{LL}\delta_{AB}\\
\slashed{\nabla}_L\mathcal{S}_A + \frac{1}{2}(\mathrm{tr}_{\slashed{g}}\mathcal{X})\mathcal{S}_A &= -(h_{BN} + \mathcal{S}_B)\hat{\mathcal{X}}_{AB} - \frac{1}{2}\mathrm{tr}_{\slashed{g}} h_{AN} - \frac{1}{2}R_{ALL\underline{L}}\\
L\,\mathrm{tr}_{\slashed{g}}\underline{\mathcal{X}} + \frac{1}{2}\mathrm{tr}_{\slashed{g}}\mathcal{X}\,\mathrm{tr}_{\slashed{g}}\underline{\mathcal{X}} &= 2\,\slashed{\mathrm{div}}\,\underline{\mathcal{S}} + k_{NN}\,\mathrm{tr}_{\slashed{g}}\underline{\mathcal{X}} - \hat{\mathcal{X}}_{AB}\underline{\hat{\mathcal{X}}}_{AB} + 2|\underline{\mathcal{S}}|^2_{\slashed{g}} + R_{A\underline{L}LA}\\
\slashed{\nabla}_{\underline{L}}\hat{\mathcal{X}}_{AB} + \frac{1}{2}\mathrm{tr}_{\slashed{g}}\underline{\mathcal{X}}\hat{\mathcal{X}}_{AB} &= -\frac{1}{2}\mathrm{tr}_{\slashed{g}}\mathcal{X}\underline{\hat{\mathcal{X}}}_{AB} + 2\slashed{\nabla}_A\mathcal{S}_B - \slashed{\mathrm{div}}\,\mathcal{S}\delta_{AB} + h_{NN}\hat{\mathcal{X}}_{AB}\\
&\quad + (2\mathcal{S}_A\mathcal{S}_B - |\mathcal{S}|^2_{\slashed{g}}\delta_{AB}) + R_{AL\underline{L}B} - \frac{1}{2}R_{CL\underline{L}C}\delta_{AB} - \underline{\hat{\mathcal{X}}}_{AC}\hat{\mathcal{X}}_{BC} + \frac{1}{2}\underline{\hat{\mathcal{X}}}_{CD}\hat{\mathcal{X}}_{CD}\delta_{AB}\\
L\mathcal{M} + \mathrm{tr}_{\slashed{g}}\mathcal{X}\mathcal{M} &= \frac{1}{2}L(\mathrm{tr}_{\slashed{g}}\mathcal{X}\,\mathrm{tr}_{\slashed{g}}\underline{\mathcal{X}}) + \frac{1}{2}(\mathrm{tr}_{\slashed{g}}\mathcal{X})^2\,\mathrm{tr}_{\slashed{g}}\underline{\mathcal{X}} - 2\slashed{\nabla}_{\underline{L}}\hat{\mathcal{X}}_{AB}\hat{\mathcal{X}}_{AB}\\
&\quad + 2(\underline{\mathcal{S}}_A - \mathcal{S}_A)\slashed{\nabla}_A\,\mathrm{tr}_{\slashed{g}}\mathcal{X} - \underline{L}R_{LL} - \underline{L}h_{NN}\,\mathrm{tr}_{\slashed{g}}\mathcal{X} - L\,\mathrm{tr}_{\slashed{g}}\mathcal{X}k_{NN}\\
\slashed{\mathrm{div}}\,\hat{\mathcal{X}}_A + \hat{\mathcal{X}}_{AB}h_{BN} &= \frac{1}{2}(\slashed{\nabla}_A\,\mathrm{tr}_{\slashed{g}}\mathcal{X} + k_{AN}\,\mathrm{tr}_{\slashed{g}}\mathcal{X}) + R_{BLBA}\\
\slashed{\mathrm{div}}\,\mathcal{S} &= \frac{1}{2}(\mathcal{M} - k_{NN}\,\mathrm{tr}_{\slashed{g}}\mathcal{X} - 2|\mathcal{S}|^2_{\slashed{g}} - |\hat{\mathcal{X}}|^2_{\slashed{g}} - 2k_{AB}\hat{\mathcal{X}}_{AB}) - \frac{1}{2}R_{A\underline{L}LA}\\
\slashed{\mathrm{curl}}\,\mathcal{S} &= \frac{1}{2}\varepsilon^{AB}\underline{\hat{\mathcal{X}}}_{AC}\hat{\mathcal{X}}_{BC} - \frac{1}{2}\varepsilon^{AB}R_{AL\underline{L}B}.
\end{aligned}
$$

Above, $R_{\alpha\beta}$ is the Ricci curvature and $R_{\alpha\beta\gamma\delta}$ the Riemann curvature of g and, according to our conventions, $R_{LL} = L^\alpha L^\beta R_{\alpha\beta}$, $R_{LAB\underline{L}} = L^\alpha e^\beta_A e^\gamma_B \underline{L}^\gamma R_{\alpha\beta\gamma\delta}$ etc.

Proof. The proof is a series of lengthy calculations. We will derive the equation.

$$
L\,\mathrm{tr}_{\slashed{g}}\mathcal{X} + \frac{1}{2}(\mathrm{tr}_{\slashed{g}}\mathcal{X})^2 = -|\hat{\mathcal{X}}|_{\slashed{g}} - k_{NN}\,\mathrm{tr}_{\slashed{g}} R_{LL},
$$

known as the Raychaudhuri equation, which is one of the main equations in the study of characteristic geometry. It plays an important role now only in global existence problems, but in many other problems (see below). It is also important in the proof of the singularity theorems in general relativity.

$$
\begin{aligned}
\slashed{\nabla}_L\mathcal{X}_{AB} &= e^\alpha_A e^\beta_B\slashed{\nabla}_L\mathcal{X}_{\alpha\beta} = e^\alpha_A e^\beta_B\left(\slashed{\Pi}\nabla_L\mathcal{X}\right)_{\alpha\beta}\\
&= e^\alpha_A e^B_A\slashed{\Pi}^\gamma_\alpha\slashed{\Pi}^\delta_\beta\nabla_L\mathcal{X}_{\gamma\delta} = L^\sigma e^\alpha_A e^\beta_A\slashed{\Pi}^\gamma_\alpha\slashed{\Pi}^\delta_\beta\nabla_\alpha\mathcal{X}_{\gamma\delta},
\end{aligned}
$$

But $e^\alpha_A\,\slashed{\Pi}^\gamma_\alpha = e^\alpha_A\big(\delta^\gamma_\alpha + \frac{1}{2}L^\gamma\underline{L}_\alpha + \underline{L}^\gamma L_\alpha\big) = e^\gamma_A$, so

$$
\begin{aligned}
&= L^\sigma e^\gamma_A e^\delta_B\nabla_\sigma\mathcal{X}_{\gamma\delta} = e^\gamma_A e^\delta_B\nabla_L\mathcal{X}_{\gamma\delta}\\
&= \nabla_L(e^\gamma_A e^\delta_B\mathcal{X}_{\gamma\delta}) - (\nabla_L e^\gamma_A)e^\delta_B\mathcal{X}_{\gamma\delta} - e^\gamma_A(\nabla_L e^\delta_B)\mathcal{X}_{\gamma\delta}\\
&= \nabla_L\mathcal{X}_{AB} - \nabla_L e^\gamma_A e^\delta_B\mathcal{X}_{\gamma\delta} - e^\gamma_A\nabla_L e^\delta_B\mathcal{X}_{\gamma\delta}.
\end{aligned}
$$

Since $\mathcal{X}$ vanishes when contracted with vectors not tangent to $S_{t,u}$, $\mathcal{X}_{\gamma\delta}\nabla_L e^\gamma_A = \mathcal{X}_{\gamma\delta}\slashed{\nabla}_L e^\gamma_A$, so

$$
= \nabla_L\mathcal{X}(e_A, e_B) - \mathcal{X}(\slashed{\nabla}_L e_A, e_B) - \mathcal{X}(e_A, \slashed{\nabla}_L e_B).
$$

But

$$\begin{aligned}\nabla_L \mathcal{X}(e_A, e_B) &= \nabla_L\, g(\nabla_A L, e_B) \\ &= g(\nabla_L \nabla_A L, e_B) + g(\nabla_A L, \nabla_L e_B)\end{aligned}$$

and

$$\begin{aligned}g(e_B, \nabla_L \nabla_A L) &= g(e_B, \nabla_A \nabla_L L + \nabla_{[L,e_A]} L + \mathrm{Riem}(L, e_A) L) \\ &= g(e_B, \nabla_A \nabla_L L) + g(e_B, \nabla_{[L,e_A]} L) + \underbrace{g(e_B, \mathrm{Riem}(L, e_A) L)}_{=\mathrm{Riem}(e_B, L, L, e_A)}.\end{aligned}$$

Thus

$$\begin{aligned}\not\nabla_L \mathcal{X}_{AB} &= g(e_B, \nabla_A \nabla_L L) + g(e_B, \nabla_{[L,e_A]} L) + g(\nabla_A L, \nabla_L e_B) \\ &\quad + \mathrm{Riem}(e_B, L, L, e_A) - \mathcal{X}(\not\nabla_L e_A, e_B) - \mathcal{X}(e_A, \not\nabla_L e_B).\end{aligned}$$

Using, from the proposition above

$$\begin{aligned}\nabla_A L &= \mathcal{X}_{AB} e_B - h_{AN} L \\ \nabla_L L &= (-h_{NN} + g(\nabla_T T, L)) L \\ \nabla_B e_A &= \not\nabla_B e_A + \frac{1}{2} \mathcal{X}_{AB} \underline{L} + \frac{1}{2} \underline{\mathcal{X}}_{AB} L\end{aligned}$$

and

$$\begin{aligned}[L, e_A] &= \not\nabla_L e_A + \underline{\mathcal{S}}_A L - \mathcal{X}_{AB} e_B + h_{AN} L \\ &= \not\nabla_L e_A - \mathcal{X}_{AB} e_B,\end{aligned}$$

which follows from $[X, Y] = \nabla_X Y - \nabla_Y X$ and our previous relations, we find

$$\begin{aligned}g(e_B, \nabla_A \nabla_L L) &= -\mathcal{X}_{AB} h_{NN} \\ g(e_B, \nabla_{[L,e_A]} L) &= \mathcal{X}(\not\nabla_L e_A, e_B) - \mathcal{X}_{AC} \mathcal{X}_{CB} \\ g(\nabla_A L, \nabla_L e_B) &= \mathcal{X}(e_A, \not\nabla_L e_B).\end{aligned}$$

Thus

$$\begin{aligned}\not\nabla_L \mathcal{X}_{AB} &= g(e_B, \nabla_A \nabla_L L) + g(e_B, \nabla_{[L,e_A]} L) + g(\nabla_A L, \nabla_L e_B) \\ &\quad + \mathrm{Riem}(e_B, L, L, e_A) - \mathcal{X}(\not\nabla_L e_A, e_B) - \mathcal{X}(e_A, \not\nabla_L e_B)\end{aligned}$$

becomes ($\mathcal{X}(\not\nabla_L e_A, e_B)$ and $\mathcal{X}(e_A, \not\nabla_L e_B)$ cancel out):

$$\not\nabla_L \mathcal{X}_{AB} = -\mathcal{X}_{AB} h_{NN} - \mathcal{X}_{AC} \mathcal{X}_{CB} + R_{BLLA}$$

Taking the trace and using that, after some algebra

$$-\not g^{AB} \mathcal{X}_{AC} \mathcal{X}_{CB} = -|\hat{\mathcal{X}}|^2_{\not g} - \frac{1}{2} (\mathrm{tr}_{\not g} \mathcal{X})^2$$

we find

$$L\, \mathrm{tr}_{\not g} \mathcal{X} + \frac{1}{2} (\mathrm{tr}_{\not g} \mathcal{X})^2 = -|\hat{\mathcal{X}}|_{\not g} - k_{NN}\, \mathrm{tr}_{\not g} - R_{LL},$$

which is the desired result.

□

Let us now give some intuition of why these geometric constructions are important for the problem of decay. We should warn, however, that the following discussion is very heuristic, and the goal is only to give some idea why the geometric formation introduced above is important for the study of quasilinear wave equations and the study of their decay properties and global existence in particular. A more precise discussion would require a more detailed exposition.

In our standard energy estimates, the energy we control arises from an integration by parts, e.g., for the linear wave equation,

$$-\partial_t^2\varphi + \Delta\varphi = 0,$$

multiply by $\partial_t\varphi$,

$$-\partial_t\varphi\,\partial_t^2\varphi + \partial_t\varphi\,\Delta\varphi = -\frac{1}{2}\partial_t(\partial_t\varphi)^2 + \partial_t\varphi\Delta\varphi = 0,$$

integrating,

$$-\frac{1}{2}\int_{\mathbb{R}^3}\partial_t(\partial_t\varphi)^2 + \underbrace{\int_{\mathbb{R}^3}\partial_t\varphi\Delta\varphi}_{\text{by parts}=-\int_{\mathbb{R}^3}\partial_t\nabla\varphi\cdot\nabla\varphi=-\frac{1}{2}\int_{\mathbb{R}^3}\partial_t|\nabla\varphi|^2}$$

$$= -\frac{1}{2}\partial_t\int_{\mathbb{R}^3}((\partial_t)^2 + |\nabla\varphi|^2) = 0$$

Thus

$$E(t) \leq E(0), \quad E(t) = \frac{1}{2}\int_{\mathbb{R}^3}\left((\partial_t\varphi(t,x))^2 + |\nabla\varphi(t,x,)|^2\right)dx.$$

We saw that we can consider higher versions of this, where we control derivatives of φ, and that this generalizes also to the quasilinear case.

If we want to obtain decay for solutions, it is natural to try to control weighted energies, e.g., expressions of the form:

$$E_w(t) = \int_{\mathbb{R}^3} w\left((\partial_t\varphi)^2 + |\nabla\varphi|^2\right)dx,$$

where w is some "weight". For example, if $w = 1+t$ and we can show something similar to $E_w(t) \leq CE_w(0)$, this would mean that

$$\int_{\mathbb{R}^3}\left(\partial_t\varphi)^2 + |\nabla\varphi|^2\right)dx \leq \frac{C}{1+t}.$$

This is decay of integrals of φ, and we want point-wise decay. But we know that if we control integrals of enough derivatives of φ then we control φ point-wise by Sovolev embedding, thus we seek to bound something like

$$E_w = \int_{\mathbb{R}^3} w\left(|\partial_t D^k\varphi|^2 + |D^{k+1}\varphi|^2\right)dx.$$

In reality, this heuristic is too crude, and the precise way of obtaining pointwise decay from suitable integrals requires more tools, one of which is known as the **Klainerman-Sobolev inequality**, which is a generalization of the Sobolev inequality that involves decay. But even if we take the above heuristics for granted, it remains the question of how to bound the weighted energy, i.e., how to prove $E_w(t) \leq CE_w(0)$. Since our goal is to show that good derivatives have better decay, this should be encoded in our energies. That means that instead of considering the energies obtained by multiplying the equation by $\partial_t\varphi$ and integrating by parts, we should multiply by $X_w\varphi$ and integrate by parts, where X_w is some specific vector field that also carries weight. For example, since we want to show the decay of $L\varphi$, we can choose $X_w\varphi = wL\varphi$, where the weight has to be a suitable power of r or t (for technical reasons r weights, and not only t weights have to be considered). But then, when we integrate by parts, we will pick derivatives of X_w, which need to be estimated. Since X depends on the characteristics (e.g. $X \sim L$), estimating it depends on the estimates for the solution itself. This can be accomplished by decomposing ∂X_w relative to a null frame: the coefficients of this decomposition will be connection coefficients, for example,

$$\nabla L = \underbrace{g(\nabla_{e_A}L, e_B)e_B}_{=\mathcal{X}_{AB}} + \ldots$$

The connection coefficients, in turn, can be estimated from the null structure equations.

In this way, we can get decay estimates, or at least decay estimates for the good derivatives. What about the bad derivative $\underline{L}\varphi$? In general, it is not possible to show that is has integrable decay. However, if the equation has certain special structural conditions (e.g. the null condition previously mentioned), we can show that the bad derivative terms that spoil the integrability in time are in fact absent, leading then to global existence.

Let is finish with the following remark. Even though we introduced the study of the characteristic geometry motivated by the problem of global existence, it turned out that these geometric techniques find applications in many other problems related to the study of quasilinear wave equations. These include the study of shocks and low regularity solutions (by which we mean local well-posedness for data $(\varphi_0, \varphi_1) \in H^{k+1} \times H^k$ with $k \leq \frac{3}{2}$; the classical well-posedness theory, which we developed here, treats only $k > \frac{3}{2}$; or $k > \frac{3}{2} + 1$ if the metric g also depends on $\partial\varphi$).