Lecture Notes 2008

Sergiu Klainerman

Department of Mathematics, Princeton University, Princeton NJ 08544

E-mail address: seri@math.princeton.edu

CHAPTER 1

Basic Tools of Analysis

1. Distribution Theory

This is a very short summary of distribution theory, for more exposure to the subject I suggest F.G. Friedlander and M. Joshi's excellent book *Introduction to the Theory of Distributions*, [**3**]. Hörmander's first volume of *The Analysis of Linear Partial Differential Operators*, [**5**], in Springer can also be useful.

Notation. Throughout these notes we use the notation $A \lesssim B$ to mean $a \leq cB$ where c is a numerical constant, independent of A, B.

1.1. Test Functions. Distributions.

We start with some standard notation. We denote vectors in $\mathbb{R}^n$ by $x = (x_1, \ldots, x_n)$ and set $\lambda x = (\lambda x_1, \ldots, \lambda x_n)$, $x + y = (x_1 + y_1, \ldots, x_n + y_n)$. We denote by $x \cdot y$ the standard scalar product and by $|x| = (x \cdot x)^{\frac{1}{2}}$ the Euclidean length of x. Given a function $f : \Omega \to \mathbb{C}$ we denote by $\operatorname{supp}(f)$ the closure in Ω of the set where $f(x) \neq 0$. We denote by $\mathcal{C}^k(\Omega)$ the set of complex valued functions on Ω which are k times continuously differentiable and by $\mathcal{C}^k_0(\Omega)$ the subset of those which are also compactly supported. We also denote by $\mathcal{C}^\infty(\Omega) = \cap_{k \in \mathbb{N}} \mathcal{C}^k(\Omega)$ the space of infinitely differentiable functions; $\mathcal{C}^\infty_0(\Omega)$ the subset of those which also have compact support. The latter plays a particularly important role in the theory of distributions; it is called the space of test functions on Ω.

Let $\Omega \subset \mathbb{R}^n$ and $f \in \mathcal{C}^\infty(\Omega)$. We denote by $\partial_i f$ the partial derivative $\frac{\partial f}{\partial x_i}$, $i = 1, \ldots, n$. For derivatives of higher order we use the standard multi-index notation. A multi-index α is an n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers with length $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Set $\alpha + \beta = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$. We denote by $\alpha!$ the product of factorials $\alpha_1! \cdots \alpha_n!$. Now set $\partial^\alpha f = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f$. Clearly $\partial^{\alpha+\beta} f = \partial^\alpha \partial^\beta f$. Given two smooth functions u, v we have the Leibnitz formula,

$$\partial^\alpha (u \cdot v) = \sum_{\beta + \gamma = \alpha} \frac{\alpha!}{\beta! \gamma!} \partial^\beta u \partial^\gamma v.$$

Taylor's formula, around the origin, for a smooth function $f : \mathbb{R}^n \to \mathbb{C}$ can be written as follows,

$$f(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} \partial^\alpha f(0) x^\alpha + O(|x|^{k+1}) \qquad \text{as} \quad x \to 0.$$

Here x^α denotes the monomial $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

PROPOSITION 1.2. *Let $f \in \mathcal{C}_0^k(\mathbb{R}^n)$, $0 \le k < \infty$. Let ρ be a test function, i.e. $\rho \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ with supp$(\rho) \subset B(0,1)$, the ball centered at the origin of radius 1, and $\int \rho(x)dx = 1$. We set $\rho_\epsilon(x) = \epsilon^{-n}\rho(x/\epsilon)$ and let*

$$f_\epsilon(x) = f{*}\rho_\epsilon(x) = \epsilon^{-n}\int f(y)\rho(\frac{x-y}{\epsilon})dy = \int f(x-\epsilon z)\rho(z)dz.$$

We have:

1. *The functions f_ϵ are in $\mathcal{C}_0^\infty(\mathbb{R}^n)$ and supp$(f_\epsilon) \subset$ supp$(f) + B(0,\epsilon)$.*
2. *We have $\partial^\alpha f_\epsilon \longrightarrow \partial^\alpha f$ uniformly as $\epsilon \to 0$.*

Proof : The first part of the proposition follows immediately from the definition since the statement about supports is immediate and, by integration by parts, we can transfer all derivatives of f_ϵ on the smooth part of the integrand ρ_ϵ. To prove the second statement we simply write,

$$\partial^\alpha f_\epsilon(x) - \partial^\alpha f(x) = \int \Big(\partial^\alpha f(x-\epsilon z) - \partial^\alpha f(x)\Big)\rho(z)dz.$$

Therefore, for $|\alpha| \le k$,

$$\begin{aligned} |\partial^\alpha f_\epsilon(x) - \partial^\alpha f(x)| &\le \int |\partial^\alpha f(x-\epsilon z) - \partial^\alpha f(x)||\rho(z)|\mathrm{d}z \\ &\le \int |\rho(z)|\mathrm{d}z \sup_{|z|\le 1} |\partial^\alpha f(x-\epsilon z) - \partial^\alpha f(x)| \\ &\lesssim \sup_{|z|\le 1} |\partial^\alpha f(x-\epsilon z) - \partial^\alpha f(x)| \end{aligned}$$

The proof follows now easily in view of the uniform continuity of the functions $\partial^\alpha f$.
■

As a corollary of the Proposition one can easily check that the space of test functions $\mathcal{C}_0^\infty(\Omega)$ is dense in the spaces $\mathcal{C}^k(\Omega)$ as well as $L^p(\Omega)$, $1 \le p < \infty$.

DEFINITION 1.3. A distribution $u \in \mathcal{D}'(\Omega)$ is a linear functional $u : \mathcal{C}_0^\infty(\Omega) \to \mathbb{C}$ verifying the following property:

For any compact set $K \subset \Omega$ there exists an integer N and a constant $C = C_{K,N}$ such that for all $\phi \in \mathcal{C}_0^\infty(\Omega)$, with supp$(\phi) \subset K$ we have

$$|<u,\phi>| \le C \sum_{|\alpha|\le N} \sup |\partial^\alpha \phi|.$$

Equivalently a distribution u is a linear functional $u : \mathcal{C}_0^\infty(\Omega) \to \mathbb{C}$ which is continuous $(***)$ in a certain nonmetrizable locally convex topology defined on $\mathcal{C}_0^\infty(\Omega)$ [1]. In this topology a sequence ϕ_j converges to 0 in $\mathcal{C}_0^\infty(\Omega)$ if all the supports of ϕ_j

[1] This topology can be constructed as an inductive limit topology of Fréchet spaces $\mathcal{C}_K$, where $K \subseteq \Omega$ is compact and $\mathcal{C}_K$ is the space of all smooth functions supported in K, endowed with a Fréchet space structure by the seminorms $\phi \mapsto \sup_K |\partial^\alpha \phi|$ for all multi-indices α. We do not, however, need the precise definition.

are included in a compact subset of Ω and, for each multi-index α, $\partial^\alpha \phi_j \to 0$ in the uniform norm. We have in fact the following characterization of distributions:

PROPOSITION 1.4. *A linear form* $u : \mathcal{C}_0^\infty(\Omega) \longrightarrow \mathbb{C}$ *is a distribution in* $\mathcal{D}'(\Omega)$ *iff* $\lim_{j\to\infty} u(\phi_j) = 0$ *for every sequence of test functions* ϕ_j *which converges to* 0*, in* $\mathcal{C}_0^\infty(\Omega)$*, as* $j \to \infty$*.*

Proof : This proof can be found in Friedlander, section 1.3, Theorem 1.3.2. ■

Example 1: Any locally integrable function $f \in L^1_{\text{loc}}(\Omega)$ defines a distribution,

$$< f, \phi >= \int f\phi, \qquad \forall \phi \in \mathcal{C}_0^\infty(\Omega).$$

We can thus identify $L^1_{\text{loc}}(\Omega)$ as a subspace of $\mathcal{D}'(\Omega)$. This is true in particular for the space $\mathcal{C}^\infty(\Omega) \subset L^1_{\text{loc}}(\Omega)$.

Example 2: The Dirac measure with mass 1 supported at $x_0 \in \mathbb{R}^n$ is defined by,

$$< \delta_{x_0}, \phi >= \phi(x_0).$$

Remark: We shall often denote the action of a distribution u on a test function by $u(\phi)$ instead of $< u, \phi >$. Thus $\delta_{x_0}(\phi) = \phi(x_0)$.

DEFINITION 1.5. A sequence of distributions $u_j \in \mathcal{D}'(\Omega)$ is said to converge, *weakly*, to a distribution $u \in \mathcal{D}'(\Omega)$ if, $u_j(\phi) \to u(\phi)$ for all $\phi \in \mathcal{C}_0^\infty(\Omega)$.

For example the sequence $u_m = e^{imx}$ converges weakly to 0 in $\mathcal{D}'(\mathbb{R})$ as $m \to \infty$. Also if $f \in L^1(\mathbb{R}^n)$, with $\int_{\mathbb{R}^n} f(x)dx = 1$, the family of functions $f_\lambda(x) = \lambda^n f(\lambda x)$ converges weakly to δ_0 as $\lambda \to \infty$.

1.6. Operations with distributions. The advantage of working with the space of distributions is that while this space is much larger than the space of smooth functions most important operations on test functions can be carried over to distributions.

1. Multiplication with smooth functions: Given $u \in \mathcal{D}'(\Omega)$ and $f \in \mathcal{C}^\infty(\Omega)$ we define,

$$< fu, \phi >=< u, f\phi >, \qquad \forall \phi \in \mathcal{C}_0^\infty(\Omega).$$

It is easily verified that multiplication with a smooth function is a continuous endomorphism of the space of distributions.

2. Convolution with a test-function: Consider, $u \in \mathcal{D}'(\mathbb{R}^n), \phi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$. Generalizing the convolution of 2 functions in a natural way, we define

$$u{*}\phi(x) =< u_y, \phi(x-y) >,$$

the subscript specifying that u is understood to be acting on functions of the variable y. Observe that the definition coincides with the usual one if u is a locally integrable function, $u \in L^1_{\text{loc}}(\mathbb{R}^n)$.

Remark: Observe that for every distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ we have $u*\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$. Indeed, e.g. letting e_k denote a standard unit vector,

$$\begin{aligned}\frac{u*\phi(x+he_k) - u*\phi(x)}{h} &= h^{-1} < u_y, \phi(x+he_k-y) - \phi(x-y) > \\ &= < u_y, \int_0^1 \partial_k\phi(x+the_k-y)dt > .\end{aligned}$$

Now if $x \in K$, for some compact set $K \subset \mathbb{R}^n$, then for every sequence $h_i \to 0$, the associated sequence of functions $y \mapsto \int_0^1 \partial_k\phi(x+th_ie_k-y)dt$, together with all its derivatives, converge uniformly toward $\partial_k\phi(x-y)$ and its corresponding derivatives. Moreover they are all compactly supported with supports contained in some compact set K'. Therefore,

$$\lim_{h\to 0}\frac{u*\phi(x+he_k) - u*\phi(x)}{h} = u*\partial_k\phi(x).$$

and thus $u*\phi$ has continuous partial derivatives. We can continue in this manner and conclude that in fact $u*\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$.

3. Differentiation of distributions: For every distribution $u \in \mathcal{D}'(\Omega)$ we define

$$< \partial^\alpha u, \phi >= (-1)^{|\alpha|} < u, \partial^\alpha\phi > .$$

Again, it is easily verified that we have thus defined a continuous endomorphism of the space of distributions. Of course, the operations above were defined so as to extend the usual operations on smooth functions.

We can now define the action of a general linear partial differential operator on distributions. Indeed let,

$$P(x,\partial) = \sum_{|\alpha|\le m} a_\alpha\partial^\alpha, \qquad a_\alpha \in \mathcal{C}^\infty(\Omega),$$

be such an operator. Then,

$$< P(x,\partial)u, \phi >=< u, P(x,\partial)^\dagger\phi >,$$

where $P(x,\partial)^\dagger$ is the *formal adjoint operator*,

$$P(x,\partial)^\dagger v = \sum_{|\alpha|\le m}(-1)^{|\alpha|}\partial^\alpha(a_\alpha v).$$

Observe that if $u_j \in \mathcal{D}'(\Omega)$ converges weakly to $u \in \mathcal{D}'(\Omega)$ then $P(x,\partial)u_j$ converges weakly to $P(x,\partial)u$.

Exercise. Show that for all $u \in \mathcal{D}'(\Omega)$ there exists a sequence $u_j \in \mathcal{C}_0^\infty(\Omega)$ such that $u_j \to u$ as $j \to \infty$ in the sense of distributions(weak convergence). Thus $\mathcal{C}_0^\infty(\Omega)$ is dense in $\mathcal{D}'(\Omega)$, with respect to the weak topology of the latter.

1.7. Example of distributions on the real line.

1.) The simplest nontrivial distribution is the Dirac function $\delta_0 = \delta_0(x)$, defined by $< \delta_0(x), \phi >= \phi(0)$.

2.) Another simple example is the Heaviside function $H(x)$ equal to 1 for $x > 0$ and zero for $x \leq 0$. Or, using the standard identification between locally integrable functions and distributions,

$$< H(x), \phi >= \int_0^\infty \phi(x)dx.$$

Observe that $H'(x) = \delta_0(x)$.

3.) A more elaborate example is $pv(\frac{1}{x})$, or simply $\frac{1}{x}$, called the principal value distribution,

$$< \frac{1}{x}, \phi >= \lim_{\epsilon\to 0} \big(\int_{-\infty}^{-\epsilon} \frac{1}{x}\phi(x)dx + \int_{\epsilon}^{\infty} \frac{1}{x}\phi(x)dx \big).$$

Observe that $\log|x|$ is locally integrable and thus a distribution by the standard identification. It is easy to check that $\frac{d}{dx}\log|x| = \mathrm{pv}(\frac{1}{x})$.

Exercise. Let, for $z \in \mathbb{C}$ with $0 < \arg(z) < \pi$, $\log z = \log|z| + i\arg(z)$. We can regard $x \to \log z = \log(x+iy)$ as a family of distributions depending on $y \in \mathbb{R}^+$. For $x \neq 0$ we have $\lim_{y\to 0^+}\log z = \log|x| + i\pi\big(1 - H(x)\big)$. Show that as $y \to 0$ in $\mathbb{R}^+$, $\partial_x \log z$ converges weakly to a distribution $\frac{1}{x+i0}$ and,

$$\frac{1}{x+i0} = x^{-1} - i\pi\delta_0(x).$$

We now define an important family of distributions χ_+^z, with $z \in \mathbb{C}$, by analytic continuation. For this we first recall the definition of the Gamma function,

Definition 1.8. For $\mathrm{Re}(z) > 0$ we define

$$\Gamma(z) = \int_0^\infty e^{-t}t^{z-1}dt \tag{1}$$

as well as the Beta function,

$$B(a,b) = \int_0^1 s^{a-1}(1-s)^{b-1}ds \tag{2}$$

Clearly $\Gamma(a) = a\Gamma(a-1)$ and $\Gamma(0) = 1$. Thus $\Gamma(n) = n!$. Recall that the following identity holds:

$$B(a,b) = \frac{\Gamma(a)\cdot\Gamma(b)}{\Gamma(a+b)} \tag{3}$$

We also record for future applications,

$$\Gamma(a)\Gamma(1-a) = B(a,1-a) = \frac{\pi}{\sin(\pi a)} \tag{4}$$

In particular $\Gamma(1/2) = \pi^{1/2}$.

Exercise. Prove formulas (12) and (13). For help see Hörmander, [**5**] section 3.4.

Definition 1.9. For $\mathrm{Re}(a) > 0$, we denote by $j_a(\lambda)$ the locally integrable function which is identically zero for $\lambda < 0$ and

$$j_a(\lambda) = \frac{1}{\Gamma(a)}\lambda^{a-1}, \qquad \lambda > 0. \tag{5}$$

The following proposition is well known,

PROPOSITION 1.10. *For all a, b, $Re(a)$, $Re(b) > 0$,*

$$j_a * j_b = j_{a+b}$$

Proof : We have,

$$\begin{aligned} j_a * j_b(\lambda) &= \frac{1}{\Gamma(a)}\frac{1}{\Gamma(b)}\int_0^\lambda \mu^{a-1}(\lambda-\mu)^{b-1}d\mu \\ &= \frac{1}{\Gamma(a)}\frac{1}{\Gamma(b)}\lambda^{a+b-1}\int_0^1 s^{a-1}(1-s)^{b-1}ds \\ &= \frac{B(a,b)}{\Gamma(a)\cdot\Gamma(b)}\lambda^{a+b-1} = \frac{1}{\Gamma(a+b)}\lambda^{a+b-1} = j_{a+b}(\lambda) \end{aligned}$$

■

PROPOSITION 1.11. *There exists a family of distribution j_a, defined for all $a \in \mathbb{C}$, which coincide with the functions j_a for $Re(a) > 0$, such that, $j_a * j_b = j_{a+b}$, $\frac{d}{d\lambda}j_a(\lambda) = j_{a-1}(\lambda)$ and $j_0 = \delta_0$, the Dirac delta function at the origin. Moreover for all positive integers m, $j_{-m}(x) = \partial_x^m\delta_0(x)$.*

Proof : The proof is based on the observation that $\frac{d}{d\lambda}j_a(\lambda) = j_{a-1}(\lambda)$. Thus, for a test function ϕ,

$$\int_{\mathbb{R}} j_{a-1}(\lambda)\phi(\lambda)d\lambda = -\int_{\mathbb{R}} j_a(\lambda)\phi'(\lambda)d\lambda$$

Based on this observation we define, for every $a \in \mathbb{C}$ such that $\mathrm{Re}(a) + m > 0$ as distribution

$$< j_a, \phi >= (-1)^m \int_0^\infty j_{a+m}(\lambda)\,\phi^{(m)}(\lambda)d\lambda$$

In particular,

$$< j_0, \phi >= -\int_0^\infty j_1(\lambda)\phi'(\lambda)d\lambda = -\int_0^\infty \phi'(\lambda)d\lambda = \phi(0)$$

Hence $j_0 = \delta_0$. It is also easy to see that $j_a * j_b = j_{a+b}$ for all $a, b \in \mathbb{C}$. ■

Remark: In applications one often sees the family of distributions $\chi_+^a = j_{a+1}$. Clearly $\chi_+^a * \chi_+^b = \chi_+^{a+b+1}$ and $\chi_+^{-1} = \delta_0$. Observe also that χ_+^a is homogeneous of degree a, i.e. , $\chi_+^a(t\lambda) = t^a\chi_+^a(\lambda)$, for any positive constant t. This clearly makes sense for $\mathrm{Re}(a) > -1$ when χ_+^a is a function. Can you also make sense of it for all $a \in \mathbb{C}$?

1.12. Support of a distribution. The support of a distribution can be easily derived as follows:

DEFINITION 1.13. For $u \in \mathcal{D}'(\Omega)$, we define the complement of the support of u,

$$\Omega\backslash\mathrm{supp}(u) = \{x \in \Omega \mid \exists\, V_x \ni x \text{ open, such that} < u, \phi >= 0 \quad \forall \phi \in \mathcal{C}_0^\infty(V_x)\}.$$

LEMMA 1.14. *If $u \in \mathcal{D}'(\Omega)$ and ϕ is a test function with $supp(\phi) \subset \Omega \backslash supp(u)$, then $< u, \phi >= u(\phi) = 0$.*

Proof : This follows easily by a partition of unity argument. The argument can be found in Friedlander, section 1.4. ■

PROPOSITION 1.15. *A distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ has compact support $K \subset \mathbb{R}^n$ iff there exists $N \in \mathbb{N}$ such that ,$\forall \phi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ we have*

$$|u(\phi)| \leq C \sup_{x \in U} \sum_{|\alpha| \leq N} |\partial^\alpha \phi(x)|,$$

where U is an arbitrary open neighborhood of K.

Proof : This is seen by using a cutoff function which is identically 1 on the support of the distribution. ■

Remark: Note that if we endow $\mathcal{C}^\infty(\mathbb{R}^n)$ with the Frechet topology induced by the family of seminorms given by $\phi \to \sup_{K_i} |\partial^\alpha \phi|$, with $\alpha \in \mathbb{N}^n$ and K_i running over a countable collection of compact sets exhausting $\mathbb{R}^n$, then the space of compactly supported distributions can be identified with $\mathcal{C}^\infty(\mathbb{R}^n)^*$, i.e. the space dual to $\mathcal{C}^\infty(\mathbb{R}^n)$.

We have the following useful fact concerning the structure of distributions supported at one point.

PROPOSITION 1.16. *Let $u \in \mathcal{D}'(\mathbb{R}^n)$ and assume that $supp(u) \subset \{0\}$. Then we have $u = \sum_{|\alpha| \leq N} a_\alpha \partial^\alpha(\delta_0)$, for some integer N, complex numbers a_α and δ_0 the Dirac measure in $\mathbb{R}^n$ supported at 0.*

Proof : See Friedlander, [**3**], Theorem 3.2.1 or Hörmander, [**5**], Theorem 2.3.4. ■

In this context, it is important to observe that the convolution of two distributions cannot be defined in general, but only when certain conditions on the support of the distributions are satisfied. We note in particular the fact that if $u_1, u_2 \in \mathcal{D}'(\mathbb{R}^n)$ one of which is compactly supported, then the convolution $u_1 * u_2$ can be defined. Indeed, assuming u_2 to be compactly supported, we simply define, $(***)$

$$(u_1 * u_2) * \phi = u_1 * (u_2 * \phi), \qquad \forall \phi \in \mathcal{C}_0^\infty(\mathbb{R}^n).$$

Here, $\text{supp}(u_2 * \phi) \subset \{x + y : x \in \text{supp}(u_2),\, y \in \text{supp}(\phi)\}$, hence a compact set. This definition extends the classical convolution for functions.

1.17. Pull back of distributions. Consider first the case of a $\mathcal{C}^\infty$ diffeomorphism $f : \Omega \to \Omega'$ and let u a distribution on Ω'. Then the pull-back f^*u is a distribution in Ω defined by,

$$< f^*u, \phi >=< u(y), g^*\phi(y) |\det Jg(y)| >, \qquad \phi \in \mathcal{C}_0^\infty(\Omega)$$

where $g = f^{-1}$ and $g^*\phi(y) = \phi(g(y))$ and $Jg(y)$ is the jacobian of the map $y \to g(y)$. It is easy to see that this definition is meaningful and that it coincides with the standard change of variable rule when u is a smooth function. Moreover the derivatives of f^*u can be computed by the standard chain rule.

Next we consider the pull back corresponding to a function $f : \Omega \to \mathbb{R}$. This procedure allows us to use the definition of some distributions on the real line to obtain interesting distributions in $\mathbb{R}^n$.

Definition 1.18. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth map with surjective differential everywhere. If $u \in \mathcal{D}'(\mathbb{R})$ we can define its pull-back $f^*(u)$ as follows:

Let $x \in \mathbb{R}^n$ such that[2] $\partial_{x_1} f(x) \neq 0$ on a neighborhood $U \ni x$. Hence the map $y \in U \to (f(y_1, y'), y') \in \mathbb{R}^n$, with $y' = (y_2, \dots, y_n)$, is a local diffeomorphism. Now we set, for every test function ϕ supported in U,

$$f^*(u)(\phi) = u_{y_1}\big(\int \phi(f(y_1, y'), y')|\partial_{y_1} f(y_1, y')|^{-1} dy'\big),$$

In this definition, u_{y_1} indicates that u operates on functions depending on the y_1-variable. Since we can proceed in this fashion for every point in $\mathbb{R}^n$, we can define the pullback of u via f globally by patching the local definitions together via a partition of unity.

Example: If f is as above, then we can explicitly obtain the pullback of the delta function δ_0, namely $f^*(\delta_0) = \frac{1}{|\nabla f|} d\sigma$. Here, $d\sigma$ denotes the canonical surface measure on the embedded sub-manifold $f^{-1}(0) \subset \mathbb{R}^n$ and ∇f denoted the gradient of f.

In connection with the above example, it is useful to observe that if f, g are two smooth functions on $\mathbb{R}^n$ with non-vanishing differential everywhere, then the following equality holds in the sense of distributions for all $a, b \in \mathbb{R}^n$:

$$\int \delta_0(f(a) - x)\delta_0(g(b) - x)dx = \delta_0(f(a) - g(b)).$$

Both sides are to be interpreted as distributions on $\mathbb{R}^n \times \mathbb{R}^n$. To check this, one completes the map $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n \to f(a) - g(b) \in \mathbf{R}$ to a local diffeomorphism, e.g. assuming that $\partial_{a_1} f(a) \neq 0, \partial_{b_1} g(b) \neq 0$, as follows: $(a, b) \to (f(a) - g(b), g(b), a', b')$, where a', b' denote $(a_2, \dots, a_n), (b_2, \dots, b_n)$. Using the above definition of the pull-back of distributions and the fact that the determinant of the Jacobian of this map is the product of the Jacobians of the maps $a \to (f(a), a'), b \to (g(b), b')$, the claim easily follows.

Remark. One cannot define, in general, a meaningful, associative, product of distributions. Why not? Produce an example of three distributions on the real line whose product, if defined, could not be associative.

[2] by surjectivity of the differential, we may always assume this.

1.19. Fundamental solutions. Given a linear partial differential operator with constant coefficients $P(\partial) = \sum_{|\alpha|\le k} a_\alpha \partial^\alpha$, with $a_\alpha \in \mathbb{C}$, we say that a distribution E is a fundamental solution if it verifies $P(\partial)E = \delta_0$. If this is the case then we can always find solution of the equation $P(\partial)u = f$, where $f \in \mathcal{D}'(\mathbb{R}^n)$ is a compactly supported distribution, by setting $u = E*f$. This follows easily from the following proposition together with the observation that $\delta_0*u = u$ for any $u \in \mathcal{D}'(\mathbb{R}^n)$.

PROPOSITION 1.20. *Assume $u, v \in \mathcal{D}'(\mathbb{R}^n)$ one of which is compactly supported. Then,*

$$P(\partial)(u*v) = P(\partial)u*v = u*P(\partial)v.$$

$(***)$ The question of the existence of such fundamental solutions was answered (independently) by Malgrange and Ehrenpreis:

THEOREM 1.21 (Ehrenpreis, Malgrange). *Any linear partial differential operator $P(\partial)$ with constant coefficients has a fundamental solution.*

We omit the proof, which involves elementary Fourier and functional analysis.

In what follows we shall calculate the fundamental solution for some special important differential operators such as the Laplacian $\Delta = \sum_{i=1}^n \partial_i^2$ in $\mathbb{R}^n$, and the D'Alembertian $\Box = -\partial_t^2 + \Delta$ in $\mathbb{R}^{n+1}$. We also consider the Heat operator $\partial_t - \Delta$ and Schrödinger operator $i\partial_t + \Delta$.

1.) $(***)$ *Laplace Operator* Δ. The Laplace operator Δ is invariant under translations and rotations, that is the group of rigid motions. Thus, it makes sense to look for spherically symmetric solutions.

PROPOSITION 1.22. *Define, for all $n \ge 3$, $K_n(x) = \big((2-n)\omega_n\big)^{-1}|x|^{2-n}$ while, for $n = 2$, $K_2(x) = (2\pi)^{-1}\log|x|$. Here ω_n denotes the area of the unit sphere $\mathbb{S}^{n-1}$. Then, for all $n \ge 2$,*

$$\Delta K_n = \delta_0.$$

Remark. By a direct calculation, ΔK_n vanishes away from the origin and therefore can be expressed as a sum of derivatives of δ_0. By homogeneity considerations we can easily infer that $\Delta K_n(x) = c\delta_0$ for some constant c.

Proof : We prove the case $n > 2$; the proof for $n = 2$ is completely analogous. By definition, we have for each test function ϕ,

$$\langle \Delta K_n, \phi\rangle = \langle K_n, \Delta\phi\rangle = \lim_{\epsilon\searrow 0}\int_{r\ge\epsilon} K_n(x)\Delta\phi(x)dx$$

where $r = |x|$. Letting I_ϵ denote the integral under the limit, then integration by parts yields

$$\begin{aligned} I_\epsilon &= -\int_{r=\epsilon} K_n \partial_r \phi dS_\epsilon + \int_{r\geq\epsilon} \nabla K_n \cdot \nabla\phi \\ &= \int_{r=\epsilon} (-K_n\partial_r\phi + \partial_r K_n \cdot \phi) dS_\epsilon + \int_{r\geq\epsilon} \Delta K_n \cdot \phi \\ &= \int_{r=\epsilon} (-K_n\partial_r\phi + \partial_r K_n \cdot \phi) dS_\epsilon \end{aligned}$$

where dS_ϵ denotes the volume element of the sphere $r = \epsilon$. The last step follows since by direct calculation, ΔK_n vanishes away from the origin.

Letting dS denote the volume element of the unit sphere $\mathbb{S}^{n-1}$, a rescaling yields

$$\begin{aligned} I_\epsilon &= \epsilon^{n-1} \int_{\mathbb{S}^{n-1}} (-K_n(\epsilon\omega) \cdot \partial_r\phi(\epsilon\omega) + \partial_r K_n(\epsilon\omega) \cdot \phi(\epsilon\omega)) dS(\omega) \\ &= ((2-n)\omega_n)^{n-1} \epsilon \int_{\mathbb{S}^{n-1}} \partial_r \phi(\epsilon\omega) dS(\omega) + \omega_n^{-1} \int_{\mathbb{S}^{n-1}} \phi(\epsilon\omega) dS(\omega) \end{aligned}$$

Letting $\epsilon \searrow 0$, the first term vanishes, while the second term goes to $\phi(0)$, as desired. ■

With some basic knowledge of differential geometry, we can shorten the above computations. In polar coordinates $x = r\omega, r > 0, |\omega| = 1$, Δ takes the form,

$$\Delta = \partial_r^2 + \frac{n-1}{r}\partial_r + r^{-2}\Delta_{\mathbb{S}^{n-1}},$$

where $\Delta_{\mathbb{S}^{n-1}}$ is the Laplace -Beltrami operator on the unit sphere $\mathbb{S}^{n-1}$.

Exercise. Recall that the Laplace-Beltrami operator on a Riemannian manifold with metric g is given, in local coordinates x^i by

$$\Delta_g \phi = \frac{1}{\sqrt{|g|}} \partial_i \big(g^{ij}\sqrt{|g|}\partial_j\phi\big).$$

Here g^{ij} are the components of the inverse metric g^{-1} relative to the coordinates x^i. The volume element dS_g on M is given, in local coordinates, by $dS_g = \sqrt{|g|}dx^1dx^2\dots dx^n$. Observe that, on compact manifold M,

$$\int_M \Delta_g u\, v dS_g = \int_M u\, \Delta_g v dS_g.$$

Exercise 2. Calculate the Laplace-Beltrami operator for the unit sphere $\mathbb{S}^{n-1}$ and check the polar decomposition formula for Δ. For the particular case $n = 3$, relative to the coordinates $x^1 = r\cos\theta^1, x^2 = r\sin\theta^1\cos\theta^2, x^3 = r\sin\theta^1\sin\theta^2$, $\theta^1 \in [0,\pi), \theta^2 \in [0, 2\pi)$ show that,

$$\Delta_{\mathbb{S}^2} = \partial_{\theta^1}^2 + \text{cotan}\theta^1\, \partial_{\theta_1} + \frac{1}{\sin^2\theta^1}\partial_{\theta^2}^2.$$

Moreover the area element dS_ω takes the form, $dS_\omega = r^2 \sin\theta^1 d\theta^1 d\theta^2$.

Proof (geometric derivation): For a smooth function $\phi(x) = \phi(r\omega)$, in polar coordinates $r = |x|$, $\omega \in \mathbb{S}^{n-1}$ unit sphere in $\mathbb{R}^n$, we have

$$\begin{aligned}\Delta\phi &= \big(\partial_r^2 + \frac{n-1}{r}\partial_r + r^{-2}\Delta_{\mathbb{S}^{n-1}}\big)\phi \\ &= r^{-(n-1)}\partial_r\big(r^{n-1}\partial_r\phi\big) + r^{-2}\Delta_{\mathbb{S}^{n-1}}\phi\end{aligned}$$

Thus passing to polar coordinates $x = r\omega$, with $dx = r^{n-1}drdS_\omega$, in the integral,

$$\begin{aligned}< \Delta K_n, \phi > &= < K_n, \Delta\phi > \\ &= \int_{|\omega|=1}\int_0^\infty K_n(r)\partial_r\big(r^{n-1}\partial_r\phi\big)drdS_\omega + \int_{|\omega|=1}\int_0^\infty K_n(r)\Delta_{\mathbb{S}^{n-1}}\phi drdS_\omega \\ &= \big((2-n)\omega_n\big)^{-1}\int_{|\omega|=1}\int_0^\infty r^{-n+2}\partial_r\big(r^{n-1}\partial_r\phi\big)drdS_\omega \\ &= -\int_0^\infty r^{-n+1}\big(r^{n-1}\partial_r\phi\big)dr = -\int_0^\infty \partial_r\phi = \phi(0)\end{aligned}$$

we infer that, for $n \ge 3$, $\Delta K_n = \delta_0$ as desired. The case $n = 2$ can be treated in the same manner.

Remark : Observe that, up to a constant, the expression of $K_n(x)$ can be easily guessed by looking for spherically symmetric solutions $K = K(|x|)$. Indeed, the equation $\Delta K = 0$ reduces to the ODE, $K''(r) + \frac{n-1}{r}K'(r) = 0$.

∎

According to the general theory we can now solve the Poisson equation $\Delta u = f$, for any smooth compactly supported f, by the formula,

$$u(x) = \int_{\mathbb{R}^n} K_n(x-y)f(y)dy = \int_{\mathbb{R}^n} K_n(y)f(x-y)dy. \tag{6}$$

For $n \ge 3$ we observe that the solution given by (15) decays to zero as $|x| \to \infty$. Indeed, for large $|x|$ we can write (15) in the form

$$u(x) = c_n|x|^{-(n-1)}\int_{\mathbb{R}^n}(1 - \frac{|y|}{|x|})^{-(n-1)}f(y)dy \lesssim |x|^{-(n-1)},$$

due to the fact that f has compact support. We claim that the equation $\Delta u = f$ has a unique solutions $u(x)$ which decays at ∞ as $x \to \infty$ and therefore it must be represented by the integral formula (15). For $n = 2$, on the other hand, we only have $|u(x)| \lesssim \log|x|$. Observe however that

$$|\partial_i u(x)| \lesssim \int_{\mathbb{R}^2} |\partial_i K_2(x-y)||f(y)|dy \lesssim |x|^{-1}$$

since $|\partial K_2(x-y)| \lesssim |x-y|^{-1}$.

PROPOSITION 1.23. *For any $f \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, $n \ge 3$ the equation $\Delta u = f$ has a unique smooth solution which vanishes at infinity, i.e. tends to zero as $|x| \to \infty$. The solution is represented by* (15). *For $n = 2$ the same equation has a unique smooth solution $u(x)$ with $\lim_{|x|\to\infty}\frac{|u(x)|}{|x|} = 0$ and $|\partial u(x)| \to 0$ as $|x| \to \infty$. The solution is represented by* (15)*, up to an additive constant.*

Proof : By linearity it suffices to take $f = 0$. For $n \geq 3$ we have to show that $\Delta u = 0$, with u vanishing at infinity, implies that $u = 0$. The result is an easy consequence of Liouville's theorem which states that every bounded harmonic[3] function in $\mathbb{R}^n$ is constant. Liouville's theorem follows from the maximum principle for Δ according to which the extreme values of a harmonic function, i.e. a solution to $\Delta u = 0$, in a domain D must be attained at the boundary of D. We shall return to both Liouville's theorem and the maximum principle later. However you can try to prove directly the version of the maximum principle needed here. In the case $n = 2$ we can use the same argument to show that the derivatives of a solution $u(x)$ of $\Delta u = 0$, with the properties mentioned in the proposition, must vanish. ■

We shall now give an alternative, direct, proof of the fact that the function $u(x)$ defined by (15) is a solution of $\Delta u = f$. Indeed,

$$\Delta u(x) = \int_{\mathbb{R}^n} K_n(y)\Delta_y f(x-y)dy.$$

We would like to integrate by parts and make use of the fact that $\Delta K_n(x) = 0$ on $\mathbb{R}^n \setminus \{0\}$. We cannot do it directly because the singularity at the origin. We circumvent this difficulty by the standard trick of decomposing the integral $I(x)$ on the right into a regular part $R_\epsilon(x) = \int_{\mathbb{R}^n \setminus B_e} K_n(y)\Delta_y f(x-y)dy$ and a singular part $S_\epsilon = \int_{B_\epsilon} K_n(y)\Delta_y f(x-y)dy$ where $\epsilon > 0$ is an arbitrary small number and B_ϵ is the closed ball of radius ϵ centered at the origin. For the singular part S_ϵ we have, for $n \geq 3$,

$$|S_\epsilon(x)| \lesssim \epsilon^2 \|\partial^2 f\|_{L^\infty}$$

and therefore converges to zero as $\epsilon \to 0$.

For the regular part,

$$R_\epsilon(x) = \int_{\mathbb{R}^n \setminus B_e} K_n(y)\Delta_y f(x-y)dy$$

we are allowed to integrate by parts. Doing it carefully by keeping track of the boundary terms on ∂B_ϵ and powers of ϵ we easliy infer that $|R_\epsilon(x) - f(x)|$ tends to zero as $\epsilon \to 0$, for all values of x.

2.) *D'Alembertian operator* $\Box$. We shall next look of a fundamental solution for the wave operator,

$$\Box = -\partial_t^2 + \Delta = -\partial_t^2 + \partial_r^2 + \frac{n-1}{r}\partial_r + r^{-2}\Delta_{\mathbb{S}^{n-1}}$$

in $\mathbb{R}^{n+1}$. We look for solutions of the form[4] $\phi(t,x) = f(\rho)$ where $\rho = (t^2 - |x|^2)^{1/2}$, in the region $|x| < t$. By a simple calculation we find $f''(\rho) + \frac{n}{r}f'(\rho) = 0$ with solutions $f(\rho) = a\rho^{-\frac{n-1}{2}} + b$. Therefore a good candidate for a fundamental solution must have the form $E = (t^2 - |x|^2)^{-\frac{n-1}{2}}$ in the region $t > |x|$. To extend this definition to all space $\mathbb{R}^{n+1}$ and derive a distribution supported in the region $\{(t,x) : |x| \leq t\}$ we are led to look at the pull back $f^*(\chi_+^{-\frac{n-1}{2}})$ of the one dimensional distribution

[3]Solutions to $\Delta u = 0$ are called harmonic.

[4]In other words we look for solutions invariant under Lorentz transformations. We shall discuss later and in more detail the geometric significance of the wave operator and its symmetries.

$\chi_+^{-\frac{n-1}{2}}$, where f is the map $f(t,x) = t^2 - |x|^2$. For simplicity we write this as $\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2)$.

Note that the expression $\chi_+^{-\frac{n-1}{2}}(t^2-|x|^2)$ is not exactly rigorous, since the gradient of $t^2-|x|^2$ vanishes at the origin, and hence $\chi_+^{-\frac{n-1}{2}}(t^2-|x|^2)$ defines a distribution only on $\mathbf{R}^{n+1} - \{0\}$. A rigorous formulation requires a bit more care. For clarity and convenience, in what follows, we will adopt a heuristic and informal approach to deriving the fundamental solutions.

To make sure that the proposed fundamental distribution is supported in $|x| \le t$, we set (informally)

$$E_+^{(n+1)}(t,x) = c_n H(t)\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2) \tag{7}$$

with $H(t)$ the Heavyside function supported on $t \ge 0$ and c_n a normalizing constant to be determined. In fact $c_n = -\frac{1}{2}\pi^{\frac{1-n}{2}}$.

Proposition 1.24. *The distribution $E_+^{(n+1)}$ is supported in $|x| \le t$ and verifies $\Box E_+^{(n+1)} = \delta_0$.*

In the following heuristic "proofs" of the proposition, we treat $\chi_+^{-\frac{n-1}{2}}(t^2-|x|^2)$ as locally integrable functions on $\mathbf{R}^{n+1}$. In reality, this is clearly not the case, and hence the heuristic computations performed here will need to be rigorously justified.

Proof [for $n = 3$]: In this case we have to check that

$$E_+(t,x) \quad = \quad -\frac{1}{2}\pi^{-1}H(t)\delta_0(t^2 - |x|^2) = -\frac{1}{4\pi}r^{-1}\delta(t-r)$$

with $r = |x|$. Thus, since $\Box\phi = -r^{-1}(\partial_t + \partial_r)(\partial_t - \partial_r)(r\phi) + \Delta_{\mathbb{S}^2}\phi$, we have with $\psi(t, r\omega) = (\partial_t - \partial_r)\big(r\phi(t, r\omega)\big)$,

$$\begin{aligned} < E_+, \Box\phi > &= \frac{1}{4\pi}\int_{\mathbb{S}^2}\int_0^\infty\int_0^\infty \delta(t-r)(\partial_t+\partial_r)\psi dt dr dS_\omega \\ &= \frac{1}{4\pi}\int_{\mathbb{S}^2} dS_\omega\big(\int_0^\infty \frac{d}{dr}\psi(r,r)dr\big) \\ &= -\psi(0,0) = \phi(0) \end{aligned}$$

Thus, $\Box E_+ = \delta_0$ as desired. ∎

We shall now consider the general case. Let $E(t,x) = H(t)\chi_+^{-(n-1)/2}$. We write, for an arbitrary test function $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^{n+1})$,

$$< \Box E, \phi >= \int_0^\infty \int_{\mathbb{R}^n} E(t,x)\Box\phi dt dx = \lim_{\epsilon\to 0}\int_\epsilon^\infty\int_{\mathbb{R}^n} E(t,x)\Box\phi dt dx$$

We integrate by parts in the slab region $[\epsilon, \infty) \times \mathbb{R}^n$,

$$\begin{aligned}
\int_\epsilon^\infty \int_{\mathbb{R}^n} E \cdot (-\partial_t^2 + \sum_{i=1}^n \partial_i^2)\phi &= +\int_\epsilon^\infty \int_{\mathbb{R}^n} (\partial_t E \partial_t \phi - \partial_i E \cdot \partial_i \phi) + \int_{\mathbb{R}^n} E \partial_t \phi(\epsilon, x) \\
&= \int_\epsilon^\infty \int_{\mathbb{R}^n} \Box E \cdot \phi - \int_{\mathbb{R}^n} \partial_t E \phi(\epsilon, x) + \int_{\mathbb{R}^n} E \partial_t \phi(\epsilon, x) \\
&= -\int_{\mathbb{R}^n} \partial_t E \phi(\epsilon, x) + \int_{\mathbb{R}^n} E \partial_t \phi(\epsilon, x)
\end{aligned}$$

since, away from from the tip $t = |x| = 0$, we have $\Box\big(\chi_+^{-(n-1)/2}(t^2 - |x|^2)\big) = 0$. Why?

Now, making the change of variables $x = \epsilon y$ and using the homogeneity[5] of $\chi_+^{-(n-1)/2}$,

$$\begin{aligned}
\int_{\mathbb{R}^n} E \partial_t \phi(\epsilon, x) &= \int_{\mathbb{R}^n} \chi_+^{-(n-1)/2}(\epsilon^2 - |x|^2) \partial_t \phi(\epsilon, x) dx \\
&= \int_{\mathbb{R}^n} \chi_+^{-(n-1)/2}\big(\epsilon^2(1 - |y|^2)\big) \partial_t \phi(\epsilon, \epsilon y) \epsilon^n dy \\
&= \epsilon \int_{\mathbb{R}^n} \chi_+^{-(n-1)/2}(1 - |y|^2) \partial_t \phi(\epsilon, \epsilon y) dy \to 0 \qquad \text{as} \quad \epsilon \to 0
\end{aligned}$$

On the other hand,

$$\partial_t \chi_+^{(n-1)/2}(t^2 - |x|^2) = 2t \chi_+^{-(n+1)/2}(t^2 - |x|^2).$$

Hence,

$$\begin{aligned}
\int_{\mathbb{R}^n} \partial_t E \cdot \phi(\epsilon, x) &= 2\epsilon \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(\epsilon^2 - |x|^2) \phi(\epsilon, x) dx \\
&= 2\epsilon \int_{\mathbb{R}^n} \epsilon^{-(n+1)} \chi_+^{-(n+1)/2}(1 - |y|^2) \phi(\epsilon, \epsilon y) \epsilon^n dy \\
&= 2 \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(1 - |y|^2) \phi(\epsilon, \epsilon y) dy
\end{aligned}$$

Now observe that the distibution $\chi_+^{-(n+1)/2}(1-|y|^2)$ is supported in $|y| \le 1$. Choose a test function $\psi(y)$ in $\mathbb{R}^n$ equal to 1 for $|y| \le 2$ and supported in $|y| \le 4$. Clearly,

$$\begin{aligned}
\lim_{\epsilon \to 0} \int_{\mathbb{R}^n} \partial_t E \cdot \phi(\epsilon, x) &= 2 \lim_{\epsilon \to 0} \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(1 - |y|^2) \phi(\epsilon, \epsilon y) \psi(y) dy \\
&= 2\phi(0) \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(1 - |y|^2) \psi(y) dy
\end{aligned}$$

Therefore we conclude that,

$$< \Box E, \phi >= -2 J_n \phi(0)$$

where $J_n = \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(1 - |y|^2)\psi(y) dy$. To finish we only have to calculate J.

LEMMA 1.25. *For a function $\psi \in \mathcal{C}_0^\infty$ which is identically 1 in a neighborhood of the origin, we have*

$$J_n = \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(1 - |y|^2)\psi(y) dy = \frac{1}{2} c_n^{-1}$$

[5] It is simple to check that, as distributions, $\chi_+^s(\lambda t) = \lambda^s \chi_+^s(t)$.

where $c_n = 1/2\pi^{(1-n)/2}$.

Proof : We consider the cases $n = 2, n = 3$. For $n = 3$,

$$
\begin{aligned}
J_3 &= \int_{\mathbb{S}^2}\int_0^\infty \chi_+^{-2}(1-r^2)r^2\psi(r\omega)dr = -2^{-1}\int_{\mathbb{S}^2}\int_0^\infty \frac{d}{dr}\big(\chi_+^{-1}(1-r^2)\big)r\psi(r\omega)dr \\
&= 2^{-1}\int_{\mathbb{S}^2}\int_0^\infty \delta(1-r^2)\frac{d}{dr}(r\psi(r\omega))dr = 4^{-1}\int_{\mathbb{S}^2}\int_0^\infty \delta(1-r)\frac{d}{dr}(r\psi(r\omega))dr \\
&= \pi
\end{aligned}
$$

as desired.

For $n = 2$, since $\chi_+^{-1/2}(1-s^2) = \frac{1}{\Gamma(1/2)}(1-s^2)^{-1/2} = \pi^{-1/2}(1-s^2)^{-1/2}$ and the derivatives of ψ vanish for $r \leq 2$,

$$
\begin{aligned}
J_2 &= -2^{-1}\int_{|\omega|=1}\int_0^\infty \frac{d}{dr}\big(\chi^{-1/2}(1-r^2)\big)\psi(r\omega)dr \\
&= 2^{-1}\cdot 2\pi\cdot\chi^{-1/2}(0)\psi(0) = \pi^{1/2}.
\end{aligned}
$$

■

Remark: As mentioned before, one needs to be more careful in order to express the above fundamental solutions formally. For example, in the case $n = 3$, where $\chi_+^{-\frac{n-1}{2}} = \delta_0$, we can write $E_+^{(3+1)}$ rigorously as follows:

$$\langle E_+^{(3+1)}, \phi\rangle = -\int_{\mathbf{R}^3}\frac{\phi(|x|, x)}{4\pi|x|}dx$$

Since r^{-1} defines a locally integrable function in $\mathbf{R}^3$, one can show that the above defines a valid distribution on $\mathbf{R}^{3+1}$.

Finally, we can construct an analogous fundamental solution $E_-^{(n-1)}$ supported the past cone $|x| \leq -t$. In particular, fundamental solutions are not unique.

3.) *Heat Operator* $\mathcal{H}$. We consider the heat operator $\mathcal{H} = \partial_t - \Delta$ acting on functions defined on $\mathbb{R}\times\mathbb{R}^n = \mathbb{R}^{n+1}$. It makes sense to look for spherically symmetric solutions to $\mathcal{H}u = 0$, that is functions $u(t,x) = u(t,|x|) = u(t,r)$. It is easy to find in this way the class of locally integrable solutions $E_c(t,x) = cH(t)\, t^{-\frac{n}{2}}e^{-|x|^2/4t}$, with $H(t)$ the heaviside function. Indeed $\mathcal{H}(E_c) = 0$ for all (t,x) with $t \neq 0$. We show below that, in the whole space, $\mathcal{H}(E_c)$ is proportional to δ_0 and that we can determine the constant $c = c_n = 2^{-n}\pi^{-\frac{n}{2}}$ such that the corresponding $E = E_c$ is a fundamental solution of $\mathcal{H}$, i.e. $\mathcal{H}(E) = \delta_0$.

Indeed, if $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^{n+1})$,

$$\begin{aligned}
<\mathcal{H}(E), \phi> &= < E, \mathcal{H}^t\phi >= -\int E(t,x)(\partial_t + \Delta)\phi(t,x)dxdt \\
&= -\lim_{\epsilon\to 0^+}\int_\epsilon^\infty \int_{\mathbb{R}^n} E(t,x)(\partial_t+\Delta)\phi(t,x)dxdt \\
&= \lim_{\epsilon\to 0^+}\int_\epsilon^\infty \int_{\mathbb{R}^n} (\partial_t+\Delta)E(t,x)\phi(t,x)dxdt + \lim_{\epsilon\to 0^+}\int_{\mathbb{R}^n} E(\epsilon,x)\phi(\epsilon,x)dx \\
&= \lim_{\epsilon\to 0^+}\int_{\mathbb{R}^n} E(\epsilon,x)\phi(x,\epsilon)dx = c_n \lim_{\epsilon\to 0^+}\epsilon^{-n/2}\int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon}\phi(x,\epsilon)dx
\end{aligned}$$

We now perform the change of variables $x = 2\epsilon^{1/2}y$,

$$\begin{aligned}
<\mathcal{H}(E), \phi> &= 2^n c_n \lim_{\epsilon\to 0^+}\int_{\mathbb{R}^n}\phi(\epsilon, 2\epsilon^{1/2}y)e^{-|y|^2}dy = 2^n c_n \phi(0,0)\int_{\mathbb{R}^n} e^{-|y|^2}dy \\
&= \phi(0,0)
\end{aligned}$$

Exercise. Check that $\int_{\mathbb{R}^n} e^{-|y|^2}dy = \pi^{n/2}$.

This proves that

$$E(t,x) = (4\pi t)^{-n/2}H(t)\ t^{-\frac{n}{2}}e^{-|x|^2/4t} \tag{8}$$

is a fundamental solution for $\mathcal{H}$.

4.) *Schrödinger equation* $\mathcal{S}$. The Schrödinger operator, $\mathcal{S} = i\partial_t + \Delta$ has a fundamental solution which looks, superficially, exactly like that of the Heat operator,

$$E(t,x) = (4\pi i t)^{-n/2}H(t)\ e^{i|x|^2/4t} \tag{9}$$

Yet, of course, the presence of i in the exponential factor $e^{-i|x|^2/4t}$ makes a world of difference.

Exercise Show that the locally integrable function E is indeed a fundamental solution for $\mathcal{S}$.

2. Distribution Theory

This is a very short summary of distribution theory, for more exposure to the subject I suggest F.G. Friedlander and M. Joshi's excellent book *Introduction to the Theory of Distributions*, [**3**]. Hörmander's first volume of *The Analysis of Linear Partial Differential Operators*, [**5**], in Springer can also be useful.

Notation. Throughout these notes we use the notation $A \lesssim B$ to mean $a \le cB$ where c is a numerical constant, independent of A, B.

2.1. Test Functions. Distributions.

We start with some standard notation. We denote vectors in $\mathbb{R}^n$ by $x = (x_1, \ldots, x_n)$ and set $\lambda x = (\lambda x_1, \ldots, \lambda x_n)$, $x + y = (x_1 + y_1, \ldots, x_n + y_n)$. We denote by $x \cdot y$ the standard scalar product and by $|x| = (x \cdot x)^{\frac{1}{2}}$ the Euclidean length of x. Given a function $f : \Omega \to \mathbb{C}$ we denote by supp(f) the closure in Ω of the set where $f(x) \neq 0$. We denote by $\mathcal{C}^k(\Omega)$ the set of complex valued functions on Ω which are k times continuously differentiable and by $\mathcal{C}_0^k(\Omega)$ the subset of those which are also compactly supported. We also denote by $\mathcal{C}^\infty(\Omega) = \cap_{k\in\mathbb{N}}\mathcal{C}^k(\Omega)$ the space of infinitely differentiable functions; $\mathcal{C}_0^\infty(\Omega)$ the subset of those which also have compact support. The latter plays a particularly important role in the theory of distributions; it is called the space of test functions on Ω.

Let $\Omega \subset \mathbb{R}^n$ and $f \in \mathcal{C}^\infty(\Omega)$. We denote by $\partial_i f$ the partial derivative $\frac{\partial f}{\partial x_i}$, $i = 1, \ldots, n$. For derivatives of higher order we use the standard multi-index notation. A multi-index α is an n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers with length $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Set $\alpha + \beta = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$. We denote by $\alpha!$ the product of factorials $\alpha_1! \cdots \alpha_n!$. Now set $\partial^\alpha f = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f$. Clearly $\partial^{\alpha+\beta} f = \partial^\alpha \partial^\beta f$. Given two smooth functions u, v we have the Leibnitz formula,

$$\partial^\alpha (u \cdot v) = \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!} \partial^\beta u \partial^\gamma v.$$

Taylor's formula, around the origin, for a smooth function $f : \mathbb{R}^n \to \mathbb{C}$ can be written as follows,

$$f(x) = \sum_{|\alpha|\leq k} \frac{1}{\alpha!} \partial^\alpha f(0) x^\alpha + O(|x|^{k+1}) \qquad \text{as} \quad x \to 0.$$

Here x^α denotes the monomial $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

PROPOSITION 2.2. *Let $f \in \mathcal{C}_0^k(\mathbb{R}^n)$, $0 \leq k < \infty$. Let ρ be a test function, i.e. $\rho \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ with supp$(\rho) \subset B(0,1)$, the ball centered at the origin of radius 1, and $\int \rho(x)dx = 1$. We set $\rho_\epsilon(x) = \epsilon^{-n}\rho(x/\epsilon)$ and let*

$$f_\epsilon(x) = f * \rho_\epsilon(x) = \epsilon^{-n} \int f(y) \rho(\frac{x-y}{\epsilon}) dy = \int f(x - \epsilon z) \rho(z) dz.$$

We have:

(1) *The functions f_ϵ are in $\mathcal{C}_0^\infty(\mathbb{R}^n)$ and supp$(f_\epsilon) \subset$ supp$(f) + B(0,\epsilon)$.*
(2) *We have $\partial^\alpha f_\epsilon \longrightarrow \partial^\alpha f$ uniformly as $\epsilon \to 0$.*

Proof : The first part of the proposition follows immediately from the definition since the statement about supports is immediate and, by integration by parts, we can transfer all derivatives of f_ϵ on the smooth part of the integrand ρ_ϵ. To prove the second statement we simply write,

$$\partial^\alpha f_\epsilon(x) - \partial^\alpha f(x) = \int \Big(\partial^\alpha f(x - \epsilon z) - \partial^\alpha f(x)\Big) \rho(z) dz.$$

Therefore, for $|\alpha| \leq k$,

$$\begin{aligned}
|\partial^\alpha f_\epsilon(x) - \partial^\alpha f(x)| &\leq \int |\partial^\alpha f(x - \epsilon z) - \partial^\alpha f(x)| |\rho(z)| \mathrm{d}z \\
&\leq \int |\rho(z)| \mathrm{d}z \sup_{|z|\leq 1} |\partial^\alpha f(x - \epsilon z) - \partial^\alpha f(x)| \\
&\lesssim \sup_{|z|\leq 1} |\partial^\alpha f(x - \epsilon z) - \partial^\alpha f(x)|
\end{aligned}$$

The proof follows now easily in view of the uniform continuity of the functions $\partial^\alpha f$. ■

As a corollary of the Proposition one can easily check that the space of test functions $\mathcal{C}_0^\infty(\Omega)$ is dense in the spaces $\mathcal{C}^k(\Omega)$ as well as $L^p(\Omega)$, $1 \le p < \infty$.

Definition 2.3. A distribution $u \in \mathcal{D}'(\Omega)$ is a linear functional $u : \mathcal{C}_0^\infty(\Omega) \to \mathbb{C}$ verifying the following property:

For any compact set $K \subset \Omega$ there exists an integer N and a constant $C = C_{K,N}$ such that for all $\phi \in \mathcal{C}_0^\infty(\Omega)$, with supp$(\phi) \subset K$ we have

$$| < u, \phi > | \le C \sum_{|\alpha| \le N} \sup |\partial^\alpha \phi|.$$

Equivalently a distribution u is a linear functional $u : \mathcal{C}_0^\infty(\Omega) \to \mathbb{C}$ which is continuous if the space of test functions is endowed with the *standard Frechet space structure*[6]. In this topology a sequence ϕ_j converges to 0 in $\mathcal{C}_0^\infty(\Omega)$ if all the supports of ϕ_j are included in a compact subset of Ω and, for each multi-index α, $\partial^\alpha \phi_j \to 0$ in the uniform norm. We have in fact the following characterization of distributions:

Proposition 2.4. *A linear form $u : \mathcal{C}_0^\infty(\Omega) \longrightarrow \mathbb{C}$ is a distribution in $\mathcal{D}'(\Omega)$ iff $\lim_{j\to\infty} u(\phi_j) = 0$ for every sequence of test functions ϕ_j which converges to 0, in $\mathcal{C}_0^\infty(\Omega)$, as $j \to \infty$.*

Proof : This proof can be found in Friedlander, section 1.3, Theorem 1.3.2. ■

Example 1: Any locally integrable function $f \in L^1_{\text{loc}}(\Omega)$ defines a distribution,

$$< f, \phi >= \int f\phi, \qquad \forall \phi \in \mathcal{C}_0^\infty(\Omega).$$

We can thus identify $L^1_{\text{loc}}(\Omega)$ as a subspace of $\mathcal{D}'(\Omega)$. This is true in particular for the space $\mathcal{C}^\infty(\Omega) \subset L^1_{\text{loc}}(\Omega)$.

Example 2: The Dirac measure with mass 1 supported at $x_0 \in \mathbb{R}^n$ is defined by,

$$< \delta_{x_0}, \phi >= \phi(x_0).$$

Remark: We shall often denote the action of a distribution u on a test function by $u(\phi)$ instead of $< u, \phi >$. Thus $\delta_{x_0}(\phi) = \phi(x_0)$.

Definition 2.5. A sequence of distributions $u_j \in \mathcal{D}'(\Omega)$ is said to converge, *weakly*, to a distribution $u \in \mathcal{D}'(\Omega)$ if, $u_j(\phi) \to u(\phi)$ for all $\phi \in \mathcal{C}_0^\infty(\Omega)$.

[6]This is the topology induced by the countable family of seminorms $\phi \mapsto \sup_{K_i} |\partial^{(\alpha)}\phi|$, where K_i is a countable family of compact sets exhausting Ω, and α ranges over all natural multi-indices. We do not need however the precise definition.

For example the sequence $u_m = e^{imx}$ converges weakly to 0 in $\mathcal{D}'(\mathbb{R})$ as $m \to \infty$. Also if $f \in L^1(\mathbb{R}^n)$, with $\int_{\mathbb{R}^n} f(x)dx = 1$, the family of functions $f_\lambda(x) = \lambda^n f(\lambda x)$ converges weakly to δ_0 as $\lambda \to \infty$.

2.6. Operations with distributions. The advantage of working with the space of distributions is that while this space is much larger than the space of smooth functions most important operations on test functions can be carried over to distributions.

1. Multiplication with smooth functions: Given $u \in \mathcal{D}'(\Omega)$ and $f \in \mathcal{C}^\infty(\Omega)$ we define,

$$< fu, \phi >=< u, f\phi >, \qquad \forall \phi \in \mathcal{C}_0^\infty(\Omega).$$

It is easily verified that multiplication with a smooth function is a continuous endomorphism of the space of distributions.

2. Convolution with a test-function: Consider, $u \in \mathcal{D}'(\mathbb{R}^n), \phi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$. Generalizing the convolution of 2 functions in a natural way, we define

$$u*\phi(x) =< u_y, \phi(x-y) >,$$

the subscript specifying that u is understood to be acting on functions of the variable y. Observe that the definition coincides with the usual one if u is a locally integrable function, $u \in L^1_{\text{loc}}(\mathbb{R}^n)$.

Remark: Observe that for every distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ we have $u*\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$. Indeed, e.g. letting e_k denote a standard unit vector,

$$\begin{aligned} \frac{u*\phi(x+he_k) - u*\phi(x)}{h} &= h^{-1} < u_y, \phi(x+he_k-y) - \phi(x-y) > \\ &= < u_y, \int_0^1 \partial_k\phi(x+the_k-y)dt > . \end{aligned}$$

Now if $x \in K$, for some compact set $K \subset \mathbb{R}^n$, then for every sequence $h_i \to 0$, the associated sequence of functions $y \mapsto \int_0^1 \partial_k\phi(x+th_ie_k-y)dt$, together with all its derivatives, converge uniformly toward $\partial_k\phi(x-y)$ and its corresponding derivatives. Moreover they are all compactly supported with supports contained in some compact set K'. Therefore,

$$\lim_{h\to 0} \frac{u*\phi(x+he_k) - u*\phi(x)}{h} = u*\partial_k\phi(x).$$

and thus $u*\phi$ has continuous partial derivatives. We can continue in this manner and conclude that in fact $u*\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$.

3. Differentiation of distributions: For every distribution $u \in \mathcal{D}'(\Omega)$ we define

$$< \partial^\alpha u, \phi >= (-1)^{|\alpha|} < u, \partial^\alpha\phi > .$$

Again, it is easily verified that we have thus defined a continuous endomorphism of the space of distributions. Of course, the operations above were defined so as to extend the usual operations on smooth functions.

We can now define the action of a general linear partial differential operator on distributions. Indeed let,

$$P(x,\partial) = \sum_{|\alpha|\le m} a_\alpha \partial^\alpha, \qquad a_\alpha \in \mathcal{C}^\infty(\Omega),$$

be such an operator. Then,

$$< P(x,\partial)u, \phi >=< u, P(x,\partial)^\dagger \phi >,$$

where $P(x,\partial)^\dagger$ is the *formal adjoint operator*,

$$P(x,\partial)^\dagger v = \sum_{|\alpha|\le m} (-1)^{|\alpha|}\partial^\alpha(a_\alpha v).$$

Observe that if $u_j \in \mathcal{D}'(\Omega)$ converges weakly to $u \in \mathcal{D}'(\Omega)$ then $P(x,\partial)u_j$ converges weakly to $P(x,\partial)u$.

Exercise. Show that for all $u \in \mathcal{D}'(\Omega)$ there exists a sequence $u_j \in \mathcal{C}_0^\infty(\Omega)$ such that $u_j \to u$ as $j \to \infty$ in the sense of distributions(weak convergence). Thus $\mathcal{C}_0^\infty(\Omega)$ is dense in $\mathcal{D}'(\Omega)$, with respect to the weak topology of the latter.

2.7. Example of distributions on the real line.

1.) The simplest nontrivial distribution is the Dirac function $\delta_0 = \delta_0(x)$, defined by $< \delta_0(x), \phi >= \phi(0)$.
2.) Another simple example is the Heaviside function $H(x)$ equal to 1 for $x > 0$ and zero for $x \le 0$. Or, using the standard identification between locally integrable functions and distributions,

$$< H(x), \phi >= \int_0^\infty \phi(x)dx.$$

Observe that $H'(x) = \delta_0(x)$.
3.) A more elaborate example is $pv(\frac{1}{x})$, or simply $\frac{1}{x}$, called the principal value distribution,

$$< \frac{1}{x}, \phi >= \lim_{\epsilon\to 0}\big(\int_{-\infty}^{-\epsilon} \frac{1}{x}\phi(x)dx + \int_\epsilon^\infty \frac{1}{x}\phi(x)dx\big).$$

Observe that $\log|x|$ is locally integrable and thus a distribution by the standard identification. It is easy to check that $\frac{d}{dx}\log|x| = \mathrm{pv}(\frac{1}{x})$.

Exercise. Let, for $z \in \mathbb{C}$ with $0 < \arg(z) < \pi$, $\log z = \log|z| + i\arg(z)$. We can regard $x \to \log z = \log(x+iy)$ as a family of distributions depending on $y \in \mathbb{R}^+$. For $x \neq 0$ we have $\lim_{y\to 0^+}\log z = \log|x| + i\pi\big(1 - H(x)\big)$. Show that as $y \to 0$ in $\mathbb{R}^+$, $\partial_x \log z$ converges weakly to a distribution $\frac{1}{x+i0}$ and,

$$\frac{1}{x+i0} = x^{-1} - i\pi\delta_0(x).$$

We now define an important family of distributions χ_+^z, with $z \in \mathbb{C}$, by analytic continuation. For this we first recall the definition of the Gamma function, Definition 2.8. For $\mathrm{Re}(z) > 0$ we define

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt \tag{10}$$

as well as the Beta function,

$$B(a,b) = \int_0^1 s^{a-1}(1-s)^{b-1} ds \tag{11}$$

Clearly $\Gamma(a) = a\Gamma(a-1)$ and $\Gamma(0) = 1$. Thus $\Gamma(n) = n!$. Recall that the following identity holds:

$$B(a,b) = \frac{\Gamma(a)\cdot\Gamma(b)}{\Gamma(a+b)} \tag{12}$$

We also record for future applications,

$$\Gamma(a)\Gamma(1-a) = B(a, 1-a) = \frac{\pi}{\sin(\pi a)} \tag{13}$$

In particular $\Gamma(1/2) = \pi^{1/2}$.

Exercise. Prove formulas (12) and (13). For help see Hörmander, [**5**] section 3.4.

Definition 2.9. For $\mathrm{Re}(a) > 0$, we denote by $j_a(\lambda)$ the locally integrable function which is identically zero for $\lambda < 0$ and

$$j_a(\lambda) = \frac{1}{\Gamma(a)}\lambda^{a-1}, \qquad \lambda > 0. \tag{14}$$

The following proposition is well known,

Proposition 2.10. *For all a, b, $Re(a)$, $Re(b) > 0$,*

$$j_a * j_b = j_{a+b}$$

Proof : We have,

$$\begin{aligned}
j_a * j_b(\lambda) &= \frac{1}{\Gamma(a)}\frac{1}{\Gamma(b)}\int_0^\lambda \mu^{a-1}(\lambda-\mu)^{b-1} d\mu \\
&= \frac{1}{\Gamma(a)}\frac{1}{\Gamma(b)}\lambda^{a+b-1}\int_0^1 s^{a-1}(1-s)^{b-1} ds \\
&= \frac{B(a,b)}{\Gamma(a)\cdot\Gamma(b)}\lambda^{a+b-1} = \frac{1}{\Gamma(a+b)}\lambda^{a+b-1} = j_{a+b}(\lambda)
\end{aligned}$$

■

Proposition 2.11. *There exists a family of distribution j_a, defined for all $a \in \mathbb{C}$, which coincide with the functions j_a for $Re(a) > 0$, such that, $j_a * j_b = j_{a+b}$, $\frac{d}{d\lambda} j_a(\lambda) = j_{a-1}(\lambda)$ and $j_0 = \delta_0$, the Dirac delta function at the origin. Moreover for all positive integers m, $j_{-m}(x) = \partial_x^m \delta_0(x)$.*

Proof : The proof is based on the observation that $\frac{d}{d\lambda} j_a(\lambda) = j_{a-1}(\lambda)$. Thus, for a test function ϕ,

$$\int_{\mathbb{R}} j_{a-1}(\lambda)\phi(\lambda)d\lambda = -\int_{\mathbb{R}} j_a(\lambda)\phi'(\lambda)d\lambda$$

Based on this observation we define, for every $a \in \mathbb{C}$ such that $\mathrm{Re}(a) + m > 0$ as distribution

$$< j_a, \phi >= (-1)^m \int_0^\infty j_{a+m}(\lambda)\, \phi^{(m)}(\lambda)d\lambda$$

In particular,

$$< j_0, \phi >= -\int_0^\infty j_1(\lambda)\phi'(\lambda)d\lambda = -\int_0^\infty \phi'(\lambda)d\lambda = \phi(0)$$

Hence $j_0 = \delta_0$. It is also easy to see that $j_a * j_b = j_{a+b}$ for all $a, b \in \mathbb{C}$. ■

Remark: In applications one often sees the family of distributions $\chi_+^a = j_{a+1}$. Clearly $\chi_+^a * \chi_+^b = \chi_+^{a+b+1}$ and $\chi_+^{-1} = \delta_0$. Observe also that χ_+^a is homogeneous of degree a, i.e. , $\chi_+^a(t\lambda) = t^a \chi_+^a(\lambda)$, for any positive constant t. This clearly makes sense for $\mathrm{Re}(a) > -1$ when χ_+^a is a function. Can you also make sense of it for all $a \in \mathbb{C}$?

2.12. Support of a distribution. The support of a distribution can be easily derived as follows:

Definition 2.13. For $u \in \mathcal{D}'(\Omega)$, we define the complement of the support of u,

$$\Omega \backslash \mathrm{supp}(u) = \{x \in \Omega \,|\exists\, V_x \ni x \text{ open, such that } < u, \phi >= 0 \quad \forall \phi \in \mathcal{C}_0^\infty(V_x)\}.$$

Lemma 2.14. *If $u \in \mathcal{D}'(\Omega)$ and ϕ is a test function with $supp(\phi) \subset \Omega \backslash supp(u)$, then $< u, \phi >= u(\phi) = 0$.*

Proof : This follows easily by a partition of unity argument. The argument can be found in Friedlander, section 1.4. ■

Proposition 2.15. *A distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ has compact support $K \subset \mathbb{R}^n$ iff there exists $N \in \mathbb{N}$ such that ,$\forall \phi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ we have*

$$|u(\phi)| \leq C \sup_{x \in U} \sum_{|\alpha| \leq N} |\partial^\alpha \phi(x)|,$$

where U is an arbitrary open neighborhood of K.

Proof : This is seen by using a cutoff function which is identically 1 on the support of the distribution. ■

Remark: Note that if we endow $\mathcal{C}^\infty(\mathbb{R}^n)$ with the Frechet topology induced by the family of seminorms given by $\phi \to \sup_{K_i} |\partial^\alpha \phi|$, with $\alpha \in \mathbb{N}^n$ and K_i running over a countable collection of compact sets exhausting $\mathbb{R}^n$, then the space of compactly supported distributions can be identified with $\mathcal{C}^\infty(\mathbb{R}^n)^*$, i.e. the space dual to $\mathcal{C}^\infty(\mathbb{R}^n)$.

We have the following useful fact concerning the structure of distributions supported at one point.

PROPOSITION 2.16. *Let $u \in \mathcal{D}'(\mathbb{R}^n)$ and assume that supp$(u) \subset \{0\}$. Then we have $u = \sum_{|\alpha|\le N} a_\alpha \partial^\alpha(\delta_0)$, for some integer N, complex numbers a_α and δ_0 the Dirac measure in $\mathbb{R}^n$ supported at 0.*

Proof : See Friedlander, [**3**], Theorem 3.2.1 or Hörmander, [**5**], Theorem 2.3.4. ∎

In this context, it is important to observe that the convolution of two distributions cannot be defined in general, but only when certain conditions on the support of the distributions are satisfied. We note in particular the fact that if $u_1, u_2 \in \mathcal{D}'(\mathbb{R}^n)$ one of which is compactly supported, then the convolution $u_1 * u_2$ can be defined. Indeed, assuming u_2 to be compactly supported, we simply define,

$$< u_1 * u_2, \phi >=< u_1, u_2 * \phi >, \qquad \forall \phi \in \mathcal{C}_0^\infty(\mathbb{R}^n).$$

Here, $\text{supp}(u_2 * \phi) \subset \{x + y : x \in \text{ supp}(u_2),\, y \in \text{ supp}(\phi)\}$, hence a compact set. This definition extends the classical convolution for functions.

2.17. Pull back of distributions. Consider first the case of a $\mathcal{C}^\infty$ diffeomorphism $f : \Omega \to \Omega'$ and let u a distribution on Ω'. Then the pull-back f^*u is a distribution in Ω defined by,

$$< f^*u, \phi >=< u(y), g^*\phi(y)|\det Jg(y)| >, \qquad \phi \in \mathcal{C}_0^\infty(\Omega)$$

where $g = f^{-1}$ and $g^*\phi(y) = \phi(g(y))$ and $Jg(y)$ is the jacobian of the map $y \to g(y)$. It is easy to see that this definition is meaningful and that it coincides with the standard change of variable rule when u is a smooth function. Moreover the derivatives of f^*u can be computed by the standard chain rule.

Next we consider the pull back corresponding to a function $f : \Omega \to \mathbb{R}$. This procedure allows us to use the definition of some distributions on the real line to obtain interesting distributions in $\mathbb{R}^n$.

DEFINITION 2.18. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth map with surjective differential everywhere. If $u \in \mathcal{D}'(\mathbb{R})$ we can define its pull-back $f^*(u)$ as follows:

Let $x \in \mathbb{R}^n$ such that[7] $\partial_{x_1} f(x) \neq 0$ on a neighborhood $U \ni x$. Hence the map $y \in U \to (f(y_1, y'), y') \in \mathbb{R}^n$, with $y' = (y_2, \dots, y_n)$, is a local diffeomorphism. Now we set, for every test function ϕ supported in U,

$$f^*(u)(\phi) = u_{y_1}\Big(\int \phi(f(y_1, y'), y')|\partial_{y_1} f(y_1, y')|^{-1} dy'\Big),$$

In this definition, u_{y_1} indicates that u operates on functions depending on the y_1-variable. Since we can proceed in this fashion for every point in $\mathbb{R}^n$, we can define the pullback of u via f globally by patching the local definitions together via a partition of unity.

[7] by surjectivity of the differential, we may always assume this.

Example: If f is as above, then we can explicitly obtain the pullback of the delta function δ_0, namely $f^*(\delta_0) = \frac{1}{|\nabla f|} d\sigma$. Here, $d\sigma$ denotes the canonical surface measure on the embedded sub-manifold $f^{-1}(0) \subset \mathbb{R}^n$ and ∇f denoted the gradient of f.

In connection with the above example, it is useful to observe that if f, g are two smooth functions on $\mathbb{R}^n$ with non-vanishing differential everywhere, then the following equality holds in the sense of distributions for all $a, b \in \mathbb{R}^n$:

$$\int \delta_0(f(a) - x)\delta_0(g(b) - x)dx = \delta_0(f(a) - g(b)).$$

Both sides are to be interpreted as distributions on $\mathbb{R}^n \times \mathbb{R}^n$. To check this, one completes the map $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n \to f(a) - g(b) \in \mathbf{R}$ to a local diffeomorphism, e.g. assuming that $\partial_{a_1} f(a) \neq 0, \partial_{b_1} g(b) \neq 0$, as follows: $(a, b) \to (f(a) - g(b), g(b), a', b')$, where a', b' denote $(a_2, \dots, a_n), (b_2, \dots, b_n)$. Using the above definition of the pullback of distributions and the fact that the determinant of the Jacobian of this map is the product of the Jacobians of the maps $a \to (f(a), a'), b \to (g(b), b')$, the claim easily follows.

Remark. One cannot define, in general, a meaningful, associative, product of distributions. Why not? Produce an example of three distributions on the real line whose product, if defined, could not be associative.

2.19. Fundamental solutions. Given a linear partial differential operator with constant coefficients $P(\partial) = \sum_{|\alpha| \leq k} a_\alpha \partial^\alpha$, with $a_\alpha \in \mathbb{C}$, we say that a distribution E is a fundamental solution if it verifies $P(\partial)E = \delta_0$. If this is the case then we can always find solution of the equation $P(\partial)u = f$, where $f \in \mathcal{D}'(\mathbb{R}^n)$ is a compactly supported distribution, by setting $u = E * f$. This follows easily from the following proposition together with the observation that $\delta_0 * u = u$ for any $u \in \mathcal{D}'(\mathbb{R}^n)$.

PROPOSITION 2.20. *Assume $u, v \in \mathcal{D}'(\mathbb{R}^n)$ one of which is compactly supported. Then,*

$$P(\partial)(u*v) = P(\partial)u*v = u*P(\partial)v.$$

In what follows we shall calculate the fundamental solution for some special important differential operators such as the Laplacean $\Delta = \sum_{i=1}^n \partial_i^2$ in $\mathbb{R}^n$, and the D'Alembertian $\Box = -\partial_t^2 + \Delta$ in $\mathbb{R}^{n+1}$. We also consider the Heat operator $\partial_t - \Delta$ and Schrödinger operator $i\partial_t + \Delta$.

1.) *Laplace Operator* Δ. The Laplace operator Δ is invariant under translations and rotations, that is the group of rigid motions. In polar coordinates $x = r\omega, r > 0, |\omega| = 1$, it takes the form,

$$\Delta = \partial_r^2 + \frac{n-1}{r}\partial_r + r^{-2}\Delta_{\mathbb{S}^{n-1}},$$

where $\Delta_{\mathbb{S}^{n-1}}$ is the Laplace -Beltrami operator on the unit sphere $\mathbb{S}^{n-1}$. We denote by dS_ω the area element of the hypersurface $\mathbb{S}^{n-1}$ and by ω_n the total area of the unit sphere.

Exercise. Recall that the Laplace-Beltrami operator on a Riemannian manifold with metric g is given, in local coordinates x^i by

$$\Delta_g \phi = \frac{1}{\sqrt{|g|}} \partial_i \big(g^{ij} \sqrt{|g|} \partial_j \phi \big).$$

Here g^{ij} are the components of the inverse metric g^{-1} relative to the coordinates x^i. The volume element dS_g on M is given, in local coordinates, by $dS_g = \sqrt{|g|} dx^1 dx^2 \ldots dx^n$. Observe that, on compact manifold M,

$$\int_M \Delta_g u \, v dS_g = \int_M u \, \Delta_g v dS_g.$$

Exercise 2. Calculate the Laplace-Beltrami operator for the unit sphere $\mathbb{S}^{n-1}$ and check the polar decomposition formula for Δ. For the particular case $n = 3$, relative to the coordinates $x^1 = r \cos \theta^1, x^2 = r \sin \theta^1 \cos \theta^2, x^3 = r \sin \theta^1 \sin \theta^2$, $\theta^1 \in [0, \pi), \theta^2 \in [0, 2\pi)$ show that,

$$\Delta_{\mathbb{S}^2} = \partial^2_{\theta^1} + \text{cotan}\theta^1 \, \partial_{\theta_1} + \frac{1}{\sin^2 \theta^1} \partial^2_{\theta^2}.$$

Moreover the area element dS_ω takes the form, $dS_\omega = r^2 \sin \theta^1 d\theta^1 d\theta^2$.

Proposition 2.21. *Define, for all $n \geq 3$, $K_n(x) = \big((2-n)\omega_n\big)^{-1} |x|^{2-n}$ while, for $n = 2$, $K_2(x) = (2\pi)^{-1} \log |x|$. Here w_n denotes the area of the unit sphere $\mathbb{S}^{n-1}$. Then, for all $n \geq 2$,*

$$\Delta K_n(x) = \delta_0.$$

Proof : Observe that $\Delta K_n(x) = 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$. Thus, in the whole space ΔK_n is supported at the origin and therefore can be expressed as a sum of derivatives of δ_0. By homogeneity considerations we can easily infer that $\Delta K_n(x) = c\delta_0$ for some constant c. Now, for a smooth function $\phi(x) = \phi(r\omega)$, in polar coordinates $r = |x|$, $\omega \in \mathbb{S}^{n-1}$ unit sphere in $\mathbb{R}^n$, that is $|\omega| = 1$, we have

$$\begin{aligned} \Delta \phi &= \big(\partial_r^2 + \frac{n-1}{r} \partial_r + r^{-2} \Delta_{\mathbb{S}^{n-1}} \big) \phi \\ &= r^{-(n-1)} \partial_r \big(r^{n-1} \partial_r \phi \big) + r^{-2} \Delta_{\mathbb{S}^{n-1}} \phi \end{aligned}$$

where $\Delta_{\mathbb{S}^{n-1}}$ is the Laplace -Beltrami operator on $\mathbb{S}^{n-1}$. Thus passing to polar coordinates $x = r\omega$, with $dx = r^{n-1} dr dS_\omega$, in the integral,

$$\begin{aligned} < \Delta K_n, \phi > &= < K_n, \Delta \phi > \\ &= \int_{|\omega|=1} \int_0^\infty K_n(r) \partial_r \big(r^{n-1} \partial_r \phi \big) dr dS_\omega + \int_{|\omega|=1} \int_0^\infty K_n(r) \Delta_{\mathbb{S}^{n-1}} \phi dr dS_\omega \\ &= \big((2-n)\omega_n \big)^{-1} \int_{|\omega|=1} \int_0^\infty r^{-n+2} \partial_r \big(r^{n-1} \partial_r \phi \big) dr dS_\omega \\ &= - \int_0^\infty r^{-n+1} \big(r^{n-1} \partial_r \phi \big) dr = - \int_0^\infty \partial_r \phi = \phi(0) \end{aligned}$$

we infer that, for $n \geq 3$, $\Delta K_n = \delta_0$ as desired. The case $n = 2$ can be treated in the same manner.

Remark : Observe that, up to a constant, the expression of $K_n(x)$ can be easily guessed by looking for spherically symmetric solutions $K = K(|x|)$. Indeed, the equation $\Delta K = 0$ reduces to the ODE, $\quad K''(r) + \frac{n-1}{r}K'(r) = 0.$

■

According to the general theory we can now solve the Poisson equation $\Delta u = f$, for any smooth compactly supported f, by the formula,

$$u(x) = \int_{\mathbb{R}^n} K_n(x-y)f(y)dy = \int_{\mathbb{R}^n} K_n(y)f(x-y)dy. \tag{15}$$

For $n \geq 3$ we observe that the solution given by (15) decays to zero as $|x| \to \infty$. Indeed, for large $|x|$ we can write (15) in the form

$$u(x) = c_n|x|^{-(n-1)} \int_{\mathbb{R}^n} (1 - \frac{|y|}{|x|})^{-(n-1)} f(y)dy \lesssim |x|^{-(n-1)},$$

due to the fact that f has compact support. We claim that the equation $\Delta u = f$ has a unique solutions $u(x)$ which decays at ∞ as $x \to \infty$ and therefore it must be represented by the integral formula (15). For $n = 2$, on the other hand, we only have $|u(x)| \lesssim \log |x|$. Observe however that

$$|\partial_i u(x)| \lesssim \int_{\mathbb{R}^2} |\partial_i K_2(x-y)||f(y)|dy \lesssim |x|^{-1}$$

since $|\partial K_2(x-y)| \lesssim |x-y|^{-1}$.

Proposition 2.22. *For any $f \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, $n \geq 3$ the equation $\Delta u = f$ has a unique smooth solution which vanishes at infinity, i.e. tends to zero as $|x| \to \infty$. The solution is represented by* (15)*. For $n = 2$ the same equation has a unique smooth solution $u(x)$ with $\lim_{|x|\to\infty} \frac{|u(x)|}{|x|} = 0$ and $|\partial u(x)| \to 0$ as $|x| \to \infty$. The solution is represented by* (15)*, up to an additive constant.*

Proof : By linearity it suffices to take $f = 0$. For $n \geq 3$ we have to show that $\Delta u = 0$, with u vanishing at infinity, implies that $u = 0$. The result is an easy consequence of Liouville's theorem which states that every bounded harmonic[8] function in $\mathbb{R}^n$ is constant. Liouville's theorem follows from the maximum principle for Δ according to which the extreme values of a harmonic function, i.e. a solution to $\Delta u = 0$, in a domain D must be attained at the boundary of D. We shall return to both Liouville's theorem and the maximum principle later. However you can try to prove directly the version of the maximum principle needed here. In the case $n = 2$ we can use the same argument to show that the derivatives of a solution $u(x)$ of $\Delta u = 0$, with the properties mentioned in the proposition, must vanish. ■

We shall now give an alternative, direct, proof of the fact that the function $u(x)$ defined by (15) is a solution of $\Delta u = f$. Indeed,

$$\Delta u(x) = \int_{\mathbb{R}^n} K_n(y)\Delta_y f(x-y)dy.$$

[8]Solutions to $\Delta u = 0$ are called harmonic.

We would like to integrate by parts and make use of the fact that $\Delta K_n(x) = 0$ on $\mathbb{R}^n \setminus \{0\}$. We cannot do it directly because the singularity at the origin. We circumvent this difficulty by the standard trick of decomposing the integral $I(x)$ on the right into a regular part $R_\epsilon(x) = \int_{\mathbb{R}^n \setminus B_e} K_n(y)\Delta_y f(x-y)dy$ and a singular part $S_\epsilon = \int_{B_\epsilon} K_n(y)\Delta_y f(x-y)dy$ where $\epsilon > 0$ is an arbitrary small number and B_ϵ is the closed ball of radius ϵ centered at the origin. For the singular part S_ϵ we have, for $n \ge 3$,

$$|S_\epsilon(x)| \lesssim \epsilon^2 \|\partial^2 f\|_{L^\infty}$$

and therefore converges to zero as $\epsilon \to 0$.

For the regular part,

$$R_\epsilon(x) = \int_{\mathbb{R}^n \setminus B_e} K_n(y)\Delta_y f(x-y)dy$$

we are allowed to integrate by parts. Doing it carefully by keeping track of the boundary terms on ∂B_ϵ and powers of ϵ we easliy infer that $|R_\epsilon(x) - f(x)|$ tends to zero as $\epsilon \to 0$, for all values of x.

2.) *D'Alembertian operator* $\Box$ *.* We shall next look of a fundamental solution for the wave operator,

$$\Box = -\partial_t^2 + \Delta = -\partial_t^2 + \partial_r^2 + \frac{n-1}{r}\partial_r + r^{-2}\Delta_{\mathbb{S}^{n-1}}$$

in $\mathbb{R}^{n+1}$. We look for solutions of the form[9] $\phi(t,x) = f(\rho)$ where $\rho = (t^2 - |x|^2)^{1/2}$, in the region $|x| < t$. By a simple calculation we find $f''(\rho) + \frac{n}{r}f'(\rho) = 0$ with solutions $f(\rho) = a\rho^{-\frac{n-1}{2}} + b$. Therefore a good candidate for a fundamental solution must have the form $E = (t^2 - |x|^2)^{-\frac{n-1}{2}}$ in the region $t > |x|$. To extend this definition to all space $\mathbb{R}^{n+1}$ and derive a distribution supported in the region $\{(t,x) : |x| \le t\}$ we are led to look at the pull back $f^*(\chi_+^{-\frac{n-1}{2}})$ of the one dimensional distribution $\chi_+^{-\frac{n-1}{2}}$, where f is the map $f(t,x) = t^2 - |x|^2$. For simplicity we write this distribution as $\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2)$. To make sure that we have a distribution supported in $|x| \le t$ we set,

$$E_+^{(n+1)}(t,x) = c_n H(t)\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2), \tag{16}$$

with $H(t)$ the Heavyside function supported on $t \ge 0$ and c_n a normalizing constant to be determined. In fact $c_n = -\frac{1}{2}\pi^{\frac{1-n}{2}}$.

Proposition 2.23. *The distribution* $E_+^{(n+1)}$ *is supported in* $|x| \le t$ *and verifies* $\Box E_+^{(n+1)} = \delta_0$.

Proof [for $n = 3$]: We first prove the proposition for the particular case of dimension $n = 3$. In that case we have to check that

$$E_+(t,x) \quad = \quad -\frac{1}{2}\pi^{-1}H(t)\delta_0(t^2 - |x|^2) = -\frac{1}{4\pi}r^{-1}\delta(t-r)$$

[9]In other words we look for solutions invariant under Lorentz transformations. We shall discuss later and in more detail the geometric significance of the wave operator and its symmetries.

with $r = |x|$. Thus, since $\Box\phi = -r^{-1}(\partial_t + \partial_r)(\partial_t - \partial_r)(r\phi) + \Delta_{\mathbb{S}^2}\phi$, we have with $\psi(t, r\omega) = (\partial_t - \partial_r)\big(r\phi(t, r\omega)\big)$,

$$\begin{aligned} < E_+, \Box\phi > &= \frac{1}{4\pi}\int_{\mathbb{S}^2}\int_0^\infty\int_0^\infty \delta(t-r)(\partial_t + \partial_r)\psi dt dr dS_\omega \\ &= \frac{1}{4\pi}\int_{\mathbb{S}^2} dS_\omega\big(\int_0^\infty \frac{d}{dr}\psi(r,r)dr\big) \\ &= -\psi(0,0) = \phi(0) \end{aligned}$$

Thus, $\Box E_+ = \delta_0$ as desired. ■

We shall now consider the general case. Let $E(t,x) = H(t)\chi_+^{-(n-1)/2}$. We write, for an arbitrary test function $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^{n+1})$,

$$< \Box E, \phi >= \int_0^\infty\int_{\mathbb{R}^n} E(t,x)\Box\phi dt dx = \lim_{\epsilon\to 0}\int_\epsilon^\infty\int_{\mathbb{R}^n} E(t,x)\Box\phi dt dx$$

Remark. Properly speaking the integral in the above identity does not make sense since E is not a locally integrable function. To be completely correct one has to write,

$$< \Box E, \phi >=< H(t)\chi_+^{-(n-1)/2}(t^2 - |x|^2), \Box\phi >= \lim_{\epsilon\to 0} < H(t-\epsilon)\chi_+^{-(n-1)/2}(t^2 - |x|^2), \Box\phi >$$

and then follow the same steps as below with the understanding that $\partial_t H(t-\epsilon) = \delta(t-\epsilon)$ and, for any test function ψ,

$$< \delta(t-\epsilon)\chi_+^{-(n-1)/2}(t^2 - |x|^2), \psi(t,x) >=< \chi_+^{-(n-1)/2}(\epsilon^2 - |x|^2), \psi(\epsilon, x) > .$$

We integrate by parts in the slab region $[\epsilon, \infty) \times \mathbb{R}^n$,

$$\begin{aligned} \int_\epsilon^\infty\int_{\mathbb{R}^n} E\cdot(-\partial_t^2 + \sum_{i=1}^n \partial_i^2)\phi &= +\int_\epsilon^\infty\int_{\mathbb{R}^n}(\partial_t E\partial_t\phi - \partial_i E\cdot\partial_i\phi) + \int_{\mathbb{R}^n} E\partial_t\phi(\epsilon,x) \\ &= \int_\epsilon^\infty\int_{\mathbb{R}^n}\Box E\cdot\phi - \int_{\mathbb{R}^n}\partial_t E\phi(\epsilon,x) + \int_{\mathbb{R}^n} E\partial_t\phi(\epsilon,x) \\ &= -\int_{\mathbb{R}^n}\partial_t E\phi(\epsilon,x) + \int_{\mathbb{R}^n} E\partial_t\phi(\epsilon,x) \end{aligned}$$

since, away from from the tip $t = |x| = 0$, we have $\Box\big(\chi_+^{-(n-1)/2}(t^2 - |x|^2)\big) = 0$. Why?

Now, making the change of variables $x = \epsilon y$ and using the homogeneity[10] of $\chi_+^{-(n-1)/2}$,

$$\begin{aligned} \int_{\mathbb{R}^n} E\partial_t\phi(\epsilon,x) &= \int_{\mathbb{R}^n}\chi_+^{-(n-1)/2}(\epsilon^2 - |x|^2)\partial_t\phi(\epsilon,x)dx \\ &= \int_{\mathbb{R}^n}\chi_+^{-(n-1)/2}\big(\epsilon^2(1-|y|^2)\big)\partial_t\phi(\epsilon,\epsilon y)\epsilon^n dy \\ &= \epsilon\int_{\mathbb{R}^n}\chi_+^{-(n-1)/2}(1-|y|^2)\partial_t\phi(\epsilon,\epsilon y)dy \to 0 \qquad \text{as} \quad \epsilon\to 0 \end{aligned}$$

[10]It is simple to check that, as distributions, $\chi_+^s(\lambda t) = \lambda^s\chi_+^s(t)$.

On the other hand,

$$\partial_t \chi_+^{(n-1)/2}(t^2 - |x|^2) = 2t\chi_+^{-(n+1)/2}(t^2 - |x|^2).$$

Hence,

$$\begin{aligned}
\int_{\mathbb{R}^n} \partial_t E \cdot \phi(\epsilon, x) &= 2\epsilon \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(\epsilon^2 - |x|^2)\phi(\epsilon, x)dx \\
&= 2\epsilon \int_{\mathbb{R}^n} \epsilon^{-(n+1)}\chi_+^{-(n+1)/2}(1 - |y|^2)\phi(\epsilon, \epsilon y)\epsilon^n dy \\
&= 2 \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(1 - |y|^2)\phi(\epsilon, \epsilon y)dy
\end{aligned}$$

Now observe that the distibution $\chi_+^{-(n+1)/2}(1-|y|^2)$ is supported in $|y| \leq 1$. Choose a test function $\psi(y)$ in $\mathbb{R}^n$ equal to 1 for $|y| \leq 2$ and supported in $|y| \leq 4$. Clearly,

$$\begin{aligned}
\lim_{\epsilon\to 0} \int_{\mathbb{R}^n} \partial_t E \cdot \phi(\epsilon, x) &= 2 \lim_{\epsilon\to 0} \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(1 - |y|^2)\phi(\epsilon, \epsilon y)\psi(y)dy \\
&= 2\phi(0) \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(1 - |y|^2)\psi(y)dy
\end{aligned}$$

Therefore we conclude that,

$$< \Box E, \phi >= -2J_n\phi(0)$$

where $J_n = \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(1 - |y|^2)\psi(y)dy$. To finish we only have to calculate J.

LEMMA 2.24. *For a function $\psi \in \mathcal{C}_0^\infty$ which is identically 1 in a neighborhood of the origin, we have*

$$J_n = \int_{\mathbb{R}^n} \chi_+^{-(n+1)/2}(1 - |y|^2)\psi(y)dy = \frac{1}{2}c_n^{-1}$$

where $c_n = 1/2\pi^{(1-n)/2}$.

Proof : We consider the cases $n = 2, n = 3$. For $n = 3$,

$$\begin{aligned}
J_3 &= \int_{\mathbb{S}^2}\int_0^\infty \chi_+^{-2}(1 - r^2)r^2\psi(r\omega)dr = -2^{-1}\int_{\mathbb{S}^2}\int_0^\infty \frac{d}{dr}\big(\chi_+^{-1}(1 - r^2)\big)r\psi(r\omega)dr \\
&= 2^{-1}\int_{\mathbb{S}^2}\int_0^\infty \delta(1 - r^2)\frac{d}{dr}(r\psi(r\omega))dr = 4^{-1}\int_{\mathbb{S}^2}\int_0^\infty \delta(1 - r)\frac{d}{dr}(r\psi(r\omega))dr \\
&= \pi
\end{aligned}$$

as desired.

For $n = 2$, since $\chi_+^{-1/2}(1 - s^2) = \frac{1}{\Gamma(1/2)}(1 - s^2)^{-1/2} = \pi^{-1/2}(1 - s^2)^{-1/2}$ and the derivatives of ψ vanish for $r \leq 2$,

$$\begin{aligned}
J_2 &= -2^{-1}\int_{|\omega|=1}\int_0^\infty \frac{d}{dr}\big(\chi^{-1/2}(1 - r^2)\big)\psi(r\omega)dr \\
&= 2^{-1} \cdot 2\pi \cdot \chi^{-1/2}(0)\psi(0) = \pi^{1/2}.
\end{aligned}$$

■

3.) *Heat Operator $\mathcal{H}$.* We consider the heat operator $\mathcal{H} = \partial_t - \Delta$ acting on functions defined on $\mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{n+1}$. It makes sense to look for spherically symmetric solutions to $\mathcal{H}u = 0$, that is functions $u(t,x) = u(t,|x|) = u(t,r)$. It is easy to find in this way the class of locally integrable solutions $E_c(t,x) = cH(t)\ t^{-\frac{n}{2}} e^{-|x|^2/4t}$, with $H(t)$ the heaviside function. Indeed $\mathcal{H}(E_c) = 0$ for all (t,x) with $t \neq 0$. We show below that, in the whole space, $\mathcal{H}(E_c)$ is proportional to δ_0 and that we can determine the constant $c = c_n = 2^{-n}\pi^{-\frac{n}{2}}$ such that the corresponding $E = E_c$ is a fundamental solution of $\mathcal{H}$, i.e. $\mathcal{H}(E) = \delta_0$.

Indeed, if $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^{n+1})$,

$$\begin{aligned}
< \mathcal{H}(E), \phi > &= < E, \mathcal{H}^t \phi > = -\int E(t,x)(\partial_t + \Delta)\phi(t,x)dxdt \\
&= -\lim_{\epsilon \to 0^+} \int_\epsilon^\infty \int_{\mathbb{R}^n} E(t,x)(\partial_t + \Delta)\phi(t,x)dxdt \\
&= \lim_{\epsilon \to 0^+} \int_\epsilon^\infty \int_{\mathbb{R}^n} (\partial_t + \Delta)E(t,x)\phi(t,x)dxdt + \lim_{\epsilon \to 0^+} \int_{\mathbb{R}^n} E(\epsilon,x)\phi(\epsilon,x)dx \\
&= \lim_{\epsilon \to 0^+} \int_{\mathbb{R}^n} E(\epsilon,x)\phi(x,\epsilon)dx = c_n \lim_{\epsilon \to 0^+} \epsilon^{-n/2} \int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon}\phi(x,\epsilon)dx
\end{aligned}$$

We now perform the change of variables $x = 2\epsilon^{1/2}y$,

$$\begin{aligned}
< \mathcal{H}(E), \phi > &= 2^n c_n \lim_{\epsilon \to 0^+} \int_{\mathbb{R}^n} \phi(\epsilon, 2\epsilon^{1/2}y)e^{-|y|^2}dy = 2^n c_n \phi(0,0) \int_{\mathbb{R}^n} e^{-|y|^2}dy \\
&= \phi(0,0)
\end{aligned}$$

Exercise. Check that $\int_{\mathbb{R}^n} e^{-|y|^2}dy = \pi^{n/2}$.

This proves that

$$E(t,x) = (4\pi t)^{-n/2} H(t)\ t^{-\frac{n}{2}} e^{-|x|^2/4t} \tag{17}$$

is a fundamental solution for $\mathcal{H}$.

4.) *Schrödinger equation $\mathcal{S}$.* The Schrödinger operator, $\mathcal{S} = i\partial_t + \Delta$ has a fundamental solution which looks, superficially, exactly like that of the Heat operator,

$$E(t,x) = (4\pi i t)^{-n/2} H(t)\ e^{i|x|^2/4t} \tag{18}$$

Yet, of course, the presence of i in the exponential factor $e^{-i|x|^2/4t}$ makes a world of difference.

Exercise Show that the locally integrable function E is indeed a fundamental solution for $\mathcal{S}$.

3. Fourier transform

3.1. Basic properties. Recall that if $f \in L^1(\mathbb{R}^n)$, then the Fourier transform $\mathcal{F}(f) = \hat{f}$ is defined as

$$\hat{f}(\xi) = \int f(x)e^{-ix\xi}dx \tag{19}$$

In case that $\hat{f} \in L^1(\mathbb{R}^n)$, we have the inversion formula

$$f(x) = (2\pi)^{-n} \int \hat{f}(\xi) e^{ix\xi} d\xi, \tag{20}$$

whose proof we shall indicate later. The inversion formula takes particularly concrete form in the case of the gaussian function $G(x) = e^{-|x|^2/2}$.

LEMMA 3.2. *The following calculation holds true for functions of one variable and $a, b \in \mathbb{R}, b > 0$,*

$$\int_{-\infty}^{\infty} e^{iax} e^{-bx^2} = \left(\frac{\pi}{b}\right)^{1/2} e^{-a^2/4b} \tag{21}$$

Thus in $\mathbb{R}^n$, for $t > 0$

$$\int_{\mathbb{R}^n} e^{ix\cdot y} e^{-ty^2} = \left(\frac{\pi}{t}\right)^{n/2} e^{-|x|^2/4t} \tag{22}$$

In particular

$$\mathcal{F}(G)(\xi) = (2\pi)^{n/2} G(\xi) \tag{23}$$

Proof : Make the change of variables in the complex domain, $z = b^{1/2}x - \frac{a}{2b^{1/2}} i$, and denote by Γ the contour $\operatorname{Im}(z) = -\frac{a}{2b^{1/2}}$,

$$\int_{-\infty}^{\infty} e^{iax} e^{-bx^2} dx = \frac{e^{-a^2/4b}}{b^{1/2}} \int_\Gamma e^{-z^2} dz = \frac{e^{-a^2/4b}}{b^{1/2}} \int_{-\infty}^{\infty} e^{-x^2} dx$$

by a standard contour deformation argument. Now recall[11] that the integral $J = \int_{-\infty}^{\infty} e^{-x^2} dx = \pi^{1/2}$ which proves (21). Formula (22) now follows immediately.

■

The Fourier transform is linear and verifies the following simple properties:

- Fourier transform takes translations in physical space $T_{x_0} f(x) = f(x - x_0)$ into modulations in frequency space $\mathcal{F}(T_{x_0} f)(\xi) = e^{-i\xi\cdot x_0} \hat{f}(\xi)$.
- Fourier transform takes modulations in physical space $M_{\xi_0} f(x) = e^{ix\cdot\xi_0} f(x)$ into translation in frequency space $\mathcal{F}(M_{\xi_0} f)(\xi) = \hat{f}(\xi - \xi_0)$.
- Fourier transform takes scaling in physical space $S_\lambda f(x) = f(\lambda x)$ into a dual scaling in Fourier space, $\mathcal{F}(S_\lambda f)(\xi) = \lambda^{-n} \hat{f}(\xi/\lambda)$. Observe that $S_\lambda(f)$ preserves size, i.e. $\|S_\lambda f\|_{L^\infty} = \|f\|_{L^\infty}$ while the dual scaling $S_\lambda^* f = \lambda^{-n} f(x/\lambda)$ preserves mass, that is $\|S_\lambda^* f\|_{L^1} = \|f\|_{L^1}$.
- Fourier transform takes conjugation in physical space into conjugation and reflection in frequency, i.e. $\mathcal{F}(\bar{f})(\xi) = \bar{\hat{f}}(-\xi)$.
- Fourier transform takes convolution in physical space into multiplication in frequency space, $\widehat{f*g} = \hat{f}\,\hat{g}$.

[11]For a quick proof of this observe that $J^2 = \int_{\mathbb{R}^2} e^{-|x|^2} dx = \pi$ by passing to polar coordinates.

- Fourier transform takes partial derivatives in physical space into multiplication in frequency space, $\mathcal{F}(\partial_{x_j} f)(\xi) = i\xi_j \hat{f}(\xi)$.
- Fourier transform takes multiplication by x_j in physical space into the partial derivative ∂_{ξ_j} in frequency space, $\mathcal{F}(x_j f)(\xi) = i\partial_{\xi_j}\hat{f}(\xi)$.
- We also have the simple self duality relation,

$$\int f(x)\hat{g}(x)dx = \int \hat{f}(x)g(x)dx.$$

Let $G_{\lambda,x_0,\xi_0}(x) = e^{ix\cdot\xi_0}G((x-x_0)/\sqrt{\lambda})$ be a translated, modulated, rescaled Gaussian. Then,

$$\begin{aligned}\mathcal{F}(G_{\lambda,x_0,\xi_0})(\xi) &= \lambda^{n/2}e^{-i(\xi-\xi_0)\cdot x_0}\int e^{-i\sqrt{\lambda}y\cdot(\xi-\xi_0)}G(y)dy \\ &= (\pi\lambda)^{n/2}G\big(\sqrt{\lambda}(\xi-\xi_0)\big)\end{aligned}$$

We can interpret this result as saying that G_{λ,x_0,ξ_0} is localized at spatial position x_0, with spatial spread $\Delta x \approx \sqrt{\lambda}$, and at frequency position ξ_0 with frequency spread $\Delta\xi = 1/\sqrt{\lambda}$. Observe that $\Delta x \cdot \Delta\xi \approx 1$, corresponding to the uncertainty principle.

Proposition 3.3 (Riemann Lebesgue). *Given an arbitrary $f \in L^1(\mathbb{R}^n)$ we have, $\|\hat{f}\|_{L^\infty} \lesssim \|f\|_{L^1}$. Moreover, $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$.*

Proof : Only the last statement requires an argument. Observe that if $f \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, then we can use integration by parts to conclude that $\hat{f}$ decays rapidly. Indeed for any multi-index α, $|\alpha| = k \in \mathbb{N}$,

$$\begin{aligned}\xi^\alpha\hat{f}(\xi) &= i^k\int \partial_x^\alpha e^{-ix\xi}f(x)dx = (-i)^k\int e^{-ix\xi}\partial_x^\alpha f(x)dx \\ |\xi^\alpha\hat{f}(\xi)| &\lesssim \int|\partial_x^\alpha f(x)dx| \le C_\alpha\end{aligned}$$

for some constant C_α. Thus, $|\hat{f}(\xi)| \lesssim (1+|\xi|)^{-k}$ which proves the statement in this case. For general $f \in L^1(\mathbb{R}^n)$, given $\epsilon > 0$, we can choose $g \in \mathcal{C}_0^\infty$ such that $||f-g||_{L^1} \le \frac{\epsilon}{2}$. From the preceding, we know that $|\hat{g}(\xi)| \le \frac{\epsilon}{2}$ if $|\xi| > M = M_\epsilon$ sufficiently large and therefore,

$$\sup_{|\xi|>M}|\hat{f}(\xi)| \le \|f-g\|_{L^1(\mathbb{R}^n)} + \sup_{|\xi|>M}|\hat{g}(\xi)| \le \epsilon$$

■

The Fourier transform converts constant coefficient linear partial differential operators into multiplication with polynomials, as immediate consequence of the relations $\widehat{\partial_{x_j}f}(\xi) = i\xi_j\hat{f}(\xi)$, $\widehat{x_j f}(\xi) = i\partial_{\xi_j}\hat{f}(\xi)$. We would like to extend Fourier transforms to distributions. However, since the space of test functions, i.e. $\mathcal{C}_0^\infty$, is not preserved by the Fourier transform, we need to restrict ourselves to a more limited class of distributions, namely the dual of a space of test functions that is preserved under the Fourier transform.

DEFINITION 3.4. A function $\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$ is said to be rapidly decreasing if for all multi indices α, β we have

$$\sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta \phi(x)| < \infty.$$

This so-called Schwarz space $\mathcal{S}(\mathbb{R}^n)$ of rapidly decreasing functions is endowed in the usual way with a natural Frechet topology. A sequence of functions ϕ_j converges to zero in this topology if, for all multi-indices α, β, $x^\alpha \partial^\beta \phi_j$ converges uniformly to zero. Note that $\mathcal{S}(\mathbb{R}^n)$ contains the compactly supported functions $\mathcal{C}_0^\infty(\mathbb{R}^n)$. Since this is dense in the $L^p(\mathbb{R}^n)$ spaces, for $1 \le p < \infty$, so is $\mathcal{S}(\mathbb{R}^n)$. It is also easy to check that $\mathcal{C}_0^\infty(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$.

We have the following important fact, which is the reason for considering the Schwarz space in our context:

PROPOSITION 3.5. *The Fourier transform is an isomorphism of $\mathcal{S}(\mathbb{R}^n)$ onto itself with inverse given by the inversion formula* (20)*. Moreover we have the Plancherel identity, for all $f, g \in \mathcal{S}(\mathbb{R}^n)$,*

$$(f,g)_{L^2} = \int_{\mathbb{R}^n} f(x)\bar{g}(x)dx = (2\pi)^{-n}(\hat{f},\hat{g})_{L^2} \tag{24}$$

In particular we have the Parseval identity $\|f\|_{L^2} = (2\pi)^{-n/2}\|\mathcal{F}(f)\|_{L^2}$.

Proof : Observe that $|\xi^\alpha \partial^\beta \hat{\phi}(\xi)| = |\widehat{x^\beta \partial^\alpha \phi}|$ and that $\partial^\alpha \phi(x)$ decays faster than $|x|^{-|\beta|-n-1}$. Thus we easily infer that $\mathcal{F}$ maps $\mathcal{S}(\mathcal{R}^n)$ into itself. Let $Rf(x) = f(-x)$ and define $T = R\mathcal{F}^2$. Observe that T commutes with partial derivatives ∂_j and multiplications by x_j. Indeed, for all $j = 1, \dots n$,

$$T(\partial_j f) = \partial_j(Tf), \quad T(x_j f) = x_j(Tf) \tag{25}$$

LEMMA 3.6. *An linear, continuous*[12]*, operator $T : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n)$ which verifies* (25) *must be of the form $T\phi = c\phi$ for some constant c.*

Proof : Exercise. ■

To determine the constants we only have to remark that, in view of lemma 3.2 we have $T(G) = \big((2\pi)^{n/2}\big)^2 G = (2\pi)^n G$. Hence the constant $c = (2\pi)^n$ which ends the proof of the inversion formula, and the proposition, for Schwartz functions. The constant could also be determined directly by observing that $G(x) = e^{-|x|^2/2}$ verifies the equation $(x_j + \partial_{x_j})G = 0$ and therefore also $(\xi_j + \partial_{\xi_j})\hat{G} = 0$. Hence, by uniqueness, $\hat{G}(\xi) = aG(\xi)$ for some constant a. Therefore, $a = \hat{G}(0) = (2\pi)^{n/2}$. The Plancherel and Parseval identities are immediate consequences of the inversion formula. ■

[12] That is $T(\phi_j) \to 0$ whenever $\phi_j \to 0$ in $\mathcal{S}(\mathbb{R}^n)$

Corollary 3.7. *The following properties hold for all functions in $\mathcal{S}$:.*

$$\begin{aligned}\int \hat{\phi}\psi dx &= \int \phi\hat{\psi} dx \\ \int \phi\bar{\psi} dx &= (2\pi)^{-n}\int \hat{\phi}\bar{\hat{\psi}} dx \\ \widehat{\phi * \psi} &= \hat{\phi}\hat{\psi} \\ \widehat{\phi\psi} &= (2\pi)^{-n}\hat{\phi}*\hat{\psi}\end{aligned}$$

As a corollary to the Parseval and Plancherel formulas we can extend our definition of Fourier to $L^2(\mathbb{R}^n)$ functions by a simple density argument. Indeed for any $u \in L^2$ we can choose a sequence of $\mathcal{S}(\mathbb{R}^n) \subset L^1$ functions u_j converging gto u in the L^2 norm. By Plancherel, $\|\mathcal{F}(u_j) - \mathcal{F}(u_k)\|_{L^2} \lesssim \|u_j - u_k\|_{L^2}$. Hence the sequence $\mathcal{F}(u_j)$ forms a Cauchy sequence in L^2 and therefore converges to a limit which we may call $\hat{u}$. Clearly this definition does not depend on the particular sequence. Moreover one can easily check that the Parseval identity extends to all L^2 functions. Thus we have proved,

Theorem 3.8. *The Fourier transform is an isometry of the Hilbert space $L^2(\mathbb{R}^n)$ into itself.*

We can extend the Fourier transform even further to a special class of distributions defined on $\mathbb{R}^n$.

Definition. We define a tempered distribution to be an element in the dual space of the Schwarz space. Note that the tempered distributions embed continuously into the space of ordinary distributions defined earlier. In analogy to the properties of ordinary distributions, for every tempered distribution u, there exists a natural number N and a constant C such that

$$| < u, \phi > | \le C \sum_{|\alpha|,\, |\beta| \le N} \sup |x^\alpha \partial^\beta \phi|, \phi \in \mathcal{S}(\mathbb{R}^n)$$

We can now easily define the Fourier transform of a tempered distribution, namely,

$$< \hat{u}, \phi > = < u, \hat{\phi} > .$$

One easily checks that this defines a tempered distribution $\hat{u}$ for every tempered u. Moreover all the properties of the Fourier transform, which have been verified for Schwartz functions in $\mathcal{S}(\mathbb{R}^n)$ can be easily extended to all tempered distributions. In particular, since all L^p spaces are included in $\mathcal{S}'(\mathbb{R}^n)$ we have a definition of Fourier transform for all such spaces. Observe that, in the case of L^1 this definition coincides with the definition given in (19).

The following simple, and very useful, formulas for the Fourier transform of the Dirac measure δ_0 make now sense:

$$\mathcal{F}(\delta_0) = 1, \qquad \mathcal{F}(1) = (2\pi)^n \delta_0 \tag{26}$$

Observe also that if we denote by $\text{sign}(x)$ the one dimensional tempered distribution given by the locally integrable function $\frac{x}{|x|}$ we have,

$$\widehat{\text{sign}}(\xi) = -2i\text{pv}(\xi) \tag{27}$$

Indeed $\text{sign}'(x) = 2\delta_0$. Hence, $i\xi\,\widehat{\text{sign}}(\xi) = 2$. Therefore, for any rapidly decreasing ϕ, we have

$$i\int \text{sign}(x)\widehat{x\phi}(x)dx = 2\hat{\phi}(0) = 2\int \phi(x)dx.$$

Also, observe that $\widehat{\text{sign}(x)}$ is an odd distribution in the sense that if ϕ is even, $\phi(x) = \phi(-x)$, then $< \widehat{\text{sign}}, \phi >= 0$. Now given a general test function ϕ, write $\phi = \frac{1}{2}(\phi(x) + \phi(-x)) + \frac{1}{2}(\phi(x) - \phi(-x)) = \phi_{ev} + \phi_{odd}$. Hence, from the preceding, we infer that

$$< \widehat{\text{sign}}, \phi >=< \widehat{\text{sign}}, x\big(\frac{1}{x}\phi_{odd}\big) >= -2i < pv(\frac{1}{x}), \phi >$$

as desired.

3.9. Applications to the basic PDE's.

3.10. Uncertainty principle and localization. On the real line let the operators X, D defined by,

$$Xf(t) = tf(t), \qquad Df(t) = -if'(t)$$

Observe that,

$$[D, X]f = DXf - XDf = -if$$

This lack of commutation is responsible for the following:

PROPOSITION 3.11 (Heisenberg uncertainty principle). *The following inequality holds,*

$$\|Xf\|_{L^2} \cdot \|Df\|_{L^2} \geq \frac{1}{2}\|f\|_{L^2}^2$$

Proof : Observe, using the commutator relation above,

$$0 \leq \|(aX + ibD)f\|_{L^2}^2 = a^2\|Xf\|_{L^2}^2 + b^2\|Df\|_{L^2}^2 - ab\|f\|_{L^2}^2$$

Now, pick $a = \|Df\|_{L^2}$ and $b = \|Xf\|_{L^2}$. ■

The uncertainty principle, which can informally described as $\Delta x \cdot \Delta\xi \geq 1/2$ places a limit on how accurately we can localize a function, or any other relevant object, simultaneously in both space and frequency. Let us investigate these localizations in more details.

1.) *Physical space localization.* If we want to localize a function f to a domain $D \subset \mathbb{R}^n$ we may simply multiply f by the characteristic function χ_D. The problem with this localization is that the resulting function $\chi_D f$ is not smooth even if f is. To correct for this we choose $\phi_D \in \mathcal{C}_0^\infty(D)$ in such a way that ϕ_D is not too different from χ_D. In the particular case when D is a ball $B(x_0, R)$ centered at x_0

we can choose ϕ_D to be 1 on the ball $B(x_0, R)$ and zero outside the ball $B(x_0, 2R)$. This leads to the following bounds for the derivatives of ϕ_D,

$$|\partial^\alpha \phi_D| \lesssim R^{-|\alpha|}.$$

In general given a domain D to which we can associate a length scale R (such as its diameter or distance from a fixed point in its interior), we can find a function $\phi_D \in \mathcal{C}_0^\infty(D)$ such that,

$$|\partial^\alpha \phi_D| \lesssim R^{-|\alpha|}, \tag{28}$$

for all multi-indices $\alpha \in \mathbb{N}^n$.

2.) *Frequency space localization.* Just like before we can localize a function to a domain $D \subset \mathbb{R}^n$ in frequency space by $\mathcal{F}^{-1}(\chi_D \hat{f})$. Once more, it often pays to use a smoother version of cut-off, thus we set,

$$\widehat{P_D f}(\xi) = \phi_D \hat{f}(\xi).$$

P_D is an example of a Fourier multiplier operator, that is an operator of the type:

$$\widehat{T_m f}(\xi) = m(\xi)\hat{f}(\xi). \tag{29}$$

with $m = m(\xi)$ a given function called the symbol of the operator. Clearly,

$$T_m f(x) = f * K(x) = \int f(x-y)K(y)dy \tag{30}$$

where K, the kernel of T, is the inverse Fourer transform of m,

$$K(x) = (2\pi)^{-n} \int e^{ix\cdot\xi} m(\xi) d\xi.$$

Clearly any linear differential operator $P(\partial)$ is a multiplier with symbol $P(i\xi)$.

To compare the action, in physical space, between rough and smooth cut-off operators it suffices to look at the corresponding kernels K. Let $I = [-1, 1] \subset \mathbb{R}$ and χ_I the rough cut-off (while ignoring the 2π constants). The corresponding kernel

$$K(x) = \int_{-1}^{1} e^{ix\cdot\xi} d\xi = 2\frac{sinx}{x}$$

decays very slowly as $|x| \to \infty$. Because of this the operator

$$\mathcal{F}^{-1}(\chi_I \hat{f})(x) = 2\int \frac{sin(x-y)}{(x-y)} f(y)dy$$

has very poor localization properties. Indeed, the operator spreads around to the whole $\mathbb{R}$ any function supported in some set $J \subset \mathbb{R}$. This situation corresponds to a perfect localization in frequency space and a very bad one in physical space. The exact opposite situation occurs when we do the rough cut-off localization $\chi_I f$ in physical space.

Now let us consider the frequency cut-off operator $P_I f = \mathcal{F}^{-1}(\phi_I \hat{f})$ whose kernel is

$$K(x) = \int_{\mathbb{R}} e^{ix\cdot\xi} \phi_I(\xi) d\xi.$$

Though we cannot explicitly calculate $K(x)$, as before, we can nevertheless get a good handle on its properties. Clearly, to start with, $K(x) \lesssim 1$. This bound is as good as we expect for $|x| \lesssim 1$. For $|x| \geq 1$ we can do much better by exploiting the rapid oscillations of the phase function $e^{ix\cdot\xi}$ and the smoothness of ϕ_I. Integrating by parts, for $|x| \geq 1$,

$$K(x) = \int_{\mathbb{R}} (\frac{1}{ix})^j (\frac{d}{d\xi})^j e^{ix\cdot\xi} \phi_I(\xi) d\xi = \int_{\mathbb{R}} (\frac{-1}{ix})^j e^{ix\cdot\xi} (\frac{d}{d\xi})^j \phi_I(\xi) d\xi$$

Thus, since all derivatives of ϕ_I are bounded, see (28), we have for all positive j,

$$|K(x)| \lesssim |x|^{-j}.$$

that is $K(x)$ is rapidly decreasing , unlike our previous case of the rough cut-off. Returning to $P_I f$ we can now prove the following:

LEMMA 3.12. *Let $I = [-1, 1]$, ϕ_I a smooth cut-off on I and $P_I f = \mathcal{F}^{-1}(\phi_I \hat{f})$. Then, if f is any L^2 function supported on a set $D \subset \mathbb{R}$,*

$$|P_I(f)(x)| \lesssim C_j \|f\|_{L^1} \big(1 + dist(x, D)\big)^{-j}$$

for all $j \in \mathbb{N}$.

Thus P_I spreads the support of any function f by a distance $O(1)$ plus a rapidly decreasing tail.

Exercise. Show that there exists no non-trivial function ϕ such that both ϕ and $\mathcal{F}(\phi)$ are compactly supported.

The above discussion can be easily extended to higher dimensions. In particular we can get a qualitative description of functions in $\mathbb{R}^n$ whose Fourier support is restricted to a ball $B_R = B(0, R)$ centered at the origin. Let ϕ_R be a smooth cut-off for B_R, that is $\sup_\xi |\partial_\xi^\alpha \phi_R(\xi)| \lesssim R^{-|\alpha|}$ for any multi-index α. Observe that we can in fact first pick ϕ a smooth cut-off for B_1 and define $\phi_R(\xi) = \phi(\xi/R)$ If f is a function whose support is restricted to B_R then $\hat{f} = \phi_R \hat{f}$. Hence,

$$f(x) = \int_{\mathbb{R}^n} f(y) K_R(x - y) dy \tag{31}$$

where $K(x) = \mathcal{F}^{-1}(\phi_R)$ i.e.,

$$\begin{aligned} K_R(x) &= \int_{\mathbb{R}^n} e^{ix\cdot\xi} \phi_R(\xi) d\xi = \int_{\mathbb{R}^n} (\frac{-1}{ix})^\alpha \partial_\xi^\alpha (e^{ix\cdot\xi}) \phi_R(\xi) d\xi \\ &= \int_{\mathbb{R}^n} (\frac{1}{ix})^\alpha e^{ix\cdot\xi} \partial_\xi^\alpha \phi_R(\xi) d\xi \end{aligned}$$

Thus, for any α, $|\alpha| = N$, denoting by $|B_R| = c_n R^n$ the volume of B_R,

$$|x|^N |K_R(x)| \quad \lesssim \quad \int_{\mathbb{R}^n} |\partial_\xi^\alpha \phi_R(\xi)| \lesssim R^{-N} |B_R| \lesssim R^{-N+n}$$

Hence, $|K_R(x)| \leq C_N R^n (|x| R)^{-N}$, for some constant C_N which may depend on N. On the other hand, for $|x| \lesssim R^{-1}$, $|K_R(x)| \lesssim R^n$. Hence, for every $N \in \mathbb{N}$,

$$|K_R(x)| \lesssim C_N R^n (1 + |x| R)^{-N}.$$

It is easy to check also that each derivative of K_R costs us a factor of R, that is,

$$|\partial_x^\alpha K_R(x)| \lesssim C_N R^{|\alpha|} R^n (1+|x|R)^{-N}, \quad \alpha \in \mathbb{N}^n \tag{32}$$

Now back to (31) we have

$$\begin{aligned} |\partial^\alpha f(x)| = |\int_{\mathbb{R}^n} f(y)\, \partial^\alpha K_R(x-y)dy| &\lesssim R^{|\alpha|+n} \int_{\mathbb{R}^n} |f(y)|(1+R|x-y|)^{-N} dy \\ &\lesssim R^{|\alpha|+n} \|f\|_{L^1} \end{aligned}$$

Also, by Cauchy -Scwartz with $\frac{1}{p} + \frac{1}{p'} = 1$,

$$\begin{aligned} |\partial^\alpha f(x)| &\lesssim \|f\|_{L^p} \|\partial^\alpha K_R\|_{L^{p'}} \lesssim R^{|\alpha|} R^n R^{-n/p'} \|f\|_{L^p} \\ &\lesssim R^{|\alpha|+n/p} \|f\|_{L^p} \end{aligned}$$

We have just proved the following version ($L^p - L^\infty$ version) of the very important Bernstein inequality,

PROPOSITION 3.13. *Assue that f is an L^p function which has its fourier transform supported in the ball $B_R = B(0,R)$. Then f has infinitely many derivatives bounded in L^∞ and we have,*

$$\|\partial^\alpha f\|_{L^\infty(\mathbb{R}^n)} \lesssim R^{n/p+|\alpha|} \|f\|_{L^p}.$$

Remark. Observe that the proposition could have been proved by reducing it to the particular case of $R = 1$. More precisely assume that the result is true for $R = 1$ and consider a function f whose Fourier transform is supported in B_R. Let $g(x) = R^{-n} f(R^{-1}x)$ and observe that, $\operatorname{supp} \hat{g}(\xi) = \operatorname{supp} \hat{f}(R\xi) \subset B_1$ and therefore we have, $\|\partial^\alpha g\|_{L^\infty(\mathbb{R}^n)} \lesssim \|g\|_{L^1} = R^{-n} R^{n/p} \|f\|_{L^p}$. Thus, $\|\partial^\alpha f\|_{L^\infty(\mathbb{R}^n)} \lesssim R^{n/p+|\alpha|} \|f\|_{L^p}$.

4. Basic interpolation theory

4.1. Introduction. Consider the Fourier transform as a linear operator $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$. According to the Plancherel identity we have $\|\mathcal{F}(f)\|_{L^2} \leq 2\pi^{n/2} \|f\|_{L^2}$. On the other hand, we have $\|\mathcal{F}(f)\|_{L^\infty} \leq \|f\|_{L^1}$. Can we get other bounds of the type $\|\mathcal{F}(f)\|_{L^q} \lesssim \|f\|_{L^p}$? It turns out that such estimates can be easily established by interpolating between the two estimates mentioned above. Complex interpolation allows us to conclude an L^p to L^q estimate for any values of p and q such that $p^{-1} + q^{-1} = 1$ and $q \geq 2$. This is known as the Young-Hausdorff inequality. Interpolation theory is particularly useful for linear multiplier operators of the form

$$\widehat{T_m f}(\xi) = m(\xi) \hat{f}(\xi)$$

with bounded multipler m. In view of Parseval's identity it is very easy to check the $L^2 - L^2$ estimate, $\|T_m f\|_{L^2} \lesssim \|f\|_{L^2}$. To obtain additional estimates we typically use the integral representation (30) $T_m f(x) = f * K(x) = \int f(x-y)K(y)dy$ where K is the inverse Fourier transform of m. If, for example, we can establish that $K \in L^1$ than we easily deduce that $\|T_m f\|_{L^1} \lesssim \|f\|_{L^1}$, since $\|f*K\|_{L^1} \leq \|f\|_{L^1} \cdot \|K\|_{L^1}$. We thus have both $L^1 - L^1$ and $L^2 - L^2$ estimates for T_m. and it is tempting to conclude we might have an $L^p - L^p$ estimate for all $1 \leq p \leq 2$. Such an estimate is indeed true and follows by interpolation. If on the other hand we can establish that $K \in L^\infty$ then $\|f*K\|_{L^\infty} \lesssim \|f\|_{L^1}$ and thus can prove, by interpolation, the same $L^p - L^q$ estimate as in the Hausdorff-Young inequality.

4.2. Review of L^p spaces. Given a measurable subset $\Omega \subset \mathbb{R}^n$ the space $L^p(\Omega)$, $1 \leq p < \infty$, consists in all measurables functions $f : \Omega \to \mathbb{C}$ with finite L^p norm,

$$\|f\|_{L^p} = \left(\int_\Omega |f(x)|^p \, dx\right)^{1/p} < \infty.$$

The space $L^\infty(\Omega)$ consists of all measurable functions, bounded almost everywhere, that is,

$$\|f\|_{L^\infty} = \text{ess sup}_{x\in\Omega} |f(x)| < \infty.$$

For all values of $1 \leq p \leq \infty$ the spaces $L^p(\Omega)$ are Banach spaces. The following is called Hölder's inequality

$$\|fg\|_{L^p} \leq \|f\|_{L^q} \|g\|_{L^r}, \tag{33}$$

whenever $1/p = 1/q + 1/r$. In particular, for $p = 1$,

$$\|fg\|_{L^1} \leq \|f\|_{L^q} \|g\|_{L^{q'}}$$

where q' verifying $\frac{1}{q'} = 1 - \frac{1}{q}$ is the exponent dual to q. For all $1 \leq q < \infty$ the space $L^{q'}(\Omega)$ is dual to $L^q(\Omega)$ while the dual of $L^\infty(\Omega)$ consists on the space of finite Borel masures on Ω, which includes $L^1(\Omega)$.

Exercise. Show that $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$ for all $1 \leq p < \infty$.

Given a measurable function f and a positive number α, denote by $\Lambda(f, \alpha)$ the *distribution function* of f defined by

$$\Lambda(f, \alpha) = |\{x \in \Omega : |f(x)| > \alpha\}|.$$

For $1 \leq p < \infty$ we have the obvious *Chebyschev*'s inequality

$$\Lambda(f, \alpha) \leq \alpha^{-p} \|f\|_{L^p}^p. \tag{34}$$

We can write the L^p norm of f in terms of its distribution function. Indeed, the integral $\int |f|^p$ is the measure of the set $\{(\beta, x) : 0 < \beta < |f(x)|^p\}$, hence

$$\int |f(x)|^p dx = \int_0^\infty \Lambda(|f|^p, \beta) d\beta = p \int_0^\infty \alpha^{p-1} \Lambda(f, \alpha) d\alpha, \tag{35}$$

where the last integral is obtained from the substitution $\beta = \alpha^p$.

A measurable function $f : \Omega \to \mathbb{C}$ is said to be simple if its range consists of a finite number of points in $\mathbb{C}$, that is $f = \sum_{i=1}^N a_i \chi_{A_i}$ for $a_i \in \mathbb{C}$ and $A_i \subset \Omega$ measurable. In this section we denote by $\mathcal{S}(\Omega)$ the set of all simple functions in Ω. Recall that $\mathcal{S}(\Omega)$ is dense in $L^p(\Omega)$ for all $1 \leq p \leq \infty$.

Exercise. Let $f(x,y)$ be a measurable function on $\Omega_1 \times \Omega_2 \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Prove the following version of the *Minkowski*'s inequality,

$$\left\| \int_{\Omega_2} f(x,y)dy \right\|_{L^p_x(\Omega_1)} \leq \int_{\Omega_2} \|f(x,y)\|_{L^p_x(\Omega_1)} dy,$$

for $1 \leq p \leq \infty$.

4.3. Three lines lemma. The method of analytic interpolation, for linear operators acting on L^p spaces, is based on a variant of the maximum modulus theorem for a strip-like domain called the three lines lemma. Consider the strip-like domain,

$$D = \{z \in \mathbb{C} : 0 < \mathrm{Re}(z) < 1\}.$$

We will denote by $\mathcal{A}_{BC}$ the set of bounded continuous functions on the closure of D which are analytic on D.

LEMMA 4.4 (Three lines lemma). *Let $f \in \mathcal{A}_{BC}$ such that*

$$|f(0+ib)| \leq M_0, \quad |f(1+ib)| \leq M_1,$$

for all $b \in \mathbb{R}$. Then for all $0 < a < 1$ and $b \in \mathbb{R}$,

$$|f(a+ib)| \leq M_0^{1-a} M_1^a.$$

Proof : We may assume that $M_0, M_1 > 0$. Let $\varepsilon > 0$ and define the analytic function

$$F_\varepsilon(z) = e^{-\varepsilon(1-z)z} \frac{f(z)}{M_0^{1-z} M_1^z}.$$

Because of the exponential factor, $F_\varepsilon(z)$ decays rapidly to 0 as $\mathrm{Im}(z) \to \pm\infty$, uniformly in D; it is then possible to find $L = L(\varepsilon) > 0$ such that $|F_\varepsilon(z)| \leq 1$ when $|\mathrm{Im}(z)| \geq L$. Since we also have $|F_\varepsilon(z)| \leq 1$ when $\mathrm{Re}(z) = 0$ or $\mathrm{Re}(z) = 1$, it follows, from the maximum modulus principle applied to the rectangle $D_L = D \cap \{|\mathrm{Im}(z)| \leq L\}$, that $|F_\varepsilon(z)| \leq 1$ for every $z \in D_L$ and therefore in D. This means

$$|f(z)| \leq \left| e^{\varepsilon(1-z)z} M_0^{1-z} M_1^z \right| = e^{\varepsilon \mathrm{Re}((1-z)z)} M_0^{1-\mathrm{Re}(z)} M_1^{\mathrm{Re}(z)}.$$

but f is independent of ε and when $\varepsilon \to 0$ we obtain the result.

■

4.5. Stein-Riesz-Thorin interpolation.

DEFINITION 4.6. We say that a family of linear operators T_z, indexed by $z \in D$, is an *analytic family of operators* if,

(1) T_z maps simple functions into measurable functions;
(2) For any pair of simple functions $f, g \in \mathcal{S}(\Omega)$, the map $z \mapsto \int g(x) T_z f(x) dx$ belongs to $\mathcal{A}_{BC}$.

REMARK 4.7. The reason for choosing simple functions as test functions in the previous definition is because they are easy to manipulate and they make a dense set in L^p for every $p \in [1, \infty]$.

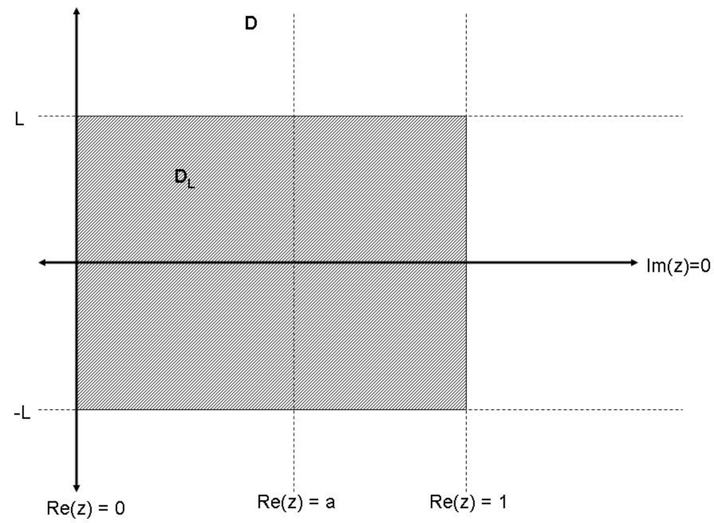

FIGURE 1. Three Lines Lemma

THEOREM 4.8. *Let T_z be an analytic family of operators and assume there are positive constants M_0, M_1 such that, for every $b \in \mathbf{R}$,*

$$\|T_{ib}f\|_{L^{q_0}} \leq M_0 \|f\|_{L^{p_0}}, \quad \|T_{1+ib}f\|_{L^{q_1}} \leq M_1 \|f\|_{L^{p_1}},$$

with $1 \leq q_0, p_0, q_1, p_1 \leq \infty$. Then, for $z = a + ib \in D$, T_z extends to a bounded operator from L^p to L^q and

$$\|T_z f\|_{L^q} \leq M_0^{1-a} M_1^a \|f\|_{L^p},$$

where

$$\frac{1}{p} = \frac{1-a}{p_0} + \frac{a}{p_1}, \quad \frac{1}{q} = \frac{1-a}{q_0} + \frac{a}{q_1}.$$

Proof : Adopting a bilinear formulation we have to prove that

$$\left| \int g(x) T_z f(x) dx \right| \leq M_0^{1-a} M_1^a, \tag{36}$$

for every pair of simple functions f, g with $\|f\|_{L^p} = \|g\|_{L^{q'}} = 1$. Fix such a pair f, g and consider the related (analytic) families of simple functions

$$f_z(x) = |f(x)|^{\frac{p}{p(z)}-1} f(x), \quad g_z(x) = |g(x)|^{\frac{q'}{q'(z)}-1} g(x),$$

with the exponents,

$$\frac{1}{p(z)} = \frac{1-z}{p_0} + \frac{z}{p_1}, \quad \frac{1}{q'(z)} = \frac{1-z}{q'_0} + \frac{z}{q'_1}.$$

We can easily check that

$$|f_{ib}| \leq |f|^{p/p_0}, \quad |f_{1+ib}| \leq |f|^{p/p_1}, \quad |g_{ib}| \leq |g|^{q'/q'_0}, \quad |g_{1+ib}| \leq |g|^{q'/q'_1}.$$

Here we use the convention that $1/\infty = 0$, and in particular if $p_0 = p_1 = \infty$ then $p = p(z) = \infty$ and $f_z \equiv f$, similarly $q'_0 = q'_1 = \infty$ then $q' = q'(z) = \infty$ and $g_z \equiv g$.

It is immediate to verify that $\|f_z\|_{L^{\mathrm{Re}(p(z))}} = \|f\|_{L^p} = 1$ and $\|g_z\|_{L^{\mathrm{Re}(q'(z))}} = \|g\|_{L^{q'}} = 1$.

Now consider the map defined on D,

$$h(z) = \int g_z(x) T_z f_z(x) dx.$$

It is not difficult to see from our construction and the linearity and analyticity properties of T_z, that $h \in \mathcal{A}_{BC}$. By hypothesis (and Cauchy-Schwarz) we have that $|h(ib)| \leq M_0$ and $|h(1+ib)| \leq M_1$ for every $b \in \mathbf{R}$. It follows from the three-lines lemma that $|h(z)| \leq M_0^{1-\mathrm{Re}(z)} M_1^{\mathrm{Re}(z)}$ and in particular (36). ∎

4.9. Young inequality. We often need to estimate integral operators of the form

$$Tf(x) = \int k(x,y) f(y) \mathrm{d}y, \tag{37}$$

The simplest result of this type is given by Young's theorem below.

THEOREM 4.10 (Young). *Let $k(x,y)$ be a measurable function and assume that for some $1 \leq r \leq \infty$ we have*

$$\sup_x \|k(x,\cdot)\|_{L^r} \lesssim 1, \quad \sup_y \|k(\cdot,y)\|_{L^r} \lesssim 1.$$

Then, for $1 \leq p \leq r'$ and

$$1 + \frac{1}{q} = \frac{1}{r} + \frac{1}{p}, \tag{38}$$

we have

$$\|Tf\|_{L^q} \leq \|f\|_{L^p}. \tag{39}$$

Proof : By Hölder inequality,

$$\|Tf\|_{L^\infty} \leq \|f\|_{L^{r'}}. \tag{40}$$

On the other hand the dual operator T^* has the same form as T,

$$T^* g(y) = \int \overline{k(x,y)} g(x) \mathrm{d}x,$$

and hence,

$$\|T^* g\|_{L^\infty} \leq \|g\|_{L^{r'}},$$

which by duality gives the other endpoint

$$\|Tf\|_{L^r} \leq \|f\|_{L^1}. \tag{41}$$

Now, we can use Theorem 4.8, with $T_z \equiv T$, to interpolate between (40) and (41) and obtain (39). ∎

As an immediate consequence, when k is translation invariant, $k(x,y) = k(x-y)$, we obtain the well known estimate for convolutions:

$$\|k * f\|_{L^q} \leq \|k\|_{L^r} \|f\|_{L^p}, \tag{42}$$

whenever the exponents $1 \le p, q, r \le \infty$ satisfy (38).

Exercise. Prove, using complex interpolation, the Hausdorff-Young inequality for the Fourier transform $\mathcal{F}$,

$$\|\mathcal{F}(f)\|_{L^q} \lesssim \|f\|_{L^p}, \qquad \text{for all} \quad q \ge 2, \quad 1/q + 1/p = 1.$$

4.11. Marcinkiewicz interpolation. A slightly weaker condition than L^p integrability for a function f is the so called *weak-L^p* property.

DEFINITION 4.12. For $1 \le p < \infty$, we say that f belongs to weak-L^p if $\Lambda(f, \alpha) \lesssim \alpha^{-p}$, for every $\alpha > 0$. If $p = \infty$ we let weak-L^∞ coincide with L^∞.

By Chebyschev's inequality (34), any function in L^p is also in weak-L^p. The following is the simplest example of real interpolation. It applies to sublinear operators, that is,

$$|T(f+g)(x)| \lesssim |Tf(x)| + |Tg(x)|,$$

THEOREM 4.13. *Consider a sublinear operator T mapping measurable functions on X to measurable functions on Y. Assume that T maps $L^{p_i}(X)$ into weak-$L^{p_i}(Y)$, with bound*

$$\Lambda(Tf, \alpha) \lesssim \alpha^{-p_i} \|f\|_{L^{p_i}}^{p_i},$$

for $i = 1, 2$ and $1 \le p_1 < p_2 \le \infty$. Then, for any p, $p_1 < p < p_2$, T maps $L^p(X)$ into $L^p(Y)$, with the bound

$$\|Tf\|_{L^p} \lesssim \|f\|_{L^p},$$

.

Proof : Given $f \in L^p(X)$ and $\alpha > 0$ we write $f = f^\alpha + f_\alpha$, where $f^\alpha(x) = f(x)$ if $|f(x)| > \alpha$ and $f_\alpha(x) = f(x)$ if $|f(x)| \le \alpha$. In particular $f^\alpha \in L^{p_1}$ and $f_\alpha \in L^{p_2}$.

Consider first the case $p_2 < \infty$. By our assumptions on T we have

$$\Lambda(Tf, 2\alpha) \lesssim \Lambda(Tf^\alpha, \alpha) + \Lambda(Tf_\alpha, \alpha) \lesssim \alpha^{-p_1} \|f^\alpha\|_{L^{p_1}}^{p_1} + \alpha^{-p_2} \|f_\alpha\|_{L^{p_2}}^{p_2}. \tag{43}$$

Using formula (35) and Fubini's theorem, we infer that

$$\int |Tf(x)|^p \,\mathrm{d}x \lesssim \iint_{0<\alpha<|f(x)|} |f(x)|^{p_1} \alpha^{p-p_1-1} \mathrm{d}\alpha \mathrm{d}x + \iint_{|f(x)| \le \alpha} |f(x)|^{p_2} \alpha^{p-p_2-1} \mathrm{d}\alpha \mathrm{d}x.$$

But $\int_0^{|f(x)|} \alpha^{p-p_1-1} \mathrm{d}\alpha \simeq |f(x)|^{p-p_1}$, since $p - p_1 - 1 > -1$, and $\int_{|f(x)|}^\infty \alpha^{p-p_2-1} \mathrm{d}\alpha \simeq |f(x)|^{p-p_2}$, since $p - p_2 - 1 < -1$, and the conclusion follows.

In the case of $p_2 = \infty$ the proof is actually simpler. We only have to observe that $|Tf(x)| \gg \alpha$ implies $|Tf^\alpha(x)| \gg \alpha$, since $|Tf_\alpha(x)| \lesssim \|f_\alpha\|_{L^\infty} \le \alpha$. Hence we can replace (43) by

$$\Lambda(Tf, C\alpha) \lesssim \Lambda(Tf^\alpha, \alpha) \lesssim \alpha^{-p_1} \|f^\alpha\|_{L^{p_1}}^{p_1},$$

where C is some positive constant, and the proof proceeds as before. ∎

5. Maximal function, fractional integration and applications

5.1. Maximal Function. A function f which is in $L^p(\mathbb{R}^n)$, for some $1 \le p \le \infty$, may possess very bad regularity properties. Given $\alpha > 0$, the set of points x where $|f(x)| > \alpha$ may merely be any measurable set (with finite measure if $p < \infty$). It is often desirable to replace f with a positive function which has (almost) the same integrability properties of f but better local regularity. This is achieved by considering maximal averages of f.

DEFINITION 5.2. Given a measurable function on $\mathbb{R}^n$ we define its *maximal function* by

$$\mathcal{M}f(x) = \sup_{x \in B} \frac{1}{|B|} \int_B |f(y)| \mathrm{d}y.$$

Here the supremum is taken over all possible euclidean balls B containing x.

REMARK 5.3. It follows immediately from the definition that $\mathcal{M}f$ is lower semicontinuous. Indeed, for every $\alpha \ge 0$, the sets $E_\alpha = \{x \in \mathbb{R}^n : \mathcal{M}f(x) > \alpha\}$ are always open: if $x \in E_\alpha$ then there exists a ball B containing x such that

$$\frac{1}{|B|} \int_B |f(y)| \mathrm{d}y > \alpha, \tag{44}$$

and this also means that $\mathcal{M}f(y) > \alpha$ for every $y \in B$, hence $B \subset E_\alpha$.

By the triangle inequality we also see that $f \mapsto \mathcal{M}f$ is a *subadditive* operator,

$$\mathcal{M}(f+g)(x) \le \mathcal{M}f(x) + \mathcal{M}g(x). \tag{45}$$

The averaging process may improve local regularity, but, because of the supremum, it is not clear whether $\mathcal{M}f$ preserves the integrability properties of f. If f is essentially bounded, then $\mathcal{M}f$ is bounded and

$$\|\mathcal{M}f\|_{L^\infty} \le \|f\|_{L^\infty}. \tag{46}$$

But, if f is an integrable function, it doesn't follow that $\mathcal{M}f$ is integrable. Take for example $f = \chi_B \in L^1$, the characteristic function of a ball, then $\mathcal{M}f(x) \gtrsim (1+|x|)^{-n}$ which barely fails to be in L^1. Fortunately, the maximal function still retains most of the information about the integrability properties of f.

THEOREM 5.4. *If $f \in L^1$ then $\mathcal{M}f$ is weakly in L^1, in the sense that for $\alpha > 0$ we have*

$$|E_\alpha| = \Lambda(\mathcal{M}f(x), \alpha) \lesssim \frac{1}{\alpha} \|f\|_{L^1} \tag{47}$$

If $f \in L^p$ with $1 < p \le \infty$ then $\mathcal{M}f \in L^p$ and we have

$$\|\mathcal{M}f\|_{L^p} \lesssim \|f\|_{L^p}. \tag{48}$$

Proof : The second part of the statement follows from the first and the L^∞ boundedness of the maximal operator by Marcinkiewicz interpolation, Theorem 4.13. Hence, we only need to prove (47).

Let $f \in L^1$ and fix $\alpha > 0$. By the discussion in Remark 5.3 we can find a family of balls $\mathcal{B} = \{B\}$, such that $E_\alpha = \cup_{B\in\mathcal{B}} B$ and each ball B satisfies (44). If these balls were all disjoint then it would be easy to conclude, since in that case

$$|E_\alpha| \leq \sum_{B\in\mathcal{B}} |B| < \frac{1}{\alpha}\sum_B \int_B |f(y)|\mathrm{d}y \leq \frac{1}{\alpha}\int_{R^n} |f(y)|\mathrm{d}y.$$

In general these balls are not disjoint and we have to be more careful.

Let K be a compact subset of E_α, then it is possibile to select a finite subfamily $\mathcal{B}'$ of balls in $\mathcal{B}$ that cover K. (This is sometimes known as the Vitali Covering Lemma.) Using the covering lemma proved below, Lemma 5.5, we can select among the balls in $\mathcal{B}'$ another finite subfamily $\mathcal{B}''$ made of disjoint balls such that

$$|\cup_{B'\in\mathcal{B}'} B'| \lesssim \sum_{B''\in\mathcal{B}''} |B''|.$$

Then, proceeding as above, we find

$$|K| \lesssim \frac{1}{\alpha}\, \|f\|_{L^1}\,,$$

and taking the supremum over all possible compact sets K we finally obtain (47). ■

LEMMA 5.5. *Let $B_1, \ldots, B_N$ be a finite collection of balls in $\mathbb{R}^n$, then it is possible to select a subcollection $B_{j_1}, \ldots, B_{j_M}$, $M \leq N$, of disjoint balls such that*

$$\left|\cup_{j=1}^N B_j\right| \lesssim \sum_{k=1}^M |B_{j_k}|.$$

Proof : We can assume that the balls $B_j = B(x_j, r_j)$ are labeled so that the radii are in nonincreasing order, $r_1 \geq r_2 \geq \cdots \geq r_N$.

Take $j_1 = 1$, so that B_{j_1} is the ball with largest radius. Then by induction, define j_{k+1} to be the minimum index among those of the balls B_j which don't intersect with the previously chosen balls $B_{j_1}, \ldots, B_{j_k}$; if there are no such balls then stop at step k.

With this construction we have that each ball B_j intersects one of the chosen balls B_{j_k} with $r_j \leq r_{j_k}$, hence $B_j \subset B(x_{j_k}, 3r_{j_k})$. This implies that

$$\left|\cup_{j=1}^N B_j\right| \leq \left|\cup_{k=1}^M B(x_{j_k}, 3r_{j_k})\right| \leq 3^n \sum_{k=1}^M |B_{j_k}|.$$

■

5.6. Lebesgue differentiation theorem. If a function f is continuous then, clearly,

$$\lim_{r\to 0} \frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)\mathrm{d}y = f(x). \tag{49}$$

As an application of Theorem 5.4 we can show that this property continue to hold for locally integrable functions.

COROLLARY 5.7 (Lebesgue's differentiation theorem). *If* $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ *then* (49) *holds for almost every* x.

Proof : Since the statement is local we can assume that $f \in L^1$.

Let A_r be the averaging operator defined by $A_r f(x) = |B(x,r)|^{-1} \int_{B(x,r)} f(y)\mathrm{d}y$. The proof consist of two steps. First we prove that $A_r f \to f$ in L^1 as $r \to 0$, and then it will be enough to show that $\lim_{r\to 0} A_r f(x)$ exists almost everywhere.

For the first step, given $\varepsilon > 0$, using the density of C_0 in L^1, we can always find a compactly supported continuous function g which approximates f in L^1 and have $\|A_r f - A_r g\|_{L^1} \le \|f - g\|_{L^1} < \varepsilon$ uniformly in r. Then by the uniform continuity of g, we know that $A_r g \to g$ in L^1 as $r \to 0$, hence there exists an r_ε such that

$$\|A_r f - f\|_{L^1} \le \|A_r f - A_r g\|_{L^1} + \|A_r g - g\|_{L^1} + \|f - g\|_{L^1} \le 3\varepsilon,$$

for $r < r_\varepsilon$.

For the second step, we define the *oscillation* of an L^1 function f by

$$\Omega f(x) = \limsup_{r\to 0} A_r f(x) - \liminf_{r\to 0} A_r f(x).$$

The oscillation is a subadditive operator, $\Omega(f+g) \le \Omega f + \Omega g$ and is bounded by the maximal function operator, $\Omega f \le 2\mathcal{M}f$, moreover the oscillation of a continuous function vanishes. If g is a continuous function which appoximate f in L^1 then we have that

$$\Omega f \le \Omega(f-g) + \Omega g = \Omega(f-g) \le 2\mathcal{M}(f-g).$$

We can apply now the weak-L^1 property of the maximal function, and for any positive α we find that

$$|\{x : \Omega f(x) > \alpha\}| \le |\{x : \mathcal{M}(f-g)(x) > \alpha/2\}| \lesssim \frac{1}{\alpha} \|f-g\|_{L^1}.$$

Since $\|f-g\|_{L^1}$ can be arbitrarily small, we infer that set of points where the oscillation of f is positive is of measure zero. ∎

5.8. Fractional integration. Let T be an integral operator acting on functions defined over $\mathbb{R}^n$ with kernel k as in (37). If the only information that we have on $k(x,y)$ is a decay estimate of the type

$$|k(x,y)| \lesssim |x-y|^{-\gamma},$$

for some $\gamma > 0$, then Young's inequality, Theorem 4.10, does not allow us to recover a good control on Tf, since the function $|\,x\,|^{-\gamma}$ fails, barely, to be in $L^{n/\gamma}$. However, the convolution has smoothing properties that imply some positive results which are contained in the following important theorem, originally proved by Hardy and Littlewood for $n = 1$ and then extended by Sobolev to $n > 1$.

THEOREM 5.9 (Hardy-Littlewood-Sobolev inequality). *Let $0 < \gamma < n$ and $1 < p < q < \infty$ such that*

$$1 - \frac{\gamma}{n} = \frac{1}{p} - \frac{1}{q}, \tag{50}$$

then

$$\left\| |\cdot|^{-\gamma} * f \right\|_{L^q(\mathbb{R}^n)} \lesssim \|f\|_{L^p(\mathbb{R}^n)}. \tag{51}$$

Proof : We can split the convolution with the singular kernel into two parts:

$$I_\gamma f(x) = |\cdot|^{-\gamma} * f(x) = \int_{|y|\geq R} \frac{f(x-y)}{|y|^\gamma} \mathrm{d}y + \int_{|y|<R} \frac{f(x-y)}{|y|^\gamma} \mathrm{d}y,$$

where the radius R is a positive constant to be chosen later We estimate the first term simply by Hölder's inequality,

$$\left| \int_{|y|\geq R} \frac{f(x-y)}{|y|^\gamma} \mathrm{d}y \right| \leq \|f\|_{L^p} \left(\int_{|y|\geq R} |y|^{-\gamma p'} \mathrm{d}y \right)^{1/p'} \lesssim R^{\frac{n}{p'}-\gamma} \|f\|_{L^p},$$

where we need the integrability condition $\gamma p' > n$, which by (50) is equivalent to $q < \infty$.

For the second part we perform a dyadic decomposition around the singularity and get an estimate in terms of the maximal function,

$$\begin{aligned} \left| \int_{|y|<R} \frac{f(x-y)}{|y|^\gamma} \mathrm{d}y \right| &\leq \sum_{k=0}^\infty \int_{2^{-k-1} \leq \frac{|y|}{R} \leq 2^{-k}} \frac{|f(x-y)|}{|y|^\gamma} \mathrm{d}y \lesssim \\ &\lesssim \sum_{k=0}^\infty \frac{1}{(2^{-k}R)^\gamma} \int_{|y|\leq 2^{-k}R} |f(x-y)| \mathrm{d}y \lesssim \\ &\lesssim \sum_{k=0}^\infty (2^{-k}R)^{n-\gamma} \mathcal{M}f(x) \simeq R^{n-\gamma} \mathcal{M}f(x), \end{aligned}$$

where we need $\gamma < n$ for the convergence of the last geometric series.

At this point we have found that for every $x \in \mathbb{R}^n$ and every $R > 0$,

$$\left| |\cdot|^{-\gamma} * f(x) \right| \lesssim R^{\frac{n}{p'}-\gamma} \|f\|_{L^p} + R^{n-\gamma} \mathcal{M}f(x),$$

with constants independent of R and x. We optimize this inequality choosing, for each x, a radius $R = R(x)$ such that the two terms on the right hand side are equal,

$$R^{\frac{n}{p'}-\gamma} \|f\|_{L^p} = R^{n-\gamma} \mathcal{M}f(x),$$

i.e.,

$$R(x) = \left(\frac{\|f\|_{L^p}}{\mathcal{M}f(x)} \right)^{p/n},$$

and since $(n-\gamma)p/n = 1 - p/q$, we have

$$|I_\gamma f(x)| \lesssim \|f\|_{L^p}^{1-\frac{p}{q}} \mathcal{M}f(x)^{\frac{p}{q}}.$$

Then take the L^q norm on both sides,

$$\|I_\gamma f\|_{L^q} \lesssim \|f\|_{L^p}^{1-\frac{p}{q}} \|\mathcal{M}f\|_{L^p}^{\frac{p}{q}}.$$

If $p > 1$ we can conclude using the estimates for the maximal function (48). ■

Remark. The Hardy-Littlewood-Sobolev inequality has an equivalent bilinear formulation, which reads

$$\iint \frac{f(x)g(y)}{|x-y|^\gamma} \mathrm{d}x\mathrm{d}y \lesssim \|f\|_{L^{p_1}} \|g\|_{L^{p_2}},$$

for $0 < \gamma < n$ and $1 < p_1, p_2 < \infty$ such that

$$\frac{1}{p_1'} + \frac{1}{p_2'} = \frac{\gamma}{n}.$$

Remark. Using the Hardy-Littlewood-Sobolev inequality, we now show that it is possible to give a very short proof of the Sobolev inequality,

$$\|f\|_{L^q} \lesssim \|\partial f\|_{L^p},$$

for $n/q = n/p - 1$, in the non sharp regime $p > 1$. Assume $f \in C_0^\infty(\mathbb{R}^n)$. For every unit vector ω we have

$$f(x) = -\int_0^\infty \frac{d}{dr} f(x + \omega r) \mathrm{d}r,$$

hence, if we integrate over the unit sphere, recalling that the volume element in $\mathbb{R}^n$ in polar coordinates is $\mathrm{d}y = r^{n-1}\mathrm{d}r\mathrm{d}\sigma_\omega$, we find that

$$|f(x)| \lesssim \int \frac{|\partial f(y)|}{|x-y|^{n-1}} \mathrm{d}y = \left(|\cdot|^{1-n} * |\partial f|\right)(x).$$

We take the L^q norm and use (51) to get

$$\|f\|_{L^q} \lesssim \left\| |\cdot|^{1-n} * |\partial f| \right\|_{L^q} \lesssim \|\, \partial f \,\|_{L^p},$$

whenever $p > 1$ and

$$1 - \frac{n-1}{n} = \frac{1}{p} - \frac{1}{q}.$$

Exercise. Prove the Hilbert inequality,

$$\int_0^\infty \int_0^\infty \frac{f(x)g(y)}{x+y} dxdy \lesssim \|f\|_{L^p} \|g\|_{L^q}, \qquad 1/p + 1/q = 1, \quad p, q \neq 1$$

5.10. Sobolev Inequalities. In the previous section we have seen how to estimate the $L^q(\mathbb{R}^n)$ norm of a function in terms of an L^p norm, $1 - \frac{n-1}{n} = \frac{1}{p} - \frac{1}{q}$, $p > 1$, of the gradient of f. We shall prove now a stronger version of this.

THEOREM 5.11 (Galgliardo-Nirenberg-Sobolev). *The inequality*

$$\|f\|_{L^q(\mathbb{R}^n)} \lesssim \|\partial^m f\|_{L^p(\mathbb{R}^n)}, \quad f \in C_0^\infty(\mathbb{R}^n), \tag{52}$$

holds for

$$\frac{1}{q} = \frac{1}{p} - \frac{m}{n} > 0, \quad m \in \mathbb{N}, \quad (1 \le p < q < \infty). \tag{53}$$

While for $q = \infty$, we have

$$\|f\|_{L^\infty(\mathbb{R}^n)} \lesssim \sum_{k=0}^m \left\|\partial^k f\right\|_{L^p(\mathbb{R}^n)}, \quad f \in C_0^\infty(\mathbb{R}^n), \tag{54}$$

when $m > n/p$.

Remark. We don't need to remember the precise condition (53); it can be deduced by a simple dimensional analysis. Since the estimate is homogeneous, it has to be invariant under dilations, and (53) simply says that both sides in (52) have the same scaling.

Remark. The following non-sharp version of estimate (52) also holds for all $1 \leq p < q < \infty$ and $1/p - m/n < 1/q$,

$$\|f\|_{L^q(\mathbb{R}^n)} \lesssim \sum_{|\alpha| \leq m} \|\partial^\alpha f\|_{L^p} \quad f \in C_0^\infty(\mathbb{R}^n), \tag{55}$$

Exercise. Show by an example that the inequality (54) fails to be true for $m = n/p$. Prove (55) for $m = 1$, using the results of theorem 5.11.

Exercise. Show by a scaling argument that if the inequality (55) holds true for $1/p = 1/q - m/n < 0$ then the homogeneous inequality (52) is also true.

Proof [Proof of (52)]: We obtain the cases with $m > 1$ by repeated iterations of the case $m = 1$. Hence, we can assume $m = 1$ and, by (53),

$$1 \leq p < n, \quad \frac{n}{n-1} \leq q = \frac{np}{n-p} < \infty.$$

Once we have the estimate for $p = 1$ and $q = n/(n-1)$, then we get the cases with $p > 1$ and $q > n/(n-1)$ by simply applying Hölder inequality. Indeed, let $q = \lambda n/(n-1)$, for some $\lambda > 1$, then

$$\|f\|_{L^q}^\lambda = \left\||f|^\lambda\right\|_{L^{\frac{n}{n-1}}} \lesssim \left\||f|^{\lambda-1}\partial f\right\|_{L^1} \leq \left\||f|^{\lambda-1}\right\|_{L^{p'}} \|\partial f\|_{L^p},$$

and we just have to check that

$$(\lambda - 1)p' = \frac{\frac{n-1}{n}q - 1}{1 - \frac{1}{n} - \frac{1}{q}} = q.$$

It only remains to prove the special case $m = 1$, $p = 1$, $q = n/(n-1)$. Following Nirenberg, [**15**], one can show that for $f \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ we have

$$\|f\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \lesssim \prod_{j=1}^n \|\partial_j f\|_{L^1(\mathbb{R}^n)}^{1/n}. \tag{56}$$

When $n = 1$, this comes easily from writing

$$f(x) = \int_{-\infty}^x f'(y)\mathrm{d}y.$$

When $n = 2$, we do the same with respect to to each variable and then multiply and integrate:

$$\begin{aligned}\iint |f(x_1,x_2)|^2 \mathrm{d}x_1\mathrm{d}x_2 &\leq \iint \int |\partial_1 f(y_1,x_2)|\mathrm{d}y_1 \int |\partial_2 f(x_1,y_2)|\mathrm{d}y_2\mathrm{d}x_1\mathrm{d}x_2 \\ &= \|\partial_1 f\|_{L^1} \|\partial_2 f\|_{L^1}.\end{aligned}$$

When $n \geq 3$ things become more tricky and, to separate the variables, we have to make a repeated use of Hölder inequality. Let just look at the case $n = 3$. To ease the notation set $f_j = \partial_j f$ and $\int \phi(x)\mathrm{d}x_j = \int_j \phi(\hat{x}_j)$. We start with

$$|f(x)|^{\frac{3}{2}} \leq \left(\int_1 |f_1(\cdot, x_2, x_3)|\right)^{\frac{1}{2}} \left(\int_2 |f_2(x_1, \cdot, x_3)|\right)^{\frac{1}{2}} \left(\int_3 |f_3(x_1, x_2, \cdot)|\right)^{\frac{1}{2}}.$$

Then integrate with respect to x_1. The first factor on the right hand side doesn't depend on x_1, while we use Hölder to separate the second from the third,

$$\int_1 |f(\cdot, x_2, x_3)|^{\frac{3}{2}} \leq \left(\int_1 |f_1(\cdot, x_2, x_3)|\right)^{\frac{1}{2}} \left(\int_{1,2} |f_2(\cdot, \cdot, x_3)|\right)^{\frac{1}{2}} \left(\int_{1,3} |f_3(\cdot, x_2, \cdot)|\right)^{\frac{1}{2}}.$$

Proceed similarly with the integration with respect to x_2,

$$\int_{1,2} |f(\cdot, \cdot, x_3)|^{\frac{3}{2}} \leq \left(\int_{1,2} |f_1(\cdot, \cdot, x_3)|\right)^{\frac{1}{2}} \left(\int_{1,2} |f_2(\cdot, \cdot, x_3)|\right)^{\frac{1}{2}} \left(\int_{1,2,3} |f_3(\cdot)|\right)^{\frac{1}{2}},$$

and finally do the same with x_3,

$$\int_{1,2,3} |f(\cdot)|^{\frac{3}{2}} \leq \left(\int_{1,2,3} |f_1(\cdot)|\right)^{\frac{1}{2}} \left(\int_{1,2,3} |f_2(\cdot)|\right)^{\frac{1}{2}} \left(\int_{1,2,3} |f_3(\cdot)|\right)^{\frac{1}{2}}.$$

When $n > 3$ the procedure is exacly the same. ■

Proof [Proof of (54)]: It clearly suffices to look at the case $m = 1$, since the cases $m > 1$ will follow from it applying (52). Assume thus $m = 1$ and $p > n$, we want to prove that

$$|f(0)| \lesssim \|f\|_{L^p} + \|Df\|_{L^p}.$$

Suppose first that f has support contained in the unit ball $B = \{|x| < 1\}$, then

$$f(0) = -\int_0^1 \frac{d}{dr} f(r\omega)\mathrm{d}r, \quad \omega \in \mathbb{S}^{n-1}. \tag{57}$$

Integrate with respect to ω and then apply Hölder,

$$|f(0)| \lesssim \int_B \frac{|\partial f(x)|}{|x|^{n-1}}\mathrm{d}x \lesssim \|\partial f\|_{L^p} \left(\int_B \frac{\mathrm{d}x}{|x|^{(n-1)p'}}\right)^{1/p'} \lesssim \|\partial f\|_{L^p}, \tag{58}$$

where the integrability condition needed here is $(n-1)p' < n$, which is precisely $p > n$.

In general, fix a cutoff function $\phi \in C_0^\infty$ with support in B and $\phi(0) = 1$, then in view of the above, $|f(0)| = |\phi(0)f(0)| \lesssim \|\partial(\phi f)\|_{L^p} \lesssim \|f\|_{L^p} + \|\partial f\|_{L^p}$. ■

5.12. Classical Sobolev spaces. The Sobolev inequalities of theorem (5.11) lead us to the introduction of Sobolev spaces.

Definition 5.13. Let Ω be an open subset of $\mathbb{R}^n$. Fix $1 \leq p \leq \infty$ and let $s \in \mathbb{N}$ be a non-negative integer. The space $W^{s,p}(\mathbb{R}^n)$ consists of all locally integrable, real (or complex) valued functions u on Ω such that for all multiindex α with $|\alpha| \leq s$

the weak[13] derivatives $\partial^\alpha u$ belong to $L^p(\Omega)$. These spaces come equiped with the norms,

$$\begin{aligned} \|u\|_{W^{s,p}(\Omega)} &= \Big(\sum_{|\alpha|\le s} \|\partial^\alpha u\|^p_{L^p(\Omega)}\Big)^{1/p}, \qquad \text{for} \quad 1 \le p < \infty \\ \|u\|_{W^{s,\infty}(\Omega)} &= \sum_{|\alpha|\le s} \|\partial^\alpha u\|_{L^\infty(\Omega)} \end{aligned}$$

We also denote by $W_0^{k,p}(\Omega)$ the closure of $\mathcal{C}_0^\infty(\Omega)$ in $W^{k,p}(\Omega)$.

In the particular case $p = 2$ we write $H^s(\Omega) = W^{s,2}(\Omega)$. Clearly $H^0(\Omega) = L^2(\Omega)$. We also write $H_0^s(\Omega) = W_0^{s,2}(\Omega)$.

In the particular case $p = \infty$ we work with the smaller space $C^s(\overline{\Omega}) \subset W^{s,\infty}(\Omega)$, the set of functions which are s times continuously differentiable and have bounded $\| \ \|_{W^{s,\infty}}$ norm.

Exercise. Show that for each $s \in \mathbb{N}$ and $1 \le p \le \infty$ the spaces $W^{s,p}(\Omega)$ are Banach spaces.

There is a lot more to be said about Sobolev spaces in domains $\Omega \subset \mathbb{R}^n$. We refer the reader to Evans, [**1**], chapter 5. For the time being we specialize to the case $\Omega = \mathbb{R}^n$.

Exercise. Show that the spaces $W^{k,p}(\mathbb{R}^n)$ and $W_0^{k,p}(\mathbb{R}^n)$ coincide. That means that $\mathcal{C}_0^\infty$ is dense in $W^{k,p}(\mathbb{R}^n)$.

The Sobolev inequalities proved in the previous subsection can be interpreted as embedding theorems. Indeed (52) and (55) can be interpreted as saying that the Sobolev space $W^{m,p}(\mathbb{R}^n)$ is included in the Lebesgue space $L^q(\mathbb{R}^n)$ as long as $\frac{1}{p} - \frac{m}{n} \le \frac{1}{q}$.

PROPOSITION 5.14. *The following inclusions are continuous*

$$W^{m,p}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n), \qquad \textit{if} \quad \frac{1}{p} - \frac{m}{n} \le \frac{1}{q}.$$

Moreover, for $q = \infty$, $W^{m,p}(\mathbb{R}^n)$ embeds into the space of bounded continuous functions on $\mathbb{R}^n$ provided that $m > n/p$.

Proof : Follows from theorem 5.11 and the density of $\mathcal{C}_0^\infty(\mathbb{R}^n)$ in $W^{m,p}(\mathbb{R}^n)$. ■

5.15. Hölder spaces. Together with Sobolev spaces Hölder spaces play a very important role in Analysis, especially in connection to elliptic equations. Before introducing these spaces we recall the definitions of the spaces $C^m(\overline{\Omega})$ of m times

[13] That is derivatives in the sense of distributions.

continuously differentiable functions $u : \Omega \to \mathbb{R}$ on an open domain Ω for which the $W^{s,\infty}$ norm is bounded,

$$\|u\|_{C^m(\overline{\Omega})} = \sum_{|\alpha| \leq m} \|\partial^\alpha u(x)\|_{L^\infty(\Omega)} < \infty.$$

Definition 5.16. Let Ω be an open domain in $\mathbb{R}^n$ We say that a function $u : \Omega \to \mathbb{R}$ is Hölder continuous with exponent $0 < \gamma \leq 1$ if,

$$[u]_{C^{0,\gamma}(\overline{\Omega})} = \sup_{x \neq y \in \Omega} \frac{|u(x) - u(y)|}{|x - y|^\gamma} < \infty. \tag{59}$$

The Hölder space $C^{k,\gamma}(\overline{\Omega})$ consists of all functions $u \in C^k(\overline{\Omega})$ for which the norm,

$$\|u\|_{C^{k,\gamma}(\overline{\Omega})} = \|u\|_{C^k(\overline{\Omega})} + \sum_{|\alpha| = k} [\partial^\alpha u]_{C^{0,\gamma}(\overline{\Omega})}. \tag{60}$$

is finite.

Exercise. The space $C^{k,\gamma}(\overline{\Omega})$ is a Banach space.

The following stronger version of the Sobolev embedding in L^∞ is important in elliptic theory.

Theorem 5.17 (Morrey's inequality). *Assume $n < p \leq \infty$. Then, for all $u \in C_0^\infty(\mathbb{R}^n)$,*

$$\|u\|_{C^{0,\gamma}(\mathbb{R}^n)} \lesssim \|u\|_{W^{1,p}(\mathbb{R}^n)} \tag{61}$$

provided that $\gamma = 1 - n/p$.

Proof : See Evans, *Partial Differential Equations*, section 5.6.2. [**1**] ■

5.18. Fractional H^s- Sobolev spaces. Consider the Sobolev space

$$H^s(\mathbb{R}^n) = \{u \in L^2 : \partial^\alpha u \in L^2, \quad \forall\, |\alpha| \leq s\}.$$

Proposition 5.19. *The Sobolev space $H^s(\mathbb{R}^n)$ coincides with the set of all distributions $u \in \mathcal{S}'(\mathbb{R}^N)$ for which $\hat{u}$ is locally integrable and,*

$$\|u\|^2_{H^s} = \int_{\mathbb{R}^N} (1 + |\xi|^2)^s |\hat{u}(\xi)|^2 < \infty \tag{62}$$

Proof : Follows easily from the Parseval identity. ■

Observe that the equivalent definition of proposition 5.19 makes sense not only for positive integers but for all real numbers s. We can thus talk about Sobolev spaces H^s for all real values of s. We shall also make use of the following homogeneous Sobolev norm, for all $s \geq 0$,

$$\|u\|^2_{\dot{H}^s} = \int_{\mathbb{R}^N} |\xi|^{2s} |\hat{u}(\xi)|^2 < \infty \tag{63}$$

Exercise. For $s \in (0,1)$ the space $H^s(\mathbb{R}^n)$ coincides with the space of locally integrable functions such that,

$$\left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(x) - u(x+y)|^2}{|y|^{n+2s}} dxdy + \|u\|^2_{L^2(\mathbb{R}^n)} \right)^{1/2} < \infty \tag{64}$$

Exercise. Prove that, for $s > n/2$ the Sobolev space $H^s(\mathbb{R}^n)$ embeds in the space of bounded continuous functions.

5.20. A Trace Theorem. The following theorem can be found, for example, in Renardy and Rogers, [**16**], Section 6.4.8.

THEOREM 5.21. *Let $s > 1/2$ be real. Then there exists a continuous linear map $T : H^s(\mathbb{R}^n) \to H^{s-1/2}(\mathbb{R}^{n-1})$ called the trace operator, with the property that for any smooth f, we have*

$$Tf(x_1, \ldots, x_n) = f(x_1, \ldots, x_{n-1}, 0) \tag{65}$$

Tf is the restriction of f to the hyperplane $x_n = 0$.

Proof Take f smooth and $g(x') = f(x', 0)$. Let $\tilde{f}$ be the Fourier transform of f in x_n only, and $\hat{f}$, $\hat{g}$ be the Fourier transforms of f and g in $\mathbb{R}^n$ and $\mathbb{R}^{n-1}$, respectively. I.e.

$$\tilde{f}(x', \xi_n) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x', x_n) e^{-ix_n\xi_n} dx_n$$

By applying Fourier inversion (with $x_n = 0$) and then the Fourier transform, we get

$$g(x') = f(x', 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(x', \xi_m)$$

$$\hat{g}(\xi') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(x', \xi_m)$$

We can then see, using our knowledge of fractional H^s spaces:

$$\begin{aligned}
\|g\|_{H^{s-1/2}} &\lesssim \int_{\mathbb{R}^{n-1}} |\hat{g}(\xi')|^2 (1 + |\xi'|^2)^{s-1/2} d\xi' \\
&\lesssim \int_{\mathbb{R}^{n-1}} \left| \int_{-\infty}^{\infty} \hat{f}(\xi) d\xi_n \right|^2 (1 + |\xi'|^2)^{s-1/2} d\xi' \\
&\lesssim \int_{\mathbb{R}^{n-1}} (1 + |\xi'|^2)^{s-1/2} \left(\int_{-\infty}^{\infty} |\hat{f}(\xi)|^2 (1 + |\xi|^2)^s d\xi_n \right) \left(\int_{-\infty}^{\infty} (1 + |\xi|^2)^{-s} d\xi_n \right) d\xi'
\end{aligned}$$

And since $s > 1/2$, we have

$$\begin{aligned}
\int_{-\infty}^{\infty} (1 + |\xi|^2)^s d\xi_n &= \int_{-\infty}^{\infty} (1 + |\xi'|^2 + |\xi_n|^2)^{-s} d\xi_n \\
&= (1 + |\xi'|^2)^{-s+1/2} \int_{-\infty}^{\infty} (1 + y^2)^{-s} d\xi_n
\end{aligned}$$

Plugging this into our above estimate for $\|g\|_{H^{s-1/2}}$ proves the result. ∎

Later on, we will see a strengthening of this result which uses Littlewood-Paley theory.

6. Littlewood-Paley theory

In its simplest manifestation Littlewood-Paley theory is a systematic and very useful method to understand various properties of functions f, defined on $\mathbb{R}^n$, by decomposing them in infinite dyadic sums $f = \sum_{k\in\mathbb{Z}} f_k$, with frequency localized components f_k, i.e. $\widehat{f_k}(\xi) = 0$ for all values of ξ outside the dyadic annulus $2^{k-1} \le |\xi| \le 2^{k+1}$. Such a decomposition can be easily achieved by choosing a test function $\chi(\xi)$ in Fourier space, supported in $\frac{1}{2} \le |\xi| \le 2$, and such that, for all $\xi \neq 0$,

$$\sum_{k\in\mathbb{Z}} \chi(2^{-k}\xi) = 1. \tag{66}$$

Indeed choose $\phi(\xi)$ to be a real radial bump function supported in $|\xi| \le 2$ which equals 1 on the ball $|\xi| \le 1$. Then the function $\chi(\xi) = \phi(\xi) - \phi(2\xi)$ verifies the desired properties.

We now define

$$\widehat{P_k f}(\xi) = \chi(\xi/2^k)\hat{f}(\xi) \tag{67}$$

or, in physical space,

$$P_k f = f_k = m_k * f \tag{68}$$

where $m_k(x) = 2^{nk} m(2^k x)$ and $m(x)$ the inverse Fourier transform of χ. Clearly, from (66)

$$f = \sum_{k\in\mathbb{Z}} P_k f \tag{69}$$

as desired. Observe that the Fourier transform of $P_k f$ is supported in the dyadic interval $2^{k-1} \le |\xi| \le 2^{k+1}$ and therefore,

$$P_{k'} P_k f = 0, \qquad \forall\, k, k' \in \mathbb{Z}, \quad |k - k'| > 2.$$

Therefore,

$$P_k f = \sum_{k'\in\mathbb{Z}} P_{k'}(P_k f) = \sum_{|k-k'|\le 1} P_{k'} P_k f$$

Thus, since P_{k-1}, P_k, P_{k+1} do not differ much between themselves we can write $P_k = \sum_{|k-k'|\le 1} P_{k'} P_k \approx P_k^2$. It is for this reason that the cut-off operators P_k are called, improperly, LP projections.

Denote $P_J = \sum_{k\in J} P_k$ for all intervals $J \subset \mathbb{Z}$. We write, in particular, $P_{\le k} = P_{(-\infty,k]}$ and $P_{<k} = P_{\le k-1}$. Clearly, $P_k = P_{\le k} - P_{<k}$.

The following properties of these LP projections lie at the heart of the classical LP theory:

THEOREM 6.1. *The LP projections verify the following properties:*

LP 1. *Almost Orthogonality. The operators P_k are selfadjoint and verify $P_{k_1}P_{k_2} = 0$ for all pairs of integers such that $|k_1 - k_2| \geq 2$. In particular,*

$$\|f\|_{L^2}^2 \approx \sum_k \|P_k f\|_{L^2}^2 \tag{70}$$

LP 2. *L^p-boundedness: For any $1 \leq p \leq \infty$, and any interval $J \subset \mathbb{Z}$,*

$$\|P_J f\|_{L^p} \lesssim \|f\|_{L^p} \tag{71}$$

LP 3. *Finite band property. We can write any partial derivative $\partial P_k f$ in the form $\partial P_k f = 2^k \tilde{P}_k f$ where $\tilde{P}_k$ is a cut-off operator*[14] *which verifies property LP2. In particular, for any $1 \leq p \leq \infty$*

$$\|\partial P_k f\|_{L^p} \lesssim 2^k \|f\|_{L^p} \tag{72}$$

$$2^k \|P_k f\|_{L^p} \lesssim \|\partial f\|_{L^p} \tag{73}$$

LP 4. *Bernstein inequalities. For any $1 \leq p \leq q \leq \infty$ we have the Bernstein inequalities,*

$$\|P_k f\|_{L^q} \lesssim 2^{kn(1/p - 1/q)} \|f\|_{L^p}, \qquad \forall\, k \in \mathbb{Z} \tag{74}$$

$$\|P_{\leq 0} f\|_{L^q} \lesssim \|f\|_{L^p}. \tag{75}$$

In particular,

$$\|P_k f\|_{L^\infty} \lesssim 2^{kn/p} \|f\|_{L^p}.$$

LP5. *Commutator estimates Consider the commutator*

$$[P_k\,,\, f] \cdot g = P_k(f \cdot g) - f \cdot P_k g$$

with $f, g \in \mathcal{C}_0^\infty(\mathbb{R}^n)$. We have,

$$\|\, [P_k\,,\, f] \cdot g\|_{L^p} \lesssim 2^{-k} \|\nabla f\|_{L^\infty} \|g\|_{L^p}.$$

LP6. *Square function inequalities. Let $\mathbf{S}f$ be the vector valued function $\mathbf{S}f = (P_k f)_{k \in \mathbb{Z}}$. The quantity*

$$Sf(x) = |\mathbf{S}f(x)| = \big(\sum_{k \in \mathbb{Z}} |P_k f(x)|^2\big)^{1/2} \tag{76}$$

is known as the Littlewood-Paley square function. For every $1 < p < \infty$ there exists constant(s), depending on p, such that for all $f \in \mathcal{C}_0^\infty$

$$\|f\|_{L^p} \lesssim \|Sf\|_{L^p} \lesssim \|f\|_{L^p} \tag{77}$$

[14] Associated with a slightly different test function $\tilde{\chi}$ which remains supported in $\frac{1}{2} \leq |\xi| \leq 2$, but may fail to satisfy (66).

Proof : Only the proof of LP6 is not straightforward and we postpone it until next section. The proof of LP1 is immediate. Indeed we only have to check (70). Clearly,

$$\begin{aligned}
\|f\|_{L^2}^2 &= \|\sum_k P_k f\|_{L^2}^2 = \sum_{|k-k'|\le 1} < P_k f, P_{k'} f >_{L^2} \\
&\le \sum_{|k-k'|\le 1} \|P_k f\|_{L^2} \|P_{k'} f\|_{L^2} \\
&\lesssim \sum_k \|P_k f\|_{L^2}^2
\end{aligned}$$

To show that $\sum_k \|P_k f\|_{L^2}^2 \lesssim \|f\|_{L^2}^2$ we only need to use Parseval's identity together with the definition of the projections P_k.

It suffices to prove LP2 for intervals of the form $J = (-\infty, k] \subset \mathbb{Z}$, that is to prove L^p boundedness for $P_{\le k}$. If $\chi(\xi) = \phi(\xi) - \phi(2\xi)$ then $\widehat{P_{\le k} f} = \phi(\xi/2^k)\hat{f}(\xi)$. Thus

$$P_{\le k} f = \bar{m}_k * f,$$

where $\bar{m}_k(x) = 2^{nk}\bar{m}(2^k x)$ and $\bar{m}(x)$ is the inverse Fourier transform of ϕ. Observe that $\|\bar{m}_k\|_{L^1} = \|\bar{m}\|_{L^1} \lesssim 1$. Thus, using the convolution inequality (42),

$$\|P_{\le k} f\|_{L^p} \le \|\bar{m}_k\|_{L^1} \|f\|_{L^p} \lesssim \|f\|_{L^p}$$

as desired.

To prove LP3 we write $\partial_i(P_k f) = 2^k(\partial_i m)_k * f$ where $(\partial_i m)_k(x) = 2^{nk}\partial_i m(2^k x)$. Clearly $\|(\partial_i m)_k\|_{L^1} = \|\partial_i m\|_{L^1} \lesssim 1$. Hence,

$$\|\partial_i(P_k f)\|_{L^p} \lesssim 2^k \|f\|_{L^p}$$

which establishes (72). To prove (73) we write $\hat{f}(\xi) = \sum_{j=1}^n \frac{\xi_j}{i|\xi|^2}\widehat{\partial_{x_j} f}(\xi)$. Hence,

$$2^k \widehat{P_k f}(\xi) = \sum_{j=1}^n 2^k \frac{\xi_j}{i|\xi|^2}\chi(\xi/2^k)\widehat{\partial_{x_j} f}(\xi) = \sum_{j=1}^n 2^k \psi_j(\xi/2^k)\widehat{\partial_{x_j} f}(\xi)$$

where $\psi_j(\xi) = \frac{\xi_j}{i|\xi|^2}\chi(\xi)$. Hence, in physical space,

$$2^k P_k f = \sum_{j=1}^n (\overline{{}^j m})_k * \partial_j f$$

with $(\overline{{}^j m})_k(x) = 2^{nk} \cdot \overline{{}^j m}(2^k x)$ and $\overline{{}^j m}$ the inverse Fourier transform of ψ_j. Thus, as before,

$$2^k \|P_k f\|_{L^p} \lesssim \sum_{j=1}^n \|\partial_j f\|_{L^p} = \|\partial f\|_{L^p}$$

as desired.

Property LP4 is an immediate consequence of the physical space representation (68) and the convolution inequality (42).

$$\|P_k f\|_{L^q} = \|m_k * f\|_{L^q} \lesssim \|m_k\|_{L^r} \|f\|_{L^p}$$

where $1+q^{-1}=r^{-1}+p^{-1}$. Now,

$$\|m_k\|_{L^r}=2^{nk}\big(\int_{\mathbb{R}^n}|m(2^kx)|^r dx\big)^{1/r}=2^{nk}2^{-nk/r}\|m\|_{L^r}\lesssim 2^{nk(1-1/r)}\lesssim 2^{nk(1/p-1/q)}$$

It only remains to prove LP5. In view of (68) we can write,

$$P_k(fg)(x)-f(x)P_kg(x) = \int_{\mathbb{R}^n} m_k(x-y)\big(f(y)-f(x)\big)g(y)dy$$

On the other hand,

$$\begin{aligned}|f(y)-f(x)| &\lesssim \big|\int_0^1 \frac{d}{ds}f(x+s(y-x))ds\big|\\ &\lesssim |x-y|\|\partial f\|_{L^\infty}\end{aligned}$$

Hence,

$$|P_k(fg)(x)-f(x)P_kg(x)| \lesssim 2^{-k}\|\partial f\|_{L^\infty}\int_{\mathbb{R}^n}|\bar{m}_k(x-y)||g(y)|dy$$

where $\bar{m}_k(x)=2^{nk}\bar{m}(2^kx)$ and $\bar{m}(x)=|x|m(x)$. Thus,

$$\|P_k(fg)-fP_kg\|_{L^p}\lesssim 2^{-k}\|\partial f\|_{L^\infty}\|g\|_{L^p}$$

We leave the proof of property LP6 for the next section. ■

Definition. We say that a Fourier multiplier operator $\tilde{P}_k$ is similar to a standard LP projection P_k if its symbol $\tilde{\chi}_k$ is a bump function adapted to the dyadic region $|\xi|\sim 2^k$. More precisely we can write $\tilde{\chi}_k(\xi)=\tilde{\chi}(\frac{\xi}{2^k})$ for some bump function $\tilde{\chi}$ supported in the region $c^{-1}2^k\lesssim|\xi|\le c2^k$ for some fixed $c>0$.

Remark. Observe that the inequality $\|P_kf\|_{L^p}\lesssim\|f\|_{L^p}$ holds for every other operator $\tilde{P}_k$ similar to P_k. The same holds true for the properties LP3, LP4 and LP5.

Remark: We have the following pointwise relation of the operator $\tilde{P}_k$ with the maximal function:

$$|\tilde{P}_{\le k}f|\lesssim \mathcal{M}f(x) \tag{78}$$

Indeed we have, as before,

$$\tilde{P}_{\le k}f=\tilde{m}_k * f,$$

where $\tilde{m}_k(x) = 2^{nk}\tilde{m}(2^k x)$ and $\tilde{m}(x) \in \mathcal{S}(\mathbb{R}^n)$. Therefore,

$$\begin{aligned}
|\tilde{P}_{\le k} f| &\lesssim 2^{nk}\int |f(y)|\tilde{m}\big(2^k(x-y)\big)|dy \lesssim 2^{nk}\int |f(y)|(1+2^k|x-y|)^{-n-1}dy \\
&\lesssim 2^{nk}\int_{B(x,2^{-k})} |f(y)|(1+2^k|x-y|)^{-n-1}dy \\
&+ 2^{nk}\sum_{j=0}^{\infty}\int_{2^j\le 2^k|x-y|\le 2^{j+1}} |f(y)|(1+2^k|x-y|)^{-n-1}dy \\
&\lesssim 2^{nk}\big(\int_{B(x,2^{-k})}|f(y)|dy + \sum_{j\ge 0} 2^{-(n+1)j}\int_{|x-y|\le 2^{j+1-k}}|f(y)|dy\big) \\
&\lesssim \mathcal{M}f(x) + \sum_{j>0} 2^{-(n+1)j}2^{nk}2^{n(j+1-k)}\frac{1}{|B(x,2^{-k+j+1})|}\int_{B(x,2^{-k+j+1})}|f(y)|dy \\
&\lesssim \mathcal{M}f(x) + 2^n\sum_{j>0}2^{-j}\mathcal{M}f(x) \lesssim \mathcal{M}f(x)
\end{aligned}$$

as desired.

Properties LP3-LP4 go a long way to explain why LP theory is such a useful tool for partial differential equations. The finite band property allows us to replace derivatives of the dyadic components f_k by multiplication with 2^k. The $L^p \to L^\infty$ Bernstein inequality is a dyadic remedy for the failure of the embedding of the Sobolev space $W^{\frac{n}{p},p}(\mathbb{R}^n)$ to $L^\infty(\mathbb{R}^n)$. Indeed, in view of the finite band property, the Bernstein inequality does actually imply the desired Sobolev inequality for each LP component f_k, the failure of the Sobolev inequality for f is due to the summation $f = \sum_k f_k$.

In what follows we give a few applications of LP -calculus.

1.) *Interpolation inequalities.* The following inequality holds true for arbitrary functions in $\mathcal{C}_0^\infty(\mathbb{R}^n)$ and any integers $0 \le i \le m$:

$$\|\partial^i f\|_{L^p} \lesssim \|f\|_{L^p}^{1-i/m}\|\partial^m f\|_{L^p}^{i/m} \tag{79}$$

To prove it we decompose $f = P_{\le k}f + P_{>k}f = f_{\le k} + f_{>k}$. Now, using LP2-LP4, for any fixed value of $k \in \mathbb{Z}$,

$$\begin{aligned}
\|\partial^i f\|_{L^p} &\le \|\partial^i f_{\le k}\|_{L^p} + \|\partial^i f_{>k}\|_{L^p} \\
&\le 2^{ki}\|f\|_{L^p} + 2^{k(i-m)}\|\partial^m f\|_{L^p}
\end{aligned}$$

Thus,

$$\|\partial^i f\|_{L^p} \le \lambda^i\|f\|_{L^p} + \lambda^{i-m}\|\partial^m f\|_{L^p}$$

for any $\lambda \in 2^{\mathbb{Z}}$. To finish the proof we would like to choose λ such that the two terms on the right hand side are equal to each other, i.e.,

$$\lambda_0 = \left(\frac{\|\partial^m f\|_{L^p}}{\|f\|_{L^p}}\right)^{1/m}$$

since we are restricted to $\lambda \in 2^{\mathbb{Z}}$ we choose the dyadic number $\lambda \in 2^{\mathbb{Z}}$ such that, $\lambda \le \lambda_0 \le 2\lambda$ Hence,

$$\|\partial^i f\|_{L^p} \quad \le \quad \lambda_0^i \|f\|_{L^p} + \big(\frac{2}{\lambda_0}\big)^{m-i} \|\partial^m f\|_{L^p} \lesssim \|f\|_{L^p}^{1-i/m} \|\partial^m f\|_{L^p}^{i/m}.$$

2.) *Non-sharp Sobolev inequalities.* We shall prove the following slightly improved version of the inequality (55), for functions $f \in C_0^\infty(\mathbb{R}^n)$ and exponents $1 \le p < q < \infty$ with $1/p - m/n < 1/q$,

$$\|f\|_{L^q(\mathbb{R}^n)} \lesssim \|f\|_{L^p} + \|\partial^m f\|_{L^p}$$

We decompose $f = P_{\le 0} f + \sum_{k \in \mathbb{N}} P_k f = f_{<0} + \sum_{k>0} f_k$. Thus, using LP4 and then LP3,

$$\begin{aligned}
\|f\|_{L^q} \quad &\le \quad \|f_{<0}\|_{L^q} + \sum_{k>0} \|f_k\|_{L^q} \lesssim \|f\|_{L^p} + \sum_{k>0} 2^{kn(1/p-1/q)} \|f\|_{L^p} \\
&\lesssim \quad \|f\|_{L^p} + \sum_{k>0} 2^{kn(m/n-\epsilon)} \|f\|_{L^p} \lesssim \|f\|_{L^p} + \sum_{k>0} 2^{-kn\epsilon} \|\partial^m f\|_{L^p} \\
&\lesssim \quad \|f\|_{L^p} + \|\partial^m f\|_{L^p}
\end{aligned}$$

3. *Spaces of functions.* The Littlewood-Paley theory can be used both to give alternative descriptions of Sobolev spaces and introduce new, more refined, spaces of functions. We first remark that, in view of the almost orthogonality property LP1,

$$\begin{aligned}
\|f\|_{L^2}^2 \quad &= \quad \|\sum_{k \in \mathbb{Z}} P_k f\|_{L^2}^2 \lesssim \sum_{k \in \mathbb{Z}} \|P_k f\|_{L^2}^2 \\
\sum_{k \in \mathbb{Z}} \|P_k f\|_{L^2}^2 \quad &\lesssim \quad \|f\|_{L^2}
\end{aligned}$$

We can thus give an LP description of the homogeneous Sobolev norms $\| \ \|_{\dot{H}^s(\mathbb{R}^n)}$

$$\|f\|_{\dot{H}^s}^2 \approx \sum_{k \in \mathbb{Z}} 2^{2ks} \|P_k f\|_{L^2}^2 \tag{80}$$

For $k \in \mathbb{Z}^+$, define operator $\Delta_k = P_k$ if $k > 0$, and $\Delta_0 = P_{\le 0}$. Also for the H^s norms,

$$\|f\|_{H^s}^2 \approx \sum_{k=0}^{\infty} 2^{2ks} \|\Delta_k f\|_{L^2}^2 \tag{81}$$

The Littlewood- Paley decompositions can be used to define new spaces of functions such as Besov spaces.

Definition: The Besov space $B^s_{p,q}(\mathbb{R}^n)$ is the closure of $\mathcal{C}_0^\infty(\mathbb{R}^n)$ relative to the norm:

$$\|f\|_{B^s_{p,q}} = \big(\sum_{k=0}^{\infty} 2^{ksq} \|\Delta_k f\|_{L^p}^q\big)^{\frac{1}{q}} \tag{82}$$

The corresponding homogeneous Besov norm is defined by,

$$\|f\|_{\dot{B}^s_{p,q}} = \big(\sum_{k \in \mathbb{Z}} 2^{sqk} \|P_k f\|_{L^p}^q\big)^{\frac{1}{q}}, \tag{83}$$

One similarly define Triebel space $F^s_{p,q}$ by reversing the L^p norm and l^q norm in (82). Thus, for example, the H^s norm is equivalent with the Besov norm $B^s_{2,2}$. Observe that, $H^s \subset B^s_{2,1}$. One reason why the larger space $B^s_{2,1}$ is useful is because of the following

$$\|f\|_{L^\infty} \lesssim \|f\|_{\dot{B}^{n/2}_{2,1}} \tag{84}$$

which follows trivially from the Bernstein inequality LP4. (84) will play a key role in the following section. Another reason to use the Besov norms $B^s_{2,1}$ will become transparent in the next section where we discuss product estimates.

6.2. Product estimates. The LP calculus is particularly useful for nonlinear estimates. Let f, g be two functions on $\mathbb{R}^n$. Consider,

$$P_k(fg) = \sum_{k',k''\in\mathbb{Z}} P_k(P_{k'} f P_{k''} g) \tag{85}$$

Now, since $P_{k'} f$ has Fourier support in the set $D' = 2^{k'-1} \le |\xi| \le 2^{k'+1}$ and $P_{k''} f$ has Fourier support in $D" = 2^{k''-1} \le |\xi| \le 2^{k''+1}$ it follows that $P_{k'} f P_{k''} g$ has Fourier support in $D' + D''$. We only get a nonzero contribution in the sum (85) if $D' + D''$ intersects $2^{k-1} \le |\xi| \le 2^{k+1}$. Therefore, writing $f_k = P_k f$ and $f_{<k} = P_{<k} f$, and $f_J = P_J f$ for any interval $J \subset \mathbb{Z}$ we derive,

LEMMA 6.3. *Given functions f, g we have the following decomposition:*

$$\begin{aligned}
P_k(f\cdot g) &= HH_k(f,g) + LL_k(f,g) + LH_k(f,g) + HL_k(f,g) \qquad (86)\\
HH_k(f,g) &= \sum_{k',k''>k+5,|k'-k''|\le 3} P_k\big(f_{k'}\cdot P_{k''} g\big)\\
LL_k(f,g) &= P_k\big(f_{[k-5,k+5]}\cdot g_{[k-5,k+5]}\big)\\
LH_k(f,g) &= P_k\big(f_{\le k-5}\cdot g_{[k-3,k+3]}\big)\\
HL_k(f,g) &= P_k\big(f_{[k-3,k+3]}\cdot g_{\le k-5}\big)
\end{aligned}$$

The term $HH_k(f,g)$ corresponds to high-high interactions. That is each term in the sum defining $HH_k(f,g)$ have frequence $\sim 2^m$ for some $2^m >> 2^k$. We shall write schematically,

$$HH_k(f,g) = P_k\big(\sum_{m>k} f_m\cdot g_m\big) \tag{87}$$

The term $LL_k(f,g)$ consists of a finite number of terms which can be typically ignored. Indeed they can be treated, in any estimates, like either a finite number of HH terms or a finite number of LH and HL terms. We write, schematically,

$$LL_k(f,g) = 0 \tag{88}$$

Finally the LH_k and HL_k terms consist of low high, respectively high-low, interactions. We shall write schematically,

$$\begin{aligned}
LH_k(f,g) &= P_k\big(f_{<k}\cdot g_k\big) \qquad (89)\\
HL_k(f,g) &= P_k\big(f_k\cdot g_{<k}\big) \qquad (90)
\end{aligned}$$

Remark. In the correct expression of LH_k given by (86) the terms of the form $f_{\le k-5}\cdot g_{k''}$, $k''\in[k-3,k+3]$, have Fourier supports in the dyadic region $\sim 2^k$. Thus P_k can be safely ignored and we can write,

$$LH_k(f,\,g\,)\sim f_{<k}\cdot g_k.$$

We have thus established, the famous *trichotomy formula*,

$$P_k(f\cdot g)=LH_k(f,g)+HL_k(f,g)+HH_k(f,g) \tag{91}$$

which is the basis of *paradifferential calculus*. In practice whenever we apply formula (91) we have to recall that formulas (88)–(90) are only appproximate; the correct definitions are given by (86). However in any estimates we can safely ignore the additional terms as they are estimated precisely in the same way as the terms we keep.

We shall now make use of the trichotomy formula to prove a product estimate.

THEOREM 6.4. *The following estimate holds true for all $s>0$.*

$$\|fg\|_{H^s}\lesssim\|f\|_{L^\infty}\|g\|_{H^s}+\|g\|_{L^\infty}\|f\|_{H^s} \tag{92}$$

Thus for all $s>n/2$,

$$\|fg\|_{H^s}\lesssim\|f\|_{H^s}\|g\|_{H^s} \tag{93}$$

Proof: Recall the characterization (81) of the H^s norm using the LP projections. Since $s>0$ we only need to look at the positive frequencies $P_k(fg)$ with $k>0$. We need to estimate the L^2 norm of the square function $\big(\sum_{k>0}|2^{sk}P_k(fg)|^2\big)^{1/2}$. Clearly,

$$\big(\sum_{k>0}|2^{sk}P_k(fg)|^2\big)^{1/2}\lesssim\big(\sum_{k>0}|2^{sk}LH_k|^2\big)^{1/2}+\big(\sum_{k>0}|2^{sk}HL_k|^2\big)^{1/2}+\big(\sum_{k>0}|2^{sk}HH_k|^2\big)^{1/2}$$

Now, using the pointwise bound (78)

$$\big(\sum_{k>0}|2^{sk}LH_k|^2\big)^{1/2}=\big(\sum_{k>0}|f_{<k}|^2|2^{sk}g_k|^2\big)^{1/2}\lesssim|\mathcal{M}f|\big(\sum_{k>0}|2^{sk}g_k|^2\big)^{1/2}$$

Hence,

$$\|\big(\sum_{k>0}2^{2sk}|LH_k|^2\big)^{1/2}\|_{L^2}\lesssim\|\mathcal{M}f\|_{L^\infty}\|\big(\sum_{k>0}2^{2sk}|g_k|^2\big)^{1/2}\|_{L^2}\lesssim\|f\|_{L^\infty}\|g\|_{H^s}$$

By symmetry we also have,

$$\|\big(\sum_{k>0}2^{2sk}|HL_k|^2\big)^{1/2}\|_{L^2}\lesssim\|\mathcal{M}g\|_{L^\infty}\|\big(\sum_{k>0}2^{2sk}|f_k|^2\big)^{1/2}\lesssim\|g\|_{L^\infty}\|f\|_{H^s}$$

It only remains to estimate the high-high term. Using the Minkowski inequality for l^2 sequences,

$$\begin{aligned}\big(\sum_{k>0}|2^{sk}HH_k|^2\big)^{1/2}&=\big(\sum_{k>0}|2^{sk}\sum_{a\ge0}P_k(f_{k+a}g_{k+a})|^2\big)^{1/2}\\&\le\sum_{a>0}\big(\sum_{k>0}|2^{sk}P_k(f_{k+a}g_{k+a})|^2\big)^{1/2}\end{aligned}$$

By Minkowski inequality in L^2,

$$\|\big(\sum_{k>0}|2^{sk}HH_k|^2\big)^{1/2}\|_{L^2}\lesssim\sum_{a>0}\|\big(\sum_{k>0}|2^{sk}P_k(f_{k+a}g_{k+a})|^2\big)^{1/2}\|_{L^2}$$

Now, using once more the pointwise inequality (78)

$$\begin{aligned}|P_k(f_{k+a}g_{k+a})| &\le \mathcal{M}(|f_{k+a}g_{k+a}|)\lesssim\mathcal{M}\big(\mathcal{M}f\cdot|g_{k+a}|\big)\\ \big(\sum_{k>0}|2^{sk}P_k(f_{k+a}g_{k+a})|^2\big)^{1/2} &\lesssim \mathcal{M}\Big(\mathcal{M}f\cdot\big(\sum_{k>0}|2^{sk}g_{k+a}|^2\big)^{1/2}\Big)\end{aligned}$$

Thus, using the L^2 boundedness of the maximal function,

$$\begin{aligned}\|\big(\sum_{k>0}|2^{sk}HH_k|^2\big)^{1/2}\|_{L^2} &\lesssim \sum_{a>0}\|\mathcal{M}\Big(\mathcal{M}f\cdot\big(\sum_{k>0}|2^{sk}g_{k+a}|^2\big)^{1/2}\Big)\|_{L^2}\\ &\lesssim \|f\|_{L^\infty}\sum_{a>0}\|\big(\sum_{k>0}|2^{sk}g_{k+a}|^2\big)^{1/2}\|_{L^2}\\ &\lesssim \|f\|_{L^\infty}\sum_{a>0}2^{-as}\|\big(\sum_{k>0}|2^{s(k+a)}g_{k+a}|^2\big)^{1/2}\|_{L^2}\\ &\lesssim \|f\|_{L^\infty}\|g\|_{H^s}\end{aligned}$$

Therefore,

$$\|\big(\sum_{k>0}|2^{sk}P_k(fg)|^2\big)^{1/2}\|_{L^2}\lesssim\|f\|_{L^\infty}\|g\|_{H^s}+\|g\|_{L^\infty}\|f\|_{H^s}$$

as desired. ■

Exercise. Give a rigorous proof of theorem 6.4.

The proof given above can be generalized, using LP6, to $W^{s,p}$ spaces. In what follows we give a somewhat simpler proof of theorem (6.4) which is very instructive. The proof[15] shows that it is sometimes better not to rely on the full decomposition (86) but rather using decompositions sparingly whenever needed. Indeed, we write,

$$\|fg\|^2_{\dot H^s}\;\lesssim\;\sum_k 2^{2ks}\|P_k(fg)\|^2_{L^2}\lesssim\sum_k 2^{2ks}\|P_k(f_{<k}g)\|^2_{L^2}+\sum_k 2^{2ks}\|P_k(f_{\ge k}g)\|^2_{L^2}$$

Now,

$$\begin{aligned}\sum_k 2^{2ks}\|P_k(f_{\ge k}g)\|^2_{L^2} &\lesssim \|g\|^2_{L^\infty}\sum_k 2^{2ks}\;\|f_{\ge k}\|^2_{L^2}\\ &\lesssim \|g\|^2_{L^\infty}\sum_k\sum_{k'\ge k}2^{2(k-k')s}\|2^{k's}f_{k'}\|^2_{L^2}\\ &= \|g\|^2_{L^\infty}\sum_{k'}\big(\sum_{k\le k'}2^{2(k-k')s}\big)\|2^{k's}f_{k'}\|^2_{L^2}\\ &\lesssim \|g\|^2_{L^\infty}\|f\|^2_{\dot H^s}\end{aligned}$$

To estimate $\sum_k 2^{2ks}\|P_k(f_{\le k}g)\|^2_{L^2}$ we shall decompose further, proceeding as in the decomposition (86). But first observe that the term $\sum_k 2^{2ks}\|P_k(f_{[k-3,k]}g)\|^2_{L^2}$ can

[15] I thank Igor Rodnianski for pointing the argument to me.

be treated precisely as $\sum_k 2^{2ks}\|P_k(f_{>k}g)\|_{L^2}^2$. Indeed we might as well estimated $\sum_k 2^{2ks}\|P_k(f_{>k-3}g)\|_{L^2}^2$ instead. Now,

$$\begin{aligned} P_k(f_{\le k-3}g) &= \sum_{k'} P_k(f_{\le k-3}g_{k'}) = \sum_{k'<k-2} P_k(f_{\le k-3}g_{k'}) + \sum_{k-2\le k'\le k+2} P_k(f_{\le k-3}g_{k'}) \\ &+ \sum_{k'>k+2} P_k(f_{\le k-3}g_{k'}) \end{aligned}$$

Observe that the first and last term are zero, therefore,

$$P_k(f_{\le k-3}g) = \sum_{k-2\le k'\le k+2} P_k(f_{\le k-3}g_{k'}) \approx P_k(f_{\le k-3}g_k).$$

Often, for simplicity, we simply write,

$$P_k(f_{<k}g) \approx f_{<k}\cdot g_k \tag{94}$$

Of course this formula is not quite right, but is morally right. Now,

$$\begin{aligned} \sum_k 2^{2ks}\|P_k(f_{<k}g)\|_{L^2}^2 &= \sum_k 2^{2ks}\|f_{<k}g_k\|_{L^2}^2 \\ &\lesssim \|f\|_{L^\infty}^2 \sum_k 2^{2ks}\|g_k\|_{L^2}^2 = \|f\|_{L^\infty}^2\|g\|_{\dot H^s}^2 \end{aligned}$$

as desired.

Remark. In view of (94) we have the following partial decomposition formula,

$$P_k(fg) = f_{<k}g_k + P_k(f_{\ge k}g) = LH_k(f,g) + P_k(f_{\ge k}g) \tag{95}$$

Contrast this with the full trichotomy decomposition (91).

Similar estimates, easier to prove, hold in Besov spaces. Indeed, for every $s>0$ we have,

$$\|fg\|_{H^{s,1}} \lesssim \|f\|_{L^\infty}\|g\|_{H^{s,1}} + \|g\|_{L^\infty}\|f\|_{H^{s,1}} \tag{96}$$

Exercise. Prove estimate (96).

7. Wente's Inequality

In this section we prove Wente's inequality as an application of Littlewood-Paley theory. In what follows given two functions f, g in $\mathbb{R}^2$ we consider the bilinear expression $(df\wedge dg)^* = \partial_x f\partial_y g - \partial_y f\partial_x g$, where $*$ denotes the trivial Hodge duality in $\mathbb{R}^2$. By abuse of language we drop the dual sign below and write simply $df\wedge dg$.

THEOREM 7.1. *On $\mathbb{R}^2$, assume $f,\ g\in H^1(\mathbb{R}^2)$, $\Delta u = (df\wedge dg)$. Then $u\in L^\infty$, in fact continuous.*

Remark. In fact $df\wedge dg$ If $\wedge$ is replaced by ordinary multiplication, then the best we can get is $df\cdot dg\in L^1$. This is obviously not enough to obtain that $u\in L^\infty$. It turns out however that $df\wedge dg$ has special structure which allows us to derive the desired estimate.

Proof : It is easy to see from finite band property that Δ is a isometric operator from $\dot{B}^s_{p,1}$ to $\dot{B}^{s-2}_{p,1}$. In fact we shall work with $p = 2$, In view of the Sobolev inequality (84), it suffices to show that $df \wedge dg \in \dot{B}^{-1}_{2,1}(\mathbb{R}^2)$. Using the trichotomy formula and the fact that the LP projections P_k commute with d we write,

$$\begin{aligned} I &= df \wedge dg = LH_k + HL_k + HH_k \\ LH_k &= dP_{<k}f \wedge dP_k g \\ HL_k &= dP_k \wedge dP_{<k} g \\ HH_k &= P_k(\sum_{m\geq k}(dP_m f \wedge dP_m g) \end{aligned}$$

By symmetry we only need to deal with LH and HH. The LH term is trivial to estimate, without using the special structure of the wedge product. Using the Bernstein inequality we write,

$$\begin{aligned} 2^{-k}\|LH_k\|_{L^2} &\lesssim 2^{-k}\sum_{l<k}\|dP_l f\|_{L^\infty}\|dP_k(g)\|_{L^2} \\ &\lesssim \sum_{l<k} 2^{l-k}\|DP_l f\|_{L^2}\|DP_k f\|_{L^2} \end{aligned}$$

The proof now follows with the following discrete version of the Young inequality.

Lemma 7.2. *Let $f(k) \in l^1(\mathbb{Z})$ and $g(k), h(k) \in l^2(\mathbb{Z})$. Then,*

$$\sum_{k,l} f(k-l)\mathbf{g}(l)h(k) \leq \|f\|_{l^1}\|g\|_{L^2}\|h\|_{l^2}.$$

Using the lemma, we derive,

$$\begin{aligned} \sum_k 2^{-k}\|LH_k\|_{L^2} &\lesssim \big(\sum_l \|DP_l f\|^2_{L^2}\big)^{1/2}(\sum_k \|DP_k f\|^2_{L^2})^{1/2} \\ &\lesssim \|Df\|_{L^2}\|Dg\|_{L^2} \end{aligned}$$

We now consider HH_k. It is here that we need to use the special structure of the wedge product. In fact we shall simply use the identity, $df \wedge dg = d(f \wedge dg)$. Thus,

$$\begin{aligned} HH_k &= \sum_{m\geq k} P_k(dP_m f \wedge dP_m g) \\ &= \sum_{m\geq k} dP_k(P_m f \wedge dP_m g) \end{aligned}$$

Thus, using the finite band property and Bernstein inequality,

$$\begin{aligned} \|HH_k\|_{L^2} &\lesssim 2^{2k}\|P_m f \wedge dP_m g\|_{L^1} \\ &\lesssim 2^{2k}\|P_m f\|_{L^2}\|DP_m g\|_{L^2} \\ &\lesssim 2^{2k-m}\|DP_m f\|_{L^2}\|DP_m g\|_{L^2} \end{aligned}$$

Therefore,

$$2^{-k}\|HH_k\|_{L^2} \lesssim 2^{k-m}\|DP_m f\|_{L^2}\|DP_m g\|_{L^2}$$

Thus, again, using the discrete Young inequality of the lemma above,

$$\sum_k 2^{-k}\|LH_k\|_{L^2} \lesssim \|Df\|_{L^2}\|Dg\|_{L^2}$$

as desired. ■

8. An LP Trace Theorem

In this section, we provide another application of LP theory: a stronger version of the the Trace Theorem, in Besov spaces. It is taken from Klainerman and Rodnianski, "Sharp Trace Theorems for Null Hypersurfaces on Einstein Metrics with Finite Curvature Flux", see [**9**].

For simplicity, let $I = [0,1]$ and consider $I \times \mathbb{R}^2$. We will use the mixed norm notation:

$$\|f\|_{L_t^q L_x^p} = \left(\int_0^1 \|f(t,\cdot)\|^q_{L_x^p(\mathbb{R}^2)} dt \right)^{\frac{1}{q}}$$
$$\|f\|_{L_x^p L_t^q} = \left(\int_{\mathbb{R}^2} \|f(\cdot,x)\|^p_{L_t^q(I)} dx \right)^{\frac{1}{p}}$$

with the obvious modifications if $p = \infty$ or $q = \infty$.

We will get the following trace-like estimate:

$$\| \int_I |\partial_t f|^2 dt \|_{B^1_{2,1}} \lesssim \|f\|^2_{H^2(I\times\mathbb{R}^2)} \tag{97}$$

We observe that

$$\|g\|_{B^1_{2,1}} \lesssim \|\nabla g\|_{B^0_{2,1}} + \|g\|_{L^2}$$

Thus, (97) follows from the "sharp bilinear trace" theorem below.

THEOREM 8.1. *For any smooth, scalar functions g,h on $I \times \mathbb{R}^2$, we have*

$$\| \int_I \partial_t g \cdot h dt \|_{B^0_{2,1}} \lesssim \|g\|_{H^1(I\times\mathbb{R}^2)} \cdot \|h\|_{H^1(I\times\mathbb{R}^2)} \tag{98}$$

Proof Immediately we see:

$$\begin{aligned} \| \int_I \partial_t g \cdot h dt \|_{B^0_{2,1}} &= \sum_{k\geq 0} \|P_k \int_0^1 \partial_t g \cdot h dt\|_{L^2_x} + \|P_{<0} \int_0^1 \partial_t g \cdot h dt\|_{L^2_x} \\ &\lesssim \sum_{k\geq 0} \|P_k \int_0^1 \partial_t g \cdot h dt\|_{L^2_x} \end{aligned}$$

We will then decompose g and h with respect to x; $g = \sum_k P_k g = \sum_k g_k$, $h = \sum_k P_k h = \sum_k h_k$. Then we can decompose $P_k \int_0^1 (\partial_t g \cdot h) = A_k + B_k + C_k + D_k$, where

$$A_k = P_k \int_0^1 (\partial_t g)_{<k} \cdot h_{\geq k}$$
$$B_k = P_k \int_0^1 (\partial_t g)_{\geq k} \cdot h_{<k}$$

$$C_k = P_k \int_0^1 (\partial_t g)_{<k} \cdot h_{<k}$$

$$D_k = P_k \int_0^1 (\partial_t g)_{\geq k} \cdot h_{\geq k}$$

As in the Trichotomy Formula, C_k is essentially zero (with the exception of finitely many terms which can be subsumed in A_k, B_k, or D_k).

We now briefly sketch how to estimate each of A_k, B_k, D_k, leaving the details to be filled in. Note that P_k trivially commutes with the integrals $\int_0^1 dt$ and any partial derivatives ∂_t.

To estimate A_k, note that we can write (using **LP2**):

$$\|A_k\|_{L^2_x} \lesssim \sum_{k'<k\leq k''} \int_0^1 \|(\partial_t g)_{k'} \cdot h_{k''}\|_{L^2_x} dt$$

We can then use Bernstein inequality **LP4** and property **LP3** on h to pull out the power $2^{k'-k''}$. Writing $2^{k'-k''} \lesssim 2^{(k'-k)/2+(k-k'')/2}$, using **LP1**, and summing over k, we can then get:

$$\sum_{k\geq 0} \|A_k\|_{L^\infty_t L^2_x} \lesssim \|\partial_t g\|_{L^\infty_t L^2_x} \cdot \|\nabla h\|_{L^\infty_t L^2_x}$$

To estimate $D_k = P_k \int_0^1 (\partial_t g)_{\geq k} \cdot h_{\geq k}$, write

$$D_k = D_k^1 + D_k^2 = \sum_{k\leq k'\leq k''} P_k \int_0^1 (\partial_t g)_{k'} \cdot h_{k''} + \sum_{k\leq k'\leq k''} P_k \int_0^1 (\partial_t g)_{k''} \cdot h_{k'}$$

D_k^1 can be estimated straightforwardly, without integration by parts. Use **LP4** and **LP3** to write

$$\|D_k^1\|_{L^2_x} \lesssim 2^{k-k'} \|\partial_t g\|_{L^2_t L^2_x} \cdot \|\nabla h\|_{L^2_t L^2_x}$$

Then sum over k and use **LP1** to get:

$$\sum_{k\geq 0} \|D_k^1\|_{L^2_x} \lesssim \|\partial_t g\|_{L^2_t L^2_x} \cdot \|\nabla h\|_{L^2_t L^2_x}$$

To estimate D_k^2 we use integration by parts to transfer the ∂_t from the high-frequency $g_{k''}$ to the low-frequency $h_{k'}$. After integrating by parts we treat the result exactly as D_k^1. Thus, we need only estimate the boundary terms: $\|I_k(1) - I_k(0)\|_{L^2_x} \lesssim \|I_k\|_{L^\infty_t L^2_x}$, where

$$I_k = \sum_{k\leq k'<k''} P_k(g_{k''} \cdot h_{k'})$$

We use the following lemma to do so:

LEMMA 8.2. *For any k,k',k" we have*

$$\|P_k(g_{k'} \cdot h_{k''})\| \lesssim 2^{-\frac{1}{4}(|k'-k|+|k''-k|)} \|g_{k'}\| \|h_{k''}\|$$

Using this lemma, we integrate by parts and bound D_k^2 just as D_k^1 plus the boundary term, and eventually get:

$$\sum_k \|D_k^2\|_{L_x^2} \lesssim \|g\|_{H^1} \cdot \|h\|_{H^1}$$

Now we estimate B_k by similarly decomposing to $B_k = \sum_{k'<k\le k''} P_k \int_0^1 (\partial_t g)_{k''} \cdot h_{k'}$. As above, we integrate by parts and use the lemma to estimate the boundary terms $J_k = \sum_{k'<k\le k''} P_k(g_{k''}) \cdot h_{k'})$. It is then not hard to manipulate and sum over k to get

$$\sum_k \|B_k\|_{L_x^2} \lesssim \|g\|_{H^1} \cdot \|h\|_{H^1}$$

Combining all the estimates for A_k, B_k, and D_k completes the proof of the theorem.

It only remains to prove the above Lemma which helped us estimate the boundary terms. Without going into all the details, this is done by considering the three cases:

$$k' \ge k'' \ge k,\ k' \ge k > k'',\ k > k' \ge k''$$

We note that the third ("low-low") case is impossible. The other two cases are bounded using **LP3** and the the following (simple) calculus inequality:

$$\|f\|_{L_t^\infty L_x^2} \lesssim \|\partial_t f\|_{L_t^2 L_x^2}^{\frac{1}{2}} \cdot \|f\|_{L_t^2 L_x^2}^{\frac{1}{2}} + \|f\|_{L_t^2 L_x^2} \tag{99}$$

Estimating $\|P_k(g_{k'} \cdot h_{k''})\|_{L_t^\infty L_x^2}$ using (99) and **LP3** yields the estimate in the lemma. ■

Exercise. Fill in the missing steps in the proof of the above theorem.

9. Calderon-Zygmund theory

The following L^2 identity

$$\sum_{i,j=1}^n \|\partial_i \partial_j u\|_{L^2}^2 = \|\Delta u\|_{L^2}^2.$$

for any $u \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ can be easily established by integration by parts, see below in (103). Thus,

$$\|\partial^2 u\|_{L^2} \lesssim \|\Delta u\|_{L^2} \tag{100}$$

It is natural to ask whether such estimate still holds true for other L^p norms. It turns out that the problem can be reduced to that of study the L^p boundedness properties for a very important class of linear operators called Calderon-Zygmund.

Definition 9.1. A linear operator T acting on $L^2(\mathbb{R}^n)$ is called a *Calderon-Zygmund* operator if:

(1) T is bounded from L^2 to L^2.

(2) There exists a measurable kernel k such that for every $f \in L^2$ with compact support and for $x \notin \operatorname{supp} f$, we have

$$Tf(x) = \int_{\mathbb{R}^n} k(x-y)f(y)\mathrm{d}y,$$

where the integral converges absolutely for all x in the complement of $\operatorname{supp} f$.

(3) There exists constants $C > 1$ and $A > 0$ such that

$$\int_{|x|\geq C|y|} |k(x-y) - k(x)|\, \mathrm{d}x \leq A, \tag{101}$$

uniformly in y. For simplicity one can take $C = 2$.

Proposition 9.2. *Assume that the kernel $k(x)$ verifies, for all $x \neq 0$,*

$$|k(x)| \lesssim |x|^{-n}, \qquad |\partial k(x)| \lesssim |x|^{-n-1} \tag{102}$$

Then k verifies the cancellation condition (101).

Exercise. Prove the proposition.

Example 1. Hilbert transform $Hf(x) = \int e^{ix\cdot\xi} \operatorname{sign} \xi\, \hat{f}(\xi)\mathrm{d}\xi$. By Plancherel it is easy to check that H is a bounded linear operator on L^2. On the other hand we know that the inverse Fourier transform of $\operatorname{sign} \xi$ is proportional to the principal value distribution $\mathrm{pv}(1/x)$. Hence, if $x \notin \operatorname{supp} f$,

$$Hf(x) = c\int_{-\infty}^{+\infty} \frac{1}{x-y} f(y)dy.$$

It is easy to check that the kernel $k(x) = \frac{1}{x}$ verifies condition 3 above.

Example 2. Consider the equation $\Delta u = f$ in $\mathbb{R}^n$, $n \geq 3$, for f, smooth, compactly supported. Recall, see (15), that any solution u, vanishing at[16] ∞, can be represented in the form, $u = K_n * f$ where $K_n(x) = c_n|x|^{2-n}$. Thus, if $x \notin \operatorname{supp} f$, it makes sense to differentiate under the integral sign and derive,

$$\partial_i\partial_j u = \partial_i\partial_j K_n * f = \int_{\mathbb{R}^n} \partial_i\partial_j K_n(x-y)f(y)dy.$$

It is easy to check that the kernel $k(x) = \partial_i\partial_j K_n(x)$ verifies condition 3. To show that the operators $R_{ij}f(x) = \int_{\mathbb{R}^n} \partial_i\partial_j K_n(x-y)f(y)dy$ are Calderon-Zygmund operators, it only remains to check the L^2 -boundedness property. This follows easily from the equation $\Delta u = f$. Indeed $u = K_n * f$ is the unique solution of the equation vanishing at ∞. Moreover $|u(x)| \lesssim |x|^{2-n}$, $|\partial u(x)| \lesssim |x|^{1-n}$ and $R_{ij}f = \partial_i\partial_j u(x)$. Thus we can integrate by parts in the expression,

$$\begin{aligned}
\int_{\mathbb{R}^n} |f(x)|^2 dx &= \int_{\mathbb{R}^n} \Delta u(x)\Delta u(x) dx = \sum_{i,j=1}^n \int_{\mathbb{R}^n} |\partial_i\partial_j u(x)|^2 dx \\
&= \sum_{i,j=1}^n \int_{\mathbb{R}^n} |R_{ij}f(x)|^2 dx
\end{aligned} \tag{103}$$

[16] In the case of $n = 2$ any solution whose first derivatives vanish at ∞.

Hence for each pair $1 \leq i, j \leq n$,

$$\|R_{ij} f\|_{L^2} \leq \|f\|_{L^2}.$$

Thus the operators R_{ij} are Calderon-Zygmund. We shall write schematically $R_{ij} = \partial_i \partial_j (-\Delta)^{-1}$.

THEOREM 9.3. *Calderon-Zygmund operators are bounded from L^1 into weak-L^1.*

As a consequence we derive,

COROLLARY 9.4. *Calderon-Zygmund operators are bounded from L^p into L^p, for any $1 < p < \infty$. They are not bounded, in general, for $p = 1$ and $p = \infty$.*

Proof : The boundedness over L^p for $1 < p < 2$ follows from the weak-L^1 and the L^2 boundedness by Marcinkiewicz interpolation. The cases $p > 2$ follow by duality from the fact that the dual of a Calderon-Zygmund operator, with kernel $k(x)$, is again a Calderon-Zygmund operator, with kernel $k(-x)$. More precisely, if f, g have disjoint supports,

$$\int_{\mathbb{R}^n} Tf(x)g(x)dx = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} k(x-y)f(y)g(x)dx = \int_{\mathbb{R}^n} f(y)\, T^* g(y)dy$$

where

$$T^* g(y) = \int_{\mathbb{R}^n} k(-y+x)g(x)dx, \qquad \forall y \notin \ \text{supp} g.$$

On the other hand $\|T^* f\|_{L^2} = \|Tf\|_{L^2} \lesssim \|f\|_{L^2}$. Hence T^* is indeed a CZ operator. Now, using the duality between L^p and $L^{p'}$, $1/p + 1/p' = 1$ and the fact that T^* is $L^{p'}$ bounded for $p' \leq 2$,

$$\begin{aligned}
\|Tf\|_{L^p} &= \sup_{\|g\|_{L^{p'}} \leq 1} \left| \int_{\mathbb{R}^n} Tf(x)g(x)dx \right| = \sup_{\|g\|_{L^{p'}} \leq 1} \left| \int_{\mathbb{R}^n} f(x)\, T^* g(x)dx \right| \\
&= \sup_{\|g\|_{L^{p'}} \leq 1} \|f\|_{L^p} \cdot \|T^* g\|_{L^{p'}} \lesssim \|f\|_{L^p}.
\end{aligned}$$

■

We shall prove the main theorem 9.3 in the next two subsections.

9.5. Calderon-Zygmund decompositions.

DEFINITION 9.6. We define a *dyadic cube* in $\mathbb{R}^n$ to be a cube Q of the form

$$Q = [2^k a_1, 2^k(a_1+1)] \times \cdots \times [2^k a_n, 2^k(a_n+1)],$$

where $k, a_1, \ldots, a_n \in \mathbb{Z}$. We then say that $\text{size}(Q) = 2^k$. If Q is a dyadic cubes then its *parent* is the only dyadic cube Q^* such that $Q \subset Q^*$ and $\text{size}(Q^*) = 2\,\text{size}(Q)$ and we say that Q is a *child* of Q^*.

LEMMA 9.7 (Whitney decomposition). *Any proper open set Ω in $\mathbb{R}^n$ can be covered by a family $\mathcal{Q} = \{Q\}$ of disjoint dyadic cubes*

$$\Omega = \cup_{Q \in \mathcal{Q}} Q,$$

where each cube $Q \in \mathcal{Q}$ satisfies the property

$$\textit{size}\,(Q) \approx \textit{dist}\,(Q, \partial\Omega). \tag{104}$$

Proof : For each $x \in \Omega$ denote by Q_x the largest dyadic cube containing x with the property: $\operatorname{dist}(Q_x, \partial\Omega) > \operatorname{size}(Q_x)$. If Q^* denotes the parent of Q_x then $\operatorname{dist}(Q^*, \partial\Omega) \leq \operatorname{size}(Q^*)$. By the triangular inequality it follows that

$$\operatorname{dist}(Q_x, \delta\Omega) \leq \sqrt{n}\, \operatorname{size}(Q_x) + \operatorname{dist}(Q^*, \delta\Omega) \leq \left(\sqrt{n} + 2\right) \operatorname{size}(Q_x).$$

Hence, Q_x verifies (104). If $y \in Q_x$ then, by the maximality property of Q_x and Q_y, we necessarily have $Q_y = Q_x$. Hence, the family $\mathcal{Q} = \{Q_x\}_{x\in\Omega}$ is formed of disjoint cubes and covers Ω. ■

PROPOSITION 9.8 (Calderon-Zygmund decomposition). *Let $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$. Then it is possible to find a countable family of disjoint dyadic cubes $\mathcal{Q} = \{Q\}$ and a decomposition $f = g + \sum_{Q\in\mathcal{Q}} b_Q$, such that:*

$$\|g\|_{L^\infty} \lesssim \alpha, \tag{105a}$$

$$\textit{supp}\ b_Q \subseteq Q, \tag{105b}$$

$$\int b_Q(x)\,dx = 0, \tag{105c}$$

$$\|b_Q\|_{L^1} \lesssim \alpha |Q|, \tag{105d}$$

$$\sum_Q |Q| \lesssim \frac{1}{\alpha} \|f\|_{L^1}\,. \tag{105e}$$

Proof : Let $\mathcal{Q}$ be the Whitney decomposition of the open set $\Omega = \{\mathcal{M}f(x) > \alpha\}$ as indicated in Lemma 9.7. For each Q, define $f_Q = |Q|^{-1} \int_Q f(x)\mathrm{d}x$. Let

$$g(x) = \begin{cases} f(x), & \text{if } x \notin \Omega, \\ f_Q, & \text{if } x \in Q, \end{cases}$$

and $b_Q(x) = \chi_Q(x)(f(x) - f_Q)$ with χ_Q the characteristic function of the cube Q. Of course we have $f = g + \sum_Q b_Q$. The important property, which follows from (104), is that each cube Q is contained inside a ball B which is not entirely contained in Ω and with $|Q| \approx |B|$. Let $x \in B \setminus \Omega$, we have

$$|f_Q| \leq \frac{1}{|Q|} \int_Q |f(y)|\,\mathrm{d}y \lesssim \frac{1}{|B|} \int_B |f(y)|\,\mathrm{d}y \leq \mathcal{M}f(x) \leq \alpha. \tag{106}$$

We check now that this decomposition has the desired properties. For almost every x outside Ω, by Lebesgue's differentiation theorem, Corollary 5.7, we have $|g(x)| \leq \mathcal{M}f(x) \leq \alpha$. When $x \in \Omega$ it follows from (106) that $g(x) \lesssim \alpha$. Hence (105a) is satisfied. Properties (105b) and (105c) are immediate consequences of the definition of h_Q. Property (105d) is implied by (106). Finally, (105e) is nothing but the weak L^1 property for $\mathcal{M}f$ proved in Theorem 5.4. ■

9.9. Proof of Theorem 9.3. Consider $f \in L^1$ and $\alpha > 0$. Let $f = g + \sum_Q b_Q = g + b$ be the Calderon-Zygmund decomposition of f according to Theorem 9.8. Since

$$\{|Tf(x)| > \alpha\} \subseteq \{|Tg(x)| > \alpha/2\} \cup (\{|Tb(x)| > \alpha/2\})$$

and in view of (105e) it is enough to prove separately that

$$|\{|Tg(x)| > \alpha/2\}| \lesssim \frac{1}{\alpha} \|f\|_{L^1}, \tag{107}$$

$$|\{|Tb(x)| > \alpha/2\}| \lesssim \frac{1}{\alpha} \|f\|_{L^1} \tag{108}$$

Estimate (107) follows from Chebyschev's inequality, the boundedness of T on L^2 and the uniform bound on g,

$$\begin{aligned} |\{|Tg(x)| > \alpha/2\}| &\lesssim \frac{1}{\alpha^2} \|Tg\|_{L^2}^2 \lesssim \frac{1}{\alpha^2} \|g\|_{L^2}^2 \lesssim \frac{1}{\alpha} \|g\|_{L^1} \leq \\ &\leq \frac{1}{\alpha}\left(\|f\|_{L^1} + \sum_Q \|b_Q\|_{L^1} \right) \lesssim \frac{1}{\alpha} \|f\|_{L^1} + \sum_Q |Q| \lesssim \frac{1}{\alpha} \|f\|_{L^1}. \end{aligned}$$

It remains to derive (108). Since the family $\mathcal{Q}$ is countable we denote its members by Q_j, $j \in \mathbb{N}$. For each Q_j let $y_{(j)}$ be its center and take $\hat{Q}_j$ to be the cube with the same center but with the sides expanded by $2n^{1/2}$, such that for all x in the complement of $\hat{Q}_j$,

$$|x - y_{(j)}| \geq 2 \max_{y \in Q_j} |y - y_{(j)}|$$

Let $\Omega = \cup_j \hat{Q}_j$ and F its complement. We denote $b_j = b_{Q_j}$. Since $\int b_j dy = 0$ we write, for $x \in F$,

$$T(b_j)(x) = \int_{Q_j} \left(k(x-y) - k(x - y_{(j)})\right) b_j(y) \mathrm{d}y,$$

or, since the cubes Q_j are disjoint,

$$T(b_j)(x) = \int_{Q_j} \left(k(x-y) - k(x - y_{(j)})\right) b(y) \mathrm{d}y,$$

Thus, in view of (101),

$$\begin{aligned} \int_F |T(b)(x)| dx &\leq \sum_j \int_F |T(b)(x)| dx \lesssim \sum_j \int_{x \in \mathbb{R}^n \setminus \hat{Q}_j} \int_{y \in Q_j} |k(x-y) - k(x - y_{(j)})| \, |b(y)|, \\ &= \sum_j \int_{y \in Q_j} |b_j(y)| \int_{x \in \mathbb{R}^n \setminus \hat{Q}_j} |k(x-y) - k(x - y_{(j)})| \\ &\leq \sum_j \int_{y \in Q_j} |b(y)| \int_{x \in \mathbb{R}^n \setminus \{\hat{Q}_j - y_{(j)}\}} |k(x - (y - y_{j)}) - k(x)| \\ &\lesssim \sum_j \int_{y \in Q_j} |b(y)| \int_{|x| \geq 2|(y - y_{j)}|} |k(x - (y - y_{j)}) - k(x)| \\ &\lesssim A \sum_j \int_{y \in Q_j} |b(y)| \lesssim \|f\|_{L^1} \end{aligned}$$

Therefore,

$$|\{x \in F : |Tb(x)| > \alpha/2\}| \lesssim \alpha^{-1} \|f\|_{L^1}$$

On the other hand, the measure of the complement of F, i.e. $\Omega = \cup \hat{Q}_j$ is given by,

$$|\Omega| \leq \sum_j |\hat{Q}_j| \lesssim \sum_j Q_j \lesssim \alpha^{-1} \|f\|_{L^1}.$$

Hence,

$$|\{x \in \mathbb{R}^n : |Tb(x)| > \alpha/2\}| \lesssim \alpha^{-1} \|f\|_{L^1}$$

as desired.

9.10. Michlin-Hörmander theorem. An important class of CZ operators can bedefined by means of Fourier multiplier operators. Recall that these are defined by Fourier transform,

$$\widehat{Tf}(\xi) = m(\xi) \widehat{f}(\xi), \tag{109}$$

where m is a bounded function, called the multiplier. We can view these operators as convolution operators, $Tf = k * f$, where $\widehat{k} = m$.

THEOREM 9.11. *Let $l > n/2$. Suppose m is a Fourier multiplier of class C^l on $\mathbb{R}^n \setminus 0$, such that*

$$\left|\partial_\xi^\alpha m(\xi)\right| \lesssim |\xi|^{-|\alpha|}, \qquad \forall \xi \in \mathbb{R}^n \setminus 0$$

for every multiindex α with $|\alpha| \leq l$. Then the operator defined by (109) *is a Calderon-Zygmund operator.*

Proof : Consider the same dyadic partition of unity as that used in the LP projections,

$$1 = \sum_{\lambda \in 2^{\mathbb{Z}}} \chi_\lambda(\xi) \qquad \text{for} \quad \xi \in \mathbb{R}^n \setminus 0,$$

generated by $\chi \in C_0^\infty$ with $\operatorname{supp}\chi \subseteq \{1/2 \leq |\xi| \leq 2\}$, and $\chi_\lambda(\xi) = \chi(\xi/\lambda)$.

Decompose m into dyadic pieces, $m = \sum_\lambda m_\lambda$, where $m_\lambda = \chi_\lambda m$. Since $|\partial^\gamma m(\xi)| \lesssim |\xi|^{-|\gamma|}$ and all derivatives of $\chi(\xi)$ are bounded,

$$\left|\partial_\xi^\alpha m_\lambda(\xi)\right| \leq \sum_{|\beta|+|\gamma| \leq |\alpha|} \left|\partial^\beta \chi_\lambda |\xi|^{-\gamma}\right| \lesssim \sum_{|\beta|+|\gamma| \leq |\alpha|} \lambda^{-|\beta|} \lambda^{-|\gamma|} \approx \lambda^{-|\alpha|}. \tag{110}$$

Let k_λ be the inverse Fourier transform of m_λ. Since m_λ has compact support k_λ is a smooth function. Moreover, for any integer N we have[17]

$$|k_\lambda(x)| \lesssim |x|^{-N} \left\|\partial^N m_\lambda\right\|_{L^1} \lesssim |x|^{-N} \lambda^{n-N}.$$

Now take $N > n$ and sum over $\lambda \in 2^{\mathbb{Z}}$. Observe that $\sum_\lambda k_\lambda$ converges to a well defined measurable function k on $\mathbf{R}^n \setminus 0$, and it easy to see that k satisfies property 2 of Definition 9.1.

[17] Recall that, by integration by parts, we have $|\mathcal{F}^{-1} f(x)| \leq |x|^{-N} \left\|\partial_\xi^N f\right\|_{L^1}$,

The boundedness of T on L^2 follows immediately from the boundedness of m on $\mathbb{R}^n$.

For $0 \leq j \leq l$, by Plancherel's theorem and (110) we obtain

$$\int |x|^{2j} |k_\lambda(x)|^2 \,\mathrm{d}x \simeq \sum_{|\alpha|=j} \int \left|\partial_\xi^\alpha m_\lambda(\xi)\right|^2 \mathrm{d}\xi \lesssim \lambda^{n-2j}.$$

Let $R > 0$, using the case $j = 0$ we find that

$$\int_{|x|\leq R} |k_\lambda(x)| \,\mathrm{d}x \lesssim \left(\int |k_\lambda(x)|^2 \,\mathrm{d}x\right)^{1/2} R^{n/2} \lesssim (\lambda R)^{n/2}, \tag{111}$$

while using the case $j = l$ we find that

$$\int_{|x|\geq R} |k_\lambda(x)| \,\mathrm{d}x \lesssim \left(\int |x|^{2l} |k_\lambda(x)|^2 \,\mathrm{d}x\right)^{1/2} \left(\int_{|x|>R} \frac{\mathrm{d}x}{|x|^{2l}}\right)^{1/2} \lesssim (\lambda R)^{n/2-l}. \tag{112}$$

If we choose $R = 1/\lambda$, summing (111) and (112) we obtain $\|k_\lambda\|_{L^1} \lesssim 1$ uniformly in λ. We can apply the same procedure to ∂k_λ, which has symbol $\xi m_\lambda \approx \lambda m_\lambda$, to prove that $\|\partial k_\lambda\|_{L^1} \lesssim \lambda$. Hence,

$$\begin{aligned}
\int_{|x|\gg|y|} |k_\lambda(x-y) - k_\lambda(x)| \,\mathrm{d}x &\leq \int\int_0^{|y|} |\partial k_\lambda(x - ty/|y|)| \,\mathrm{d}t\mathrm{d}x && (113)\\
&= |y| \cdot \|\partial k_\lambda\|_{L^1} \lesssim \lambda|y|, && (114)
\end{aligned}$$

but also, by (112),

$$\int_{|x|\gg|y|} |k_\lambda(x-y) - k_\lambda(x)| \,\mathrm{d}x \leq 2\int_{|x|\geq|y|} |k_\lambda(x)| \,\mathrm{d}x \lesssim (\lambda|y|)^{n/2-l}. \tag{115}$$

We sum over λ using (113) when $\lambda|y| \leq 1$ and (115) when $\lambda|y| > 1$, and obtain[18]

$$\int_{|x|\gg|y|} |k(x-y) - k(x)| \,\mathrm{d}x \lesssim |y| \sum_{\lambda\leq|y|^{-1}} \lambda + |y|^{n/2-l} \sum_{\lambda>|y|^{-1}} \lambda^{n/2-l} \lesssim 1.$$

as desired. ∎

9.12. Square function estimates. We recall property LP6 for the square function, $Sf = \left(\sum_k |P_k f|^2\right)^{1/2}$,

THEOREM 9.13 (Littlewood-Paley). *We have,*

$$\|f\|_{L^p} \lesssim \|Sf\|_{L^p} \lesssim \|f\|_{L^p} \tag{116}$$

for all $1 < p < \infty$.

---

[18]Here we used the following summation properties, in dyadic notation, for geometric series, $\sum_{\lambda\leq L} \lambda^\alpha \simeq L^\alpha$ and $\sum_{\lambda\geq L} \lambda^{-\alpha} \simeq L^{-\alpha}$ for $\alpha > 0$.

We give two proofs of this estimate.

Proof [first proof]: First we show using duality arguments that the first inequality in (116) follows from the second one. Indeed using Plancherel's theorem, the fact that $P_k P_{k'} = 0$ unless $k \sim k'$, and Cauchy-Schwartz inequality we obtain

$$\begin{aligned}
\int f(x)g(x)\mathrm{d}x &\simeq \int \sum_{k\approx k'} P_k f(x) P_{k'} g(x)\mathrm{d}x \\
&\lesssim \int \left(\sum_k |P_k f(x)|^2\right)^{1/2} \left(\sum_{k'} |P_{k'} g(x)|^2\right)^{1/2} \mathrm{d}x \leq \\
&\lesssim \|Sf\|_{L^p} \|Sg\|_{L^{p'}} \lesssim \|Sf\|_{L^p} \|g\|_{L^{p'}} .
\end{aligned}$$

The left inequality in (116) now follows by taking the sup over all g with $\|g\|_{L^{p'}} = 1$.

To prove the right inequality in (116) we need to introduce the Rademacher functions $r_k(t)$ defined on $\mathbb{R}$ as follows: for every $k \geq 0$, $k \in \mathbb{Z}$ and $t \in \mathbb{R}$ set $r_k(t) = r_0(2^k t)$, where $r_0(t)$ is the periodic function, $r_0(t+1) = r_0(t)$, such that $r_0(t) = 1$ for $0 \leq t < 1/2$ and $r_0(t) = -1$ for $1/2 \leq t < 1$. These Rademacher functions form an orthonormal sequence in $L^2[0,1]$ and they form a sequence of independent identically distributed random variables. The basic property that we need is that the L^p norm of a linear combination of Rademacher function is equivalent to the l^2 norm of its coefficients.

LEMMA 9.14. *Given a sequence of real numbers $\{a_k\}$ satisfying $\sum_{k=0}^{\infty} a_k^2 < \infty$, define*

$$F(t) = \sum_{k=0}^{\infty} a_k r_k(t).$$

Then $F \in L^2([0,1])$ with $\|F\|_{L^2} = (\sum_{k=0}^{\infty} a_k^2)^{1/2}$. In addition, $F \in L^p([0,1])$ for $1 < p < \infty$, and there exist constants A_p so that

$$A_p^{-1} \|F\|_{L^p} \leq \|F\|_{L^2} \leq A_p \|F\|_{L^p} .$$

For a proof of this lemma see Stein, [**18**, Appendix D].

Define the operator T_t so that

$$T_t f = \sum_{k=0}^{\infty} r_k(t) P_k f$$

Clearly T_t is the Fourier multiplier operator with symbol $m_t(\xi) = \sum_k r_k(t)\chi(2^{-k}\xi)$, where χ is the smooth cut-off function used to define the LP projections. For $\xi \neq 0$, at most three of the terms in the sum defining $m_t(\xi)$ can be non-zero. We can then easily verify that m_t verifies the condition of Thm. 9.11. That is, that

$$|\partial_\xi^\alpha m_t(\xi)| \leq C_\alpha |\xi|^{-|\alpha|},$$

with constants C_α independent of t. Thus, by Calderon-Zygmund theory (specifically Corollary 9.4), we have:

$$\|T_t f\|_{L^p} \lesssim \|f\|_{L^p}$$

And so,

$$\left(\int_0^1 \|T_t f\|_{L^p}^p dt\right)^{1/p} \lesssim \|f\|_{L^p}$$

In addition, we can use Lemma 9.14 to see that:

$$\begin{aligned}\int_0^1 \|T_t f\|_{L^p}^p dt &= \int_0^1 \int_{\mathbb{R}} |\sum_k r_k(t)(P_k f)(x)|^p dx dt \\ &\gtrsim \int_{\mathbb{R}} \left(\sum_k |(P_k f)(x)|^2\right)^{p/2} dx\end{aligned}$$

And so combining our results we get:

$$\|Sf\|_{L^p} \lesssim \|f\|_{L^p}$$

(Note that this argument proves the theorem only in the one-dimensional case, $n = 1$. It can, however, be extended to $\mathbb{R}^n$ as in Stein, *Singular Integrals*, Ch. IV, Section 5.)

■

Proof [second proof]: We recall the definition for the vector-valued function,

$$\mathbf{S}f(x) = \big(P_k f(x)\big)_{k\in\mathbb{Z}}.$$

Clearly, if $f \in \mathcal{S}(\mathbb{R}^n)$, for every $x \in \mathbb{R}^n$, $\mathbf{S}f(x) \in l^2$ and $Sf(x) = |\mathbf{S}f(x)|$ denotes the l^2 norm of $\mathbf{S}f(x)$. We claim that

$$\mathbf{S}f(x) = \int \mathbf{K}(x-y)f(y)dy$$

is a an l^2 -valued Calderon-Zygmund operator with the l^2-valued kernel defined by,

$$\mathbf{K}(x) = \big(K_k(x)\big)_{k\in\mathbb{Z}}, \qquad K_k(x) = 2^{nk}\hat{\chi}(2^k x)$$

Denote $|\mathbf{K}(x)| = \left(\sum_k |K_k(x)|^2\right)^{1/2}$, $|\partial\mathbf{K}(x)| = \left(\sum_k |\partial K_k(x)|^2\right)^{1/2}$. We easily check that the l^2- valued version of the condition (102) is verified,

$$|\mathbf{K}(x)| \lesssim |x|^{-n} \qquad |\partial\mathbf{K}(x)| \lesssim |x|^{-(n+1)}, \qquad \text{for} \quad x \neq 0. \tag{117}$$

On the other hand,

$$\|\mathbf{S}f\|_{L^2} := \|Sf\|_{L^2} \lesssim \|f\|_{L^2}.$$

Thus $\mathbf{S}$ is indeed an l^2 valued C-Z operator and therefore, in view of a straightforward extension of Theorem 9.3 and its corollary, we infer that,

$$\|\mathbf{S}f\|_{L^p} := \||\mathbf{S}f|\|_{L^p} = \|Sf\|_{L^p} \lesssim \|f\|_{L^p}$$

In view of the beginning of the first proof of our theorem we infer that also,

$$\|f\|_{L^p} \lesssim \|Sf\|_{L^p}.$$

■

Remark that, according to theorem 9.13, $|\sum_k P_k f| \approx \big(\sum_k |P_k f|^2\big)^{1/2}$. A more general principle asserts that if a sequence of functions $f_1, f_2, \ldots f_k \ldots$ oscillate at different rates, that is any two phases are different, then $|\sum_k f_k| \approx \big(\sum_k |f_k|^2\big)^{1/2}$.

The following version of the property LP6, and theorem 9.13, also holds true for LP projections $\tilde{P}_k \sim P_k$. More precisely,

$$\|\big(\sum_k |\tilde{P}_k f|^2\big)^{1/2}\|_{L^p} \lesssim \|f\|_{L^p}, \qquad 1 < p < \infty. \tag{118}$$

This can be proved in the same manner as the inequality $\|Sf\|_{L^p} \lesssim \|f\|_{L^p}$ by introducing the l^2 valued operator, $\tilde{\mathbf{S}}f = (\tilde{P}_k f)_{k\in\mathbb{Z}}$, and proceeding exactly as in the second proof of theorem 9.13. Given an l^2 valued vector function $\mathbf{g} = (g_k)_{k\in\mathbb{Z}}$ observe that

$$< \tilde{\mathbf{S}}f, \mathbf{g} >= \int_{\mathbb{R}^n} \tilde{\mathbf{S}}f(x) \cdot \bar{\mathbf{g}}(x)dx = \int_{\mathbb{R}^n} \sum_k \tilde{P}_k f(x)\overline{g_k}(x)dx = \int_{\mathbb{R}^n} f(x)\overline{\sum_k \tilde{P}_k g_k(x)}dx$$

Thus,

$$\tilde{\mathbf{S}}^*\mathbf{g} = \sum_k \tilde{P}_k g_k \tag{119}$$

and therefore the estimate dual to (118) has the form, $\|\tilde{\mathbf{S}}^*\mathbf{g}\|_{L^{p'}} \lesssim \|\mathbf{g}\|_{L^{p'}}$, for $1/p + 1/p' = 1$. In other words,

$$\|\sum_k \tilde{P}_k g_k\|_{L^p} \lesssim \|\big(\sum_k |g_k|^2\big)^{1/2}\|_{L^p}, \qquad 1 < p < \infty. \tag{120}$$

The following is an easy consequence of theorem 9.13.

Corollary 9.15. *For $2 \le p < \infty$ we have*

$$\|f\|_{L^p}^2 \lesssim \sum_{k\in\mathbb{Z}} \|P_k f\|_{L^p}^2 . \tag{121}$$

For $1 < p \le 2$ we have

$$\sum_{k\in\mathbb{Z}} \|P_k f\|_{L^p}^2 \lesssim \|f\|_{L^p}^2 . \tag{122}$$

Proof : Recall that $Sf(x)^2 = \sum_{k\in\mathbb{Z}} |P_k f|^2$. If $p/2 \ge 1$, in view of LP6 and Minkowski inequality, we have

$$\|f\|_{L^p}^2 \lesssim \|Sf\|_{L^p}^2 = \|\sum_k |P_k f|^2\|_{L^{p/2}} \le \sum_k \big\| |P_k f|^2 \big\|_{L^{p/2}} = \sum_k \|P_k f\|_{L^p}^2 .$$

If $p/2 \le 1$, we make use instead of the reverse Minkowski inequality,

$$\|f\|_{L^p}^2 \gtrsim \|\sum_k |P_k f|^2\|_{L^{p/2}} \ge \sum_k \big\| |P_k f|^2 \big\|_{L^{p/2}} = \sum_k \|P_k f\|_{L^p}^2 .$$

The reverse Minkowski inequality we have used here states that for $0 < q \le 1$ and a sequence of positive functions $(f_k)_{k\in\mathbb{Z}}$

$$\|\sum_k |f_k|\|_{L^q} \ge \sum_k \|f_k\|_{L^q} . \tag{123}$$

We briefly sketch a proof of (123); it can be found in many books (e.g. Garling, *Inequalities* or DiBenedetto, *Real Analysis*, from which we take this particular proof).

One way is to first prove a reverse Hölder inequality: For $0 < p < 1$, $q < 0$, $\frac{1}{p} + \frac{1}{q} = 1$, $f \in L^p$, $g \in L^q$, we have $\int |fg| \geq \|f\|_{L^p} \|g\|_{L^q}$. This can be easily shown by writing $\|f\|_{L^p} = \left(\int \frac{|fg|^p}{|g|^p}\right)^{1/p}$ and applying the usual Hölder inequality with the exponents $\tilde{p} = 1/p > 1$ and $\tilde{q} = 1/(1-p) > 1$.

With this in hand, the reverse Minkowski inequality in two terms ($\||f| + |g|\|_{L^q} \geq \|f\|_{L^q} + \|g\|_{L^q}$ for $0 < q \leq 1$) follows (writing $\frac{1}{q'} = 1 - \frac{1}{q}$):

$$\begin{aligned}
\||f| + |g|\|^q_{L^q} &= \int (|f| + |g|)^{q-1}(|f| + |g|) \\
&\geq \left(\int (|f| + |g|)^{(q-1)q'}\right)^{1/q'} (\|f\|_{L^q} + \|g\|_{L^q}) \\
&\geq \||f| + |g|\|^{q-1}_{L^q} (\|f\|_{L^q} + \|g\|_{L^q})
\end{aligned}$$

■

9.16. $W^{s,p}$- Sobolev spaces. We recall that we have defined the $W^{s,p}$ norm of a function by,

$$\|f\|_{W^{s,p}} = \sum_{j=0}^{s} \|\partial^j f\|_{L^p}.$$

We claim the following

LEMMA 9.17. *For any $j \geq 0$, $1 < p < \infty$ we have,*

$$\|\partial^j f\|_{L^p} \approx \|\big(\sum_k |2^{jk} P_k f|^2\big)^{1/2}\|_{L^p}$$

Proof : We first write,

$$\|\partial^j f\|_{L^p} \lesssim \|\sum_k \partial^j P_k f\|_{L^p}$$

As in the proof of the property LP5, we can express $\nabla^j P_k f = 2^{jk} \tilde{P}_k P_k f$ for some P_k similar to P_k. Hence, using the estimate (120)

$$\|\partial^j f\|_{L^p} \lesssim \|\sum_k 2^{jk} \tilde{P}_k P_k f\|_{L^p} \lesssim \|\big(\sum_k |2^{jk} P_k f|^2\big)^{1/2}\|_{L^p}.$$

On the other hand, we can also write $2^{jk} P_k f = \tilde{P}_k \partial^j f$ for some other similar LP projection. Then, in view of (118),

$$\|\big(\sum_k |2^{jk} P_k f|^2\big)^{1/2}\|_{L^p} \lesssim \|\big(\sum_k |\tilde{P}_k \partial^j f|^2\big)^{1/2}\|_{L^p} \lesssim \|\partial^j f\|_{L^p}$$

■

Using the lemma we can now find an equivalent definition using LP projections:

PROPOSITION 9.18. *For any $1 < p < \infty$ and any $s \in \mathbb{N}$ we have,*

$$\|f\|_{W^{s,p}} \approx \|\sum_k (1+2^k)^s P_k f\|_{L^p}. \tag{124}$$

Moreover, for the homogeneous $W^{s,p}$ norm $\|f\|_{\dot{W}^{s,p}} = \|\partial^s f\|_{L^p}$,

$$\|f\|_{\dot{W}^{s,p}} \approx \|\sum_k 2^{ks} P_k f\|_{L^p}. \tag{125}$$

Observe that the expressions on the right hand side of (124) and (125) make sense for every value $s \in \mathbb{R}$. We can thus extend the definitions of $W^{s,p}$, and $\dot{W}^{s,p}$ spaces to all real values s.

Additional characterizations of the homogeneous Sobolev norms $\| \ \|_{\dot{W}^{s,p}}$ can be given using the following,

PROPOSITION 9.19. *For $2 \le p < \infty$ and any s we have,*

$$\left(\sum_k 2^{kps} \|P_k f\|_{L^p}^p\right)^{1/p} \lesssim \|f\|_{\dot{W}^{s,p}} \lesssim \left(\sum_k 2^{2ks} \|P_k f\|_{L^p}^2\right)^{1/2}. \tag{126}$$

For $1 < p \le 2$ and $s \in \mathbf{R}$ we have

$$\left(\sum_k 2^{2ks} \|P_k f\|_{L^p}^2\right)^{1/2} \lesssim \|f\|_{\dot{W}^{s,p}} \lesssim \left(\sum_k 2^{kps} \|P_k f\|_{L^p}^p\right)^{1/p}. \tag{127}$$

Proof : If $p/2 \ge 1$, by Theorem 9.13 and Minkowski inequality we have

$$\|f\|_{L^p}^2 \lesssim \left\|\sum_k |P_k f|^2\right\|_{L^{p/2}} \le \sum_k \left\||P_k f|^2\right\|_{L^{p/2}} = \sum_k \|P_k f\|_{L^p}^2 .$$

If $p/2 \le 1$, by Theorem 9.13 and the reverse Minkowski inequality we have

$$\|f\|_{L^p}^2 \gtrsim \left\|\sum_k |P_k f|^2\right\|_{L^{p/2}} \ge \sum_k \left\||P_k f|^2\right\|_{L^{p/2}} = \sum_k \|P_k f\|_{L^p}^2 .$$

The remaining details should be clear to fill in. ∎

10. Midterm Exam

Problem 1.[Distributions in $\mathbb{R}$] In $\mathbb{R}^2$ we set $z = x + iy$, $\partial_z = \frac{1}{2}(\partial_x - i\partial_y)$ and $\partial_{\bar{z}} = \frac{1}{2}(\partial_x + i\partial_y)$. Let $E = \pi^{-1}\frac{1}{z}$. Show that E is a fundamental solution for the operator $\partial_{\bar{z}}$. Establish a connection bewteen this fact and the Cauchy formula for analytic functions.

Let $f(z)$ be a an analytic function in the domain $D_+ = \{z \in \mathbb{C} : 0 < \mathrm{Im}(z) < \epsilon\}$ such that $|f(z)| \lesssim |\mathrm{Im}(z)|^{-N}$ for all $z \in D$. Show that there exists a distribution $f_+ = f(\cdot + i0)$ such that for every $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$,

$$\lim_{y \to 0, y > 0} \int_{\mathbb{R}} f(x + iy)\phi(x)dx \quad = \quad < f_+, \phi >,$$

Similarly, for analytic functions defined on $D_- = \{z \in \mathbb{C}/) - \epsilon < \mathrm{Im}(z) < 0\}$ we can define a distribution $f_- = f(\cdot - i0)$,

$$\lim_{y \to 0, y < 0} \int_{\mathbb{R}} f(x + iy)\phi(x)dx \quad = \quad < f_-, \phi >$$

This defines, in particular when $f = \frac{1}{z} = \frac{1}{x+iy}$, the distributions $(x + i0)^{-1}$ and $(x - i0)^{-1}$. Prove the formulas,

$$(x + i0)^{-1} - (x - i0)^{-1} \quad = \quad -2\pi i \delta_0(x).$$

Show also that,

$$(x + i0)^{-1} = x^{-1} - i\pi\delta_0(x)$$

where $\frac{1}{x}$ is the principal value distribution defined in the text.

Problem 2.[Fundamental solutions] Consider the operator $Lu = \Delta u + u$ in $\mathbb{R}^3$. Find all solutions of $Lu = 0$ with spherical symmetry. Show that

$$K(x) = -\frac{\cos|x|}{4\pi|x|}$$

is a fundamental solution for L.

Problem 3.[Initial value problem] Consider the initial value problems for the following, four evolution equations in $\mathbb{R} \times \mathbb{R}^n$,

$$\partial_t u = \Delta u, \qquad u(0, x) = f(x) \tag{128}$$

$$\partial_t u = i\Delta u, \qquad u(0, x) = f(x) \tag{129}$$

$$\partial_t^2 u = \Delta u, \qquad u(0, x) = f(x), \quad \partial_t u(0, x) = g(x) \tag{130}$$

$$\partial_t^2 u = -\Delta u, \qquad u(0, x) = f(x), \quad \partial_t u(0, x) = g(x) \tag{131}$$

In each of these cases write down solutions using the Fourier transform method. In other words take the Fourier transform of each equation, set

$$\hat{u}(t, \xi) = \int e^{-ix\cdot\xi} u(t, x)dx,$$

and solve the resulting differential equation in t. Compare the results for the last two equations. Show that (130) has solutions for all $f, g \in \mathcal{S}(\mathcal{R}^n)$ while (131) does not. Show however that if we only prescribe $u(0, x) = f$ (this is the Dirichlet problem for the Laplacian $\partial_t^2 + \Delta$ in $\mathbb{R}^{n+1}$), then the problem has a unique solution u, which decays to zero as $|t| + |x| \to \infty$, for all functions $f \in \mathcal{S}(\mathbb{R}^n)$. In all cases express[19] the resulting solutions as integral operators applied to the initial data(in physical space).

Problem 4.[Extension operator] Let H be the half space $x_n > 0$ in $\mathbb{R}^n$ and $1 \le p \le \infty$. Show that there exists an extension operator, that is a bounded linear operator $E : W^{1,p}(H) \to W^{1,p}(\mathbb{R}^n)$ such that for all $u \in W^{1,p}(H)$ we have $Eu = u$ a.e. in H and

$$\|Eu\|_{W^{1,p}(\mathbb{R}^n)} \lesssim \|u\|_{W^{1,p}(H)}.$$

Extend the result to any $s \in \mathbb{N}$. Can you extend the result to arbitrary domains $U \subset \mathbb{R}^n$? What about domains with smooth boundaries ?

Problem 5.[Trace theorems] Let $\mathbb{R}^{n-1}$ be a hyperplane in $\mathbb{R}^n$, for example $x_n = 0$. For any $f \in \mathcal{S}(\mathbb{R}^n)$ let Rf denote the restriction of f to $\mathbb{R}^{n-1}$.

i. Prove that, for any $s > \frac{1}{2}$,

$$\|Rf\|_{L^2(\mathbb{R}^{n-1})} \lesssim \|f\|_{H^s(\mathbb{R}^n)} \tag{132}$$

ii. Show that the result is not true for $s \le 1/2$. Show however that the following sharp trace theorem holds for all $s > 0$,

$$\|Rf\|_{H^s(\mathbb{R}^{n-1})} \lesssim \|f\|_{H^{s+1/2}(\mathbb{R}^n)} \tag{133}$$

iii. Show that f is a function with Fourier support in the ball $|\xi| \lesssim 2^k$ for some integer k then, for all $1 \le p \le \infty$ and $s > 1/p$,

$$\|f\|_{L^p(\mathbb{R}^{n-1})} \lesssim 2^{k/p} \|f\|_{W^{s,p}(\mathbb{R}^n)}$$

Can you deduce from here a trace result, in L^p norms, generalizing that of (132) ? What about (133) ?

iv. Let H be the half space $x_n > 0$. According to the above considerations we can talk about the trace of a function in $W^{1,p}(H)$ to the hyperplane $x_n = 0$(Prove this !). Show that a function $f \in W^{1,p}(H)$ belongs[20] to $W_0^{1,p}(H)$ if and only if its trace to $x_n = 0$ is zero.

Problem 6[Littlewood-Paley] Consider the spaces $\Lambda_\gamma = C^{0,\gamma}(\mathbb{R}^n)$ with norm

$$\|f\|_{\Lambda_\gamma} = \|f\|_{L^\infty(\mathbb{R}^n)} + \sup_{x \ne y \in \mathbb{R}^n} \frac{|u(x) - u(y)|}{|x-y|^\gamma}$$

i. Show, using the Littlewood-Paley projections P_k, that

$$\|f\|_{\Lambda^\gamma} \approx \|P_{\le 0} f\|_{L^\infty} + \sup_{k>0} 2^{k\gamma} \|P_k\|_{L^p}.$$

ii. Define the Zygmund class Λ_* of functions with norm,

$$\|f\|_{\Lambda^*} = \|f\|_{L^\infty} + \sup_{x \in \mathbb{R}^n,\ 0 \le h \le 1} \frac{|f(x+h) + f(x-h) - 2f(x)|}{h}$$

---

[19]You will have to perform the inverse Fourier tarnsform, $u(t,x) = \mathcal{F}^{-1}\hat{u}(t,\xi)$. For the wave equation this is more difficult, in general, but you can do it for dimension $n = 3$.

[20]recall that $W_0^{1,p}(H)$ is the closure of $\mathcal{C}_0^\infty(H)$ in $W^{1,p}(H)$

Show that

$$\|f\|_{\Lambda_*} \approx \|P_{\leq 0} f\|_{L^\infty} + \sup_{k>0} 2^k \|P_k\|_{L^p}.$$

iii. Prove the product estimate in Besov spaces $B^s = H^{s,1}$, $s > 0$.

$$\|fg\|_{B^s} \lesssim \|f\|_{L^\infty} \|g\|_{B^s} + \|g\|_{L^\infty} \|f\|_{B^s}.$$

Problem 7. Read on your own the section on Calderon-Zygmund operators. Indicate how the theory can be extended to operators valued in a given Hilbert space, such as l^2.

11. Restriction Theorems

It is well known that when $f \in L^1(\mathbb{R}^n)$ then its Fourier transform $\hat{f}$ is a bounded and continuous function, thus the restriction of $\hat{f}$ to any hypersurface is perfectly well defined. On the other hand, if $f \in L^2(\mathbb{R}^n)$ then $\hat{f}$ may be any function in L^2, hence defined only almost everywhere and completely arbitrary on sets of measure zero like hypersurfaces.

Can one make sense of the restriction of $\hat{f}$ to a smooth hypersurface S when f belongs to some L^p with $1 < p < 2$? This is a basic question in modern Fourier analysis, which, as we shall see, turns out to be intimately tied to regularity properties of solutions to wave equations.

If we take S to be a hyperplane, we immediately see that the answer is negative. Indeed, let $f(x_1, x') = u(x_1)v(x')$, $\hat{f}(\xi_1, \xi') = \hat{u}(\xi_1)\hat{v}(\xi')$, with $x_1, \xi_1 \in \mathbb{R}$ and $x', \xi' \in \mathbb{R}^{n-1}$. The restriction of $\hat{f}$ to the hyperplane $\xi_1 = 0$ is well defined only when $\hat{u}(0) = \int u(x)\mathrm{d}x$ is well defined. For any $p > 1$ it is always possible to find $u \in L^p(\mathbb{R})$ such that $\int u\mathrm{d}x$ doesn't make sense. We deduce that the restriction of the Fourier transform on hyperplanes cannot be defined when $p > 1$.

The answer is different if we consider hypersurfaces which have non vanishing curvature. For simplicity we consider the model case of the sphere.

11.1. The Stein-Tomas theorem.

The following type of result was first proved by Stein [], then extended by Tomas [] and given its final form again by Stein [].

Theorem 11.2 (Stein-Tomas). *Let $\mathbb{S} = \mathbb{S}^{n-1}$ be the standard unit sphere in $\mathbb{R}^n$ and $d\sigma$ its standard volume element. Let $f \in L^p(\mathbb{R}^n)$ with*

$$1 \leq p \leq \frac{2(n+1)}{n+3}.$$

Then $\mathcal{R}f = \hat{f}\Big|_S \in L^2(\mathbb{S})$ and

$$\|\mathcal{R}f\|_{L^2(\mathbb{S})} \lesssim \|f\|_{L^p(\mathbb{R}^n)}.$$

This theorem has an equivalent dual formulation. Define the Stein operator to be the dual of the Fourier restriction operator $\mathcal{R}f = \hat{f}\Big|_{\mathbb{S}}$,

$$\mathcal{S}g(x) = \mathcal{R}^* g(x) = \int_{\mathbb{S}} e^{ix\cdot\xi} g(\xi) \mathrm{d}\sigma_\xi \simeq (g\mathrm{d}\sigma)^\vee(x),$$

where now g is a function defined on the sphere.

THEOREM 11.3. *Let $f \in L^2(\mathbb{S})$ and*

$$\frac{2(n+1)}{n-1} \leq p \leq \infty.$$

Then $\mathcal{S}f \in L^p(\mathbb{R}^n)$ and

$$\|\mathcal{S}f\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{L^2(\mathbb{S})}. \tag{134}$$

REMARK 11.4. It suffices to prove Theorem 11.3 for $p = p_* = 2(n+1)/(n-1)$. Indeed for $p > p_*$, by Sobolev inequality we have

$$\|\mathcal{S}f\|_{L^p} \lesssim \|D^s \mathcal{S}f\|_{L^{p*}}$$

for $s = n(1/p_* - 1/p) > 0$, where $(D^s u)^\wedge(\xi) = |\xi|^s \hat{u}(\xi)$. But here

$$D^s \mathcal{S}f = \mathcal{S}(|\cdot|^s f) = \mathcal{S}f.$$

Thus, if we can prove the theorem when $p = p_*$ then

$$\|\mathcal{S}f\|_{L^p} \lesssim \|\mathcal{S}f\|_{L^{p*}} \lesssim \|f\|_{L^2(\mathbb{S})}$$

REMARK 11.5. The result remains true if we replace $\mathrm{d}\sigma$ by $\mathrm{d}\mu = \psi\mathrm{d}\sigma$, with $\psi \in C_0^\infty(\mathbb{R}^n)$, since the theorem implies

$$\|(f\mathrm{d}\mu)^\vee\|_{L^p} \lesssim \|f\psi\|_{L^2(\mathbb{S})} \lesssim \|f\|_{L^2(\mathbb{S})}.$$

Moreover, using a partition of unity, it suffices to prove Theorem 11.3 just for $\mathcal{S}f = (f\mathrm{d}\mu)^\vee$, with $\mathrm{d}\mu = \psi\mathrm{d}\sigma$ and $\psi \in C_0^\infty(\mathbb{R}^n)$ supported in a small neighborhood of a point on the sphere. Though obvious, it is a very important fact that we can localize the restriction estimate as we shall see in the future.

11.6. Knapp counterexample. The result of theorem 11.3 is false for any $p < p_*$ in virtue of the following counterexample ([**?**]).

Define, for some small $\delta > 0$, the region in phase space

$$D = \left\{\xi \in \mathbb{R}^n : |\xi_1 - 1| < \delta^2, |\xi'| < \delta\right\}.$$

Let now $f = \chi_{\mathbb{S}\cap D}$ be the characteristic function of the cap $\mathbb{S} \cap D$, then

$$\|f\|_{L^2(\mathbb{S})} = |\mathbb{S} \cap D|^{1/2} \sim \delta^{(n-1)/2}.$$

We can write

$$\mathcal{S}f(x) = e^{ix_1} \int_{\mathbb{S}\cap D} e^{i\phi(x,\xi)} \mathrm{d}\sigma_\xi,$$

with phase $\phi(x, \xi) = x_1(\xi_1 - 1) + x' \cdot \xi'$. It then possible to fix a region in physical space,

$$R = \left\{x \in \mathbb{R}^n : |x_1| < \frac{\pi}{6}\delta^{-2}, |x'| < \frac{\pi}{6}\delta^{-1}\right\},$$

such that for $x \in R$ and $\xi \in D$ we have $|\phi(x,\xi)| \leq \pi/3$, hence, when $x \in R$,

$$|\mathcal{S}f(x)| \geq \mathrm{Re}(e^{-ix_1}\mathcal{S}f(x)) = \int_{\mathbb{S}\cap D} \cos(\phi(x,\xi))\mathrm{d}\sigma_\xi \geq \frac{1}{2}|\mathbb{S}\cap D|.$$

This implies that

$$\frac{\|\mathcal{S}f\|_{L^p}}{\|f\|_{L^2}} \gtrsim |\mathbb{S}\cap D|^{1/2}|R|^{1/p} \sim \delta^{\frac{n-1}{2}-\frac{n+1}{p}}.$$

For small values of δ, an estimate like (134) will necessarily require $\frac{n-1}{2} - \frac{n+1}{p} \geq 0$, which is possible only if $p \geq p_* = 2(n+1)/(n-1)$.

This example suggests that there is some sort of parabolic scaling property in the structure of the operator $\mathcal{S}$ which comes from the nonvanishing curvature of the sphere.

11.7. The importance of curvature. The restriction theorem and its dual counterpart remain true if we replace the standard sphere $\mathbb{S}^{n-1}$ by a compact hypersurface $H \subset \mathbb{R}^n$ with non-vanishing Gauss curvature. The importance of non-vanishing Gauss curvature is illustrated by the following result.

LEMMA 11.8. *Let $H \subset \mathbb{R}^n$ be a compact hypersurface with non-vanishing Gauss curvature (i.e. with all its principal curvatures different from zero) and volume element $d\sigma$. Then, for any smooth function ψ, we have,*

$$|(\psi d\sigma)^\vee(x)| \lesssim (1+|x|)^{-\frac{n-1}{2}} \tag{135}$$

If exactly one principal curvature vanishes then we have instead,

$$|(\psi d\sigma)^\vee(x)| \lesssim (1+|x|)^{-\frac{n-2}{2}}$$

Proof The general proof is based on the method of stationary phase, see Stein's Harmonic Analysis book. For the particular case of the standard sphere $H = \mathbb{S}^{n-1}$ and odd n the proof can be done by a direct computation in polar coordinates. ■

Exercise Prove the lemma for $\mathbb{S}^2 \subset \mathbb{R}^3$.

REMARK 11.9. Another interesting observation links these restiction theorems to partial differential equations. Indeed if $u = \mathrm{d}\sigma^\vee * f$, then u is a solution of the linear elliptic equation

$$\Delta u + u = 0,$$

as we can be easily seen taking the Fourier transform,

$$\mathcal{F}(u+\Delta u)(\xi) \simeq (1-|\xi|^2)\delta(1-|\xi|)\hat{f}(\xi) = 0.$$

where δ is the Dirac distribution.

11.10. TT^* **principle.** The following simple functional analysis result plays an important role in restriction and Strichartz type estimates. Let B be a Banach space and denote by B' its dual. Let H be an Hilbert space with inner product denoted by $\langle \cdot, \cdot \rangle$. Consider a linear operator $T : H \to B'$. Since we can identify H with its dual, we can consider T to be the adjoint of the operator $T^* : B \to H$ defined by

$$\langle h, T^*(x) \rangle = Th(x).$$

Actually, T^* is the adjoint of T when B is reflexive, but for our purposes we shall keep calling T^* the adjoint of T.

The TT^* principle states that the boundedness of T is equivalent to the boundedness of TT^*. More precisely we have:

PROPOSITION 11.11. *The following statements are equivalent:*

(i) $T : H \to B'$ *is bounded and* $\|T\| = M$*;*
(ii) $T^* : B \to H$ *is bounded and* $\|T^*\| = M$*;*
(iii) $TT^* : B \to B'$ *is bounded and* $\|TT^*\| = M^2$*;*
(iv) *the bilinear form* $(x, y) \mapsto \langle T^*x, T^*y \rangle$ *is bounded on* $B \times B$ *with norm* M^2*.*

The proof is a standard exercise in functional analysis.

11.12. TT^* **formulation of the restriction theorem.** The TT^* formulation for the Stein operator corresponds to a convolution with the (inverse) Fourier transform of the measure on the sphere. Formally, we have,

$$SS^*f(x) = SRf(x) = \int_{\mathbb{S}} e^{ix\cdot\xi} \hat{f}(\xi) \mathrm{d}\sigma_\xi = \int_{\mathbb{R}^n} \int_{\mathbb{S}} e^{i(x-y)\cdot\xi} \mathrm{d}\sigma_\xi \, f(y) \mathrm{d}y = \mathrm{d}\sigma^\vee * f(x).$$

We are thus led to the following equivalent form of the restriction theorem,

$$\|\mathrm{d}\sigma^\vee * f\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{L^{p'}(\mathbb{R}^n)}, \tag{136}$$

for $p \geq p^*$.

One can give three distinct proofs of Theorem 11.3. We shall sketch the first proof based on analytic interpolation. This is essentially the original proof of Stein and Tomas. The second proof, based on introducing a time parameter and treating $\mathcal{S}f$ as an evolution operator allows us to regard the restriction theorem as part of a more general framework which includes Strichartz estimates for various linear PDE with constant coefficients. Finally the third approach, which only applies for specific exponents, will allow us to to connect with bilinear estimates.

11.13. First proof: analytic interpolation. According to Remark 11.12 and Remark 11.4 it suffices to prove that $Uf = \mathrm{d}\sigma^\vee * f$ verifies

$$\|Uf\|_{L^{p_*}(\mathbb{R}^n)} \lesssim \|f\|_{L^{p'_*}(\mathbf{R}^n)}, \tag{137}$$

where $p_* = 2(n+1)/(n-1)$ and $p'_* = 2(n+1)/(n+3)$.

In general, to obtain $L^{p'} - L^p$ estimates directly is usually very complicated and we don't know any direct proof except in cases where p is a nice exponent like $p = 4, 6$ (which happens only for $n = 2$ or $n = 3$). We would feel more comfortable with $L^2 - L^2$ type estimates, where Plancherel's theorem is a powerful tool, or with $L^1 - L^\infty$ type estimates, since pointwise decay estimates of oscillatory integrals can be obtained from stationary phase methods. This suggests to use some interpolation theory for L^p spaces. But, an $L^2 - L^2$ estimate for the operator U is ruled out by the Knapp counterexample and a $L^\infty - L^1$ one is too trivial and doesn't answer to our question. It is here that the Stein interpolation theorem, Thm. 4.8, shows its power, since it allows us to obtain the $L^{p'} - L^p$ estimate for U from $L^2 - L^2$ and $L^\infty - L^1$ estimates for other (reasonable) operators different from U.

We will accomplish this by constructing a family of convolution operators $U_z f = \mu_z^\vee * f$, with μ_z being distributions depending analytically in z. The parameter z will essentially reflect the degree of homogeneity of the distribution μ_z. For this reason it is natural to place our target at $z = -1$, requiring $U_{-1} = U$ or $\mu_{-1} = \mathrm{d}\sigma$, since $\mathrm{d}\sigma$ can be written as the pullback of a delta distribution (which is homogeneous of degree -1) on the sphere: $\mathrm{d}\sigma \simeq \delta(1 - |\xi|)\mathrm{d}\xi$.

An $L^2 - L^2$ estimate for U_z will follow if μ_z coincides with a bounded function, indeed, by Plancherel's theorem, we have

$$\|U_z f\|_{L^2} \simeq \|(U_z f)^\wedge\|_{L^2} \simeq \left\|\mu_z \cdot \hat{f}\right\|_{L^2} \lesssim \|\mu_z\|_{L^\infty} \|f\|_{L^2} . \tag{138}$$

To have $\mu_z(\xi)$ bounded we must require that $\mu_z(\xi)$ is essentially homogeneous of degree 0, hence when z lies on the line $\mathrm{Re}(z) = 0$.

An $L^1 - L^\infty$ estimate for U_z will follow instead when $\mu_z^\vee$ coincides with a bounded function, since we directly have

$$\|U_z f\|_{L^\infty} \lesssim \|\mu_z^\vee\|_{L^\infty} \|f\|_{L^1} . \tag{139}$$

To obtain (137) from the analytic interpolation of (138) and (139), we would like the latter to happen on the line $\mathrm{Re}(z) = a$, where a is chosen so that

$$-1 = \theta a + (1-\theta)0, \quad \frac{1}{p_*} = \frac{\theta}{\infty} + \frac{1-\theta}{2}, \quad \frac{1}{p'_*} = \frac{\theta}{1} + \frac{1-\theta}{2},$$

and this happens precisely when $\mathrm{Re}(z) = a = -(n+1)/2$.

This argument leads to the precise version of the Stein analytic interpolation theorem that we are going to use.

PROPOSITION 11.14. *Let U_z be an analytic family of linear operators such that:*

(i) $U_{-1} = U$;
(ii) $\|U_z f\|_{L^2} \lesssim \|f\|_{L^2}$, *uniformly on the line Re*$(z) = 0$;
(iii) $\|U_z f\|_{L^\infty} \lesssim \|f\|_{L^1}$, *uniformly on the line Re*$(z) = -(n+1)/2$.

Then it follows that

$$\|Uf\|_{L^{p_*}} \lesssim \|f\|_{L^{p'_*}} .$$

The above discussion showed that, when we write $U_z f$ as the convolution $\mu_z^\vee * f$, then the hypothesis of the proposition are fulfilled whenever μ_z is an analytic family of distribution such that

(i') $\mu_{-1} = \mathrm{d}\sigma$;
(ii') $\mu_z(\xi)$ coincides with a bounded function, with a uniform bound on the line $\mathrm{Re}(z) = 0$;
(iii') $\mu_z^\vee(x)$ coincides with a bounded function, with a uniform bound on the line $\mathrm{Re}(z) = -(n+1)/2$.

It thus remains to define the distributions μ_z and verify these properties.

Inspired by the identity $\delta = \chi_+^{-1}$ and $\mathrm{d}\sigma_\xi \simeq \delta(1 - |\xi|)$, we define our family of distributions as

$$\mu_z(\xi) = e^{z^2} \chi_+^z (1 - |\xi|) \psi(|\xi|), \tag{140}$$

where $\psi \in C_0^\infty(\mathbf{R})$ is a cut-off function supported in a small neighborhood of 1, say $[1/2, 3/2]$, and $\psi(1) = 1$.

We recall that the homogeneous distributions χ_+^z, when $\mathrm{Re}(z) > -1$, coincide with the functions:

$$\chi_+^z(t) = \begin{cases} t^z/\Gamma(z+1) & \text{if } t \geq 0, \\ 0 & \text{if } t < 0, \end{cases}$$

where the Gamma function is defined by $\Gamma(z+1) = \int_0^\infty t^z e^{-t} \mathrm{d}t$. From the identity $\Gamma(z+1) = z\Gamma(z)$, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t} \chi_+^z(t) = \chi_+^{z-1}(t). \tag{141}$$

Using this formula, χ_+^z can be analytically continued for all $z \in \mathbb{C}$ by performing repeated integrations by parts. To do this we first observe that for $\mathrm{Re}(z) > -1$ and $\phi \in C_0^\infty$ we have

$$\int \chi_+^z(t)\phi(t)\mathrm{d}t = -\int \chi_+^{z+1}(t)\phi'(t)\mathrm{d}t = \ldots = (-1)^m \int \chi_+^{z+m}(t)\phi^{(m)}(t)\mathrm{d}t.$$

Thus integrating by parts sufficiently many times we can make sense of $\int \chi_+^z \phi \mathrm{d}t$ when $\mathrm{Re}(z) > -1 - m$ for any m, and hence for all z. To see that $\chi_+^{-1} = \delta$ it takes just an integration by parts, indeed

$$\int \chi_+^{-1} \phi \mathrm{d}t = -\frac{1}{\Gamma(1)} \int_0^\infty \phi'(t)\mathrm{d}t = \phi(0).$$

For more information about χ_+^z and distribution theory one can consult the books by Gel'fand and Shilov [**4**] or Hormander [**5**].

The factor e^{z^2} in the definition of μ_z is chosen in order to garantee a uniform boundedness of our operators for large $\mathrm{Im}(z)$, indeed e^{z^2} decreases exponentially as $\mathrm{Im}(z) \to \infty$, uniformly on the strip $-(n+1)/2 \leq \mathrm{Re}(z) \leq 0$. This permits to allow the various constants in the following inequalities to have a polynomial growth in terms of $b = \mathrm{Im}(z)$.

Clearly $\mu_{-1} \simeq \delta(1-|\xi|)\psi(|\xi|) \simeq \mathrm{d}\sigma$. This verifies (i').

Condition (ii') is immediately verified, since χ_+^{-z} is always a bounded function when $\mathrm{Re}(z) = 0$. Condition (iii') will follow from stationary phase arguments, more generally we have:

Proposition 11.15.

$$|\mu_z^\vee(x)| \lesssim (1+|x|)^{-Re(z)-1-\frac{n-1}{2}}. \tag{142}$$

11.16. Second proof: evolution operators approach. In this section we make the following assumption on f:

$$f \in C^\infty(\mathbb{S}), \qquad \mathrm{supp} f \subset \{\xi_1 > 1/2\}. \tag{143}$$

With this assumption we can relabel $x_1 = t$ as a time parameter and rewrite $\mathcal{S}f$ as

$$\begin{aligned}\mathcal{S}f(t,x') &= \int_{|\xi'|<\sqrt{3}/2} e^{it\sqrt{1-|\xi'|^2}} e^{ix'\cdot\xi'} f(\sqrt{1-|\xi'|^2},\xi') \frac{\mathrm{d}\xi'}{\sqrt{1-|\xi'|^2}} \\ &= \int e^{it\sqrt{1-|\xi'|^2}} e^{ix'\cdot\xi'} \beta(|\xi'|) g(\xi') \mathrm{d}\xi'.\end{aligned}$$

with $\beta \in C_0^\infty$ supported in $|\xi'| < 1$ and $g(\xi') = f(\sqrt{1-|\xi'|^2},\xi')/\sqrt{1-|\xi'|^2}$. Observe that

$$\int |g(\xi')|^2 \mathrm{d}\xi' = \int_{\mathbb{S}} \frac{|f(\xi)|^2}{|\xi_1|^2} \mathrm{d}\sigma_\xi \simeq \|f\|^2_{L^2(\mathbb{S})}$$

by the assumption on the support of f.

Theorem 11.17. *Let $\beta \in C_0^\infty(\mathbf{R}^{n-1})$ be supported in the unit ball $\{\xi \in \mathbf{R}^{n-1} : |\xi| < 1\}$ and consider the operator*

$$Tg(t,x) = \int_{\mathbf{R}^{n-1}} e^{it\sqrt{1-|\xi|^2}} e^{ix\cdot\xi} \beta(\xi) g(\xi) d\xi, \quad t \in \mathbf{R},\ x \in \mathbf{R}^{n-1}.$$

Let q, r be Lebesgue exponents verifing the conditions:

$$0 \le \frac{2}{q} \le \min\{1, \gamma(r)\}, \tag{144}$$

$$\left(\frac{2}{q}, \gamma(r)\right) \ne (1,1), \tag{145}$$

where $\gamma(r) = (n-1)(1/2 - 1/r)$. Then the following estimate holds true for all $g \in C_0^\infty(\mathbf{R}^{n-1})$,

$$\|Tg\|_{L_t^q L_x^r(\mathbf{R}\times\mathbf{R}^{n-1})} \lesssim \|g\|_{L^2(\mathbf{R}^{n-1})}. \tag{146}$$

where we use the mixed norm notation defined in section 7.

By Remark 11.5, Theorem 11.3 follows from the special case $q = r = 2\frac{n+1}{n-1}$.

Remark 11.18. We can run again the Knapp example to prove the necessity of condition (144), when $q \ge 2$. Indeed let $D \subset \mathbb{R}^{n-1}$ be the disk defined by $|\xi| \le \delta$, for sufficiently small $\delta > 0$, and take $g = \chi_D$ to be the characteristic function of D. We write,

$$Tg(t,x) \quad = \quad e^{it}\int_D e^{it(\sqrt{1-|\xi|^2}-1)}e^{ix\cdot\xi}\beta(\xi)d\xi$$

and observe that for $|t| \leq \delta^{-2}$ and $|x| \leq \delta^{-1}$ we have, with a fixed constant $c > 0$, $|Tg(t,x)| \geq c$. Indeed this follows easily from $\xi| \leq \delta$ and $|\sqrt{1-|\xi|^2}-1| \lesssim \delta^2$. Therefore, if (146) holds true, we must have, for all sufficiently small $\delta > 0$,

$$c\delta^{-\frac{2}{q}}\delta^{-\frac{n-1}{r}} \lesssim \|Tg\|_{L^q_tL^r_x} \lesssim \|\chi_D\|_{L^2} \lesssim \delta^{-\frac{n-1}{2}}$$

from which (144), $q \geq 2$ follows.

REMARK 11.19. The end-point restriction (145) can be removed when $n \neq 3$, due to a well known result by Keel and Tao [**14**] ("Endpoint Strichartz Inequalities"). The other restriction $q \geq 2$, implicit in (144) will be discussed in the next chapter.

We start by calculating T^* and TT^*.

$$\begin{aligned}< T^*F, g > =< F, Tg > &= \iint F\overline{Tg}\mathrm{d}t\mathrm{d}x = \\ &= \iint F(t,x)\int e^{-it\sqrt{1-|\xi|^2}}e^{-ix\cdot\xi}\overline{\beta(\xi)g(\xi)}\mathrm{d}\xi\mathrm{d}t\mathrm{d}x = \\ &= \int \overline{g(\xi)\beta(\xi)}\left(\iint e^{-it\sqrt{1-|\xi|^2}}e^{-ix\cdot\xi}F(t,x)\mathrm{d}t\mathrm{d}x\right)\mathrm{d}\xi.\end{aligned}$$

Hence

$$T^*F(\xi) = \overline{\beta(\xi)}\iint e^{-it\sqrt{1-|\xi|^2}}e^{-ix\cdot\xi}F(t,x)\mathrm{d}t\mathrm{d}x,$$

and

$$\begin{aligned}TT^*F(t,x) &= \int e^{it\sqrt{1-|\xi|^2}}e^{ix\cdot\xi}\beta(\xi)T^*F(\xi)\mathrm{d}\xi \\ &= \iint e^{i(t-s)\sqrt{1-|\xi|^2}}e^{ix\cdot\xi}|\beta(\xi)|^2\hat{F}(s,\xi)\mathrm{d}\xi\mathrm{d}s,\end{aligned}$$

where $\hat{F}(s,\xi) = \int e^{-ix\cdot\xi}F(s,x)\mathrm{d}x$. If we introduce the family of operators

$$U(t)f(x) = \int e^{it\sqrt{1-|\xi|^2}}e^{ix\cdot\xi}|\beta(\xi)|^2\hat{f}(\xi)\mathrm{d}\xi,$$

we can write TT^* as a convolution operator,

$$TT^*F(t,\cdot) = \int U(t-s)F(s,\cdot)\mathrm{d}s. \tag{147}$$

By Proposition 11.11, to show that T is a bounded operator from $L^q_tL^r_x(\mathbf{R}^n)$ to $L^2(\mathbf{R}^{n-1})$ it suffices to prove that TT^* is a bounded operator from $L^{q'}_tL^{r'}_x(\mathbf{R}^n)$ to $L^q_tL^r_x(\mathbf{R}^n)$.

We shall first prove an estimate for $U(t)$.

PROPOSITION 11.20. *Let $2 \leq r \leq \infty$ and $\gamma(r) = (n-1)(1/2 - 1/r)$. Then $U(t)$ verifies the estimate*

$$\|U(t)f\|_{L^r(\mathbf{R}^{n-1})} \lesssim (1+|t|)^{-\gamma(r)}\|f\|_{L^{r'}(\mathbf{R}^{n-1})}. \tag{148}$$

Proof Once we have proved the two extreme cases $r = 2$ and $r = \infty$,

$$\|U(t)f\|_{L^2(\mathbf{R}^{n-1})} \lesssim \|f\|_{L^2(\mathbf{R}^{n-1})} \tag{149}$$

$$\|U(t)f\|_{L^\infty(\mathbf{R}^{n-1})} \lesssim (1+|t|)^{-(n-1)/2} \|f\|_{L^1(\mathbf{R}^{n-1})} \tag{150}$$

then the estimate follows from the standard Riesz interpolation theorem.

We obtain (149) immediately using Plancherel formula, since

$$(U(t)f)^\wedge(\xi) \simeq e^{it\sqrt{1-|\xi|^2}}|\beta(\xi)|^2\hat{f}(\xi).$$

To prove (150) we write

$$U(t)f(x) = \int K_t(x-y)f(y)\mathrm{d}y,$$

where

$$\begin{aligned} K_t(x) &= \int e^{ix\cdot\xi}e^{it\sqrt{1-|\xi|^2}}|\beta(\xi)|^2\mathrm{d}\xi \\ &\simeq \iint e^{ix\cdot\xi}e^{it\tau}\delta(1-\tau^2-|\xi|^2)\sqrt{1-|\xi|^2}|\beta(\xi)|^2\mathrm{d}\tau\mathrm{d}\xi \\ &\simeq \iint e^{i(t,x)\cdot(\tau,\xi)}\delta(1-|(\tau,\xi)|)\beta_1(\tau,\xi)\mathrm{d}\tau\mathrm{d}\xi, \qquad \left(\beta_1(\tau,\xi) = \tau|\beta(\xi)|^2\right), \\ &= (\beta_1\mathrm{d}\sigma_{n-1})^\vee(t,x). \end{aligned}$$

Hence K_t is just the Fourier transform of a measure supported on the sphere $\mathbb{S}^{n-1}$, for which we have the decay estimate

$$|K_t(x)| \lesssim (1+|t|+|x|)^{-(n-1)/2},$$

which implies (150). ∎

We next apply Proposition 11.20 to (147),

$$\|TT^*F(t,\cdot)\|_{L^r_x} \lesssim \int \frac{1}{(1+|t-s|)^{\gamma(r)}}\|F(s,\cdot)\|_{L^{r'}_x}\mathrm{d}s. \tag{151}$$

Finally, we are in a position to apply the Hardy-Littlewood-Sobolev inequality and, if $0<\gamma(r)<1$, we obtain

$$\|TT^*F\|_{L^q_tL^r_x} \lesssim \|F\|_{L^{q'}_tL^{r'}_x},$$

when $-\gamma(r)+1+1/q = 1/q'$, hence $\gamma(r) = 2/q$. Therefore we proved Theorem 11.17 in the case $0<\gamma(r)=2/q<1$.

On the other hand if $q=2$ and $\gamma(r)>1$ we have from (151),

$$\|TT^*F\|_{L^2_tL^r_x} \lesssim \|F\|_{L^2_tL^{r'}_x},$$

by an application of the standard Hausdorff-Young inequality.

Finally, if $2/q<1$ and $\gamma(r)>2/q$ the result follows from the case $\gamma(r)=2/q$ using Sobolev inequalities.

11.21. Third proof: bilinear forms ($n = 2$ and $n = 3$). We present now another method to prove the restriction theorem for the sphere that works for the special cases $n = 2$, $p = 6$ or $n = 3$, $p = 4$. The idea is that when p is an even integer, the restriction theorem can be viewed as an L^2 estimate for a multilinear form, which, through the Fourier transform, has a convolution structure that provides some smoothing effects. The proofs given below are at the root of the so called bilinear trilinear estimates, which play a fundamental role in the modern theory of nonlinear wave and dispersive equations.

Let us see the case $n = 3$ first. We consider the Stein operator $\mathcal{S}f = (f\mathrm{d}\sigma)^\vee$, and use the fact that $(\mathcal{S}f \cdot \mathcal{S}f)^\wedge \simeq (f\mathrm{d}\sigma) * (f\mathrm{d}\sigma)$. Let $B(f,g) = \mathcal{S}f \cdot Sg$, then an L^4 estimate for $\mathcal{S}f$ corresponds to an L^2 estimate for $B(f,f)$. We have

$$\hat{B}(f,g)(\xi) \simeq (f\mathrm{d}\sigma) * (g\mathrm{d}\sigma)(\xi) = \int_{\mathbf{R}^3} \delta(1-|\xi-\eta|)\delta(1-|\eta|)f(\xi-\eta)g(\eta)\mathrm{d}\eta,$$

and applying Cauchy-Schwarz with respect to the measure $\delta(1-|\xi-\eta|)\delta(1-|\eta|)\mathrm{d}\eta$ we find

$$|\hat{B}(f,g)(\xi)|^2 \leq \hat{B}(1,1)(\xi)\hat{B}\left(|f|^2,|g|^2\right)(\xi).$$

Integrating with respect to ξ, we obtain

$$\|B(f,g)\|^2_{L^2(\mathbf{R}^3)} \lesssim A\,\|f\|^2_{L^2(\mathbb{S}^2)}\,\|g\|^2_{L^2(\mathbb{S}^2)}\,, \tag{152}$$

with

$$A = \sup_\xi |\hat{B}(1,1)(\xi)| = \sup_\xi \int \delta(1-|\xi-\eta|)\delta(1-|\eta|)\mathrm{d}\eta. \tag{153}$$

Thus, to prove the theorem in this case it suffices to check that A is finite. It is useful to carry out the explicit calculation of $A(\xi) = \hat{B}(1,1)(\xi)$. For any dimension $n \geq 2$ we have:

LEMMA 11.22.

$$A(\xi) = \int_{\mathbf{R}^n} \delta(1-|\xi-\eta|)\delta(1-|\eta|)d\eta \simeq \frac{1}{|\xi|}\left(4-|\xi|^2\right)_+^{\frac{n-3}{2}}. \tag{154}$$

Proof

$$A(\xi) = \int \delta(1-|\xi-\eta|)\delta(1-|\eta|)\mathrm{d}\eta \simeq \int_{|\eta|=1} \delta(1-|\xi-\eta|^2)\mathrm{d}\sigma_\eta =$$
$$= \int_{|\eta|=1} \delta(|\xi|^2 - 2\xi\cdot\eta)\mathrm{d}\sigma_\eta \simeq \frac{1}{|\xi|}\int_{|\eta|=1} \delta\left(\frac{|\xi|}{2} - \frac{\xi}{|\xi|}\cdot\eta\right)\mathrm{d}\sigma_\eta.$$

Because of the rotational symmetry, we may assume that $\xi = (|\xi|, 0, \ldots, 0)$, so that

$$A(\xi) \simeq \frac{1}{|\xi|}\int_0^\pi \delta\left(\frac{|\xi|}{2} - \cos\theta\right)(\sin\theta)^{n-2}\mathrm{d}\theta =$$
$$= \frac{1}{|\xi|}\int_{-1}^1 \delta\left(\frac{|\xi|}{2} - u\right)(1-u^2)^{\frac{n-3}{2}}\mathrm{d}u = \frac{1}{|\xi|}\left(1 - \frac{|\xi|^2}{4}\right)^{\frac{n-3}{2}},$$

when $|\xi|/2 \in [-1,1]$. ■

When $n=3$, $A(\xi)\simeq 1/|\xi|$ is singular only at $\xi=0$, but we can avoid this difficulty by assuming that f and g are supported in a small neighborhood of a point in S^2 (recall that without loss of generality we can localize the estimate on a small cap on the sphere). Then the supremum in (153) can be taken over just all $\xi\in \operatorname{supp}(f)+\operatorname{supp}(g)$, which is a set bounded away from 0. Hence we may restrict to $|\xi|\geq C>0$ in (153) and the singularity disappears leaving $A<\infty$.

From the L^2 estimate (152) of the bilinear form $B(f,g)$, it follows the L^4 estimate for the Stein operator $\mathcal{S}f$:

$$\|\mathcal{S}f\|^2_{L^4(\mathbf{R}^3)}=\|B(f,f)\|_{L^2}\simeq A^{1/2}\,\|f\|^2_{L^2(\mathbb{S}^2)}\,,$$

with the assumption that f is supported in a small cap on the sphere.

In the case $n=2$ what we want is an L^6 estimate for $\mathcal{S}f$. Since $6=3\times 2$ we can try to repeat the same calculation using this time a trilinear form, $T(f,g,h)=\mathcal{S}f\cdot Sg\cdot Sh$, and the fact that $\|\mathcal{S}f\|^3_{L^6}=\|T(f,f,f)\|_{L^2}$. We have

$$\begin{aligned}\hat{T}(f,g,h)(\xi)&\simeq(f\mathrm{d}\sigma)*(g\mathrm{d}\sigma)*(h\mathrm{d}\sigma)(\xi)=\\&=\iint_{\mathbb{R}^2\times\mathbb{R}^2}\delta(1-|\xi-\eta-\zeta|)\delta(1-|\eta|)\delta(1-|\zeta|)f(\xi-\eta)g(\eta)h(\zeta)\mathrm{d}\eta\mathrm{d}\zeta,\end{aligned}$$

and applying Cauchy-Schwarz with respect to the measure $\delta(1-|\xi-\eta|)\delta(1-|\eta|)\delta(1-|\zeta|)\mathrm{d}\eta\mathrm{d}\zeta$ we find

$$|\hat{T}(f,g,h)(\xi)|^2\leq\hat{T}(1,1,1)(\xi)\hat{T}\left(|f|^2,|g|^2,|h|^2\right)(\xi).$$

Integrating with respect to ξ, we obtain

$$\|T(f,g,h)\|^2_{L^2(\mathbb{R}^2)}\lesssim A\,\|f\|^2_{L^2(\mathbb{S}^1)}\,\|g\|^2_{L^2(\mathbb{S}^1)}\,\|h\|^2_{L^2(\mathbb{S}^1)}\,,\tag{155}$$

with

$$A=\sup_\xi|\hat{T}(1,1,1)(\xi)|=\sup_\xi\iint\delta(1-|\xi-\eta|)\delta(1-|\eta|)\delta(1-|\zeta|)\mathrm{d}\eta\mathrm{d}\zeta.\tag{156}$$

The convolution structure allows us to restrict ξ to the set $\operatorname{supp}f+\operatorname{supp}g+\operatorname{supp}h$, and, if we make the hypothesis of f,g,h supported in a small cap of the sphere, we can assume $1\leq|\xi|\leq 3$. Using Lemma 11.22 we can evaluate $T(1,1,1)$ and show that A is bounded,

$$\begin{aligned}T(1,1,1)(\xi)&=\int B(1,1)(\xi-\zeta)\,\delta(1-|\zeta|)\mathrm{d}\zeta\sim\\&\sim\int_{|\xi-\zeta|<2}\frac{\delta(1-|\zeta|)}{(4-|\xi-\zeta|^2)^{1/2}}\mathrm{d}\zeta=\int_{\substack{\zeta\in\mathbb{S}^1\\|\xi-\zeta|<2}}\frac{\mathrm{d}\sigma_\zeta}{(3-2\xi\cdot\zeta+|\xi|^2)^{1/2}}\simeq\\&\simeq\int_{a(\xi)}^1\frac{\mathrm{d}a}{(3-|\xi|^2+2|\xi|a)^{1/2}(1-a^2)^{1/2}}\sim\int_{a(\xi)}^1\frac{\mathrm{d}a}{(a-a(\xi))^{1/2}(1-a)^{1/2}}\simeq 1,\end{aligned}$$

where $a(\xi)=-\frac{3-|\xi|^2}{2|\xi|}$. From the L^2 estimate (155) of the trilinear form $T(f,g,h)$, it follows the L^6 estimate for the Stein operator $\mathcal{S}f$:

$$\|\mathcal{S}f\|^3_{L^6(\mathbb{R}^2)}=\|T(f,f,f)\|_{L^2}\simeq A^{1/2}\,\|f\|^3_{L^2(\mathbb{S}^2)}\,.$$

We can also try to repeat the bilinear argument for $n = 2$. As before, for $B(f, g) = \mathcal{S}f \cdot Sg$ we have

$$|\hat{B}(f,g)(\xi)|^2 \le \hat{B}(1,1)(\xi)\hat{B}\left(|f|^2, |g|^2\right)(\xi).$$

Integrate with respect to ξ, and use Lemma 11.22 to evaluate $\hat{B}(1,1)$,

$$\|B(f,g)\|^2_{L^2(\mathbb{R}^2)} \lesssim \iint \frac{\delta(1-|\xi-\eta|)\delta(1-|\eta|)}{|\xi|(4-|\xi|^2)^{1/2}} |f(\xi-\eta)|^2 |g(\eta)|^2 \mathrm{d}\eta \mathrm{d}\xi.$$

Change variable, $\xi \to \zeta = \xi - \eta$, and observe that when $|\eta| = |\zeta| = 1$ we have

$$|\xi| = |\eta + \zeta| \simeq (1 + \eta \cdot \zeta)^{1/2},$$
$$(4 - |\xi|^2)^{1/2} = (4 - |\eta + \zeta|^2)^{1/2} \simeq (1 - \eta \cdot \zeta)^{1/2},$$

hence

$$\|B(f,g)\|^2_{L^2(\mathbb{R}^2)} \lesssim \iint_{\mathbb{S}^1 \times \mathbb{S}^1} \frac{|f(\zeta)|^2 |g(\eta)|^2}{\left(1 - (\eta \cdot \zeta)^2\right)^{1/2}} \mathrm{d}\sigma_\eta \mathrm{d}\sigma_\zeta. \tag{157}$$

This is an interesting formula. Observe that if the supports of f and g on $\mathbb{S}^1$ are projectionally disjoint, i.e. don't contain points in the same direction, then the quantity $1 - (\eta \cdot \zeta)^2$ is bounded below by a positive constant and in this case we obtain the bilinear restriction estimate

$$\|B(f,g)\|_{L^2(\mathbb{R}^2)} \lesssim \|f\|_{L^2(\mathbb{S}^1)} \|g\|_{L^2(\mathbb{S}^1)}.$$

We can consider also other types of bilinear forms which have a special structure that cancel the singularity in the denominator. Take for example $Q(f, g) = \partial_1 \mathcal{S}f \partial_2 Sg - \partial_2 \mathcal{S}f \partial_1 Sg$, then taking the Fourier transform and proceeding as before we see that

$$\begin{aligned} \|Q(f,g)\|^2_{L^2(\mathbb{R}^2)} &\lesssim \iint_{\mathbb{S}^1 \times \mathbb{S}^1} \frac{|\eta_1\zeta_2 - \eta_2\zeta_1|^2}{\left(1 - (\eta \cdot \zeta)^2\right)^{1/2}} |f(\zeta)|^2 |g(\eta)|^2 \mathrm{d}\sigma_\eta \mathrm{d}\sigma_\zeta \\ &\lesssim \|f\|^2_{L^2(\mathbb{S}^1)} \|g\|^2_{L^2(\mathbb{S}^1)}, \end{aligned}$$

since we have the identity $|\eta_1\zeta_2 - \eta_2\zeta_1|^2 = 1 - (\eta \cdot \zeta)^2 \le 1$.

12. Strichartz inequalities for the wave equation

Strichartz inequalities are an important tool in the study of linear and nonlinear wave equations. They are intimately tied to restriction theorems. In this chapter we shall only consider the case of the standard linear wave equation. Similar inequalities hold true however for linear dispersive equations such as the Schrödinger, linear KdV etc.

12.0.1. *Homogeneous wave equation.* Consider solutions $u = u(t, x), t \in \mathbb{R}, x \in \mathbb{R}^n$ to the equation

$$\Box u = F, \tag{158}$$
$$u(0,x) = f(x), \quad \partial_t u(0,x) = g(x), \tag{159}$$

with $\Box$ the wave operator $\Box = -\partial_t^2 u + \Delta$. Clearly, a solution to eqrefeq:genwave can be written as a superposition between a solution to the homogeneous wave equation,

$$\Box u = 0, \tag{160}$$

verifying the initial condition (159) at time $t = 0$, and a solution to the purely inhomogeneous wave equation

$$\Box u = F, \tag{161}$$

with zero initial data

$$u(0,x) = 0, \quad \partial_t u(x,0) = 0.$$

We denote by $W(t)h$ the fundamental solution of the homogeneous problem (160), i.e. $u(t,x) = \big(W(t)h\big)(x)$ is the unique solution of (160) which verifies the initial conditions

$$u(0,x) = 0, \quad \partial_t u(0,x) = h(x)$$

By Duhamel's principle any solution of the inhomogeneous equation can itself be written as a superposition of solutions to the homogeneous equation according to the formula,

$$u(t) = \int_0^t W(t-t')F(t')\mathrm{d}t', \tag{162}$$

Before stating the main result of this section we make the following definition.

DEFINITION 12.1. We say that the pair of real numbers (q,r) is an admissible wave pair if they satisfy the conditions

$$\begin{array}{rcl} q & \geq & 2, \\ \dfrac{2}{q} & \leq & (n-1)\left(\dfrac{1}{2} - \dfrac{1}{r}\right), \\ (q,r,n) & \neq & (2,\infty,3). \end{array}$$

We are now ready to state the following.

THEOREM 12.2. *Suppose that $n \geq 2$ and (q,r) is a wave admissible pair* [21] *with $r < \infty$.*

(1) *Assume the dimensional condition, $\frac{1}{q} + \frac{n}{r} = \frac{n}{2} - \gamma$. Then, if u verifies the homogeneous equation* (160) *with initial conditions* (159)*,*

$$\|u\|_{L_t^q L_x^r} + \|u\|_{L_t^\infty \dot H_x^\gamma} + \|\partial_t u\|_{L_t^\infty \dot H_x^{\gamma-1}} \quad \lesssim \quad \|f\|_{\dot H^\gamma} + \|g\|_{\dot H^{\gamma-1}} \tag{163}$$

(2) *Assume the dimensional condition,*[22] *$\frac{1}{q} + \frac{n}{r} = \frac{n}{2} - \gamma = \frac{1}{q'} + \frac{n}{r'} - 2$, with q' dual to q and r' dual to r. Then, if u verifies the purely inhomogeneous problem* (12.0.1) *with zero initial conditions, then on a finite time interval $[0,T]$:*

$$\|u\|_{L^q([0,T];L^r)} + \|u\|_{C([0,T];\dot H^\gamma)} + \|\partial_t u\|_{C([0,T];\dot H^{\gamma-1})} \quad \lesssim \quad \|F\|_{L^{q'}([0,T];L^{r'})} \tag{164}$$

[21] The case when $r = \infty$ can also be included provided that we modify the spaces on the left of the estimates below to appropriate Besov spaces.

[22] Thus, in fact, $\gamma = 1/2$.

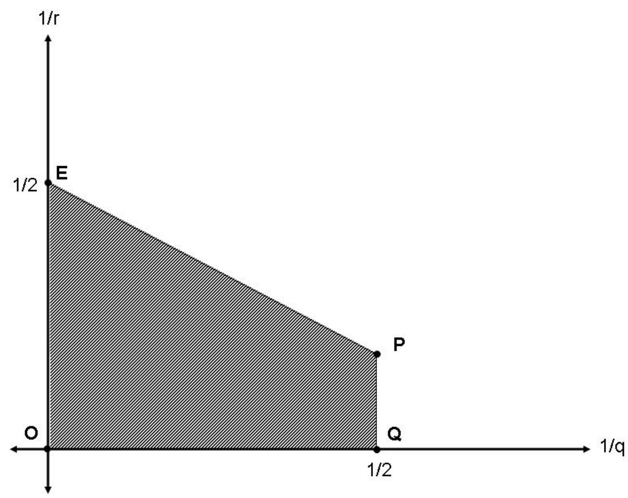

FIGURE 2. Admissable exponents for $n \geq 4$

(3) *We also have the following more general version of* (164) *for admissible pairs* (q_1, r_1), (q_2, r_2) *with* $r_1, r_2 < \infty$ *verifying the dimensional condition,*

$$\frac{1}{q_1} + \frac{n}{r_1} = \frac{n}{2} - \gamma = \frac{1}{q_2'} + \frac{n}{r_2'} - 2$$

Then,

$$\|u\|_{L^{q_1}([0,T];L^{r_1})} + \|u\|_{C([0,T];\dot{H}^\gamma)} + \|\partial_t u\|_{C([0,T];\dot{H}^{\gamma-1})} \lesssim \|F\|_{L^{q_2'}([0,T];L^{r_2'})} \tag{165}$$

REMARK 12.3. For $n \geq 4$, the region of admissable exponents corresponds to a quadrilateral $OEPQ$ in the plane $(1/q, 1/r)$ with vertices $O = (1/\infty, 1/\infty)$, $E = (1/\infty, 1/2)$, $P = (1/2, \frac{n-3}{2(n-1)})$ and $Q = (1/2, 1/\infty)$. When $n = 3$ the point P coincides with Q and the region reduces to the triangle OEQ. When $n = 2$ we have a smaller triangle OEQ_2 where $Q_2 = (1/4, 1/\infty)$.

For $n = 3$, the boundary of the triangular region is allowed except for the endpoint P. For $n \geq 4$, the boundary of the quadrilateral region is entirely allowed, as we will note below.

The interesting cases are the ones on the segment EP and the ones on PQ close to P, since all the others can be deduced from these using Sobolev embeddings. The point E corresponds to the energy estimates. There are counterexamples that exclude the point P when $n = 3$, while the inclusion of P in higher dimensions were recently obtained by Keel and Tao [**14**].

The standard Strichartz estimate[23] corresponds to the point $S = (\frac{n-1}{2(n+1)}, \frac{n-1}{2(n+1)})$.

REMARK 12.4. We remark that in even though the end-point case $n = 3, q = \infty, r = 2$ is forbidden, the estimates holds in the spherically symmetric case. Indeed let ϕ be a solution of the homogeneous wave equation $\Box\phi = 0$ in $\mathbb{R}^{3+1}$ subject to the initial conditions

$$\phi(0, x) = 0, \quad \partial_t\phi(0, x) = f(x)$$

and assume that f is spherically symmetric i.e. $f(x) = f(|x|)$. Then,

$$\int_0^\infty \|\phi(t, \cdot)\|^2_{L^\infty(\mathbb{R}^3)} dt \le c\|f\|^2_{L^2}. \tag{166}$$

The proof is an immediate consequence of the Hardy–Littlewood maximal theorem[24] in view of the fact that, for spherically symmetric f,

$$\phi(x, t) = \frac{c}{|x|} \int_{\left||x|-t\right|}^{|x|+t} \lambda f(\lambda)\, d\lambda.$$

REMARK 12.5. We give an elementary example below to illustrate how the end point result $n = 3, q = \infty, r = 2$ fails in the general case due to possible concentrations along null rays. We show below that there exists a sequence of functions f_n in $C_0^\infty(\mathbb{R}^3)$, with $\|f_n\|_{L^2} = 1$ such that for the corresponding solutions ϕ_n,

$$\int_0^\infty |\phi_n(t, t, 0, 0)|^2\, dt \ge n. \tag{167}$$

assume by contradiction that in fact, $J := \int_0^\infty \phi(t, t, 0, 0)\varphi(t)\, dt < C$ for all $f \in C_0^\infty(\mathbb{R}^3)$ with $\|f\|_{L^2} = 1$ and some $\varphi \in \mathcal{S}(\mathbb{R}), \varphi \not\equiv 0$. In view of the formula (see section on the fundamental solution of $\Box$ in $\mathbb{R}^{3+1}$),

$$\phi(t, x) = (4\pi)^{-1} t \int_{|\xi|=1} f(x + t\xi)\, d\xi$$

we find that,

$$J = (4\pi)^{-1} \int_{\mathbb{R}^3} |y|^{-1} f_1(y_1 + |y|, y_2, y_3)\varphi(|y|)\, dy$$

or, changing the variables $z = y + (|y|, 0, 0)$

$$J = (4\pi)^{-1} \int\limits_{z_1>0} \frac{1}{z_1} f(z)\varphi\Big(\frac{|z|^2}{2z_1}\Big) dz < c.$$

Since f is an arbitrary $C_0^\infty(\mathbb{R}^3)$ function, $\|f\|_{L^2} = 1$, we must have that,

$$z \to \frac{1}{z_1}\varphi\Big(\frac{|z|^2}{2z_1}\Big)$$

is in $L^2(\mathbb{R}^3_+)$ which is false whenever $\varphi \not\equiv 0$. In fact,

$$\int_{\mathbb{R}^3_+} \frac{1}{z_1^2}\varphi^2\Big(\frac{|z|^2}{2z_1}\Big) dz \quad = \quad \int_{\mathbb{R}^3} \frac{1}{(y_1 + |y|)|y|}\varphi^2(|y|) dy = 2\pi \int_0^\infty \varphi^2(\lambda) \int_0^\pi \frac{\sin\theta}{1 + \cos\theta}\, d\theta$$

diverges logarithmically if $\varphi \not\equiv 0$.

[23] i.e. the one actually proved by Strichartz.

[24] This is obviously so in the region $r \le t$ while for $r \ge t$ the argument is elementary.

12.6. Fourier representation of solutions. We can solve the homogeneous problem (160) by the Fourier method. To recall, If we apply the Fourier transform with respect to the space variables, the initial value problem (160), (159) becomes a Cauchy problem for an ordinary differential equation:

$$\partial_t^2 \widehat{u} + |\xi|^2 \widehat{u} = 0, \quad \widehat{u}(0,\xi) = \hat{f}(\xi), \quad \partial_t \widehat{u}(0,\xi) = \hat{g}(\xi),$$

which can be solved explicitly:

$$\widehat{u}(t,\xi) = \cos(t|\xi|)\, \hat{f}(\xi) + \sin(t|\xi|)\, \frac{\hat{g}(\xi)}{|\xi|} \tag{168}$$

Thus the fundamental solution $W(t)h$, defined above, takes the form,

$$W(t)h(x) = \int_{\mathbb{R}^n} e^{ix\cdot\xi} \frac{\sin(t|\xi|)}{|\xi|} \hat{h}(\xi) d\xi. \tag{169}$$

By Duhamel principle, see (162), the general solution of the inhomogeneou equation $\Box u = F$ can be expressed in the form,

$$u(t) = \partial_t W(t) f + W(t) g + \int_0^t W(t-s) F(s) ds. \tag{170}$$

let $D = (-\Delta)^{1/2}$ be the operator whose symbol in Fourier space is given by $|\xi|$. Observe that,

$$(DW(t))f(x) \quad = \quad (W(t)Df)(x) = \int_{\mathbb{R}^n} e^{ix\cdot\xi} \sin t|\xi|)\hat{f}(\xi) d\xi.$$

Since $\sin t|\xi|$ and $\cos t|\xi|$ are bounded the operators $\partial_t W(t)$ and $DW(t)$ map $H^s(\mathbb{R}^n)$ in itself. In particular, solutions u of (160), (159) preserves the (Sobolev) regularity of the initial data f and g. More precisely, If $f, D^{-1}g \in H^s$ for some $s \in \mathbb{R}$, then $u(t), D^{-1}\partial_t u(t) \in H^s$ uniformly for $t \in \mathbb{R}$. We can also write,

$$\|u\|_{L_t^\infty \dot{H}_x^\gamma} + \|\partial_t u\|_{L_t^\infty \dot{H}_x^{\gamma-1}} \quad \lesssim \quad \|f\|_{\dot{H}^\gamma} + \|g\|_{\dot{H}^{\gamma-1}}$$

which provides the easy part of estimate[25] (163). Therefore to prove (163) it suffices to prove,

$$\|u\|_{L_t^q L_x^r} \quad \lesssim \quad \|f\|_{\dot{H}^\gamma} + \|g\|_{\dot{H}^{\gamma-1}} \tag{171}$$

for and wave admissible pair (q, r).

We also remark that,

$$\partial_t W(t)h(x) \quad = \quad \int_{\mathbb{R}^n} e^{ix\cdot\xi} \cos(t|\xi|)\hat{h}(\xi) d\xi$$

and,

$$D^{-1} W(t)h(x) \quad = \quad \int_{\mathbb{R}^n} e^{ix\cdot\xi} \frac{\cos(t|\xi|)}{|\xi|} \hat{h}(\xi) d\xi$$

We can rewrite (168) as

$$\widehat{u}(t,\xi) = e^{it|\xi|} \hat{f}^+(\xi) + e^{-it|\xi|} \hat{f}^-(\xi),$$

[25]Another derivation, based on energy identities, is given in the next subsection.

where $f^{\pm} = \frac{1}{2}\left(f \pm D^{-1}g\right)$. It follows that $u = u^+ + u^-$ where

$$u^{\pm} = \int e^{i(x\cdot\xi \pm t|\xi|)} \hat{f}_{\pm}(\xi) d\xi$$

Observe that to prove (171) it suffices to prove,

$$\|u^+\|_{L^q_t L^r_x} \lesssim \|f^+\|_{\dot{H}^\gamma} \tag{172}$$

and a similar estimate for f^-.

12.7. Energy estimates. We will derive a simple L^2 estimate for general solutions of $\Box u = F$ by integration by parts. It all follows from the simple algebraic identity:

$$-\frac{1}{2}\partial_t(|\partial_t u|^2 + |\nabla u|^2) + \partial_i(\partial_t u \partial_i u) = \partial_t u \cdot F \tag{173}$$

where $|\nabla u|^2 = \sum_{i=1}^n (\partial_i u)^2$ and $\partial_i = \partial_{x^i}$. Integrating with respect to x, and assuming that u and its derivatives vanish[26] at infinity we derive,

$$\partial_t \int_{\mathbb{R}^n} \big((\partial_t u)^2 + |\nabla u|^2\big) dx = 2\int_{\mathbb{R}^n} \partial_t u \cdot F dx$$

Thus integrating in t,

$$\|\partial_t u(t)\|^2_{L^2} + \|\nabla u(t)\|^2_{L^2} \leq \|\partial_t u(0)\|^2_{L^2} + \|\nabla u(0)\|^2_{L^2} + 2\int_0^t \int_{\mathbb{R}^n} \partial_t u \cdot F dx ds$$

which we rewrite, with $|\partial u|^2 = |\partial_t u|^2 + |\nabla u|^2$,

$$\|\partial u(t)\|^2_{L^2} = \|\partial u(0)\|^2_{L^2} + 2\int_0^t \int_{\mathbb{R}^n} \partial_t u \cdot F dx ds. \tag{174}$$

In particular, applying Hölder,

$$\|\partial u(t)\|^2_{L^2} \leq \|\partial u(0)\|^2_{L^2} + 2\int_0^t \|\partial_t u(s)\|_{L^2} \|F(s)\|_{L^2} ds$$

from which we derive the inhomogeneous energy estimate,

$$\sup_{t\in[0,T]} \|\partial u(t)\|_{L^2} \lesssim \|\partial u(0)\|_{L^2} + \int_0^T \|F(s)\|_{L^2} ds \tag{175}$$

Now let D^s be the operator $D^s = (-\Delta)^{s/2}$ whose symbol in Fourier space is given by $|\xi|^s$. Since D^s commutes with $\Box$ we easily derive,

$$\|\partial D^s u(t)\|^2_{L^2} = \|\partial D^s u(0)\|^2_{L^2} + 2\int_{\mathbb{R}^n} \partial_t D^s u \cdot D^s F dx$$

We can write, using Plancherel with respect to the x variables,

$$\int_{\mathbb{R}^n} \partial_t D^s u \cdot D^s F dx = \int_{\mathbb{R}^n} \partial_t D^{2s} u \cdot F dx,$$

Therefore, by Hölder, in the slab $\mathcal{D}_T = [0, T] \times \mathbb{R}^n$,

$$\sup_{t\in[0,T]} \|\partial D^s u(t)\|^2_{L^2} \leq \|\partial D^s u(0)\|^2_{L^2} + 2\|D^{2s}\partial_t u\|_{L^q_t L^r_x(\mathcal{D}_T)} \|F\|_{L^{q'}_t L^{r'}_x(\mathcal{D}_T)}$$

[26] This can easily be justified by the finite propagation speed property of solutions to the wave equation

Choosing $s=-1$ we infer that,

$$\sup_{t\in[0,T]}\|\partial D^{-1/2}u(t)\|_{L^2}^2\le\|\partial D^{-1/2}u(0)\|_{L^2}^2+2\|D^{-1}\partial_t u\|_{L_t^qL_x^r(\mathcal{D}_T)}\|F\|_{L_t^{q'}L_x^{r'}(\mathcal{D}_T)}$$

We apply this energy estimate to solution of the inhomogeneous problem (12.0.1) with zero initial conditions. We also assume that the dimensional condition $\frac{1}{q}+\frac{n}{r}=\frac{n}{2}-\gamma=\frac{1}{q'}+\frac{n}{r'}-2$ is verified. That implies $\gamma=\frac{1}{2}$. We thus have,

$$\sup_{t\in[0,T]}\|\partial D^{-1/2}u(t)\|_{L^2}^2\le 2\|D^{-1}\partial_t u\|_{L_t^qL_x^r(\mathcal{D}_T)}\|F\|_{L_t^{q'}L_x^{r'}(\mathcal{D}_T)}$$

Assume for a moment that we can prove the estimate,

$$\|D^{-1}\partial_t u\|_{L_t^qL_x^r(\mathcal{D}_T)}\lesssim\|F\|_{L_t^{q'}L_x^{r'}(\mathcal{D}_T)}\tag{176}$$

Then,

$$\sup_{t\in[0,T]}\|\partial D^{-1/2}u(t)\|_{L^2}\lesssim\|F\|_{L_t^{q'}L_x^{r'}(\mathcal{D}_T)}$$

which is equivalent to,

$$\sup_{t\in[0,T]}\big(\|u(t)\|_{\dot H^\gamma)}+\|\partial_t u\|_{\dot H^{\gamma-1})}\big)\quad\lesssim\quad\|F\|_{L^{q'}([0,T];L^{r'})}$$

thus proving half of estimate (164). Therefore the inhomogeneous estimate (164) reduces to proving,

$$\|u\|_{L^q([0,T];L^r)}+\|D^{-1}\partial_t u\|_{L^q([0,T];L^r)}\quad\lesssim\quad\|F\|_{L^{q'}([0,T];L^{r'})}\tag{177}$$

12.8. Homogenous Case. In this section we prove estimate (172) and thus complete the proof for the homogeneous Strichartz estimate of theorem 12.2. Using the space-time Fourier transform, i.e. Fourier transform with respect to both t and x,

$$\widetilde{u}_+(\tau,\xi)=\delta(\tau-|\xi|)\hat f_+(\xi),\quad \widetilde{u}_-(\tau,\xi)=-\delta(\tau+|\xi|)\hat f_-(\xi),\tag{178}$$

These are the components of $\widetilde{u}$ living on the forward null cone $C_+=\{\tau=|\xi|\}$ and on the backward null cone $C_-=\{\tau=-|\xi|\}$, respectively. Thus we can interpret (172) from the point of view of a restriction theorem for the half light cones C_+ or C_-. We next show that it suffices to prove (172) for the case when $\hat f_+$ is included in fixed dyadic piece. More precisely, dropping the label + it suffices to show that,

$$\|u_k^+\|_{L_t^qL_x^r}\quad\lesssim\quad 2^{k\gamma}\|f_k^+\|_{L^2}\tag{179}$$

where $u^+=\sum_{k\in 2^{\mathbb{Z}}}u_k^+$, $u_k^+=P_ku^+$, $f_k^+=P_kf^+$ and P_k the standard LP projections with respect to the spatial variables x.

To show that (180) implies (172) is highly nontrivial[27] as we need to rely on corollary 9.15 adapted to the mixed norms $L_t^qL_x^r$ with both q and r larger than 2. Thus,

$$\|u^+\|_{L_t^qL_x^r}^2\quad\lesssim\quad\sum_{k\in\mathbb{Z}}\|u_k^+\|_{L_t^qL_x^r}^2\lesssim\sum_{k\in\mathbb{Z}}2^{2k\gamma}\|f_k^+\|_{L^2}^2\lesssim\|f^+\|_{\dot H^\gamma}$$

[27]Without using corollary 9.15 we would only derive a weaker estimate with the Besov norm $\dot B^\gamma_{2,1}$ replacing $\dot H^\gamma$ norm on the right.

Finally we observe, using a simple scaling argument, that (180) follows from,

$$\|u_0^+\|_{L_t^q L_x^r} \quad \lesssim \quad \|f_0^+\|_{L^2} \tag{180}$$

We now define the *truncated cone operator* C to be the operator

$$Cf(t,x) = \int e^{it|\xi|} e^{ix\cdot\xi} \chi(\xi) \hat{f}(\xi) \mathrm{d}\xi, \tag{181}$$

where χ is a cut-off function supported in $1.2 \le |\xi| \le 2$, such as the one used in the definition of the LP projections, see (66). The operator C can be viewed as the adjoint of the restriction of the Fourier transform to a truncated cone,

$$\widehat{C^* F}(\xi) = \overline{\chi(\xi)}\ \widetilde{F}(|\xi|, \xi)$$

Estimate (180) is an immediate consequence of the following theorem.

THEOREM 12.9. *Let $(q,r), (q_1,r_1), (q_2,r_2)$ be admissable pairs of exponents. Then we have the estimates*

$$\|Cf\|_{L_t^q L_x^r} \lesssim \|f\|_{L^2}\,, \tag{182}$$

and also

$$\|CC^* F\|_{L_t^{q_1} L_x^{r_1}} \lesssim \|F\|_{L_t^{q_2'} L_x^{r_2'}}\,. \tag{183}$$

Composing C with C^* we derive,

$$CC^* F(t,x) \simeq \int e^{i[(t-s)|\xi| + (x-y)\cdot\xi]} |\beta(\xi)|^2 F(s,y) \mathrm{d}s\mathrm{d}y\mathrm{d}\xi,$$

which can be rewritten as the convolution

$$CC^* F(t,\cdot) = \int U(t-s) F(s,\cdot) \mathrm{d}s, \tag{184}$$

with the evolution operator

$$U(t)f(x) = \int e^{i(t|\xi| + x\cdot\xi)} |\chi(\xi)|^2 \hat{f}(\xi) \mathrm{d}\xi. \tag{185}$$

(Observe that U is essentially the same operator as C !) By the TT^* principle, we know that the estimate (182) is equivalent to the following estimate for CC^*,

$$\|CC^* F\|_{L_t^q L_x^r} \lesssim \|F\|_{L_t^{q'} L_x^{r'}}\,. \tag{186}$$

which is also equivalent to the polarized form (183). Thus, to prove the theorem it suffices to prove (186). As in the second proof of the restriction theorem presented in the previous section to prove (186) we need to prove the following properties for the evolution operators $U(t)$.

PROPOSITION 12.10. *Let $\chi(\xi)$ be a fixed $C_0^\infty(\mathbb{R}^n)$ function supported in $1/2 \le |\xi| \le 2$ and,*

$$U(t)f(x) = \int e^{i(t|\xi| + x\cdot\xi)} \chi(\xi) \hat{f}(\xi) d\xi. \tag{187}$$

Then,

$$\|U(t)f\|_{L^2} \quad \lesssim \quad C\|f\|_{L^2} \tag{188}$$

$$\|U(t)f\|_{L^\infty} \quad \lesssim \quad (1+|t|)^{-\frac{n-1}{2}} \|f\|_{L^1} \tag{189}$$

from which, interpolating, for all $2 \le r \le \infty$,

$$\|U(t)f\|_{L^r} \quad \lesssim \quad (1+|t|)^{-\frac{n-1}{2}(1-\frac{2}{r})}\|f\|_{L^{r'}} \tag{190}$$

Moreover, if in addition, $\chi = \chi_\mu$ *is supported in a cube of size* μ*, then* (189) *can be strengthened to*

$$\|U(t)f\|_{L^\infty} \quad \lesssim \quad \mu(1+|t|)^{-\frac{n-1}{2}}\|f\|_{L^1} \tag{191}$$

Proof We prove directly the stronger version (191). We only need to check (**??**). We write,

$$U(t)f = K_t * f, \qquad K_t(x) = \int e^{i(x\cdot\xi + t|\xi|)}\chi_\mu(\xi)d\xi$$

It suffices to show that,

$$|K_t(x)| \lesssim \mu\frac{1}{(1+|t|+|x|)}$$

In the regions $|x| < |t|/2$ and $|x| \ge 2|t|$ we integrate by parts k times with respect to the operator $L = -i\sum_j \frac{x_j + t\frac{\xi_j}{|\xi}}{|x+t\frac{\xi}{|\xi|}|^2}\partial_{\xi_j}$, such that $L(e^{i(x\cdot\xi+t|\xi|)}) = e^{i(x\cdot\xi+t|\xi|)}$. We also make use of the straightforward estimate, $|\partial_\xi^\alpha \chi_\mu(\xi)| \lesssim \mu^{-|\alpha|}$ to derive, $|K_t(x)| \lesssim (1+|t|)^{-k}\mu^{n-k}$ or, choosing $k = \frac{n-1}{2}$,

$$|K_t(x)| \lesssim (1+|t|)^{-\frac{n-1}{2}}\mu^{\frac{n+1}{2}}.$$

On the other hand, in the region $|t| \approx |x|$, we write, with $\beta(|\xi|)$ vanishing on the support of h_μ,

$$K_t(x) \quad = \quad \int_{1-2\mu}^{1+2\mu} e^{it\lambda}\chi(\lambda)\int_{|\xi|=\lambda} e^{ix\cdot\xi}h_\mu(\xi)d\sigma(\xi)$$

We now need to rely on the following estimate,

$$\sup_{1/2\le\lambda\le 2} |\int_{|\xi|=\lambda} e^{ix\cdot\xi}h(\xi)d\sigma(\xi)| \lesssim (1+|x|)^{-\frac{n-1}{2}} \tag{192}$$

which follows easily from the decay of the Fourier transform of measures supported on $\mathbb{S}^{n-1}$ discussed in the previous section, see lemma 11.8. Therefore, for $|t| \sim |x|$,

$$|K_t(x)| \lesssim \mu(1+|x|)^{-\frac{n-1}{2}} \lesssim \mu(1+|t|)^{-\frac{n-1}{2}}$$

as desired. ■

We are now ready to prove (186) by following the same argument as in the second proof of the restriction theorem. Indeed, in view of (184) and (190) we derive,

$$\|CC^*F\|_{L^r_x}(t) \quad \lesssim \quad \int_{-\infty}^{+\infty}(1+|t-s|)^{-\gamma(r)}\|F(s)\|_{L^{r'}_x}ds \tag{193}$$

where $\gamma(r) = -\frac{n-1}{2}(1-\frac{2}{r})$. We are now precisely in the same situation as in the second proof of the restriction theorem, see the argument following formula (151). If $0 < \gamma(r) < 1$ we can apply the Hardy-Littlewood-Sobolev inequality to obtain

$$\|CC^*F\|_{L^q_tL^r_x} \lesssim \|F\|_{L^{q'}_tL^{r'}_x},$$

when $-\gamma(r) + 1 + 1/q = 1/q'$, hence $\gamma(r) = 2/q$. This proves (180), and thus theorem 12.9, in the case $0 < \gamma(r) = 2/q < 1$. If $q = 2$ and $\gamma(r) > 1$ we have from (193),

$$\|CC^*F\|_{L^2_t L^r_x} \lesssim \|F\|_{L^2_t L^{r'}_x},$$

by an application of the standard Hausdorff-Young inequality.

Finally, if $2/q < 1$ and $\gamma(r) > 2/q$ the result follows from the case $\gamma(r) = 2/q$ using Sobolev inequalities. Due to the fact that one of the principal curvatures of the light cone vanishes, the Strichartz estimates for the wave equation is not as strong as it could be. Using the improved dispersive estimate (191) we can however derive a stronger statement ,which is very useful in applications.

Proposition 12.11. *Let $0 < \mu < 1$. Let f be an L^2 function with Fourier transform supported in a cube of size μ at a distance 1 from the origin. Let (q, r) be an admissable pair of exponents for the Strichartz estimates. Then*

$$\|Cf\|_{L^q_t L^r_x} \lesssim \mu^{\left(\frac{1}{2}-\frac{1}{r}\right)} \|f\|_{L^2}. \tag{194}$$

The proof is based on the improved dispersive estimate (191). Interpolating it with (188) we derive,

$$\|U(t)f\|_{L^r} \quad \lesssim \quad \mu^{1-\frac{2}{r}}(1+|t|)^{-\frac{n-1}{2}(1-\frac{2}{r})}\|f\|_{L^{r'}}$$

The proof the continues exactly as above to derive,

$$\|CC^*F\|_{L^q_t L^r_x} \lesssim \mu^{1-\frac{2}{r}} \|F\|_{L^{q'}_t L^{r'}_x},$$

and therefore, by the TT^* argument, $\|Cf\|_{L^q_t L^r_x} \lesssim \mu^{\left(\frac{1}{2}-\frac{1}{r}\right)} \|f\|_{L^2}$, as desired. As a straightforward corollary to the proposition we derive:

Theorem 12.12. *Consider a general solution of $\Box u = 0$ with data f, g supported, in Fourier space, on a cube of size μ situated in a dyadic shell of size λ, with λ much larger than μ, say $\lambda \geq 8\mu$. Then,*

$$\|u\|_{L^q_t L^r_x} \lesssim \mu^{1-\frac{2}{r}} \big(\|f\|_{\dot{H}^\gamma} + \|f\|_{\dot{H}^{\gamma-1}}\big) \tag{195}$$

Proof The proof follows easily by a scaling argument from the proposition above. ■

Finally we state below another result, which follows easily from the decay estimate (189).

Theorem 12.13. *Let u be a free wave, i.e. solution of the homogeneous equation $\Box u = 0$, with initial data (f, g). Then,*

$$\begin{aligned} \|u(t)\|_{L^\infty} \quad &\lesssim \quad |t|^{-\frac{n-1}{2}} \sum_{\lambda \in 2^{\mathbb{Z}}} \left(\lambda^{\frac{n+1}{2}} \|f_\lambda\|_{L^1} + \lambda^{\frac{n-1}{2}} \|g_\lambda\|_{L^1}\right) \\ &= \quad |t|^{-\frac{n-1}{2}} \left(\|f\|_{\dot{B}^{n+1/2}_{1,1}} + \|g\|_{\dot{B}^{n-1/2}_{1,1}}\right). \end{aligned}$$

The uniform decay rate $|t|^{-\frac{n-1}{2}}$, for large t, plays a very important role in the study of nonlinear perturbations of the standard wave equation.

12.14. Inhomogeneous Strichartz estimates. We have already reduced the inhomogeneous Strichartz estimate (164) of theorem 12.2 to estimate (177). Proceeding as in the case of the homogeneous estimates we can now reduce (177) to the case when the spatial Fourier transform of F is supported in the unit dyadic ring $1/2 \le |\xi| \le 2$. Moreover, decomposing u as before in the $\pm$ parts it suffices to prove the estimates separately for u_+ and u_-. Therefore we need to prove,

$$\|u^+|_{L^q([0,T];L^r)} + \|D^{-1}\partial_t u^+\|_{L^q([0,T];L^r)} \quad \lesssim \quad \|F\|_{L^{q'}([0,T];L^{r'})} \tag{196}$$

We have,

$$\begin{aligned} u_+(t,\cdot) \quad &= \quad \int_0^t U(t-s)F(s,\cdot)ds \\ D^{-1}\partial_t u_+(t,\cdot) \quad &= \quad \int_0^t \partial_t D^{-1}U(t-s)F(s,\cdot)ds \end{aligned}$$

Since, in view of the dyadic restriction, $\partial_t D^{-1}U(t) \sim U(t)$ it suffices to prove the estimate for $\|u^+|_{L^q([0,T];L^r)}$. Clearly, u^+ differs from CC^*F in (184) only by the restriction of the interval of integration to $[0,t]$. In view of this fact we write $u_+ = (CC^*)_R F$. We are thus led to the following theorem, from which (196) and thus (164).

THEOREM 12.15. *Let $U(t)$ defined as in* (187) *and let*

$$(CC^*)_R F(t,\cdot) \quad = \quad \int_0^t U(t-s)F(s,\cdot)ds$$

Then, for all admissible pairs (q_1,r_1), (q_2,r_2),

$$\|(CC^*)_R F\|_{L_t^{q_1}L_x^{r_1}([0,T]\times\mathbb{R}^n)} \quad \lesssim \quad \|F\|_{L_t^{q_2'}L_x^{r_2'}([0,T]\times\mathbb{R}^n)} \tag{197}$$

Proof The proof is straightforward in the case $(q_1,r_1) = (q_2,r_2) = (q,r)$. Indeed in this case we can simply repeat the proof of estimate (186) and just take into account the limits of integration. We have also treated the case when $q_1 = \infty$, $r_1 = 2$, see the subsection on energy estimates. The other non-diagonal case cases are a little more difficult and will be treated in the more general abstract setting discuss later in this section. The proof we have given covers however the most interesting case of estimate (164). We have thus given complete proofs for the first two parts of theorem 12.2 ■

12.16. Necessity of the admissibility conditions. To understand what is the optimal range of exponents q and r we consider the analog of the Knapp counterexample in the context of the truncated cone operator C defined in (181).

For some small $\delta > 0$, let

$$D = \{\xi \in \mathbb{R}^n : |\xi_1 - 1| < 1/2, |\xi'| < \delta\},$$

and consider $f = \chi_D$. We have

$$Cf(t,x) = e^{i(t+x_1)} \int_D e^{i[t(|\xi|-\xi_1)+(t+x_1)(\xi_1-1)+x'\cdot\xi']} \mathrm{d}\xi,$$

and observe that

$$|\xi| - \xi_1 = \frac{|\xi'|^2}{|\xi| + \xi_1} \lesssim \delta^2.$$

We can then choose a region of space-time R defined by

$$|t| \lesssim \delta^{-2}, \quad |t + x_1| \lesssim 1, \quad |x'| \lesssim \delta^{-1},$$

such that, when $(t, x) \in R$ and $\xi \in D$, then the oscillatory factor inside the last integral can be treated as a constant. Hence, $|Cf(t, x)| \gtrsim |D|$ for $(t, x) \in R$ and we have

$$\frac{\|Cf\|_{L_t^q L_x^r}}{\|f\|_{L^2}} \gtrsim \frac{|D| \, \|\chi_R\|_{L_t^q L_x^r}}{|D|^{1/2}} \sim \delta^{\frac{n-1}{2} - \frac{2}{q} - \frac{n-1}{r}}.$$

In the limit $\delta \to 0$, an estimate of the form (186) will necessarily imply that q and r satisfy the condition

$$\frac{2}{q} \leq (n-1)\left(\frac{1}{2} - \frac{1}{r}\right). \tag{198}$$

The other restriction on the range for q, i.e. $q \geq 2$ is a consequence of the invariance of the operator CC^* under time translations. Indeed for translation invariant operators we have the following general result due to Hörmander, [**7**].

PROPOSITION 12.17. *Let $T : L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$ be a (non trivial) linear operator which commutes with translations, in the sense that $(Tf) \circ \tau_y = T(f \circ \tau_y)$, where $\tau_y(x) = x + y$, for $x, y \in \mathbb{R}^n$. If T is bounded from L^p to L^q then we necessarily have $q \geq p$.*

The proof is based on the following lemma.

LEMMA 12.18. *Let $f \in L^p(\mathbb{R}^n)$, then*

$$\lim_{|y| \to \infty} \|f + f \circ \tau_y\|_{L^p} = 2^{1/p} \|f\|_{L^p}.$$

Proof For every $R > 0$ consider the decomposition $f = g_R + h_R$, where $g_R(x) = f(x)$ if $|x| < R$ and 0 if $|x| \geq R$, and $h_R(x) = 0$ if $|x| < R$ and $f(x)$ if $|x| \geq R$. Then

$$\lim_{R \to \infty} \|g_R\|_{L^p} = \|f\|_{L^p}, \quad \lim_{R \to \infty} \|h_R\|_{L^p} = 0.$$

For $R = |y|/2$ we have

$$f + f \circ \tau_y = g_R + g_R \circ \tau_y + h_R + h_R \circ \tau_y.$$

The functions g_R and $g_R \circ \tau_y$ have disjoint supports, so that

$$\|g_R + g_R \circ \tau_y\|_{L^p}^p = \|g_R\|_{L^p}^p + \|g_R \circ \tau_y\|_{L^p}^p = 2 \|g_R\|_{L^p}^p,$$

while

$$\lim_{|y| \to \infty} \|h_R + h_R \circ \tau_y\|_{L^p} \leq \lim_{|y| \to \infty} 2 \|h_R\|_{L^p} = 0,$$

hence

$$\lim_{|y| \to \infty} \|f + f \circ \tau_y\|_{L^p} = \lim_{|y| \to \infty} 2^{1/p} \|g_R\|_{L^p} = 2^{1/p} \|f\|_{L^p}.$$

∎

Proof [Proof of Proposition 12.17] Let $C > 0$ be the optimal constant for the estimate

$$\|Tf\|_{L^q} \leq C \|f\|_{L^p}, \quad \forall f \in L^p.$$

Then by linearity and the translation invariance,

$$\|Tf + (Tf) \circ \tau_y\|_{L^q} \leq C \|f + f \circ \tau_y\|_{L^p}.$$

When $|y| \to \infty$, applying the lemma we obtain

$$2^{1/q} \|Tf\|_{L^q} \leq C 2^{1/p} \|f\|_{L^p}, \quad \forall f \in L^p.$$

The optimality of C implies that $2^{\frac{1}{p}-\frac{1}{q}} \geq 1$, hence $q \geq p$. ■

The proposition generalizes easily to vector valued L^p spaces and if we consider CC^* as an operator from $L^{q'}(\mathbb{R}; L_x^{r'})$ to $L^q(\mathbb{R}; L_x^r)$, then we must have $q \geq q'$, which is the condition $q \geq 2$.

12.19. A general, abstract framework. It turns out that the method of proving Strichartz estimates described above applies to many other equations, such as Schrödinger, KdV etc. It thus pays to have a general framework which applies to all these cases.

Let $(X, \mathrm{d}\mu)$ be a measure space and H a Hilbert space. Consider a family $(U(t))_{t\in\mathbf{R}}$ of operators $U(t) : H \to L^2(X)$, which describes the *evolution* of some system with *data* in H. We assume that this evolution satisfies the following two properties:

- for all $t \in \mathbf{R}$ and $f \in H$ we have the energy estimate:
$$\|U(t)f\|_{L^2(X)} \lesssim \|f\|_H ; \tag{199}$$
- for all $t \neq s$ and $g \in L^1(X)$ we have the dispersive inequality:
$$\|U(t)U^*(s)g\|_{L^\infty(X)} \lesssim |t-s|^{-\gamma_0} \|g\|_{L^1(X)}, \tag{200}$$
for some $\gamma_0 > 0$.

Interpolating between (199) and (200) we obtain the estimate

$$\|U(t)U^*(s)g\|_{L^r(X)} \lesssim |t-s|^{-\gamma(r)} \|g\|_{L^{r'}(X)}, \tag{201}$$

for $r \geq 2$, where

$$\gamma(r) = \gamma_0 \left(1 - \frac{2}{r}\right).$$

THEOREM 12.20. *If the evolution operator $U(t)$ satisfies* (199) *and* (200)*, then the estimates*

$$\|U(t)f\|_{L_t^q L_X^r} \lesssim \|f\|_H, \tag{202}$$

hold for all $q, r \geq 2$ verifing:

$$\frac{2}{q} = \gamma(r), \quad (q, r, \gamma_0) \neq (2, \infty, 1). \tag{203}$$

REMARK 12.21. This form of the Strichartz inequalities applies to linear dispersive equations such as Schrödinger.

Proof If we consider the operator $T : H \to L^q_t L^r_X$ defined by $Tf(t,x) = (U(t)f)(x)$ then it is easy to verify that the dual of T is the operator $T^* : L^{q'}_t L^{r'}_X \to H$ given by $T^*F = \int U^*(s)F(s,\cdot)\mathrm{d}s$. By the TT^* method, (202) is then equivalent to the estimate

$$\left\| \int U(t)U^*(s)F(s)\mathrm{d}s \right\|_{L^q_t L^r_X} \lesssim \|F\|_{L^{q'}_t L^{r'}_X}. \tag{204}$$

By duality and symmetry considerations, this is in turn equivalent to

$$|B(F,G)| \lesssim \|F\|_{L^{q'}_t L^{r'}_X} \|G\|_{L^{q'}_t L^{r'}_X}, \tag{205}$$

where $B(F,G)$ is the bilinear form

$$B(F,G) = \iint_{s<t} \langle U^*(t)F(t), U^*(s)G(s)\rangle \mathrm{d}t\mathrm{d}s. \tag{206}$$

From the bilinear version of (201) we have that

$$|B(F,G)| \lesssim \iint \frac{\|F(t)\|_{L^{r'}} \|G(s)\|_{L^{r'}}}{|t-s|^{\gamma(r)}} \mathrm{d}s\mathrm{d}t. \tag{207}$$

If $\gamma(r) < 1$, we can apply the Hardy-Littlewood-Sobolev inequality and obtain (205). This concludes the proof for the cases $q = 2/\gamma(r) > 2$.

The endpoint case, corresponding to $\gamma(r) = 2/q = 1$, is allowed when $r < \infty$. Its proof will be described in the next section. ■

REMARK 12.22. If we strengthen the dispersive condition (200) to

$$\|U(t)U^*(s)g\|_{L^\infty(X)} \lesssim (1 + |t-s|)^{-\gamma_0} \|g\|_{L^1(X)}, \tag{208}$$

then (207) can be improved to

$$|B(F,G)| \lesssim \iint \frac{\|F(t)\|_{L^{r'}} \|G(s)\|_{L^{r'}}}{(1+|t-s|)^{\gamma(r)}} \mathrm{d}s\mathrm{d}t. \tag{209}$$

Now we can obtain (205) from Young's inequality when $2/q = 1/p$ and $(1 + |t|)^{-\gamma(r)} \in L^p(\mathbf{R})$, i.e. $\gamma(r)p > 1$. Hence, (208) allows us to extend the Strichartz estimates (202) in Theorem 12.20 to the range

$$\frac{2}{q} \le \gamma(r), \quad (q, r, \gamma_0) \neq (2, \infty, 1). \tag{210}$$

This case applies to the linear wave equations.

REMARK 12.23. We observe that there is a natural scaling associated to the objects in this abstract formulation. More precisely, the estimates (202) in Theorem 12.20 are invariant under the change of scale defined by

$$U(t) \leftarrow U(t/\lambda), \quad U^*(s) \leftarrow U^*(s/\lambda), \quad \mathrm{d}\mu \leftarrow \lambda^{\gamma_0}\mathrm{d}\mu, \quad \langle f, g\rangle_H \leftarrow \lambda^{\gamma_0}\langle f, g\rangle_H. \tag{211}$$

We can also consider the endpoint case.

$$q = 2, \quad r = \frac{2\gamma_0}{\gamma_0 - 1}, \quad \gamma_0 > 1.$$

This, in fact, is more difficult than the previous non-endpoint case, and requires a two-parameter estimate which is better than the one-parameter family given by the interpolation (201). This proof is presented in the previously mentioned paper by Keel and Tao, "Endpoint Strichartz Estimates". We omit it here.

12.24. Inhomogeneous estimates. Saying that an operator T maps the Hilbert space H into $L_t^q L_X^r$, is equivalent to saying that its dual T^* maps $L_t^{q'} L_X^{r'}$ into H, and is also equivalent to saying that the TT^* operator maps $L_t^{q'} L_X^{r'}$ into $L_t^q L_X^r$. If the pair (q, r) is allowed to vary in a set E of admissable exponents, we can view TT^* as a composition of two operators associated with different pairs of exponents. It follows that TT^* actually satisfies a larger set of mapping properties, since it maps $L_t^{\tilde{q}'} L_X^{\tilde{r}'}$ into $L_t^q L_X^r$, for any couple of pairs $(q, r), (\tilde{q}, \tilde{r}) \in E$.

The operator $Tf(t) = U(t)f$ defined in the previous subsection can be viewed as the solution of some homogenous, translation invariant, linear evolution equation. The solution of the corresponding inhomogenoues problem, using Duhamel's principle, would be represented by the retarded operator

$$RF(t) = \int_{s<t} U(t)U^*(s)F(s)\mathrm{d}s.$$

Observe that operator R looks very similar to the TT^* operator, which is given by

$$TT^*F(t) = \int U(t)U^*(s)F(s)\mathrm{d}s.$$

The restriction $s < t$ in the definition of R, however, destroys the composition structure of TT^*. Fortunately, all the mapping properties of TT^*, which we have derived above, can be transfered to R.

THEOREM 12.25. *The operator R maps $L_t^{\tilde{q}'} L_X^{\tilde{r}'}$ into $L_t^q L_X^r$, for any couple of pairs $(q, r), (\tilde{q}, \tilde{r})$ for which the Strichartz estimate 202 holds.*

Proof First of all observe that in the proof of theorem 12.20 we have actually proved the diagonal case $(q, r) = (\tilde{q}, \tilde{r})$. Indeed, the bilinear form defined in (206) can be written as $B(F, G) = \iint R(F) \cdot G \mathrm{d}x\mathrm{d}t$ and (205) is the dual formulation of the mapping property for R.

The non diagonal cases with $\frac{1}{q} + \frac{1}{\tilde{q}} < 1$ follow from the mapping properties of TT^* by using a general argument about integral operators due to Christ and Kiselev (see [] and []) which we summarize in Proposition 12.27 below.

It remains to consider the cases with $q = \tilde{q} = 2$ and $r \neq \tilde{r}$, under the assumption that the evolution $U(t)$ satisfies the stronger dispersive inequality (208) with $\gamma_0 > 1$. Since, we have already proved the case $r = \tilde{r}$, by interpolation it is enough to consider the extreme case: $r = r_* = \frac{2\gamma_0}{\gamma_0 - 1}$, $\tilde{r} = \infty$, and show that

$$|B(F, G)| \lesssim \|F\|_{L_t^2 L_X^{r_*'}} \|G\|_{L_t^2 L_X^1}.$$

This estimate follows by decomposing $B(F,G)$ into dyadic pieces, $B = \sum_{\lambda\in 2^{\mathbb{Z}}} B_\lambda$, where

$$B_\lambda(F,G) = \iint_{\lambda/2\le|t-s|\le 2\lambda} \langle U^*(t)F(t), U^*(s)G(s)\rangle \mathrm{d}t\mathrm{d}s, \tag{212}$$

The desired conclusion follows immediately from the lemma below. ■

LEMMA 12.26. *Let $B_\lambda(F,G)$ be the bilinear form defined in* (212). *Then, there exists an $\varepsilon > 0$ such that*

$$|B_\lambda(F,G)| \lesssim \min\left\{\lambda, \lambda^{-1}\right\}^\varepsilon \|F\|_{L_t^2 L_X^{r_*'}} \|G\|_{L_t^2 L_X^1}.$$

Proof We may assume that F and G are supported on disjoint time intervals of length $O(\lambda)$ separated by a distance $O(\lambda)$. Then $B_\lambda(F,G) = \langle T^*F, T^*G\rangle_H$. We use the energy estimate to bound $\|T^*F\|_H$ and the Strichartz estimate with $q=2$ and $r=\infty$ to bound $\|T^*G\|_H$, so that

$$|B_\lambda(F,G)| \lesssim \|F\|_{L_t^1 L_X^2} \|G\|_{L_t^2 L_X^1}.$$

We then apply Holder inequality and use the assumption on the support of F to obtain

$$|B_\lambda(F,G)| \lesssim \lambda^{1/2} \|F\|_{L_t^2 L_X^2} \|G\|_{L_t^2 L_X^1}.$$

We can also write $B_\lambda(F,G) = \iiint F(t)\cdot U(t)U^*(s)G(s)\mathrm{d}x\mathrm{d}s\mathrm{d}t$ and make use of the dispersive inequality,

$$|B_\lambda(F,G)| \lesssim (1+\lambda)^{-\gamma_0} \|F\|_{L_t^1 L_X^1} \|G\|_{L_t^1 L_X^1}.$$

Again, we apply Holder inequality and use the assumption on the support of F and G to obtain

$$|B_\lambda(F,G)| \lesssim \frac{\lambda}{(1+\lambda)^{\gamma_0}} \|F\|_{L_t^2 L_X^1} \|G\|_{L_t^2 L_X^1}.$$

Hence, B_λ is bounded on $L_t^2 L_X^2 \times L_t^2 L_X^1$ with constant $\lambda^{1/2}$ and on $L_t^2 L_X^1 \times L_t^2 L_X^1$ with constant $\frac{\lambda}{(1+\lambda)^{\gamma_0}}$. By standard interpolation of L^p spaces we obtain that B_λ is bounded on $L_t^2 L_X^{r_*'} \times L_t^2 L_X^1$ with constant C_λ, where

$$C_\lambda = \lambda^{\theta/2}\left(\frac{\lambda}{(1+\lambda)^{\gamma_0}}\right)^{1-\theta}, \quad \frac{1}{r_*'} = \frac{\theta}{2} + \frac{1-\theta}{1}, \quad r_* = \frac{2\gamma_0}{\gamma_0 - 1}.$$

Simplyfing the expression we find that

$$C_\lambda = \frac{\lambda^{\frac{\gamma_0+1}{2\gamma_0}}}{1+\lambda} \lesssim \min\left\{\lambda, \lambda^{-1}\right\}^\varepsilon,$$

with

$$\varepsilon = \min\left\{\frac{\gamma_0+1}{2\gamma_0}, 1 - \frac{\gamma_0+1}{2\gamma_0}\right\} = \frac{\gamma_0 - 1}{2\gamma_0} = \frac{1}{r_*} > 0.$$

■

12.26.1. *Integral operators with restricted kernel.* In this subsection we give a self contained exposition of the results of Christ-Kisselev mentioned above. Consider an integral operator with a measurable kernel $K(s,t)$,

$$Tf(t) = \int_{\mathbb{R}} K(s,t) f(s) \mathrm{d}s,$$

and its restricted version associated with the kernel $K(s,t)\chi(s<t)$,

$$Rf(t) = \int_{s<t} K(s,t) f(s) \mathrm{d}s.$$

If T maps L^p into L^q and $1 \le p < q \le \infty$ then we have that R also maps L^p into L^q. An equivalent formulation of this fact is given in the following proposition.

Proposition 12.27. *Let $K(s,t)$ be a measurable function on $\mathbf{R} \times \mathbf{R}$. Let $B(f,g)$ be the bilinear form with kernel K,*

$$B(f,g) = \iint K(s,t) f(s) g(t) \, ds\, dt,$$

and $\widetilde{B}(f,g)$ the bilinear form with kernel restricted to the region $s<t$,

$$\widetilde{B}(f,g) = \iint_{s<t} K(s,t) f(s) g(t) \, ds\, dt.$$

Let $p, q \ge 1$, with the condition

$$\frac{1}{p} + \frac{1}{q} > 1. \tag{213}$$

If B is bounded on $L^p \times L^q$,

$$|B(f,g)| \lesssim \|f\|_{L^p} \|g\|_{L^q},$$

then $\widetilde{B}$ is also bounded on $L^p \times L^q$,

$$\left|\widetilde{B}(f,g)\right| \lesssim \|f\|_{L^p} \|g\|_{L^q}.$$

Remark 12.28. There are cases for which equality in condition (213) is not allowed. Consider for the example the case of the Hilbert transform, which corresponds to the kernel $K(s,t) = \frac{1}{s-t}$, with $p = q = 2$.

Proof Let $f \in L^p$ and $g \in L^q$ with $\|f\|_{L^p} = \|g\|_{L^q} = 1$.

Define $F(t) = \int_{s<t} |f(s)|^p \mathrm{d}s$. F is a continuous non-decreasing function which maps $[-\infty, +\infty]$ onto $[0,1]$. In particular, the inverse image of an interval of the type $I = [a,b] \subset [0,1]$ will be an interval of the same type, $F^{-1}(I) = [A,B]$, with $F(A) = a$, $F(B) = b$, and $\int_A^B |f(s)|^p \mathrm{d}s = F(B) - F(A) = b - a$. Hence,

$$\|f\|_{L^p(F^{-1}(I))} = |I|^{1/p}. \tag{214}$$

Consider now a Whitney decomposition of the set $\Omega = \left\{(x,y) \in \mathbb{R}^2 : x < y\right\}$ into disjoint dyadic squares, as in Lemma 9.7, $\Omega = \cup_Q Q$, where each square $Q = I \times J$ has the property

$$\operatorname{dist}(I,J) \approx |I| = |J| = \lambda, \tag{215}$$

for some dyadic value of λ. If we look only at those squares needed to cover the triangle $\Omega \cap [0,1]^2$, then $\lambda \le 1/2$.

Observe that $s < t$ implies that either $F(s) < F(t)$ or $f \equiv 0$ almost everywhere on the interval $[s,t]$. Hence, we can write

$$\widetilde{B}(f,g) = \iint_{F(s)<F(t)} K(s,t) f(s) g(t) \mathrm{d}s\mathrm{d}t = \sum_Q B(\chi_{F^{-1}(I)} f, \chi_{F^{-1}(J)} g).$$

Using the boundedness of B on $L^p \times L^q$ we obtain

$$\left|\widetilde{B}(f,g)\right| \lesssim \sum_Q \|f\|_{L^p(F^{-1}(I))} \|g\|_{L^q(F^{-1}(J))} .$$

Now we use (214), (215) and the fact that, for each given dyadic interval J, the number of intervals I for which $I \times J$ is one of the squares in the decomposition of Ω is bounded by a universal constant. Hence,

$$\left|\widetilde{B}(f,g)\right| \lesssim \sum_{\lambda \le 1/2} \lambda^{\frac{1}{p}} \sum_{|J|=\lambda} \|g\|_{L^q(F^{-1}(J))} .$$

Next, we apply Hölder's inequality to the summation over the dyadic intervals J of length λ and since there are λ^{-1} of them in $[0,1]$ we have

$$\left|\widetilde{B}(f,g)\right| \lesssim \sum_{\lambda \le 1/2} \lambda^{\frac{1}{p}} \lambda^{-\frac{1}{q'}} \|g\|_{L^q} = \sum_{\lambda \le 1/2} \lambda^{\frac{1}{p}+\frac{1}{q}-1} \lesssim 1.$$

■

13. L^2 bilinear estimates

13.1. Bilinear proofs of some Strichartz estimates. Consider the homogeneous wave equation $\Box u = 0$ in $\mathbb{R}^{1+3}$. The Strichartz estimate (163) with $q = r = 4$ and $\gamma = 1/2$. Takes the form,

$$\|u\|_{L^4(\mathbb{R}^{1+3})} \quad \lesssim \quad \|f\|_{\dot{H}^{1/2}} + \|g\|_{\dot{H}^{-1/2}}$$

Writing $u = u^+ + u^-$ it suffices to prove,

$$\|u^+\|_{L^4(\mathbb{R}^{1+3})} \quad \lesssim \quad \|f^+\|_{\dot{H}^{1/2}} \tag{216}$$

where

$$u_+(t,x) = \int e^{ix\cdot\xi + t|\xi|} \hat{f}(\xi) d\xi$$

Clearly,

$$\|u^+\|^2_{L^4(\mathbb{R}^{1+3})} \quad = \quad \|u^+ \cdot u^+\|_{L^2} = \|\widetilde{u^+} * \widetilde{u^+}\|_{L^2}$$

Now, recalling (178), and dropping the index $+$,

$$\begin{aligned} \tilde{u} * \tilde{u}(\tau,\xi) &= \int\int \delta(\tau - \lambda - |\xi - \eta|) \hat{f}(\xi - \eta) \delta(\lambda - |\eta|) \hat{f}(\eta) d\lambda d\eta \\ &= \int \delta((\tau - |\eta| - |\xi - \eta|) \hat{f}(\eta) \hat{f}(\xi - \eta) d\eta \end{aligned}$$

Clearly, (216) follows from the following:

THEOREM 13.2. *The bilinear operator,*

$$B(F,G) \quad = \quad \int \delta(\tau-|\eta|-|\xi-\eta|)\frac{F(\xi-\eta)}{|\xi-\eta|^{1/2}}\frac{G(\eta)}{|\eta|^{1/2}}d\eta.$$

verifies the estimate,

$$\|B(F,G)\|_{L^2(\mathbb{R}^{1+3})} \lesssim \|F\|_{L^2(\mathbb{R}^3)}\|G\|_{L^2(\mathbb{R}^{1+3})} \tag{217}$$

Proof By Cauchy-Schwartz,

$$\begin{aligned} |B(F,G)(\tau,\xi)|^2 &\lesssim J(\tau,\xi)\int \delta(\tau-|\eta|-|\xi-\eta|)|F(\xi-\eta)|^2|G(\eta)|^2 d\eta \\ J(\tau,\xi) &= \int \delta(\tau-|\eta|-|\xi-\eta|)\frac{1}{|\xi-\eta|}\frac{1}{|\eta|}d\eta \end{aligned}$$

It suffices to show that J is uniformly bounded. Indeed, if that is the case,

$$\begin{aligned} \|B(F,G)\|_{L^2(\mathbb{R}^{1+3})} &\lesssim \sup_{\tau,\xi} J(\tau,\xi)\int\int \delta(\tau-|\eta|-|\xi-\eta|)F(\xi-\eta)|^2|G(\eta)|^2 d\eta d\tau d\xi \\ &\lesssim \sup_{\tau,\xi} J(\tau,\xi)\|F\|_{L^2}^2\|G\|_{L^2}^2 \end{aligned}$$

Therefore the bilinear estimate is an immediate consequence of the uniform boundedness of J. This follows from the following more general lemma below. ■

LEMMA 13.3. *Let F be an arbitrary function of two variables and J_F the integral*

$$J_F^{\mp}(\tau,\xi) \quad = \quad \int_{\mathbb{R}^n} \delta(\tau-|\eta|\mp|\xi-\eta|)F(|\eta|,|\xi-\eta|)$$

Then,

$$J_F^{-}(\tau,\xi) \quad = \quad (\tau^2-|\xi|^2)^{\frac{n-3}{2}}\int_{-1}^{1} F\big(\frac{\tau+s|\xi|}{2},\frac{\tau+s|\xi|}{2}\big)(\tau^2-x^2|\xi|^2)(1-|x|^2)^{\frac{n-3}{2}}dx, \tag{218}$$

$$J_F^{+}(\tau,\xi) \quad = \quad (\tau^2-|\xi|^2)^{\frac{n-3}{2}}\int_{1}^{\infty} F\big(\frac{\tau+s|\xi|}{2},\frac{\tau+s|\xi|}{2}\big)(\tau^2-x^2|\xi|^2)(1-|x|^2)^{\frac{n-3}{2}}dx \tag{219}$$

Proof : Observe that in the case $\mp = -$ the measure $\delta(\tau-|\eta|-|\xi-\eta|)$ is supported on the *ellipsoid* of revolution with foci at 0 and ξ, $\mathcal{E}(\tau,\xi)=\{\eta\in\mathbb{R}^n : |\eta|+|\xi-\eta|=\tau\}$,. In this case $|\xi|\le\tau$. In the $\mp = +$ the measure $\delta(\tau-|\eta|+|\xi-\eta|)$ is supported in the *hyperboloid* of revolution with foci at 0 and ξ, $\mathcal{H}(\tau,\xi)=\{\eta\in\mathbb{R}^n : |\eta|-|\xi-\eta|=\tau\}$, which is an unbounded hypersurface with infinite volume. In this case $|\xi|^2\le\tau^2$. In the sense of distributions, we have the identity

$$\begin{aligned} \delta(\tau-|\eta|\mp|\xi-\eta|) &= \delta\big(\frac{(\tau-|\eta|)^2-|\xi-\eta|^2}{2(\tau-|\eta|)}\big) \\ &= 2(\tau-|\eta|)\delta\big((\tau-|\eta|)^2-|\xi-\eta|^2\big) \\ &= 2(\tau-|\eta|)\delta\big(\tau^2-|\xi|^2-2\tau\lambda+2\lambda\xi\cos\theta\big) \\ &= 2(\tau-|\eta|)\delta\big(\tau^2-|\xi|^2-2\tau\lambda+2a|\xi|\big) \end{aligned}$$

with a the cosine of the angle between η and ξ. Thus, for fixed τ and ξ we must have, on the support of the measure,

$$a \quad = \quad -\frac{\tau^2-|\xi|^2-2\tau\lambda}{2|\xi|\lambda} \tag{220}$$

Observe that in the ellipsoidal case a can take any values in the interval $[-1,1]$ and thus, since $\lambda=\frac{\tau^2-|\xi|^2}{2(\tau-a|\xi|)}$, we have $\frac{\tau-|\xi|}{2}\le\lambda\le\frac{\tau+|\xi|}{2}$. On the other hand, in the hyperboloidal case when $|\xi|^2>\tau^2$, we must also have the restriction,

$$\frac{\tau}{|\xi|}\le a.$$

and thus, $\lambda=\frac{-\tau^2+|\xi|^2}{2(-\tau+a|\xi|)}\ge\frac{\tau+|\xi}{2}$.

Thus, since $d\eta=\lambda^{n-1}d\lambda dS_\omega=(1-a^2)^{\frac{n-3}{2}}\lambda^{n-1}d\lambda dS_{\omega'}$,

$$\begin{aligned}
J_F^- \quad &= \quad \frac{1}{|\xi|}\int_{\frac{\tau-|\xi|}{2}}^{\frac{\tau+|\xi|}{2}} F(\lambda,\tau-\lambda)(\tau-\lambda)\lambda^{n-2}\big[1-\big(\frac{\tau^2-|\xi|^2-2\tau\lambda}{2|\xi|\lambda}\big)\big]^{\frac{n-3}{2}}d\lambda\\
&= \quad \frac{(\tau^2-|\xi|^2)^{\frac{n-3}{2}}}{|\xi|^{n-2}}\int_{\frac{\tau-|\xi|}{2}}^{\frac{\tau+|\xi|}{2}} F(\lambda,\tau-\lambda)(\tau-\lambda)\lambda\big[\big(\frac{\tau+|\xi|}{2}-\lambda\big)\big(\lambda-\frac{\tau-|\xi|}{2}\big)\big]^{\frac{n-3}{2}}
\end{aligned}$$

At last we perform the change of variables $x=\frac{2\lambda-\tau}{|\xi|}$ to derive the desired formula (218). The proof for (219) follows in the same manner.

∎

13.4. Improved Bilinear Strichartz. Consider two solutions of the homogeneous wave equations, $\Box u=\Box v=0$. For simplicity, and without loss of generality, we assume that u,v verify the reduced initial data at $t=0$,

$$u(0,x)=f(x),\ v(0,x)=g(x),\ \partial_t u(0,x)=\partial_t v(0,x)=0.$$

We consider estimates of the form,

$$\|D^{-b}(uv)\|_{L_t^{q/2}L_x^{r/2}} \quad \lesssim \quad \|f\|_{\dot H^a}\|g\|_{\dot H^a}$$

with (q,r) an acceptable pair. By dimensional analysis and recalling the exponent $\gamma=n((\frac12-\frac1r))-\frac1q$ in (163), we must have,

$$2a \quad = \quad -b+2\big(n(\frac12-\frac1r)-\frac1q\big)=-b+2\gamma \tag{221}$$

We decompose the product $u\cdot v$ by the trichotomy formula,

$$\begin{aligned}
u\cdot v \quad &= \quad \sum_{\mu<\lambda}u_\mu v_\lambda+\sum_{\mu<\lambda}v_\mu u_\lambda+\sum_{\mu\le\lambda}P_\mu(u_\lambda v_\lambda)\\
&= \quad (u\cdot v)_{LH}+(u\cdot v)_{HL}+(u\cdot v)_{HH}
\end{aligned}$$

Here $\mu,\lambda\in 2^{\mathbb{Z}}$, $u_\lambda=P_\lambda u$ and P_λ the usual LP projections. Now,

$$\|D^{-b}(uv)_{LH}\|_{L_t^{q/2}L_x^{r/2}} \quad \le \quad \sum_{\mu\le\lambda}\lambda^{-b}\|u_\mu v_\lambda\|_{L_t^{q/2}L_x^{r/2}}\le\sum_{\mu\le\lambda}\lambda^{-b}\|u_\mu\|_{L_t^qL_x^r}\|v_\lambda\|_{L_t^qL_x^r}$$

in view of the Strichartz estimates of the previous section

$$\|u_\mu\|_{L^q_t L^r_x} \lesssim \mu^{(\gamma-a)}\|f_\mu\|_{\dot H^a} = \mu^{b/2}\|f_\mu\|_{\dot H^a}$$
$$\|v_\lambda\|_{L^q_t L^r_x} \lesssim \lambda^{(\gamma-a)}\|g_k\|_{\dot H^a} = \lambda^{b/2}\|g_\lambda\|_{\dot H^\alpha}$$

and therefore, for $b > 0$,

$$\begin{aligned}\|D^{-b}(uv)_{LH}\|_{L^{q/2}_t L^{r/2}_x} &\lesssim \sum_{\mu\le\lambda}\big(\frac{\mu}{\lambda}\big)^b\|f_\mu\|_{\dot H^a}\|g_\lambda\|_{\dot H^a}\\ &\lesssim \|f\|_{\dot H^a}\|g\|_{\dot H^a}\end{aligned}$$

By symmetry,

$$\|D^{-b}(uv)_{LH}\|_{L^{q/2}_t L^{r/2}_x} \lesssim \|f\|_{\dot H^a}\|g\|_{\dot H^a}$$

It thus only remains to estimate the high-high term $\|(u\cdot v)_{HH}\|_{L^{q/2}_t L^{r/2}_x}$. This requires a more subtle argument based on theorem **??**. We write,

$$\|D^{-b}(u\cdot v)_{HH}\|_{L^{q/2}_t L^{r/2}_x} \lesssim \sum_{\mu\le\lambda}\mu^{-b}\|P_\mu(u_\lambda v_\lambda)\|_{L^{q/2}_t L^{r/2}_x}$$

If we use the standard Strichartz estimate, i.e.,

$$\begin{aligned}\|P_\mu(u_\lambda v_\lambda)\|_{L^{q/2}_t L^{r/2}_x} &\lesssim \|u_\lambda\|_{L^q_t L^r_x}\|v_\lambda\|_{L^q_t L^r_x}\lambda^{2(\gamma-a)}\|f\|_{\dot H^a}\|g\|_{\dot H^a}\\ &= \lambda^b\|f\|_{\dot H^a}\|g\|_{\dot H^a}\end{aligned}\tag{222}$$

we would derive,

$$\|D^{-b}(u\cdot v)_{HH}\|_{L^{q/2}_t L^{r/2}_x} \lesssim \sum_{\mu\le\lambda}\lambda^b\mu^{-b}\|f\|_{\dot H^a}\|g\|_{\dot H^a}$$

which diverges. We need to replace (222) by a stronger estimate which takes into account the presence of P_μ in front of $u_\lambda v_\lambda$. To achieve this, we need first to exploit some orthogonality properties. We decompose the the data f_λ, g_λ, in Fourier space, into pieces supported on cubes of size μ, $f_\lambda = \sum_Q f_Q$, $g_\lambda = \sum_Q g_Q$ and denote by u_Q, v_Q the corresponding solutions. Clearly the decomposition commutes with the wave operator $\Box$. Thus, $u_\lambda \sim \sum_Q u_Q$, $v_\lambda \sim \sum_Q v_Q$ and

$$P_\mu(u_\lambda\cdot v_\lambda) \sim \sum_{Q_1,Q_2} P_\mu(u_{Q_1}v_{Q_2})$$

Observe that $P_\mu(u_{Q_1}u_{Q_2}) \neq 0$ only if Q_1+Q_2 intersects the region of frequencies of size μ where P_μ is supported. For each cube Q_1, of size μ, there are only a finite number (which depends only on n) of cubes Q_2 for which this happens. Morally, by enlarging the cubes if necessary we may assume that $Q_2 = -Q_1$ and thus,

$$P_\mu(u_\lambda\cdot v_\lambda) \sim \sum_Q u_Q v_{-Q}.$$

Hence,

$$\|P_\mu(u_\lambda v_\lambda)\|_{L^{q/2}_t L^{r/2}_x} \lesssim \sum_Q \|u_Q v_{-Q}\|_{L^q_t L^r_x} \lesssim \sum_Q \|u_Q\|_{L^q_t L^r_x}\|v_{-Q}\|_{L^q_t L^r_x}.$$

We are now in a position to apply theorem 12.12. Thus,

$$\|u_Q\|_{L^q_t L^r_x} \lesssim \big(\frac{\mu}{\lambda}\big)^{\frac12-\frac1r}\|f_Q\|_{\dot H^\gamma}$$

and similarly for v_{-Q}. Hence,

$$\begin{aligned}\|P_\mu(u_\lambda v_\lambda)\|_{L_t^{q/2}L_x^{r/2}} &\lesssim (\frac{\mu}{\lambda})^{1-\frac{2}{r}}\sum_Q \|f_Q\|_{\dot H^\gamma}\|g_Q\|_{\dot H^\gamma}\\ &\lesssim (\frac{\mu}{\lambda})^{1-\frac{2}{r}}\|f_\lambda\|_{\dot H^\gamma}\|g_\lambda\|_{\dot H^\gamma}\\ &\lesssim (\frac{\mu}{\lambda})^{1-\frac{2}{r}}\lambda^{2\gamma-2a}\|f_\lambda\|_{\dot H^a}\|g_\lambda\|_{\dot H^a}\\ &\lesssim (\frac{\mu}{\lambda})^{1-\frac{2}{r}}\lambda^{b}\|f_\lambda\|_{\dot H^a}\|g_\lambda\|_{\dot H^a}\end{aligned}$$

and, consequently,

$$\begin{aligned}\|D^{-b}(u\cdot v)_{HH}\|_{L_t^{q/2}L_x^{r/2}} &\lesssim \sum_{\mu<\lambda}(\frac{\mu}{\lambda})^{1-\frac{2}{r}-b}\|f_\lambda\|_{\dot H^a}\|g_\lambda\|_{\dot H^a}\\ &\lesssim \|f\|_{\dot H^a}\|g\|_{\dot H^a}\end{aligned}$$

provided that $b < 1-\frac{2}{r}$. We have just proved the following bilinear estimate, see [**13**].

THEOREM 13.5. *The following estimate* [28] *holds for solutions* $\Box u = \Box v = 0$*, any admissible pair* (q,r) *and any* $0 \le b < 1-\frac{2}{r}$,

$$\|D^{-b}(u\cdot v)_{L_t^{q/2}L_x^{r/2}} \lesssim \|u[0]\|_{\dot H^a}\|v[0]\|_{\dot H^a} \tag{223}$$

provided that the dimensional condition,

$$a = -\frac{b}{2}+\gamma, \qquad \gamma = n(\frac{1}{2}-\frac{1}{r})-\frac{1}{q} \tag{224}$$

13.6. Bilinear estimates for null forms. In this subsection we discuss the simplest bilinear estimates for null quadratic forms, see [**8**], [**11**], [**12**] and [**2**].

DEFINITION 13.7. Let u, v be two smooth solutions of $\Box = \Box v = 0$ on $\mathbb{R}^{n+1}$. The standard null quadratic forms are $Q_0(u,v) = -\partial_t u\,\partial_t v + \sum_{i=1}^n \partial_i u\partial_i v$, as well as $Q_{ij}(u,v) = \partial_i u\partial_j v - \partial_i v\partial_j u$, and $Q_{0i}(u,v) = \partial_i u\partial_t v - \partial_i v\partial_t u$ for $i,j = 1,\dots,n$.

THEOREM 13.8. *For any null form* Q *and any solutions to* $\Box = \Box v = 0$ *on* $\mathbb{R}^{n+1}$, $n \ge 2$*, we have,*

$$\|Q(u,v)\|_{L^2(\mathbb{R}^{n+1})} \lesssim \|u[0]\|_{\dot H^1(\mathbb{R}^n)}\|v[0]\|_{H^{\frac{n+1}{2}}(\mathbb{R}^n)} \tag{225}$$

REMARK 13.9. Without loss of generality, it suffices to consider the reduced initial value problems

$$u(0,x) = f(x),\ v(0,x) = g(x),\ \partial_t u(0,x) = \partial_t v(0,x) = 0 \tag{226}$$

In what follows we show how to deduce the estimate (13.8) from a more general form of bilinear estimates presented in the next section.

DEFINITION 13.10. Let D^α, D_+^α and D_-^α be the operators in $\mathbb{R}^{n+1}$ defined by the multipliers with symbols, respectively

$$|\xi|^\alpha, \quad (|\tau|+|\xi|)^\alpha, \quad \big||\tau|-|\xi|\big|^\alpha.$$

[28] Here $\|u[0]\|_{\dot H^a} = \|u(0)\|_{\dot H^a} + \|\partial_t u(0)\|_{\dot H^a}$

Observe that we can write, for any smooth functions u, v,

$$2Q_0(u,v) = \Box(uv) - \Box uv - u\Box u$$

Thus, if $\Box u = \Box v = 0$, using Plancherel,

$$\begin{aligned}\|Q_0(u,v)\|_{L^2(\mathbb{R}^{n+1})} &\le \frac{1}{2}\|\Box(uv)\|_{L^2(\mathbb{R}^{n+1})} = \frac{1}{2}(2\pi)^{-n}\|(\tau^2-|\xi|^2)\widetilde{uv}\|_{L^2(\mathbb{R}^{n+1})}\\ &\lesssim \|D_+D_-(uv)\|_{L^2(\mathbb{R}^{n+1})}\end{aligned}$$

Therefore,

$$\|Q_0(u,v)\|_{L^2(\mathbb{R}^{n+1})} \le \|D_+D_-(uv)\|_{L^2(\mathbb{R}^{n+1})} \tag{227}$$

Thus, in the case of the null form Q_0, theorem 13.8 reduces to,

$$\|D_+D_-(uv)\|_{L^2(\mathbb{R}^{n+1})} \lesssim \|u[0]\|_{\dot H^1(\mathbb{R}^n)}\|v[0]\|_{H^{\frac{n+1}{2}}(\mathbb{R}^n)} \tag{228}$$

which is a special case of theorem 13.15.

Below we show that similar estimates hold true for the other null forms, Q_{ij}, Q_{0i}.

Remark 13.11. Given a solution u of $\Box u = 0$ with initial data $u(0,x) = f(x)$, $\partial_t u(0,x) = 0$ we denote by u' the solution of the same equation with data $u'(0,x) = f'(x)$, $\partial_t u'(0,x) = 0$ where $f' = \mathcal{F}^{-1}(|\hat f|)$. Observe, of course, that $\|f'\|_{\dot H^a} = \||f\|_{\dot H^a}$ and thus, from the point of view of the L^2 type estimates we are considering u and u' are indistinguishable.

Proposition 13.12. *Let u, v be smooth solutions of the homogeneous wave equation with initial . The following estimates hold true:*

$$\|Q_{ij}(u,v)\|_{L^2(\mathbb{R}^{n+1})} \lesssim \|D^{1/2}D_-^{1/2}(D^{1/2}u'\cdot D^{1/2}v')\|_{L^2(\mathbb{R}^{n+1})} \tag{229}$$

$$\|Q_{0i}(u,v)\|_{L^2(\mathbb{R}^{n+1})} \lesssim \|D_+^{1/2}D_-^{1/2}(D^{1/2}u'\cdot D^{1/2}v')\|_{L^2(\mathbb{R}^{n+1})} \tag{230}$$

Proof : We first decompose, as before, $u = u^+ + u^-$, $v = v^+ + v^-$ We write, in Fourier variables,

$$Q_{ij}\widetilde{(u^+, v^\pm)}(\tau,\xi) = \int q_{ij}(\eta,\xi-\eta)\delta(\tau - |\eta| \pm |\xi-\eta|)\hat f(\eta)\hat g(\xi-\eta)d\eta$$

where $q_{ij}(\eta,\xi-\eta) = \eta_i(\xi-\eta)_j - \eta_j(\xi-\eta)_i = (\xi\wedge\eta)_{ij}$ We now rely on the following simple lemma.

Lemma 13.13. *The following inequalities hold true,*

$$|\xi\wedge\eta| \lesssim |\xi|^{1/2}|\eta|^{1/2}|\xi+\eta|^{1/2}(|\xi|+|\eta|-|\xi+\eta|)^{1/2} \tag{231}$$

$$|\xi\wedge\eta| \lesssim |\xi|^{1/2}|\eta|^{1/2}|\xi+\eta|^{1/2}(|\xi+\eta|-\big||\xi|-|\eta|\big|)^{1/2} \tag{232}$$

We have indeed,

$$\begin{aligned}4|\xi\wedge\eta|^2 &= 4(|\xi||\eta| - \xi\cdot\eta)(|\xi||\eta|+\xi\cdot\eta)\\ &= ((|\xi|+|\eta|-|\xi+\eta|)((|\xi|+|\eta|+|\xi+\eta|)\\ &\quad (|\xi+\eta| - \big||\xi|-|\eta|\big|)(|\xi+\eta|+\big||\xi|-|\eta|\big|)\end{aligned}$$

from which the lemma immediately follows.

Therefore, in both cases, $|Q_{ij}\widetilde{(u^+, v^\pm)}(\tau,\xi)|$ can be bounded by the expression,

$$\begin{aligned}
&\int |q_{ij}(\eta,\xi-\eta)|\ \delta(\tau-|\eta|\pm|\xi-\eta|)|\hat{f}(\eta)|\,|\hat{g}(\xi-\eta)|d\eta \\
&\lesssim \big||\tau|-|\xi|\big|^{1/2}|\xi|^{1/2}\int \delta(\tau-|\eta|\pm|\xi-\eta|)\,|\eta|^{1/2}|\xi-\eta|^{1/2}|\hat{f}(\eta)|\,|\hat{g}(\eta)|d\eta \\
&= D^{1/2}D_-^{1/2}(D^{1/2}u'D^{1/2}v')
\end{aligned}$$

as desired.

■

According to proposition 13.12, theorem 13.8 reduces, for $Q = Q_{ij}$, resp. $Q = Q_{0i}$, to the statements,

$$\begin{aligned}
\|D^{1/2}D_-^{1/2}(u\cdot v)\|_{L^2(\mathbb{R}^{n+1})} &\lesssim \|u[0]\|_{\dot{H}^{1/2}}\cdot\|u[0]\|_{\dot{H}^{n/2}} \\
\|D_+^{1/2}D_-^{1/2}(u\cdot v)\|_{L^2(\mathbb{R}^{n+1})} &\lesssim \|u[0]\|_{\dot{H}^{1/2}}\cdot\|u[0]\|_{\dot{H}^{n/2}}
\end{aligned}$$

which are particular cases of theorem 13.15.

13.14. General Bilinear Estimates. In this section we investigate the space-time regularity properties of products of solutions to the homogeneous wave equation. Let D^α, D_+^α and D_-^α be the multipliers with symbols

$$|\xi|^\alpha, \quad (|\tau|+|\xi|)^\alpha, \quad \big||\tau|-|\xi|\big|^\alpha$$

respectively. We are interested in blinear estimates of the form

$$\left\|D^{\beta_0}D_+^{\beta_+}D_-^{\beta_-}(uv)\right\|_{L^2(\mathbf{R}^{1+n})} \lesssim \|D^{\alpha_1}u[0]\|_{L^2}\,\|D^{\alpha_2}v[0]\|_{L^2} \tag{233}$$

*** discuss history of this type of estimates with references

THEOREM 13.15. *Estimate* (233) *holds true, for arbitrary solutions of the homogeneous equations $\Box u = \Box v = 0$, in any space dimensions $n \geq 2$. if and only if the exponents $\alpha_1, \alpha_2, \beta_0, \beta_+, \beta_-$ satisfy the following conditions:*

$$\beta_0+\beta_++\beta_- = \alpha_1+\alpha_2-\frac{n-1}{2}, \tag{234}$$

$$\beta_- \geq -\frac{n-3}{4}, \tag{235}$$

$$\beta_0 > -\frac{n-1}{2}, \tag{236}$$

$$\alpha_i \leq \beta_- + \frac{n-1}{2}, \quad i=1,2, \tag{237}$$

$$\alpha_1+\alpha_2 \geq \frac{1}{2}, \tag{238}$$

$$(\alpha_i,\beta_-) \neq \left(\frac{n+1}{4}, -\frac{n-3}{4}\right), \quad i=1,2, \tag{239}$$

$$(\alpha_1+\alpha_2,\beta_-) \neq \left(\frac{1}{2}, -\frac{n-3}{4}\right). \tag{240}$$

*** discuss special cases

As before, it suffices to prove the theorem for the case when u, v verify the standard initial value problem

$$u(0,x) = f(x),\ v(0,x) = g(x),\ \partial_t u(0,x) = \partial_t v(0,x) = 0.$$

We also need to decompose, $u = u^+ + u^-$, $v = v^+ + v^-$ with, $\tilde{u}_\pm(\tau,\xi) = \delta(\tau \mp |\xi|)\hat{f}(\xi)$, $\tilde{v}_\pm(\tau,\xi) = \delta(\tau \mp |\xi|)\hat{g}(\xi)$ Thus, taking the spacetime Fourier transform,

$$\begin{aligned}
\widetilde{u_+ v_+}(\tau,\xi) &= \int \delta(\tau - \eta - |\xi - \eta|)\hat{f}(\xi - \eta)\hat{g}(\eta)d\eta && (241)\\
\widetilde{u^+ v^-}(\tau,\xi) &= \int \delta((\tau + |\eta| - |\xi - \eta|)\hat{f}(\eta)\hat{g}(\xi - \eta)d\eta && (242)\\
\widetilde{u^- v^+}(\tau,\xi) &= \int \delta((\tau - |\eta| + |\xi - \eta|)\hat{f}(\eta)\hat{g}(\xi - \eta)d\eta\\
\widetilde{u^- v^-}(\tau,\xi) &= \int \delta((\tau + |\eta| + |\xi - \eta|)\hat{f}(\eta)\hat{g}(\xi - \eta)d\eta
\end{aligned}$$

In the proof below, by symmetry, it will suffice to consider the cases $u^+v^\pm$. The two integrals look similar but have different behaviors: (241) is an integration over the *ellipsoid* of revolution with foci at 0 and ξ,

$$\mathcal{E}(\tau,\xi) = \{\eta \in \mathbb{R}^n : |\eta| + |\xi - \eta| = \tau\}, \tag{243}$$

which is a compact manifold; (242) is an integration over the *hyperboloid* of revolution with foci at 0 and ξ,

$$\mathcal{H}(\tau,\xi) = \{\eta \in \mathbb{R}^n : |\eta| - |\xi - \eta| = \tau\}, \tag{244}$$

which is an unbounded manifold with infinite volume. Also, notice that $\widetilde{u^+v^+}$ is supported on the region $\tau \geq |\xi|$, while $\widetilde{u^+v^-}$ is supported on the region $|\tau| \leq |\xi|$.

We decompose uv by the trichotomy formula,

$$\begin{aligned}
u^+v^\pm &\sim \sum_{\mu<\lambda} u^+_\mu v^\pm_\lambda + \sum_{\mu<\lambda} u^+_\lambda v^\pm_\mu + \sum_{\mu\leq\lambda} P_\mu(u^+_\lambda v^\pm_\lambda) && (245)\\
&= \Sigma_1 + \Sigma_2 + \Sigma_3 && (246)
\end{aligned}$$

With the exception of some end points the bilinear estimates of theorem 13.15 follow from their following dyadic version.

THEOREM 13.16. *Let $0 < \mu \lesssim \lambda$ and $\gamma > -\frac{n-3}{4}$. Then,*

$$\begin{aligned}
\left\|D^\gamma_-\left(u^+_\lambda v^\pm_\mu\right)\right\|_{L^2(\mathbb{R}^{1+3})} &\lesssim \mu^{(\gamma+\frac{n-1}{2})}\|f_\lambda\|\cdot\|g_\mu\| && (247)\\
\left\|D^\gamma_- P_\mu(u^+_\lambda v^+_\lambda)\right\|_{L^2(\mathbb{R}^{1+3})} &\lesssim \mu^{\frac{n-1}{2}}\lambda^\gamma\|f_\lambda\|\cdot\|g_\lambda\| && (248)\\
\left\|D^\gamma_- P_\mu(u^+_\lambda v^-_\lambda)\right\|_{L^2(\mathbb{R}^{1+3})} &\lesssim \mu^{\gamma+\frac{n-2}{2}}\lambda^{\frac{1}{2}}\|f_\lambda\|\cdot\|g_\lambda\| && (249)
\end{aligned}$$

Assuming theorem 13.16 to be true we prove below a slightly weaker version of theorem 13.15. In fact we replace the main non-scaling conditions (235)– (238) by the following,

$$\beta_- > -\frac{n-3}{4}, \tag{250}$$

$$\beta_0 > -\frac{n-1}{2}, \tag{251}$$

$$\alpha_i < \beta_- + \frac{n-1}{2}, \quad i = 1, 2, \tag{252}$$

$$\alpha_1 + \alpha_2 > \frac{1}{2}. \tag{253}$$

Proof : We have, $\Sigma_1 = \sum_{\mu<\lambda} u_\mu^+ v_\lambda^\pm$. Clearly,

$$\|D^{\beta_0} D_+^{\beta_+} D_-^{\beta_-} \Sigma_1\|_{L^2} \quad \lesssim \quad \sum_{\mu<\lambda} \lambda^{\beta_0+\beta_+} \|D_-^{\beta_-}(u_\mu v_\lambda)\|_{L^2}$$

Using the dyadic estimate (247) with $\gamma = \beta_-$ we derive,

$$\|D_-^{\beta_-}(u_\mu v_\lambda)\|_{L^2} \quad \lesssim \quad \sum_{\mu<\lambda} \|u_\mu v_\lambda\|_{L^2} \lesssim \sum_{\mu<\lambda} \mu^{\beta_- + \frac{n-1}{2}} \lambda^{-\alpha_1} \mu^{-\alpha_2} \|D^{\alpha_1} f_\lambda\| \|D^{\alpha_2} g_\mu\|$$

Therefore,

$$\|D^{\beta_0} D_+^{\beta_+} D_-^{\beta_-} \Sigma_1\|_{L^2} \quad \lesssim \quad \sum_{\mu<\lambda} \lambda^{\beta_0+\beta_+-\alpha_1} \mu^{\frac{(n-1)}{2}+\beta_- - a_2} \|D^{\alpha_1} f_\lambda\| \, \|D^{\alpha_2} g_\mu\|$$

We now apply condition (252) which we write in the form, $\varepsilon := \frac{(n-1)}{2} + \beta_- - \alpha_2 > 0$. According to the dimensional condition (234),

$$\beta_0 + \beta_+ - \alpha_1 = -\big(\frac{(n-1)}{2} + \beta_- - \alpha_2\big) = -\varepsilon$$

Therefore,

$$\|D^{\beta_0} D_+^{\beta_+} D_-^{\beta_-} \Sigma_1\|_{L^2} \quad \lesssim \quad \sum_{\mu<\lambda} \big(\frac{\mu}{\lambda}\big)^\varepsilon \|D^{\alpha_1} f_\lambda\| \|D^{\alpha_2} g_\mu\| \lesssim \|D^{\alpha_1} f\| \, \|D^{\alpha_2} g\|$$

as desired. The term Σ_2 can be estimated in precisely the same manner.

To estimate Σ_3 we write,

$$\|D^{\beta_0} D_+^{\beta_+} D_-^{\beta_-} \Sigma_3\|_{L^2} \quad \lesssim \quad \sum_{\mu\le\lambda} \mu^{\beta_0} \|D_+^{\beta_+} D_-^{\beta_-} P_\mu(u_\lambda^+ v_\lambda^\pm)\|_{L^2}$$

The operator D_+ behaves differently in the cases $++$ and $+-$. We first estimate in the $++$ case , when the symbol $|\tau| + |\xi|$ is dominated by $|\tau$ which is no better in size than λ. Thus, applying (248) with $\gamma = \beta_-$,

$$\|D_+^{\beta_+} D_-^{\beta_-} P_\mu(u_\lambda^+ v_\lambda^+)\|_{L^2} \quad \lesssim \quad \lambda^{\beta_+} \lambda^{\beta_-} \mu^{\frac{n-1}{2}} \lambda^{-\alpha_1} \lambda^{-\alpha_2} \|D^{\alpha_1} f_\lambda\| \, \|D^{\alpha_2} g_\lambda\|$$

from which,

$$\|D^{\beta_0} D_+^{\beta_+} D_-^{\beta_-} \Sigma_3\|_{L^2} \quad \lesssim \quad \sum_{\mu\le\lambda} \mu^{\beta_0+\frac{n-1}{2}} \lambda^{\beta_+ + \beta_- - \alpha_1 - \alpha_2} \|D^{\alpha_1} f_\lambda\| \, \|D^{\alpha_2} g_\lambda\|$$

We now use (251), i.e. $\beta_0 > -\frac{n-1}{2}$, to set $\varepsilon = \beta_0 + \frac{n-1}{2} > 0$, and the scaling condition (234) to write,

$$\varepsilon := \beta_0 + \frac{n-1}{2} = -\beta_+ - \beta_- - \alpha_1 - \alpha_2$$

Thus,

$$\begin{aligned}\|D^{\beta_0} D_+^{\beta_+} D_-^{\beta_-} \Sigma_3\|_{L^2} &\lesssim \sum_{\mu\le\lambda} \big(\frac{\mu}{\lambda}\big)^{\varepsilon} \|D^{\alpha_1} f_\lambda\|\,\|D^{\alpha_2} g_\lambda\| \lesssim \sum_\lambda \|D^{\alpha_1} f_\lambda\|\,\|D^{\alpha_2} g_\lambda\| \\ &\le \|D^{\alpha_1} f\|\,\|D^{\alpha_2} g\|\end{aligned}$$

as desired.

In the $(+-)$ case $\widetilde{P_\mu(u_\lambda^+ v_\lambda^-)}$ is supported in the region $|\tau|\le|\xi|$ where the symbol $|\tau|+|\xi|$ is dominated by $|\xi|\sim\mu$. Hence, applying (249) with $\gamma=\beta_-$,

$$\|D_+^{\beta_+} D_-^{\beta_-} P_\mu(u_\lambda^+ v_\lambda^-)\|_{L^2} \lesssim \mu^{\beta_+}\lambda^{1/2}\mu^{\beta_- \frac{n-2}{2}}\lambda^{-\alpha_1}\lambda^{-\alpha_2}\|D^{\alpha_1} f_\lambda\|\,\|D^{\alpha_2} g_\lambda\|$$

and thus,

$$\|D^{\beta_0} D_+^{\beta_+} D_-^{\beta_-} \Sigma_3\|_{L^2} \lesssim \sum_{\mu\le\lambda} \mu^{\beta_0+\beta_++\beta_- \frac{n-2}{2}}\lambda^{\frac12-\alpha_1-\alpha_2}\|D^{\alpha_1} f_\lambda\|\,\|D^{\alpha_2} g_\lambda\|$$

We now make use of condition (253) according to which $\varepsilon=\alpha_1+\alpha_2-\frac12>0$. Hence, in view of the scaling condition,

$$\beta_0+\beta_++\beta_-\frac{n-2}{2}=\varepsilon$$

and thus,

$$\|D^{\beta_0} D_+^{\beta_+} D_-^{\beta_-} \Sigma_3\|_{L^2} \lesssim \sum_{\mu\le\lambda} \big(\frac{\mu}{\lambda}\big)^{\varepsilon}\|D^{\alpha_1} f_\lambda\|\,\|D^{\alpha_2} g_\lambda\| \lesssim \|D^{\alpha_1} f|\,\|D^{\alpha_2} g\|$$

as desired.

∎

We give below a sketch of the proof of theorem 13.16

Proof : We start with a proof of (247). By Cauchy -Schwartz,

$$\begin{aligned}|D_-^{\gamma}(u_\lambda\cdot v_\mu)|^2 &\lesssim J_{\lambda,\mu}(\tau,\xi)\cdot\int \delta(\tau-|\eta|\mp|\xi-\eta|)\,|\hat f_\lambda(\eta)|^2\,|\hat g_\mu(\xi-\eta)|^2\,d\eta \\ J_{\lambda,\mu}(\tau,\xi) &:= ||\tau|-|\xi||^{2\gamma}\int \delta(\tau-|\eta|\mp|\xi-\eta|)\chi_\lambda(\eta)\,\chi_\mu(\xi-\eta)d\eta\end{aligned}$$

where χ_μ is a fixed C^∞ function, $0\le|\chi_\mu|\le1$, supported in the dyadic piece $|\xi|\sim\mu$. We claim that the proof of (247) reduces to the proof of the bound,

$$|J_{\lambda,\mu}(\tau,\xi)|\lesssim\mu^{2(\gamma+1)} \tag{254}$$

Indeed, assuming this to be true, we deduce,

$$\begin{aligned}\|D_-^{\gamma}(u_\lambda\cdot v_\mu)\|_{L^2}^2 &\lesssim \mu^{2(\gamma+1)}\int\int \delta(\tau-|\eta|\mp|\xi-\eta|)\,|\hat f_\lambda|^2(\eta)|\hat g_\mu(\xi-\eta)|^2 d\eta d\xi d\tau \\ &\lesssim \mu^{2(\gamma+1)}\cdot\|f_\lambda\|_{L^2}^2\cdot\|g_\mu\|_{L^2}^2\end{aligned}$$

To prove estimate (254) we observe that, on the support of the corresponding integral we have $|\eta|\pm|\xi-\eta|=\tau$. Thus,

$$\begin{aligned}\big|\tau-|\xi|\big| &= \big||\eta|\pm|\xi-\eta|-|\xi|\big|\le|\xi-\eta|+\big||\eta|-|\xi|\big| \\ &\le 2|\xi-\eta|\lesssim\mu.\end{aligned}$$

Thus,

$$|J_{\lambda,\mu}(\tau,\xi)| \lesssim \mu^{\gamma} \int_{|\xi-\eta|\lesssim\mu, |\eta|\sim\lambda} \delta(\tau - |\eta| \pm |\xi-\eta|)d\eta$$

Since $\mu << \lambda$ we also must have $\xi \sim \lambda$ on the support of the integral. We can always rescale the integral and thus assume that $\mu << 1$ and $|\xi| \sim 1$.

Estimate (254) follows easily from the following,

LEMMA 13.17. *Let $|\mu << 1$. Then, for all ξ, such that $1/2 \leq |\xi| \leq 2$, and all $\tau \neq |\xi|$,*

$$\int_{|\eta|\lesssim\mu} \delta(\tau - |\eta| \pm |\xi-\eta|)d\eta \lesssim \big||\tau| - |\xi|\big|^{\frac{n-3}{2}} \mu^{\frac{n+1}{2}} \tag{255}$$

Proof : Consider first the $++$ case, when $|\xi| \leq \tau \leq 4$. We can Introduce polar coordinates and write the integral (255) in the form,

$$I \quad := \quad \int_{\mathbb{S}^{n-1}} \int_0^{\mu} \delta(\tau - \lambda - |\xi - \lambda\omega|)\lambda^{n-1} d\lambda dA(\omega)$$

Clearly, on the support of the measure, $|\xi - \lambda\omega|^2 = (\tau-\lambda)^2$. We can thus solve for λ and find, $\lambda = \frac{\tau^2-|\xi|^2}{\tau-\xi\cdot\omega}$. We are led to

Thus, writing $\tau - \xi\cdot\omega = \tau - x|\xi|$ with $x = \cos\theta$ the cosine of the angle between ξ and ω,

$$I \quad = \quad \int_{-1}^{1} \left(\min(\mu, \frac{\tau^2-|\xi|^2}{\tau - x|\xi|}) \right)^{n-1} (1-x^2)$$

We split $\lambda^{n-1} = \lambda^{\frac{n+1}{2}} \cdot \lambda^{\frac{n-3}{2}}$. Thus,

$$\begin{aligned} I \quad &\lesssim \quad \mu^{\frac{n+1}{2}} (\tau^2 - |\xi|^2|)^{\frac{n-3}{2}} \int_{\mathbb{S}^{n-1}} \frac{1}{(\tau - \xi\cdot\omega)^{\frac{n-3}{2}}} d\omega \\ &\lesssim \quad \mu^{\frac{n+1}{2}} (\tau^2 - |\xi|^2|)^{\frac{n-3}{2}} \lesssim \mu^{\frac{n+1}{2}} (\tau - |\xi|)^{\frac{n-3}{2}} \end{aligned}$$

since $\tau > |\xi|$.

In the $+-$ we have $1 \leq |\tau| \leq |\xi|$ and the integral (255) becomes,

$$I \quad := \quad \int_{\mathbb{S}^{n-1}} \int_0^{\mu} \delta(\tau - \lambda + |\xi - \lambda\omega|)\lambda^{n-1} d\lambda dA_{\omega}$$

Proceeding as before, $|\xi - \lambda\omega|^2 = (\tau-\lambda)^2$, thus $\lambda(\xi,\omega) = \frac{\tau^2-|\xi|^2}{\tau-\xi\cdot\omega}$ and,

$$I \lesssim \mu^{\frac{n+1}{2}} (|\xi| - |\tau|)^{\frac{n-3}{2}} \int_{\omega\in\mathbb{S}^{n-1};\ \lambda(x,\omega)\leq\mu} \frac{1}{(|\tau - \xi\cdot\omega|)^{\frac{n-3}{2}}} d\omega$$

■

■

CHAPTER 2

BASIC TOOLS IN LINEAR PDE

1. Laplace Equation in $\mathbb{R}^n$

The Laplace operator $\Delta = \Delta_{\mathbb{R}^n} = \partial_1^2+\partial_2^2+\ldots+\partial_n^2$ is the Laplace Beltrami operator of the euclidean space $\mathbb{R}^n$. Recall that the latter comes equipped with the standard coordinates $x = (x^1, x^2, \ldots x^n)$ relative to which the euclidean metric has the form,

$$ds^2 = (dx^1)^2 + (dx^2)^2 + \ldots (dx^n)^2.$$

Recall that the form of the euclidean metric is invariant relative to translations

$$\mathbf{T}_{x_0}(x) = x + x_0, \qquad x_0 \in \mathbb{R}^n$$

and rotations,

$$\mathbf{O}(x) = O_{ij}x^j, \qquad O \cdot O^t = I.$$

Thus $\mathbf{T}$ and $\mathbf{O}$ are isometries of the euclidean metric. In addition to these the Euclidean space admits as conformal isometries the dilations $\mathbf{S}_\lambda x = \lambda x$ and the inversion $\mathbf{R}x = |x|^{-2}x$.

Exercise: For any function $u : \mathbb{R}^n \to \mathbb{R}$ let $\mathbf{S}_\lambda^* u(x) = u(\mathbf{S}_\lambda x)$ and $\mathbf{R}^* u(x) = |x|^{2-n}u(\mathbf{R}\, x)$. Check that,

$$\begin{aligned} \Delta(\mathbf{S}_\lambda^* u) &= \lambda^2 \mathbf{S}_\lambda^*(\Delta u), \quad \forall x \in \mathbb{R}^n \\ |x|^{2-n}\Delta(\mathbf{R}^* u)(x) &= \mathbf{R}^*(\Delta u)(x), \quad \forall x \in \mathbb{R}^n \setminus 0. \end{aligned}$$

In particular, if u is harmonic, i.e. $\Delta u = 0$, so are $\mathbf{S}_\lambda^* u$ and $\mathbf{R}^* u$. Recall from Ch. 1 that the fundamental solution of Δ is given by,

$$\begin{aligned} K_n(x) &= \big((2-n)\omega_n\big)^{-1}|x|^{2-n}, \quad \text{for} \quad n \geq 3 \\ K_2(x) &= (2\pi)^{-1}\log|x|. \end{aligned}$$

We gather together the elementary properties of harmonic functions in the following:

Theorem 1.1. *Let $D \subset \mathbb{R}^n$ be a bounded, connected open set.*

i.) *Mean Value Property. Let $u \in \mathcal{C}^2(D)$. If u is harmonic then, for each ball $B(x, R) \subset D$ with boundary $S(x, r)$,*

$$u(x) = |S(x,r)|^{-1}\int_{S(x,r)} u(y)dA(y) \tag{256}$$

$$= |B(x,r)|^{-1}\int_{B(x,r)} u(y)dy. \tag{257}$$

Conversely, if (256) *is verified, for all $B(x,R)\subset D$, then u is harmonic.*

ii.) *Strong Maximum Principle. If $u\in\mathcal{C}^2(D)\cap\mathcal{C}^0(\overline{D})$, is harmonic in D then,*

$$\max_{\bar D} u=\max_{\partial D} u.$$

Moreover if the maximum is reached at some interior point $x_0\in D$ then u is constant in D. A similar statement holds for the minimum of u.

iii.) *Uniqueness of Dirichlet Problem. The Dirichlet problem in D,*

$$\Delta u=f,\qquad u|_{\partial D}=g,$$

with $f\in\mathcal{C}(D)$ and $g\in\mathcal{C}(\partial D)$ has a unique solution $u\in\mathcal{C}^2(D)\cap\mathcal{C}^0(\bar D)$.

iv.) *Local regularity estimate. If u is harmonic in D and $B=B(x_0,r)\subset D$,*

$$|\partial^\alpha u(x_0)|\lesssim r^{-n-|\alpha|}\|u\|_{L^1(B)}. \tag{258}$$

As a consequence we deduce that any harmonic function in $u\in\mathcal{C}^2(D)$ must in fact be smooth, $u\in\mathcal{C}^\infty(D)$. By keeping track of the precise constants in (258) *one can in fact show that in fact u is real analytic in D. Another consequence of* (258) *is Liouville's theorem, according to which any bounded harmonic function $u:\mathbb{R}^n\to\mathbb{R}$ must be constant.*

v.) *Harnack inequality. If $K\subset D$ is compact, then there exists a constant C depending on K such that, for all non-negative harmonic functions u in D*

$$\sup_K u\le C\inf_K u.$$

Proof : To prove i.) let

$$\begin{aligned}
\phi_x(r) &= |S(x,r)|^{-1}\int_{S(x,r)} u(y)dS_y=|S(0,1)|^{-1}\int_{S(0,1)} u(x+rz)dA_z\\
\frac{d}{dr}\phi_x(r) &= |S(0,1)|^{-1}\int_{S(0,1)}\partial u(x+rz)\cdot z dA_z
\end{aligned}$$

On the other hand, by Green's formula,

$$\begin{aligned}
\int_{B(x,r)}\Delta u(y)dy &= \int_{S(x,r)}\partial u(y)\cdot\frac{y-x}{r}dS_y=r^{n-1}\int_{S(0,1)}\partial u(x+rz)\cdot z dA_z\\
&= |S(0,1)|r^{n-1}\frac{d}{dr}\phi_x(r)
\end{aligned}$$

Hence,

$$\frac{d}{dr}\phi_x(r)=|S(x,r)|^{-1}\int_{B(x,r)}\Delta u(y)dy=0. \tag{259}$$

So $\phi_x(r)$ is a constant, so $\phi_x(r)=\lim_{t\to0}\phi_x(t)=u(x)$. For the other statement,

$$\int_{B(x,r)} u(y)dy = \int_0^r\Big(\int_{S(x,s)} u(y)dS_y\Big)ds=u(x)\,\omega_n\int_0^r s^{n-1}ds=|B(x,r)|u(x)$$

as desired.

To prove ii.) assume that $u(x_0) = \sup_{\bar{D}} u$ for some $x_0 \in D$. Then, for any $0 < r < d(x_0, \partial D)$, the mean value property implies,

$$M = u(x_0) = |B(x,r)|^{-1} \int_{B(x,r)} u(y)dy \le M$$

with equality holding only if $u \equiv M$ in $B(x,R)$. From this, we see that $\{x : u(x) = M\}$ is both open and closed in D, and therefore equal to D. This proves the strong "moreover" statement, from which the weaker first statement follows. For the analogous result for minima, replace u by $-u$.

Statement iii.) is an immediate consequence of and ii.) and the linearity of the Dirichlet problem.

To prove iv.) we proceed by induction with respect to $|\alpha|$. The case $|\alpha| = 0$ follows easily from (257). Indeed, for every $B(y,r') \subset D$,

$$|u(y)| \le |B(y,r')|^{-1} \|u\|_{L^1(B(y,r'))} = \omega_n^{-1}(r')^{-n} \|u\|_{L^1(B(y,r'))} \tag{260}$$

To understand how the induction works it suffices to understand the case $|\alpha| = 1$. Note that if u is harmonic, then so is $\partial_i u$. Apply (257) to $\partial_i u$ and any $r > 0$ for which $B(x_0, r) \subset D$,

$$\partial_i u(x_0) = \frac{1}{|B(x_0, r/2)|} \int_{B(x_0,r/2)} \partial_i u(y)dy = \frac{1}{|B(x_0,r/2)|} \int_{S(x_0,r/2)} n_i(y)u(y)dS_y$$

with $n_i(y)$ the exterior unit normal to $y \in S(x_0, r/2)$. Hence,

$$\begin{aligned} |\partial_i u(x_0)| &\le n\omega_n^{-1}(2/r)^n w_n (r/2)^{n-1} \|u\|_{L^\infty(S(x_0,r/2))} \\ &\le \frac{2n}{r} \|u\|_{L^\infty(S(x_0,r/2))} \end{aligned}$$

Now, since for any $y \in S(x_0, r/2)$ we have $B(y, r/2) \subset B(x,r) \subset D$, we make use of estimate (260) with $r' = r/2$ to infer that,

$$|\partial_i u(x_0)| \le \frac{2n}{r} \omega_n^{-1} (r/2)^{-n} \|u\|_{L^1(B(x_0,r))} = cr^{-n-1} \|u\|_{L^1(B(x_0,r))}$$

with the constant $c = \frac{n2^{n+1}}{\omega_n}$. The general case can be done by induction in the same way. The fact that u is smooth then follows easily. The analyticity of u can be shown by simply writing down the Taylor series and noting its convergence using these bounds and the exact constants (see Evans [**1**] section 2.2 for details). Liouville's theorem follows by letting $r \to \infty$ with $|\alpha| = 1$.

It remains to prove v), the Harnack inequality. Let r denote $\frac{1}{3} d(K, \partial D)$. Let $x, y \in K$ with $|x - y| \le r$. According to (257), since u is non-negative and $B(x, 2r) \subset D$,

$$\begin{aligned} u(x) &= |B(x,2r)|^{-1} \int_{B(x,2r)} u(z)dz \ge \frac{|B(y,r)|}{|B(x,2r)|} \big(|B(y,r)|^{-1} \int_{B(y,r)} u(z)dz \big) \\ &= 2^{-n} u(y) \end{aligned}$$

Hence, for all $x, y \in K$ with $|x - y| \le r$, we must have $2^n u(x) \ge u(y)$. Since K is compact we can cover it by a chain of finitely many balls $B_1, \dots, B_N$ of radius r such that $B_i \cap B_{i+1} \ne \varnothing$. Thus, recursively,

$$u(x) \ge 2^{-nN} u(y), \quad \forall x, y \in K.$$

■

Exercise. (Evans, [**1**], Ch. 2). We say $v \in \mathcal{C}^2(D)$ is *subharmonic* if $-\Delta v \leq 0$ in D.

(a) Prove for subharmonic v that

$$v(x) \leq |B(x,r)|^{-1} \int_{B(x,r)} v(y) dy$$

(b) Prove that therefore $\max_{\bar{D}} v = \max_{\partial D} v$

(c) Let $\phi{:}\mathbb{R} \to \mathbb{R}$ be smooth and convex (i.e. $\phi(tx+(1-t)y) \leq tf(x)+(1-t)f(y)$, $\forall x, y \in \mathbb{R}, t \in [0,1]$). Assume u is harmonic and $v = \phi(u)$. Prove v is harmonic.

(d) Prove $v := |\partial u|^2$ is subharmonic, wheneven u is harmonic.

1.2. Representation formulae. The formula $u = K_n * f$ with K_n the fundamental solution of the Laplacian, allows us to solve the Poisson's equation $\Delta u = f$ in the whole space $\mathbb{R}^n$. Can we get similar formulas for other domains $D \in \mathbb{R}^n$? We first check the following integration by parts formula, called Green's identity

$$\int_D \big(v\Delta u - \Delta v u\big) dx = \int_{\partial D} \big(v \frac{du}{dn} - u \frac{dv}{dn}\big) \tag{261}$$

where $\frac{du}{dn}$ denotes the derivative with respect to the exterior unit normal n to ∂D. We apply the idenitity to $K(y) = K_n(y - x_0)$ and make use of the fact that $\Delta_y K_n(x_0 - y) = \delta_{x_0}$ to derive[1],

$$u(x_0) = \int_D K(y - x_0) \Delta u(y) dy - \int_{\partial D} \big(K(y - x_0) \frac{du}{dn_y}(y) - u(y) \partial_{n_y} K(y - x_0)\big) dS_y \tag{262}$$

for any $x_0 \in D$ and any function $u \in \mathcal{C}^2(\bar{D})$.

Assume that $\Delta u = f$ and that the boundary values of u on ∂D are given. We need to eliminate the term on the right hand side of (**??**) which contains the normal derivative of u; without that term, (262) would allow us to solve for u. We can do that by introducing, as correction, a harmonic function $\psi_{x_0}(y)$ which such that the *Green's function for* D,

$$G(x_0, y) = K_n(y - x_0) + \psi_{x_0}(y) \tag{263}$$

verifies

$$\Delta_y G(x_0, y) = \delta_{x_0}, \qquad G(x_0, y) = 0 \quad \text{on} \quad \partial D \tag{264}$$

Thus, using formula (262) with $K_n(y - x_0)$ replaced by $G(x_0, y)$ we infer that,

$$u(x_0) = \int_D G(x_0, y) \Delta u(y) dy + \int_{\partial D} \frac{d}{dn_y} G(x_0, y) u(y) dS_y u(y)\big) \tag{265}$$

[1] To prove it we need to show that the singularity of $K(y - x_0)$ at $y = x_0$ does not create problems. One does that by replacing D with $D \setminus B(x_0, \epsilon)$ and then let $\epsilon \to 0$. See Evans, [**1**], section 2.2 for details.

Recall that $\frac{d}{dn_y}G(x_0, y)$ is the derivative in the direction to the exterior normal n_y at a point $p \in \partial D$. In practice it is not at all easy to find such corrections. There are however two important examples when this can be done by symmetry consideratins.

1.) *Dirichlet problem for a half space.* Let ,

$$\mathbb{R}^n_+ = \{x = (x^1, x^2, \dots, x^n)/\, x_n > 0\}$$

Let $x \in \mathbb{R}^n_+$ and consider its reflection $\overline{x}$ relative to the hyperspace $x_n = 0$. It is then easy to show that $G(x, y) = K(y - x) - K(y - \bar{x})$ is a Green function for $\mathbb{R}^n_+$. Thus, since the exterior normal derivative at $x_n = 0$ is given by ∂_n we easily find the *Poisson's Kernel for* $\mathbb{R}^n_+$

$$P_+(x, y) = \partial_n G(x, y) = \frac{2x_n}{\omega_n}|x - y|^{-n} \tag{266}$$

Thus,

$$u(x) = \int_{x_n=0} P_+(x, y)g(y)dy \tag{267}$$

is a solution to the Dirichlet problem in $\mathbb{R}^n_+$:

$$\Delta u = 0, \qquad u = g \quad \text{on} \quad x_n = 0. \tag{268}$$

PROPOSITION 1.3. *Assume* $g \in \mathcal{C}^0(\mathbb{R}^{n-1}) \cap L^\infty(\mathbb{R}^{n-1})$. *The the function* u *defined by* (267) *is a bounded harmonic function in* $\mathbb{R}^n_+$ *and verifies*

$$u(x) \to g(x_0) \quad as \quad x \to x_0 \quad \forall x_0 \in \partial\mathbb{R}^n_+.$$

Exercise: Prove proposition (1.3) by observing that $P_+(x, y)$ is a positive harmonic function in y, for all $x \in \mathbb{R}^n_+$ and $y \in \partial\mathbb{R}^n_+$. Moreover, for all $x \in \mathbb{R}^n_+$, we have $\int_{\partial\mathbb{R}^n_+} P_+(x, y)dy = 1$.

Exercise. Rederive formula (267) using the Fourier transform.

2.) *Dirichlet problem for a ball.* Let $D = B(0, a)$, the ball centered at 0 of radius a. Let x_0 be an arbitrary point of D. Let $x_0^* = a^2\frac{x_0}{|x_0|^2}$ be the inverse of x_0 relative to the sphere $|x| = a$. Observe that for any x on the boundary of D we have, $\frac{|x - x_0^*|}{|x - x_0|} = \frac{a}{|x_0|}$. Thus,

$$G(x_0, x) = K(x_0 - x) - \big(\frac{a}{|x_0|}\big)^{2-n}K(x_0^* - x) \tag{269}$$

vanishes for $x \in \partial D$. Moreover the *correction* $\big(\frac{a}{|x_0|}\big)^{2-n}K(x_0^* - x)$ is clearly harmonic in the domain $D = B(0, a)$. After a simple computation we infer from (265) that,

$$u(x) = \int_{|y|=a} H(x, y)g(y)dS_y, \qquad H(x, y) = \frac{1}{a\omega_n}\frac{a^2 - |x|^2}{|y - x|^n} \tag{270}$$

is a solution to the Dirichlet problem,

$$\Delta u = 0 \quad \text{in} \quad B(0, a), \qquad u = g \quad \text{on} \quad S(0, a).$$

PROPOSITION 1.4. *Let g be continuous on $S(0,a)$. Then the function $u(x)$ defined by* (270) *for $|x| < a$, is continuous for $|x| \le a$ and harmonic in $|x| < a$.*

Exercise. Prove the above proposition by taking advantage of the fact that H is a positive harmonic function in $|x| < a$ for all $y \in S(0,a)$. We also have, $\int_{|y|=a} H(x,y)dS_y = 1$.

1.5. A-priori estimates for Δ in $\mathbb{R}^n$.

First recall the L^2 identity,

$$\sum_{i,j=1}^n \|\partial_i\partial_j u\|_{L^2}^2 = \|\Delta u\|_{L^2}^2, \tag{271}$$

for any $u \in \mathcal{C}_0^\infty(\mathbb{R}^n)$. According to the Calderon-Zygmund theory (see the discussion at the beginning of chapter 1, section 9) we also have, for any $1 < p < \infty$,

$$\|\partial_i\partial_j u\|_{L^p} \lesssim \|\Delta u\|_{L^p}. \tag{272}$$

The cases $p = 1$ and ∞ are exceptional. It turns out, in particular, that the estimate (272) is false for $p = \infty$. This is due to a logarithmic loss of derivatives in the estimate and can be circumvented in various ways. The simplest[2], introduced by Schauder, is based on the Hölder norms with fractional exponents $0 < \gamma < 1$

$$[f]_{\mathcal{C}^{0,\gamma}} = \sup_{x\neq y} \frac{f(x) - f(y)}{|x-y|^\gamma},$$

see chapter 1, section 5.15. Using these norms one finds the Schauder estimate,

$$[\partial^2 u]_{\mathcal{C}^{0,\gamma}} \le c_\alpha [\Delta u]_{\mathcal{C}^{0,\gamma}}. \tag{273}$$

The proof of (273) can be derived from the identity,

$$\partial_i\partial_j u(x) = \int_{\mathbb{R}^n} \partial_i\partial_j K(x-y)\big(f(y) - f(x)\big)dy \tag{274}$$

where $f = \Delta u$.

Exercise. Prove formula (274) and the Schauder estimate (273).

We can also derive first derivative estimates applying the Hardy-Littlewood-Sobolev inequalities of Theorem 5.9 to the representation,

$$\partial u = \int_{\mathbb{R}^n} \partial K_n(x-y)\Delta u(y)dy,$$

Thus, since $|\partial K_n(x-y)| \lesssim |x-y|^{1-n}$, we derive for $1 < p < q < \infty$,

$$\|\partial u\|_{L^q} \lesssim \|\Delta u\|_{L^p}, \qquad 1/q = 1/p - 1/n. \tag{275}$$

---

[2] Other refinements, which also work for L^1, are based on more complicated spaces such BMO, Hardy or Besov spaces.

1.6. Dirichlet problem for general domains. The methods developed in the treatment of the Dirichlet problem in a given domain D have had a huge impact throughout the field of partial differential equations. There are four major approaches to the Dirichlet problem in a given domain. These are known under the following names:
A. Variational method (Dirichlet Principle),
B. Perron's method(subsolutions and supersolutions)
C. Method of continuity
D. Potential theory

1.7. Energy methods and Dirichlet Principle. Consider the Dirichlet boundary value problem,

$$-\Delta u = f \quad \text{in} \quad D, \qquad u|_{\partial D} = g. \tag{276}$$

We have already proved uniqueness with the help of the maximum principle. In what follows we give an alternative prove of uniqueness based on integration by parts, or *energy method*. Consider two $\mathcal{C}^2(\bar{D})$ solutions u_1, u_2 and set $v = u_1 - u_2$. Then clearly,

$$\Delta v = 0, \quad v|\partial D = 0.$$

Therefore, by integration by parts,

$$0 = \int_D v\Delta v = \int_D |\partial v|^2$$

Thus v must be constant in $\bar{D}$ and zero on the boundary; that is $v = 0$.

The energy metod can also be used to construct solutions to (276). This is based on the idea that solutions of (276) are minimizer of a functional. To see this we define the Dirichlet Integral,

$$I[w] = \int_D (\frac{1}{2}|\partial w|^2 - wf)dx. \tag{277}$$

with w belonging to the set of admissible functions,

$$\mathcal{A} = \{w \in \mathcal{C}^2(\bar{D}) : w|_{\partial D} = g\}.$$

THEOREM 1.8. *A function $u \in \mathcal{A}$ is a solution of the Dirichlet problem* (276) *if and only if u minimizes the Dirichlet integral among all functions in $\mathcal{A}$,*

$$I[u] = \min_{w \in \mathcal{A}} I[w] \tag{278}$$

Proof : Assume that u is a solution of (276) and $w \in \mathcal{A}$. Since $(u - w)|_{\partial D} = 0$ we derive by integration by parts,

$$\begin{aligned} 0 &= -\int_D (\Delta u - f)(u - w) = \int_D \big(\partial u \cdot \partial(u - w) - f(u - w)\big)dx \\ &= \int_D (|\partial u|^2 - uf)dx - \int_D \partial u \cdot \partial w dx + \int_D wf dx \end{aligned}$$

Hence, using the inequality $|\partial u \cdot \partial w| \le \frac{1}{2}|\partial u|^2 + \frac{1}{2}|\partial w|^2$,

$$\begin{aligned}\int_D (|\partial u|^2 - uf)dx &= \int_D \partial u \cdot \partial w dx - \int_D wf dx \\ &\le \frac{1}{2}\int_D |\partial u|^2 + \frac{1}{2}\int_D |\partial w|^2 - \int_D wf dx\end{aligned}$$

Thus, $I[u] \le I[W]$ as desired.

Conversely assume that (278) holds and consider the function $J(\epsilon) = I[u + \epsilon w]$. Since $J(0)$ is a minimum value for J we must have $J'(0) = 0$. By a simple integration by parts we derive $0 = J'(0) = \int_D(-\Delta u - f)w dx$. Since this is true for all $w \in \mathcal{C}_0^\infty(D)$ we infer that $-\Delta u = f$ in D. ■

It turns out however that the functional $I[w]$ cannot be easily minimized in the class $\mathcal{A}$ of admissible functions. The avoidance of this difficulty has led to some of the most exciting developments in PDE last century. Here are the main ideas.

Step 1. It is easy to see that the general solution of (276) can be reduced to the case $g = 0$.

Step 2. Instead of the admissible set $\mathcal{A}$, with $g = 0$, we consider the Sobolev space $H_0^1(D)$. Consider also the bilinear form,

$$(u, v)_{H_0^1(D)} =< u, v >= \int_D \partial u \cdot \partial v dx. \tag{279}$$

Observe that $H_0^1(D)$ is a Hilbert space relative to the scalar product $< u, v >= (u, v)_{H_0^1(D)}$. Clearly, if u is a $\mathcal{C}^2(\bar{D})$ solution of (276) then, for every $v \in H_0^1(D)$,

$$< u, v >= (f, v).$$

with $(f, v) = \int f(x)v(x)dx$ denoting the standard inner product in $L^2(D)$.

Definition. We say that $u \in H_0^1(D)$ is a weak solution of (276) if,

$$< u, v >= (f, v) \tag{280}$$

for all $v \in H_0^1(D)$.

Step 3. To find a weak solution of the Dirichlet problem we only need to use a little bit of Hilbert space theory. The idea is to consider the linear functional $F[v] = (f, v) = \int_D f(x)v(x)dx$ defined on the Hilbert space $H_0^1(D)$. According to the Riesz representation theorem in Hilbert spaces to find a weak solution of our Dirichlet problem it suffices to show that our linear functional $F[v]$ is bounded on $H_0^1(D)$. This reduces to a simple functional inequality,

$$\|v\|_{L^2(D)} \lesssim \|v\|_{H_0^1(D)}, \tag{281}$$

called the Poincaré inequality.

THEOREM 1.9 (Poincaré inequality). *Let D be a bounded open set in $\mathbb{R}^n$ and $u \in W_0^{1,p}(D)$, $1 \le p < n$. Then we have the estimate,*

$$\|u\|_{L^q(D)} \lesssim \|\partial u\|_{L^p(D)}. \tag{282}$$

for each $q \in [1, p^*]$ *with* $p^* = \frac{np}{n-p}$.

Proof : By definition, there exists a sequence $u_k \in \mathcal{C}_0^\infty(D)$ which converges to u in $W^{1,p}(D)$. We extend each function u_m to be zero on $\mathbb{R}^n \setminus \bar{D}$. According to the sharp Gagliardo-Nirenberg-Sobolev inequality of Theorem 5.11 we have,

$$\|u_m\|_{L^{p^*}(D)} \lesssim \|u_m\|_{L^{p^*}(\mathbb{R}^n)} \lesssim \|\partial u_m\|_{L^p(D)}$$

Taking $m \to \infty$, we infer that

$$\|u\|_{L^{p^*}(D)} \lesssim \|\partial u\|_{L^p(D)}$$

Since $|D| < \infty$, this holds by the Hölder inequality for $q \in [1, p^*]$.

■

One can prove this inequality for functions $v \in \mathcal{C}_0^\infty(D)$. Thus, $F[v]$ is a bounded linear functional on ${\mathcal{C}_0}^\infty(D)$ and therefore can be extended by density to the Hilbert space $H_0^1(D)$.

Step 4. We now have a weak solution $u \in H_0^1(D)$ of our Dirichlet problem. Clearly u is a distribution in D, $u \in \mathcal{D}'(D)$, and we have

$$-\Delta u = f,$$

in the sense of distributions. We expect to be able to show that u is in fact better than $H_0^1(D)$. In fact, recalling the regularity results of the previous paragraph, we expect that if $f \in L^2(D)$ then $u \in H^2_{\text{loc}}(D)$.

THEOREM 1.10 (Interior regularity). *Assume that* $f \in L^2(D)$ *and that* $u \in H^1(D)$ *is a weak solution of* $-\Delta u = f$ *in* D, *i.e.* $< u, v >= (f, v)$ *for all* $v \in H_0^1(D)$. *Then* $u \in H^2_{loc}(D)$ *and, for every* $V \subset D$,

$$\|u\|_{H^2(V)} \lesssim \|f\|_{L^2(D)} + \|u\|_{L^2(D)} \tag{283}$$

Proof : Choose open sets $V \subset\subset W \subset D$ and a test function $0 \le \zeta \le 1$ equal to one on V and zero on $\mathbb{R}^n \setminus W$. Since u is a weak solution we have

$$\int_D \partial_i u \partial_i v dx = \int_D f v dx \tag{284}$$

We introduce the difference quotients,

$$\partial_k^{(h)} u = \frac{u(x + he_k) - u(x)}{h}, \qquad h \neq 0.$$

Observe that for all $w \in L^2(D)$, supported in $\bar{W}$ we have,

$$\int_D v(x) \partial_k^{(-h)} w(x) \quad = \quad - \int_D \partial_k^{(h)} v(x) w(x)$$

for all sufficiently small $h \neq 0$.

Now set $v = -\partial_k^{(-h)}(\zeta^2 \partial_k^{(h)} u)$ in (283). Thus,

$$\begin{aligned}
\int_D \partial_i u \partial_i v dx &= -\int_D \partial_i u \partial_i \big(\partial_k^{(-h)}(\zeta^2 \partial_k^{(h)} u)\big)\\
&= \int_D \partial_i \partial_k^{(h)} u \partial_i \big((\zeta^2 \partial_k^{(h)} u)\big)\\
&= 2\int_D \zeta \partial_i \zeta\, \partial_k^{(h)} \partial_i u\, \partial_k^{(h)} u + \int_D \zeta^2\, \partial_k^{(h)} \partial_i u \partial_k^{(h)} \partial_i u = I_1 + I_2\\
I_2 &= \int_D \zeta^2 |\partial_k^{(h)} \partial u|^2\\
I_2 &\le C\int_D \zeta |\partial_k^{(h)} \partial u|\, |\partial_k^{(h)} u|\\
&\le c\epsilon \int_D \zeta^2 |\partial_k^{(h)} \partial u|^2 c\epsilon^{-1} \int_D |\partial_k^{(h)} u|^2 \le \int_D \zeta^2 |\partial_k^{(h)} \partial u|^2 + c\epsilon^{-1}\int_D |u|^2
\end{aligned}$$

Therefore, chosing ϵ such that $C\epsilon = \frac{1}{2}$,

$$\int_D \partial_i u \partial_i v dx \ge \frac{1}{2}\int_D \zeta^2 |\partial_k^{(h)} \partial u|^2 - C\int_D |\partial u|^2$$

Thus, in view of (284), and our choice of ζ, we deduce

$$\int_V |\partial_k^{(h)} \partial u|^2 \le \int_D \zeta^2 |\partial_k^{(h)} \partial u|^2 \lesssim \int_D |\partial u|^2 + \int_D |f|^2 \tag{285}$$

for all $k = 1, \dots n$ and all sufficiently small $h \neq 0$. Using (285) it is easy to conclude that $\partial u \in H^1(V)$ and therefore $u \in H^2(V)$ as desired. Moreover,

$$\|u\|_{H^2(V)} \lesssim \|f\|_{L^2(D)} + \|u\|_{H^1(D)} \tag{286}$$

To end the proof of theorem 1.10 we only need to replace $\|u\|_{H^1(D)}$ in (286) by $\|u\|_{L(D)}$. We first remark that we can replace the right hand side in (286) with $\|f\|_{L^2(W)} + \|u\|_{H^1(W)}$. To eliminate $\|u\|_{H^1(W)}$ we choose a new cut-off $0 \le \zeta \le 1$ supported in D and equal to 1 on W. Setting $v = \zeta^2 u$ in (284), $\int_D \partial_i u \partial_i v dx = \int_D f v dx$, we easily check that

$$\int_D |\zeta|^2 |\partial u|^2 dx \lesssim \|f\|^2_{L^2(D)} + \|u\|^2_{L^2(D)}$$

Hence,

$$\|u\|_{H^1(W)} \lesssim \|f\|_{L^2(D)} + \|u\|_{L^2(D)},$$

as desired. ■

Step 5. Having proved that $f \in L^2(D)$ implies $u \in H^2_{\mathrm{loc}}(D)$ we would like to show that if f is more regular so is u.

THEOREM 1.11 (Higher interior regularity). *Assume that u is a weak solution of $-\Delta u = f$ in D and $f \in H^m(D)$. Then $u \in H^{m+2}_{loc}(D)$ and we have the estimate,*

$$\|u\|_{H^{2+m}(D)} \lesssim \|f\|_{H^m(D)} + \|u\|_{L^2(D)} \tag{287}$$

Proof : Consider again (284) and take $v = (-1)^{|\alpha|}\partial^\alpha \tilde{v}$ with $\tilde{v} \in \mathcal{C}_0^\infty(W)$ and $|\alpha| = m$. As before $V \subset\subset W \subset D$. Clearly, integrating by parts,

$$< \tilde{u}, \tilde{v} >= (\tilde{f}, \tilde{v}$$

where $\tilde{u} = (-1)^{|\alpha|}\partial^\alpha u$, $\tilde{f} = (-1)^{|\alpha|}\partial^\alpha f$. According to theorem 1.10, $\tilde{u} \in H^2(V)$ and,

$$\|\partial^\alpha u\|_{H^2(V)} \lesssim \|\tilde{f}\|_{L^2(W)} + \|\tilde{u}\|_{L^2(W)} \lesssim \|f\|_{H^{m+1}(D)} + \|\mathbf{u}\|_{H^m(D)}.$$

Hence,

$$\|u\|_{H^{m+2}(V)} \lesssim \|f\|_{H^m(D)} + \|u\|_{H^m(D)},$$

and the proof of the theorem proceeds now by induction on m. ■

Step 6. So far we have established interior regularity but have no informations about the behavior of u on the boundary of D. In particular we cannot yet show that $u|_{\partial D} = 0$ in the traditional sense. Clearly, to achieve this, we need more regualrity information about the boundary of D.

THEOREM 1.12 (Boundary regularity). *Assume that $u \in H_0^1(D)$ is a weak solution of $-\Delta u = f, u|_{\partial D} = 0$ with $f \in L^2(D)$. Assume also that ∂D is C^2 regular. Then $u \in H^2(D)$ and*

$$\|u\|_{H^2(D)} \lesssim \|f\|_{L^2(D)} + \|u\|_{L^2(D)} \tag{288}$$

Moreover if $f \in H^m(D)$ and ∂D is C^{m+2} then $u \in H^{m+2}(D)$ and,

$$\|u\|_{H^{2+m}(D)} \lesssim \|f\|_{H^m(D)} + \|u\|_{L^2(D)} \tag{289}$$

Proof : We only sketch the proof for the particular case when D is a half ball $U = B(0,1) \cap \mathbb{R}^n_+$, where $\mathbb{R}^n_+ = \{(x_1, \ldots, x_n)/x_n \geq 0\}$. Proceeding exactly in a similar manner as for the interior estimates of theorem 1.10 we can first derive estimates for the tangential derivatives finite difference derivatives of u i.e.,

$$\int_V |\partial_k^{(h)} \partial u|^2 dx \lesssim \int_D |f|^2 dx + \int_D |\partial u|^2 dx \tag{290}$$

where $V = B(0, 1/3) \cap \mathbb{R}^n_+$ and $k = 1, 2, \ldots n-1$. This can be achieved with the help of the smooth cutoff function $0 \leq \zeta \leq 1$, $\zeta = 1$ on $B(0, 1/3)$ and $\zeta = 0$ on $\mathbb{R}^n \setminus B(0, 2/3)$ and choosing $v = -\partial_k^{(-h)}(\zeta^2 \partial_k^{(h)} u)$ in the identity $< u, v >= (f, v)$. One can easily infer from (291) that,

$$\sum_{i,j=1}^{n-1} \|\partial_i \partial_j u\|_{L^2(V)} + \sum_{i=1}^{n-1} \|\partial_i \partial_n u\|_{L^2(V)} \lesssim \|f\|_{L^2(D)} + \|u\|_{H^2(D)}.$$

To derive the remaining estimate for $\partial_n^2 u$ we only have to observe that, since $-\delta u = f$ we have,

$$\|\partial_n^2 u\|_{L^2(V)} \lesssim \| \sum_{i,j=1}^{n-1} \|\partial_i \partial_j u\|_{L^2(V)} \lesssim \|f\|_{L^2(D)} + \|u\|_{H^2(D)}.$$

Hence, in fact,

$$\|u\|_{H^2(V)} \lesssim \|f\|_{L^2(D)} + \|u\|_{H^1(D)} \tag{291}$$

We can then proceed, as we did for the interior estimates, to eliminate $\|u\|_{H^1(D)}$ in favor of $\|u\|_{L^2(D)}$.

The higher derivatives estimate (289) can be proved in a similar manner, see proof of theorem 1.11. ■

2. Dirichlet problem on compact Riemannian manifolds

Let $\mathbf{M}$ be a compact Riemannian manifold and consider the problem,

$$-\Delta_{\mathbf{M}} u = f \tag{292}$$

Let $\mathcal{C}^\infty(\mathbf{M})$ denote the space of smooth functions on $\mathbf{M}$. For two such functions u, v, we have,

$$-\int_{\mathbf{M}} \Delta u v dv_{\mathbf{M}} = \int_{\mathbf{M}} D^i u D_i v \, dv_{\mathbf{M}} :=< u, v > \tag{293}$$

Observe that $< u, u >= 0$ if and only if u is a constant. We say that two continuous functions are equivalent if they differ by a constant. We consider the space of classes of equivalence of $\mathcal{C}^\infty(\mathbf{M})$ functions on $\mathbf{M}$ modulo constants. Let $\dot{H}^1(\mathbf{M})$ be the completion of this space relative to the scalar product $< u, v >$. We also introduce the Sobolev space $H^1(\mathbf{M})$ which is defined as the completion of $\mathcal{C}^\infty(\mathbf{M})$ relative to the norm

$$\|u\|^2_{H^1(\mathbf{M})} = (u, u) + < u, u > .$$

Definition. We say that $u \in \dot{H}^1(\mathbf{M})$ is a weak solution of (292), for $f \in L^2(\mathbf{M})$, if, for all $v \in \dot{H}^1(\mathbf{M})$,

$$< u, v >= (f, v),$$

with $(f, v) = (f, v)_{\mathbf{M}} = \int_{\mathbf{M}} f v dv_{\mathbf{M}}$.

Clearly weak solutions must be unique. Indeed if u_1, u_2 are two solutions and $u = u_1 - u_2$ then $< u, v >= 0$ for all $v \in \dot{H}^1(\mathbf{M})$, hence $< u, u >= 0$ and thus u_1 and u_2 are equivalent.

To prove existence we have to show that the linear functional $v \to (f, v)$ is continuous on the Hilbert space $\dot{H}^1(\mathbf{M})$. Since

$$|(f, v)| \lesssim \|f\|_{L^2(\mathbf{M})} \|v\|_{L^2(\mathbf{M})},$$

we need to check an inequality of the form,

$$\|v - \bar{v}\|_{L^2(\mathbf{M})} \lesssim \|Dv\|_{L^2(\mathbf{M})} \tag{294}$$

where $\bar{v}$ the average of v defined by,

$$\bar{v} = \frac{1}{|\mathbf{M}|} \int_{\mathbf{M}} v \, dv_{\mathbf{M}},$$

and $|M|$ the volume of M. The proof of this version of the Poincaré inequality is based on the Rellich compactness theorem.

THEOREM 2.1 (Rellich compactness). *The embedding of $H^1(\mathbf{M}) \subset L^2(\mathbf{M})$ is compact operator, i.e every bounded sequence in $H^1(\mathbf{M})$ has an accumulation point in $L^2(\mathbf{M})$.*

We use Rellich's theorem to prove estimate (294).

Proof of (294). In view of the definition of $\dot{H}^1(\mathbf{M})$ it suffices to prove (294) for functions $v \in \mathcal{C}^\infty(\mathbf{M})$. By contradiction assume that (294) is false. Thus there exists functions $v_k \in \mathcal{C}^\infty(\mathbf{M})$ veifying

$$\|v_k - \bar{v_k}\|_{L^2(\mathbf{M})} > k\|Dv_k\|_{L^2(\mathbf{M})}$$

Let

$$w_k = \frac{v_k - \bar{v_k}}{\|v_k - \bar{v_k}\|_{L^2(\mathbf{M})}}.$$

Clearly $\bar{w_k} = 0$ and $\|w_k\|_{L^2(\mathbf{M})} = 1$. Moreover,

$$\|Dw_k\|_{L^2(\mathbf{M})} < 1/k \tag{295}$$

Thus w_k is a vounded sequence of functions in $H^1(\mathbf{M})$. By the Rellich theorem there exists a subsequence $w_j = w_{k_j}$ which converges to a function w in $L^2(\mathbf{M})$. Clearly, $\bar{w} = 0$ and $\|w\|_{L^2(\mathbf{M})} = 1$. On the other hand, according to (295), for any $\phi \in \mathcal{C}^\infty(\mathbf{M})$ and any smooth one form A, using the integrations by formula,

$$\| \int_{\mathbf{M}} f \text{div } A dv_{\mathbf{M}} = -\int_{\mathbf{M}} Df \cdot A dv_{\mathbf{M}}, \tag{296}$$

$$\int_{\mathbf{M}} w \text{div } A = \lim_{j\to\infty} \int_{\mathbf{M}} w_j \text{div } A = -\lim_{j\to\infty} \int_{\mathbf{M}} Dw_j \cdot A$$

On the other hand,

$$| \int_{\mathbf{M}} Dw_j \cdot A \lesssim \|Dw_j\|_{L^2(\mathbf{M})} \|A\|_{L^2(\mathbf{M})} \lesssim 1/k \|A\|_{L^2(\mathbf{M})}$$

Hence,

$$\lim_{j\to\infty} \int_{\mathbf{M}} Dw_j \cdot A = 0$$

Thus, for every smooth one form A,

$$0 = \int_{\mathbf{M}} w \text{div } A$$

and therefore w must be a constant scalar function. Since $\bar{w} = 0$ it follows that $w = 0$ which is in contradiction with $\|w\|_{L^2(\mathbf{M})} = 1$.

Exercise. Prove Rellich's theorem. ■

2.2. Regularity theorey. We start with an a-priori energy estimate on manifolds which is the exact analogue of (271). We shall prove the following,

LEMMA 2.3 (Bochner identity). *The following identity holds for a scalar function $u \in \mathcal{C}^\infty(\mathbf{M})$,*

$$\int_{\mathbf{M}} |D^2 u|^2 + \int_{\mathbf{M}} R^{ij} D_i u D_j u = \int_{\mathbf{M}} |\Delta u|^2 \tag{297}$$

with $R_{ij} = g^{ab} R_{iajb}$ the Ricci curvature of $\mathbf{M}$.

Proof :

$$\begin{aligned} D_a(\Delta u) &= D_a(D_c D_c u) = D_c D_a D_c u + [D_a, D_c] D_c u \\ &= D_c D_c D_a u + R_{cdac} D_d u \\ &= \Delta(D_a u) - R_{ad} D_d u \end{aligned}$$

Thus,

$$\begin{aligned} \int_{\mathbf{M}} |\Delta u|^2 &= -\int_{\mathbf{M}} D_a(\Delta u) \cdot D_a u = -\int_{\mathbf{M}} \Delta D_a u \cdot D_a u + R_{ab} D_a u D_b u \\ &= \int_{\mathbf{M}} |D^2 u|^2 + \int_{\mathbf{M}} R_{ab} D_a u D_b u \end{aligned}$$

as desired. ∎

Remark. If $\mathbf{M}$ is 2-dimensional we have, $R_{ab} = g_{ab} K$ with K the Gauss curvature of $\mathbf{M}$. Thus, in that case,

$$\int_{\mathbf{M}} |D^2 u|^2 + \int_{\mathbf{M}} K |Du|^2 = \int_{\mathbf{M}} |\Delta u|^2.$$

subsectionMaximum Principle for second order elliptic equations We consider a second order elliptic operator in the form,

$$Lu = -a^{ij} \partial_i \partial_j u + b^i \partial_i u + cu \tag{298}$$

where the coefficients a, bc are continuous and verify the ellipticity condition,

$$a^i j(x) \xi_i \xi_j \geq m |\xi|^2 \tag{299}$$

We also assume that a^{ij} are symmetric, i.e. $a^{ij} = a^{ji}$.

THEOREM 2.4 (Weak maximum principle). *Assume D is a open bounded doamin in $\mathbb{R}^n$ and $u \in C^2(D) \cap C^0(\bar{D})$ such that, for $c = 0$, u is a subsolution*

$$Lu(x) \leq 0, \qquad x \in D.$$

Then,

$$\max_{\bar{D}} u = \max_{\partial D} u.$$

Proof Assume first that $LU < 0$ in D and there exists $x_0 \in D$ such that $u(x_0) = \max u_{\bar{D}}$. Since x_0 is a point of maximum we must have $\partial u(x_0) = 0$ and, as a matrix, the hessian $\partial^2 u(x_0)$ is negative definite, i.e. $\partial^2 u(x_0) \leq 0$. Since the matrix $A = (a^{ij})_{i,j=1\ldots n}$ is positive definite it is diagonalizable. Let $O = (O^i_j)_{i,j=1,\ldots j}$ be an orthogonal matrix such that, $OAO^T = D$ with D the diagonal matrix with strictly positive entries $d_1, \dots d_n$. Writing, $y = x_0 + O(x - x_0)$, or in components $y^i = x^i_{(0)} + O^i_j(x^j - x^j_{(0)})$, we derive $\partial_{x_i} u = O^a_i \partial_{y_a} u$, $\partial^2_{x^i x^j} u = O^a_i O^b_j \partial^2_{y^a y^b} u$. Hence at x_0,

$$\begin{aligned} a^{ij} \partial^2_{x^i x^j} u &= a^{ij} O^a_i O^b_j \partial^2_{y^a y^b} u = (O \cdot A \cdot O^T)^{ab} \partial^2_{y^a y^b} u \\ &= D^{ab} \partial^2_{y^a y^b} u = \sum_k d_k \partial^2_{y^k y^k} u \le 0 \end{aligned}$$

since for each k we have, at x_0, $d_k \ge 0$ and $\partial^2_{y^k y^k} \le 0$. Consequently at the point x_0,

$$Lu = -a^{ij} \partial^2_{ij} u + b^i \partial_i u \ge 0,$$

which contradicts our assumption.

To treat the general case let

$$u^\epsilon(x) = u(x) + \epsilon e^{\lambda x^1}$$

where $\epsilon > 0$ and $\lambda > 0$ sufficeintly large. According to the uniform ellipticity condition we have $a^{ii}(x) \ge m > 0$. Now, at all points of U,

$$\begin{aligned} Lu^\epsilon &= Lu + \epsilon L(e^{\lambda x^1} \le \epsilon e^{\lambda x^1} \big(- \lambda^2 a^{11} + \lambda b^1 \big) \\ &\le \epsilon e^{\lambda x^1} \big(- \lambda^2 m + \lambda \|b\|_{L^\infty} \big) < 0 \end{aligned}$$

provided $\lambda > 0$ sufficiently large.

■

3. Minkowski space

3.1. Basic definitions. The n+1 dimensional Minkowski space, which we denote by $\mathbb{R}^{n+1}$, consists of the manifold $\mathbb{R}^{n+1}$ together with a Lorentz metric $\mathbf{m}$ and a distinguished system of coordinates x^α, $\alpha = 0, 1, \dots n$, called inertial, relative to which the metric has the diagonal form $\mathbf{m}_{\alpha\beta} = \mathrm{diag}(-1, 1, \dots, 1)$. Two inertial systems of coordinates are connected to each other by translations or Lorentz transformations. We use standard geometric conventions of lowering and raising indices relative to $\mathbf{m}$, and its inverse $\mathbf{m}^{-1} = \mathbf{m}$, as well as the usual summation convention over repeated indices. The coordinate vectorfields $\frac{\partial}{\partial x^\alpha}$ are denoted by ∂_α, an arbitrary vectorfield is denoted by $X = X^\alpha \partial_\alpha$ with $X^\alpha = X^\alpha(x^0, \dots, x^n)$. Observe that by lowering indices relative to $\mathbf{m}$ we get $X_0 = -X^0$ and $X_i = X^i$ for all $i = 1, \dots, n$. We denote by D the flat covariant derivative of $\mathbb{R}^{n+1}$, that is $D_\alpha \omega_\beta = \partial_\alpha \omega_\beta$ for an arbitrary 1- form $w = \omega_\alpha dx^\alpha$. We also split the spacetime coordinates x^α into the time component $x^0 = t$ and space components $x = x^i, \dots x^n$. Note that $t_0 = -t$ and $x^i = x_i$ for $i = 1, \dots, n$.

A vector X is said to be timelike, null or spacelike according to whether $\mathbf{m}(X, X)$ is < 0, $= 0$ or > 0. Accordingly a smooth curve $x^\alpha(s)$ is said to be timelike, null or spacelike if its tangent vector $\frac{dx^\alpha}{ds}$ is timelike, null or spacelike at every one of its points. A causal curve may be timelike or null. Similarly a hypersurface $u(x^0, \dots x^n) = 0$ is said to be spacelike, null or timelike if its normal $N^\alpha = -\mathbf{m}^{\alpha\beta}\partial_\beta u$ is, respectively, timelike, null or spacelike. The metric induced by $\mathbf{m}$ on a spacelike hypersurface is necessarily positive definite, that is Riemannian. A function $\mathbf{t}(x^0, x^1, \dots, x^n)$ is said to be a time function if its level hypersurfaces $\mathbf{t} = t$ are spacelike. On a null hypersurface the induced metric is degenerate relative to the normal direction, i.e. $\mathbf{m}(N, N) = 0$. In particular function $\mathbf{u} = \mathbf{u}(x^0, \dots x^n)$ whose level surfaces $\mathbf{u} = u$ are null must verify the Eikonal equation

$$\mathbf{m}^{\alpha\beta}\partial_\alpha \mathbf{u}\partial_\beta \mathbf{u} = 0 \tag{300}$$

Equation (300) can also be written in the form $D_N N = 0$. We call N a *null geodesic generator* of the level hypersurfaces of $\mathbf{u}$.

A causal curve can be either timelike and null at any of its points. The canonical time orientation of $\mathbb{R}^{n+1}$ is given by the vectorfield $T_0 = \partial_0$. A timelike vector X is said to be *future oriented* if $\mathbf{m}(X, T_0) < 0$ and *past oriented* if $\mathbf{m}(X, T_0) > 0$. The causal future $J^+(S)$ of a set S consists of all points in $\mathbb{R}^{n+1}$ which can be connected to S by a future directed causal curve. The causal past $\mathcal{J}^-(S)$ is defined in the same way. Thus, for a point $p = (t, x)$, $\mathcal{J}^+(p) = \{(t \geq t_0, x)/|x - x_0| \leq t - t_0\}$. Given a smooth domain D, its future set $\mathcal{J}^+(D)$ may, in general, have a nonsmooth boundary, due to caustics.

We consider *conservative domains* $\mathcal{J}^+(D_1) \cap \mathcal{J}^-(D_2)$ with $D_1 \subset \Sigma_1$, $D_2 \subset \Sigma_2$, spacelike hypersurfaces. The domain is *regular* if both D_1, D_2 are regular and its non- spacelike boundaries $\mathcal{N}_1 \subset \partial(\mathcal{J}^+(D_1)) \setminus D_1$ and $\mathcal{N}_2 \subset \partial(\mathcal{J}^-(D_2)) \setminus D_2$ are smooth. In the particular case, when $D_1 = \Sigma_1$ and $D = D_2 \subset \Sigma_2$, we obtain $\mathcal{J}^+(\Sigma_1) \cap \mathcal{J}^-(D)$, called *domain of dependence* of D relative to Σ_1, consisting of all points in the causal past of $D \subset \Sigma_2$, to the future of Σ_1. Similarly $\mathcal{J}^+(D) \cap \mathcal{J}^-(\Sigma_2)$, with $D \subset \Sigma_1$ is called the domain of *dependence of influence* of D relative to Σ_2. Particularly useful examples are given in terms of a time function $\mathbf{t}$ with $\Sigma_1 = \{(t, x)/\mathbf{t}(t, x) = t_1\}$, $\Sigma_2 = \{(t, x)/\mathbf{t}(t, x) = t_1\}$ two, nonintersecting, level hypersurfaces, Σ_2 lying in the future of Σ_1.

A pair of null vectorfields $L, \underline{L}$ form a *null pair* if $\mathbf{m}(L, \underline{L}) = -2$. A null pair $e_n = L, e_{n+1} = \underline{L}$ together with vectorfields $e_1, \dots e_{n-1}$ such that $\mathbf{m}(L, e_a) = \mathbf{m}(\underline{L}, e_a) = 0$ and $\mathbf{m}(e_a, e_b) = \delta_{ab}$, for all $a, b = 1, \dots, n-1$, is called a *null frame.* The null pair,

$$\mathbf{L} = \partial_t + \partial_r, \quad \underline{\mathbf{L}} = \partial_t - \partial_r, \tag{301}$$

with $r = |x|$ and $\partial_r = x^i/r\partial_i$, is called *canonical.* Simmilarly a null frame $e_1, \dots e_{n+1}$ with $e_n = \mathbf{L}, e_{n+1} = \underline{\mathbf{L}}$ is called a canonical null frame. In that case $e_1, \dots, e_{n-1}$ form, at any point, an orthonormal basis for the the sphere $S_{t,r}$, of constant t and r, passing through that point. Observe also that $\mathbf{L}$ is the null geodesic generator associated to $\mathbf{u} = t - r$ while $\underline{\mathbf{L}}$ the null geodesic of $\underline{u} = t + r$.

3.2. Conformal Killing vectorfields. Let x^μ be an inertial coordinate system of Minkowski space $\mathbb{R}^{n+1}$. The following are all the isometries and conformal isometries of $\mathbb{R}^{n+1}$.

1. Translations: for any given vector $a = (a^0, a^1, ..., a^n) \in \mathbb{R}^{n+1}$,

$$x^\mu \to x^\mu + a^\mu$$

2. Lorentz rotations: Given any $\Lambda = \Lambda^\rho_\sigma \in \mathbf{O}(1, n)$,

$$x^\mu \to \Lambda^\mu_\nu \, x^\nu$$

3. Scalings: Given any real number $\lambda \neq 0$,

$$x^\mu \to \lambda x^\mu$$

4. Inversion: Consider the transformation $x^\mu \to I(x^\mu)$, where

$$I(x^\mu) = \frac{x^\mu}{(x, x)}$$

defined for all points $x \in \mathbb{R}^{n+1}$ such that $(x, x) \neq 0$.

The first two sets of transformations are isometries of $\mathbb{R}^{n+1}$, the group generated by them is called the Poincarè group. The last two type of transformations are conformal isometries. the group generated by all the above transformations is called the Conformal group. In fact the Liouville theorem, whose infinitesimal version will be proved later on, states that it is the group of all the conformal isometries of $\mathbb{R}^{n+1}$.

We next list the Killing and conformal Killing vector fields which generate the above transformations.

i. The generators of translations in the x^μ directions, $\mu = 0, 1, ..., n$:

$$\mathbf{T}_\mu = \frac{\partial}{\partial x^\mu}$$

ii. The generators of the Lorentz rotations in the (μ, ν) plane:

$$\mathbf{L}_{\mu\nu} = x_\mu \partial_\nu - x_\nu \partial_\mu$$

iii. The generators of the scaling transformations:

$$\mathbf{S} = x^\mu \partial_\mu$$

iv. The generators of the inverted translations [3]:

$$\mathbf{K}_\mu = 2 x_\mu x^\rho \frac{\partial}{\partial x^\rho} - (x^\rho x_\rho) \frac{\partial}{\partial x^\mu}$$

[3]Observe that the vector fields $\mathbf{K}_\mu$ can be obtained applying I_* to the vector fields $\mathbf{T}_\mu$.

We also list below the commutator relations between these vector fields,

$$\begin{aligned}
&[\mathbf{L}_{\alpha\beta}, \mathbf{L}_{\gamma\delta}] = \eta_{\alpha\gamma}\mathbf{L}_{\beta\delta} - \eta_{\beta\gamma}\mathbf{L}_{\alpha\delta} + \eta_{\beta\delta}\mathbf{L}_{\alpha\gamma} - \eta_{\alpha\delta}\mathbf{L}_{\beta\gamma}\\
&[\mathbf{L}_{\alpha\beta}, \mathbf{T}_{\gamma}] = \eta_{\alpha\gamma}\mathbf{T}_{\beta} - \eta_{\beta\gamma}\mathbf{T}_{\alpha}\\
&[\mathbf{T}_{\alpha}, \mathbf{T}_{\beta}] = 0\\
&[\mathbf{T}_{\alpha}, \mathbf{S}] = \mathbf{T}_{\alpha}\\
&[\mathbf{T}_{\alpha}, \mathbf{K}_{\beta}] = 2(\eta_{\alpha\beta}\mathbf{S} + \mathbf{L}_{\alpha\beta})\\
&[\mathbf{L}_{\alpha\beta}, \mathbf{S}] = [\mathbf{K}_{\alpha}, \mathbf{K}_{\beta}] = 0\\
&[\mathbf{L}_{\alpha\beta}, \mathbf{K}_{\gamma}] = \eta_{\alpha\gamma}\mathbf{K}_{\beta} - \eta_{\beta\gamma}\mathbf{K}_{\alpha}
\end{aligned} \tag{302}$$

Denoting $\mathcal{P}(1,n)$ the Lie algebra generated by the vector fields $\mathbf{T}_{\alpha}, \mathbf{L}_{\beta\gamma}$ and $\underline{\mathcal{K}}(1,n)$ the Lie algebra generated by all the vector fields $\mathbf{T}_{\alpha}, \mathbf{L}_{\beta\gamma}, \mathbf{S}, \mathbf{K}_{\delta}$ we state the following version of the Liouville theorem,

THEOREM 3.3. *The following statements hold true.*

1) $\mathcal{P}(1,n)$ is the Lie algebra of all Killing vector fields in $\mathbb{R}^{n+1}$.

2) If $n > 1$, $\underline{\mathcal{K}}(1,n)$ is the Lie algebra of all conformal Killing vector fields in $\mathbb{R}^{n+1}$.

3) If $n = 1$, the set of all conformal Killing vector fields in $\mathbb{R}^{1+1}$ is given by the following expression

$$f(x^0 + x^1)(\partial_0 + \partial_1) + g(x^0 - x^1)(\partial_0 - \partial_1)$$

where f, g are arbitrary smooth functions of one variable.

Proof: The proof for part 1 of the theorem follows immediately, as a particular case, from Proposition (0.15). From (399) as $\mathbf{R} = 0$ and X is Killing we have

$$D_{\mu}D_{\nu}X_{\lambda} = 0\ .$$

Therefore, there exist constants $a_{\mu\nu}, b_{\mu}$ such that $X^{\mu} = a_{\mu\nu}x^{\nu} + b_{\mu}$. Since X is Killing $D_{\mu}X_{\nu} = -D_{\nu}X_{\mu}$ which implies $a_{\mu\nu} = -a_{\nu\mu}$. Consequently X can be written as a linear combination, with real coefficients, of the vector fields $T_{\alpha}, L_{\beta\gamma}$.

Let now X be a conformal Killing vector field. There exists a function Ω such that

$${}^{(X)}\pi_{\rho\sigma} = \Omega\eta_{\rho\sigma} \tag{303}$$

From (399) and (400) it follows that

$$D_{\mu}D_{\nu}X_{\lambda} = \frac{1}{2}\left(\Omega_{,\mu}\eta_{\nu\lambda} + \Omega_{,\nu}\eta_{\mu\lambda} - \Omega_{,\lambda}\eta_{\nu\mu}\right) \tag{304}$$

Taking the trace with respect to μ, ν, on both sides of (304) we infer that

$$\begin{aligned}
\Box X_{\lambda} &= -\frac{n-1}{2}\Omega_{\lambda}\\
D^{\mu}X_{\mu} &= \frac{n+1}{2}\Omega
\end{aligned} \tag{305}$$

and applying D^{λ} to the first equation, $\Box$ to the second one and subtracting we obtain

$$\Box\Omega = 0 \tag{306}$$

Applying D_μ to the first equation of (305) and using (306) we obtain

$$\begin{aligned}(n-1)D_\mu D_\lambda \Omega &= \frac{n-1}{2}(D_\mu D_\lambda \Omega + D_\lambda D_\mu \Omega) = -\Box(D_\mu X_\lambda + D_\lambda X_\mu)\\ &= -(\Box\Omega)\eta_{\mu\lambda} = 0 \end{aligned} \qquad (307)$$

Hence for $n \neq 1$, $D_\mu D_\lambda \Omega = 0$. This implies that Ω must be a linear function of x^μ. We can therefore find a linear combination, with constant coefficients, $cS + d^\alpha K_\alpha$ such that the deformation tensor of $X - (cS + d^\alpha K_\alpha)$ must be zero. This is the case because ${}^{(S)}\pi = 2\eta$ and ${}^{(K_\mu)}\pi = 4x_\mu \eta$. Therefore $X - (cS + d^\alpha K_\alpha)$ is Killing which, in view of the first part of the theorem, proves the result.

Part 3 can be easily derived by solving (303). Indeed posing $X = a\partial_0 + b\partial_1$, we obtain $2D_0X_0 = -\Omega$, $2D_1X_1 = \Omega$ and $D_0X_1 + D_1X_0 = 0$. Hence a, b verify the system

$$\frac{\partial a}{\partial x^0} = \frac{\partial b}{\partial x^1}\ ,\ \frac{\partial b}{\partial x^0} = \frac{\partial a}{\partial x^1}\ .$$

Hence the one form $adx^0 + bdx^1$ is exact, $adx^0 + bdx^1 = d\phi$, and $\frac{\partial^2 a}{\partial x^{0\,2}} = \frac{\partial^2 b}{\partial x^{1\,2}}$, that is $\Box\phi = 0$. In conclusion

$$X = \frac{1}{2}\left(\frac{\partial\phi}{\partial x^0} + \frac{\partial\phi}{\partial x^1}\right)(\partial_0 + \partial_1) + \frac{1}{2}\left(\frac{\partial\phi}{\partial x^0} - \frac{\partial\phi}{\partial x^1}\right)(\partial_0 - \partial_1)$$

which proves the result.

Remark. Expresse relative to the canonical null pair,

$$\mathbf{T}_0 = 2^{-1}(\mathbf{L} + \underline{\mathbf{L}}), \quad \mathbf{S} = 2^{-1}(\underline{u}\,\mathbf{L} + \mathbf{u}\,\underline{\mathbf{L}}), \quad \mathbf{K}_0 = 2^{-1}(\underline{u}^2\,\mathbf{L} + \mathbf{u}^2\,\underline{\mathbf{L}}). \qquad (308)$$

Both $\mathbf{T}_0 = \partial_t$ and $\mathbf{K}_0 = (t^2 + |x|^2)\partial_t + 2tx^i\partial_i$ are causal. This makes them important in deriving energy estimates. Observe that $\mathbf{S}$ is causal only in $\mathcal{J}^+(0) \cup \mathcal{J}^-(0)$.

3.4. Null hypersurfaces. Null hypersurfaces are particularly important as they correspond to the propagation fronts of solutions to the wave or Maxwell equation in Minkowski space[4]. The simplest way to describe the geometry of a null hypersurfaces is to start with a codimension one hypersurface $S_0 \subset \Sigma_0$, where Σ_0 is a fixed spacelike hypersurface of $\mathbb{M}^{n+1}$. At every point $p \in S_0$ there are precisely two null directions ortogonal to the tangent space $T_p(S_0)$. Let L denote a smooth null vectorfield orthogonal to S_0 and consider the congruence of null geodesics[5] generated by the integral curves of L. As long as these null geodesics do not intersect the congruence forms a smooth null hypersurface $\mathcal{N}$. We can also extend L, by parallel transport, to all points of $\mathcal{N}$. Clearly $D_L L = 0$, $\mathbf{m}(L, L) = 0$, moreover $\mathbf{m}(L, X) = 0$ for every vector X tangent to $\mathcal{N}$. Observe also that L is uniquely defined up to multiplication by a conformal factor depending only on S_0. Define, for all vectorfields X, Y tangent to $\mathcal{N}$,

$$\gamma(X, Y) = \mathbf{m}(X, Y), \qquad \chi(X, Y) = \mathbf{m}(D_X L, Y) \qquad (309)$$

They are both symmetric tensors, called, respectively, the first and second null fundamental forms of $\mathcal{N}$. Observe that χ is uniquely defined up to the same conformal

[4]Or more generally on a Lorentz spacetime.

[5]These are in fact straight lines in Minkowski space.

factor associated to L. Clearly $\gamma(L,X)=\chi(L,X)=0$ for all X tangent to $\mathcal{N}$, therefore they both depend, at a fixed $p\in\mathcal{N}$, only on a fixed hyperplane transversal to L_p. Define s, called affine parameter, by the condition $L(s)=1$, $s=0$ on S_0. Its level surfaces defines the *geodesic foliation* of $\mathcal{N}$. Given coordinates $w=(\omega^a)$, $a=1,\dots n-1$ on S_0 we can parametrize points on S_s by the flow $x^\mu(s,\omega)$ defined by $\frac{dx^\mu}{ds}=L^\mu$ with $x^\mu(0,\omega)$ the point on S_0 of coordinates w. Let,

$$\gamma_{ab}=\gamma(\frac{\partial}{\partial\omega^a},\frac{\partial}{\partial\omega^a}),\quad \chi_{ab}=\chi(\frac{\partial}{\partial\omega^a},\frac{\partial}{\partial\omega^b})$$

denote the components of γ and χ relative to these coordinates. One can easily check that $\frac{d}{ds}\gamma_{ab}=2\chi_{ab}$. The volume element of S_s is given by

$$da_{S_s}=\sqrt{|\gamma|}d\omega^1\dots dw^{n-1}$$

with γ the determinant of the metric γ. Observe that $\frac{d}{ds}\log|\gamma|=\gamma^{ab}\frac{d}{ds}\gamma_{ab}=2\mathrm{tr}\chi$, with $\mathrm{tr}\chi=\gamma^{ab}\chi_{ab}$ the *expansion coefficient* of the null hypersurface. Thus,

$$\frac{d}{ds}\sqrt{|\gamma|}=\mathrm{tr}\chi\sqrt{|\gamma|}.$$

The rate of change of the total volume $|S_s|$ is given by the following formula,

$$\frac{d}{ds}|S_s|=\int_{S_s}\mathrm{tr}\chi da_{S_s}. \tag{310}$$

We also remark that χ verifies the following Ricatti type equation,

$$\frac{d}{ds}\chi+\chi^2=0 \tag{311}$$

which can be explicitely integrated. Thus one can verify that $\mathrm{tr}\chi(s,\omega_0)$ may become $-\infty$ at a finite value of $s>0$ if $\mathrm{tr}\chi(0,\omega_0)<0$ at some point of S_0. This occurence corresponds to the formation of a *caustic*.

An arbitrary foliation S_v on $\mathcal{N}$ can be parametrized by $v(s,\omega)$ with (s,ω) the geodesic coordinates defined above. We call $\Omega=\frac{dv}{ds}$ the null lapse function of the foliation and denote by γ' and χ' the restiction of γ,χ to S_v. If X is a vectorfield tangent to the geodesic foliation S_s then $X'=X-\Omega^{-1}X(v)L$ is tangent to S_v. Thus, if X,Y are tangent to S_s then $\gamma(X,Y)=\gamma(X',Y')$ and $\chi(X',Y')=\chi(X,Y)$. Relative to the coordinates (v,ω) we have

$$\gamma'_{ab}=\gamma_{ab},\qquad \chi'_{ab}=\chi_{ab}.$$

To define the volume element on a null hypersurface $\mathcal{N}$ we choose an arbitrary foliation v with null lapse function $\frac{dv}{ds}=\Omega$ and induced metric γ and set

$$da_{\mathcal{N}}=\Omega^{-1}da_{S_v}dv \tag{312}$$

where da_{S_v} denotes the area element of S_v induced by γ. The definition does not depend on the particular foliation.

3.5. Energy momentum tensor. An energy momentum tensor in $\mathbb{R}^{n+1}$ is a symmetric two tensor $Q_{\alpha\beta}$ verifying the *positive energy condition*,

$$Q(X,Y) \geq 0,$$

for all X, Y causal, future oriented. We say that Q is *divergenceless* if,

$$D^\beta Q_{\alpha\beta} = 0 \tag{313}$$

Given an arbitrary vectorfield X,

$$D^\alpha(Q_{\alpha\beta}X^\beta) = Q^{\alpha\beta}D_\alpha X_\beta = \frac{1}{2}Q^{\alpha\beta}\,{}^{(X)}\pi_{\alpha\beta},$$

where ${}^{(X)}\pi = \mathcal{L}_X\mathbf{m}$ denotes the deformation tensor of X. Recall that ${}^{(X)}\pi_{\alpha\beta} = \partial_\alpha X_\beta + \partial_\beta X_\alpha$. In the particular case when X is a Killing vectorfield, that is ${}^{(X)}\pi = 0$, we derive

$$D^\alpha(Q_{\alpha\beta}X^\beta) = 0. \tag{314}$$

The same identity holds if X is conformal Killing and Q is *traceless*, that is $\mathbf{m}^{\alpha\beta}Q_{\alpha\beta} = 0$.

A typical conservation law is obtained when we integrate the latter identity, and apply Stokes theorem, on a *regular conservative* spacetime domain(see section 3.1) $\mathcal{J}^+(D_1) \cap \mathcal{J}^-(D_2)$ with smooth spacelike boundaries $D_i \subset \Sigma_i$ and null boundaries $\mathcal{N}_i$, $i = 1, 2$. We denote by T_1, T_2 the future unit normals to the spacelike hypersurfaces Σ_1, Σ_2 and chose the null normals L_1, L_2 such that $\mathbf{m}(L_i, T_i) = -1$ along the boundaries $D_i \subset \Sigma_i$, $i = 1, 2$. For simplicity we denote both timelike normals by T and both null normals by L whenever there is no possibility of confusion.

Proposition 3.6. *Assume that $Q_{\alpha\beta}$ is a divergenceless energy momentum tensor and X a Killing vectorfield in a neighborhood of the regular conservative domain $\mathcal{J}(D_1, D_2)$ as above. Then,*

$$\int_{\mathcal{N}_2} Q(X,L) + \int_{D_2} Q(X,T) = \int_{\mathcal{N}_1} Q(X,L) + \int_{D_1} Q(X,T) \tag{315}$$

The integrals are taken with respect to the area elements $da_{\mathcal{N}}$ along the null hypersurfaces $\mathcal{N}_1, \mathcal{N}_2$ and the area elements of the Riemannian metrics induced by $\mathbf{m}$ on Σ_1, Σ_2. Observe that all integrands are positive if X is causal. The identity (315) *remains valied if X is conformal Killing and Q is traceless.*

Proof : Let $P_\alpha = Q_{\alpha\beta}X^\beta$. According to eqrefeq:cons-law1 we have $\mathbf{D}^\alpha P_\alpha = 0$. ■

The result simplifies for domains of dependence $\mathcal{J}^+(\Sigma_1) \cap \mathcal{J}^-(D \subset \Sigma_2)$, or influence $\mathcal{J}^+(D \subset \Sigma_1) \cap \mathcal{J}^-(\Sigma_2)$, with Σ_2 in the future of Σ_1. We normalize L by the condition $\mathbf{m}(L, T) = -1$ on $\partial D \subset \Sigma_2$ where T denotes the unit normal to Σ_1, Σ_2.

Corollary 3.7. *If Q is divergenceless, X is Killing and $D \subset \Sigma_2$,*

$$\int_{\mathcal{N}} Q(X,L) + \int_{D\subset\Sigma_2} Q(X,T) = \int_{\mathcal{J}^-(D)\cap\Sigma_1} Q(X,T) \tag{316}$$

Similarily, if $D \subset \Sigma_1$,

$$\int_{\mathcal{N}} Q(X,L) + \int_{D\subset\Sigma_1} Q(X,T) = \int_{\mathcal{J}^+(D)\cap\Sigma_2} Q(X,T) \tag{317}$$

The identity remains true if X is conformal Killing and Q is traceless.

4. Wave Equation in $\mathbb{R}^{n+1}$

We rely on the notations and results of section 3.1. The wave operator in Minkowski space $\mathbb{R}^{n+1}$ is defined by $\Box = \mathbf{m}^{\alpha\beta}\partial_\alpha\partial_\beta = -\partial_t^2 + \sum_{i=1}^n \partial_i^2$. It is the simplest scalar operator invariant with respect to the Poincaré group, consisting of both translations and Lorentz transformations, i.e. the group of isometries of $\mathbb{R}^{n+1}$. To solve the wave equation means to find solutions $\phi(t,x)$ which verify $\Box\phi = 0$. The Cauchy problem[6] for $\Box$ consists in finding solutions to $\Box\phi = 0$ with prescribed ϕ and normal derivative of ϕ on a given spacelike hypersurface Σ_0. In the particular case when Σ_0

Definition. The energy momentum tensor (see section 3.5) of a solution $\Box\phi = 0$ is given by,

$$Q_{\alpha\beta} = Q_{\alpha\beta}[\phi] = \partial_\alpha\phi\partial_\beta\phi - \frac{1}{2}\mathbf{m}_{\alpha\beta}\big(\mathbf{m}^{\mu\nu}\partial_\mu\phi\partial_\nu\phi\big). \tag{318}$$

Proposition 4.1. *The tensor Q is symmetric and divergenceless, $\partial^\beta Q_{\alpha\beta} = 0$. Moreover, for any time-like or null(that is causal), future oriented, vectorfields X, Y, we have,*

$$Q(X,Y) > 0.$$

Proof : The only part which is not immediate is the positivity of Q. Take X, Y arbitrary future oriented causal vectors. The 2-plane which they generate intersects the light cone through the origin along two distinct null directions. Choose $\mathbf{L}, \underline{\mathbf{L}}$ two null, future oriented, vectors along the these directions such that $< \mathbf{L}, \underline{L} >= -2$. Choose also vectors $(e_a)_{1=1,\dots n-1}$ such that they form a null frame together with $L, \underline{L}$. Observe that,

$$Q(L,L) = |L(\phi)|^2, \quad Q(\underline{L},\underline{L}) = |\underline{L}(\phi)|^2, \quad Q(L,\underline{L}) = |\nabla\phi|^2 = \sum_{a=1,\dots n-1} |e_a(\phi)|^2$$

On the other hand both X, Y are linear combinations of $L, \underline{L}$ with positive coefficients. ■

It is easy to observe that We are thus in a position to apply proposition 3.6 and its corollary, see section 3.5, concerning conservation laws associated to Q. In particular we derive the following,

[6]more generally one may consider, in addition to the Cauchy problem on Σ_0 a boundary condition on the timelike boundary of a spacetime domain $\subset \mathbb{R}^{n+1}$.

THEOREM 4.2 (Noëther theorem). *Consider an arbitrary solution of* $\square\phi = 0$, *a Killing vectorfield* X *and any domain of dependence* $J^-(D \subset \Sigma_2) \cap \mathcal{J}^+(\Sigma_1) \subset \mathbb{R}^{n+1}$ *with* Σ_1, Σ_2 *spacelike hypersurfaces,* $\Sigma_2 \subset \mathcal{J}^+(\Sigma_1)$, *and regular null boundary* $\mathcal{N}$. *Then, with* $Q = Q[\phi]$ *as above and* L, T *as in corollary 3.7 (section 3.5),*

$$\int_{\mathcal{N}} Q(X, L) + \int_D Q(X, T) = \int_{\mathcal{J}^-(D)\cap\Sigma_1} Q(X, T) \tag{319}$$

When $X = \mathbf{T}_0 = \partial_t$ *we obtain the law of conservation of energy. For* $X = \mathbf{T}_i = \partial_i$, $i = 1, \dots, n$ *we derive conservation of linear momentum while with* $X = \mathbf{O}_{ij} = x_i\partial_j - x_j\partial_i$ *(see section* **??***) we derive the conservation law of angular momentum.*

Observe that,

$$Q(\mathbf{T}_0, T) = \frac{1}{2}\big(|\partial_t\phi|^2 + |D\phi|^2\big).$$

where $|D\phi|$ denotes the norm of the gradient of ϕ along Σ_t. Also $Q(L, T_0) = \frac{1}{2}\big(|L\phi|^2 + |\nabla\phi|^2\big)$ with $|\nabla\phi|$ the norm of the gradient of ϕ restricted to the $n-1$ dimensional surfaces $\Sigma_t \cap \mathcal{N}$.

COROLLARY 4.3. *Consider* $D \subset \Sigma_2 \subset \mathcal{J}^+(\Sigma_1)$. *Assume that* ϕ *and its normal derivative* $T(\phi)$ *vanish on* $J^-(D) \cap \Sigma_1$ *and that* $\square\phi = 0$ *in a neighborhood of the domain of dependence* $J^-(D) \cap \mathcal{J}^+(\Sigma_1)$. *Then* $\phi \equiv 0$ *in* $J^-(D) \cap \mathcal{J}^+(\Sigma_1)$.

COROLLARY 4.4 (Huygens Principle). *Any solution of* $\square\phi = 0$ *with initial data supported in the closure of a domain in* $D \subset \Sigma_1$ *is supported in* $\mathcal{J}^-(F) \cup \mathcal{J}^+(F)$.

4.5. Representation formulas. The above uniqueness results applies in particular to the *standard* initial value problem (i.v.p.) for the equation,

$$\square\phi = F, \qquad \phi(0, x) = f(x),\ \partial_t\phi(0, x) = g(x) \tag{320}$$

According to the results of the previous section any two solutions of (320) must coincide. By the principle of superposition[7] to solve (320) it suffices to consider, separately,

Case 1. $F = 0$ and f, g arbitrary,

Case 2. $f = g = 0$ and F arbitrary.

Case 1 can be further reduced to what is called *reduced i.v.p.*,

$$\square\phi = 0, \qquad \phi(0, x) = f(x),\ \partial_t\phi(0, x) = 0. \tag{321}$$

Exercise. Show how to deduce the general homogeneous solution of case 1 from the reduced problem.

We have already found a fundamental solution for $\square$,

$$E_+^{(n+1)}(t, x) = c_n H(t)\chi_+^{-\frac{n-1}{2}}(t^2 - x^2), \tag{322}$$

[7]that is linearity of $\square$

We can now show, using the results of the previous section, that $E_+^{(n+1)}$ is the unique fundamental solution of $\Box$, supported in the upper half plane $t \geq 0$[8].

The fundamental solution takes a particularly simply form for $n = 3$ and $n = 2$. Indeed, for $n = 3$, $\chi_+^{-1} = \delta_0$, the one dimensional Dirac measure supported at the origin. In that case the solution to the reduced initial value problem takes the form,

$$\phi(t, x) = \frac{1}{4\pi t} \int_{|x-y|=t} g(y) dS_y \tag{323}$$

For $n = 2$ we have $\chi_+^{-1/2}(\lambda) = \lambda_+^{-1/2}$, with λ_+ the positive part of λ. In all other odd dimensions, $n \geq 3$, the fundamental solution E_+ can be expressed in terms of derivatives of δ_0. The case of even dimensions can be reduced to odd dimensions by the so called *method of descent*. In particular, for two space dimensions the solution to the reduced Initial value problem takes the form,

$$\phi(t, x) = \frac{1}{2\pi t} \int_{|x-y| \leq t} \frac{1}{\sqrt{t^2 - |x-y|^2}} g(y) dy \tag{324}$$

Exercise. Derive (324) from (323) by interpreting solutions $\phi(t, x^1, x^2)$ of $\Box\phi = 0$ in $\mathbb{R}^{1+2}$ as solutions $\phi(t, x^1, x^2, x^3)$ of $\Box\phi = 0$ in $\mathbb{R}^{1+3}$ which are constant in x^3.

Remark. It is a remarkable fact that in all odd space dimensions[9] the fundamental solution is supported on the boundary of the future null cone of the origin, $\{(t \geq 0, x)/t^2 - |x|^2 = 0\}$. This is called *Strong Huygens Principle.*

The fundamental solution allows us to find explicit representations for (321). There are three other known methods of solving directly (321), without the a-priori knowledge of the fundamental solution.

Fourier transform. The best known method is based on taking the Fourier transform of equation (321) with respect to the space variables. Thus, denoting by $\hat{\phi}(t, \xi)$ the Fourier transform of $\phi(t, x)$ in x, one derives $\partial_t^2 \hat{\phi} + |\xi|^2 \hat{\phi} = 0$ and $\hat{\phi}(0) = \hat{f}$, $\partial_t \hat{\phi}(0) = 0$. Hence, solving the differential equation and using the inversion formula for the Fourier transform,

$$\phi(t, x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \cos(t|\xi|) e^{ix\cdot\xi} \hat{f}(\xi) d\xi \tag{325}$$

Plane waves. The method is based on the observation that if $g_\omega(x) = g(x \cdot \omega)$, for $\omega \in \mathbb{S}^{n-1}$, then $\phi(t, x) = 2^{-1}\big(g_\omega(x \cdot \omega + t) + g_\omega(x \cdot \omega - t)\big)$ verifies (321) with $f = g_\omega$. On the other hand, for odd $n \geq 3$, an arbitrary smooth function f can be expressed in the form [10] $f(x) = c_n \int_{|\omega|=1} g_\omega(x) dS_\omega$ with $g_\omega(x) = \int_{\mathbb{R}^n} |(x - y) \cdot \omega| \, \Delta_y^{(n+1)/2} f(y) dy$. Alternatively one can reexpress (325) using polar coordinates.

---

[8] in fact it is supported in the future null cone with vertex at the origin, $|x| \leq t$.

[9] while for even dimensions the support of the fundamental solution extends to the interior of the cone

[10] For some constant c_n. Indeed $\int_{\mathbb{R}^n} |(x - y)\omega| = a_n |x - y|$ for some constant a_n. Also, using the fundamental solution of Δ, $\Delta^{(n+1)/2}|x - y| = b_n \delta_0(x - y)$ for another constant b_n.

Thus, for odd n,

$$\begin{aligned}
\phi(t,x) &= (2\pi)^{-n}\int_{|\omega|=1} dS_\omega \int_0^\infty \cos(t\lambda)\mathrm{e}^{i\lambda(x\cdot\omega)}\hat{f}(\lambda\omega)\lambda^{n-1}d\lambda \\
&= \frac{1}{2}(2\pi)^{-n}\int_{|\omega|=1} dS_\omega \int_{-\infty}^\infty \cos(t\lambda)\mathrm{e}^{i\lambda(x\cdot\omega)}\hat{f}(\lambda\omega)\lambda^{n-1}d\lambda \\
&= \frac{1}{4}(2\pi)^{-n}\int_{|\omega|=1} dS_\omega \int_{-\infty}^\infty (\mathrm{e}^{i\lambda(t+x\cdot\omega)}+\mathrm{e}^{i\lambda(t-x\cdot\omega)})\hat{f}(\lambda\omega)\lambda^{n-1}d\lambda \\
&= \frac{1}{4}(2\pi)^{-n}\int_{\mathbb{R}^n} f(y)dy\int_{|\omega|=1} dS_\omega\Big(\int_{-\infty}^\infty (\mathrm{e}^{i\lambda(t+(x-y)\cdot\omega)}+\mathrm{e}^{i\lambda(t-(x-y)\cdot\omega)}\big)\lambda^{n-1}d\lambda\Big) \\
&= \frac{1}{4}(2\pi)^{-n+1}\int_{|\omega|=1}\int_{\mathbb{R}^n}\big(\delta_0^{(n-1)}(t+(x-y)\cdot\omega)+\delta_0^{(n-1)}(t-(x-y)\cdot\omega)\big)f(y)dy
\end{aligned}$$

where $\delta_0^{(n-1)}$ denotes the $n-1$ derivative[11] of the Dirac measure δ_0. Therefore,

$$\phi(t,x) = \int_{|\omega|=1}\frac{d^{n-1}}{dt^{n-1}}\big(p_+(f,\omega)+(-1)^{n-1}p_-(f,\omega)\big)(t,x)dS_\omega \tag{326}$$

where $p_\pm(f,\omega)$ define the plane waves, $p_\pm(f,\omega)(t,x) = 4^{-1}(2\pi)^{-n+1}\int_{(x-y)\cdot\omega=\mp t} f(y)\,dS_y$. In the particular case of dimension $n=1$ we derive

$$\phi(t,x) = 2^{-1}\big(f(x-t)+f(t+x)\big). \tag{327}$$

Spherical means. One considers the spherical means of a function g in $\mathbb{R}^n$, $M_g(x,r) = |S(x,r)|^{-1}\int_{S(x,r)} g(y)dS_y$ with $S(x,r)$ the sphere of radius r centered at x and $|S(x,r)|$ its area. It is easy to see that $M_g(x,r)$ verifies the Darboux equation $(\partial_r^2+\frac{n-1}{r}\partial_r)M_g = \Delta M_g$. If ϕ verifies (321) then $M_\phi(t,r,x)$ verifies the Euler -Poisson-Draboux equation

$$\partial_t^2(M_\phi) = (\partial_r^2+\frac{n-1}{r}\partial_r)M_\phi, \quad M_\phi(0,r,x) = M_f(r,x),\ \partial_t M_\phi(0,r,x) = 0.$$

This can be explicitely solved for odd values of n. In the particular case $n=3$,[12]

$$\phi(t,x) = \partial_t\big((4\pi t)^{-1}\int_{|x-y|=t} f(y)dS_y\big) \tag{328}$$

Formulas (325)–(328) can be easily extended to $\Box\phi = 0$, $\phi(0,x) = f$, $\partial_t\phi(0,x) = g(x)$. To solve the inhomogeneous problem $\Box\phi = F$ one needs to rely on the following,

Duhamel Principle. The solution to $\Box\phi = F, \phi(0) = \phi_t(0) = 0$ can be expressed in the form, $\phi(t,x) = \int_0^t \Phi_s(t,x)ds$ where, for every $0\le s\le t$, $\Phi_s(t,x)$ verifies $\Box\Phi_s = 0$ with initial data at time s, $\Phi_s(s,x) = 0$, $\partial_t\Phi_s(s,x) = F(t,x)$.

[11] in the sense of distributions

[12] Clearly (328) can also be derived from (325), by evaluating $\int\cos(t|\xi|)e^{i(x-y)\cdot\xi}d\xi$.

4.6. A-priori estimates. We can see from both representation formulas (326) and (328) that the solutions $\phi(t,x)$ of $\square\phi = 0$ in $\mathbb{R}^{n+1}$, $n > 1$, lose derivatives in the uniform L^∞ norm relative to the space variables x. One can show that this phenomenon, due to focusing of waves, holds true any L^p norm with $p \neq 2$. For $p = 2$, on the other hand, the law of conservation of energy gives,

$$\|\partial_t\phi(t)\|^2_{L^2(\mathbb{R}^n)} + \sum_{i=1}^n \|\partial_i\phi(t)\|^2_{L^2(\mathbb{R}^n)} = \sum_{i=1}^n \|\partial_i f\|^2_{L^2(\mathbb{R}^n)} \tag{329}$$

This follows easily from theorem 4.2 applied to $D = \Sigma_2$ and Σ_1, Σ_2 level hypersurfaces of the standard time function $t = x^0$. This global energy identity can also be derived, by Plancherel formula, from the Fourier representation formula (325).

In particular we have the energy inequalities

$$\|\partial_t\phi(t)\|_{L^2}, \|\nabla\phi(t)\|_{L^2} \le \|\nabla f\|_{L^2}.$$

Thus, if $f \in H^1(\mathbb{R}^n)$ the solution ϕ remains in $H^1(\mathbb{R}^n)$ for any later time $t \ge 0$. Morever, using the fact that all partial derivatives ∂_i commute with $\square$, one can easily show that,

$$\sup_{t\ge 0} \|\partial\phi(t)\|_{H^s} \le \|f\|_{H^{s+1}} \tag{330}$$

In particular $f \in H^s(\mathbb{R}^n)$ implies $\phi(t) \in H^s(\mathbb{R}^n)$. Also, for every positive integer k, $\partial_t^k\phi(t) \in H^{s-k}(\mathbb{R}^n)$. Thus, in particular, $f \in C^\infty$ implies $\phi \in C^\infty$. Singularities of f, however, *propagate*, along null hypersurfaces, to all spacetime. This fact is in sharp contrast to solutions of the boundary value problem for the Laplace equation(see section ?) $\Delta\phi = 0$, in a regular open domain $D \subset \mathbb{R}^n$, which are automatically in $C^\infty(D)$, independent of the regularity at the boundary of D. Precise information about the propagation of singularities can be given using *wave front sets* and *bicharacteristics.*

Estimate (330) to derive a global uniform bound for ϕ. Indeed, using the Sobolev inequality in $\mathbb{R}^n$, $\|g\|_{L^\infty} \lesssim \|g\|_{H^s(\mathbb{R}^n)}$, with $s > n/2$, we infer that,

$$\|\partial\phi\|_{L^\infty(\mathbb{R}^{n+1})} \lesssim \|f\|_{H^{s+1}(\mathbb{R}^n)}, \quad s > n/2 \tag{331}$$

Thus L^2 bounds for sufficiently many derivatives of the initial data f assures the uniform boundedness of solution ϕ of (321). What is significant in this derivation of uniform boundedness is its a-priori character, that is we did not need to appeal to the exact form of solutions. This plays a fundamental role in dealing with more complicated situations, when the exact form solutions is impossible to establish. In fact one can use an extension of the method presented above, called *invariant vectorfield method* to derive not just uniform boundedness but also uniform decay. Indeed one can see from the explicit representation in terms of spherical means that solutions $\phi(t,x)$ to (321), corresponding to sufficiently smooth, compactly supported, data, decay uniformly in time like $t^{-(n-1)/2}$. One can derive this fact, by a-priori estimates, observing that $\square$ commutes not only with the coordinate derivatives $\mathbf{T}_\alpha = \partial_\alpha$ but also with the Killing vectorfields $\mathbf{O}_{\alpha\beta}$, that is $[\square, \mathbf{O}_{\alpha\beta}] = 0$. Morever, $[\square, \mathbf{S}] = -2\square$. Thus if Γ^k denotes any product of k vectorfields $\mathbf{T}, \mathbf{O}, \mathbf{S}$,

$$\square\phi = 0 \Rightarrow \square\Gamma^k\phi = 0.$$

As in the derivation[13] of (330), we infer that

$$\|\partial\Gamma^k\phi(t)\|_{L^2} \le I_k(f),$$

for a constant depending on f and k. Denoting

$$\mathcal{E}_s[\partial\phi](t) = \sum_{\Gamma, 0\le k\le s} \|\partial\Gamma^k\phi(t)\|_{L^2},$$

we infer that,

$$\mathcal{E}_s[\partial\phi](t) \lesssim I_s(f).$$

Finally, using a *global Sobolev inequality* for $s > n/2$, $t \ge 0$,

$$|\partial\phi(t,x)| \lesssim (1+t+|x|)^{-(n-1)/2}(1+|t-|x||)^{-1/2}\, I_s(f). \tag{332}$$

In particular, if $I_s(f)$ is finite,

$$\|\partial\phi(t)\|_{L^\infty} \lesssim (1+t)^{-(n-1)/2}$$

as desired. In fact (332) provides more information, most of the energy of ϕ propagates along the boundary of the outgoing null cones $t-|x| = \mathbf{u}$, for $t \ge 0$. Moreover one can easily show that, relative to a canonical null frame $\mathbf{L}, \underline{\mathbf{L}}$, e_a, $a = 1, \dots, n-1$, the derivatives $\mathbf{L}(\phi)$, $e_a(\phi)$ decay as $t^{-(n+1)/2}$ as $t \to \infty$, while $\underline{\mathbf{L}}(\phi)$ improves only by a power of the degenerate weight $\mathbf{u}$. This simple fact explains the improved behavior of *null forms*,

$$Q_{\alpha\beta}(\phi,\psi) = \partial_\alpha\phi\partial_\beta\psi - \partial_\beta\phi\partial_\alpha\psi, \quad Q_0(\phi,\psi) = \mathbf{m}^{\alpha\beta}\partial_\alpha\phi\partial_\beta\psi. \tag{333}$$

One can easily show that, for any solutions $\square\phi = \square\psi = 0$ and any null form Q, we have $\|Q(\phi,\psi)(t)\|_{L^2(\mathbb{R}^{n+1})} = O(t^{-(n+1)/2})$ as $t \to \infty$.[14]

[13] taking into account that f is smooth, compactly supported. One only needs, in fact, bounds for some weighted Sobolev norms of f.

[14] This distinguishes null forms from typical bilinear expressions in $\partial\phi, \partial\psi$ for which the corresponding decay rate is only $O(t^{-(n-1)/2})$.

CHAPTER 3

Equations Derived by the Variational Principle

1. Basic Notions

In this section we will discuss some basic examples of nonlinear wave equations which arise variationally from a relativistic Lagrangian. The fundamental objects of a relativistic field theory are

- **Space-time** $(\mathbf{M}, \mathbf{g})$ which consists of an $n+1$ dimensional manifold $\mathbf{M}$ and a Lorentz metric $\mathbf{g}$; i.e . a nondegenerate quadratic form with signature $(-1, 1, \dots, 1)$ defined on the tangent space at each point of $\mathbf{M}$. We denote the coordinates of a point in $\mathbf{M}$ by x^α, $\alpha = 0, 1, \dots, n$.

  Throughout most of this chapter the space-time will in fact be the simplest possible example - namely, the Minkowski space-time in which the manifold is $\mathbb{R}^{n+1}$ and the metric is given by
$$ds^2 = \mathbf{m}_{\alpha\beta} dx^\alpha dx^\beta = -dt^2 + \left(dx^1\right)^2 + \cdots + \left(dx^n\right)^2 \tag{334}$$
  with $t = x^0, m_{\alpha\beta} = \mathrm{diag}(-1, 1, \dots, 1)$. Recall that any system of coordinates for which the metric has the form (334) is called inertial. Any two inertial coordinate systems are related by Lorentz transformations.
- **Collection of fields** $\psi = \psi^{(1)}, \psi^{(2)}, \dots, \psi^{(p)}$ which can be scalars, tensors, or some other geometric objects[1] such as spinors, defined on $\mathbf{M}$.
- **Lagrangian density** L which is a scalar function on $\mathbf{M}$ depending only on the tensorfields ψ and the metric[2] $\mathbf{g}$.

We then define the corresponding *action* $\mathcal{S}$ to be the integral,
$$\mathcal{S} = \mathcal{S}[\psi, \mathbf{g} : \mathcal{U}] = \int_{\mathcal{U}} L[\psi] dv_{\mathbf{g}}$$
where $\mathcal{U}$ is any relatively compact set of $\mathbf{M}$. Here $dv_{\mathbf{g}}$ denotes the volume element generated by the metric $\mathbf{g}$. More precisely, relative to a local system of coordinates x^α, we have
$$dv_{\mathbf{g}} = \sqrt{-\mathbf{g}} dx^0 dx^1 \cdots dx^n = \sqrt{-\mathbf{g}} dx$$
with g the determinant of the matrix $(\mathbf{g}_{\alpha\beta})$.

By a compact variation of a field ψ we mean a smooth one-parameter family of fields $\psi_{(s)}$ defined for $s \in (-\epsilon, \epsilon)$ such that,

[1] For simplicity we restrict ourselves to covariant tensors.
[2] as well as its inverse $\mathbf{g}^{-1}$

(1) At $s=0$, $\psi_{(0)}=\psi$.
(2) At all points $p\in\mathbf{M}\setminus\mathcal{U}$ we have $\psi_{(s)}=\psi$.

Given such a variation we denote $\delta\psi:=\dot{\psi}:=\frac{d\psi_{(s)}}{ds}\Big|_{s=0}$. Thus, for small s,

$$\psi_{(s)}=\psi+s\dot{\psi}+O(s^2)$$

A field ψ is said to be *stationary* with respect to $\mathcal{S}$ if, for any compact variation $(\psi_{(s)},\mathcal{U})$ of ψ, we have

$$\frac{d}{ds}\mathbf{S}(s)\Big|_{s=0}=0$$

where,

$$\mathbf{S}(s)=\mathbf{S}[\psi_{(s)},\mathbf{g};\mathcal{U}]$$

We write this in short hand notation as

$$\frac{\delta\mathbf{S}}{\delta\psi}=0$$

Action Principle, also called the Variational Principle, states that an acceptable solution of a physical system must be stationary with respect to a given Lagrangian density called the Lagrangian of the system. The action principle allows us to derive partial differential equations for the fields ψ called the Euler-Lagrange equations. Here are some simple examples:

1. Scalar Field Equations :

One starts with the Lagrangian density

$$L[\phi]=-\frac{1}{2}\mathbf{g}^{\mu\nu}\partial_\mu\phi\partial_\nu\phi-V(\phi)$$

where ϕ is a complex scalar function defined on $(\mathbf{M},\mathbf{g})$ and $V(\phi)$ a given real function of ϕ.

Given a compact variation $(\phi_{(s)},\mathcal{U})$ of ϕ, we set $\mathcal{S}(s)=\mathcal{S}[\phi_{(s)},\mathbf{g};\mathcal{U}]$. Integration by parts gives,

$$\begin{aligned}\frac{d}{ds}\mathcal{S}(s)\Big|_{s=0} &= \int_{\mathcal{U}}[-\mathbf{g}^{\mu\nu}\partial_\mu\dot{\phi}\partial_\nu\phi-V'(\phi)\dot{\phi}]\sqrt{-\mathbf{g}}dx\\ &= \int_{\mathcal{U}}\dot{\phi}[\Box_{\mathbf{g}}\phi-V'(\phi)]dv_{\mathbf{g}}]\end{aligned}$$

where $\Box_{\mathbf{g}}$ is the D'Alembertian,

$$\Box_{\mathbf{g}}\phi=\frac{1}{\sqrt{-\mathbf{g}}}\partial_\mu\Big(\mathbf{g}^{\mu\nu}\sqrt{-\mathbf{g}}\,\partial_\nu\phi\Big)$$

.

In view of the action principle and the arbitrariness of $\dot{\phi}$ we infer that ϕ must satisfy the following Euler-Lagrange equation

$$\Box_{\mathbf{g}}\phi-V'(\phi)=0, \tag{335}$$

Equation (335) is called the *scalar wave equation* with potential $V(\phi)$.

CONFORMAL PROPERTIES **2. Wave Maps** :

The wave map equations will be defined in the context of a space-time $(\mathbf{M}, \mathbf{g})$, a Riemannian manifold N with metric h, and a mapping

$$\phi : \mathbf{M} \longrightarrow N.$$

We recall that if X is a vectorfield on $\mathbf{M}$ then $\phi_* X$ is the vectorfield on N defined by $\phi_* X(f) = X(f \circ \phi)$. If ω is a 1-form on N its pull-back $\phi^* \omega$ is the 1-form on $\mathbf{M}$ defined by $\phi^* \omega(X) = \omega(\phi_* X)$, where X is an arbitrary vectorfield on $\mathbf{M}$. Similarly the pull-back of the metric h is the symmetric 2-covariant tensor on $\mathbf{M}$ defined by the formula $(\phi^* h)(X, Y) = h(\phi_* X, \phi_* Y)$. In local coordinates x^α on $\mathbf{M}$ and y^a on N, if ϕ^a denotes the components of ϕ relative to y^a, we have,

$$(\phi^* h)_{\alpha\beta}(p) = \frac{\partial \phi^a}{\partial x^\alpha} \frac{\partial \phi^b}{\partial x^\beta} h_{ab}(\phi(p)) = \langle \frac{\partial \phi}{\partial x^\alpha}, \frac{\partial \phi}{\partial x^\beta} \rangle$$

where $< \cdot, \cdot >$ denotes the Riemannian scalar product on N.

Consider the following Lagrangian density involving the map ϕ,

$$L = -\frac{1}{2} \mathrm{Tr}_{\mathbf{g}}(\phi^* h)$$

where $\mathrm{Tr}_{\mathbf{g}}(\phi^* h)$ denotes the trace relative to $\mathbf{g}$ of $\phi^* h$. In local coordinates,

$$L[\phi] = -\frac{1}{2} \mathbf{g}^{\mu\nu} h_{ab}(\phi) \frac{\partial \phi^a}{\partial x^\mu} \frac{\partial \phi^b}{\partial x^\nu}.$$

By definition *wave maps* are the stationary points of the corresponding action. Thus by a a straightforward calculation,

$$\begin{aligned} 0 &= \frac{d}{ds} \mathcal{S}(s) \Big|_{s=0} = I_1 + I_2 & (336) \\ I_1 &= -\frac{1}{2} \int_{\mathcal{U}} \mathbf{g}^{\mu\nu} \frac{\partial h_{ab}(\phi)}{\partial \phi^c} \dot{\phi}^c \, \partial_\mu \phi^a \partial_\nu \phi^b \sqrt{-\mathbf{g}} dx \\ I_2 &= -\int_{\mathcal{U}} \mathbf{g}^{\mu\nu} h_{ab}(\phi) \partial_\mu \dot{\phi}^a \partial_\nu \phi^b \sqrt{-\mathbf{g}} dx \end{aligned}$$

After integrating by parts, relabelling and using the symmetry in b, c, we can rewrite I_2 in the form,

$$\begin{aligned} I_2 &= \int_{\mathcal{U}} \dot{\phi}^a \left(h_{ab}(\phi) \Box_{\mathbf{g}} \phi^b + \mathbf{g}^{\mu\nu} \frac{\partial h_{ab}}{\partial \phi^c} \partial_\mu \phi^c \partial_\nu \phi^b \right) dv_{\mathbf{g}} & (337) \\ &= \int_{\mathcal{U}} \dot{\phi}^a \left(h_{ab}(\phi) \Box_{\mathbf{g}} \phi^b + \frac{1}{2} \mathbf{g}^{\mu\nu} \left(\frac{\partial h_{ab}}{\partial \phi^c} + \frac{\partial h_{ac}}{\partial \phi^b} \right) \partial_\mu \phi^b \partial_\nu \phi^c \right) dv_{\mathbf{g}} \end{aligned}$$

Also, relabelling indices

$$I_1 = -\frac{1}{2} \int_{\mathcal{U}} \mathbf{g}^{\mu\nu} \frac{\partial h_{bc}}{\partial \phi^a} \dot{\phi}^a \, \partial_\mu \phi^b \partial_\nu \phi^c \, dv_{\mathbf{g}}.$$

Therefore,

$$\begin{aligned}
0 &= I_1 + I_2 \\
&= \int_{\mathcal{U}} \dot{\phi}^a \left(h_{ab} \Box_{\mathbf{g}} \phi^b + \partial_\mu \phi^b \partial_\nu \phi^c \mathbf{g}^{\mu\nu} \frac{1}{2} \left(\frac{\partial h_{ab}}{\partial \phi^c} + \frac{\partial h_{ac}}{\partial \phi^b} - \frac{\partial h_{bc}}{\partial \phi^a} \right) \right) dv_{\mathbf{g}} \\
&= \int_{\mathcal{U}} \dot{\phi}^a \left(h_{ad} \Box_{\mathbf{g}} \phi^d + \partial_\mu \phi^b \partial_\nu \phi^c \mathbf{g}^{\mu\nu} \frac{1}{2} h^{ds} h_{ad} \cdot \left(\frac{\partial h_{sb}}{\partial \phi^c} + \frac{\partial h_{sc}}{\partial \phi^b} - \frac{\partial h_{bc}}{\partial \phi^s} \right) \right) dv_{\mathbf{g}} \\
&= \int_{\mathcal{U}} \dot{\phi}^a h_{ad} \left(\Box_{\mathbf{g}} \phi^d + \partial_\mu \phi^b \partial_\nu \phi^c \mathbf{g}^{\mu\nu} \Gamma^d_{bc} \right) dv_{\mathbf{g}}
\end{aligned}$$

where $\Gamma^d_{bc} = \frac{1}{2} h^{ds} \left(\frac{\partial h_{sb}}{\partial \phi^c} + \frac{\partial h_{sc}}{\partial \phi^b} - \frac{\partial h_{bc}}{\partial \phi^s} \right)$ are the Christoffel symbols corresponding to the Riemannian metric h. The arbitrariness of $\dot{\phi}$ yields the following equation for wave maps,

$$\Box_{\mathbf{g}} \phi^a + \Gamma^a_{bc} \, \mathbf{g}^{\mu\nu} \partial_\mu \phi^b \partial_\nu \phi^c = 0 \tag{338}$$

Example: Let N be a two dimensional Riemannian manifold endowed with a metric h of the form,

$$ds^2 = dr^2 + f(r)^2 d\theta^2$$

Let ϕ be a wave map from $\mathbf{M}$ to N with components ϕ^1, ϕ^2, relative to the r, θ coordinates. Then, $\Gamma^1_{11} = \Gamma^2_{11} = \Gamma^1_{12} = \Gamma^2_{22} = 0$ and $\Gamma^1_{22} = -f'(r) f(r)$, $\Gamma^2_{12} = \frac{f'(r)}{f(r)}$. Therefore,

$$\begin{aligned}
\Box_{\mathbf{g}} \phi^1 &= f'(r) f(r) \mathbf{g}^{\mu\nu} \partial_\mu \phi^2 \partial_\nu \phi^2 \\
\Box_{\mathbf{g}} \phi^2 &= -\frac{f'(r)}{f(r)} \mathbf{g}^{\mu\nu} \partial_\mu \phi^1 \partial_\nu \phi^2
\end{aligned}$$

The equations of wave maps can be given a simpler formulation when N is a submanifold of the Euclidean space $\mathbb{R}^m$. In this case, the metric h is the Euclidean metric so the first term in (336) vanishes.

$$\begin{aligned}
\frac{d}{ds} \mathcal{S}(s) \Big|_{s=0} &= -\int_{\mathcal{U}} \mathbf{g}^{\alpha\beta} \langle \frac{\partial \phi}{\partial x^\alpha}, \frac{\partial \dot{\phi}}{\partial x^\beta} \rangle dv_{\mathbf{g}} \\
&= \int_{\mathcal{U}} < \Box \phi, \dot{\phi} > dv_{\mathbf{g}}
\end{aligned}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean scalar product and $\Box$ the D'Alembertian operator on $\mathbf{M}$. Thus the Euler-Lagrange equations take the form,

$$(\Box \phi(p))^T = 0 \tag{339}$$

where T here means the projection onto the tangent space of N at $\phi(p)$.

In the special case when $N \subset \mathbb{R}^m$ is a hypersurface, we can rewrite (339) in a more concrete form. Let ν be the unit normal on N and k the second fundamental form $k(X, Y) = \langle \mathbf{D}_X \nu, Y \rangle$, with $\mathbf{D}_X$ the standard covariant derivative of Euclidean space. The hypersurface N is defined (locally) as the level set of some real valued f. Differentiating the equation $f(\phi(x)) = 0$ with respect to local coordinates x^μ on

$\mathbf{M}$ yields $0 = <\nu(\phi), \partial_\mu \phi>$ along $\mathbf{M}$. Hence,

$$\begin{aligned} 0 &= \partial^\mu <\nu(\phi), \partial_\mu \phi> = <\Box\phi, \nu> + \mathbf{g}^{\mu\nu} <\partial_\nu \nu(\phi), \partial_\mu \phi> \\ &= <\Box\phi, \nu> + \mathbf{g}^{\mu\nu} <\nabla_{\phi_*(E_\nu)} \nu, \phi_*(E_\mu)> \end{aligned}$$

Where $\phi_*(E_\mu) = \frac{\partial \phi^i}{\partial x^\mu} \frac{\partial}{\partial y^i}$ is the pushforward of $E_\mu = \frac{\partial}{\partial x^\mu}$. In particular, $\phi_*(E_\mu)$ is tangent to N. Therefore,

$$<\Box\phi, \nu> = -k(\phi_*(E^\alpha), \phi_*(E_\alpha)) \tag{340}$$

In view of (**??**) the equation for wave maps becomes,

$$\Box\phi = -k(\phi_*(E^\alpha), \phi_*(E_\alpha))N$$

In the case when N is the standard sphere $S^{m-1} \subset \mathbb{R}^m$, $k(X,Y) = -<X,Y>$ and the equation for wave maps becomes, in coordinates x^α, y^a,

$$\Box\phi^a = -\phi^a \mathbf{g}^{\alpha\beta} <\frac{\partial \phi}{\partial x^\alpha}, \frac{\partial \phi}{\partial x^\beta}>$$

3. Maxwell equations:

An *electromagnetic field* F is an exact two form on a four dimensional manifold $\mathbf{M}$. That is, F is an antisymmetric tensor of rank two such that

$$F = dA \tag{341}$$

where A is a one-form on $\mathbf{M}$ called a *gauge potential* or connection 1-form. Note that A is not uniquely defined - indeed if χ is an arbitrary scalar function then the transformation

$$A \longrightarrow \tilde{A} = A + d\chi \tag{342}$$

yields another gauge potential $\tilde{A}$ for F. This degree of arbitrariness is called gauge freedom, and the transformations (342) are called gauge transformations.

The Lagrangian density for electromagnetic fields is

$$L[F] = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}.$$

Any compact variation $(F_{(s)}, \mathcal{U})$ of F can be written in terms of a compact variation $(A_{(s)}, \mathcal{U})$ of a gauge potential A, so that $F_{(s)} = dA_{(s)}$. Write

$$\dot{F} = \frac{d}{ds} F_{(s)} \Big|_{s=0}, \qquad \dot{A} = \frac{d}{ds} A_{(s)} \Big|_{s=0}$$

so that relative to a coordinate system x^α we have $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and therefore $\dot{F}_{\mu\nu} = \partial_\mu \dot{A}_\nu - \partial_\nu \dot{A}_\mu$. The action principle gives

$$\begin{aligned} 0 &= \frac{d}{ds} \mathcal{S}(s) \Big|_{s=0} = -\frac{1}{2} \int_{\mathbf{M}} \dot{F}_{\mu\nu} F^{\mu\nu} dv_{\mathbf{g}} \\ &= -\frac{1}{2} \int_{\mathcal{U}} (\partial_\mu \dot{A}_\nu - \partial_\nu \dot{A}_\mu) F^{\mu\nu} dv_{\mathbf{g}} \\ &= -\int_{\mathcal{U}} \partial_\mu \dot{A}_\nu F^{\mu\nu} dv_{\mathbf{g}} = \int_{\mathcal{U}} \dot{A}_\nu \left(\frac{1}{\sqrt{-\mathbf{g}}} \partial_\nu \left(\sqrt{-\mathbf{g}} F^{\mu\nu} \right) \right) dv_{\mathbf{g}} \end{aligned}$$

Note that the second factor in the integrand is just $\mathbf{D}_\mu F^{\mu\nu}$ where $\mathbf{D}$ is the covariant derivative on $\mathbf{M}$ corresponding to $\mathbf{g}$. Hence the Euler-Lagrange equations take the form

$$\mathbf{D}_\nu F^{\mu\nu} = 0. \tag{343}$$

Together, (341) and (343) constitute the Maxwell equations.

Exercise. Given a vector field X^α on $\mathbf{M}$, show

$$\mathbf{D}_\alpha X^\alpha = \frac{1}{\sqrt{-\mathbf{g}}}\partial_\alpha\left(\sqrt{-\mathbf{g}}X^\alpha\right)$$

We can write the Maxwell equations in a more symmetric form by using the Hodge dual of F,

$${}^\star F_{\mu\nu} = \frac{1}{2}\in_{\mu\nu\alpha\beta} F^{\alpha\beta}$$

and by noticing that (343) is equivalent to $d\,{}^\star F = 0$. The Maxwell equations then take the form

$$dF = 0, \qquad d\,{}^\star F = 0 \tag{344}$$

or, equivalently,

$$\mathbf{D}_\nu F^{\mu\nu} = 0, \qquad \mathbf{D}_\nu\,{}^\star F^{\mu\nu} = 0 \tag{345}$$

Note that since Lorentz transformations commute with both the Hodge dual and exterior differentiation, the Lorentz invariance of the Maxwell equations is explicit in (344).

Definition. Given X an arbitrary vector field, we can define the contractions

$$\begin{aligned} E_\alpha &= (i_X F)_\alpha &&= X^\mu F_{\alpha\mu} \\ H_\alpha &= (i_X\,{}^\star F)_\alpha &&= X^\mu\,{}^\star F_{\alpha\mu} \end{aligned}$$

called, respectively, the electric and magnetic components of F. Note that both these one-forms are perpendicular to X.

We specialize to the case when $\mathbf{M}$ is the Minkowski space and $X = \frac{d}{dx^0} = \frac{d}{dt}$. As remarked, E, H are perpendicular to $\frac{d}{dt}$, so $E_0 = H_0 = 0$. The spatial components are by definition

$$\begin{aligned} E_i &= F_{0i} \\ H_i &= {}^\star F_{0i} = \frac{1}{2}\in_{0ijk} F^{jk} = \frac{1}{2}\in_{ijk} F^{jk} \end{aligned}$$

We now use (344) to derive equations for E and H from above, which imply

$$\mathbf{D}_\nu\,{}^\star F^{\mu\nu} = 0 \tag{346}$$

and (343), respectively. Setting $\mu = 0$ in both equations of (345) we derive,

$$\partial^i E_i = 0, \qquad \partial^i H_i = 0 \tag{347}$$

Setting $\mu = i$ and observing that $F_{ij} = \in_{ijk} H^k$, $\quad {}^\star F_{ij} = -\in_{ijk} E^k$ we write

$$\begin{aligned} 0 &= -\partial^0 E_i + \partial^j F_{ij} = \partial_0 E_i + \in_{ijk}\partial^j H^k = \partial_t E_i + (\nabla\times H)_i \\ 0 &= \partial_t H_i - \in_{ijk}\partial_j E_k = \partial_t H_i - (\nabla\times E)_i \end{aligned}$$

Therefore,

$$\begin{aligned} \partial_t E + \nabla \times H &= 0 && (348) \\ \partial_t H - \nabla \times E &= 0 && (349) \end{aligned}$$

Alongside (348) and (349) we can assign data at time $t = 0$,

$$E_i(0, x) = E_i^{(0)}, \qquad H_i(0, x) = H_i^{(0)}$$

Exercise. Show that the equations (347) are preserved by the time evolution of the system (348)-(349). In other words if $E^{(0)}, H^{(0)}$ satisfy (347) then they are satisfied by E, H for all times $t \in \mathbb{R}$.

4. Yang-Mills equations :

The Lagrangians of all classical field theories exhibit the symmetries of the space-time. In addition to these space-time symmetries a Lagrangian can have symmetries called internal symmetries of the field. A simple example is the complex scalar Lagrangian,

$$L = -\frac{1}{2}\mathbf{m}^{\alpha\beta}\partial_\alpha\phi\overline{\partial_\beta\phi} - V(|\phi|)$$

where ϕ is a complex valued scalar defined on the Minkowski space-time $\mathbb{R}^{n+1}$, $\bar{\phi}$ its complex conjugate. We note that L is invariant under the transformations $\phi \to e^{i\theta}\phi$ with θ a fixed real number. It is natural to ask whether the Lagrangian can be modified to allow more general, *local* phase transformations of the form $\phi(x) \to e^{i\theta(x)}\phi(x)$. It is easy to see that under such transformations, the Lagrangian fails to be invariant, due to the term $\mathbf{m}^{\alpha\beta}\partial_\alpha\phi\overline{\partial_\beta\phi}$. To obtain an invariant Lagrangian one replaces the derivatives $\partial_\alpha\phi$ by the covariant derivatives $D_\alpha^{(A)}\phi \equiv \phi_{,\alpha} + iA_\alpha\phi$ depending on a gauge potential A_α. We can now easily check that the new Lagrangian

$$L = -\frac{1}{2}\mathbf{m}^{\alpha\beta}D_\alpha^{(A)}\phi\bar{D}_\beta^{(A)}\phi - V(|\phi|)$$

is invariant relative to the local transformations,

$$\phi(x^\alpha) \to e^{i\theta(x)}\phi(x^\alpha) \quad , \quad A_\alpha \to A_\alpha - \theta_{,\alpha}$$

called gauge transformations.

Remark that the gauge transformations introduced above fit well with the definition of the electromagnetic field F. Indeed setting $F = dA$ we notice that F is invariant. This allows us to consider a more general Lagrangian which includes F,

$$L = -\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} - \frac{1}{2}m^{\alpha\beta}\phi_{,\alpha}\bar{\phi}_{,\beta} - V(|\phi|)$$

called the *Maxwell-Klein-Gordon* Lagrangian.

The Yang-Mills Lagrangian is a natural generalization of the Maxwell-Klein-Gordon Lagrangian to the case when the group $SU(1)$, corresponding to the phase transformations of the complex scalar ϕ, is replaced by a more general Lie group G. In this case the role of the *gauge potential* or *connection 1-form* is taken by a $\mathcal{G}$ valued one form $A = A_\mu dx^\mu$ defined on $\mathbf{M}$. Here $\mathcal{G}$ is the Lie algebra of the Lie group G.

Let $[\cdot,\cdot]$ its Lie bracket and $<\cdot,\cdot>$ its Killing scalar product. Typically the Lie group G is one of the classical groups of matrices, i.e. a subroup of either $\mathrm{Mat}(n,\mathbb{R})$ or $\mathrm{Mat}(n,\mathbb{C})$. We pause briefly to recall some facts about the relavent Lie groups and their Lie algebras.

(1) The orthogonal groups $\mathbf{O}(p,q)$. These are the groups of linear transformations of Re^n which preserve a given nondegenerate symmetric bilinear form of signature p,q, $p+q=n$. We denote by $\mathbf{R}^n_{p,q}$ the corresponding space. The case $p=0$ is that of the Euclidean case, the group is then simply denoted by $\mathbf{O}(n)$. The case $p=1, q=n$ is that of the Minkowski space-time $\mathbb{R}^{n+1}$, the group $\mathbf{O}(1,n)$ is the Lorentz group. In general let Q be the diagonal matrix whose first p diagonal elements are -1 and the remaining ones are $+1$. Then,

$$\begin{aligned}\mathbf{O}(p,q) &= \{L \in \mathrm{M}at(n,\mathbb{R}) | L^T Q L = Q\} \\ &= \{L \in \ \mathrm{M}at(n,\mathbb{R}) | LML^T = M\}\end{aligned}$$

Note that for $L \in \mathbf{O}(p,q)$, $\det(L) = \pm 1$.

Recall that the special orthogonal groups $\mathbf{SO}(p,q)$ are defined by

$$\mathbf{SO}(p,q) = \{L \in \ \mathbf{O}(p,q) | \det\, L = 1\}.$$

They correspond to all orientation preserving isometries of $\mathbf{R}^n_{p,q}$. Both $\mathbf{O}(p,q)$ and $\mathbf{SO}(p,q)$ have as Lie algebra[3]

$$\mathcal{SO}(p,q) = \{A \in \mathrm{M}at(n,\mathbb{R}) | AQ + QA^T = 0\}.$$

and that $\dim_{\mathbf{R}}\mathbf{O}(p,q) = \dim_{\mathbf{R}}\mathbf{SO}(p,q) = n(n-1)/2$. The Lie bracket on $\mathcal{SO}(p,q)$ is the usual Lie bracket of matrices,

i.e. $[A,B] = AB - BA$ and we have the Jacobi identity

$$[A,[B,C]] + [C,[A,B]] + [B,[C,A]] = 0 \tag{350}$$

and its Killing scalar product $<A,B>= -\mathrm{Tr}(AB^T)$ (where Tr is the usual trace for matrices) enjoys the compatibility condition

$$<A,[B,C]>= -<[A,B],C> \tag{351}$$

(2) The unitary groups $\mathbf{U}(p,q)$. These are the complex analogues of the orthogonal groups. They are the groups of all linear transformations of $\mathbb{C}^n$ which preserve a given nondegenerate hermitian bilinear form. Denote by $\mathbb{C}^n_{p,q}$ the corresponding space. Then, with the matrix Q as above,

$$\mathbf{U}(p,q) = \{U \in \mathrm{M}at(n,\mathbb{C}) \mid U^* Q U = Q\}$$

and,

$$\mathbf{SU}(p,q) = \{U \in \mathbf{U}(p,q) \mid \det U = 1\},$$

The corresponding Lie algebras are,

$$\begin{aligned}\mathcal{U}(p,q) &= \{A \in \mathrm{M}at(n,\mathbb{C}) \mid AQ + QA^* = 0\}, \\ \mathcal{SU}(p,q) &= \{A \in \mathcal{U}(p,q) \mid \mathrm{tr}_M A = 0\},\end{aligned}$$

where the trace $\mathrm{tr}_Q A = Q^{ij}A_{ij}$. The Lie bracket is again the usual one for matrices. The Killing scalar product is given by $<A,B>= -\mathrm{Tr}(AB^*)$. Remark also that $\dim_{\mathbf{R}}\mathbf{U}(p,q) = n^2$, $\dim_{\mathbf{R}}\mathbf{SU}(p,q) = n^2-1$.

---

[3]Recall that the Lie algebra of a Lie group G is simply the tangent space to G at the origin.

In the Yang-Mills theory one is interested in compact Lie groups with a positive definite Killing form. This is the case for the groups $O(n), SO(n), U(n), SU(n)$.

In a given system of coordinates the connection 1-form A has the form, $A_\mu dx^\mu$ and we define the (gauge) covariant derivative of a $\mathcal{G}$-valued tensor ψ by

$$\mathbf{D}_\mu^{(A)}\psi = \mathbf{D}_\mu\psi + [A_\mu, \psi] \tag{352}$$

where $\mathbf{D}$ is the covariant derivative on $\mathbf{M}$. Observe that (352) is invariant under the following *gauge transformations*, for a given $\mathcal{G}$-valued gauge potential A and a $\mathcal{G}$- valued tensor ψ,

$$\tilde{\psi} \quad = \quad U^{-1}\psi U, \qquad \tilde{A}_\alpha = U^{-1}A_\alpha U + \left(\mathbf{D}_\alpha U^{-1}\right)U \tag{353}$$

with $U \in G$.

PROPOSITION 1.1.

$$\begin{aligned}\mathbf{D}_\mu^{(\tilde{A})}\tilde{\psi} &= U^{-1}\left(\mathbf{D}_\mu^{(A)}\psi\right)U \\ &= \widetilde{\mathbf{D}^A\psi}\end{aligned}$$

Proof : This just requires some patience. First we will show

$$\mathbf{D}_\alpha\left(U^{-1}\psi U\right) = U^{-1}\left(\mathbf{D}_\alpha\psi + [\psi, U(\mathbf{D}_\alpha U^{-1})]\right)U$$

Indeed

$$\begin{aligned}\mathbf{D}_\alpha\left(U^{-1}\psi U\right) &= \left(\mathbf{D}_\alpha U^{-1}\right)\psi U + U^{-1}\left(\mathbf{D}_\alpha\psi\right)U + U^{-1}\psi\left(\mathbf{D}_\alpha U\right) \\ &= U^{-1}\left(-\left(\mathbf{D}_\alpha U\right)U^{-1}\psi + \mathbf{D}_\alpha\psi + \psi\left(\mathbf{D}_\alpha U\right)U^{-1}\right)U \\ &= U^{-1}\left(\mathbf{D}_\alpha\psi + [\psi, \left(\mathbf{D}_\alpha U\right)U^{-1}]\right)U\end{aligned}$$

as desired. Hence

$$\begin{aligned}\mathbf{D}_\alpha^{(\tilde{A})}\tilde{\psi} &= \mathbf{D}_\alpha\tilde{\psi} + [\tilde{A}_\alpha, \tilde{\psi}] \\ &= U^{-1}\left(\mathbf{D}_\alpha\psi + [\psi, U\left(\mathbf{D}_\alpha U^{-1}\right)]\right) + \left[U^{-1}A_\alpha U + \left(\mathbf{D}_\alpha U^{-1}\right)U, U^{-1}\psi U\right] \\ &= U^{-1}\left(\mathbf{D}_\alpha\psi + [\psi, \left(\mathbf{D}_\alpha U\right)U^{-1}] + [A_\alpha, \psi] + \left[U\left(\mathbf{D}_\alpha U^{-1}\right), \psi\right]\right)U \\ &= U^{-1}\left(\mathbf{D}_\alpha\psi + [A_\alpha, \psi]\right)U = \widetilde{\mathbf{D}_\alpha^{(A)}\psi}\end{aligned}$$

■

As in Riemmanian geometry, commuting two (gauge) covariant derivatives produces a fundamental object called the curvature, here denoted by F

$$\mathbf{D}_\alpha\mathbf{D}_\beta\psi - \mathbf{D}_\beta\mathbf{D}_\alpha\psi = [F_{\alpha\beta}, \psi] \tag{354}$$

where the components $F_{\alpha\beta}$ of the curvature can be deduced by the following straightforward computation:

$$\begin{aligned}\mathbf{D}_\alpha\mathbf{D}_\beta\psi &= \mathbf{D}_\alpha\left(\mathbf{D}_\beta\psi\right) + [A_\alpha, \mathbf{D}_\beta\psi] \\ &= \mathbf{D}_\alpha\left(\mathbf{D}_\beta\psi + [A_\beta, \psi]\right) + [A_\alpha, \mathbf{D}_\beta\psi + [A_\beta, \psi]] \\ &= \mathbf{D}_\alpha\mathbf{D}_\beta\psi + [\mathbf{D}_\alpha A_\beta, \psi] + [A_\beta, \mathbf{D}_\alpha\psi] + [A_\alpha, \mathbf{D}_\beta\psi] + [A_\alpha, [A_\beta, \psi]]\end{aligned}$$

So that

$$\begin{aligned} (\mathbf{D}_\alpha \mathbf{D}_\beta - \mathbf{D}_\beta \mathbf{D}_\alpha)\,\psi &= [\mathbf{D}_\alpha A_\beta - \mathbf{D}_\beta A_\alpha, \psi] \\ &\quad + \underbrace{[A_\alpha, [A_\beta, \psi]] - [A_\beta, [A_\alpha, \psi]]}_{[[A_\alpha, A_\beta], \psi]} \end{aligned}$$

Therefore,

$$F_{\alpha\beta} = \mathbf{D}_\alpha A_\beta - \mathbf{D}_\beta A_\alpha + [A_\alpha, A_\beta] \tag{355}$$

We leave it to the reader to show that the curvature tensor F is invariant under gauge transformations. That is,

$$\widetilde{F^{(\tilde{A})}} \; \left(\equiv U^{-1} F^{(\tilde{A})} U\right) \; = F^{(A)}$$

and that F satisfies the Bianchi identity

$$\mathbf{D}_\alpha F_{\beta\gamma} + \mathbf{D}_\gamma F_{\alpha\beta} + \mathbf{D}_\beta F_{\gamma\alpha} \; = 0 \tag{356}$$

We are finally ready to present the generalization of the Maxwell theory provided by the Yang-Mills Lagrangian:

$$L[A] = -\frac{1}{4} < F^{(A)}_{\alpha\beta}, F^{(A)\,\alpha\beta} >_{\mathcal{G}} \tag{357}$$

We derive the Euler-Lagrange equations just as in the Maxwell theory,

$$\begin{aligned} 0 &= \frac{d}{ds}\mathbf{S}(s)\Big|_{s=0} = -\frac{1}{2}\int_{\mathcal{U}} < \dot{F}_{\alpha\beta}, F^{\alpha\beta} >_{\mathcal{G}} dv_{\mathbf{g}} \\ &= -\frac{1}{2}\int_{\mathcal{U}} < \mathbf{D}_\alpha \dot{A}_\beta - \mathbf{D}_\beta \dot{A}_\alpha + [\dot{A}_\alpha, A_\beta] + [A_\alpha, \dot{A}_\beta], F^{\alpha\beta} >_{\mathcal{G}} dv_{\mathbf{g}} \\ &= -\int_{\mathcal{U}} < \mathbf{D}_\alpha \dot{A}_\beta, F^{\alpha\beta} > + < [A_\alpha, \dot{A}_\beta], F^{\alpha\beta} >_{\mathcal{G}} dv_{\mathbf{g}} \\ &= \int_{\mathcal{U}} < \dot{A}_\beta, \mathbf{D}_\alpha F^{\alpha\beta} >_{\mathcal{G}} + < \dot{A}_\beta, \left[A_\alpha, F^{\alpha\beta}\right] >_{\mathcal{G}} dv_{\mathbf{g}} \end{aligned}$$

which implies

$$\mathbf{D}_\nu F^{\mu\nu} = 0 \tag{358}$$

Together, (356) and (358) form the Yang-Mills equations.

Note that the equations are invariant under the group of gauge transformations. A solution of the Yang-Mills equations, then, is an equivalence class of gauge-equivalent potentials A_α whose curvature F satisfies (358).

In our later treatment of Yang-Mills, we will almost always specify a representative of a solution's equivalence class by imposing additional constraints - called *gauge conditions* - on A. There are three standard ways of doing this, each yielding its own rendition of the Yang-Mills equations with its own faults and advantages:

- **Coulomb Gauge** is defined by,

$$\nabla^i A_i(t,x) = 0 \;\; (t,x) \in \mathbb{R}^{n+1} \tag{359}$$

To simplify notation, first write (358) in terms of the current J.

$$\mathbf{D}^{\beta}F_{\alpha\beta} = J_{\alpha} = -\left[A^{\beta}, F_{\alpha,\beta}\right] \tag{360}$$

When $\alpha = 0$ (359) allows us to write (360) as

$$J_0 \quad = \quad \partial^i F_{0i} = \partial^i\left(\partial_0 A_i - \partial_i A_0 + [A_0, A_i]\right) - \Delta A_0 + \partial^i[A_0, A_i]$$

giving us for the time component of A:

$$\Delta A_0 = 2\left[\partial_i A_0, A_i\right] + \left[A_0, \partial_t A_i\right] + \left[A_i, \left[A_0, A_j\right]\right] \tag{361}$$

When $\alpha = i$, (360) reads

$$J_i \quad = \quad -\partial_t + \partial^j F_{ij} = -\partial_t\left(\partial_i A_0 + [A_i, A_0]\right) + \partial^j\left(\partial_i A_j - \partial_j A_i + [A_i, A_j]\right)$$

and after simplifying,

$$\begin{aligned}\Box A_i \quad = \quad & -\partial_t\partial_i A_0 - 2\left[A_j, \partial_j A_i\right] + \left[A_j, \partial_i A_j\right] + \left[\partial_t A_i, A_j\right] \\ & +2\left[A_0, \partial_t A_i\right] - \left[A_0, \partial_i A_0\right] - \left[A_j, \left[A_j, A_i\right]\right] + \left[A_0, \left[A_0, A_i\right]\right]\end{aligned} \tag{362}$$

- **Lorentz Gauge** is specified by,

$$\partial^{\mu}A_{\mu}(t,x) = 0 \quad (t,x) \in \mathbb{R}^{3+1} \tag{363}$$

Appealing in its symmetric treatment of the time and space components of A, , the Lorentz gauge also allows (358) to be written as a system of wave equations:

$$\begin{aligned}\mathbf{D}^{\beta}F_{\alpha\beta} \quad &= \quad \mathbf{D}^{\beta}\left(\partial_{\alpha}A_{\beta} - \partial_{\beta}A_{\alpha} + [A_{\alpha}, A_{\beta}]\right) \\ &= \quad -\Box A_{\alpha} + \partial^{\beta}[A_{\alpha}, A_{\beta}] + [A_{\beta}, \partial_{\alpha}A_{\beta}] - [A^{\beta}, \partial_{\beta}A_{\alpha}] + \left[A_{\beta}, [A_{\alpha}, A^{\beta}]\right]\end{aligned}$$

The system can be written schematically in the form

$$\Box\Phi \quad = \quad \Phi \cdot \partial\Phi + \Phi^3$$

Again, it is not at all clear that one can transform an arbitrary solution into the Lorentz gauge. In addition, we will have a hard time finding good estimates for this purely hyperbolic system of nonlinear wave equations.

- **Temporal Gauge** is specified by the condition $A_0 = 0$.

5. The Einstein Field Equations:

According to the general relativistic variational principle the space-time metric $\mathbf{g}$ is itself stationary relative to an action,

$$\mathcal{S} = \int_{\mathcal{U}} L dv_{\mathbf{g}}.$$

Here U is a relatively compact domain of $(\mathbf{M}, \mathbf{g})$ and L, the Lagrangian, is assumed to be a scalar function on $\mathbf{M}$ whose dependence on the metric should involve no more than two derivatives[4]. It is also assumed to depend on the matterfields $\psi = \psi^{(1)}, \psi^{(2)}, \ldots \psi^{(p)}$ present in our space-time.

[4] In fact we only require that the corrsponding Euler-Lagrange equations should involve no more than two derivatives of the metric.

In fact we write,

$$\mathcal{S} = \mathcal{S}_G + \mathcal{S}_M$$

with,

$$\begin{aligned}\mathcal{S}_G &= \int_{\mathcal{U}} L_G dv_{\mathbf{g}} \\ \mathcal{S}_M &= \int_{\mathcal{U}} L_M dv_{\mathbf{g}}\end{aligned}$$

denoting, respectively, the actions for the gravitational field and matter. The matter Lagrangian L_M depends only on the matterfields ψ, assumed to be covariant tensorfields, and the inverse of the space-time metric $\mathbf{g}^{\alpha\beta}$ which appears in the contraction of the tensorfields ψ in order to produce the scalar L_M. It may also depend on additional positive definite metrics which are not to be varied [5].

Now the only possible candidate for the gravitational Lagrangian L_G, which should be a scalar invariant of the metric with the property that the corresponding Euler-Lagrange equations involve at most two derivatives of the metric, is given[6] by the scalar curvature $\mathbf{R}$. Therefore we set,

$$L_G = \mathbf{R}.$$

Consider now a compact variation $(\mathbf{g}_{(s)}, \mathcal{U})$ of the metric $\mathbf{g}$. Let $\dot{\mathbf{g}}_{\mu\nu} = \frac{d}{ds}\mathbf{g}_{\mu\nu}|_{s=0}$. Thus for small s, $\mathbf{g}_{\mu\nu}(s) = \mathbf{g}_{\mu\nu} + s\dot{\mathbf{g}}_{\mu\nu} + O(s^2)$. Also, $\mathbf{g}^{\mu\nu}(s) = \mathbf{g}^{\mu\nu} - s\dot{\mathbf{g}}^{\mu\nu} + O(s^2)$ where $\dot{\mathbf{g}}^{\mu\nu} = \mathbf{g}^{\alpha\mu}\mathbf{g}^{\beta\nu}\dot{\mathbf{g}}_{\alpha\beta}$. Then,

$$\frac{d}{ds}\mathbf{S}_G(s)\Big|_{s=0} = \int_{\mathcal{U}} \dot{\mathbf{R}} dv_{\mathbf{g}} + \int_{\mathcal{U}} \mathbf{R}\dot{dv_{\mathbf{g}}}$$

Now,

$$\dot{dv_{\mathbf{g}}} = \frac{1}{2}\mathbf{g}^{\mu\nu}\dot{\mathbf{g}}_{\mu\nu} dv_{\mathbf{g}}$$

Indeed, relative to a coordinate system, $dv_{\mathbf{g}} = \sqrt{-\mathbf{g}}dx^0 dx^1 \ldots dx^n$ Thus, the above equality follows from,

$$\dot{\mathbf{g}} = \mathbf{g}\mathbf{g}^{\alpha\beta}\dot{\mathbf{g}}_{\alpha\beta},$$

with $\mathbf{g}$ the determinant of $\mathbf{g}_{\alpha\beta}$. On the other hand, writing $\mathbf{R} = \mathbf{g}^{\mu\nu}\mathbf{R}_{\mu\nu}$ and using the formula $\frac{d}{ds}\mathbf{g}^{\mu\nu}_{(s)}\Big|_{s=0} = -\dot{\mathbf{g}}^{\mu\nu}$, we calculate, $\dot{\mathbf{R}} = -\dot{\mathbf{g}}^{\mu\nu}\mathbf{R}_{\mu\nu} + \mathbf{g}^{\mu\nu}\dot{\mathbf{R}}_{\mu\nu}$. Therefore,

$$\frac{d}{ds}\mathbf{S}_G(s)\Big|_{s=0} = -\int_{\mathcal{U}} (\mathbf{R}^{\mu\nu} - \frac{1}{2}\mathbf{g}^{\mu\nu}\mathbf{R})\dot{\mathbf{g}}_{\mu\nu} dv_{\mathbf{g}} + \int_{\mathcal{U}} \mathbf{g}^{\mu\nu}\dot{\mathbf{R}}_{\mu\nu} dv_{\mathbf{g}} \tag{364}$$

To calculate $\dot{\mathbf{R}}_{\mu\nu}$ we make use of the following Lemma,

LEMMA 1.2. *Let $\mathbf{g}_{\mu\nu}(s)$ be a family of space-time metrics with $\mathbf{g}(0) = \mathbf{g}$ and $\frac{d}{ds}\mathbf{g}(0) = \dot{\mathbf{g}}$. Set also, $\frac{d}{ds}\mathbf{R}_{\alpha\beta}(s)|_{s=0} = \dot{\mathbf{R}}_{\alpha\beta}$. Then,*

$$\dot{\mathbf{R}}_{\mu\nu} = \mathbf{D}_\alpha \dot{\Gamma}^\alpha_{\mu\nu} - \mathbf{D}_\mu \dot{\Gamma}^\alpha_{\alpha\nu}$$

[5] This is the case of the metric h in the case of wave maps or the Killing scalar product in the case of the Yang-Mills equations.

[6] up to an additive constant

where $\dot{\Gamma}$ *is the tensor,*

$$\dot{\Gamma}^{\alpha}_{\beta\gamma} = \frac{1}{2}\mathbf{g}^{\alpha\lambda}(\mathbf{D}_\beta\dot{\mathbf{g}}_{\gamma\lambda} + \mathbf{D}_\gamma\dot{\mathbf{g}}_{\beta\lambda} - \mathbf{D}_\lambda\dot{\mathbf{g}}_{\beta\gamma})$$

Proof : Since both sides of the identity are tensors it suffices to prove the formula at a point p relative to a particular system of coordinates for which the Christoffel symbols Γ vanish at p. Relative to such a coordinate system the Ricci tensor has the form $\mathbf{R}_{\mu\nu} = \mathbf{D}_\alpha\Gamma^{\alpha}_{\mu\nu} - \mathbf{D}_\mu\Gamma^{\alpha}_{\alpha\nu}$. ∎

Returning to (364) we find that since $\mathbf{g}^{\mu\nu}\dot{\mathbf{R}}_{\mu\nu}$ can be written as a space-time divergence of a tensor compactly supported in U the corresponding integral vanishes identically. We therefore infer that,

$$\frac{d}{ds}\mathbf{S}_G(s)\Big|_{s=0} = -\int_U \mathbf{E}^{\mu\nu}\dot{\mathbf{g}}_{\mu\nu}dv_{\mathbf{g}} \tag{365}$$

where $\mathbf{E}^{\mu\nu} = \mathbf{R}^{\mu\nu} - \frac{1}{2}\mathbf{g}^{\mu\nu}\mathbf{R}$. We now consider the variation of the action integral $\mathbf{S}_M$ with respect to the metric. As remarked before L_M depends on the metric $\mathbf{g}$ through its inverse $\mathbf{g}^{\mu\nu}$. Therefore if we denote $\mathbf{S}_M(s) = \mathbf{S}_M[\psi, \mathbf{g}_{(s)}; \mathcal{U}]$ we have, writing $d\dot{v}_{\mathbf{g}} = \frac{1}{2}\mathbf{g}_{\mu\nu}\dot{\mathbf{g}}^{\mu\nu}dv_{\mathbf{g}}$,

$$\begin{aligned}\frac{d}{ds}\mathbf{S}_M(s)\Big|_{s=0} &= -\int_{\mathcal{U}} \frac{\partial L_M}{\partial \mathbf{g}^{\mu\nu}}\dot{\mathbf{g}}^{\mu\nu}dv_{\mathbf{g}} + \int_{\mathcal{U}} L_M d\dot{v}_{\mathbf{g}} \\ &= -\int_{\mathcal{U}} (\frac{\partial L_M}{\partial \mathbf{g}^{\mu\nu}} - \frac{1}{2}\mathbf{g}_{\mu\nu}L_M)\dot{\mathbf{g}}^{\mu\nu}dv_{\mathbf{g}}\end{aligned}$$

Definition. The symmetric tensor,

$$\mathbf{T}_{\mu\nu} = -\left(\frac{\partial L_M}{\partial \mathbf{g}^{\mu\nu}} - \frac{1}{2}\mathbf{g}_{\mu\nu}L_M\right)$$

is called the *energy-momentum tensor* of the action $\mathbf{S}_M$.

With this definition we write,

$$\frac{d}{ds}\mathbf{S}_M(s)\Big|_{s=0} = \int_{\mathcal{U}} \mathbf{T}^{\mu\nu}\dot{\mathbf{g}}_{\mu\nu}dv_{\mathbf{g}} \tag{366}$$

Finally, combining 365 with 366, we derive for the total action $\mathbf{S}$,

$$\frac{d}{ds}\mathbf{S}(s)\Big|_{s=0} = -\int_{\mathcal{U}} (\mathbf{E}^{\mu\nu} - \mathbf{T}^{\mu\nu})\dot{\mathbf{g}}_{\mu\nu}dv_{\mathbf{g}}$$

Since $\dot{\mathbf{g}}_{\mu\nu}$ is an arbitrary symmetric 2-tensor compactly supported in U we derive the *Einstein field equation*,

$$\mathbf{E}^{\mu\nu} = \mathbf{T}^{\mu\nu}$$

Recall that the Einstein tensor $\mathbf{E}$ satisfies the twice contracted Bianchi identity,

$$\mathbf{D}^\nu\mathbf{E}_{\mu\nu} = 0$$

This implies that the energy-momentum tensor $\mathbf{T}$ is also divergenceless,

$$\mathbf{D}_\nu \mathbf{T}^{\mu\nu} = 0 \tag{367}$$

which is the concise, space-time expression for the law of conservation of energy-momentum of the matter-fields.

2. The energy-momentum tensor

The conservation law (367) is a fundamental property of a matterfield. We now turn to a more direct derivation.

We consider an arbitrary Lagrangian field theory with stationary solution ψ. Let Φ_s be the one-parameter group of local diffeomorphisms generated by a given vectorfield X. We shall use the flow Φ to vary the fields ψ according to

$$\begin{aligned}\mathbf{g}_s &= (\Phi_s)_* \mathbf{g} \\ \psi_s &= (\Phi_s)_* \psi.\end{aligned}$$

From the invariance of the action integral under diffeomorphisms,

$$\mathbf{S}(s) = \mathbf{S}[\psi_s, \mathbf{g}_s; \mathbf{M}] = \mathbf{S}_M[\psi, \mathbf{g}; \mathbf{M}].$$

So that

$$0 = \frac{d}{ds}\mathbf{S}(s)\Big|_{s=0} = \int_{\mathbf{M}} \frac{\delta \mathbf{S}}{\delta \psi} dv_{\mathbf{g}} + \int_{\mathbf{M}} \mathbf{T}^{\mu\nu} \dot{\mathbf{g}}_{\mu\nu} dv_{\mathbf{g}} \tag{368}$$

The first term is clearly zero, ψ being a stationary solution. In the second term, which represents variations with respect to the metric, we have

$$\dot{\mathbf{g}}_{\mu\nu} = \frac{d}{ds}(\mathbf{g}_s)_{\mu\nu}\Big|_{s=0} = \mathcal{L}_X \mathbf{g}_{\mu\nu} = \mathbf{D}_\mu X_\nu + \mathbf{D}_\nu X_\mu$$

Therefore

$$0 = \int_{\mathbf{M}} \mathbf{T}^{\mu\nu} \mathcal{L}_X \mathbf{g}_{\mu\nu} dv_{\mathbf{g}} = 2\int_{\mathbf{M}} \mathbf{T}^{\mu\nu} \mathbf{D}_\nu X_\mu dv_{\mathbf{g}} = -2\int_{\mathbf{M}} \mathbf{D}_\nu \mathbf{T}^{\mu\nu} X_\mu dv_{\mathbf{g}}$$

As X was arbitrary, we conclude

$$\mathbf{D}_\nu \mathbf{T}^{\mu\nu} = 0. \tag{369}$$

This is again the law of conservation of energy-momentum.

We list below the energy-momentum tensors of the field theories discussed before. We leave it to the reader to carry out the calculations using the definition.

(1) The energy-momentum for the scalar field equation is,

$$\mathbf{T}_{\alpha\beta} = \frac{1}{2}\left(\phi_{,\alpha}\phi_{,\beta} - \frac{1}{2}\mathbf{g}_{\alpha\beta}(\mathbf{g}^{\mu\nu}\phi_{,\mu}\phi_{,\nu} + 2V(\phi))\right)$$

(2) The energy-momentum for wave maps is given by,

$$\mathbf{T}_{\alpha\beta} = \frac{1}{2}\left(< \phi_{,\alpha}, \phi_{,\beta} > -\frac{1}{2}\mathbf{g}_{\alpha\beta}(\mathbf{g}^{\mu\nu} < \phi_{,\mu}, \phi_{,\nu} >)\right)$$

where $< \, , \, >$ denotes the Riemannian inner product on the target manifold.

(3) The energy-momentum tensor for the Maxwell equations is,

$$\mathbf{T}_{\alpha\beta} = F_{\alpha}^{\cdot\mu} F_{\beta\mu} - \frac{1}{4}\mathbf{g}_{\alpha\beta}(F_{\mu\nu}F^{\mu\nu})$$

(4) The energy-momentum tensor for the Yang-Mills equations is,

$$\mathbf{T}_{\alpha\beta} =< F_{\alpha}^{\cdot\mu}, F_{\beta\mu} > -\frac{1}{4}\mathbf{g}_{\alpha\beta}(< F_{\mu\nu}, F^{\mu\nu} >)$$

An acceptable notion of the energy-momentum tensor $\mathbf{T}$ must satisfy the following properties in addition of the conservation law (369),

(1) $\mathbf{T}$ is symmetric
(2) $\mathbf{T}$ satisfies the *positive energy condition* that is, $\mathbf{T}(X, Y) \geq 0$, for any future directed time-like vectors X, Y.

The symmetry property is automatic in our construction. The following proposition asserts that the energy-momentum tensors of the field theories described above satisfy the positive energy condition.

PROPOSITION 2.1. *The energy-momentum tensor of the scalar wave equation satisfies the positive energy condition if V is positive. The energy- momentum tensors for the wave maps, Maxwell equations and Yang-Mills satisfy the positive energy condition.*

Proof : To prove the positivity conditions consider two vectors X, Y, at some point $p \in \mathbf{M}$, which are both causal future oriented. The plane spanned by X, Y intersects the null cone at p along two null directions[7]. Let $L, \underline{L}$ be the two future directed null vectors corresponding to the two complementary null directions and normalized by the condition

$$< L, \underline{L} >= -2$$

i.e. they form a null pair. Since the vectorfields X, Y are linear combinations with positive coefficients of $L, \underline{L}$, the proposition will follow from showing that $\mathbf{T}(L, L) \geq 0$, $\mathbf{T}(\underline{L}, \underline{L}) \geq 0$ and $\mathbf{T}(L, \underline{L}) \geq 0$. To show this we consider a frame at p formed by the vectorfields $E_{(n+1)} = L$, $E_{(n)} = \underline{L}$ and $E_{(1)}, \ldots, E_{(n-1)}$ with the properties,

$$< E_{(i)}, E_{(n)} >=< E_{(i)}, E_{(n+1)} >= 0$$

and

$$< E_{(i)}, E_{(j)} >= \delta_{ij}$$

for all $i, j = 1, \ldots, n-1$. A frame with these properties is called a *null frame.*

[7]If X, Y are linearly dependent any plane passing through their common direction will do.

(1) We now calculate, in the case of the wave equation,

$$\begin{aligned}\mathbf{T}(L,L) &= \frac{1}{2}E(\phi)^2\\ \mathbf{T}(\underline{L},\underline{L}) &= \frac{1}{2}\underline{L}(\phi)^2.\end{aligned}$$

which are clearly non-negative. Now,

$$\mathbf{T}(L,\underline{L}) = \frac{1}{2}\left[L(\phi)\underline{L}(\phi) + (g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} + 2V(\phi))\right]$$

and we aim to express $g^{\mu\nu}\phi_{,\mu}\phi_{,\nu}$ relative to our null frame. To do this, observe that relative to the null frame the only nonvanishing components of the metric $g_{\alpha\beta}$ are,

$$g_{n(n+1)} = -2 \quad , \quad g_{ii} = 1 \;\; i = 1, \dots, n-1$$

and those of the inverse metric $g^{\alpha\beta}$ are

$$g^{n(n+1)} = -\frac{1}{2} \quad , \quad g^{ii} = 1 \;\; i = 1, \dots, n-1$$

Therefore,

$$g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} = -L(\phi)\underline{L}(\phi) + |\nabla\!\!\!/\phi|^2$$

where

$$|\nabla\!\!\!/\phi|^2 = (E_{(1)}(\phi))^2 + (E_{(2)}(\phi))^2 + \dots E_{(n-1)}(\phi)^2.$$

Therefore,

$$\mathbf{T}(L,\underline{L}) = \frac{1}{2}|\nabla\!\!\!/\phi|^2 + V(\phi).$$

(2) For wave maps we have, according to the same calculation.

$$\begin{aligned}T(E,E) &= \frac{1}{2} < E(\phi), E(\phi) >\\ T(\underline{E},\underline{E}) &= \frac{1}{2} < \underline{E}(\phi), \underline{E}(\phi) >\\ T(E,\underline{E}) &= \frac{1}{2}\sum_{i=1}^{n-1} < E_{(i)}(\phi), E_{(i)}(\phi) > .\end{aligned}$$

The positivity of T is then a consequence of the Riemannian metric h on the target manifold N.

(3) To show positivity for the energy momentum tensor of the Maxwell equations in $3+1$ dimensions we first write the tensor in the more symmetric form

$$T_{\alpha\beta} = \frac{1}{2}\left(F_{\alpha}{}^{\mu}F_{\beta\mu} + {}^\star F_{\alpha}{}^{\mu}\,{}^\star F_{\beta\mu}\right) \tag{370}$$

where ${}^\star F$ is the Hodge dual of F, i.e. ${}^\star F_{\alpha\beta} = \frac{1}{2}\in_{\alpha\beta\mu\nu} F^{\mu\nu}$.

Exercise. Check formula (370).

We introduce the following null decomposition of F at every point $p \in \mathbf{M}$,

$$\begin{aligned}\alpha_A = F_{A4} \quad &, \quad \underline{\alpha}_A = F_{A3}\\ \rho = \frac{1}{2}F_{34} \quad &, \quad \sigma = \frac{1}{2}\,{}^\star F_{34}.\end{aligned}$$

which completely determines the tensor F. Here the indices $A = 1, 2$ correspond to the directions E_1, E_2 tangent to the sphere while the indices $3, 4$ correspond to $E_3 = \underline{L}$ and $E_4 = L$. We then calculate that for ${}^\star F$,

$$
\begin{aligned}
{}^\star F_{A4} = -\,{}^\star\alpha_A = \quad &, \qquad {}^\star F_{A3} = {}^\star\underline{\alpha}_A \\
{}^\star F_{34} = 2\sigma \quad &, \qquad {}^{\star\star}F_{34} = -2\rho
\end{aligned}
$$

where ${}^\star\alpha_A = \in_{AB}\alpha_B$. Here $\in_{AB}$ is the volume form on the unit sphere, hence $\in_{AB} = \frac{1}{2}\in_{AB34}$, i.e. $\in_{11} = \in_{22} = 0$, $\in_{12} = -\epsilon_{21} = 1$. With this notation we calculate,

$$
\begin{aligned}
T(E_{(4)}, E_{(4)}) &= \frac{1}{2}\sum_{A=1}^{2}\left(F_{4A}\cdot F_{4A} + \frac{1}{4}\,{}^\star F_{4A}\cdot{}^\star F_{4A}\right) \\
&= \frac{1}{2}\sum_{A=1}^{2}\left(\alpha_A\cdot\alpha_A + {}^\star\alpha_A\cdot{}^\star\alpha_A\right) \\
&= \sum_{A=1}^{2}\alpha_A\cdot\alpha_A = |\alpha|^2 \geq 0.
\end{aligned}
$$

Similarly,

$$T(E_{(3)}, E_{(3)}) = \sum_{A=1}^{2}\underline{\alpha}_A\cdot\underline{\alpha}_A = |\underline{\alpha}|^2 \geq 0$$

and in the same vein we find

$$T(E, \underline{E}) = \rho^2 + \sigma^2 \geq 0$$

which proves our assertion.

(4) The positivity of the energy-momentum tensor of the Yang- Mills equations is proved in precisely the same manner as for the Maxwell equations, using the positivity of the Killing scalar product $< \cdot, \cdot >_{\mathcal{G}}$.

■

Another important property which the energy momentum tensor of a field theory may satisfy is the trace free condition, that is

$$\mathbf{g}_{\alpha\beta}\mathbf{T}^{\alpha\beta} = 0.$$

It turns out that this condition is satisfied by all field theories which are conformally invariant.

Definition. A field theory is said to be *conformally invariant* if the corresponding action integral is invariant under conformal transformations of the metric

$$\mathbf{g}_{\alpha\beta} \longrightarrow \tilde{\mathbf{g}}_{\alpha\beta} = \Omega\mathbf{g}_{\alpha\beta}$$

Ω a positive smooth function on the space-time.

PROPOSITION 2.2. *The energy momentum tensor* $\mathbf{T}$ *of a conformally invariant field theory is traceless.*

Proof: Consider an arbitrary smooth function f compactly supported in $\mathcal{U} \subset \mathcal{M}$. Consider the following variation of a given metric $\mathbf{g}$,

$$\mathbf{g}_{\mu\nu}(s) = e^{sf}\mathbf{g}_{\mu\nu}.$$

Let $\mathcal{S}(s) = \mathcal{S}_{\mathcal{U}}[\psi, \mathbf{g}(s)]$. In view of the covariance of $\mathcal{S}$ we have $\mathcal{S}(s) = \mathcal{S}(0)$. Hence,

$$0 = \frac{d}{ds}\mathcal{S}(s)|_{s=0} = \int_{\mathcal{U}} T^{\mu\nu}\dot{\mathbf{g}}_{\mu\nu}dv_{\mathbf{g}}$$

where

$$\dot{\mathbf{g}}_{\mu\nu} = \frac{d}{ds}\mathbf{g}_{\mu\nu}(s)\Big|_{s=0} = f\mathbf{g}_{\mu\nu}.$$

Hence, $\int_{\mathcal{U}} (T^{\mu\nu}\mathbf{g}_{\mu\nu})\, f dv_{\mathbf{g}} = 0$ and since f is arbitrary we infer that,

$$trT = g^{\mu\nu}T_{\mu\nu} \equiv 0.$$

$\blacksquare$

We can easily check that the Maxwell and the Yang-Mills equations are conformally invariant in 3×1-dimensions. The wave maps field theory is conformally invariant in dimension $1+1$, i.e. if the space-time $\mathcal{M}$ is two-dimensional[8].

Remark: The action integral of the Maxwell equations, $\mathbf{S} = \int_{\mathcal{U}} F_{\alpha\beta}F^{\alpha\beta}dv_{\mathbf{g}}$ is conformally invariant in any dimension provided that we also scale the electromagnetic field F. Indeed if $\tilde{\mathbf{g}}_{\alpha\beta} = \Omega^2\mathbf{g}_{\alpha\beta}$ then $dv_{\tilde{\mathbf{g}}} = \Omega^{n+1}dv_{\mathbf{g}}$ and if we also set $\tilde{F}_{\alpha\beta} = \Omega^{-\frac{n-3}{2}}F_{\alpha\beta}$ we get

$$\begin{aligned}
\tilde{\mathbf{S}}[\tilde{F}, \tilde{\mathbf{g}}] &= \int \tilde{F}_{\alpha\beta}\tilde{F}_{\gamma\delta}\tilde{\mathbf{g}}^{\alpha\gamma}\tilde{\mathbf{g}}^{\beta\delta}dv_{\tilde{\mathbf{g}}} \\
&= \int F_{\alpha\beta}F_{\gamma\delta}\mathbf{g}^{\alpha\gamma}\mathbf{g}^{\beta\delta}dv_{\mathbf{g}} \\
&= \mathbf{S}[F, \mathbf{g}].
\end{aligned}$$

We finish this section with a simple observation concerning conformal field theories in $1+1$ dimensions. We specialize in fact to the Minkowski space $\mathbb{R}^{1+1}$ and consider the local conservation law, $\partial^\mu \mathbf{T}_{\nu\mu} = 0$. Setting $\nu = 0, 1$ we derive

$$\partial^0\mathbf{T}_{00} + \partial^1\mathbf{T}_{01} = 0, \quad \partial^0\mathbf{T}_{01} + \partial^1\mathbf{T}_{11} = 0 \tag{371}$$

Since the energy-momentum tensor is trace-free, we get $\mathbf{T}_{00} = \mathbf{T}_{11} = A$, say. Set $\mathbf{T}_{01} = \mathbf{T}_{10} = B$. Therefore (**??**) implies that both A and B satisfy the linear homogeneous wave equation;

$$\Box A \; = \; 0 \; = \; \Box B. \tag{372}$$

Using this observation it is is easy to prove that smooth initial data remain smooth for all time.

For example, wave maps are conformally invariant in dimension $1+1$. In this case

$$A \; = \; \mathbf{T}_{00} \; = \; \frac{1}{2}\left(< \partial_t\phi, \partial_t\phi > + < \partial_x\phi, \partial_x\phi >\right),$$

---

[8]Similarly for the linear scalar wave equation

Given data in $C_0^\infty(\mathbb{R})$, (372) implies that the derivatives of ϕ remain smooth for all positive times. This proves global existence.

3. Conservation Laws

The energy-momentum tensor of a field theory is intimately connected with conservations laws. This connection is seen through Noether's principle,

Noether's Principle: *To any one-parameter group of transformations preserving the action there corresponds a conservation law.*

We illustrate this fundamental principle as follows: Let $\mathbf{S} = \mathbf{S}[\psi, \mathbf{g}]$ be the action integral of the fields ψ. Let χ_t be a 1-parameter group of isometries of $\mathbf{M}$, i.e., $(\chi_t)_*\mathbf{g} = \mathbf{g}$. Then

$$\begin{aligned} \mathbf{S}[(\chi_t)_*\psi, \mathbf{g}] &= \mathbf{S}[(\chi_t)_*\psi, (\chi_t)_*\mathbf{g}] \\ &= \mathbf{S}[\psi, \mathbf{g}]. \end{aligned}$$

Thus the action is preserved under $\psi \to (\chi_t)_*\psi$. In view of Noether's Principle we ought to find a conservation law for the corresponding Euler-Lagrange equations[9]. We derive these laws using the Killing vectorfield X which generates χ_t.

We begin with a general calculation involving the energy-momentum tensor $\mathbf{T}$ of ψ and an arbitrary vectorfield X. P the one-form obtained by contracting $\mathbf{T}$ with X.

$$P_\alpha = \mathbf{T}_{\alpha\beta} X^\beta$$

Since $\mathbf{T}$ is symmetric and divergence-free

$$\mathbf{D}^\alpha P_\alpha = (\mathbf{D}^\alpha \mathbf{T}_{\alpha\beta}) X^\beta + \mathbf{T}_{\alpha\beta} \left(\mathbf{D}^\alpha X^\beta\right) = \frac{1}{2}\mathbf{T}^{\alpha\beta}\ {}^{(X)}\pi_{\alpha\beta}$$

where ${}^{(X)}\pi_{\alpha\beta}$ is the deformation tensor of X.

$${}^{(X)}\pi_{\alpha\beta} = (\mathcal{L}_X \mathbf{g})_{\alpha\beta} = \mathbf{D}_\alpha X_\beta + \mathbf{D}_\beta X_\alpha$$

Notation. We denote the backward light cone with vertex $p = (\bar{t}, \bar{x}) \in \mathbb{R}^{n+1}$ by

$$\mathcal{N}^-(\bar{t}, \bar{x}) = \{(t, x) \,\big|\, 0 \le t \le \bar{t}; |x - \bar{x}| = \bar{t} - t\}.$$

The restriction of this set to some time interval $[t_1, t_2]$, $t_1 \le t_2 \le \bar{t}$, will be written $\mathcal{N}^-_{[t_1,t_2]}(\bar{t}, \bar{x})$. These null hypersurfaces are null boundaries of,

$$\begin{aligned} \mathcal{J}^{-1}(\bar{t}, \bar{x}) &= \{(t, x) \,\big|\, 0 \le t \le \bar{t};\ |x - \bar{x}| \le \bar{t} - t\} \\ \mathcal{J}^-_{[t_2,t_1]}(\bar{t}, \bar{x}) &= \{(t, x) \,\big|\, t_2 \le t \le t_1;\ |x - \bar{x}| \le \bar{t} - t\} \end{aligned}$$

We shall denote by $S_t = S_t(\bar{t}, \bar{x})$ and $B_t = B_t(\bar{t}, \bar{x})$ the intersection of the time slice Σ_t with $\mathcal{N}^-$, respectively $\mathcal{J}^-$.

[9] The same argument holds for conformal isometries acting on a conformally invariant field theory. We therefore also expect conservation laws in such a setting.

At each point $q = (t, x)$ along $\mathcal{N}^-(p)$, we define the null pair (E_+, E_-) of future oriented null vectors

$$\underline{L} = E_+ \quad = \quad \partial_t + \frac{x^i - \bar{x}^i}{|x - \bar{x}|}\partial_i, \qquad L = E_- = \partial_t - \frac{x^i - \bar{x}^i}{|x - \bar{x}|}\partial_i$$

Observe that both $L, \underline{L}$ are null and $< L, \underline{L} >= -2$.

The following is a simple consequence of Stoke's theorem, in the following form.

PROPOSITION 3.1. *Let P_μ be a one-form satisfying $\partial^\mu P_\mu = F$. Then*[10]*, for all $t_1 \le t_2 \le \bar{t}$,*

$$\int_{B_{t_2}} \langle P, \partial_t\rangle + \int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \langle P, E_-\rangle \quad = \quad \int_{B_{t_1}} \langle P, \partial_t\rangle - \int_{\mathcal{J}^-_{[t_1,t_2]}(p)} F dt dx \quad (373)$$

where,

$$\int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \langle P, E_-\rangle = \int_{t_1}^{t_2} dt \int_{S_t} \langle P, E_-\rangle \, da_t.$$

Applying this proposition to Stoke's theorem to (373) we get

THEOREM 3.2. *Let T be the energy-momentum tensor associated to a field theory and X an arbitrary vector field. Then*

$$\begin{aligned}\int_{B_{t_2}} \mathbf{T}(\partial_t, X) + \int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \mathbf{T}(E_-, X) \quad &= \quad \int_{B_{t_1}} \mathbf{T}(\partial_t, X) \\ &- \quad \int_{\mathcal{J}^-_{[t_1,t_2]}(p)} \mathbf{T}^{\alpha\beta\ (X)}\pi_{\alpha\beta} dt dx\end{aligned} \quad (374)$$

In the particular case when X is Killing, its deformation tensor π vanishes identically. Thus,

COROLLARY 3.3. *If X is a killing vectorfield,*

$$\int_{B_{t_2}} \mathbf{T}(\partial_t, X) + \int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \mathbf{T}(L, X) = \int_{B_{t_1}} \mathbf{T}(\partial_t, X) \quad (375)$$

Moreover (375) *remains valid if* $\mathbf{T}$ *is traceless and X is conformal Killing.*

The identity (375) is usually applied to time-like future-oriented Killing vectorfields X in which case the positive energy condition for $\mathbf{T}$ insures that all integrands in (**??**) will be positive. We know that, up to a Lorentz transformation the only Killing, future oriented timelike vectorfield is a constant multiple of ∂_t. Choosing $X = \partial_t$ (375) becomes,

$$\int_{B_{t_2}} \mathbf{T}(\partial_t, \partial_t) + \int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \mathbf{T}(E_-, \partial_t) = \int_{B_{t_1}} \mathbf{T}(\partial_t, \partial_t) \quad (376)$$

[10] The brackets $\langle \cdot, \cdot \rangle$ in (373) denote inner product with respect to the Minkowski metric.

In the case of a conformal field theory we can pick X to be the future timelike, conformal Killing vectorfield $X = K_0 = (t^2 + |x|^2)\partial_t + 2tx^i\partial_i$. Thus,

$$\int_{B_{t_2}} \mathbf{T}(\partial_t, K_0) + \int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \mathbf{T}(L, K_0) = \int_{B_{t_1}} \mathbf{T}(\partial_t, K_0) \tag{377}$$

In (376) the term $\mathbf{T}(\partial_t, \partial_t)$ is called *energy density* while $\mathbf{T}(E_-, \partial_t)$ is called *energy flux density* . The corresponding integrals are called *energy contained in B_{t_1}, and* B_{t_2} and, respectively, *flux of energy through $\mathcal{N}^-$.* The coresponding terms in (377) are called conformal energy densities, fluxes etc.

Equation (376) can be used to derive the following fundamental properties of relativistic field theories.

(1) Finite propagation speed
(2) Uniqueness of the Cauchy problem

Proof : The first property follows from the fact that, if $\int_{B_{t_1}} \mathbf{T}(\partial_t, \partial_t)$ is zero at time $t = t_1$ then both integrals $\int_{B_{t_2}} \mathbf{T}(\partial_t, \partial_t)$ and $\int_{\mathcal{N}^-_{[t_1,t_2]}} \mathbf{T}(E_-, \partial_t)$ must vanish also. In view of the positivity properties of the $\mathbf{T}$ it follows that the corresponding integrands must also vanish. Taking into account the specific form of $\mathbf{T}$, in a particular theory, one can then show that the fields do also vanish in the domain of influence of the ball B_{t_1}. Conversely, if the initial data for the fields vanish in the complement of B_{t_1}, the the fields are identically zero in the complement of the domain of influence of of B_{t_1}.

The proof of the second property follows immediately from the first for a linear field theory. For a nonlinear theory one has to work a little more. ■

Exercise. Formulate an initial value problem for each of the field theories we have encountered so far, scalar wave equation (SWE), Wave Maps (WM), Maxwell equations (ME) and Yang-Mills (YM). Proof uniqueness of solutions to the initial value problem, for smooth solutions.

The following is another important consequence of (376) and (377). To state the results we introduce the following quantities,

$$\mathcal{E}(t) = \int_{\mathbb{R}^n} \mathbf{T}\left(\partial_t, \partial_t\right)(t,x)dx \tag{378}$$

$$\mathcal{E}_c(t) = \int_{\mathbb{R}^n} \mathbf{T}\left(K_0, \partial_t\right)(t,x)dx \tag{379}$$

THEOREM 3.4 (Global Energy). *For an arbitrary field theory, if $\mathcal{E}(0) < \infty$, then*

$$\mathcal{E}(t) = \mathcal{E}(0) \tag{380}$$

Moreover, for a conformal field theory, if $\mathcal{E}_c(0) < \infty$,

$$\mathcal{E}_c(t) = \mathcal{E}_c(0) \tag{381}$$

Proof : Follows easily by applying (376) and (377) to past causal domains $\mathcal{J}^-(p)$ with $p = (\bar{t}, 0)$ between $t_1 = 0$ and $t_2 = t$ and letting $\bar{t} \to +\infty$. ∎

3.5. Energy dissipation. In this section we shall make use of the global conformal energy identity (381) to show how energy dissipates for a filed theories in Minkowski space. Consider a conformal field theory defined on all of $\mathbb{R}^{n+1}$. At each point of $\mathbb{R}^{n+1}$, with $t \geq 0$, define the standard null frame where

$$\begin{aligned} L = E_+ &= \partial_t + \partial_r \\ \underline{L} = E_- &= \partial_t - \partial_r. \end{aligned}$$

Observe that the conformal Killing vectorfield $K_0 = (t^2 + r^2)\partial_t + 2rt\partial_r$ can be expressed in the form,

$$K_0 = \frac{1}{2}\left[(t+r)^2 E_+ + (t-r)^2 E_-.\right]$$

Thus,

$$\begin{aligned} \mathcal{E}_c(t) &= \int_{\mathbb{R}^n} \frac{1}{4}(t+r)^2 \mathbf{T}_{++} + \frac{1}{4}(t-r)^2 \mathbf{T}_{--} + \underbrace{\left((t+r)^2 + (t-r)^2\right)}_{2(t^2+r^2)} \mathbf{T}_{+-} dx. \\ &= \int_{\mathbb{R}^n} \frac{1}{4}(t+r)^2 \mathbf{T}_{++} + \frac{1}{2}(t^2+r^2)\mathbf{T}_{+-} + \frac{1}{4}(t-r)^2 \mathbf{T}_{--} dx \qquad (382) \\ \mathcal{E}_c(0) &= \int_{\mathbb{R}^n} \mathbf{T}(\partial_t, K_0)(0,x) dx = \int_{\mathbb{R}^n} |x|^2 \mathbf{T}(\partial_t, \partial_t) dx \end{aligned}$$

According to (381) we have $\mathcal{E}_c(t) = \mathcal{E}_c(0)$. Assuming that $\mathcal{E}_c(0) = \int_{\mathbb{R}^n} |x|^2 \mathbf{T}(\partial_t, \partial_t) dx$ is finite we conclude that,

$$\begin{aligned} \int_{\mathbb{R}^n} \mathbf{T}_{++}(t, \cdot) dx &\lesssim \frac{\mathcal{E}_c(0)}{t^2} \\ \int_{\mathbb{R}^n} \mathbf{T}_{+-}(t, \cdot) dx &\lesssim \frac{\mathcal{E}_c(0)}{t^2}. \end{aligned}$$

The remaining term in (382) contains the factor $(t-r)^2$ which is constant along outgoing null directions $r = t + c$. Hence for any $0 < \epsilon < 1$

$$\begin{aligned} \int_{|x|>(1+\epsilon)t} \mathbf{T}_{--} &= O(t^{-2}) \\ \int_{|x|<(1-\epsilon)t} \mathbf{T}_{--} &= O(t^{-2}). \end{aligned}$$

We conclude that most of the energy of a conformal field is carried by the $\mathbf{T}_{--}$ component and propagates near the light cone.

CHAPTER 4

APPENDIX: BASIC GEOMETRIC NOTIONS

In what follows we give a short overview of the basic notions in Riemannian and Lorentzian geometry. These will allow us to extend some of the basic facts about the standard Laplace, Heat and Wave equations, to manifolds. It will also allow us later to discuss more complicated nonlinear geometric equations.

0.6. Pseudo-riemannian metrics, tensor fields. A pseudo-riemannian manifold [1], or simply a spacetime, consist of a pair $(\mathbf{M}, \mathbf{g})$ where $\mathbf{M}$ is an orientable $p+q$-dimensional manifold and $\mathbf{g}$ is a pseudo-riemannian metric defined on it, that is a smooth, a non degenerate, 2-covariant symmetric tensor field of signature (p, q). This means that at each point $p \in \mathbf{M}$ one can choose a basis of $p+q$ vectors, $\{e_{(\alpha)}\}$, belonging to the tangent space $T\mathbf{M}_p$, such that

$$\mathbf{g}(e_{(\alpha)}, e_{(\beta)}) \quad = \quad \eta_{\alpha\beta} \tag{383}$$

for all $\alpha, \beta = 0, 1, ..., n$, where η is the diagonal matrix with -1 in the first p entries and $+1$ in the last q entries. If X is an arbitrary vector at p expressed, in terms of the basis $\{e_{(\alpha)}\}$, as $X = X^\alpha e_{(\alpha)}$, we have

$$\mathbf{g}(X, X) = -(X^1)^2 - \ldots - (X^p)^2 + (X^{p+1})^2 + ... + (X^{p+q})^2 \tag{384}$$

The case when $p = 0$ and $q = n$ corresponds to Riemannian manifolds of dimension n. The other case of interest for us is $p = 1, q = n$ which corresponds to a Lorentzian manifolds of dimension $n+1$. The primary example of Riemannian manifold is the Euclidean space $\mathbb{R}^n$. Any other Riemannian manifold looks, locally, like $\mathbb{R}^n$. Similarly, the primary example of a Lorentzian manifold is the Minkowski spacetime, the spacetime of Special Relativity. It plays the same role, in Lorentzian geometry, as the Euclidean space in Riemannian geometry. In this case the manifold $\mathbf{M}$ is diffeomorphic to $\mathbb{R}^{n+1}$ and there exists globally defined systems of coordinates, x^α, relative to which the metric takes the diagonal form $-1, 1, ..., 1$. All such systems are related through Lorentz transformations and are called inertial. We shall denote the Minkowski spacetime of dimension $n+1$ by $(\mathbb{R}^{n+1}, \mathbf{m})$.

Relative to a given coordinate system x^μ, the components of a pseudo-riemannian metric take the form

$$g_{\mu\nu} = \mathbf{g}(\partial_\mu, \partial_\nu)$$

[1]We assume that our reader is already familiar with the basics concepts of differential geometry such as manifolds, tensor fields, covariant, Lie and exterior differentiation. For a short introduction to these concepts see Chapter 2 of Hawking and Ellis, "The large scale structure of space-time", [**6**]

where $\partial_\mu = \frac{\partial}{\partial x^\mu}$ are the associated coordinate vectorfields. We denote by $g^{\mu\nu}$ the components of the inverse metric g^{-1} relative to the same coordinates x, and by $|g|$ the determinant of the matrix $g_{\mu\nu}$. The volume element $dv_{\mathbf{M}}$ of $\mathbf{M}$ is expressed, in local coordinates, by $\sqrt{|g|}dx = \sqrt{|g|}dx^1 \dots dx^n$. Thus the integral $\int_{\mathbf{M}} f dv_{\mathbf{M}}$ of a function f, supported in coordinate chart $U \subset \mathbf{M}$ is defined by $\int_U f(x)\sqrt{|g(x)|}dx$. The integral on $\mathbf{M}$ of an arbitrary function f is defined by making a partition of unity subordinated to a covering of $\mathbf{M}$ by coordinate charts. One can easily check that the definition is independent of the particular system of local coordinates.

In view of (384) we see that a Lorentzian metric divides the vectors in the tangent space $T\mathbf{M}_p$ at each p, into timelike, null or spacelike according to whether the quadratic form

$$(X, X) = g_{\mu\nu}X^\mu X^\nu \tag{385}$$

is, respectively, negative, zero or positive. The set of null vectors N_p forms a double cone, called the null cone of the corresponding point p. The set of timelike vectors I_p forms the interior of this cone. The vectors in the union of I_p and N_p are called causal. The set S_p of spacelike vectors is the complement of $I_p \cup N_p$.

A frame $e_{(\alpha)}$ verifying (383) is said to be orthonormal. In the case of Lorentzian manifolds it makes sense to consider, in addition to orthonormal frames, null frames. These are collections of vectorfields[2] e_α consisting of two null vectors e_{n+1}, e_n and orthonormal spacelike vectors $(e_a)_{a=1,\dots,n-1}$ which verify,

$$\begin{aligned}
\mathbf{g}(e_n, e_n) = \mathbf{g}(e_{n+1}, e_{n+1}) = 0\ ,\ \mathbf{g}(e_n, e_{n+1}) = -2 \\
\mathbf{g}(e_n, e_a) = \mathbf{g}(e_{n+1}, e_a) = 0\ ,\ \mathbf{g}(e_a, e_b) = \delta_{ab}
\end{aligned}$$

One-forms $A = A_\alpha dx^\alpha$ are sections of the cotangent bundle of $\mathbf{M}$. We denote by $A(X)$ the natural pairing between A and a vectorfield X. We can raise the indices of A by $A^\alpha = \mathbf{g}^{\alpha\beta}\mathbf{A}_\beta$. $A' = A^\alpha\partial_\alpha$ defines a vectorfield on $\mathbf{M}$ and we have, $A(X) = \mathbf{g}(A', X)$. Covariant tensors A of order k are k-multilinear forms on $T\mathbf{M}$.

Notation: We will use the following notational conventions: We shall use boldface characters to denote important tensors such as the metric $\mathbf{g}$, and the Riemann curvature tensor $\mathbf{R}$. Their components relative to arbitrary frames will also be denoted by boldface characters. Thus, given a frame $\{e_{(\alpha)}\}$ we write $\mathbf{g}_{\alpha\beta} = \mathbf{g}(e_\alpha, e_\beta)$, $\mathbf{R}_{\alpha\beta\gamma\delta} = \mathbf{R}(e_\alpha, e_\beta, e_\gamma, e_\delta)$ and, for an arbitrary tensor T,

$$T_{\alpha\beta\gamma\delta\dots} \equiv T(e_\alpha, e_\beta, e_\gamma, e_\delta, \dots)$$

We shall not use boldface characters for the components of tensors, relative to a fixed system of coordinates. Thus, for instance, in (385) $g_{\mu\nu} = \mathbf{g}(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu})$. In the case of a Riemannian manifold we use latin letters $i, j, k, l, \dots$ to denote indices of coordinates $x^1, x^2, \dots, x^n$ or tensors. For a Lorentzian manifold we use greek letters $\alpha, \beta, \gamma, \dots$ to denote indices $0, 1, \dots, n$.

[2]We write e_α instead of $e_{(\alpha)}$ to simplify the notation, whenever there can be no confusion.

We will review the following topics below:

1.) Lie brackets of vectorfields. Frobenius theorem

2.) Lie derivative of a tensorfield

3.) Multilinear forms and exterior differentiation

4.) Connections and covariant derivatives

5.) Pseudo-riemannian metrics. Riemannian and Lorentzian geometry.

6.) Levi-Civita connection associated to a pseudo-riemannian metric.

7.) Parallel transport, geodesics, exponential map, completeness

8.) Curvature tensor of a pseudo-riemannian manifold. Symmetries. First and second Bianchi identities.

9.) Isometries and conformal isometries. Killing and conformal Killing vectorfields.

0.7. Covariant derivatives, Lie derivatives. We recall here the three fundamental operators of the differential geometry on a Riemann or Lorentz manifold: the exterior derivative, the Lie derivative, and the Levi-Civita connection with its associated covariant derivative.

0.7.1. *The exterior derivative.* Given a scalar function f its differential df is the 1-form defined by

$$df(X) = X(f)$$

for any vector field X. This definition can be extended for all differential forms on $\mathbf{M}$ in the following way:

i) d is a linear operator defined from the space of all k-forms to that of $k+1$-forms on $\mathbf{M}$. Thus for all k-forms A,B and real numbers λ, μ

$$d(\lambda A + \mu B) = \lambda dA + \mu dB$$

ii) For any k-form A and arbitrary form B

$$d(A \wedge B) = dA \wedge B + (-1)^k A \wedge dB$$

iii) For any form A,

$$d^2 A = 0\ .$$

We recall that, if Φ is a smooth map defined from $\mathbf{M}$ to another manifold $\mathbf{M}'$, then

$$d(\Phi^* A) = \Phi^*(dA)\ .$$

Finally if A is a one form and X, Y arbitrary vector fields, we have the equation

$$dA(X,Y) = \frac{1}{2}\Big(X(A(Y)) - Y(A(X)) - A([X,Y])\Big)$$

where $[X,Y]$ is the commutator $X(Y)-Y(X)$. This can be easily generalised to arbitrary k forms, see Spivak's book, Vol.I, Chapter 7, Theorem 13. [**17**]

0.7.2. *The Lie derivative.* Consider an arbitrary vector field X. In local coordinates x^μ, the flow of X is given by the system of differential equations

$$\frac{dx^\mu}{dt} = X^\mu(x^1(t),...,x^{p+q}(t)) \ .$$

The corresponding curves, $x^\mu(t)$, are the integral curves of X. For each point $p \in \mathbf{M}$ there exists an open neighborhood $\mathcal{U}$, a small $\epsilon > 0$ and a family of diffeomorphism $\Phi_t : \mathcal{U} \to \mathbf{M},\ |t| \le \epsilon$, obtained by taking each point in $\mathcal{U}$ to a parameter distance t, along the integral curves of X. We use these diffeomorphisms to construct, for any given tensor T at p, the family of tensors $(\Phi_t)_* T$ at $\Phi_t(p)$.

The Lie derivative $\mathcal{L}_X T$ of a tensor field T, with respect to X, is:

$$\mathcal{L}_X T|_p \equiv \lim_{t\to 0} \frac{1}{t}\left(T|_p - (\Phi_t)_* T|_p\right) \ .$$

It has the following properties:

i) $\mathcal{L}_X$ linearly maps (p,q)-tensor fields into tensor fields of the same type.

ii) $\mathcal{L}_X$ commutes with contractions.

iii) For any tensor fields S, T,

$$\mathcal{L}_X(S\otimes T) = \mathcal{L}_X S \otimes T + S \otimes \mathcal{L}_X T \ .$$

If X is a vector field we easily check that

$$\mathcal{L}_X Y = [X,Y]$$

by writing $(\mathcal{L}_X Y)^i = -\frac{d}{dt}((\Phi_t)_* Y)^i\Big|_{t=0}$ and expressing $(\Phi_t)_* Y)^i\Big|_p = \frac{\partial x^i(\Phi_t(q))}{\partial x^j(q)} Y^j\Big|_q$, where $q = \Phi_{-t}(p)$. (See [**6**], Hawking and Ellis, section 2.4 for details.)

If A is a k-form we have, as a consequence of the commutation formula of the exterior derivative with the pull-back Φ^*,

$$d(\mathcal{L}_X A) = \mathcal{L}_X(dA) \ .$$

For a given k-covariant tensorfield T we have,

$$\mathcal{L}_X T(Y_1,\dots,Y_k) = XT(Y_1,\dots,Y_k) - \sum_{i=1}^{k} T(Y_1,\dots,\mathcal{L}_X Y_i,\dots,Y_k)$$

We remark that the Lie bracket of two coordinate vector fields vanishes,

$$\left[\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right] = 0.$$

The converse is also true, namely, see Spivak, [**17**], Vol.I, Chapter 5, PROPOSITION 0.8. *If $X_{(0)}, ..., X_{(k)}$ are linearly independent vector fields in a neighbourhood of a point p and the Lie bracket of any two of them is zero then there exists a coordinate system x^μ, around p such that $X_{(\rho)} = \frac{\partial}{\partial x^\rho}$ for each $\rho = 0, ..., k$.*

The above proposition is the main step in the proof of Frobenius Theorem. To state the theorem we recall the definition of a k-distribution in $\mathbf{M}$. This is an arbitrary smooth assignment of a k-dimensional plane π_p at every point in a domain $\mathcal{U}$ of $\mathbf{M}$. The distribution is said to be involute if, for any vector fields X, Y on $\mathcal{U}$ with $X|_p, Y|_p \in \pi_p$, for any $p \in \mathcal{U}$, we have $[X, Y]|_p \in \pi_p$. This is clearly the case for integrable distributions[3]. Indeed if $X|_p, Y|_p \in T\mathcal{N}_p$ for all $p \in \mathcal{N}$, then X, Y are tangent to $\mathcal{N}$ and so is also their commutator $[X, Y]$. The Frobenius Theorem establishes that the converse is also true[4], that is being in involution is also a sufficient condition for the distribution to be integrable,

THEOREM 0.9. *(Frobenius Theorem) A necessary and sufficient condition for a distribution $(\pi_p)_{p\in\mathcal{U}}$ to be integrable is that it is involute.*

0.9.1. *The connection and the covariant derivative.* A connection $\mathbf{D}$ is a rule which assigns to each vectorfield X a differential operator $\mathbf{D}_X$. This operator maps vector fields Y into vector fields $\mathbf{D}_X Y$ in such a way that, with $\alpha, \beta \in \mathbb{R}$ and f, g scalar functions on $\mathbf{M}$,

$$\begin{aligned} &a) \ \ \mathbf{D}_{fX+gY} Z = f\mathbf{D}_X Z + g\mathbf{D}_Y Z \\ &b) \ \ \mathbf{D}_X(\alpha Y + \beta Z) = \alpha\mathbf{D}_X Y + \beta\mathbf{D}_X Z \\ &c) \ \ \mathbf{D}_X fY = X(f)Y + f\mathbf{D}_X Y \end{aligned} \tag{386}$$

Therefore, at a point p,

$$\mathbf{D}Y \equiv Y^\alpha_{\ ;\beta}\, \theta^{(\beta)} \otimes e_{(\alpha)} \tag{387}$$

where the $\theta^{(\beta)}$ are the one-forms of the dual basis respect to the orthonormal frame $e_{(\beta)}$. Observe that $Y^\alpha_{;\beta} = \theta^{(\alpha)}(\mathbf{D}_{e_{(\beta)}} Y)$. On the other side, from c),

$$\mathbf{D}fY = df \otimes Y + f\mathbf{D}Y$$

so that

$$\mathbf{D}Y = \mathbf{D}(Y^\alpha e_{(\alpha)}) = dY^\alpha \otimes e_{(\alpha)} + Y^\alpha \mathbf{D}e_{(\alpha)}$$

and finally, using $df(\cdot) = e_{(\alpha)}(f)\theta^{(\alpha)}(\cdot)$,

$$\mathbf{D}Y = \Big(e_{(\beta)}(Y^\alpha) + Y^\gamma \theta^{(\alpha)}(\mathbf{D}_{e_{(\beta)}} e_{(\gamma)}) \Big)\, \theta^{(\beta)} \otimes e_{(\alpha)} \tag{388}$$

Therefore

$$Y^\alpha_{;\beta} = e_{(\beta)}(Y^\alpha) + \Gamma^\alpha_{\beta\gamma} Y^\gamma$$

and the connection is, therefore, determined by its connection coefficients,

$$\Gamma^\alpha_{\beta\gamma} = \theta^{(\alpha)}(\mathbf{D}_{e_{(\beta)}} e_{(\gamma)}) \tag{389}$$

[3]Recall that a distribution π on $\mathcal{U}$ is said to be integrable if through every point $p \in \mathcal{U}$ there passes a unique submanifold $\mathcal{N}$, of dimension k, such that $\pi_p = T\mathcal{N}_p$.

[4]For a proof see Spivak, citeSpivak, Vol.I, Chapter 6.

which, in a coordinate basis, are the usual Christoffel symbols and have the expression

$$\Gamma^{\mu}_{\rho\nu} = dx^{\mu}(\mathbf{D}_{\frac{\partial}{\partial x^{\rho}}} \frac{\partial}{\partial x^{\nu}})$$

Finally

$$\mathbf{D}_X Y = \left(X(Y^{\alpha}) + \Gamma^{\alpha}_{\beta\gamma} X^{\beta} Y^{\gamma} \right) e_{(\alpha)} \tag{390}$$

In the particular case of a coordinate frame we have

$$\mathbf{D}_X Y = \left(X^{\mu} \frac{\partial Y^{\nu}}{\partial x^{\mu}} + \Gamma^{\nu}_{\rho\sigma} X^{\rho} Y^{\sigma} \right) \frac{\partial}{\partial x^{\nu}}$$

A connection is said to be a Levi-civita connection if $\mathbf{D}\mathbf{g} = 0$. That is, for any three vector fields X, Y, Z,

$$Z(\mathbf{g}(X,Y)) = \mathbf{g}(\mathbf{D}_Z X, Y) + \mathbf{g}(X, \mathbf{D}_Z Y) \tag{391}$$

A very simple and basic result of differential geometry asserts that for any given metric there exists a unique affine connection associated to it.

Proposition 0.10. *There exists a unique connection on* $\mathbf{M}$*, called the Levi-Civita connection, which satisfies* $\mathbf{D}\,\mathbf{g} = 0$*. The connection is torsion free, that is,*

$$\mathbf{D}_X Y - \mathbf{D}_Y X = [X, Y] \ .$$

Moreover, relative to a system of coordinates, x^{μ}*, the Christoffel symbol of the connection is given by the standard formula*

$$\Gamma^{\mu}_{\rho\nu} = \frac{1}{2} g^{\mu\tau} \left(\partial_{\rho} g_{\nu\tau} + \partial_{\nu} g_{\tau\rho} - \partial_{\tau} g_{\nu\rho} \right) \ .$$

Exercise: Prove the proposition yourself, without looking in a book.

So far we have only defined the covariant derivative of a a vector field. We can easily extend the definition to one forms $A = A_{\alpha} dx^{a}$ by the requirement that,

$$X(A(Y)) = \mathbf{D}_X A(Y) + A(D_X Y),$$

for all vectorfields X, Y. Given a k-covariant tensor field T we define its covariant derivative $\mathbf{D}_X T$ by the rule,

$$\mathbf{D}_X T(Y_1, \dots, Y_k) = XT(Y_1, \dots, Y_k) - \sum_{i=1}^{k} T(Y_1, \dots, \mathbf{D}_X Y_i, \dots, Y_k)$$

We can talk about $\mathbf{D}T$ as a covariant tensor of rank $k+1$ defined by,

$$\mathbf{D}T(X, Y_1, \dots, Y_k) = \mathbf{D}_X T(Y_1, \dots, Y_k).$$

Given a frame e_{α} we denote by $T_{\alpha_1 \dots, \alpha_k; \beta} = \mathbf{D}T(e_{\beta}, e_{a_1}, \dots, e_{\alpha_k})$ the components of $\mathbf{D}T$ relative to the frame. By repeated covariant differentiation we can define $\mathbf{D}^2 T, \dots \mathbf{D}^m \mathbf{T}$. Relative to a frame e_{α} we write,

$$\mathbf{D}_{\beta_1} \dots \mathbf{D}_{\beta_m} T_{\alpha_1 \dots \alpha_k} = T_{\alpha_1 \dots \alpha_k; \beta_1 \dots \beta_m} = \mathbf{D}^m T(e_{\beta_1} \dots, e_{\beta_m}, e_{\alpha_1}, ..., e_{\alpha_k}).$$

The fact that the Levi-Civita connection is torsion free allows us to connect covariant differentiation to the Lie derivative. Thus, if T is a k-covariant tensor we have, in a coordinate basis,

$$(\mathcal{L}_X T)_{\sigma_1 \ldots \sigma_k} = X^\mu T_{\sigma_1 \ldots \sigma_k ; \mu} + X^\mu{}_{;\sigma_1} T_{\mu \sigma_2 \ldots \sigma_k} + \ldots . + X^\mu{}_{;\sigma_k} T_{\sigma_1 \ldots \sigma_{k-1} \mu} \; .$$

The covariant derivative is also connected to the exterior derivative according to the following simple formula. If A is a k-form, we have[5] $A_{[\sigma_1 \ldots \sigma_k ; \mu]} = A_{[\sigma_1 \ldots \sigma_k , \mu]}$ and

$$dA = \sum A_{\sigma_1 \ldots \sigma_k ; \mu} dx^\mu \wedge dx^{\sigma_1} \wedge dx^{\sigma_2} \wedge \ldots . \wedge dx^{\sigma_k} \; .$$

Given a smooth curve $\mathbf{x} : [0,1] \to \mathbf{M}$, parametrized by t, let $T = \left(\frac{\partial}{\partial t}\right)_{\mathbf{x}}$ be the corresponding tangent vector field along the curve. A vector field X, defined on the curve, is said to be parallelly transported along it if $\mathbf{D}_T X = 0$. If the curve has the parametric equations $x^\nu = x^\nu(t)$, relative to a system of coordinates, then $T^\mu = \frac{dx^\mu}{dt}$ and the components $X^\mu = X^\mu(\mathbf{x}(t))$ satisfy the ordinary differential system of equations

$$\frac{\mathbf{D}}{dt} X^\mu \equiv \frac{dX^\mu}{dt} + \Gamma^\mu_{\rho\sigma}(\mathbf{x}(t)) \frac{dx^\rho}{dt} X^\sigma = 0 \; .$$

The curve is said to be geodesic if, at every point of the curve, $\mathbf{D}_T T$ is tangent to the curve, $\mathbf{D}_T T = \lambda T$. In this case one can reparametrize the curve such that, relative to the new parameter s, the tangent vector $S = \left(\frac{\partial}{\partial s}\right)_{\mathbf{x}}$ satisfies $\mathbf{D}_S S = 0$. Such a parameter is called an "affine parameter". The affine parameter is defined up to a transformation $s = as' + b$ for a, b constants. Relative to an affine parameter s and arbitrary coordinates x^μ the geodesic curves satisfy the equations

$$\frac{d^2 x^\mu}{ds^2} + \Gamma^\mu_{\rho\sigma} \frac{dx^\rho}{ds} \frac{dx^\sigma}{ds} = 0 \; .$$

A geodesic curve parametrized by an affine parameter is simply called a geodesic. In Lorentzian geometry timelike geodesics correspond to world lines of particles freely falling in the gravitational field represented by the connection coefficients. In this case the affine parameter s is called the proper time of the particle.

Given a point $p \in \mathbf{M}$ and a vector X in the tangent space $T_p\mathbf{M}$, let $\mathbf{x}(t)$ be the unique geodesic starting at p with "velocity" X. We define the exponential map:

$$\exp_p : T_p\mathbf{M} \to \mathbf{M} \; .$$

This map may not be defined for all $X \in T_p\mathbf{M}$. The theorem of existence and uniqueness for systems of ordinary differential equations implies that the exponential map is defined in a neighbourhood of the origin in $T_p\mathbf{M}$. If the exponential map is defined for all $T_p\mathbf{M}$, for every point p the manifold $\mathbf{M}$ is said geodesically complete. In general if the connection is a C^r connection[6] there exists an open neighbourhood $\mathcal{U}_0$ of the origin in $T_p\mathbf{M}$ and an open neighbourhood of the point

[5] $[\sigma_1 ... \sigma_k ; \mu]$ indicates the antisymmetrization with respect to all indices (i.e. $\frac{1}{k!}$(alternating sum of the tensor over all permutations of the indices)) and "$,\mu$" indicates the ordinary derivative with respect to x^μ.

[6] A C^r connection is such that if Y is a C^{r+1} vector field then $\mathbf{D}Y$ is a C^r vector field.

p in $\mathbf{M}$, $\mathcal{V}_p$, such that the map $\exp_p$ is a C^r diffeomorphism of $\mathcal{U}_0$ onto $\mathcal{V}_p$. The neighbourhood $\mathcal{V}_p$ is called a normal neighbourhood of p.

0.11. Riemann curvature tensor, Ricci tensor, Bianchi identities. In the flat spacetime if we parallel transport a vector along any closed curve we obtain the vector we have started with. This fails in general because the second covariant derivatives of a vector field do not commute. This lack of commutation is measured by the Riemann curvature tensor,

$$\mathbf{R}(X,Y)Z = \mathbf{D}_X(\mathbf{D}_Y Z) - \mathbf{D}_Y(\mathbf{D}_X Z) - \mathbf{D}_{[X,Y]}Z \tag{392}$$

or written in components relative to an arbitrary frame,

$$\mathbf{R}^{\alpha}{}_{\beta\gamma\delta} = \theta^{(\alpha)}\big((\mathbf{D}_\gamma\mathbf{D}_\delta - \mathbf{D}_\delta\mathbf{D}_\gamma)e_{(\beta)}\big) \tag{393}$$

Relative to a coordinate system x^μ and written in terms of the $g_{\mu\nu}$ components, the Riemann components have the expression

$$R^{\mu}{}_{\nu\rho\sigma} = \frac{\partial\Gamma^{\mu}_{\sigma\nu}}{\partial x^\rho} - \frac{\partial\Gamma^{\mu}_{\rho\nu}}{\partial x^\sigma} + \Gamma^{\mu}_{\rho\tau}\Gamma^{\tau}_{\sigma\nu} - \Gamma^{\mu}_{\sigma\tau}\Gamma^{\tau}_{\rho\nu} \tag{394}$$

The fundamental property of the curvature tensor, first proved by Riemann, states that if $\mathbf{R}$ vanishes identically in a neighbourhood of a point p one can find families of local coordinates such that, in a neighbourhood of p, $g_{\mu\nu} = \eta_{\mu\nu}$ [7].

The trace of the curvature tensor, relative to the metric $\mathbf{g}$, is a symmetric tensor called the Ricci tensor,

$$\mathbf{R}_{\alpha\beta} = \mathbf{g}^{\gamma\delta}\mathbf{R}_{\alpha\gamma\beta\delta}$$

The scalar curvature is the trace of the Ricci tensor

$$\mathbf{R} = \mathbf{g}^{\alpha\beta}\mathbf{R}_{\alpha\beta}\ .$$

The Riemann curvature tensor of an arbitrary spacetime $(\mathbf{M},\mathbf{g})$ has the following symmetry properties,

$$\begin{aligned}&\mathbf{R}_{\alpha\beta\gamma\delta} = -\mathbf{R}_{\beta\alpha\gamma\delta} = -\mathbf{R}_{\alpha\beta\delta\gamma} = \mathbf{R}_{\gamma\delta\alpha\beta}\\ &\mathbf{R}_{\alpha\beta\gamma\delta} + \mathbf{R}_{\alpha\gamma\delta\beta} + \mathbf{R}_{\alpha\delta\beta\gamma} = 0\end{aligned} \tag{395}$$

The second identity in (395) is called the first Bianchi identity.

It also satisfies the second Bianchi identities, which we refer to here as the *Bianchi equations* and, in a generic frame, have the form:

$$\mathbf{D}_{[\epsilon}\mathbf{R}_{\gamma\delta]\alpha\beta} = 0 \tag{396}$$

The traceless part of the curvature tensor, $\mathbf{C}$ is called the Weyl tensor, and has the following expression in an arbitrary frame,

$$\begin{aligned}\mathbf{C}_{\alpha\beta\gamma\delta} &= \mathbf{R}_{\alpha\beta\gamma\delta} - \frac{1}{n-1}\left(\mathbf{g}_{\alpha\gamma}\mathbf{R}_{\beta\delta} + \mathbf{g}_{\beta\delta}\mathbf{R}_{\alpha\gamma} - \mathbf{g}_{\beta\gamma}\mathbf{R}_{\alpha\delta} - \mathbf{g}_{\alpha\delta}\mathbf{R}_{\beta\gamma}\right)\\ &+ \frac{1}{n(n-1)}(\mathbf{g}_{\alpha\gamma}\mathbf{g}_{\beta\delta} - \mathbf{g}_{\alpha\delta}\mathbf{g}_{\beta\gamma})\mathbf{R}\end{aligned} \tag{397}$$

[7] For a thorough discussion and proof of this fact, refer to Spivak, [**17**], Vol. II.

Observe that $\mathbf{C}$ verifies all the symmetry properties of the Riemann tensor:

$$\begin{aligned}\mathbf{C}_{\alpha\beta\gamma\delta} &= -\mathbf{C}_{\beta\alpha\gamma\delta} = -\mathbf{C}_{\alpha\beta\delta\gamma} = \mathbf{C}_{\gamma\delta\alpha\beta}\\ \mathbf{C}_{\alpha\beta\gamma\delta} &+ \mathbf{C}_{\alpha\gamma\delta\beta} + \mathbf{C}_{\alpha\delta\beta\gamma} = 0\end{aligned} \tag{398}$$

and, in addition, $\quad \mathbf{g}^{\alpha\gamma}\mathbf{C}_{\alpha\beta\gamma\delta} = 0$.

We say that two metrics $\mathbf{g}$ and $\hat{\mathbf{g}}$ are conformal if $\hat{\mathbf{g}} = \lambda^2\mathbf{g}$ for some non zero differentiable function λ. Then the following theorem holds (see Hawking- Ellis, [**6**], chapter 2, section 2.6):

THEOREM 0.12. *Let* $\hat{\mathbf{g}} = \lambda^2\mathbf{g}$*,* $\hat{\mathbf{C}}$ *the Weyl tensor relative to* $\hat{\mathbf{g}}$ *and* $\mathbf{C}$ *the Weyl tensor relative to* $\mathbf{g}$*. Then*

$$\hat{\mathbf{C}}^{\alpha}_{\beta\gamma\delta} = \mathbf{C}^{\alpha}_{\beta\gamma\delta} \ .$$

Thus $\mathbf{C}$ *is conformally invariant.*

0.13. Isometries and conformal isometries, Killing and conformal Killing vector fields. Definition. A diffeomorphism $\Phi : \mathcal{U} \subset \mathbf{M} \to \mathbf{M}$ is said to be a conformal isometry if, at every point p, $\Phi_*\mathbf{g} = \Lambda^2\mathbf{g}$, that is,

$$(\Phi^*\mathbf{g})(X, Y)|_p = \mathbf{g}(\Phi_* X, \Phi_* Y)|_{\Phi(p)} = \Lambda^2\mathbf{g}(X, Y)|_p$$

with $\Lambda \neq 0$. If $\Lambda = 1$, Φ is called an isometry of $\mathbf{M}$.

Definition. A vector field K which generates a one parameter group of isometries (respectively, conformal isometries) is called a Killing (respectively, conformal Killing) vector field.

Let K be such a vector field and Φ_t the corresponding one parameter group. Since the $(\Phi_t)_*$ are conformal isometries, we infer that $\mathcal{L}_K\mathbf{g}$ must be proportional to the metric $\mathbf{g}$. Moreover $\mathcal{L}_K\mathbf{g} = 0$ if K is a Killing vector field.

Definition. Given an arbitrary vector field X we denote ${}^{(X)}\pi$ the deformation tensor of X defined by the formula

$${}^{(X)}\pi_{\alpha\beta} = (\mathcal{L}_X g)_{\alpha\beta} = \mathbf{D}_\alpha X_\beta + \mathbf{D}_\beta X_\alpha \ .$$

The tensor ${}^{(X)}\pi$ measures, in a precise sense, how much the diffeomorphism generated by X differs from an isometry or a conformal isometry. The following Proposition holds, (see Hawking-Ellis, citeHawkEll, chapter 2, section 2.6):

PROPOSITION 0.14. *The vector field* X *is Killing if and only if* ${}^{(X)}\pi = 0$*. It is conformal Killing if and only if* ${}^{(X)}\pi$ *is proportional to* $\mathbf{g}$*.*

Remark: One can choose local coordinates such that $X = \frac{\partial}{\partial x^\mu}$. It then immediately follows that, relative to these coordinates the metric $\mathbf{g}$ is independent of the component x^μ.

PROPOSITION 0.15. *On any pseudo-riemannian spacetime* $\mathbf{M}$*, of dimension* $n = p + q$*, there can be no more than* $\frac{1}{2}(p+q)(p+q+1)$ *linearly independent Killing vector fields.*

Proof: Proposition 0.15 is an easy consequence of the following relation, valid for an arbitrary vector field X, obtained by a straightforward computation and the use of the symmetries of $\mathbf{R}$.

$$\mathbf{D}_\beta \mathbf{D}_\alpha X_\lambda = \mathbf{R}_{\lambda\alpha\beta\delta} X^\delta + {}^{(X)}\Gamma_{\alpha\beta\lambda} \tag{399}$$

where

$$^{(X)}\Gamma_{\alpha\beta\lambda} = \frac{1}{2}\left(\mathbf{D}_\beta \pi_{\alpha\lambda} + \mathbf{D}_\alpha \pi_{\beta\lambda} - \mathbf{D}_\lambda \pi_{\alpha\beta}\right) \tag{400}$$

and $\pi \equiv {}^{(X)}\pi$ is the X deformation tensor.

If X is a Killing vector field equation (399) becomes

$$\mathbf{D}_\beta(\mathbf{D}_\alpha X_\lambda) = \mathbf{R}_{\lambda\alpha\beta\delta} X^\delta \tag{401}$$

and this implies, in view of the theorem of existence and uniqueness for ordinary differential equations, that any Killing vector field is completely determined by the $\frac{1}{2}(n+1)(n+2)$ values of X and $\mathbf{D}X$ at a given point. Indeed let p, q be two points connected by a curve $x(t)$ with tangent vector T. Let $L_{\alpha\beta} \equiv \mathbf{D}_\alpha X_\beta$, Observe that along $x(t)$, X, L verify the system of differential equations

$$\frac{\mathbf{D}}{dt} X = T \cdot L \quad , \quad \frac{\mathbf{D}}{dt} L = \mathbf{R}(\cdot, \cdot, X, T)$$

therefore the values of X, L along the curve are uniquely determined by their values at p.

The n-dimensional Riemannian manifold which possesses the maximum number of Killing vector fields is the Euclidean space $\mathbb{R}^n$. Simmilarily the Minkowski spacetime $\mathbb{R}^{n+1}$ is the Lorentzian manifold with the maximum numbers of Killing vectorfields.

0.16. Laplace-Beltrami operator. The scalar Laplace-Beltrami operator on a pseudo-riemannian manifold $\mathbf{M}$ is defined by,

$$\Delta_{\mathbf{M}} u(x) = g^{\mu\nu} \mathbf{D}_\mu \mathbf{D}_\nu u \tag{402}$$

where u is a scalar function on $\mathbf{M}$. Or, in local coordinates,

$$\Delta_{\mathbf{M}} u(x) = \frac{1}{\sqrt{|g(x)|}} \partial_\mu (g^{\mu\nu} \sqrt{|g(x)|} \partial_\nu) u(x) \tag{403}$$

The Laplace-Beltrami operator is called D'Alembertian in the particular case of a Lorentzian manifold, and is then denoted by $\Box_{\mathbf{M}}$. On any pseudo-riemannian manifold, $\Delta_{\mathbf{M}}$ is symmetric relative to the following scalar product for scalar functions u, v:

$$(u, v)_{\mathbf{M}} = \int u(x) v(x) dv_{\mathbf{M}}$$

Indeed the following identities are easily established by integration by parts, for any two smooth, compactly supported[8] functions u, v,

$$(-\Delta u, v)_{\mathbf{M}} = \int_{\mathbf{M}} \nabla u \cdot \nabla v \, dv_{\mathbf{M}} = (u, -\Delta v)_{\mathbf{M}} \tag{404}$$

where $\nabla u \cdot \nabla v = g^{ij}\partial_i u \partial_j v$. In the particular case when $u = v$ we derive, $(-\Delta u, v)_{\mathbf{M}} = \int_{\mathbf{M}} |\nabla u|^2$, with $|\nabla u|^2 = \nabla u \cdot \nabla u$. Thus, $-\Delta = -\Delta_{\mathbf{M}}$ is symmetric for functions $u \in \mathcal{C}_0^\infty(\mathbf{M})$. It is positive definite if the manifold $\mathbf{M}$ is Riemannian. This is not the case for Lorentzian manifolds: $\Box_{\mathbf{M}}$ is non-definite.

[8]This is automatically satisfied if the manifold $\mathbf{M}$ is compact.

Bibliography

[1] Lawrence C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, Volume 19. AMS, 1998.

[2] D. Foschi and S. Klainerman, *Bilinear space-time estimates for homogeneous wave equations*, Ann.Sci. École.Norm.Sup (4) **33** (200). no 2, 211-274.

[3] F.G. Friedlander, M.S. Joshi, *Introduction to the Theory of Distributions.* Cambridge University Press, 1998.

[4] I.M. Gelfand and G.E. Shilov *Generalized Functions Volume I: Properties and Operations.* Academic Press, 1964.

[5] Lars Hörmander, *The Analysis of Linear Partial Differential Operators*, Volume I. Springer-Verlag, 1983.

[6] S.W. Hawking, G.F.R. Ellis, *The Large Scale Structure of Space-Time* Cambridge University Press, 1973.

[7] Lars Hörmander, "Estimates for Translation Invariant Operators in L^p Spaces", Acta Math. 104(1960), 93-140.

[8] S. Klainerman and M. Machedon, *Space-Time Estimates for Null Forms and the Local Existence Theorem*, Communications on Pure and Applied Mathematics, Vol. 46, 1221-1268, (1993).

[9] Sergiu Klainerman and Igor Rodnianski, "Sharp trace theorems for null hypersurfaces on Einstein metrics with finite curvature flux". GAFA Geometric And Functional Analysis, vol. 16, no. 1, pp. 164229, 2006.

[10] Bilinear estimates on curved space-times, Jounal of Hyperbolic Di?. Eqts. vol 2, Nr 2 (2205), 279-291.

[11] S. Klainerman and S. Selberg, *Remark on the optimal regularity for equations of Wave Maps type*, Comm. in P.D.E's, Vol. 22 (5 & 6), (1997), 901-918.

[12] S. Klainerman and S. Selberg, *Bi-linear Estimates and Applications to Non-linear Wave Equations*, Communications in Contemporary Mathematics, bf 4 (2002), no 2, 223-295.

[13] S. Klainerman and D. Tataru, *On the optimal local regularity for the Yang-Mills equations in* $\mathbb{R}^{4+1}$, Journal of AMS, Vol 12, **1** 1999, 93-116.

[14] Markus Keel and Terrence Tao, "Endpoint Strichartz Estimates", *Am. J. Math.* 120 (1998), 955 980.

[15] L. Nirenberg, "On Elliptic Partial Differential Equations", *Ann. Sc. Norm. Pisa*, tome 13, no. 2 (1959), 115-162.

[16] Michael Renardy, Robert C. Rogers, *An Introduction to Partial Differential Equations.* Springer, 2004.

[17] Michael Spivak, *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, Inc.

[18] Elias Stein, *Singular Integrals and Differentiability Properties of Functions.* Princeton, 1976.