Mathematical Analysis III

Felix Chen

Contents

Teacher: Yang Shiwu

§1 Fourier analysis

§1.1 Fourier series

This concept comes from the heat equation in physics:

$$\partial_t u = \partial_{xx} u.$$

To solve this equation, first we assume $u(t,x) = A(t)B(x)$,

$$A'(t)B(x) = A(t)B''(x) \implies \frac{A'(t)}{A(t)} = \frac{B''(x)}{B(x)} = c.$$

where c must be a constant. Hence we get $A(t) = e^{ct}$. From physics knowledge we deduce $c < 0$, then we'll get a solution $u(t,x) = e^{-c^2 t}(A_c \cos(cx) + B_c \sin(cx))$. (We write $-c^2$ for original c)

If we put some requirements on boundaries, like $u(0,0) = u(0,\pi) = 0$, we'll get $A_c = 0, c \in \mathbb{Z}$.

Since the equation is linear, any linear combination of solutions are also solutions, thus the general solution can be written as

$$u(t,x) = \sum_{m=0}^{+\infty} e^{-m^2 t} B_m \sin(mx).$$

Since $u(0,x)$ can be measured in physics, we can solve all the B_m's, this completely solves the problem in physics.

This derives a problem in mathematics: Given a function $f(x)$ on $[0,2\pi]$, is it always possible to write $f(x)$ as series:

$$f(x) = \sum_{m=0}^{+\infty} (A_m \cos mx + B_m \sin mx).$$

First we assume $f(x)$ is written as this series, then

$$\int_0^{2\pi} f(x)\sin(kx)\,\mathrm{d}x = \sum_{k=0}^{\infty} A_m \int_0^{2\pi} \cos(mx)\sin(kx)\,\mathrm{d}x + B_m \int_0^{2\pi} \sin(mx)\sin(kx)\,\mathrm{d}x = \pi B_k.$$

This can be computed by tricks of trignometry functions.

Similarly,

$$\int_0^{2\pi} f(x)\cos(kx)\,\mathrm{d}x = \begin{cases} A_0 \cdot 2\pi, & k = 0; \\ A_k \cdot \pi, & k \neq 0. \end{cases}$$

To write these coefficients nicely, we'll generalize it to complex fields:

Definition 1.1.1 (Fourier series)**.** Let f be an integrable function on $[0,2\pi]$ or $[-\pi,\pi]$, define the **Fourier series** of f to be:

$$\hat{f}(k) = \frac{1}{2\pi}\int_0^{2\pi} f(x) e^{-ikx}\,\mathrm{d}x, \quad k \in \mathbb{Z}.$$

Thus we get

$$f(x) \sim \sum_{k\in\mathbb{Z}} \hat{f}(k)e^{ikx}.$$

When f is periodic on $[0, 2\pi]$ (meaning that $f(0) = f(2\pi)$), we can just view it as a periodic function on $\mathbb{R}$.

Fourier series write functions in the "frequency space" to reveal the hidden properties in the oringinal space.

§1.2 Fourier series of smooth functions

Theorem 1.2.1 (The uniqueness of Fourier coefficients)

Let f be an integrable function on $[0, 2\pi]$, and $\hat{f}(k) = 0$, then $f = 0, a.e.$.

Proof. Since polynomial functions can be approximated by trig functions uniformly, thus

$$\int_0^{2\pi} f(x)P(x)\,\mathrm{d}x = 0$$

for any polynomial $P(x)$, hence $f = 0, a.e.$. □

This means that the operation of taking Fourier coefficient is injective.

Corollary 1.2.2

If f is continuous on $[0, 2\pi]$, and $\sum_{k\in\mathbb{Z}} |\hat{f}(k)| < +\infty$, then the partial sum $S_N(f)(x) = \sum_{|k|\le N} \hat{f}(k)e^{ikx}$ uniformly converges to $f(x)$.

Proof. Let $F(x) = \sum_{k\in\mathbb{Z}} \hat{f}(k)e^{ikx}$. Then F is continuous and periodic,

$$\begin{aligned}\hat{F}(k) &= \frac{1}{2\pi}\int_0^{2\pi} F(x)e^{-ikx}\,\mathrm{d}x \\ &= \frac{1}{2\pi}\sum_{l\in\mathbb{Z}}\int_0^{2\pi} \hat{f}(l)e^{ilx}e^{-ikx}\,\mathrm{d}x \\ &= \hat{f}(k).\end{aligned}$$

Therefore by continuity $F(x) = f(x)$. □

By Riemann-Lebesgue Lemma, f intergrable $\implies$

$$\lim_{k\to\infty}\int_0^{2\pi} f(x)e^{ikx}\,\mathrm{d}x = 0 \implies \hat{f}(k) \to 0.$$

Lemma 1.2.3

$f \in C^1 \implies |\hat{f}(k)| \le \frac{C}{1+|k|}$, $f \in C^2 \implies |\hat{f}(x)| \le \frac{C}{(1+|k|)^2}$.

This implies that if $f \in C^2$, $S_N(f)(x)$ uniformly converges to $f(x)$, i.e. $S_N(f) \rightrightarrows f$.

When f is not so nice, since

$$\begin{aligned} S_N(f)(x) &= \sum_{|k|\le N} \hat{f}(k)e^{ikx} \\ &= \sum_{k=-N}^{N} \frac{1}{2\pi}\int_0^{2\pi} f(\xi)e^{-ik\xi}\,\mathrm{d}\xi e^{ikx} \\ &= \frac{1}{2\pi}\int_0^{2\pi} f(\xi)\sum_{k=-N}^{N} e^{-ik(x-\xi)}\,\mathrm{d}\xi. \end{aligned}$$

Let $D_N(x) = \sum_{k=-N}^{N} e^{ikx} = \frac{\sin(N+\frac{1}{2})x}{\sin\frac{x}{2}}$, (When $x = 2k\pi$, $D_N(x) = 2N+1$) called the *Dirichlet kernel*.

Then $S_N(f)(x) = \frac{1}{2\pi} f * D_N$, where $*$ is the convolution.

Recall the results of *approximations to the identity*, if K_ε is an approximation to the identity, we have $\lim_{\varepsilon\to 0} f * K_\varepsilon = f(x)$ for the Lebesgue point of f.

But unfortunately, D_N is not an approximation to the identity:

$$\int_0^{2\pi} \frac{1}{2\pi} D_N(x)\,\mathrm{d}x = \frac{1}{2\pi}\sum_{|k|\le N}\int_0^{2\pi} e^{ikx}\,\mathrm{d}x = 1.$$

Also $|D_N(x)| \le 2N+1$, but $|D_N(x)| \le AN^{-1}|x|^{-2}$ doesn't hold.

Hence we need to introduce a different tool: **Fejer kernel**.

§1.3 Fejer kernel

Consider the Cesaro sum

$$\sigma_N(f) := \frac{1}{N}(S_0(f) + \cdots + S_{N-1}(f)) = \frac{\sum_{k=0}^{N-1} f * D_k}{N\cdot 2\pi}.$$

This will leads to the Fejer kernel:

$$F_N(x) = \frac{1}{N}\sum_{k=0}^{N-1}\frac{\sin(k+\frac{1}{2})x}{\sin\frac{x}{2}} = \frac{1}{N}\left(\frac{\sin\frac{N}{2}x}{\sin\frac{x}{2}}\right)^2.$$

(When $x = 2k\pi$, $F_N(x) = N$)

We can prove this Fejer kernel is indeed an approximation to the identity. Hence we get

$$f \text{ integrable} \implies \lim_{\varepsilon\to 0}\|F_N * f - f\|_{L^1} = 0,$$

and for Lebesgue point of f we have

$$\lim_{N\to\infty}\sigma_N(f)(x) = f(x).$$

Note that the condition of approximation to the identity is too strong, as it applies to every L^1 function. When f has more regularity, we can loosen some conditions of the kernel.

Definition 1.3.1 (Good kernels)**.** Let $K_N(x)$ be continuous functions on $[-\pi,\pi]$, $K_N(x)$ is a **good kernel** if:

- $\int_{-\pi}^{\pi} K_N(x)\,\mathrm{d}x = 1$;
- $\exists M > 0$ s.t.
$$\int_{-\pi}^{\pi} |K_N(x)|\,\mathrm{d}x \le M, \quad \forall N;$$
- $\forall \delta > 0$,
$$\lim_{N\to+\infty} \int_{\delta\le|x|\le\pi} |K_N(x)|\,\mathrm{d}x = 0.$$

Theorem 1.3.2

Let K_N be a good kernel, f is a bounded integrable function on $[-\pi,\pi]$, if f is continuous at x, we have
$$\lim_{N\to+\infty} f * K_N(x) = f(x).$$

Proof. Let the bound of f be M_1.
$$\begin{aligned} f * K_N - f &= \int_{-\pi}^{\pi} (f(x-y) - f(x))K_N(y)\,\mathrm{d}y \\ &= \int_{|y|\le\delta} |f(x-y) - f(x)||K_N(y)|\,\mathrm{d}y + \int_{|y|\ge\delta} 2M_1|K_N(y)|\,\mathrm{d}y \to 0. \end{aligned}$$
□

Corollary 1.3.3

Let f be a continuous periodic function on $[0,2\pi]$,
$$\frac{1}{2\pi} f * F_N \rightrightarrows f.$$

Proof. Since F_N is a good kernel, f continuous $\Longrightarrow$ f uniformly continuous. Repeat the proof above and we'll get the result. □

$$F_N(x) = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{|l|\le k} e^{ilx} = \frac{1}{N}\sum_{|l|\le N-1}(N-|l|)e^{ilx}.$$

Thus
$$\sigma_N(f)(x) = \frac{1}{2\pi}(F_N * f)(x) = \frac{1}{2\pi}\sum_{|k|\le N-1}\left(1 - \frac{|k|}{N}\right)\hat{f}(k)e^{ikx}.$$

Combining with our corrolary, we can approach any continuous function uniformly using trig functions.

Now we assume f integrable and periodic on $[-\pi,\pi]$, WLOG $\int_{-\pi}^{\pi} f(x)\,\mathrm{d}x = 0$. (otherwise minus a constant on f, only $\hat{f}(0)$ will change)

Let $F(x) = \int_0^x f(t)\,\mathrm{d}t$, $F(x)$ is continuous and periodic on $[0, 2\pi]$. Consider the Fourier series

$$\hat{F}(0) = \frac{1}{2\pi}\int_0^{2\pi} F(x)\,\mathrm{d}x = \frac{1}{2\pi}\int_0^{2\pi}\int_0^x f(t)\,\mathrm{d}t\,\mathrm{d}x = \frac{1}{2\pi}\int_0^{2\pi} f(t)(2\pi - t)\,\mathrm{d}t = -\frac{1}{2\pi}\int_0^{2\pi} tf(t)\,\mathrm{d}t.$$

$$\begin{aligned}
\hat{F}(k) &= \frac{1}{2\pi}\int_0^{2\pi} F(x)e^{-ikx}\,\mathrm{d}x \\
&= \frac{1}{2\pi}\int_0^{2\pi}\int_0^x f(t)\,\mathrm{d}t\cdot e^{-ikx}\,\mathrm{d}x \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(t)\int_t^{2\pi} e^{-ikx}\,\mathrm{d}x\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(t)\frac{1-e^{-ikt}}{-ik}\,\mathrm{d}t = \frac{1}{ik}\hat{f}(k).
\end{aligned}$$

Now we have

$$\begin{aligned}
\frac{1}{2\pi}(F_N * F)(0) &= \sum_{|k|\le N-1}\left(1 - \frac{|k|}{N}\right)\hat{F}(k) \\
&= -\frac{1}{2\pi}\int_0^{2\pi} tf(t)\,\mathrm{d}t + \sum_{1\le|k|\le N-1}\left(1-\frac{|k|}{N}\right)\frac{\hat{f}(k)}{ik} \\
&= -\frac{1}{2\pi}\int_0^{2\pi} tf(t)\,\mathrm{d}t + \sum_{1\le|k|\le N-1}\frac{\hat{f}(k)}{ik} + \sum_{1\le|k|\le N-1} i\,\mathrm{sgn}(k)\frac{\hat{f}(k)}{N}.
\end{aligned}$$

Since $\hat{f}(k) \to 0$ (by Riemann-Lebesgue), and $\frac{1}{2\pi}(F_N * F)(0) \to F(0) = 0$.

$$\lim_{N\to+\infty}\sum_{1\le|k|\le N}\frac{\hat{f}(k)}{k} = \frac{i}{2\pi}\int_0^{2\pi} tf(t)\,\mathrm{d}t.$$

Thus we have a stronger condition on $\hat{f}(k)$.

Theorem 1.3.4 (Fatou)

Let

$$a_k = \begin{cases} 0, & |k| \le 1, \\ \dfrac{1}{2i\log k}, & k \ge 2, \\ -\dfrac{1}{2i\log|k|}, & k \le -2. \end{cases}$$

Then there doesn't exist an integrable function f on $[0, 2\pi]$ s.t. $\hat{f}(k) = a_k$.

Proof. We've proven $\hat{f}(k) \to 0$ and $\sum_k \frac{\hat{f}(k)}{k}$ converges, while $\{a_k\}$ does not satisfy these condition. □

§1.4 The convergence of Fourier series

In this section we come to the main problem of Fourier series, i.e. when and how does the convergence $S_N(f)(x) \to f(x)$ holds?

Lemma 1.4.1 (Localization lemma)
Let f be a integrable periodic function. For all $x_0 \in [-\pi, \pi]$, $\forall \delta > 0$,

$$\lim_{N\to+\infty} \int_{\delta\le|y|\le\pi} D_N(y) f(x-y)\, \mathrm{d}y = 0.$$

Proof.

$$\int_{|y|\ge\delta} \frac{\sin(N+\frac{1}{2})y}{\sin\frac{y}{2}} f(x_0-y)\,\mathrm{d}y \le \frac{1}{\sin\frac{\delta}{2}} \int_{|y|\ge\delta} \sin\left(N+\frac{1}{2}\right) y f(x_0-y)\,\mathrm{d}y \to 0.$$

Where the last step is by Riemann-Lebesgue Lemma. □

Therefore

$$2\pi S_N(f)(x) = D_N * f = \int_{|y|\le\delta} D_N(y) f(x-y)\,\mathrm{d}y + \int_{|y|\ge\delta} D_N(y) f(x-y)\,\mathrm{d}y,$$

§1.4.1 Point-wise convergence

Theorem 1.4.2 (Dini)
f same as above, if there exists a constant c and $\delta > 0$ s.t. at x_0 we have

$$\int_0^\delta \frac{|f(x_0-t)+f(x_0+t)-2c|}{t}\,\mathrm{d}t < +\infty.$$

Then $\lim_{N\to+\infty} S_N(f)(x_0) = c$.

Proof. By Riemann-Lebesgue,

$$\begin{aligned}
S_N(f)(x_0) - c &= \frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(y)(f(x-y)-c)\,\mathrm{d}y \\
&= \frac{1}{2\pi}\int_{-\delta}^{\delta} D_N(y)(f(x-y)-c)\,\mathrm{d}y + \frac{1}{2\pi}\int_{|y|\ge\delta} D_N(y)(f(x-y)-c)\,\mathrm{d}y \\
&= \frac{1}{2\pi}\int_0^\delta \frac{t\sin(N+\frac{1}{2})t}{\sin\frac{t}{2}} \frac{f(x-t)+f(x+t)-2c}{t}\,\mathrm{d}t + \frac{1}{2\pi}\int_{|y|\ge\delta} D_N(y)(f(x-y)-c)\,\mathrm{d}y \\
&\to 0.
\end{aligned}$$

Because the first term is the integral of a product of trig function and L^1 function, the second term approaches 0 by lemma. □

There's a counter example which is continuous function whose Fourier series doesn't converge to itself at one point. It's known as *du Bois-Reymond* counter example.

Example 1.4.3

Let $f(x) = \sum_k \frac{e^{ikx}}{k}$.

$$\sum_{1\le |k|\le N} \frac{e^{ikx}}{k} = \sum_{k=1}^{N} \frac{\sin kx}{k} = \frac{\cos\frac{x}{2} - 1 - \frac{\cos(N+\frac{1}{2}x)-1}{N} - \sum_{k=2}^{N} \frac{\cos(k-\frac{1}{2})x-1}{k(k-1)}}{2\sin\frac{x}{2}}$$

$$= \frac{-2\sin^2\frac{x}{4} + \frac{2\sin^2(\frac{N}{2}+\frac{1}{4})x}{N} + \sum_{k=2}^{N} \frac{2\sin^2(\frac{k}{2}-\frac{1}{4})x}{k(k-1)}}{2\sin\frac{x}{2}}.$$

When $x \ge \delta$, we can prove $S_N(f)(x) \to f(x)$. When $x < \delta$, since $\sin^2\frac{2k-1}{4}x \le \min\{1, (\frac{2k-1}{4}x)^2\}$,

$$\sum_{k=2}^{N} \frac{2\sin^2(\frac{k}{2}-\frac{1}{4})x}{k(k-1)} \le \sum_{k<x^{-1}} \frac{(\frac{2k-1}{4}x)^2}{k(k-1)} + \sum_{k\ge x^{-1}} \frac{1}{k(k-1)} \le cx^2\cdot x^{-1} + cx \le 2cx.$$

Thus $S_N(f)(x) \le M$, where M is independent of x and N.

Another approach is to realize $\sum_{1\le|k|\le N} \frac{e^{ikx}}{k}$ as $\frac{1}{2}\int_0^x (D_N(t)-1)\,\mathrm{d}t$. (complicated)

Define $w_K(x) = e^{2iKx}\sum_{1\le|k|\le K} \frac{e^{ikx}}{k}$. Let $K_l = 3^{l^3}$. Define

$$f(x) = \sum_{l=1}^{\infty} l^{-2} w_{K_l}(x) = \sum_{l=1}^{\infty} l^{-2} \sum_{1\le|k|\le K_l} \frac{e^{i(kx+2K_l x)}}{k}.$$

We can check that f is continuous and periodic. Since $K_l \le k + 2K_l \le 2K_{l+1}$, this is a Fourier series.

By the uniformly bounded property of $w_K(x)$ (proved in the example above),

$$S_{2K_{l_0}}(f)(0) = l_0^{-2}\sum_{k=-1}^{-K_{l_0}} \frac{1}{k} + \sum_{l=1}^{l_0-1} l^{-2} w_{K_l}(0) = -l_0^{-2}\ln K_{l_0} + O(1) = l_0\ln 3 + O(1) \not\to f(0).$$

This counter example tells us that continuity can't ensure the convergence of Fourier series.

Corollary 1.4.4

If there exists constants $0<\alpha\le 1$, $c>0$, $\delta>0$ and c_+, c_-, s.t.

$$|f(x_0+t) - c_+| + |f(x_0-t) - c_-| \le c\cdot t^\alpha, \forall 0\le t<\delta.$$

Then $S_N(f)(x_0) \to \frac{c_+ + c_-}{2}$.

In particular if f is periodic and C^α-Holder continuous ($\alpha = 1$ corresponds to Lipschitz continuous), then $S_N(f)(x) \to f(x), \forall x \in [0, 2\pi]$.

Proof. Just check the conditions of Dini's theorem. □

Theorem 1.4.5

Let f be an integrable function, $|\hat{f}(k)| \leq \frac{C}{|k|}$, then at Lebesgue point x of f we have $S_N(f)(x) \to f(x)$.

Proof. Recall that $\sigma_N(f)(x) \to f(x)$ for Lebesgue points, and let $b_k = S_k - S_{k-1}$, by a homework problem in first semester, we only need to prove kb_k is bounded. This follows immediately from the condition $|\hat{f}(k)| \leq \frac{C}{|k|}$. □

Now we move on to the BV functions. Since BV functions can be decomposed to increasing functions, we only consider increasing functions f. WLOG $f(0+) = 0$, by localization lemma, (we change $(N + \frac{1}{2})$ to N for simplicity)

$$\int_0^\delta \frac{\sin Nt}{\sin \frac{t}{2}} f(t)\,\mathrm{d}t = \int_0^\delta \left(\frac{\sin Nt}{\sin \frac{t}{2}} - \frac{\sin Nt}{\frac{t}{2}}\right) f(t)\,\mathrm{d}t + 2\int_0^\delta \frac{\sin Nt}{t} f(t)\,\mathrm{d}t.$$

By Riemann-Lebesgue lemma we can see the first term is at most $C|f(\delta)|$.

$$\int_0^\delta \frac{\sin Nt}{t} f(t)\,\mathrm{d}t = \int_0^{N\delta} \frac{\sin t}{t} f\left(\frac{t}{N}\right)\mathrm{d}t.$$

Next we divide this integral to $[2k\pi, 2(k+1)\pi]$:

$$\begin{aligned}
\left|\int_{2k\pi}^{2(k+1)\pi} \frac{\sin t}{t} f(\frac{t}{N})\,\mathrm{d}t\right| &= \left|\int_0^\pi \sin t \left(\frac{f(\frac{2k\pi+t}{N})}{2k\pi+t} - \frac{f(\frac{2k\pi+\pi+t}{N})}{2k\pi+\pi+t}\right)\mathrm{d}t\right| \\
&\leq \int_0^\pi \sin t \left(\frac{f(\frac{2k+2}{N}\pi) - f(\frac{2k\pi}{N})}{(2k+1)\pi} + \frac{f(\frac{2k\pi+\pi}{N}2\pi)}{2k\pi(2k+2)\pi}\right)\mathrm{d}t \\
&\leq \left|f\left(\frac{2k+2}{N}\pi\right) - f\left(\frac{2k\pi}{N}\right)\right| + f(\delta)\frac{1}{k(2k+2)}.
\end{aligned}$$

§1.4.2 Uniform convergence

Theorem 1.4.6 (Jordan)

Let f be a real BV function on $[0, 2\pi]$, then $\forall x_0 \in [0, 2\pi]$ we have

$$\lim_{N\to+\infty} S_N(f)(x_0) = \frac{f(x_0+) + f(x_0-)}{2}.$$

Moreover if f is continuous, this convergence is uniform.

Proof. We only prove the uniform convergence since the first part is nearly trivial by now.

Let $g(t) = f(x+t) + f(x-t) - 2f(x)$, then g is uniformly continuous on $[0, 2\pi]$, and g is also BV since the total variation is independent of x.

The computation is too complicated... □

In fact this theorem can be proved similarly to Theorem 1.4.5, since we can prove $\hat{f}(k) \leq \frac{C}{|k|}$ for BV functions.

§1.4.3 Absolute convergence

Theorem 1.4.7 (Bernstein)

Let f be a C^α continuous, $\alpha \in (\frac{1}{2}, 1]$, then $S_N(f)$ is absolutely convergent, i.e.

$$\sum_{k\in\mathbb{Z}} |\hat{f}(k)| < +\infty.$$

Lemma 1.4.8 (Paseval's equality)

Let f be a C^α (complex) function,

$$\frac{1}{2\pi}\int_0^{2\pi} |f|^2 \,\mathrm{d}x = \sum_{k\in\mathbb{Z}} |\hat{f}(k)|^2.$$

Proof. Formally, since

$$f = \sum \hat{f}(k)e^{ikx} \implies |f|^2 = \sum \hat{f}(k)\overline{\hat{f}(l)}e^{i(k-l)x},$$

thus

$$\int_0^{2\pi} |f|^2 \,\mathrm{d}x = \sum \hat{f}(k)\overline{\hat{f}(l)}\int_0^{2\pi} e^{i(k-l)x}\,\mathrm{d}x = 2\pi \sum |\hat{f}(k)|^2.$$

Now we come to the strict proof. We have $S_N(f) \rightrightarrows f$, thus

$$\int_0^{2\pi} |S_N(f)|^2 \,\mathrm{d}x = 2\pi \sum_{|k|\le N} |\hat{f}(k)|^2,$$

which implies the result by taking the limit $N \to \infty$.

In fact, this also holds for $f \in L^2$ because

$$\int_0^{2\pi} |f - S_N(f)|^2 \,\mathrm{d}x = \int_0^{2\pi} |f|^2 \,\mathrm{d}x - 2\pi \sum_{|k|\le N} |\hat{f}(k)|^2.$$

□

Proof of Bernstein's theorem. Let $f(x) = \sum \hat{f}(k)e^{ikx}$, then

$$f(x+h) - f(x-h) = \sum \hat{f}(k)e^{ikx} \cdot 2i\sin kh.$$

By Paseval's equality,

$$\sum |\hat{f}(k)|^2 4\sin^2 kh = \frac{1}{2\pi}\int_0^{2\pi} |f(x+h) - f(x-h)|^2 \,\mathrm{d}x \le C|h|^{2\alpha}.$$

Let $h = \frac{\pi}{2^{p+1}}$, we have $4\sin^2 kh \ge 2$, where $2^{p-1} \le k \le 2^p$.

$$\implies \sum_{k=2^{p-1}}^{2^p} |\hat{f}(k)| \le C|h|^{2\alpha} = C2^{-2p\alpha}.$$

$$\implies \sum_{k=2^{p-1}}^{2^p} |\hat{f}(k)| \le \left(\sum_{k=2^{p-1}}^{2^p} |\hat{f}(k)|^2\right)^{\frac{1}{2}} \left(2^{p-1}\right)^{\frac{1}{2}} \le C \cdot 2^{-p\alpha} 2^{\frac{p}{2}} = C \cdot 2^{-(\alpha - \frac{1}{2})p}.$$

□

Remark 1.4.9 — When $\alpha = \frac{1}{2}$, there are counter examples.

Theorem 1.4.10 (Zygmurd)
Let f be a periodic C^α function, if f is BV then $S_N(f)$ is absolutely convergent.

Proof. Since

$$\begin{aligned} |f((n+1)h) - f((n-1)h)|^2 &= \left|\sum \hat{f}(k) e^{iknh} 2i \sin kh\right|^2 \\ &= \sum |\hat{f}(k)|^2 4 \sin^2 kh + \sum_{k \ne l} \hat{f}(k)\overline{\hat{f}(l)} e^{in(k-l)h} \sin kh \sin lh. \end{aligned}$$

Let $h = \frac{2\pi}{N}$, and take the sum with respect to n: When $k \ne l$, $\sum_{n=0}^{N-1} e^{in(k-l)h} = 0$. (roots of unity) (When $n \mid k - l$, we need a different approach)

$$\sum_{n=0}^{N-1} |f((n+1)h) - f((n-1)h)|^2 = N \sum_k |\hat{f}(k)|^2 4 \sin^2 kh.$$

Hence

$$\sum_k |\hat{f}(k)|^2 4 \sin^2 kh \le \frac{1}{N} C h^\alpha \sum_{n=0}^{N-1} |f((n+1)h) - f((n-1)h)| \le C h^{\alpha+1} \|f\|_{BV}.$$

The rest is the same as the proof of Bernstein's theorem. □

§1.5 Hilbert spaces

People discovered that if we want the pointwise convergence of the Fourier series, we need to put many requirements to the function. Therefore we wonder if we can proof some results for general functions.

Recall that Parseval's equality gives a map from L^2 space to $l^2 = \{\{a_k\}_{k\in\mathbb{Z}} : \sum |a_k|^2 < +\infty\}$. In fact this map is bijective by Cauchy's law of convergence. This realizes L^2 as a vector space of countable dimensions, hence we introduce the general theory of Hilbert space.

Definition 1.5.1 (Hilbert space)**.** A **Hilbert space** is a separable complete inner product space. Recall that separable means it has a countable dense subset.

The space $L^2([0, 2\pi])$ is a Hilbert space since we can assign the inner product

$$(f, g) = \int_0^{2\pi} f\overline{g} \, \mathrm{d}x.$$

By Holder's inequality, it's easy to check $L^2([0, 2\pi])$ is indeed a Hilbert space under this inner product.

Example 1.5.2

In $\mathbb{R}^d$, the usual inner product gives the structure of Hilbert space.

Also $L^2(X)$ is a Hilbert space for any measure space X, with inner product $(f, g) = \int_X fg \, \mathrm{d}\mu$.

Similarly the space l^2 we've just defined is also a Hilbert space.

Let H be a Hilbert space. We say two elements $f, g \in H$ is **orthogonal** if $(f, g) = 0$, denoted by $f \perp g$. In this case we have $\|f + g\|^2 = \|f\|^2 + \|g\|^2$.

Recall the definitions of orthogonal and orthonormal basis in finite dimensional spaces in linear algebra, we can also generalize them to Hilbert space:

Definition 1.5.3 (Orthonormal basis)**.** If there is a countable set $\{e_i\}$ s.t. $\|e_i\| = 1$, $e_i \perp e_j$ and the vector space spanned by $\{e_i\}$ are dense in H, then we say $\{e_i\}$ is an **orthonormal basis** of H.

Example 1.5.4

In $\mathbb{R}^d$, $e_k = (0, \ldots, 1, \ldots, 0)$ form an orthonormal space, where the 1 is in the k-th entry.

In l^2, $\{e_k\}_{k\in\mathbb{Z}}$ is also an orthonormal basis. By applying Fourier transformation, we have $\{\frac{1}{\sqrt{2\pi}} e^{ikx}\}$ is an orthonormal basis of $L^2([0, 2\pi])$.

In $L^2(\mathbb{R})$, there is an orthonormal basis

$$\{c_k^{-\frac{1}{2}} e^{-\frac{1}{2}x^2} H_k(x), k \geq 0\},$$

where $c_k = \sqrt{\pi} 2^k k!$, $H_k(x)$ are Hermite polynomials:

$$H_k(x) = e^{x^2} \frac{\mathrm{d}^k}{\mathrm{d}x^k}(e^{-x^2}).$$

Theorem 1.5.5

Given an orthonormal system $\{e_k\}$ on a Hilbert space H, TFAE:

(1) $\{e_k\}$ is an orthonormal basis;

(2) If $f \in H$, $f \perp e_k \implies f = 0$;

(3) For all $f \in H$, $S_N(f) \to f$ (the convergence is on norm distance induced by the inner product), where $S_N(f) = \sum_{i=1}^N (f, e_i) e_i$.

(4) For all $f \in H$, the Paseval equality holds:

$$\|f\|^2 = \sum_{i=1}^{\infty} |(f, e_i)|^2.$$

Proof. (1) $\implies$ $\forall \varepsilon > 0$, $\exists a_1, \ldots, a_k \in \mathbb{R}$ s.t.

$$\|f - \sum_{i=1}^{k} a_i e_i\| < \varepsilon.$$

Hence
$$\|f\|^2 + \sum_{i=1}^{k} a_i^2 - 2\sum_{i=1}^{k}(f, e_i)a_i < \varepsilon.$$
Therefore $f \perp e_k \implies f = 0$.

(2) $\implies$ (3):
$$\|f - S_N(f)\|^2 = \|f\|^2 + \sum_{i=1}^{N} |(f, e_i)|^2 - 2\sum_{i=1}^{N} |(f, e_i)|^2 \geq 0.$$
Hence $\|f\|^2 \geq \sum_{i=1}^{\infty} |(f, e_i)|^2$.

Note that $S_N(f)$ is a Cauchy sequence ($\sum_{i=m}^{n} |(f, e_i)|^2 \to 0$), it must converge to some $\tilde{f}$.

For any j, take $N > j$,
$$(f - \tilde{f}, e_j) = (f, e_j) - (\tilde{f} - S_N(f), e_j) - (S_N(f), e_j) = |(\tilde{f} - S_N(f), e_j)| \leq \|\tilde{f} - S_N(f)\|\|e_j\|.$$
This gives $f = \tilde{f}$ by (2).

(3) $\implies$ (4) is trivial, and (4) $\implies$ (1): $\forall f \in H$, $S_N(f) \to f$, therefore the linear combination of $\{e_k\}$ is dense in H. □

Theorem 1.5.6

All Hilbert spaces have orthonormal basis.

Proof. Since H is separable, there exists $\{f_k\}$ dense in H. WLOG $\{f_k\}$ is linearly independent.

We follow the process of Schmit orthogonalization: Let $e_1 = \frac{f_1}{\|f_1\|}$.

If $f_{k+1} - \sum_{i=1}^{k}(f_{k+1}, e_i) \neq 0$, let e_{k+1} be the normalized orthogonal vector. □

Theorem 1.5.7 (Riesz)

Let T be a linear map from a Hilbert space H to $\mathbb{R}$. If there exists a constant c s.t.
$$|T(x)| \leq c|x|.$$
Then $\exists x_0 \in H$ s.t. $T(x) = (x, x_0)$.

Proof. Consider $H_0 = \ker f$. It's a closed linear subspace of H since T is continuous.

There's a unique decomposition $x = x_0 + x_1$ where $x_0 \in H_0$, $x_1 \perp H_0$: let $u \in H$ s.t. $T(u) = 1$ and $u \perp H_0$, $x_1 = T(x)u$.

TODO □

§1.6 Uniform distribution problem

Definition 1.6.1. Let $\{\xi_k\}$ be a sequence on $[0, 1)$, if $\forall a, b \in [0, 1)$,
$$\lim_{n\to\infty} \frac{|\{k \leq n \mid \xi_k \in [a, b]\}|}{n} = b - a.$$
We say $\{\xi_k\}$ is **uniformly distributed** on $[0, 1)$.

Example 1.6.2

Let $\{x\} = x - [x]$, then for any $q \in \mathbb{Q}$, $\{\{kq\}\}$ is not uniformly distributed.

The sequence $\xi_k = \{(\frac{1+\sqrt{5}}{2})^k\}$ is not uniformly distributed, because $F_{k+1} = (\frac{1+\sqrt{5}}{2})^k + (\frac{1-\sqrt{5}}{2})^k$, so ξ_k converges to 0 as $k \to +\infty$.

Theorem 1.6.3 (Weyl uniform distribution law)

The followings are equivalent:

- $\{\xi_k\}$ is uniformly distributed;
- For all Riemann integrable function f,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(\xi_k) = \int_0^1 f(x)\,\mathrm{d}x.$$

- For all $l \neq 0$ we have

$$\lim_{n\to\infty} \left(\frac{1}{n}\sum_{k=1}^{n} e^{2\pi i l \xi_k}\right) = 0.$$

Remark 1.6.4 — The generalization of this theorem is Birkhoff ergodic theorem.

Proof. It's clear that (2) can imply the other two by taking $f = \chi_{[a,b]}$ or $f(x) = e^{2\pi i l x}$.

(1) $\implies$ (2) : For $\sigma > 0$, consider a partition $x_0 = 0 < x_1 < \cdots < x_N = 1$, $|x_{i+1} - x_i| < \sigma$. Define f_+ and f_- as

$$f_+ = \sum_{k=0}^{N-1} \sup_{y\in[x_k,x_{k+1}]} f(y)\cdot \chi_{[x_k,x_{k+1}]}.$$

and f_- changes the sup to inf.

We have $f_-(x) \le f(x) \le f_+(x)$, thus by uniform distribution,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \chi_{[a,b]}(\xi_k) = b - a = \int_0^1 \chi_{[a,b]}(x)\,\mathrm{d}x.$$

By linearity we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f_\pm(\xi_k) = \int_0^1 f_\pm(x)\,\mathrm{d}x \to \int_0^1 f(x)\,\mathrm{d}x.$$

(3) $\implies$ (2) : Since $e^{2\pi i l x}$ is surjective, and by conditions we have

$$\frac{1}{n}\sum_{k=1}^{n} e^{2\pi i l \xi_k} \to 0 = \int_0^1 e^{2\pi i l x}\,\mathrm{d}x.$$

By linearity, the polynomials of trig functions also satisfies (2). Thus C^2 functions satisfies (2). (Its Fourier series uniformly converges to itself)

To prove (1), we only need to prove the step functions $\chi_{[a,b]}$ satisfies (2), which can be approximated by C^2 functions by $f_- \le \chi_{[a,b]} \le f_+$. $\square$

Example 1.6.5

Given an irrational number σ, then $\xi_k = \{k\sigma\}$ is uniformly distributed on $[0,1)$.

We can prove it using Weyl's law:

$$\frac{1}{n}\sum_{k=1}^{n} e^{2\pi ilk\sigma} = \frac{1}{n}\frac{e^{2\pi il\sigma}-e^{2\pi il(n+1)\sigma}}{1-e^{2\pi il\sigma}} \to 0.$$

Example 1.6.6

Let $\sigma \in (0,1)$, $\xi_k = \{ak^\sigma\}$ for some $a \neq 0$. $\{\xi_k\}$ is uniformly distributed since (set $b = 2\pi la$)

Proof. By Weyl's law,

$$\left|\sum_{k=1}^{n} e^{2\pi ilak^\sigma}\right| \le \left|\sum_{k=1}^{n}\int_k^{k+1} e^{ibk^\sigma}-e^{ibx^\sigma}\,\mathrm{d}x\right| + \left|\int_1^{n+1} e^{ibx^\sigma}\,\mathrm{d}x\right|$$

Note that $|e^{ibk^\sigma}-e^{ibx^\sigma}| \le |b|(x^\sigma - k^\sigma)$ is bounded,

$$\begin{aligned}
LHS &\le \sum_{k=1}^{n}\int_k^{k+1}|b|(x^\sigma-k^\sigma)\,\mathrm{d}x + \left|\int_1^{n+1}\frac{\mathrm{d}e^{ibx^\sigma}}{ibx^{\sigma-1}\sigma}\right| \\
&\le \sum_{k=1}^{n}((k+1)^\sigma - k^\sigma) + \frac{1}{|b|\sigma}|\cdots| \\
&\le |b|(n+1)^\sigma + c\left((n+1)^{1-\sigma}+1+\int_1^{n+1}(1-\sigma)x^{-\sigma}\,\mathrm{d}x\right) \\
&\le |b|(n+1)^\sigma + c(2(n+1)^{1-\sigma}+2) = o(n).
\end{aligned}$$

While $\xi_k = \{a\ln k\}$ is not uniformly distributed by similar computation. □

§2 Fourier transformation on discrete sets

§2.1 Basic theory

Given a positive integer N, let $\mathbb{Z}(N) = \{0,1,\ldots,n-1\}$ denote the residue class modulo N.

Let V be the space of complex value functions on $\mathbb{Z}(N)$, it's an N-dimensional vector space over $\mathbb{C}$. We can define the inner product on V to be

$$(f,g) = \frac{1}{N}\sum_{k=0}^{N-1} f(k)\overline{g}(k).$$

In particular $e^{\frac{2\pi i}{N}kx}$ is an orthonormal basis of V.

Therefore we define the Fourier transform to be

$$f(x) = \frac{1}{N}\sum_{k=0}^{N-1}\hat{f}(k)e^{\frac{2\pi i}{N}kx},\quad \hat{f}(k) = N(f, e^{\frac{2\pi i}{N}kx}) = \sum_{k=0}^{N-1} f(x)e^{-\frac{2\pi i}{N}kx}$$

for functions $f \in V$.

Note that the formula coincides with normal Fourier tranformations.

Futhermore, we can define the convolution on V,

$$f * g(x) = \sum_{k=0}^{N-1} f(x-k)g(k).$$

Similarly we have $\widehat{f * g}(k) = \hat{f}(k)\hat{g}(k)$.

The good thing of discreteness is that everything is finite, so we don't need to check convergence or commutativity.

In fact we have

$$\sup_k |\hat{f}(k)\hat{g}(k)| = \sup_k |\widehat{f * g}(k)| = \sup_k \left(\sum_{x=0}^{N-1} (f * g)(x) e^{-\frac{2\pi}{N} ikx} \right) \leq \sum_{x=0}^{N-1} \left| \sum_{y=0}^{N-1} f(x-y)g(y) \right|$$

which means that the "L^∞" norm is less than or equal to the convolution's "L^1" norm.

Proposition 2.1.1

Convolution preserves inner product, i.e. $(\hat{f}, \hat{g}) = N(f, g)$.

Proof. Just a bunch of computation. □

In particular, we have Paseval's equality:

Theorem 2.1.2

Let $f \in V$,

$$\frac{1}{N} \sum_{k=0}^{N-1} |\hat{f}(k)|^2 = (\hat{f}, \hat{f}) = N(f, f) = \sum_{k=0}^{N-1} |f(x)|^2$$

§2.2 Roth three-term arithmetic sequences

Theorem 2.2.1 (Roth)

Let $A \subset \mathbb{N}$, define the density of A to be

$$\overline{\rho}(A) = \limsup_{n \to \infty} \frac{|\{x \leq n | x \in A\}|}{n}.$$

If $\overline{\rho}(A) > 0$, then there exists a three-term arithmetic sequence in A, i.e. $\exists x, y, z \in A$ s.t. $x + z = 2y$.

The idea is to prove a following weaker statement:

Proposition 2.2.2

For $\forall 0 < \delta < 1$, if $N > e^{e^{50\delta^{-1}}}$, then for all $A \subset \{0, 1, \ldots, N-1\}$, if $|A| \geq \delta N$, there exists a three-term arithmetic sequence in A.

Proof. Assume by contradiction, take the smallest N s.t. $\exists 0<\delta<1$, $N>e^{e^{50\delta^{-1}}}$ and $|A|\geq \delta N$ without arithmetic sequences.

Step 1. Let $B=A\cap[\frac{N}{3},\frac{2N}{3})$. We'll prove $|B|\geq \frac{\delta}{4}N$.

Otherwise either $|A\cap[0,\frac{N}{3}]|$ or $|A\cap[\frac{2N}{3},N]|$ is at least $\frac{3}{8}\delta N$, WLOG $|A\cap[0,\frac{N}{3}]|$.

Therefore

$$\frac{|A\cap[0,\frac{N}{3}]|}{|[0,\frac{N}{3}]|}\geq \frac{\frac{3}{8}\delta N}{\frac{N}{3}+1}\geq \frac{N}{N+3}\frac{9}{8}\delta\geq \frac{10}{9}\delta.$$

Since $|[0,\frac{N}{3}]|\geq \frac{N}{3}\geq \frac{1}{3}e^{e^{50\delta^{-1}}}$, it is larger than $e^{e^{50(\frac{10}{9}\delta)^{-1}}}$.

Step 2. If $x,z\in B$, $x+y\equiv 2z(\mathrm{mod}N)$, then $x+y=2z$. ($x+y-2z=0,N,-N$, a simple inequality will yield the result)

Hence

$$\begin{aligned}
0&=\sum_{z\in B}\sum_{x\in B}\sum_{y\in A}\frac{1}{N}\sum_{k=0}^{N-1}e^{-\frac{2\pi}{N}ik(x+y-2z)}\\
&=\frac{1}{N}\sum_{k=0}^{N-1}\sum_{z\in B}e^{-\frac{2\pi}{N}ik(-2z)}\sum_{x\in B}e^{-\frac{2\pi}{N}ikx}\sum_{y\in A}e^{-\frac{2\pi}{N}iky}\\
&=\frac{1}{N}\sum_{k=0}^{N-1}\widehat{\chi_B}(-2k)\widehat{\chi_B}(k)\widehat{\chi_A}(k).
\end{aligned}$$

Note that $\widehat{\chi_A}(0)=\sum\chi_A(x)=|A|$, set $c=\max_{1\leq k\leq N-1}|\widehat{\chi_A}(k)|$.

$$\begin{aligned}
\frac{1}{N}|B|^2|A|&=\left|\frac{1}{N}\sum_{k=1}^{N-1}\widehat{\chi_B}(-2k)\widehat{\chi_B}(k)\widehat{\chi_A}(k)\right|\\
&\leq\frac{1}{N}\left|\sum_{k=1}^{N-1}\widehat{\chi_B}(-2k)\widehat{\chi_B}(k)\right|c\\
&\leq\frac{1}{N}\sum_{k=1}^{N-1}\left(|\widehat{\chi_B}(k)|^2+\frac{1}{4}|\widehat{\chi_B}(-2k)|^2\right)c.
\end{aligned}$$

Now by Paseval's equality $\frac{1}{N}\sum_{k=0}^{N-1}|\widehat{\chi_B}(k)|^2=|B|$ and $\sum|\widehat{\chi_B}(-2k)|^2\leq 2\sum|\widehat{\chi_B}(k)|$,

$$|B|^2|A|\leq\frac{3}{2}c\sum_{k=0}^{N-1}|\widehat{\chi_B}(k)|^2=\frac{3}{2}N|B|c.$$

Therefore $\exists k_0\neq 0$ s.t.

$$|\widehat{\chi_A}(k_0)|\geq\frac{2}{3}\frac{1}{N}|B||A|\geq\frac{2}{3}\frac{1}{N}\frac{\delta}{4}N\cdot\delta N=\frac{\delta^2}{6}N.$$

Step 3. $\exists 0<d\leq\sqrt{N}$ s.t. $dk_0\equiv c_0(\mathrm{mod}N)$, $0\leq c_0\leq\sqrt{N}$.

Let $l_0=\min\{[\frac{N}{6c_0}],[\frac{N-1}{2d}]\}$, consider

$$P=\{0,d,2d,\cdots,l_0d,N-d,N-2d,\cdots,N-l_0d\}.$$

We know $|P|\geq 2l_0+1\geq 2[\frac{\sqrt{N}}{6}]+1$, $c_0l_0\leq\frac{N}{6}$.

Since $|\widehat{\chi_A}(k_0)| = |\widehat{(\chi_A - c)}(k_0)|$ $(k_0 \neq 0)$,

$$\widehat{\chi_P}(k_0) = \sum_{|l|\leq l_0} e^{-\frac{2\pi i}{N}k_0 ld} = 1 + 2\sum_{l=0}^{l_0-1}\cos\left(\frac{2k_0\pi}{N}ld\right) \geq l_0 + 1.$$

We write

$$\begin{aligned}\frac{\delta^2}{6}N(l_0+1) &\leq |\widehat{(\chi_A - c)}(k_0)\widehat{\chi_P}(k_0)| \\ &\leq \sum_{x=0}^{N-1}\left|\sum_{y=0}^{N-1}(\chi_A(x-y) - c)\chi_P(y)\right| \\ &= \sum_{x=0}^{N-1}\left|\sum_{y=0}^{N-1}(\chi_A(y)\chi_P(x-y) - c|P|)\right|.\end{aligned}$$

Let $c = \frac{|A|}{N} \geq \delta$, then

$$\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\chi_A(y)\chi_P(x-y) - cN|P| = |A||P| - cN|P| = 0.$$

There exists x_0 such that

$$|A \cap (x_0 - P)| = \sum_{y=0}^{N-1}\chi_A(y)\chi_P(x_0 - y) \geq c|P| + \frac{\delta^2}{12}(l_0+1).$$

Note that $x_0 - P$ is a subset of $\mathbb{Z}(N)$ consisting of at most 2 arithmetic sequences of common difference d, say $x_0 - P = P_1 \cup P_2$, WLOG $|P_1| \geq |P_2|$, $P_1 = \{x_0 - p \geq 0\}, P_2 = \{x_0 - p < 0\}$.

- When $|P_2| \leq \frac{\delta^2}{48}|P|$, note $|P| = 2l_0 + 1$,

$$|A \cap P_1| \geq \delta|P| + \frac{\delta^2}{12}(l_0+1) - \frac{\delta^2}{48}|P| \geq \left(\delta + \frac{\delta^2}{48}\right)|P_1|.$$

  Since $A \cap P_1$ does not contain arithmetic sequences, and

$$|P_1| \geq \frac{47}{48}|P| \geq \frac{47}{48}\left(2\left[\frac{\sqrt{N}}{6}\right] + 1\right) \geq \frac{5}{16}\sqrt{N} \geq \frac{5}{16}e^{\frac{1}{2}e^{50\delta^{-1}}} > e^{e^{50(\delta+\frac{\delta^2}{48})^{-1}}}.$$

  We've find a smaller $N' = |P_1|$ and $\delta' = \delta + \frac{\delta^2}{48}$ contradicting the assumption.

- $|P_1| \geq |P_2| \geq \frac{\delta^2}{48}|P|$, $\exists T$ $(T = P_1$ or $T = P_2)$ s.t.

$$|A \cap T| \geq \left(\delta + \frac{\delta^2}{24}\right)|T|.$$

  By similar process in previous case, we can prove $N' = |T|$ and $\delta' = \delta + \frac{\delta^2}{24}$ suffices.

□

§2.3 Fourier transform on finite abelian groups

The set $\{0, 1, \dots, N-1\}$ can be viewed as a cyclic group of order N. In this section we'll generalize the idea to finite abelian groups, which can be viewed as $\mathbb{Z}$-modules.

Let G be a finite abelian group, here we use the multiplication convention of group operation and denote the identity as 1. The complex-valued functions defined on G form a vector space V of dimension $|G|$.

Define the inner product

$$(f, g) = \frac{1}{|G|} \sum_{a \in G} f(a)\overline{g(a)}.$$

The next step is to find an orthonormal basis. In the case of cyclic groups, this was easy; In general cases, recall that G is isomorphic to a direct product of cyclic groups, so we can do similar things.

However, we won't proceed so here as this approach is somewhat complicated. Instead, we'll use the charaters of G.

Definition 2.3.1 (Characters)**.** Let $e : G \to S^1$ be a homomorphism, i.e.

$$e(ab) = e(a)e(b), \quad \forall a, b \in G.$$

We say e is a **character** on G. The constant function 1 is called the trivial character.

Theorem 2.3.2

Let G be a finite abelian group, then all the characters on G form an orthonormal basis on V. Moreover if e is a nontrivial character, we have

$$\sum_{b \in G} e(b) = 0.$$

Proof. If $e_1 \neq e_2$ are two characters, then there exists $a \in G$ s.t.

$$e_1(a)\overline{e_2(a)} \neq 1.$$

(since $e_1(a)\overline{e_1(a)} = 1$ for all $a \in G$)

Now

$$\begin{aligned}
(e_1, e_2) &= \frac{1}{|G|} \sum_{b \in G} e_1(b)\overline{e_2(b)} \\
&= \frac{1}{|G|} \sum_{b \in G} e_1(ab)\overline{e_2(ab)} \\
&= \frac{1}{|G|} \sum_{b \in G} e_1(a)\overline{e_2(a)}e_1(b)\overline{e_2(b)} = e_1(a)\overline{e_2(a)}(e_1, e_2).
\end{aligned}$$

Thus the characters are pairwise orthogonal.

Hence it's sufficient to prove the number of characters equals $|G|$. (Actually the characters form a group $\widehat{G}$ called the dual group of G.)

By the classification theorem of finite abelian groups, $G = G_1 \times G_2 \times \cdots \times G_k$, where G_i's are cyclic groups. Since we can prove $\widehat{\mathbb{Z}_n} \simeq \mathbb{Z}_n$ and $\widehat{G \times H} \simeq \widehat{G} \times \widehat{H}$, it's clear that $|\widehat{G}| = |G|$. (In fact this is a homework problem of my algebra course) $\square$

Now all the preparations are done, we can define the Fourier transformation:

$$f = \sum_{e \in \widehat{G}} \widehat{f}(e)e, \quad \widehat{f}(e) = (f, e) = \frac{1}{|G|} \sum_{b \in G} f(b)\overline{e(b)}.$$

Similarly, $(f, g) = \sum_{e \in \widehat{G}} \widehat{f}(e)\overline{\widehat{g}(e)}$. This implies $(f, f) = \sum_{e \in \widehat{G}} |\widehat{f}(e)|^2$.

The convolution is defined as

$$(f * g)(a) = \frac{1}{|G|} \sum_{b \in \widehat{G}} f(b)g(ab^{-1}).$$

The goal of developing Fourier transformation on finite abelian groups is to prove the famous *Dirichlet's theorem*, which is often used in high school math olympiads.

Theorem 2.3.3 (Dirichlet's theroem)

Let q, l be two coprime integers. Then the arithmetic sequence $\{l + nq\}_{n=1}^{\infty}$ contains infinitely many primes.

The proof is very long, so sometimes we'll skip some of the computational details.

Consider the finite abelian group

$$G = \{1 \leq n < q \mid \gcd(n, q) = 1\} = \mathbb{Z}_q^{\times}.$$

For all $e \in \widehat{G}$, we can extend e to $\mathbb{Z}$ periodically, namely

$$\chi(m) = \begin{cases} e(n), & n \equiv m (\bmod q) \\ 0, & \text{otherwise.} \end{cases}$$

We call this χ **Dirichlet character**, satisfying $\chi(mn) = \chi(m)\chi(n)$.

Define the l-identicator $\delta_l(n) = \chi_{\{n: q | n-l\}}$. Apply the Fourier transform on δ_l,

$$\delta_l(n) = \sum_{e \in \widehat{G}} \frac{1}{|G|} \sum_{a \in \widehat{G}} \delta_l(a)\overline{e(a)}e(n) = \frac{1}{|G|} \sum_{e \in \widehat{G}} \overline{e(l)}e(n), \quad \forall n \in G.$$

Thus

$$\delta_l(n) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \overline{\chi(l)}\chi(n), \quad \forall n \in \mathbb{Z}.$$

For $s > 1$ and p prime,

$$|G| \sum_{p \equiv l (\bmod q)} \frac{1}{p^s} = |G| \sum_p \frac{\delta_l(p)}{p} = \sum_p \sum_{\chi \in \widehat{G}} \frac{\overline{\chi(l)}\chi(p)}{p^s} = \sum_{\chi \in \widehat{G}} \overline{\chi(l)} \sum_p \frac{\chi(p)}{p^s}.$$

Here the sum $\sum_p$ is taken over all primes not dividing q. By putting the trivial character outside the sum, we get

$$\sum_p \frac{1}{p^s} + \sum_{\chi \in \widehat{G}, \chi \neq 1} \overline{\chi(l)} \sum_p \frac{\chi(p)}{p^s}.$$

Now we'll prove two things:

- $\sum \frac{1}{p^s} \to +\infty$ as $s \to 1$.
- For nontrivial character χ, $\sum_p \frac{\chi(p)}{p^s}$ is finite.

These two things will imply $\sum_{p\equiv l(\bmod q)} \frac{1}{p^s} \to +\infty$, which is sufficient to prove Dirichlet's theorem.

Lemma 2.3.4

$$\lim_{s\to 1+} \sum_p \frac{1}{p^s} = +\infty.$$

Proof. By Euler's formula, for $s > 1$,

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1-p^{-s}} = \prod_p \left(\sum_{k=0}^{\infty} p^{-ks} \right).$$

Taking logarithm on both sides,

$$\ln \zeta(s) = -\sum_p \ln(1-p^{-s}) = \sum_p p^{-s} + \sum_p \frac{p^{-2s}}{2} + \cdots$$

Since for $k \geq 2$,

$$\sum_p \frac{p^{-ks}}{k} \leq \frac{1}{k} \sum_{n\geq 2} n^{-ks} \leq \frac{C}{k} 2^{1-ks}.$$

But since $\zeta(s) \to \infty$ as $s \to 1+$, $\sum_p p^{-s} \to +\infty$. □

Remark 2.3.5 — This implies $\sum_p \frac{1}{p} = +\infty$, which is an analytic proof of infinitely many primes.

The second step is much harder. For character χ, define $\mathcal{L}$ function

$$\mathcal{L}(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

Since for Dirichlet character χ, $|\chi| \leq 1$, so $\mathcal{L}(s,\chi)$ is convergent.

The computation below may not be strict, they just provide a perspective.

By Euler's formula,

$$\mathcal{L}(s,\chi) = \prod_p \frac{1}{1-\chi(p)p^{-s}}.$$

$$\ln \mathcal{L}(s,\chi) = -\sum \ln(1-\chi(p)p^{-s}) \approx \sum \frac{\chi(p)}{p^s}$$

and we need to prove it is bounded as $s \to 1+$.

Note that

$$\sum_{n=1}^{q} \chi(n) = \sum_{n\in G} e(n) = 0 \implies \left| \sum_{n=1}^{m} \chi(n) \right| < M, \quad \forall m \in \mathbb{Z}.$$

Thus by Abel's criterion $\mathcal{L}(s,\chi)$ is convergent for all $s > 0$.

Theorem 2.3.6

Let χ be a nontrivial character, $\mathcal{L}(1,\chi) \neq 0$.

We'll prove this later. As for now, we assume it is true.

We take the s derivative of $\mathcal{L}(s,\chi)$,

$$\mathcal{L}'(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \ln n.$$

We need to check RHS is uniformly convergent for $s \in [\delta, 2]$, $\forall \delta > 0$, so that the equality holds.

$$|\mathcal{L}(s,\chi) - 1| = \left| \sum_{n=2}^{\infty} \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right) \sum_{l \leq n} \chi(l) \right| \leq M2^{-s}.$$

Similarly we can prove $|\mathcal{L}'(s,\chi)| \leq M2^{-s}$.

Next we're going to explain why we're able to take logarithm. We can define

$$\ln \mathcal{L}(s,\chi) = -\int_s^{+\infty} \frac{\mathcal{L}'(t,\chi)}{\mathcal{L}(t,\chi)} \, \mathrm{d}t.$$

When $t > 1$, $\mathcal{L}(t,\chi) \neq 0$. Now we'll check $e^{\ln \mathcal{L}(s,\chi)} = \mathcal{L}(s,\chi)$. This follows by taking derivatives. When $s \to \infty$, they clearly equals to each other.

TODO

Proof of Theorem 2.3.6. TODO □

§3 Fourier transformation on $\mathbb{R}^n$

Recall that Fourier transform is to express a function by a linear combination of orthonormal basis of a Hilbert space (like $L^2([0,2\pi])$).

In $\mathbb{R}^n$, when function f has a compact support,

$$\widehat{f}(\xi) = \int_K f(x) e^{-2\pi i x \xi} \, \mathrm{d}x.$$

Thus by taking K to the limit, we'll get Fourier transformation on $\mathbb{R}^n$.

Definition 3.0.1. For $f \in L^1(\mathbb{R}^n)$, define

$$\mathcal{F}(f)(\xi) := \widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x \xi} \, \mathrm{d}x.$$

The number ξ is called the frequency, note the difference that ξ is continuous instead of discrete. Also we have the inverse transformation,

$$\mathcal{F}^{-1}(f)(x) := \overset{\vee}{f}(x) = \int_{\mathbb{R}^n} f(\xi) e^{2\pi i x \xi} \, \mathrm{d}\xi.$$

Observe that

$$|\widehat{f}(\xi)| \leq \int_{\mathbb{R}^n} |f(x)| \, \mathrm{d}x = \|f\|_{L^1}.$$

Theorem 3.0.2 (Riemann-Lebesgue lemma)
For $f \in L^1$,

$$\lim_{|\xi|\to+\infty} |\widehat{f}(\xi)| = 0.$$

Proof. Take $f_k \to f$ in L^1 s.t. f_k has compact support. □

Similarly for convolution we have

$$\widehat{f * g}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi).$$

Since $f, g \in L^1$, their convolution is also in L^1, thus by Fubini's theorem it's the same as before.

Remark 3.0.3 — Recall that we've proved there doesn't exist a function e s.t. $e * f = f$ for all f. Here we can use Fourier transform to give a simple proof: $e * f = f \implies \widehat{e} = 1$, which contradicts with Riemann-Lebesgue lemma.

Note that $\widehat{f}$ is not necessarily in L^1 space. If we look at the differential properties of $\widehat{f}$,

- $\mathcal{F}(\tau_{x_0} f)(\xi) = e^{2\pi i x_0 \xi}\widehat{f}(\xi)$, where $\tau_{x_0} f = f(x + x_0)$.
- $\mathcal{F}(D_\lambda f) = |\lambda|^{-d} D_{\lambda^{-1}}(\mathcal{F}(f))$, where $D_\lambda f = f(\lambda x)$.
- More generally, let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a nondegenerate linear transformation,

$$\mathcal{F}(f \circ A)(\xi) = |\det A|^{-1}\widehat{f}((A^{-1})^t \xi)$$

- $\widehat{\partial_k \varphi}(\xi) = 2\pi i \xi_k \widehat{\varphi}(\xi)$. This can be proved by integration by parts.
- $(\widehat{-2\pi i x_k}\varphi)(\xi) = \partial_{\xi_k}\widehat{\varphi}(\xi)$. This implies that $f \in C_0^\infty(\mathbb{R}^n) \implies \widehat{f} \in C^\infty(\mathbb{R}^n)$.
- The smooth properties of physical space is equivalent to the attenuation properties of frequency space.

The downside of Fourier transform is that whenever $f \in C_0^\infty$, we must have $\widehat{f} \notin C_0^\infty$.

Remark 3.0.4 — The proof of this fact requires knowledge of complex analysis. Let $\widehat{f}(z) = \int f(x)e^{-2\pi i x z}\, dx$ which is complex analytic, $\widehat{f} \in C_0^\infty(\mathbb{R})$ means it is zero on both ends of the real axis, so by maximal principle, $\widehat{f}$ must be zero on the whole complex plane.

Therefore we need to introduce a new function space:

§3.1 Schwartz space

Definition 3.1.1. If $f \in C^\infty$ satisfies

$$\sup_{x\in\mathbb{R}^n} \left| x^\alpha \left(\frac{\partial}{\partial x}\right)^\beta f(x) \right| < +\infty, \quad \forall \alpha, \beta,$$

then we say $f \in S(\mathbb{R}^n)$.

If $f \in S(\mathbb{R}^n)$, $\widehat{f} \in S(\mathbb{R}^n)$ as well.

Proposition 3.1.2

Let $f, g \in S(\mathbb{R}^n)$, we have

$$\int_{\mathbb{R}^n} \widehat{f}(\xi) g(\xi) \,\mathrm{d}\xi = \int_{\mathbb{R}^n} f(x) \widehat{g}(x) \,\mathrm{d}x.$$

Proof.

$$\begin{aligned}&\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x) e^{-2\pi i x\xi} \,\mathrm{d}x g(\xi) \,\mathrm{d}\xi \\ &= \int_{\mathbb{R}^n} f(x) \left(\int_{\mathbb{R}^n} g(\xi) e^{-2\pi i x\xi} \,\mathrm{d}\xi \right) \mathrm{d}x \\ &= \int_{\mathbb{R}^n} f(x) \widehat{g}(x) \,\mathrm{d}x.\end{aligned}$$

□

Theorem 3.1.3 (Fourier transform of Gaussian distribution)

Let $g(x) = e^{-\pi|x|^2}$ on $\mathbb{R}^n$, we have

$$\mathcal{F}(g)(\xi) = g(\xi) = e^{-\pi|\xi|^2}.$$

Proof. The one-dimensional case is easy to proof using complex analysis.

Let $G(\xi) = e^{\pi\xi^2} \widehat{g}(\xi)$ on $\mathbb{R}$. Then $G(0) = \widehat{g}(0) = \int_{\mathbb{R}} g(x) \,\mathrm{d}x = 1$.

$$G'(\xi) = e^{\pi\xi^2}(2\pi\xi)\widehat{g}(\xi) - e^{\pi\xi^2} i \int_{\mathbb{R}} e^{-2\pi i x\xi}(-2\pi i x\xi) g(\xi) = 0.$$

TODO □

Theorem 3.1.4

Let $f \in S(\mathbb{R}^n)$, then $\mathcal{F}^{-1}(\widehat{f}) = f$.

Proof. Note that here we can't simply apply Fubini's theorem as the function may not be integrable. Alternatively, we'll use Gaussian distribution to approach f.

For $\varepsilon > 0$, define $g_\varepsilon(x) = \varepsilon^{-d} g(\varepsilon^{-1}x)$. Where $g(x) = e^{-\pi|x|^2}$. Hence

$$\mathcal{F}(g_\varepsilon)(\xi) = \varepsilon^{-d}(\varepsilon^{-1})^{-d} \widehat{g}(\varepsilon\xi) = g(\varepsilon\xi).$$

$$\mathcal{F}(g(\varepsilon\cdot))(x) = \varepsilon^{-d} \widehat{g}(\varepsilon^{-1}x) = g_\varepsilon(x).$$

We have g_ε is an approximation to identity,

$$\lim_{\varepsilon\to 0} \|f * g_\varepsilon - f\|_{L^1} = 0, \quad \lim_{\varepsilon\to 0} (f * g_\varepsilon)(x) = f(x).$$

Therefore for function $f * g_\varepsilon$,

$$\int_{\mathbb{R}^n} \mathcal{F}(f * g_\varepsilon)(\xi) \,\mathrm{d}\xi = \int_{\mathbb{R}^n} f(x) g_\varepsilon(x) \,\mathrm{d}x = (f * g_\varepsilon)(0)$$

TODO □

Theorem 3.1.5 (Plancherel's formula)
Let $f \in S(\mathbb{R}^n)$, then $\|f\|_{L^2} = \|\widehat{f}\|_{L^2}$.

Proof. Let $g(x) = \overline{\widehat{f}(x)} = \int_{\mathbb{R}^n} \overline{f(\xi)} e^{2\pi i x\xi} \,\mathrm{d}\xi = \mathcal{F}^{-1}(\overline{f})(\xi)$. We have

$$\int_{\mathbb{R}^n} \widehat{f} g \,\mathrm{d}x = \int_{\mathbb{R}^n} f \widehat{g} \,\mathrm{d}x = \int_{\mathbb{R}^n} f \mathcal{F}\mathcal{F}^{-1}(\overline{f}) \,\mathrm{d}x = \|f\|_{L^2}^2.$$

□

Theorem 3.1.6
$\mathcal{F}^4 = \mathrm{id}$.

Proof. Since

$$\mathcal{F}(-\xi) = \mathcal{F}^{-1}(f)(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x(-\xi)},$$

we have $\mathcal{F}^2(f)(-\xi) = f(\xi)$, hence $\mathcal{F}^4(f)(\xi) = f(\xi)$. □

Note that $\mathcal{F} : S \to S$ is a linear map, we can talk about the eigenvalues of $\mathcal{F}$, by $\mathcal{F}^4 = \mathrm{id}$ we know the eigenvalues can only be ± 1 or $\pm i$.

Theorem 3.1.7 (Poisson summation formula)
Let $f \in S(\mathbb{R})$, we have

$$\sum_{n=-\infty}^{+\infty} f(x+n) = \sum_{n=-\infty}^{+\infty} \widehat{f}(n) e^{2\pi i n x}.$$

In paricular when $x = 0$ we get $\sum f(n) = \sum \widehat{f}(n)$.

Proof. Let $F(x) = \sum_{n=-\infty}^{+\infty} f(x+n)$ is a periodic function, hence $F(x) = \sum_{n=-\infty}^{+\infty} \widehat{F}(n) e^{2\pi i n x}$.
Note that

$$\begin{aligned}
\widehat{F}(n) &= \int_0^1 F(x) e^{-2\pi i n x} \,\mathrm{d}x \\
&= \int_0^1 \sum_{k=-\infty}^{+\infty} f(x+k) e^{-2\pi i n x} \,\mathrm{d}x = \widehat{f}(n).
\end{aligned}$$

□

Theorem 3.1.8 (Heisenberg uncertainty principle)
Let $f \in S(\mathbb{R})$, and $\|f\|_{L^2} = 1$.

$$4\pi \|xf\|_{L^2} \cdot \|\xi \widehat{f}(\xi)\|_{L^2} \geq 1.$$

The equality holds iff $f = Ae^{-Bx^2}$, $B > 0, A^2 = \sqrt{2B\pi^{-1}}$. More generally we have

$$4\pi \|(x - x_0) f\|_{L^2} \cdot \|(\xi - \xi_0)\widehat{f}(\xi)\|_{L^2} \geq 1.$$

Proof. Need to show

$$\int_{\mathbb{R}} x^2 f^2 \, \mathrm{d}x \int_{\mathbb{R}} \xi^2 |\widehat{f}|^2 \, \mathrm{d}\xi \geq \frac{1}{16\pi^2}.$$

Write

$$\begin{aligned}\int_{\mathbb{R}} \xi^2 |\widehat{f}|^2 \, \mathrm{d}\xi &= \frac{1}{4\pi^2} \int_{\mathbb{R}} |\widehat{(f')}(\xi)|^2 \, \mathrm{d}\xi \\ &= \frac{1}{4\pi^2} \int_{\mathbb{R}} |f'|^2(x) \, \mathrm{d}x.\end{aligned}$$

Where the last equality is by Plancherel's formula. The first equality is by

$$\widehat{f}'(\xi) = \int f'(x) e^{-2\pi i x\xi} \, \mathrm{d}x = 2\pi\xi \int f(x) e^{-2\pi i x \xi} \, \mathrm{d}x = 2\pi\xi \widehat{f}(\xi).$$

Now by Cauchy-Schwarz inequality,

$$LHS \geq \frac{1}{16\pi^2} \left(\int_{\mathbb{R}} x(f^2)' \, \mathrm{d}x \right)^2 = \frac{1}{16\pi^2}.$$

The equality holds when $xf = -2Bf'$, this gives the desired result. □

§3.2 Radon transformation

The Fourier transformation has applications in medicine like CT, MR. The mathematics behind it is "Fourier transformation" onto a plane.

In $\mathbb{R}^3$, let $\omega \in S^2$ be a unit vector, $t \in \mathbb{R}$. Define the plane

$$P_{t,\omega} = \{x \in \mathbb{R}^3 \mid x \cdot \omega = t\}.$$

Essentially t and ω is the distance and direction of the plane wrt the origin.

Definition 3.2.1. Let $f \in S(\mathbb{R}^3)$, define the **Radon transformation**

$$\mathcal{R}(f)(t,\omega) = \int_{P_{t,\omega}} f \, \mathrm{d}x.$$

i.e. the integral of f on a plane.

Fix ω, we can take $e_1, e_2 \in \mathbb{R}^3$ s.t. $\{\omega, e_1, e_2\}$ forms an orthonormal basis. Then

$$\int_{P_{t,\omega}} f \, \mathrm{d}x = \int_{\mathbb{R}^2} f(t\omega + u_1 e_1 + u_2 e_2) \, \mathrm{d}u_1 \, \mathrm{d}u_2.$$

Radon transformation has some relations with Fourier transformation.

Theorem 3.2.2 (Central Slice Theorem)
Let $r \in \mathbb{R}$, $\omega \in S^2$.

$$\widehat{f}(r\omega) = \int_{\mathbb{R}} \mathcal{R}(f)(t,\omega) e^{-2\pi i t r} \, \mathrm{d}t.$$

Proof.

$$\begin{aligned}
\int_{\mathbb{R}} \mathcal{R}(f)(t,\omega)e^{-2\pi itr}\,\mathrm{d}t &= \int_{\mathbb{R}}\int_{\mathbb{R}^2} f(t\omega+u_1e_1+u_2e_2)\,\mathrm{d}u_1\,\mathrm{d}u_2 e^{-2\pi itr}\,\mathrm{d}t \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}^2} f(t\omega+u_1e_1+u_2e_2)e^{-2\pi i(t\omega+u_1e_1+u_2e_2)\cdot r\omega}\,\mathrm{d}t\,\mathrm{d}u_1\,\mathrm{d}u_2 \\
&= \int_{\mathbb{R}^3} f(x)e^{-2\pi ix\cdot r\omega}\,\mathrm{d}x.
\end{aligned}$$

Where the last equality is a substitution $x = t\omega + u_1e_1 + u_2e_2$. □

Now we hope to rebuild f using the information of $\mathcal{R}(f)$. Since

$$\begin{aligned}
f(x) &= \int_{\mathbb{R}^3} \widehat{f}(\xi)e^{2\pi ix\cdot\xi}\,\mathrm{d}\xi \\
&= \frac{1}{2}\int_{-\infty}^{+\infty}\int_{S^2} \widehat{f}(r\omega)e^{2\pi ix\cdot r\omega}r^2\,\mathrm{d}r\,\mathrm{d}\omega \\
&= \frac{1}{2}\int_{-\infty}^{+\infty}\int_{S^2}\int_{\mathbb{R}} \mathcal{R}(f)(t,\omega)e^{-2\pi itr}r^2\,\mathrm{d}t\cdot e^{2\pi ix\cdot r\omega}\,\mathrm{d}r\,\mathrm{d}\omega.
\end{aligned}$$

Note that by Fourier transformation of derivatives,

$$\int_{\mathbb{R}} \mathcal{R}(f)(t,\omega)e^{-2\pi itr}r^2\,\mathrm{d}t = -\frac{1}{4\pi^2}\mathcal{F}_t(\partial_t^2\mathcal{R}(f))(r,\omega)$$

Therefore

$$f(x) = -\frac{1}{8\pi^2}\int_{S^2} \partial_t^2\mathcal{R}(f)(x\cdot\omega,\omega)\,\mathrm{d}\omega.$$

By an inverse Fourier transformation. This formula is useful in application since it's easy to compute.

§4 Sobolev space

Recall that for $\Omega \subset \mathbb{R}^n$, if $m(\Omega) < +\infty$, we have $L^p(\Omega) \subset L^q(\Omega)$ for $p \geq q$. We have a sequence of function spaces,

$$L^1 \supseteq L^2 \supseteq \cdots \supseteq L^\infty \supseteq C^0 \supseteq \cdots \supseteq C^\infty.$$

This gives us the insight that the more derivatives the function has, or the integral of higher power exists, the better properties it has. Sobolev's theorem reveals some "commutativity" of this derivatives and integrals.

> **Remark 4.0.1 —** In this section we won't state or prove everything strictly, since we didn't introduce the distribution, we have to state everything in the language of classical theory.

§4.1 Sobolev space

Let $\Omega \subset \mathbb{R}^n$ be an open set. For $1 \leq p < \infty$, recall that

$$L^p(\Omega) = \left\{ f : \left(\int_\Omega |f|^p\,\mathrm{d}x\right)^{\frac{1}{p}} < +\infty \right\}.$$

and the norm

$$\|f\|_{L^p} = \left(\int_\Omega |f|^p \, \mathrm{d}x \right)^{\frac{1}{p}}.$$

When $p = +\infty$,

$$\|f\|_{L^\infty} = \inf\{M : |\{|f| \geq M\}| = 0\}.$$

We know that $L^p(\Omega)$ is a separable Banach space for $1 \leq p < \infty$, while $L^\infty(\Omega)$ is not separable. Now we're adding the informations of derivatives.

Definition 4.1.1 (Sobolev space)**.** First we define the norm

$$\|f\|_{W^{k,p}} = \left(\sum_{|\alpha| \leq k} \int_\Omega |\partial^\alpha f|^p \, \mathrm{d}x \right)^{\frac{1}{p}}.$$

As usual we would naturally define $W^{k,p}$ as the functions with finite norm. Clearly $C^k \subset W^{k,p}$. However, if we do so it's hard to prove the space is complete, and we don't know what the other functions look like.

Note that if $\varphi \in C_0^\infty$, by integration by parts,

$$\int \partial^\alpha f \varphi \, \mathrm{d}x = (-1)^\alpha \int f \partial^\alpha \varphi \, \mathrm{d}x.$$

Definition 4.1.2 (Weak derivatives)**.** Let $f \in L^p(\Omega)$, for all $\varphi \in C_0^\infty$ we define $\partial^\alpha f$ to be a linear operator

$$\partial^\alpha f(\varphi) := (-1)^\alpha \int_\Omega f \partial^\alpha \varphi \, \mathrm{d}x = \int_\Omega \partial^\alpha f \varphi \, \mathrm{d}x.$$

In some cases, we can realize $\partial^\alpha f$ as a function s.t. the latter equality holds, i.e. realize this linear operator as an inner product.

Therefore we'll use the weak derivatives for the definition of Sobolev space, and denote $W_0^{k,p}(\Omega)$ as the completion of C_0^∞ under the $W^{k,p}$ norm. (This is to say $\partial^\alpha f$ exists and in $L^p(\Omega)$, moreover $\partial^\alpha f|_{\partial\Omega} = 0$.)

When $p = 2$, we also write $W^{k,2} = H^k$, and $W_0^{k,2} = H_0^k$.

By Plancheral's equality,

$$\sum_{|\alpha|=k} \|\partial^\alpha f\|_{L^2}^2 = \sum_{|\alpha|=k} \|\widehat{\partial^\alpha f}\|_{L^2}^2 = \sum_{|\alpha|=k} \|(2\pi \xi i)^\alpha \widehat{f}\|_{L^2}^2 = (2\pi)^k \sum_{|\alpha|=k} \||\xi|^k \widehat{f}\|_{L^2}^2.$$

So

$$\|f\|_{H^k}^2 = \sum_{|\alpha| \leq k} \|\partial^\alpha f\|_{L^2}^2 = \int (1 + |\xi|^2 + \cdots + |\xi|^{2k}) |\widehat{f}|^2 \, \mathrm{d}\xi \approx \|(1 + |\xi|^2)^{\frac{k}{2}} \widehat{f}\|_{L^2}^2.$$

Therefore we can define H^s space for $s \in \mathbb{R}$ as

$$\|f\|_{H^s} = \|(1 + |\xi|^2)^{\frac{s}{2}} \widehat{f}\|_{L^2}.$$

We also write $\dot{H}^s$, the homogenous space as

$$\|f\|_{\dot{H}^s} = \||\xi|^s \widehat{f}\|_{L^2}.$$

The space H^s is a complete Banach space, this is essentially from the completeness of L^2. We can think of $H^s(\Omega)$ as the restriction of H^s function on Ω s.t. and $f|_\Omega, \ldots, \partial^{[s]} f|_\Omega = 0$. (or more generally if Ω is not so good, the completion of $C_0^\infty(\Omega)$ function on H^s norm)

Remark 4.1.3 — If $\partial\Omega$ is a C^1 manifold, then the restriction definition coincides with the original definition.

Note that $H^0 = L^2$ and $H^{s'} \subset H^s$ for $s' \geq s$.

Theorem 4.1.4 (Sobolev embedding of H^s)
When $s \geq \frac{n}{2}$, $H^s(\mathbb{R}^n)$ can be continuously embedded into $L^\infty(\mathbb{R}^n)$, i.e. $\exists C$ independent of f s.t.

$$\|f\|_{L^\infty} \leq C\|f\|_{H^s}, \quad \forall f \in H^s.$$

Moreover, if $s - \frac{n}{2}$ is not an integer,

$$\|f\|_{C^{[s-\frac{n}{2}],\{s-\frac{n}{2}\}}} \leq C\|f\|_{\dot{H}^s}, \forall f \in \dot{H}^s(\mathbb{R}^n)$$

Proof. WLOG $0 < s - \frac{n}{2} < 1$. If $|x-y| \leq 1$, by Fourier transformation,

$$\begin{aligned}
|f(x) - f(y)| &\leq \int_{\mathbb{R}^n} |\widehat{f}(\xi)||e^{2\pi i(x-y)\xi} - 1| \, \mathrm{d}\xi \\
&\leq \left(\int_{\mathbb{R}^n} |\widehat{f}(\xi)||\xi|^{2s} \, \mathrm{d}\xi\right)^{\frac{1}{2}} \left(\int |\xi|^{-2s}|e^{2\pi i(x-y)\xi} - 1|^2 \, \mathrm{d}\xi\right)^{\frac{1}{2}} \\
&= \|\widehat{f}\|_{\dot{H}^s} \cdot 2\left(\int |\xi|^{-2s} \sin^2 \pi(x-y)\xi \, \mathrm{d}\xi\right)^{\frac{1}{2}}.
\end{aligned}$$

Note that the integral can be estimated as

$$\int_{|\xi|>\delta} |\xi|^{-2s} \, \mathrm{d}\xi + \int_{|\xi|\leq\delta} |\xi|^{-2s}\pi^2|x-y|^2|\xi|^2 \, \mathrm{d}\xi \leq \delta^{-2s+n} + |x-y|^2\delta^{2-2s+n} = |x-y|^{2s-n}.$$

Where we take $\delta = |x-y|^{-1}$. □

Here we take a little time to review.

- The space $W^{k,p}$ consists of functions whose weak derivatives $\partial^\alpha f$ are in L^p for all $\alpha \leq k$.
- The norm is defined as

$$\|f\|_{W^{k,p}(\Omega)} = \left(\sum_{0\leq\alpha\leq k} \int_\Omega |\partial^\alpha f|^p \, \mathrm{d}x\right)^{\frac{1}{p}}$$

- The space $W^{k,p}(\Omega)$ is a Banach space, this is derived from the completeness of $L^p(\Omega)$ and the definition of weak derivatives.
- $W_0^{k,p}(\Omega)$ is the completion of $C_0^\infty(\Omega)$ functions under the $W^{k,p}(\Omega)$ norm.
  If $\partial\Omega \in C^k$ (i.e. a C^k manifold), then $f \in W_0^{k,p} \iff \partial^\alpha f|_{\partial\Omega} = 0$, for all $\alpha \leq k-1$.

Proposition 4.1.5

The smooth functions $C^\infty(\Omega)$ is dense in $W^{k,p}(\Omega)$.

Sketch of proof. When $\Omega = \mathbb{R}^n$, this is easy, since we can take an approximation to identity K_ϵ, and $f * K_\epsilon \to f$.

When Ω is not $\mathbb{R}^n$, we can also define convolution on $\Omega \subset \mathbb{R}^n$ as follows:

First note that if $\operatorname{supp} f \subset \Omega_\epsilon = \{x \in \Omega : d(x, \partial\Omega) > \epsilon\}$, then $f * K_\epsilon$ is a function on Ω.

Therefore if we take a unit decompostion φ_k, $f\varphi_k$ supports on a compact subset of Ω, hence it's at least ϵ away from $\partial\Omega$.

Thus we can define $f\varphi_k * K_{\epsilon_k}$ for each k, and when $\epsilon_k \to 0$ uniformly, $\sum_k f\varphi_k * K_\epsilon \to \sum_k f\varphi_k = f$. □

§4.1.1 Extensions

Next we discuss the extension of f. Note that if $f \in W^{k,p}(\mathbb{R}^n)$, then $f|_\Omega \in W^{k,p}(\Omega)$. The inverse statement is not true, since the boundary of Ω can get wery complicated.

This question is related to *Whitney extension*. Given a function $f \in W^{k,p}(\Omega)$, we ask whether f can be extended to g on $\mathbb{R}^n$ s.t.

$$\|g\|_{W^{k,p}} \le C\|f\|_{W^{k,p}}, \quad g|_\Omega = f.$$

This is true when $\partial\Omega$ is a C^k manifold.

§4.1.2 Restrictions

Another question is about the restriction of functions. When $f \in C(\mathbb{R}^n)$, clearly $f(0, x) \in C(\mathbb{R}^{n-1})$. But if $f \in L^1(\mathbb{R}^n)$, the function $f(0, x)$ may not be measurable at all.

The question is under what conditions can we take the restrictions and get reasonably good properties.

The so-called *trace theorem* gives an answer to this question. When $f \in H^s(\mathbb{R}^n)$, $s \ge \frac{1}{2}$, then $f(0, x) \in H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$.

§4.2 The H^s function space

- In $H^s(\mathbb{R}^n)$, by Plancherel's formula,

$$\|f\|_{H^s} = \|(1 + |\xi|^2)^{\frac{s}{2}} \widehat{f}(\xi)\|_{L^2}, \quad \forall s \in \mathbb{R}.$$

  This is equivalent to the map

$$T : L^2(\mathbb{R}^n) \to H^s(\mathbb{R}^n), \quad f \mapsto \mathcal{F}^{-1}((1 + |\xi|^2)^{-\frac{s}{2}} f), \quad s \ge 0$$

  is an isometry of Banach spaces.

  For the case $s < 0$, think of $H^s(\mathbb{R}^n)$ as the dual space of $H^{-s}(\mathbb{R}^n)$.

- Consider the space $H^s(\Omega)$. When s is an integer, $H^s(\Omega) = W^{s,2}(\Omega)$.

- When $0 < s < 1$, define the H^s norm

$$\|f\|_{H^s(\Omega)}^2 = \|f\|_{L^2(\Omega)}^2 + \iint_{\Omega\times\Omega} \frac{|f(x) - f(y)|^2}{|x - y|^{n+2s}} \, \mathrm{d}x \, \mathrm{d}y.$$

  This norm will make $H^s(\Omega)$ a Banach space.

- For general $s > 0$, define
$$\|f\|_{H^s} = \|f\|_{H^{[s]}(\Omega)} + \iint_{\Omega\times\Omega} \frac{|\partial^k f(x) - \partial^k f(y)|^2}{|x-y|^{n+2\{s\}}} \,\mathrm{d}x\,\mathrm{d}y.$$

- At last we define $H^{-s}(\Omega)$ as the dual space of $H_0^s(\Omega)$. (won't use it in this course)

Recall the embedding theorem of H^s,

Theorem 4.2.1
When $s > \frac{n}{2}$, we have
$$\|f\|_{L^\infty(\mathbb{R}^n)} \leq C\|f\|_{H^s(\mathbb{R}^n)}, \quad \forall f \in H^s(\mathbb{R}^n).$$
If $s - \frac{n}{2}$ is not an integer,
$$\|f\|_{C^{[s-\frac{n}{2}],\{s-\frac{n}{2}\}}} \leq C\|f\|_{\dot{H}^s}, \quad \forall f \in H^s(\mathbb{R}^n).$$

Proof. To prove the first one,
$$\begin{aligned}\|f\|_{L^\infty} &\leq \|\widehat{f}\|_{L^1}\\ &\leq \|(1+|\xi|^2)^{\frac{s}{2}}\widehat{f}\|_{L^2} \cdot \|(1+|\xi|^2)^{-\frac{s}{2}}\|_{L^2}\\ &\leq C_{s,n}\|f\|_{H^s(\mathbb{R}^n)}.\end{aligned}$$

The second can be proved similarly using Fourier transformation and Holder's inequality. □

Let $f_\lambda(x) = f(\lambda x)$ be a scaling of f.

- $\|f_\lambda\|_{L^\infty} = \|f\|_{L^\infty}$.
- $\|f_\lambda\|_{\dot{H}^s} \approx \|\partial^s f_\lambda\|_{L^2} = \lambda^{s-\frac{n}{2}}\|f\|_{\dot{H}^s}$.
- $\|f_\lambda\|_{C^{0,\gamma}} = \sup \frac{|f_\lambda(x)-f_\lambda(y)|}{|x-y|^\gamma} = \lambda^\gamma \|f\|_{C^{0,\gamma}}$.

By replacing f to f_λ in the above theorem we can see why s must be strictly larger than $\frac{n}{2}$.

Theorem 4.2.2
When $s > \frac{n}{2}$, for all $f, g \in H^s(\mathbb{R}^n)$ we have
$$\|fg\|_{H^s} \leq C\|f\|_{H^s}\|g\|_{H^s}.$$

Proof. By Minkowski's inequality, (or by Fourier transformation)
$$\|f * g\|_{L^2} \leq C\|f\|_{L^1}\|g\|_{L^2}.$$

For all $s > 0$, we have
$$(1+|\xi|^2)^s \leq 2^{2s}((1+|\xi-\eta|^2)^s + (1+|\eta|^2)^s).$$

Hence

$$\begin{aligned}\|fg\|_{H^s} &= \|(1+|\xi|^2)^{\frac{s}{2}}\widehat{f} * \widehat{g}\|_{L^2} \\ &\le C\left\|\int (1+|\xi-\eta|^2)^s \widehat{f}(\xi-\eta)\widehat{g}(\eta)\,\mathrm{d}\eta\right\|_{L^2} + C\left\|\int (1+|\eta|^2)^s \widehat{g}(\eta)\widehat{f}(\xi-\eta)\,\mathrm{d}\eta\right\|_{L^2} \\ &\le C\|\widehat{g}\|_{L^1}\|(1+|\xi|^2)^s\widehat{f}(\xi)\|_{L^2} + \cdots \\ &\le C\|g\|_{H^s}\|f\|_{H^s}.\end{aligned}$$

□

Returning to the weak derivatives, since it's defined by integrals, we can change its value on any null sets.

The trace theorem tells us that for $f \in W^{k,p}(\Omega)$, there exists an operator $T : f \mapsto Tf \in L^p(\partial\Omega)$ s.t. if $f \in W^{k,p}(\Omega) \cap C(\overline{\Omega})$, then $Tf = f|_{\partial\Omega}$, with a bounded L^p norm. Note that here we put some more requirements on f, so this do not contradicts with the fact above.

§4.3 Sobolev embedding theorem

Earlier we saw a simple embedding theorem of H^s, in this section we'll handle general situations.

First recall that the **weak L^p space** L^p_w is defined as

$$L^p_w = \{f : \|f\|_{L^p_w} := \sup_\alpha(\alpha\Lambda^{\frac{1}{p}}(f,\alpha)) < +\infty\}.$$

Where $\Lambda(f,\alpha) := m(\{|f| > \alpha\})$.

If $f \in L^p$,

$$\begin{aligned}\|f\|^p_{L^p} &= p\int_0^{+\infty} \alpha^{p-1}\Lambda(f,\alpha)\,\mathrm{d}\alpha \\ &\ge \int_{|f|>\alpha} |f|^p\,\mathrm{d}x \ge \alpha^p\Lambda(f,\alpha)\end{aligned}$$

Therefore $\|f\|_{L^p} \ge \|f\|_{L^p_w}$.

Keep in mind that $\|\cdot\|_{L^p_w}$ is NOT a norm.

Theorem 4.3.1 (Marcinkiewicz)

Suppose T is a sub-linear operator, i.e. $\exists C$ constant s.t.

$$|T(f+g)| \le C(|T(f)| + |T(g)|), \quad |T(\lambda f)| = |\lambda T(f)|.$$

If T satisfies that

$$\|Tf\|_{L^{p_0}_w} \le M_0\|f\|_{L^{p_0}}, \quad \|Tf\|_{L^{p_1}_w} \le M_1\|f\|_{L^{p_1}},$$

then we have

$$\|Tf\|_{L^p} \le \gamma M_0^\theta M_1^{1-\theta}\|f\|_{L^p}.$$

Where $1 \le p_0 < p < p_1 \le +\infty$, and $\frac{1}{p} = \frac{\theta}{p_0} + \frac{1-\theta}{p_1}$, γ is a constant.

Proof. WLOG $C = 1$. We can decompose f into

$$f = f\chi_{|f|>\alpha} + f\chi_{|f|\le\alpha} =: f_1 + f_2.$$

It's clear that $f_1 \in L^{p_0}$, $f_2 \in L^{p_1}$.

Hence

$$\Lambda(Tf, 2\alpha) \le \Lambda(Tf_1, \alpha) + \Lambda(Tf_2, \alpha) \le M_0^{p_0}\alpha^{-p_0}\|f_1\|_{L^{p_0}}^{p_0} + M_1^{p_1}\alpha^{-p_1}\|f_2\|_{L^{p_1}}^{p_1}.$$

$$\begin{aligned}\int \alpha^{-p_0}\|f_1\|_{L^{p_0}}^{p_0}\alpha^{p-1}\,\mathrm{d}\alpha &= \int \alpha^{-p_0+p-1}\int_{|f|>\alpha}|f|^{p_0}\,\mathrm{d}x\,\mathrm{d}\alpha \\ &= \int |f|^{p_0}(p-p_0)^{-1}|f|^{p-p_0}\,\mathrm{d}x \\ &= (p-p_0)^{-1}\|f\|_{L^p}^p.\end{aligned}$$

Doing the same thing with the other term and use the fact that

$$\|Tf\|_{L^p}^p = p\int_0^\infty \Lambda(Tf, 2\alpha)(2\alpha)^{p-1}\,\mathrm{d}(2\alpha)$$

we'll get the result. $\square$

Recall that the maximal function of $f \in L^1(\mathbb{R}^n)$ is defined as

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

In fact this $\mathcal{M}$ is a sub-linear operator, since $\mathcal{M}(f+g) \le \mathcal{M}(f) + \mathcal{M}(g)$.

Also note that $\|\mathcal{M}(f)\|_{L^\infty} \le \|f\|_{L^\infty}$, and $\|\mathcal{M}(f)\|_{L^1_w} \le 3^n\|f\|_{L^1}$. By the above theorem we know

$$\|\mathcal{M}(f)\|_{L^p} \le C_p\|f\|_{L^p}, \quad 1 < p < +\infty.$$

However, in most cases the index of L^{p_0} and L^{p_1} are not the same on both sides, like $\mathcal{F}$, the Fourier transformation satisfying $\|\mathcal{F}(f)\|_{L^\infty} \le \|f\|_{L^1}$ and $\|\mathcal{F}(f)\|_{L^2} = \|f\|_{L^2}$, we hope to get $\|\mathcal{F}(f)\|_{L^p} \le \|\mathcal{F}(f)\|_{L^q}$. But this is beyond our capability for now.

Let $D = \{z \mid 0 < Re(z) < 1\}$.

Lemma 4.3.2 (Hadamard three lines lemma)

Let f be a bounded analytic function on D, and continuous on $\overline{D}$. If

$$|f(0+ib)| \le M_0, \quad |f(1+ib)| \le M_1, \forall b \in \mathbb{R}.$$

Then we have

$$|f(a+ib)| \le M_0^{1-a}M_1^a, \quad \forall 0 \le a \le 1, b \in \mathbb{R}.$$

Proof. WLOG $M_0, M_1 > 0$. (Otherwise it must be zero everywhere)

Let $F_\epsilon(z) = e^{-\epsilon(1-z)z}\frac{f(z)}{M_0^{1-z}M_1^z}$. Since $F_\epsilon(z)$ is continuous on $\overline{D}$, analytic on D,

$$|F_\epsilon(a+ib)| = e^{-\epsilon a(1-a)-\epsilon b^2}\frac{|f(a+ib)|}{M_0^{1-a}M_1^a}.$$

Consider $D_N = \overline{D} \cap \{|Im(z)| \le N\}$. By analytic property, the maximum and minimum of f must be acheived at the boundary. Take N large, we can prove $F_\epsilon \le 1$ on the boundary of D_N. Therefore $F_\epsilon \le 1$ for all $\epsilon > 0$ sufficiently small. $\square$

We say a linear operator T_z is **analytic** wrt z, if it maps simple functions to measurable functions, and $\int f(x)(T_z g)(x)\,\mathrm{d}x$ is an analytic function of z on D, and bounded continuous on $\overline{D}$.

Theorem 4.3.3 (Stein-Riesz-Thorin)
Let $0 < p_0, p \leq +\infty$, $1 \leq q_0, q_1 \leq +\infty$, T_z is an analytic linear operator satisfying

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

Then we have

$$\|T_{a+ib} f\|_{L^q} \leq M_0^a M_1^{1-a} \|f\|_{L^p}.$$

Here $0 < a < 1$, and $\frac{1}{q} = \frac{a}{q_0} + \frac{1-a}{q_1}$, $\frac{1}{p} = \frac{a}{p_0} + \frac{1-a}{p_1}$.

Theorem 4.3.4 (Riesz-Thorin)
Let T be a linear operator, if

$$\|Tf\|_{L^{q_0}} \leq M_0 \|f\|_{L^{p_0}}, \quad \|Tf\|_{L^{q_1}} \leq M_1 \|f\|_{L^{p_1}}.$$

We have $\forall f \in L^{p_0} \cap L^{p_1}$,

$$\|Tf\|_{L^q} \leq M_0^\theta M_1^{1-\theta} \|f\|_{L^p}.$$

for $\frac{1}{p} = \frac{\theta}{p_0} + \frac{1-\theta}{p_1}$, $\frac{1}{q} = \frac{\theta}{q_0} + \frac{1-\theta}{q_1}$, $\theta \in (0,1)$.

Proof. Take $T_z = T$, by above theorem the conclusion holds for simple functions.
For general f, take simple functions f_k s.t.

$$\lim_{k \to \infty} \|f - f_k\|_{L^{p_i}} = 0.$$

$$\begin{aligned}
\|T(f - f_k)\|_{L^q} &\leq \|T(f - f_k)\|_{L^{q_0}}^\theta \|T(f - f_k)\|_{L^{q_1}}^{1-\theta} \\
&\leq (M_0 \|f - f_k\|_{L^{p_0}})^\theta (M_1 \|f - f_k\|_{L^{p_1}})^{1-\theta} \to 0.
\end{aligned}$$

□

Theorem 4.3.5 (Young's inequality)
Let T be a linear operator defined as

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

If there exists constant C s.t.

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

Then $\|Tf\|_{L^q} \leq C\|f\|_{L^p}$ for $1 + \frac{1}{q} = \frac{1}{r} + \frac{1}{p}, 1 \leq p \leq r', 1 \leq q, r \leq +\infty$. (Here $r' = \frac{r}{r-1}$)
When $q = r, p = 1$, this is the integral version of Minkowski's inequality.

Remark 4.3.6 — The relation between p, q, r can be implied by a scaling transformation. Let $k_\lambda = k(\lambda x, \lambda y)$, $f_\lambda(x) = f(\lambda x)$. Then $\|f_\lambda\|_{L^p} = \lambda^{-\frac{n}{p}} \|f\|_{L^p}$, similarly C changes by $\lambda^{-\frac{n}{r}}$, $\|Tf\|_{L^q}$ changes by $\lambda^{-n-\frac{n}{q}}$.

When the operator is the convolution operator, i.e. $k(x, y) = k(x - y)$, then $Tf = k * f$, we have

$$\|k * f\|_{L^q} \le \|k\|_{L^r} \|f\|_{L^p}.$$

When $T = \mathcal{F}$ is the Fourier transformation, by Riesz-Thorin interpolation inequality,

$$\|\mathcal{F}(f)\|_{L^p} \le \|f\|_{L^{p'}}, \quad p' = \frac{p}{1-p}, p \ge 2.$$

Theorem 4.3.7 (Hardy-Littlewood-Sobolev)
Let $0 < \gamma < n$, $1 < p < q < +\infty$ satisfying $1 - \frac{\gamma}{n} = \frac{1}{p} - \frac{1}{q}$. We have

$$\|f * |\cdot|^{-\gamma}\|_{L^q} \le C_{p,q,n} \|f\|_{L^p}.$$

Proof. WLOG $f \ge 0$.

$$\begin{aligned} f * |\cdot|^{-\gamma} &= \int f(x-y)|y|^{-\gamma} \, \mathrm{d}y \\ &= \int_{|y| \ge R} f(x-y)|y|^{-\gamma} \, \mathrm{d}y + \int_{|y|<R} f(x-y)|y|^{-\gamma} \, \mathrm{d}y \\ &\le \|f\|_{L^p} \||y|^{-\gamma} \chi_{|y|>R}\|_{L^{p'}} + \sum_{k=0}^{\infty} \int_{2^{-k-1}R \le |y| < 2^{-k}R} f(x-y)|y|^{-\gamma} \, \mathrm{d}y. \end{aligned}$$

Note that

$$\left(\int_{|y|>R} |y|^{-\gamma p'} \, \mathrm{d}y \right)^{\frac{1}{p'}} = c \left(\int_R^{+\infty} s^{-\gamma p' + n - 1} \, \mathrm{d}s \right)^{\frac{1}{p'}} = c R^{-\gamma + \frac{n}{p'}}$$

Also

$$\sum_{k=0}^{\infty} \int_{2^{-k-1}R \le |y| < 2^{-k}R} f(x-y)|y|^{-\gamma} \, \mathrm{d}y \le \sum_{k=0}^{\infty} (2^{-k}R)^{-\gamma} \int_{|y| \le 2^{-k}R} |f(x-y)| \, \mathrm{d}y \le C R^{n-\gamma} \mathcal{M}(f).$$

Therefore if we take R s.t. $R^{n-\gamma} \mathcal{M}(f)(x) = R^{-\gamma + \frac{n}{p'}} \|f\|_{L^p}$,

$$f * |\cdot|^{\gamma} \le c \|f\|_{L^p} R^{-\gamma + \frac{n}{p'}} + C R^{n-\gamma} \mathcal{M}(f)(x) = C(\mathcal{M}(f)(x))^{\frac{p}{q}} \|f\|_{L^p}^{1-\frac{p}{q}}.$$

Hence taking L^q norm on both sides will yield the result. □

Now we can state the main result:

Theorem 4.3.8 (Gagliardo-Nirenberg-Sobolev)
For $1 \le p < q < +\infty$, $m \le n$ are nonnegative integers. If $\frac{1}{p} - \frac{1}{q} = \frac{m}{n}$,

$$\|f\|_{L^q} \le C_{p,q,n} \|\partial^m f\|_{L^p}, \quad f \in C_0^\infty(\mathbb{R}^n).$$

When $q = +\infty$, $m > \frac{n}{p}$, we have

$$\|f\|_{L^\infty} \le C_{p,m,n,r}(\|\partial^m f\|_{L^p} + \|f\|_{L^r}), \quad \forall f \in C_0^\infty(\mathbb{R}^n), 1 \le r \le +\infty.$$

More generally, when $r \ge n, 1 \le p \le q \le +\infty$, except $p = +\infty, r = n, q < +\infty$,

$$\|f\|_{L^q} \le C(\|\nabla f\|_{L^r} + \|f\|_{L^p}).$$

Remark 4.3.9 — Again, the relation between p, q, m, n can be derived by scaling.
This theorem states that the regularity of derivatives can be passed to the original function (with L^p replaced by a higher L^q).
The second inequailty induces (when $r = p$)

$$\|f\|_{L^\infty} \le C\|f\|_{W^{m,p}}.$$

Proof of first inequality. Actually we only need to prove the case $m = 1$, since we can repeatedly apply $m = 1$ inequality to get larger m.

When $p > 1$, since f has compact support,

$$C|f(x)| \le \int_{S^{n-1}} \int_0^{+\infty} |\nabla f(x - r\omega)| \,\mathrm{d}r \,\mathrm{d}\omega = \int_{\mathbb{R}^n} |\nabla f(x-y)||y|^{-(n-1)} \,\mathrm{d}y = |\nabla f| * |\cdot|^{-(n-1)}.$$

Now by $1 - \frac{n-1}{n} = \frac{1}{p} - \frac{1}{q}$ and Hardy-Littlewood-Sobolev,

$$\|f\|_{L^q} \le C\|\nabla f\|_{L^p}.$$

When $p = 1$, we need to prove

$$\|f\|_{L^{\frac{n}{n-1}}} \le C\|\nabla f\|_{L^1}.$$

When $n = 1$, it's trivial; When $n = 2$,

$$f(x,y) \le \int_{\mathbb{R}} |\partial_x f(s,y)| \,\mathrm{d}s, \quad f(x,y) \le \int_{\mathbb{R}} |\partial_y f(x,t)| \,\mathrm{d}t.$$

Multiplying these together we'll get

$$\|f\|_{L^2}^2 \le \int |\partial_x f(s,y) \partial_y f(x,t)| \,\mathrm{d}s \,\mathrm{d}t \,\mathrm{d}x \,\mathrm{d}y = \int |\partial_x f(x,y)| \,\mathrm{d}x \,\mathrm{d}y \int |\partial_y f(x,y)| \,\mathrm{d}x \,\mathrm{d}y.$$

Hence the result is true.

For general $n = k$, we proceed by induction. Let $x \in \mathbb{R}, y \in \mathbb{R}^{k-1}$.

$$\begin{aligned}
\iint |f(x,y)|^{\frac{k}{k-1}} \,\mathrm{d}x \,\mathrm{d}y &\le \int_{\mathbb{R}^{k-1}} \left(\int |f(x,y)| \,\mathrm{d}x \right) \left(\int |\partial_x f(s,y)| \,\mathrm{d}s \right)^{\frac{1}{k-1}} \mathrm{d}y \\
&\le \left\| \int |\partial_x f(s,y)| \,\mathrm{d}s \right\|_{L_y^1}^{\frac{1}{k-1}} \cdot \left\| \int |f(x,y)| \,\mathrm{d}x \right\|_{L_y^{\frac{k-1}{k-2}}} \\
&\le \|\nabla f\|_{L^1}^{\frac{1}{k-1}} \cdot \left\| \int |\nabla_y f(x,y)| \right\|_{L_{x,y}^1}.
\end{aligned}$$

□

Proof of second inequality. When $m = 1$, $p > n$, let φ be a smooth function s.t. $\varphi(x) = 1$ when $|x| \le 1$, and $\varphi(x) = 0$ when $|x| > 2$.

$$\begin{aligned} |f(0)| = |f\varphi(0)| &\le |\nabla(f\varphi) * |\cdot|^{-(n-1)}(0)| \\ &= \left| \int \nabla(f\varphi)(y)|y|^{-(n-1)} \,\mathrm{d}y \right| \\ &= \int_{|y|\le 2} \nabla(f\varphi)(y)|y|^{-(n-1)} \,\mathrm{d}y. \end{aligned}$$

Thus

$$|f(0)| \le \|\nabla(f\varphi)\|_{L^p} \||y|^{-(n-1)}\chi_{|y|\le 2}\|_{L^{p'}} = C\|\nabla(f\varphi)\|_{L^p}.$$

Since we have

$$\|\nabla(f\varphi)\|_{L^p} \le \|\nabla f \cdot \varphi\|_{L^p} + \|f\nabla\varphi\|_{L^p} \le \|\nabla f\|_{L^p} + \|f\|_{L^r}^{\theta}\|\nabla\varphi\|_{L^q}^{1-\theta}.$$

Where $\frac{1}{p} = \frac{\theta}{r} + \frac{1-\theta}{q}$. □

In the proof we encountered an inequality that looks like this:

$$\|f\|_{L^{\frac{n}{n-1}}} \le C\|\nabla f\|_{L^1}.$$

This is related to the *isoperimetric inequality*:

$$|\Omega|^{\frac{n-1}{n}} \le C_n|\partial\Omega|, \quad \Omega \subset \mathbb{R}^n.$$

Which states the relations between the volume of a set and its "surface area". Of course here we require the boundary to have good properties so that we can define its "area".

Another related formula is the *Co-area equality*. (which we proved as homework in manifold section)

$$\int_\Omega g(x)|\nabla u(x)| \,\mathrm{d}x = \int_{-\infty}^{+\infty} \left(\int_{u^{-1}(t)} g(x) \,\mathrm{d}\sigma \right) \mathrm{d}t.$$

Here σ is the measure on the manifold $u^{-1}(t)$.

From this we can prove the equivalence of isoperimetric inequality and Sobolev inequality. Assume Sobolev inequality, let $f_\epsilon \to \chi_\Omega$, intuitively $\|f_\epsilon\|_{L^{\frac{n}{n-1}}} \to |\Omega|^{\frac{n-1}{n}}$, and $\int_\Omega |\nabla f_\epsilon| = \int_0^1 |f_\epsilon^{-1}(t)| \,\mathrm{d}t \to |\partial\Omega|$. (later when we learned distributions, we can directly take $f = \chi_\Omega$.)

Conversely, for $f \in C_0^\infty$, let $\Omega_t = \{f > t\}, \partial\Omega_t = \{f = t\}$. Isoperimetric inequality gives $|\{f > t\}|^{\frac{n-1}{n}} \le C|\{f = t\}|$. By co-area formula again,

$$C\int |\nabla f| \,\mathrm{d}x = C\int_{t\ge 0} |\{f = t\}| \,\mathrm{d}t \ge \int_{t\ge 0} |\{f > t\}|^{\frac{n-1}{n}} \,\mathrm{d}t.$$

Since

$$t|\{f > t\}|^{\frac{n-1}{n}} \le \int_0^t |\{f > s\}|^{\frac{n-1}{n}} \,\mathrm{d}s \le \int_0^{+\infty} |\{f > s\}|^{\frac{n-1}{n}} \,\mathrm{d}s \le C\|\nabla f\|_{L^1}.$$

and

$$\int |f|^{\frac{n}{n-1}} \,\mathrm{d}x = \frac{n}{n-1}\int_{t\ge 0} |\{f > t\}| t^{\frac{1}{n-1}} \,\mathrm{d}t \le \frac{n}{n-1}\int_{t\ge 0} |\{f > t\}|^{\frac{n-1}{n}} \,\mathrm{d}t \cdot C\|\nabla f\|_{L^1}^{\frac{1}{n-1}} \le C_n\|\nabla f\|_{L^1}^{\frac{n}{n-1}}.$$

We've shown the equivalence of these two inequality.

§5 Distribution theory

§5.1 Definitions and motivations

Let $\Omega \subset \mathbb{R}^n$, given a compact set $K \subset \Omega$, the function space $C_K^\infty(\Omega)$ is the space of all smooth functions whose support is in K.

$$C_0^\infty(\Omega) = \bigcup_{K \subset \Omega} C_K^\infty(\Omega) =: \mathscr{D}(\Omega)$$

is the usual compact supported smooth functions in Ω. This space is used as *test function space* later.

Definition 5.1.1 (Distributions)**.** The **distributions** on Ω (also called **generalized functions**) is a *linear functional* on $\mathscr{D}(\Omega)$

$$u : \mathscr{D}(\Omega) \to \mathbb{C}, \quad \varphi \mapsto \langle u, \varphi \rangle ,$$

satisfying

- $\forall \varphi, \psi \in \mathscr{D}(\Omega)$, $\alpha, \beta \in \mathbb{C}$ we have

$$\langle u, \alpha\varphi + \beta\psi \rangle = \alpha \langle u, \varphi \rangle + \beta \langle u, \psi \rangle .$$

- (Continuity) For all compact set $K \subset \Omega$, there exists a nonnegative integer P and constant $C(P, K)$, such that

$$| \langle u, \varphi \rangle | \leq C \sup_{|\alpha| \leq P} \|\partial^\alpha \varphi\|_{L^\infty(K)}.$$

Remark 5.1.2 — Recall that for general functionals, the continuity is defined as

$$\|u(\varphi)\|_{B_1} \leq C\|\varphi\|_{B_2}$$

But since $\mathscr{D}(\Omega)$ is not a Banach space, we need to change the norm of φ to the stated one.

If the choice of P does not depend on K, then the minimal such P is called the **order** of u.

Denote $\mathscr{D}'(\Omega)$ the set of distributions on Ω. It's the dual space of $\mathscr{D}(\Omega)$.

Definition 5.1.3 (Limits)**.** Let $u_n \in \mathscr{D}'(\Omega)$, we say u_n converges to u if

$$\lim_{n\to\infty} \langle u_n, \varphi \rangle = \langle u, \varphi \rangle , \quad \forall \varphi \in \mathscr{D}(\Omega).$$

We write $u_n \xrightarrow{\mathscr{D}'} u$.

Example 5.1.4 (Dirac function)

This is one of the motivations to develop theory of distributions.

For all $a \in \Omega$, define $\delta_a \in \mathscr{D}'(\Omega)$, $\forall \varphi \in \mathscr{D}(\Omega)$ we have $\langle \delta_a, \varphi \rangle = \varphi(a)$. Since

$$| \langle \delta_a, \varphi \rangle | = |\varphi(a)| \leq \|\varphi\|_{L^\infty(K)},$$

so δ_a is a distribution of order 0.

Example 5.1.5 (Locally integrable functions)
For all $f \in L^1_{loc}(\Omega)$, define the functional T_f to be

$$T_f : \mathscr{D}(\Omega) \to \mathbb{C}, \quad \varphi \mapsto \int_\Omega f\varphi.$$

Since $\left|\int_\Omega f\varphi\right| \le \|f\|_{L^1(K)}\|\varphi\|_{L^\infty(K)}$, take $C = \|f\|_{L^1(K)}$ and $P = 0$, we see that T_f is a distribution of order 0, thus classical functions can be viewed as distributions.

Moreover $f \mapsto T_f$ is injective (up to a.e. equality).

Example 5.1.6
Take $\phi(x) \in \mathscr{D}(\mathbb{R}^n)$ s.t.

$$\int_{\mathbb{R}^n} \phi(x)\,\mathrm{d}x = 1.$$

For $\varepsilon > 0$, define

$$\phi_\varepsilon(x) = \varepsilon^{-n}\phi(x\varepsilon^{-1})$$

It is an approximation to identity, we have $\phi_\varepsilon \xrightarrow{\mathscr{D}'} \delta_0$.

$$\langle \phi_\varepsilon, \varphi \rangle = \int_{\mathbb{R}^n} \varepsilon^{-n}\phi(x\varepsilon^{-1})\varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} \phi(x)\varphi(\varepsilon x)\,\mathrm{d}x \to \varphi(0).$$

Example 5.1.7 (Radon measure)
Let μ be a measure on $(\Omega, B(\Omega))$. If for every compact set $K \subset \Omega$, $\mu(K) < +\infty$, we say μ is a **Radon measure**.

For any Radon measure μ, define a distribution

$$T_\mu : \forall \varphi \in \mathscr{D}(\Omega), \quad \langle T_\mu, \varphi \rangle = \int_\Omega \varphi\,\mathrm{d}\mu.$$

Since for any $K \subset \Omega$, $\varphi \in C^\infty_K(\Omega)$,

$$|\langle T_\mu, \varphi \rangle| \le \mu(K)\|\varphi\|_{L^\infty(K)}.$$

T_μ is a distribution of order 0.

In fact all the distributions of order 0 are (signed) Radon measures.

Example 5.1.8

Note that $\frac{1}{x} \notin L^1_{loc}(\mathbb{R})$, so we can't realize $\frac{1}{x}$ as a distribution directly.

However, we can define

$$\left\langle pv\frac{1}{x}, \varphi \right\rangle = \int_0^{+\infty} \frac{\varphi(x) - \varphi(-x)}{x} \,\mathrm{d}x.$$

Called the *principal value distribution* of $\frac{1}{x}$.

$$\left| \int_0^{+\infty} \frac{\varphi(x) - \varphi(-x)}{x} \,\mathrm{d}x \right| = \left| \int_0^{+\infty} \int_{-1}^{1} \varphi'(xt) \,\mathrm{d}t \,\mathrm{d}x \right| \leq C(K) \|\varphi'\|_{L^\infty(K)}$$

Thus $pv\frac{1}{x}$ is a distribution of order at most 1. In fact its order is 1.

§5.2 Derivatives of distributions

Definition 5.2.1. Let $u \in \mathscr{D}'(\Omega)$, α is a multiple index. Define

$$\langle \partial^\alpha u, \varphi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \varphi \rangle .$$

We need to check that $\partial^\alpha u$ is a distribution. For any compact set K, $\varphi \in C^\infty_K(\Omega)$, then $\partial^\alpha \varphi \in C^\infty_K(\Omega)$ as well.

$$| \langle \partial^\alpha, \varphi \rangle | = | \langle u, \partial^\alpha \varphi \rangle | \leq C \sup_{|\beta| \leq P} \|\partial^\beta \partial^\alpha \varphi\|_{L^\infty(K)} \leq C \sup_{|\beta| \leq P + |\alpha|} \|\partial^\beta \varphi\|_{L^\infty(K)}.$$

Thus $\partial^\alpha u$ is a distribution of order at most $k + |\alpha|$, where k is the order of u.

Example 5.2.2 (Heaviside function)

Consider the locally integrable function

$$H(x) = \chi_{x \geq 0}$$

as a distribution, it has a derivative $H'(x) = \delta_0$ since for $\varphi \in \mathscr{D}(\mathbb{R})$,

$$\langle H', \varphi \rangle = - \langle H, \varphi' \rangle = - \int H(x) \varphi'(x) \,\mathrm{d}x = - \int_0^{+\infty} \varphi'(x) \,\mathrm{d}x = \varphi(0).$$

Another example is $pv\frac{1}{x} = (\log|x|)'$.

Since $\log|x| \in L^1_{loc}(\mathbb{R})$, by definition,

$$\langle (\log|x|)', \varphi \rangle = - \langle \log|x|, \varphi'(x) \rangle = - \int_{\mathbb{R}} \log|x| \varphi'(x) \,\mathrm{d}x = - \int_0^{+\infty} \log x (\varphi'(x) + \varphi'(-x)) \,\mathrm{d}x$$

If we view this as a generalized Riemann integral, using integration by parts,

$$\langle (\log|x|)', \varphi \rangle = - \log x (\varphi(x) - \varphi(-x)) \Big|_0^{+\infty} + \int_0^{+\infty} \frac{\varphi(x) - \varphi(-x)}{x} \,\mathrm{d}x = \left\langle pv\frac{1}{x}, \varphi \right\rangle .$$

§5.3 $C^\infty(\Omega)$-module structure of distributions

Now we've seen that distributions can perform addition and differentiation, but we can't simply define the multiplication of distributions.

since $\mathscr{D}'(\Omega)$ is the dual space of $C_0^\infty(\Omega)$, for $f \in C^\infty(\Omega)$, $u \in \mathscr{D}'(\Omega)$, we can define the multiplication fu as

$$\langle fu, \varphi\rangle = \langle u, f\varphi\rangle .$$

if u is a locally integrable function, this product coincides with the normal multiplication of functions.

Since $\forall K, \exists P, C$, s.t.

$$|\langle fu, \varphi\rangle| \le C \sum_{|\alpha|\le P} \|\partial^\alpha(f\varphi)\|_{L^\infty(K)} \le C \cdot C(f, P, K) \sum_{|\alpha|\le P} \|\partial^\alpha \varphi\|_{L^\infty(K)}.$$

fu is indeed a distribution.

Proposition 5.3.1 (Leibniz's law)

$$\partial(fu) = \partial u f + \partial f u.$$

Proof.

$$\langle \partial(fu), \varphi\rangle = -\langle fu, \partial\varphi\rangle = -\langle u, f\partial\varphi\rangle = -\langle u, \partial(f\varphi) - \partial f\varphi\rangle = \langle \partial u f + \partial f u, \varphi\rangle .$$

□

Example 5.3.2

$x \cdot pv\frac{1}{x} = 1$ since $\forall \varphi \in \mathscr{D}(\Omega)$,

$$\left\langle x \cdot pv\frac{1}{x}, \varphi\right\rangle = \left\langle pv\frac{1}{x}, x\varphi\right\rangle = \int_0^{+\infty} \frac{x\varphi(x) + x\varphi(-x)}{x}\,\mathrm{d}x = \int_0^{+\infty} (\varphi(x) + \varphi(-x))\,\mathrm{d}x = \langle 1, \varphi\rangle .$$

Next we'll study the variable substitution of distributions.

Recall that in differential manifolds, let $\Phi : \Omega_1 \to \Omega_2$ be a differential homeomorphism, then it induces a pushforward of tangent spaces $\Phi_* : T\Omega_1 \to T\Omega_2$, while in cotangent space (or 1-forms) it induces a pullback $\Phi^* : T^*\Omega_2 \to T^*\Omega_1$.

Let $u_1 \in \mathscr{D}'(\Omega_1)$, $\varphi_1 \in \mathscr{D}(\Omega_1)$, and $\Phi : \Omega_1 \to \Omega_2$ a smooth map. Formally we can write

$$\langle u_1, \varphi_1\rangle = \int u_1(x)\varphi_1(x)\,\mathrm{d}x = \int u_1(\Phi^{-1}(y))\varphi(\Phi^{-1}(y))\,\mathrm{d}\Phi^{-1}(y) = \int u_1\Phi^{-1}(y)\varphi_1\Phi^{-1}(y)|J_{\Phi^{-1}}|\,\mathrm{d}y$$

Thus we can define a new distribution "$u \circ \Phi^{-1}$" as

$$\left\langle u \circ \Phi^{-1}, \varphi_1 \circ \Phi^{-1}|J_{\Phi^{-1}}|\right\rangle = \langle u, \varphi_1\rangle .$$

Hence the pushforward and pullback is defined as

$$\langle \Phi_* u_1, \varphi_2\rangle = \langle u_1, \varphi_2 \circ \Phi(x)|J_\Phi|(x)\rangle$$
$$\langle \Phi^* u_2, \varphi_1\rangle = \left\langle u_2, \varphi_1 \circ \Phi^{-1}(y)|J_{\Phi^{-1}}|(y)\right\rangle$$

Here we won't bother to check they are indeed distributions because of the complicated computations.

Example 5.3.3

Let $x_0 \in \Omega_1$, $y_0 \in \Omega_2$, $\Phi(x_0) = y_0$, then

$$\Phi^* \delta_{y_0} = |J_\Phi(x_0)|^{-1} \delta_{x_0}.$$

Proposition 5.3.4 (Chain rule)

$$\partial_j(\Phi^* u_2) = \sum_{k=1}^n \partial_j \Phi_k \cdot \Phi^* \partial_k u_2.$$

Proof. For all $\varphi \in C_0^\infty(\Omega_1)$,

$$\begin{aligned}
\left\langle \sum_{k=1}^n \partial_j \Phi_k \Phi^* \partial_k u_2, \varphi \right\rangle &= \sum_{k=1}^n \langle \Phi^* \partial_k u_2, \partial_j \Phi_k \varphi \rangle \\
&= \sum_{k=1}^n \left\langle \partial_k u_2, (\partial_j \Phi_k \cdot \varphi) \circ \Phi^{-1} |J_{\Phi^{-1}}| \right\rangle \\
&= -\sum_{k=1}^n \left\langle u_2, \frac{\partial}{\partial y_k} (\partial_j \Phi_k \circ \Phi^{-1} \cdot (\varphi \circ \Phi^{-1}) |J_{\Phi^{-1}}|) \right\rangle
\end{aligned}$$

Note that

$$\sum_{k=1}^n \frac{\partial}{\partial y_k} \left(\frac{\partial \Phi_k}{\partial x_j} (\Phi^{-1}(y)) |J_{\Phi^{-1}}(y)| \right) = 0.$$

Since $\forall g \in C_0^\infty(\Omega_2)$, by Stokes formula, substitution and integration by parts,

$$0 = \int_{\Omega_1} \frac{\partial}{\partial x_j} (g \circ \Phi) \, \mathrm{d}x = \int_{\Omega_1} \sum_{k=1}^n \partial_k g \frac{\partial \Phi_k}{\partial x_j} \, \mathrm{d}x = -\int_{\Omega_2} \sum_{k=1}^n g \partial_k \left(\frac{\partial \Phi_k}{\partial x_j} (\Phi^{-1}(y)) |J_{\Phi^{-1}}(y)| \right) \mathrm{d}y$$

Continuing the computation,

$$\begin{aligned}
&= -\sum_{k=1}^n \left\langle u_2, \partial_j \Phi_k \circ \Phi^{-1} |J_{\Phi^{-1}}| \cdot \frac{\partial}{\partial y_k} (\varphi \circ \Phi^{-1}) \right\rangle \\
&= -\sum_{k=1}^n \left\langle \Phi^* u_2, \partial_j \Phi_k \cdot \frac{\partial \varphi \circ \Phi^{-1}}{\partial y_k} \circ \Phi \right\rangle \\
&= \langle \partial_j \Phi^* u_2, \varphi \rangle .
\end{aligned}$$

□

Stokes formula can also be generalized for distributions.

Let $\nu(x) = (\nu_1(x), \dots, \nu_n(x))$ be the outer normal unit vector of $\partial\Omega$, σ the measure on $\partial\Omega$. Stokes formula says

$$\int_\Omega \frac{\partial \varphi}{\partial x_j} \, \mathrm{d}x = \int_{\partial \Omega} \varphi \nu_j \, \mathrm{d}\sigma.$$

Hence the distribution version is

$$\langle \chi_\Omega, \partial_j \varphi \rangle = -\langle \partial_j \chi_\Omega, \varphi \rangle .$$

Theorem 5.3.5

Let Ω be a smooth region, $\nu(x)$ is the unit outer normal vector, σ is the measure on $\partial\Omega$. Then as distributions, we have

$$\partial_i \chi_\Omega = -\nu_i \,\mathrm{d}\sigma, \quad \nabla \chi_\Omega = -\nu \,\mathrm{d}\sigma.$$

In one dimensional version, $f(x)$ integrable on (a,b),

$$F(x) = \int_a^x f(t)\,\mathrm{d}t.$$

then $F'(x) = f(x)$, *a.e.* (Since absolute continuity). As distributions we have

$$F'(x) = f(x) \in \mathscr{D}'$$

Theorem 5.3.6

Given a distribution $u \in \mathscr{D}'((a,b))$. If the derivative $u' = 0$ as a distribution, then $u = c$ is a constant.

Proof. Since

$$\langle u', \varphi \rangle = -\langle u, \varphi' \rangle = 0,$$

note that $\int_a^b \varphi' \,\mathrm{d}x = 0$, for all $\psi \in \mathscr{D}((a,b))$, let

$$g(x) = \psi(x) - \left(\int_a^b \psi(y)\,\mathrm{d}y \right) \lambda(x),$$

where $\lambda(x) \in \mathscr{D}((a,b))$ and $\int_a^b \lambda \,\mathrm{d}x = 1$.

Thus $G(x) = \int_a^x g(y)\,\mathrm{d}y \in C_0^\infty((a,b))$. This gives

$$0 = \langle u, G'(x) \rangle = \left\langle u, \psi(x) - \lambda(x) \int_a^b \psi(y)\,\mathrm{d}y \right\rangle$$

$$\implies \langle u, \psi \rangle = \int_a^b \psi(y)\,\mathrm{d}y \,\langle u, \lambda \rangle = \langle \langle u, \lambda \rangle, \psi \rangle .$$

This means $u = \langle u, \lambda \rangle$ is a constant distribution. □

In higher dimensional cases, we only consider $\Omega = \mathbb{R}^n$ here. For $\phi \in C_0^\infty(\mathbb{R}^n)$, there exists $\Phi_1 \in C_0^\infty$ and $\chi \in C_0^\infty(\mathbb{R})$ s.t.

$$\phi(x_1, \ldots, x_n) = \partial_1 \Phi_1 + \chi(x_1) \int_{\mathbb{R}} \phi(s, x_2, \ldots, x_n)\,\mathrm{d}s, \quad \int \chi(x)\,\mathrm{d}x = 1.$$

Since $\int_{\mathbb{R}} \phi(s, x_2, \ldots, x_n)\,\mathrm{d}s$ is a function in $\mathbb{R}^{n-1}$, we can use induction on n.

§5.4 Supports of distributions

For a function f, $\operatorname{supp} f := \overline{\{f(x) \neq 0\}}$ is the closure of the nonzero points. But this can't be generalized to distributions, so if we rephrase this as

$$\operatorname{supp} f = (\{x \mid \exists V_x \ni x,\ s.t.\ f(x)|_{V_x} = 0\})^c$$

we can do the same for distributions:

Definition 5.4.1 (Support)**.** Let u be a distribution on Ω,

$$\operatorname{supp} u := \Omega \setminus \{x \in \Omega \mid \exists V_x\ s.t.\ \langle u, \varphi\rangle = 0, \forall \varphi \in \mathscr{D}(V_x)\}.$$

here V_x is an open neighborhood of x.

If $\operatorname{supp}(u)$ is compact, we say u is a **compact supported** distribution. Let $\mathscr{E}(\Omega)$ denote all the distributions with compact support.

Theorem 5.4.2 (Unit decomposition)
Let $K \subset \mathbb{R}^n$ be a compact set, $\{U_1, \ldots, U_N\}$ is an open covering of K. Then $\exists \chi_j \in C_0^\infty(U_j)$ s.t.

- $0 \leq \chi_j(x) \leq 1$.
- There exists an open set $V \supseteq K$, such that $\forall x \in V$,

$$\chi_1(x) + \cdots + \chi_N(x) = 1.$$

§5.5 Convolutions of distributions

For a smooth function f, when u is a function, we have

$$\langle u * f, \varphi\rangle = \iint u(x-y) f(y)\varphi(x)\,\mathrm{d}x\,\mathrm{d}y = \iint u(x)f(y)\varphi(x+y)\,\mathrm{d}x\,\mathrm{d}y = \left\langle u, \int f(-x)\varphi(y-x)\,\mathrm{d}x\right\rangle$$

Therefore let $\mathcal{R}f(x) := f(-x)$ be the reflection operator, we can define

$$\langle u * f, \varphi\rangle = \langle u, \mathcal{R}f * \varphi\rangle$$

for general distributions u.

We need to ensure f is integrable and has compact support, so that $\mathcal{R}f * \varphi \in C_0^\infty(\mathbb{R}^n)$.

Since

$$\begin{aligned}
|\langle u * f, \varphi\rangle| &\leq C \sum_{|\alpha| \leq P} \|\partial^\alpha((\mathcal{R}f) * \varphi)\|_{L^\infty(\operatorname{supp}(\mathcal{R}f)*\varphi)} \\
&\leq C \sum_{|\alpha| \leq P} \|\mathcal{R}f * \partial^\alpha \varphi\|_{L^\infty(\operatorname{supp}(\mathcal{R}f)*\varphi)} \\
&\leq C \sum_{|\alpha| \leq P} \|\mathcal{R}f\|_{L^1} \|\partial^\alpha \varphi\|_{L^\infty(\operatorname{supp}\varphi)}.
\end{aligned}$$

The distribution $u * f$ is well-defined.

Theorem 5.5.1

Let $f \in C_0^\infty(\mathbb{R}^n)$, then $u * f$ as a distribution is equivalent to the smooth function $u(\mathcal{R}\tau_x f)$ w.r.t. x.

Here τ_x is the translation operator.

Proof. First we prove that $u(f(x-y))$ is a smooth function of x.

For the continuity, $u(f(x_0-y)) - u(f(x-y)) = u(f(x_0-y) - f(x-y)) \to 0$ when $x \to x_0$ from the continuity of distributions.

The derivative is $\frac{u(f(x_0+tv-y)-f(x-y))}{t} = u(\frac{f(x_0+tv-y)-f(x-y)}{t}) \to u(\partial_v f(x_0-y))$. Hence it is indeed smooth.

Secondly, we need to prove $\langle u * f, \varphi\rangle = \langle u(\mathcal{R}\tau_x f), \varphi\rangle$.

$$\begin{aligned}\langle u, \mathcal{R}f * \varphi\rangle &= u\left(\int (\mathcal{R}f)(y-x)\varphi(x)\,\mathrm{d}x\right) \\ &= u\left(\int \mathcal{R}\tau_x f(-y)\cdot\varphi(x)\,\mathrm{d}x\right)\end{aligned}$$

Now by the definition of Riemann integrals, for all $\varepsilon > 0$, $\exists$ a partition Δ_j s.t. the finite sum approaches the integral.

Since u is commutative with finite sums, and u has continuity as operators, so u is commutative with integrals. □

Let ϕ_ε be an approximation to identity, then as distributions

$$u * \phi_\varepsilon \xrightarrow{\mathscr{D}'} u \iff \langle u * \phi_\varepsilon, \varphi\rangle = \langle u, \mathcal{R}\phi_\varepsilon * \varphi\rangle \to \langle u, \varphi\rangle\,.$$

This means that smooth functions are dense in distributions.

Definition 5.5.2 (Convolution of distributions)**.** Let $u \in \mathscr{D}'(\mathbb{R}^n)$, $c \in \mathscr{E}'(\mathbb{R}^n)$. Define $u * c$ as

$$\langle u * c, \varphi\rangle = \langle u, \mathcal{R}c * \varphi\rangle\,.$$

Where $\langle \mathcal{R}c, \varphi\rangle = \langle c, \mathcal{R}\varphi\rangle$ is the reflection of distributions.

Proof. For all K, $\exists C, P$ s.t.

$$|\langle u, \varphi\rangle| \le C \sum_{|\alpha|\le P} \|\partial^\alpha \varphi\|_{L^\infty(K)}.$$

Let $L := \operatorname{supp} \mathcal{R}c$, then $\exists C_1$ and P_1 s.t.

$$|\langle \mathcal{R}c, \varphi\rangle| \le C_1 \sum_{|\alpha|\le P_1} \|\partial^\alpha \varphi\|_{L^\infty(L)}.$$

Hence

$$
\begin{aligned}
|\langle u * c, \varphi\rangle| &= |\langle u, \mathcal{R}c * \varphi\rangle| \\
&\leq C \sum_{|\alpha| \leq P} \|\partial^\alpha(\mathcal{R}c * \varphi)\|_{L^\infty(\operatorname{supp}(\mathcal{R}c*\varphi))} \\
&\leq C \sum_{|\alpha \leq P|} \|\mathcal{R}c * \partial^\alpha \varphi\|_{L^\infty} \\
&= C \sum_{|\alpha| \leq P} \|\mathcal{R}c(\partial^\alpha \varphi(x-y))\|_{L^\infty_x} \\
&\leq \sum_{|\alpha| \leq P} \|C_1 \sup_{|\beta| \leq P_1} \|\partial^\beta \partial^\alpha \varphi(x-y)\|_{L^\infty_x} \|_{L^\infty_y} \\
&\leq CC_1 \sum_{|\alpha| \leq P, |\beta| \leq P_1} \|\partial^\beta \partial^\alpha \varphi(z)\|_{L^\infty_z(K)}.
\end{aligned}
$$

□

Example 5.5.3 (Translation operator)
Given a distribution u, define $\tau_a u$ as

$$\langle \tau_a u, \varphi\rangle = \langle u, \tau_{-a}\varphi\rangle .$$

Hence $u * \delta_a = \tau_{-a} u$.

$$\langle u * \delta_a, \varphi\rangle = \langle u, \delta_{-a} * \varphi\rangle = \langle u, \varphi(x+a)\rangle = \langle u, \tau_a \varphi\rangle .$$

Note that in fact $\delta_a = \tau_{-a}\delta_0$, this is a little surprising.

Proposition 5.5.4
Let $u \in \mathscr{D}'(\mathbb{R}^n)$, $c \in \mathscr{E}'(\mathbb{R}^n)$,

- $\operatorname{supp}(u * c) \subset \operatorname{supp} u + \operatorname{supp} c$.
- Convolution is commutative with derivatives,

$$\partial^\alpha(u * c) = \partial^\alpha u * c = u * \partial^\alpha c.$$

Proof. Since $\langle u * c, \varphi\rangle = \langle u, \mathcal{R}c * \varphi\rangle$,

$$\operatorname{supp}\varphi \cap (\operatorname{supp} c + \operatorname{supp} u) = \emptyset \implies \operatorname{supp}(\mathcal{R}c * \varphi) \cap \operatorname{supp} u = \emptyset \implies \langle u, \mathcal{R}c * \varphi\rangle = 0.$$

We get the desired.

For the second one, just compute

$$\langle \partial^\alpha(u * c), \varphi\rangle = \langle u * c, \partial^\alpha \varphi\rangle (-1)^{|\alpha|} = \langle u, \mathcal{R}c * \varphi\rangle (-1)^{|\alpha|} = \langle u, \partial^\alpha(\mathcal{R}c * \varphi)\rangle (-1)^{|\alpha|}.$$

□

Proposition 5.5.5 (Commutativity of convolutions)
Let c_1, c_2 be distributions with compact support, then $c_1 * c_2 = c_2 * c_1$.

Proof. Let $\chi_\varepsilon = \varepsilon^{-n}\chi(x\varepsilon^{-1})$ be an approximation to identity, then $\lim_{\varepsilon\to 0} c_1 * \chi_\varepsilon = c_1$. Hence

$$\begin{aligned}\langle c_1 * \chi_\varepsilon, \mathcal{R}c_2 * \varphi\rangle &= \langle \mathcal{R}c_2 * \varphi, c_1 * \chi_\varepsilon\rangle \\ &= \langle \mathcal{R}c_2, (c_1 * \chi_\varepsilon) * \mathcal{R}\varphi\rangle \\ &= \langle c_2, \mathcal{R}((c_1 * \chi_\varepsilon) * \mathcal{R}\varphi)\rangle \\ &= \langle c_2, (\mathcal{R}c_1 * \mathcal{R}\chi_\varepsilon) * \varphi\rangle \\ &= \langle c_2, \mathcal{R}c_1 * (\mathcal{R}\chi_\varepsilon * \varphi)\rangle = \langle c_2 * c_1, \mathcal{R}\chi_\varepsilon * \varphi\rangle .\end{aligned}$$

Where the last but second equality used the associativity of convolution:
Let $u \in \mathscr{D}'(\mathbb{R}^n)$, $f, g \in \mathscr{D}(\mathbb{R}^n)$. We have

$$\begin{aligned}\langle (u * f) * g, \varphi\rangle &= \langle u * f, \mathcal{R}g * \varphi\rangle \\ &= \langle u, \mathcal{R}f * (\mathcal{R}g * \varphi)\rangle \\ &= \langle u * (f * g), \varphi\rangle .\end{aligned}$$

□

Theorem 5.5.6 (Continuity of convolutions)
Let $c_k \in \mathscr{E}(\mathbb{R}^n)$, there exists a compact set K, such that $\operatorname{supp} c_k \subset K$. Let $u_k \in \mathscr{D}'(\mathbb{R}^n)$. If there exists c and u s.t.

$$c_k \xrightarrow{\mathscr{D}'} c, \quad u_k \xrightarrow{\mathscr{D}'} u.$$

Then

$$u * c_k \xrightarrow{\mathscr{D}'} u * c, \quad u_k * c \xrightarrow{\mathscr{D}'} u * c.$$

Remark 5.5.7 — Note that in general we can't write $u_k * c_k \xrightarrow{\mathscr{D}'} u * c$, due to the fact that u_k doesn't have compact supports.

Proof. Write

$$\langle u_k * c, \varphi\rangle = \langle u_k, \mathcal{R}c * \varphi\rangle \to \langle u, \mathcal{R}c * \varphi\rangle = \langle u * c, \varphi\rangle .$$

Hence the second limit holds.
The first one is because

$$\langle u * c_k, \varphi\rangle = \langle u, \mathcal{R}c_k * \varphi\rangle$$

We need to show that $\mathcal{R}c_k * \varphi \to \mathcal{R}c * \varphi$ here. This is a little annoying, so we take a different approach.
Let ψ be a truncation function s.t. $\psi = 1$ on $\operatorname{supp}\varphi - K$. Then

$$\langle u, \mathcal{R}c_k * \varphi\rangle = \langle u, \psi\mathcal{R}c_k * \psi\rangle = \langle u\psi, \mathcal{R}c_k * \varphi\rangle = \langle (u\psi) * c_k, \varphi\rangle = \langle c_k * (u\psi), \varphi\rangle \to \langle c * (u\psi), \varphi\rangle .$$

Hence $u * c_k \xrightarrow{\mathscr{D}'} u * c$. □

Corollary 5.5.8 (Associativity of convolutions)
Let $u \in \mathscr{D}'(\mathbb{R}^n)$, $c_1, c_2 \in \mathscr{E}'(\mathbb{R}^n)$.

$$(u * c_1) * c_2 = u * (c_1 * c_2).$$

Proof. When $c_1, c_2 \in \mathscr{D}(\mathbb{R}^n)$, we already proved it.

When $c_2 = f \in \mathscr{D}(\mathbb{R}^n)$, using the continuity of convolutions,

$$(u * c_1) * f = \lim_{\varepsilon \to 0} (u * (c_1 * \chi_\varepsilon)) * f = \lim_{\varepsilon \to 0} u * ((c_1 * \chi_\varepsilon) * f) = u * (c_1 * f)$$

Hence for general cases,

$$\begin{aligned}
(u * c_1) * c_2 &= \lim_{\varepsilon \to 0} (u * c_1) * (c_2 * \chi_\varepsilon) \\
&= \lim_{\varepsilon \to 0} u * (c_1 * (c_2 * \chi_\varepsilon)) \\
&= \lim_{\varepsilon \to 0} u * ((c_1 * c_2) * \chi_\varepsilon) \\
&= u * (c_1 * c_2)
\end{aligned}$$

□

Theorem 5.5.9
Let c be a distribution with compact support, $f \in C^\infty(\mathbb{R}^n)$. Then $f * c \in C^\infty(\mathbb{R}^n)$ and

$$(f * c)(x) = \langle c, f(x - \cdot) \rangle .$$

Proof. Let χ be a truncation s.t. $\chi(x) = 1$ for $|x| \le 1$ and $\operatorname{supp} \chi \subset \{|x| \le 2\}$.

Define $\chi_0 = \chi$, $\chi_k(x) = \chi(2^{-k}x) - \chi(2^{-k+1}x)$, then

$$\sum_{k=0}^{\infty} \chi_k(x) = 1, \quad \operatorname{supp} \chi_k(x) \subset \{2^{k-1} \le |x| \le 2^{k+1}\}.$$

This step is to find a unit decomposition which is locally finite. Thus

$$\begin{aligned}
\langle f * c, \varphi \rangle = \left\langle \left(\sum_{k=0}^{\infty} f \chi_k \right) * c, \varphi \right\rangle &= \left\langle \sum_{k=0}^{\infty} f \chi_k, \mathcal{R}c * \varphi \right\rangle \\
&= \sum_{k=0}^{\infty} \langle f \chi_k, \mathcal{R}c * \varphi \rangle \\
&= \sum_{k=0}^{\infty} \langle (f \chi_k) * c, \varphi \rangle = \sum_{k=0}^{\infty} \langle c * (f \chi_k), \varphi \rangle .
\end{aligned}$$

We are able to put the summation outside the inner product since it's actually a finite sum (c has compact support).

Therefore

$$f * c = \sum_{k=0}^{\infty} c * (f \chi_k) = \sum_{k=0}^{\infty} \langle c, (f \chi_k)(x - y) \rangle = \langle c, f(x - y) \rangle .$$

□

Definition 5.5.10 (Convolutable sets)**.** Let F_1, F_2 be closed set in $\mathbb{R}^n$, if $\forall R > 0$, $\exists R' > 0$ s.t. $\forall x_1 \in F_1, x_2 \in F_2$,

$$|x_1 + x_2| \leq R \implies |x_1| < R', |x_2| < R'.$$

Then we say F_1 and F_2 are **convolutable**. We can also define the same thing for a series of closed sets.

Lemma 5.5.11

Closed sets F_1, F_2 are convolutable $\implies$ $F_1 + F_2$ is closed.

Proof. $x_k + y_k$ converges $\implies$ $x_k + y_k$ bounded, thus x_k, y_k are both bounded, there exists a subsequence which is convergent, so $F_1 + F_2$ must be closed. □

Example 5.5.12

The following sets are convolutable:

- F_1 is closed, F_2 is compact.
- The sets $\{[x_i, +\infty)\}$.
- In the time-space $\mathbb{R}^{1+3}$, the solid future light cone:

  $$\widehat{C_+} := \{(t, x) \in \mathbb{R}^{1+3} \mid t \geq |x|\}$$

  and the future of the plane $t = T$:

  $$\mathbb{R}^{1+3}_{t \geq T} := \{(t, x) \in \mathbb{R}^{1+3} \mid t \geq T\}$$

  are convolutable.

  These sets are related to the wave equation in physics.

Proposition 5.5.13

Let $u, v \in \mathscr{D}'(\mathbb{R}^n)$, $\operatorname{supp} u$ and $\operatorname{supp} v$ are convolutable, then we can define the convolution as follows.

Let χ_k be truncation function s.t. $\chi_k(x) = 1$ for all $|x| \leq k$, and $\operatorname{supp} \chi_k \subset \{|x| \leq k+1\}$.

$$\langle u * v, \varphi \rangle = \lim_{k \to \infty} \langle (u\chi_k) * (v\chi_k), \varphi \rangle .$$

Proof. First we prove the limit exists.

For a fixed φ, we claim that there exists $N > 0$ s.t.

$$\langle (u\chi_k) * (v\chi_k), \varphi \rangle = \langle (u\chi_l) * (v\chi_l), \varphi \rangle , \forall k, l > N.$$

This is because if we look at the difference

$$\langle (\chi_k - \chi_l)u * (\chi_k v), \varphi \rangle = 0 \impliedby \operatorname{supp}(\varphi) \cap \operatorname{supp}((\chi_k - \chi_l)u * (\chi_k v)) = \emptyset.$$

Otherwise if $x \in \operatorname{supp}(\chi_k - \chi_l)u$, $y \in \operatorname{supp} \chi_k v$, $x + y \in \operatorname{supp} \varphi$, assume that $\operatorname{supp} \varphi \subset \{|x| \le R\}$, then $|x + y| \le R \implies |x|, |y| \le R'$.

But $|x| \ge \min\{k, l\}$, contradiction!

Next we prove that this limit is independent of the choice of χ_k. The proof is essentially the same, if we replace χ_l with χ_k'.

At last, $u * v$ is a distribution since for any compact set K, let $N > 0$ s.t. $K \subset \{|x| \le N\}$. For $\varphi \in C_K^\infty(\mathbb{R}^n)$,

$$\langle u * v, \varphi \rangle = \langle (\chi_N u) * (\chi_N v), \varphi \rangle \le C \sum_{|\alpha| \le P} \|\partial^\alpha \varphi\|_{L^\infty(K)}.$$

Here we omitted some bothering details. □

Proposition 5.5.14

Let $u, v, w \in \mathscr{D}'(\mathbb{R}^n)$, $\operatorname{supp} u, \operatorname{supp} v, \operatorname{supp} w$ are convolutable, α is any multi-index.

- $\operatorname{supp}(u * v) \subset \operatorname{supp} u + \operatorname{supp} v$.
- $u * v = v * u$.
- $u * (v * w) = (u * v) * w$.
- $\partial^\alpha (u * v) = \partial^\alpha u * v = u * \partial^\alpha v$.

§6 Differential Equations

Let

$$P = \sum_{|\alpha| \le m} a_\alpha(x) \partial^\alpha : \mathscr{D}'(\Omega) \to \mathscr{D}'(\Omega). \quad (a_\alpha \in C^\infty(\Omega))$$

It is a **linear differential operator** of degree m. If a_α are constants, we call P to be constant coefficient.

If $a_\alpha = 0$ for every $|\alpha| < m$, we say P is homogenous.

Definition 6.0.1. For a differential operator P, we can define its **adjoint operator**

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

Then $\langle Pu, \varphi \rangle = \langle u, P^* \varphi \rangle$.

Definition 6.0.2. Given a constant coefficient linear differential operator P on $\mathbb{R}^n$, if the distribution $E \in \mathscr{D}'(\Omega)$ satisfies

$$P(E) = \delta_0,$$

we say E is a **basic solution**.

Example 6.0.3

The Laplace operator on $\mathbb{R}^n$ is

$$\Delta_g = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} = \sum_{i,j} \frac{1}{\sqrt{G}} \partial_j (g^{ij}\sqrt{G}\partial_i)$$

the latter one is the general form on a Riemann manifold, and we do not require it in this course.

The basic solution of Δ is

$$E(x) = \begin{cases} \dfrac{1}{2\pi} \log |x|, & n = 2 \\ \dfrac{1}{(2-n)|S^{n-1}|} \dfrac{1}{|x|^{n-2}}, & n \geq 3. \end{cases}$$

where $|S^{n-1}|$ is the surface area of S^{n-1} in $\mathbb{R}^n$.

Proof. When $x \neq 0$, $\Delta E = 0$. This follows from the polar form of the Laplace operator:

$$\Delta = \partial_{rr} + \frac{n-1}{r}\partial_r + \frac{1}{r^2}\Delta_{S^2}.$$

Then $\Delta E = (\partial_{rr} + \frac{n-1}{r}\partial_r)E(|x|) = 0$, which we won't check it here. □

Example 6.0.4

The heat operator $\partial_t - \Delta$ on $\mathbb{R} \times \mathbb{R}^n$ has basic solution

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

Here $H(t)$ is Heaviside function.

Proof. For all $\varphi \in C_0^\infty(\mathbb{R} \times \mathbb{R}^n)$,

$$\begin{aligned}
\langle (\partial t - \Delta)E(t,x), \varphi \rangle &= -\int_0^\infty \int_{\mathbb{R}^n} E(t,x)\partial_t \varphi \, \mathrm{d}t \, \mathrm{d}x - \int_0^\infty \int_{\mathbb{R}^n} E\Delta\varphi \, \mathrm{d}t \, \mathrm{d}x \\
&= -\lim_{\varepsilon \to 0} \int_\varepsilon^\infty \int_{\mathbb{R}^n} \partial_t(E\varphi) - \partial_t E\varphi + \Delta E\varphi \, \mathrm{d}t \, \mathrm{d}x \\
&= -\lim_{\varepsilon \to 0} \int_{\mathbb{R}^n} \frac{\varphi(t,x)}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}} \, \mathrm{d}x \Big|_\varepsilon^\infty \\
&= \lim_{\varepsilon \to 0} \int_{\mathbb{R}^n} \frac{\varphi(\varepsilon, 2\sqrt{\varepsilon}x)}{\pi^{\frac{n}{2}}} e^{-|x|^2} \, \mathrm{d}x = \varphi(0,0).
\end{aligned}$$

□

If E is the basic solution of P, we have

$$P(f * E) = f$$

that's why it's called the basic solution.

§6.1 Solutions of Laplace equation

Let E be the basic solution of Δ.

Lemma 6.1.1

If $u \in \mathscr{D}'(\mathbb{R}^n)$ has support inside a compact set K, then $E * u$ is smooth on K^c.

Proof. Let $\chi(x)$ be a truncation, $\chi_\varepsilon = \chi(\varepsilon^{-1}(x))$.

Since E is a smooth function on $\mathbb{R} \setminus \{0\}$, in a neighborhood of $p \in K^c$, apply the truncation and we're done since the convolution of a smooth function is also smooth. $\square$

Lemma 6.1.2

Let u be a distribution with compact support. If $\Delta u = 0$, then $u = 0$.

Proof. For all $f \in C^\infty$,

$$0 = \langle \Delta u, f \rangle = \langle \Delta u, \chi f \rangle = \langle u, \Delta(\chi f) \rangle = \langle u, \Delta \chi f + 2\nabla \chi \cdot \nabla f + x \Delta f \rangle = \langle u, \Delta f \rangle .$$

Where χ is a truncation with $\chi|_{\operatorname{supp} u} = 1$.

So for all $\varphi \in C_0^\infty(\mathbb{R}^n)$, $\varphi * E \in C^\infty$,

$$0 = \langle u, \Delta(\varphi * E) \rangle = \langle u, \varphi \rangle .$$

$\square$

Theorem 6.1.3

Let u be a harmonic distribution on $\mathbb{R}^n$, i.e. $\Delta u = 0$, then u is a smooth function.

Proof. For all constants $K > 0$, take $\chi(x)$ which is 1 on $\{|x| \le K\}$.

Consider

$$\Delta(\chi u) = \Delta \chi u + 2\nabla \chi \nabla u = f_\chi .$$

Here f_χ is a compact supported distribution.

TODO.. $\square$

§6.2 The basic solution of wave operator

Another important differential operator in physics is the wave operator. In the time-space $\mathbb{R}^{1+n}$, $t \in \mathbb{R}$, $x \in \mathbb{R}^n$. The metric is defined as

$$m = -\,\mathrm{d}t^2 + \mathrm{d}x^2$$

This is known as *Lorentz Geometry* (The one we're familiar with is called Riemann geometry).

In Riemann geometry, the Laplace operator is

$$\Delta_g = \frac{1}{\sqrt{G}} \partial_i (g^{ij} \sqrt{G} \partial_j)$$

and its counterpart in Lorentz geometry is the wave operator

$$\Box = -\partial_{tt}^2 + \Delta_x$$

Define a distribution on $\mathbb{R}$:

$$\chi_+^a = \frac{x_+^a}{\Gamma(a+1)}, \quad x_+ = \max\{0, x\}, a > -1.$$

Note that here we have

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

Hence we can define χ_+^a for $a < -1$ using derivatives.

Lemma 6.2.1

Let k be a nonnegative integer, we have

$$\chi_+^{-k} = \delta_0^{(k-1)}, \quad \chi_+^{-k-\frac{1}{2}} = \frac{1}{\sqrt{\pi}}\left(x_+^{-\frac{1}{2}}\right)^{(k)}.$$

Here $u^{(k)}$ means the k-th derivative of u.

Proof. By definition $\chi_+^{-1} = H'(x) = \delta_0$, and $\chi_+^{-\frac{1}{2}} = x_+^{-\frac{1}{2}}\frac{1}{\sqrt{\pi}}$. ☐

Definition 6.2.2. We say E_+ is a future basic solution of wave equation, if

- E_+ is a distribution on $\mathbb{R}^{1+n}$, such that
$$\Box E_+ = 0.$$
- The support of E_+ lies inside the light cone
$$\operatorname{supp}(E_+) \subset \{(t,x) \mid 0 \le |x| \le t\}.$$

Recall that the pullback of distribution:

Definition 6.2.3 (Pullbacks)**.** Let $\Omega_1 \subset \mathbb{R}^{n+m}$, $\Omega_2 \subset \mathbb{R}^n$. Let $\varphi : \Omega_1 \to \Omega_2$ be a smooth map with rank n, then for any $u \in \mathscr{D}'(\Omega_2)$, we define the **pullback** $\varphi^* u$ as follows:

Define $\Phi : \Omega_1 \to \Omega_2 \times \mathbb{R}^m$,

$$\Phi(x,y) = (\varphi(x,y), y).$$

For $\phi \in C_0^\infty(\mathbb{R}^{n+m})$, define

$$\langle u, \phi(x,y)\rangle = \left\langle u, \int_{\mathbb{R}^m} \phi(x,y)\,\mathrm{d}y\right\rangle$$

i.e. we can view u as a distribution on $\mathbb{R}^{n+m}$. Thus we can define

$$\varphi^* u = \Phi^* u(x,y).$$

Applying the chain rule,

$$\partial_j(\Phi^* u) = \sum_{k=1}^{n+m} \partial_k u \partial_j \Phi_k = \sum_{k=1}^{n} \partial_k u \partial_j \varphi_k.$$

Proposition 6.2.4 (Basic solution of wave operator)

$$E_+ = -\frac{\pi^{\frac{1-n}{2}}}{2} H(t)\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2).$$

Remark 6.2.5 — In the expression, $H(t)$ denotes it's the "future" part, the rest is a pullback of a distribution:

$$\varphi : \mathbb{R}^{1+n} \to \mathbb{R}, \quad \varphi(t,x) = t^2 - |x|^2,$$
$$\varphi^* \chi_+^{-\frac{1-n}{2}} = \chi_+^{-\frac{1-n}{2}}(t^2 - |x|^2).$$

However, φ is not full rank at 0, thus it requires extreme caution and we won't dive too deep into this issue.

Proof. Let $s = t^2 - |x|^2$, we have $\mathrm{d}t = \frac{1}{2\sqrt{s+|x|^2}}\,\mathrm{d}s$,

$$\left\langle \chi_+^{-\frac{n-1}{2}}(s), \varphi(t,x) \right\rangle = \left\langle \chi_+^{-\frac{n-1}{2}}(s), \frac{1}{2}\int_{\mathbb{R}^n} \frac{\varphi(\sqrt{s+|x|^2}, x)}{\sqrt{s+|x|^2}}\,\mathrm{d}x \right\rangle$$

When $s > 0$, this is well-defined; When $s = 0$, let $\psi(s)$ be the function on the right, we have $\psi(s)$ is C^k at $s = 0$ with $2k+1 < n$. (Since $\partial_k \psi(s) \approx \int \varphi(s + |x^2|)^{-k-\frac{1}{2}} \approx |x|^{-2k-1}$)

Therefore we can define E_+ in $\mathbb{R}^n$.

Inside the light cone,

$$\begin{aligned}
\Box\left(\chi_+^{-\frac{n-1}{2}}(t^2 - |x|^2)\right) &= -\partial_t\left(2t\chi_+^{-\frac{1+n}{2}}(t^2 - |x|^2)\right) - \partial_i\left(2x_i\chi_+^{-\frac{1+n}{2}}(t^2 - |x|^2)\right) \\
&= -2\chi_+^{-\frac{1+n}{2}}(t^2 - |x|^2) - 4t^2\chi_+^{-\frac{3+n}{2}}(t^2 - |x|^2) \\
&- 2n\chi_+^{-\frac{1+n}{2}}(t^2 - |x|^2) + 4x_i^2\chi_+^{-\frac{3+n}{2}}(t^2 - |x|^2) \\
&= -2(n+1)\chi_+^{-\frac{1+n}{2}}(t^2 - |x|^2) - 4(t^2 - |x|^2)\chi_+^{-\frac{3+n}{2}}(t^2 - |x|^2) \\
&= 0.
\end{aligned}$$

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

Now we have $\Box E_+$ supports on the origin,

$$\implies \Box E_+ = \sum_{|\alpha| \le N} C_\alpha \partial^\alpha \delta_0.$$

Note that E_+ is a homogenous distribution, specifically

$$E_+(\lambda x) = \lambda^{\frac{1-n}{2}} E_+(x).$$

This implies $\langle \Box E_+, \phi(\lambda t, \lambda x)\rangle = \langle \Box E_+, \phi(t,x)\rangle$. But $\langle \partial^\alpha \delta_0, \phi(\lambda t, \lambda x)\rangle = (-1)^{|\alpha|}\lambda^{|\alpha|}\partial^\alpha \phi(0,0)$, view it as a polynomial of λ, we must have $\Box E_+ = C\delta_0$.

At last we need to compute C, which should be done by find a suitable test function and compute the inner product. But here we'll just do this formally. (Actually I didn't take notes of the computation since it's too complicated) $\square$

Below we state the solution of wave equation, which we won't prove in this course.

Proposition 6.2.6 (Kirchhoff's formula)
The wave equation
$$\Box\phi = 0, \quad \phi(0,x) = 0, \partial_t\phi(0,x) = \phi_1(x)$$
has solution
$$\phi(t,x) = \begin{cases} \dfrac{1}{2}\displaystyle\int_{x-t}^{x+t} \phi_1(y)\,\mathrm{d}y, & n = 1 \\ \dfrac{\pi^{\frac{1-n}{2}}}{4}\left(\dfrac{1}{2t}\dfrac{\mathrm{d}}{\mathrm{d}t}\right)^{\frac{n-3}{2}}\left(t^{n-2}\displaystyle\int_{|\omega|=1}\phi_1(x+t\omega)\,\mathrm{d}\omega\right), & n \geq 3, 2\nmid n \\ \dfrac{\pi^{-\frac{n}{2}}}{2}\left(\dfrac{1}{2t}\dfrac{\mathrm{d}}{\mathrm{d}t}\right)^{\frac{n-2}{2}}\left(t^{n-1}\displaystyle\int_{|y|\leq 1}\dfrac{\phi_1(x+ty)}{\sqrt{1-|y|^2}}\,\mathrm{d}y\right), & n \geq 2, 2 \mid n. \end{cases}$$

§7 Fourier transformations of distributions

§7.1 Tempered distributions

Like the smooth functions, $f \in C_0^\infty(\mathbb{R}^n) \implies \widetilde{f} \notin C_0^\infty(\mathbb{R}^n)$, we defined a different space $S(\mathbb{R}^n)$ for Fourier transformations.

So we'll also define a space for distributions that is closed under Fourier transformations. For $p \geq 1$, denote
$$N_p(\varphi) = \sum_{|\alpha|,|\beta|\leq p} \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta \varphi|.$$

Definition 7.1.1. We say u is a **tempered distribution** on $\mathbb{R}^n$, if u is a linear map from $S(\mathbb{R}^n)$ to $\mathbb{R}$, such that there exists C and p,
$$|\langle u, \varphi\rangle| \leq CN_p(\varphi), \quad \forall\varphi \in S(\mathbb{R}^n).$$
Write this space as $S'(\mathbb{R}^n)$.

Since $C_0^\infty(\mathbb{R}^n) \subset S(\mathbb{R}^n)$, $\mathscr{D}'(\mathbb{R}^n) \supset S'(\mathbb{R}^n)$.

Given $u \in S'(\mathbb{R}^n)$ and multi-index α, β, the distributions $\partial^\alpha u, x^\beta u \in S'(\mathbb{R}^n)$ as well. But for a function f larger than polynomials (e.g. e^x), fu may not lie in $S'(\mathbb{R}^n)$.

This means that u cannot increase faster than polynomials, that's why it is called "slowly increasing".

Check:
$$|\langle \partial^\alpha u, \varphi\rangle| = |\langle u, \partial^\alpha\varphi\rangle| \leq CN_p(\partial^\alpha\varphi) \leq CN_{p+|\alpha|}(\varphi).$$

The convergence in S' is defined as
$$u_k \xrightarrow{S'} u \iff \lim_{k\to\infty} \langle u_k, \varphi\rangle = \langle u, \varphi\rangle, \quad \forall\varphi \in S(\mathbb{R}^n).$$

Example 7.1.2

Let $1 \le p \le +\infty$, the functions in L^p are tempered distributions.

Take k s.t. $kp' \ge n$.

$$\begin{aligned} |\langle f, \varphi \rangle| = \left| \int f\varphi \, \mathrm{d}x \right| &\le \|f\|_{L^p} \|\varphi\|_{L^{p'}} \\ &\le \|f\|_{L^p} \|(1+|x|)^{-k}\|_{L^{p'}} N_k(\varphi) \\ &\le C N_k(\varphi) \end{aligned}$$

Example 7.1.3

Distributions with compact support are tempered. Let $K = \operatorname{supp} u$.

$$\begin{aligned} |\langle u, \varphi \rangle| = |\langle u, \chi\varphi \rangle| &\le C \sum_{|\alpha| \le p} \|\partial^\alpha (\chi\varphi)\|_{L^\infty(K)} \\ &\le C \sum_{|\alpha| \le p} \|\partial^\alpha \varphi\|_{L^\infty(K)} \le C N_p(\varphi). \end{aligned}$$

Example 7.1.4

The distribution $pv\frac{1}{x}$ is tempered. Just split the integral to $[0,1]$ and $(1,+\infty)$, then control each part by $N_1(\varphi)$.

Example 7.1.5

Exponential increasing distributions can also be tempered. Let $u = ie^x e^{ie^x} = (e^{ie^x})'$. Since e^{ie^x} has norm 1, it is tempered $\implies$ u as its derivative is also tempered.

§7.2 Fourier transformations

Definition 7.2.1. Let $u \in S'(\mathbb{R}^n)$, for all $\varphi \in S(\mathbb{R}^n)$, define

$$\langle \widehat{u}, \varphi \rangle = \langle u, \widehat{\varphi} \rangle, \quad \left\langle \mathcal{F}^{-1}(u), \varphi \right\rangle = \left\langle u, \mathcal{F}^{-1}(\varphi) \right\rangle.$$

Recall that $\int \widehat{f} g \, \mathrm{d}x = \int f \widehat{g} \, \mathrm{d}x$ for rapidly decreasing functions, so this coincides with the Fourier transformations of functions.

Since

$$\begin{aligned} |\langle \widehat{u}, \varphi \rangle| = |\langle u, \widehat{\varphi} \rangle| \le C N_p(\widehat{\varphi}) = C \sum_{|\alpha|,|\beta| \le p} \sup |\xi^\alpha \partial^\beta \widehat{\varphi}| \\ \le C \sum \|\widehat{x^\alpha \partial^\beta \varphi}\|_{L^\infty} \\ \le C \sum \|x^\alpha \partial^\beta \varphi\|_{L^1} \le C N_{p+n+1}(\varphi). \end{aligned}$$

The Fourier transformation $\widehat{u} \in S'(\mathbb{R}^n)$.

Theorem 7.2.2

The Fourier transformation
$$\mathcal{F} : S'(\mathbb{R}^n) \to S'(\mathbb{R}^n)$$
is a continuous linear isomorphism. Where the continuity means if $u_k \xrightarrow{S'} u$, then $\widehat{u}_k \xrightarrow{S'} \widehat{u}$.

Moreover,
$$\widehat{\partial_k u} = 2\pi\xi_k i\widehat{u}, \quad \widehat{2\pi x_k u} = i\partial_k \widehat{u}, \quad \mathcal{F}^{-1}(\widehat{u}) = u.$$

Example 7.2.3

The dirac function δ_0 satisfies $\widehat{\delta}_0 = 1$.

For all $a \in \mathbb{R}^n$ and multi-index α, we have
$$\widehat{\partial^\alpha \delta_a} = (2\pi i\xi)^\alpha \widehat{\delta}_a = (2\pi i\xi)^\alpha e^{-2\pi i a\cdot\xi}.$$
Thus $a = 0$ yields
$$\widehat{x^\alpha} = \frac{(-1)^{|\alpha|}}{(2\pi i)^\alpha}\delta_0.$$
In particular, $\widehat{1} = \delta_0$.

Proposition 7.2.4

Let $u \in S'(\mathbb{R}^n)$ be harmonic, i.e. $\Delta u = 0$, then u must be a polynomial.

Proof. By Fourier transformation,
$$\widehat{\Delta u} = |2\pi i\xi|^2 \widehat{u} = 0 \implies |\xi|^2 \widehat{u} = 0.$$
Thus $\widehat{u}$ supports on the origin.
$$\widehat{u} = \sum_{|\alpha|\le N} C_\alpha \partial^\alpha \delta_0 \implies u \text{ is a polynomial.}$$
□

Example 7.2.5

The Fourier transformation of $pv\frac{1}{x}$.
$$x \cdot pv\frac{1}{x} = 1 \implies \widehat{x \cdot pv\frac{1}{x}} = \delta_0.$$
Hence
$$i\left(\widehat{pv\frac{1}{x}}\right)' = \delta_0 2\pi \implies \frac{\mathrm{d}}{\mathrm{d}\xi}\widehat{pv\frac{1}{x}} = -2\pi i\delta_0 \implies \widehat{pv\frac{1}{x}} = -2\pi i H(x) + C.$$
Since $pv\frac{1}{x}$ is an odd distribution, i.e. $\langle u, \varphi(-x)\rangle = -\langle u, \varphi(x)\rangle$, we have $C = \pi i$.

This tells us $\widehat{H}(\xi) = \frac{1}{2}\delta_0 - \frac{i}{2\pi}pv\frac{1}{x}$.