1. The Calderon-Zygmund theory.

1.1. **Some example of singular integrals: Hilbert transform and Riesz transform.** We introduce the following tempered distribution $W_0 \in \mathcal{S}'$ as follows

$$\langle W_0, \phi \rangle = \frac{1}{\pi} \lim_{\epsilon \to 0} \int_{|x|>\epsilon} \frac{\phi(x)}{x} dx, \qquad \phi \in \mathcal{S}.$$

To prove that $W_0 \in \mathcal{S}'$ we split

$$\langle W_0, \phi \rangle = \frac{1}{\pi} \lim_{\epsilon \to 0} \int_{1>|x|>\epsilon} \frac{\phi(x) - \phi(0)}{x} dx + \frac{1}{\pi} \int_{|x|\geq 1} \frac{\phi(x)}{x} dx$$

and estimate $\langle W_0, \phi \rangle \leq \frac{2}{\pi} \|\phi'\|_{L^\infty} + \frac{2}{\pi} \|x\phi(x)\|_{L^\infty}$.

The Hilbert transform of $f \in \mathcal{S}$ is defined by

$$Hf(x) = (W_0 * f)(x) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \int_{|y|>\varepsilon} \frac{f(x-y)}{y} dy.$$

The integral above does not converge absolutely even for $f \in \mathcal{S}$! It will be important in later arguments that we do not absolute values too early so as to exploit subtle cancellations.

Recall that if $f \in \mathcal{S}$ then $W_0 * f \in C^\infty$. This limit is called the principal value integral and denoted by

$$(Hf)(x) = \frac{1}{\pi} p.v. \int_{-\infty}^{\infty} \frac{f(x-y)}{y} dy.$$

We will extend to domain of H to larger family of functions (larger than $\mathcal{S}$).

Example. By direct computation one checks that

$$H(\chi_{[a,b]})(x) = \frac{1}{\pi} \ln |\frac{x-a}{x-b}|.$$

Since Hf is a convolution on the physical side, it is expected that on the Fourier side is a product-type operator. We now make this precise by computing $\hat{W}_0$:

$$\begin{aligned}
\langle \hat{W}_0, \phi \rangle &= \langle W_0, \hat{\phi} \rangle \\
&= \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{|\xi| \geq \varepsilon} \hat{\phi}(\xi) \frac{d\xi}{\xi} \\
&= \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\frac{1}{\epsilon} \geq |\xi| \geq \varepsilon} \int_{\mathbb{R}} \phi(x) e^{-2\pi i x \xi} dx \frac{d\xi}{\xi} \\
&= \int_{\mathbb{R}} \left(\lim_{\varepsilon \to 0} \frac{1}{\pi} \int_{\frac{1}{\epsilon} \geq |\xi| \geq \varepsilon} e^{-2\pi i x \xi} \frac{d\xi}{\xi} \right) \phi(x) dx \\
&= \int_{\mathbb{R}} \left(\lim_{\varepsilon \to 0} \frac{-i}{\pi} \int_{\frac{1}{\epsilon} \geq |\xi| \geq \varepsilon} \sin(2\pi x \xi) \frac{d\xi}{\xi} \right) \phi(x) dx
\end{aligned}$$

The exercise 4.1.1 from Grafakos computes the inner integral:

$$\lim_{\varepsilon\to 0}\int_{\frac{1}{\epsilon}\geq|\xi|\geq\varepsilon}\sin(2\pi x\xi)\frac{d\xi}{\xi}=\pi sgn(x)$$

Therefore

$$\langle\hat{W}_0,\phi\rangle=\int\phi(x)(-isgn(x))dx$$

hence $\hat{W}_0(\xi)=-i\cdot sgn(\xi)$. Hence

$$\mathcal{F}(Hf)(\xi)=-i\cdot sgn(\xi)\hat{f}(\xi).$$

This is sometimes an alternative definition of the Hilbert transform. Note that we obtain for free that $\|Hf\|_{L^2}=\|f\|_{L^2}$, thus H is an isometry on L^2. Since $(-isgn(\xi))^2=-1$ it follows that $H^2=-I$ and its adjoint defined by

$$\langle Hf,g\rangle_s=\langle f,H^*g\rangle_s$$

satisfies $H^*=-H$; here we denoted by $\langle f,g\rangle_s=\int f(x)\bar{g}(x)dx$ the usual scalar product.

A fundamental question is whether H is bounded on other L^p spaces. The L^2 boundedness was easy due to the nice formulation of the Hilbert on the Fourier side and the Plancherel formula. This cannot be adapted to any other L^p space. We also saw that H is not bounded on L^∞ in the example $f=\chi_{[a,b]}$. Since $H\chi_{[a,b]}$ decays like $|x|^{-1}$ at ∞ hence it is not in L^1 either; but it is in $L^{1,\infty}$! One can settle this question in full on its own, but the Calderon-Zygmund theory of singular integral provides a general result which covers a larger class of such operators.

One would like to extend the Hilbert transform to higher dimensions. There are n operators called the Riesz transforms with similar properties.For $1\leq j\leq n$ we define W_j as follows

$$\langle W_j,\phi\rangle=\frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}\lim_{\varepsilon\to 0}\int_{|y|\geq\varepsilon}\frac{y_j}{|y|^{n+1}}\phi(y).$$

One can check that $W_j\in\mathcal{S}'$. Then we define the j'th Riesz transform of f by

$$R_j(f)(x)=(f*W_j)(x)=\frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}p.v\int\frac{x_j-y_j}{|x-y|^{n+1}}\phi(y)dy$$

for $\phi\in\mathcal{S}$. A more involved computations shows that

$$\mathcal{F}(R_jf)(\xi)=-i\frac{\xi_j}{|\xi|}\hat{f}(\xi).$$

From this it follows that R_j are bounded on L^2. There is no reason to believe that they are bounded on L^∞ or L^1 since this is not the case in the case $n=1$. The boundedness of R_j on other L^p spaces will be derived from the Calderon-Zygmund theory.

The Riesz operators satisfy

$$\sum_j R_j^2 = -I$$

and this can be verified on the Fourier side.

Example. A fundamental PDE is the Laplace's equation

$$\Delta u = f.$$

One the Fourier side this takes the form $(-4\pi^2|\xi|^2)\hat{u}(\xi) = \hat{f}(\xi)$. We can estimate all second derivatives of u by using

$$\mathcal{F}(\partial_j\partial_k u) = (2\pi i\xi_j)(2\pi i\xi_k)\hat{u} = \frac{(2\pi i\xi_j)(2\pi i\xi_k)}{-4\pi^2|\xi|^2}\hat{f} = \mathcal{F}(-R_jR_kf).$$

so $\partial_j\partial_k u = -R_jR_kf$. In particular this shows that $\|\partial_j\partial_k u\|_{L^2} \leq \|\Delta u\|_{L^2}$. But similar conclusion hold true for other L^p spaces.

Note that the above operators belong to a larger class. If Ω is an integrable function on the unit sphere $\mathbb{S}^{n-1}$ with mean value zero, then we can define the kernel

$$K_\Omega(x) = \frac{\Omega(\frac{x}{|x|})}{|x|^n}$$

which is homogeneous of degree $-n$. We define the distribution

$$\langle W_\Omega, \phi\rangle = \lim_{\varepsilon\to 0}\int_{|x|\geq\varepsilon} K_\Omega(x)\phi(x)dx.$$

Based on this we define the singular integral

$$T_\Omega(f)(x) = (f * W_\Omega)(x) = \lim_{\varepsilon\to 0}\int_{|y|\geq\epsilon} f(x-y)\frac{\Omega(\frac{y}{|y|})}{|y|^n}dy.$$

One can write down an explicit formula for $\hat{W}_\Omega$, but one needs some conditions on Ω to obtain the L^2 boundedness of T_Ω.

1.2. **The Calderon-Zygmund decomposition.** This result a key step in establishing the Calderon-Zygmund theory.

By a dyadic cube in $\mathbb{R}^n$ we mean a cube of the form $[2^km_1, 2^k(m_1+1)] \times .... \times [2^km_n, 2^k(m_n+1)]$.

Theorem 1. *Let $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$. Then there exist functions g and b on $\mathbb{R}^n$ such that*

i) $f = g + b$,

ii) $\|g\|_{L^1} \leq \|f\|_{L^1}$ and $\|g\|_{L^\infty} \leq 2^n\alpha$,

iii) $b = \sum_j b_j$, where b_j is supported in a dyadic cube Q_j. Furthermore the cubes $\{Q_j\}_{j\in\mathbb{N}}$ are mutually disjoint,

iv) $\int_{Q_j} b_j = 0$,

v) $\|b_j\|_{L^1} \leq 2^{n+1}\alpha|Q_j|$,

vi) $\sum_j |Q_j| \leq \alpha^{-1}\|f\|_{L^1}$.

For a proof see Grafakos.

1.3. **Singular integrals.** Let K be a measurable function defined on $\mathbb{R}^n \setminus \{0\}$ satisfying
i) $|K(x)| \leq C|x|^{-n}$,
ii) $|\nabla K(x)| \leq C|x|^{-n+1}$,
iii) $\lim_{\epsilon\to\infty} \int_{1\geq|x|\geq\epsilon} K(x)dx = L$.

From i) and iii) it follows that there is a tempered distribution $W \in S'$ which coincides with K away from the origin defined by

$$W(\phi) = \lim_{\varepsilon\to 0} \int K(x)\phi(x)dx$$

where the argument to show that $W \in \mathcal{S}'$ is similar to the one used for the Hilbert transform. Based on this we define the operator T by

$$Tf(x) = (W * f)(x) = p.v. \int K(y)f(x-y)dy = \lim_{\varepsilon\to 0} \int_{|y|\geq\varepsilon} K(y)f(x-y)dy.$$

The conditions imposed on K can be relaxed.
i) can be replaced by $\int_{R\leq|x|\leq 2R} |K(x)|dx \leq C$
ii) can be replaced by $\sup_{|y|>0} \int_{|x|\geq 2|y|} |K(x-y) - K(x)|dx \leq C$
iii) the full limit can be replaced by a limit on a subsequence going to 0, though this may cause the output operators to depend on the choice of the limit. See the example in Grafakos.

An easier way to see the above ambiguity is as follows. Given a linear operator T, K is the kernel of T if

$$Tf(x) = \int K(y)f(x-y)dy$$

for any f with support compact and x outside the support of f. K is uniquely determined by T. However T is not uniquely determined by K. Indeed let $Tf = cf$ then it follows that $K = 0$ regardless of the value of c. In other words $K(x) + c\delta_0$ give rise to the same operators. If T is assumed to be bounded as an operator from L^2 to L^2, this is in fact the only ambiguity which can arise when we try to reconstruct T from K. In some sense this is also at the heart of the example regarding part iii) above.

The main result of this section is the following

Theorem 2. *Assume K is defined on $\mathbb{R}^n \setminus \{0\}$ and satisfies i)-iii). Assume that the operator T constructed above is bounded on $L^2(\mathbb{R}^n)$. Then T has an extension that maps L^1 to $L^{1,\infty}$; moreover T extends to an operator from L^p to L^p for any $1 < p < \infty$.*

Such operators are called Calderon-Zygmund type operators.
See proof in Grafakos.

As an application we will prove the convergence of the partial Fourier sums in L^p. Let $f \in \mathcal{S}$ and define the partial sums of f by

$$S_N f(x) = \int_{-N}^{N} e^{2\pi i x\cdot\xi} \hat{f}(\xi) d\xi.$$

We claim that for any $1 < p < \infty$ the operators S_N are uniformly bounded on L^p, i.e.

$$\|S_N f\|_{L^p} \leq C_p \|f\|_{L^p} \tag{1}$$

The first step is to rewrite S_N in terms of more familiar operators. Since

$$\chi_{[-N,N]}(\xi) = \frac{1}{2}(sgn(\xi + N) - sgn(\xi - N)) = \frac{1}{2}(\tau^{-N} sgn(\xi) - \tau^N sgn(\xi))$$

it follows

$$\mathcal{F}(S_N f) = \chi_{[-N,N]}(\xi)\hat{f}(\xi) = \frac{1}{2}\left(\tau^{-N}(sgn(\xi)\tau^N \hat{f}) - \tau^N(sgn(\xi)\tau^{-N}\hat{f})\right)$$

where recall that $\tau^N \hat{f}(\xi) = \hat{f}(\xi - N)$. Therefore

$$S_N f = \frac{i}{2}\left(M_{-N} H M_N f - M_N H M_{-N} f\right)$$

where $(M_{\xi_0} f)(x) = e^{2\pi i x \xi_0} f(x)$. The multiplication by characters is a bounded operator on any L^p since $|e^{2\pi i x\xi_0}| = 1$. Since H is bounded on L^p for $1 < p < \infty$, it follows that S_N are uniformly bounded on L^p.

From (1) we can also conclude that if $f \in L^p$ then $\lim_{N\to\infty} S_N f = f$ in L^p. Indeed, since $\mathcal{S}$ is dense in L^p we write $f = s + g$ with $s \in S$ and $\|g\|_{L^p} \leq \epsilon$, use a soft argument to show that $\lim_{N\to\infty} \|S_N s - s\|_{L^p} = 0$ and estimate

$$\|S_N f - f\|_{L^p} \leq \|S_N s - s\|_{L^p} + \|S_N g - g\|_{L^p}$$

to obtain

$$\lim_{N\to\infty} \|S_N f - f\|_{L^p} \leq \sup_N \|S_N g\|_{L^p} + \|g\|_{L^p} \leq C_p \epsilon$$

from which the conclusion follows by taking $\epsilon \to 0$.

1.4. **BMO.** A measurable function f is of bounded mean oscillation if

$$\|f\|_{BMO} = \sup_B \frac{1}{|B|} \int_B |f(x) - f_B|$$

where the sup above is taken over the set of all balls and $f_B = \frac{1}{|B|}\int f dx$ is the average of f on B. Note that $\|\cdot\|_{BMO}$ is not a norm since all constant functions have size 0. However if functions are defined up to a constant, i.e. $f = g$ means $f(x) = g(x) + c$ then $\|\cdot\|_{BMO}$ becomes a norm.

Note that f_B can be replaced with any constant c_B; indeed, if that was the case, it follows that

$$|f_B - c_B| = \frac{1}{|B|}\int_B |f_B - c_B| dx = \frac{1}{|B|}\int_B |\frac{1}{|B|}\int (f(y) - c_B) dy| dx \leq A$$

and from this, one obtains $\|f\|_{BMO} \leq 2A$.

It is obvious that $\|f\|_{BMO} \leq 2\|f\|_{L^\infty}$. In fact BMO is strictly larger than L^∞. The standard example of a BMO function not in L^∞ is $\log|x|$ (Check this!). One can show that $\log|P(x)| \in BMO$.

Recall that $H(\chi_{[a,b]})(x) = \frac{1}{\pi}\ln|\frac{x-a}{x-b}| \in BMO$, hence one may expect that H takes L^∞ to BMO. This is the case indeed.

But first we need to define Tf since $\mathcal{S}$ is not dense in L^∞ and the usual definition by extension will not work. Recall that formally

$$Tf(x) = p.v. \int K(x-y)f(y)dy$$

and that the main problem arrises from the part of the integral near $x = y$ as well as from the part near ∞. Let B be a ball containing both 0 and x; then we define:

$$Tf(x) = T(\chi_B f)(x) + \int_{B^c} (K(x-y) - K(-y))f(y)dy$$

where now the integral is absolutely convergent and $\chi_B f \in L^p$ for any $1 \leq p \leq \infty$ where we have already defined T. The definition is unique modulo a constant in the sense that if we change B then we get the same function modulo a constant. An instructive exercise is to check that this definition of T agrees with the standard one on, let's say on $L^2 \cap L^\infty$, modulo a constant.

Theorem 3. *If T is a Calderon-Zygmund operator (see Theorem 2) then T extends as a bounded operator from L^∞ to BMO.*

Proof. Given a ball B, let $2B$ to be its double and write

$$Tf = T(\chi_{2B}f) + T(\chi_{(2B)^c}f).$$

Since T is bounded on L^2 we have

$$\|T(\chi_{2B}f)\|_{L^1(B)} \leq |B|^{\frac{1}{2}}\|T(\chi_{2B}f)\|_{L^2} \leq C|B|^{\frac{1}{2}}\|\chi_{2B}f\|_{L^2} \leq C|B|\|f\|_{L^\infty}$$

where we have used the Cauchy-Schwartz inequality several times.

Next, we have

$$T(\chi_{(2B)^c}f)(x) = \int_{(2B)^c} (K(x-y) - K(-y))f(y)dy + c$$

for some constant c. Using the property iii) of K, it follows that $|T(\chi_{(2B)^c}f)(x) - c| \leq C$, hence

$$\frac{1}{|B|}\int_B |T(\chi_{(2B)^c}f)(x) - c| \leq C.$$

□

In fact it can be shown that T maps BMO to BMO.

We end this section by listing a few more facts about the BMO space. Functions in BMO are "nearly" bounded. There are several ways in which

one can quantify this. The first way is via the inequality

$$\int \frac{|f(x)-f_{B_1}|}{\log^{2+\epsilon}(2+|x|)(1+|x|)^n}dx \le c\|f\|_{BMO}. \tag{2}$$

for some $\varepsilon > 0$ and where $B_1 = B_1(0)$. To prove this we start from the definition of BMO which gives

$$\int_{B_{2^k}} |f(x)-f_{B_{2^k}}|dx \le c2^{nk}\|f\|_{BMO}$$

where $B_{2^k} = B_{2^k}(0)$. Then we have

$$|f_{B_{2^{k+1}}} - f_{B_{2^k}}| \le |f_{B_{2^{k+1}}} - f(x)| + |f(x) - f_{B_{2^k}}|$$

for any x, which we integrate on B_{2^k} to obtain

$$2^{nk}|f_{B_{2^{k+1}}}-f_{B_{2^k}}| \le \int_{B_{2^k}} |f_{B_{2^{k+1}}}-f(x)|+\int_{B_{2^k}} |f(x)-f_{B_{2^k}}| \le c(2^{nk}+2^{n(k+1)})\|f\|_{BMO}.$$

Therefore $|f_{B_{2^{k+1}}}-f_{B_{2^k}}| \lesssim \|f\|_{BMO}$. From this it follows that $|f_{B_{2^k}}-f_{B_1}| \lesssim k\|f\|_{BMO}$ which then gives

$$\int_{B_{2^k}} |f(x)-f_{B_1}|dx \lesssim k2^{nk}\|f\|_{BMO}.$$

We then have

$$\int \frac{|f(x)-f_{B_1}|}{\log^{2+\epsilon}(2+|x|)(1+|x|)^{n+1}}dx \lesssim \sum_{k=1}^{\infty}\int_{B_{2^k}} \frac{|f(x)-f_{B_1}|}{k^{2+\epsilon}2^{nk}}dx \lesssim \sum_{k=1}^{\infty} k^{-1-\epsilon} \lesssim 1$$

(2) hints that at ∞, $f(x)$ can grow at most of log type. On the other hand, locally, (2) does not say anything more than the function is locally integrable; this is not a very strong conclusion since $|x|^{-n+\epsilon}$ is locally integrable, but it grows near the origin faster than $\log|x|$. We will be able to obtain some extra information using a different argument.

Let $f \in BMO$ with $\|f\|_{BMO} = 1$. Therefore for every ball B we have

$$\int_B |f(x)-f_B| \le |B|.$$

From this it follows that for any $\lambda \ge 1$ we have

$$|\{x \in B : |f(x)-f_B| \ge \lambda\}| \le \frac{|B|}{\lambda}.$$

One reads this as follows: f exceeds its average by λ on at most λ^{-1} fraction of the set B. Based on this one obtains the following

Theorem 4 (John-Nirenberg inequality)**.** *Let $f \in BMO$. Then for every ball B we have*

$$|\{x \in B : |f(x)-f_B| \ge \lambda\|f\|_{BMO}\}| \lesssim e^{-c_n\lambda}|B|$$

for any $\lambda > 0$ and for some constant c_n depending only on n.

Proof. Normalize $\|f\|_{BMO} = 1$. Define $A(\lambda)$ to be the best constant such that

$$|\{x \in B : f(x) \geq f_B + \lambda| \leq A(\lambda)|B|.$$

for any ball B. We already know that $A(\lambda) \leq \min(1, \lambda^{-1})$ and we want to improve this to $A(\lambda) \lesssim e^{-c_n \lambda}$. Let λ_0 be a large constant to be chosen later, and we consider $\lambda > \lambda_0$. For a ball B define $F(x) = \max((f(x) - f_B)\chi_{2B}, 0)$. We estimate

$$\|F\|_{L^1} \leq \int_{2B} |f - f_B| dx \lesssim |B|$$

by using a similar argument as in the previous theorem (the above is obvious if f_B is replaced by f_{2B} given that $\|f\|_{BMO} = 1$). We consider the set $E = B \cap \{MF > \lambda_0\}$.

For $\varepsilon > 0$ small, let K be any compact subset of E such that $|E \setminus K| \leq \varepsilon$. For each $x \in K$ there is a ball B_x such that $F_{B_x} = \lambda_0$.

Using (a variation of) the Vitali decomposition lemma, we obtain a cover

$$K \subset \cup_k B_k$$

where B_k is a finite family balls intersecting K(the finite part is obtained by using the compactness of K) and such that

$$F_{B_n} = \lambda_0, \qquad \sum_k |B_k| \lesssim_d \frac{\|F\|_{L^1}}{\lambda_0} \lesssim \frac{|B|}{\lambda_0}.$$

Since B_k intersects $K \subset B$ we can conclude that $B_k \subset 2B$ for λ_0 large enough.

Using the Lebesgue differentiation theorem, it follows

$$\begin{aligned} |\{x \in B : |F(x)| \geq \lambda\}| &= |\{x \in B : |F(x)| \geq \lambda, MF(x) \geq \lambda_0\}| \\ &\leq \sum_k |\{x \in B : |F(x)| > \lambda\} \cap B_k| + \varepsilon \end{aligned}$$

Since $B_k \subset 2B$, it follows that on each B_k we have

$$f(x) = f_B + F(x) + g(x)$$

with $g \leq 0$; hence

$$f_{B_k} \leq f_B + F_{B_k} = f_B + \lambda_0$$

Hence we obtain

$$|\{x \in B : F(x) > \lambda\} \cap B_k| \leq |\{x \in B_k : f(x) - f_{B_k} > \lambda - \lambda_0\}| \leq A(\lambda - \lambda_0)|B_k|$$

Adding these estimates with respect to k gives

$$|\{x : |F(x)| \geq \lambda\}| \leq A(\lambda - \lambda_0) \sum_k |B_k| + \varepsilon \lesssim A(\lambda - \lambda_0) \frac{|B|}{\lambda_0} + \varepsilon.$$

Since the estimates are independent of ε, we can take $\varepsilon \to 0$. Then taking supremum over all the balls B gives us

$$A(\lambda) \leq c_n \frac{A(\lambda - \lambda_0)}{\lambda_0}.$$

From this the conclusion follows. □

Using the distributional characterization of the L^p norm, one can prove the following alternative characterization of the BMO

Corollary 1. *For any $1 \leq p < \infty$, the following holds true:*

$$\sup_B \left(\frac{1}{|B|} \int |f - f_B|^p dx \right)^{\frac{1}{p}} \approx_{p,n} \|f\|_{BMO}$$

The proof is left as an exercise.

One can also prove that BMO is a good substitute for the L^∞ for interpolation purposes. For instance, if $1 \leq p < q < \infty$ the following inequality holds true:

$$\|f\|_{L^q} \lesssim \|f\|_{L^p}^{\frac{p}{q}} \|f\|_{BMO}^{1-\frac{p}{q}}.$$

2. Multiplier theory

Here I follow closely Tao's notes, see Notes 4 there.