Singular Integrals and L^p estimates
Lecture Notes

Swarnendu Sil*

Department of Mathematics
Indian Institute of Science
Bangalore, India

NCM Workshop
May 20-June 1 2024, TIFR CAM

Contents

*swarnendusil@iisc.ac.in

1 Newtonian potential

1.1 Fourier transform and L^2 estimate

Let $n \geq 2$ be an integer and let $f \in C_c^\infty(\mathbb{R}^n)$ and consider the following problem

$$-\Delta u = f \qquad \text{in } \mathbb{R}^n. \tag{1}$$

We are interested in estimating the second derivatives of u in terms of f in L^p norms. To get an idea of what we are up against, first let us gather some information. Taking Fourier transform on both sides, we arrive at

$$\hat{f} = -\left[\sum_{j=1}^{n} (i\xi_j)^2\right]\hat{u} = |\xi|^2\,\hat{u}.$$

Thus, at least formally, we have

$$\hat{u} = \frac{1}{|\xi|^2}\hat{f}. \tag{2}$$

This is quite useful in many respects. We would soon use this to write down a fundamental solution of the Laplacian. But for now, notice that by properties of Fourier transform, we have

$$\left(\frac{\partial^2}{\partial x_j \partial x_k} u\right)^{\hat{}} = (i\xi_j)(i\xi_j)\,\hat{u} = \xi_j\xi_k\hat{u} = \frac{\xi_j\xi_k}{|\xi|^2}\hat{f}.$$

Using Parseval identity, we have

$$\begin{aligned}\left\|\frac{\partial^2 u}{\partial x_j \partial x_k}\right\|^2_{L^2(\mathbb{R}^n)} &= \left\|\left(\frac{\partial^2}{\partial x_j \partial x_k}u\right)^{\hat{}}\right\|^2_{L^2(\mathbb{R}^n)} \\ &= \int_{\mathbb{R}^n} \frac{\xi_j^2 \xi_k^2}{|\xi|^4} \left|\hat{f}(\xi)\right|^2 \, \mathrm{d}\xi \\ &\leq \int_{\mathbb{R}^n} \frac{\xi_j^2 \xi_k^2}{|\xi|^4} \left|\hat{f}(\xi)\right|^2 \, \mathrm{d}\xi = \|f\|^2_{L^2(\mathbb{R}^n)}.\end{aligned}$$

So far, so good. We have at least managed our goal in this simple case for $p = 2$. However, our technique here can only work for L^2, as we have made crucial use of the Parseval identity. We have also made crucial use of the fact that the 'Fourier multiplier' is bounded, i.e.

$$m(\xi) = \frac{\xi_j \xi_k}{|\xi|^2} \in L^\infty(\mathbb{R}^n).$$

To understand the significance of this, let us look at our problem differently. Assuming the solution u to our problem is unique (this can be ensured by quite mild extra conditions, e.g. by requiring $u \in L^2(\mathbb{R}^n)$) and smooth (this we shall show momentarily) define the linear operator $T_{jk} : C_c^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ by

$$T_{jk} f := \frac{\partial^2 u}{\partial x_j \partial x_k},$$

where u is the unique solution of (1). What we have proved now is that we have the estimate

$$\|T_{jk} f\|_{L^2(\mathbb{R}^n)} \leq \|f\|_{L^2(\mathbb{R}^n)}.$$

Thus, T_{jk} extends as a bounded linear operator from $L^2(\mathbb{R}^n)$ to itself. To see what kind of an operator this is, we suppose we can find a 'function' K_{jk} such that

$$\hat{K}_{jk} = \frac{1}{(2\pi)^{\frac{n}{2}}} m(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{\xi_j \xi_k}{|\xi|^2}.$$

Then we can write, at least formally,

$$T_{jk} f = \frac{\partial^2 u}{\partial x_j \partial x_k} = \left[\left(\frac{\partial^2 u}{\partial x_j \partial x_k}\right)^{\hat{}}\right]^{\check{}} = \left((2\pi)^{\frac{n}{2}} \hat{K}_{jk} \hat{f}\right)^{\check{}} = K_{jk} * f.$$

Hence, at least formally, T_{jk} is an integral operator of convolution type, with a kernel K_{jk}. Unfortunately, K_{jk} is not really a 'function', not even a locally integrable one and the ‘convolution’ is quite problematic to define. To understand this kernel better, we first try to find a ‘kernel’ for u itself. Suppose we can find a ‘function’ $\mathcal{N}$ such that

$$\hat{\mathcal{N}} = \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{|\xi|^2}.$$

Then we can write

$$u = (\hat{u})^{\check{}} = \left(\frac{1}{|\xi|^2}\hat{f}\right)^{\check{}} = \left((2\pi)^{\frac{n}{2}} \hat{\mathcal{N}}_2 \hat{f}\right)^{\check{}} = \left([\mathcal{N} * f]^{\hat{}}\right)^{\check{}} = \mathcal{N} * f.$$

Trying to find $\mathcal{N}$ by inverse Fourier transform is also not trivial at all. The reason is very simple. There simply is no such function in $L^1(\mathbb{R}^n)$! It is easy to check that the function

$$\xi \mapsto \frac{1}{|\xi|^2}$$

is not in $L^\infty(\mathbb{R}^n)$ nor in $L^2(\mathbb{R}^n)$. So if there exists any such function $\mathcal{N}$, such a function clearly can not be either in $L^1(\mathbb{R}^n)$ or in $L^2(\mathbb{R}^n)$. Nonetheless, luckily for us, there exists a **locally integrable** function $\mathcal{N} \in L^1_{\text{loc}}(\mathbb{R}^n)$ with these properties when $n \geq 3$.

Lemma 1. *Let $\alpha \in \mathbb{R}$ such that $0 < \alpha < n$. Let*

$$f(x) = \frac{1}{|x|^\alpha} \qquad \text{for } x \in \mathbb{R}^n.$$

Then we have

$$\hat{f}(\xi) = \frac{2^{\frac{n}{2}}\Gamma\left(\frac{n-\alpha}{2}\right)}{2^\alpha \Gamma\left(\frac{\alpha}{2}\right)} \frac{1}{|x|^{n-\alpha}}.$$

The proof of the lemma is beyond the scope of this course, as this requires us to work with tempered distributions.

Returning back to our problem, we see that

$$\mathcal{N}(x) = \frac{c_2}{|x|^{n-2}}$$

for some constant $c_2 > 0$. Thus, if we define

$$u(x) := c_2 \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-2}}\, \mathrm{d}y \qquad \text{for } x \in \mathbb{R}^n,$$

this solves (1), at least in the sense of tempered distributions. Proving that this defines a strong solution is actually not immediate. First we note that

$$c_2 = \frac{1}{(n-2)\left|\mathbb{S}^{n-1}\right|}.$$

Thus, the formula for u is

$$u(x) := \frac{1}{(n-2)\left|\mathbb{S}^{n-1}\right|} \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-2}}\, \mathrm{d}y \qquad \text{for } x \in \mathbb{R}^n \tag{3}$$

Definition 2. *The kernel*

$$\mathcal{N}(x) := \frac{1}{(n-2)\left|\mathbb{S}^{n-1}\right|} \frac{1}{|x|^{n-2}} \qquad \text{for all } x \in \mathbb{R}^n \setminus \{0\},$$

is called the Newtonian kernel in n dimensions for $n \geq 3$ and the operator

$$Nf := \mathcal{N} * f$$

is called the Newtonian potential for f.

As the singularity of kernel is locally integrable around the origin, due to the fact that $n-2 < n$, the integral defining the convolution operator Nf is actually what is called a 'fractional integral' and is not a singular integral. These integrals exist and are easy to estimate.

1.2 Fundamental solution and singular integrals

Now we prove that the formal candidate formula (3) indeed defines a smooth solution of (1).

Theorem 3. *Let $n \geq 3$. For any $f \in C_c^\infty(\mathbb{R}^n)$, the function u defined by (3) is $C^\infty(\mathbb{R}^n)$ and satisfies the Poisson equation in the whole space*

$$-\Delta u = f \qquad \text{in } \mathbb{R}^n. \tag{4}$$

Proof. We need to show that the function

$$u(x) = \int_{\mathbb{R}^n} \mathcal{N}(x-y) f(y) \, \mathrm{d}y$$

solves (4) for $f \in C_c^\infty(\mathbb{R}^n)$. First we note that $\mathcal{N}$ is locally integrable. Thus, by properties of convolution and the fact that f has compact support implies that

$$D^\alpha u = \mathcal{N} * D^\alpha f$$

and thus $u \in C^\infty(\mathbb{R}^n)$. This also yields,

$$\Delta u = \mathcal{N} * \Delta f.$$

Now we want to show that the RHS is actually equal to $-f$. Now note that by simple computations, formally we have

$$\frac{\partial \mathcal{N}}{\partial x_j}(x) = -\frac{1}{\left|\mathbb{S}^{n-1}\right|} \frac{x_j}{|x|^n} \quad \text{and} \quad \frac{\partial^2 \mathcal{N}}{\partial x_j \partial x_k}(x) = -\frac{1}{\left|\mathbb{S}^{n-1}\right|} \left[\frac{\delta_{jk}}{|x|^n} - \frac{n x_j x_k}{|x|^{n+2}} \right],$$

where δ_{jk} is the Kronecker delta. From this, we deduce the following growths

$$|\nabla \mathcal{N}| \simeq \frac{c}{|x|^{n-1}} \qquad \text{and} \qquad \left|\nabla^2 \mathcal{N}\right| \simeq \frac{c}{|x|^n}.$$

So the second derivatives of $\mathcal{N}$ are not locally integrable around the origin. Since the only trouble is at the origin, to isolate the trouble, we pick an arbitrary $\varepsilon > 0$ and write

$$\begin{aligned}
\mathcal{N} * \Delta f &= \int_{\mathbb{R}^n} \mathcal{N}(y)\, \Delta_x f(x-y)\, \mathrm{d}y \\
&= \int_{B(0,\varepsilon)} \mathcal{N}(y)\, \Delta_x f(x-y)\, \mathrm{d}y + \int_{\mathbb{R}^n \setminus B(0,\varepsilon)} \mathcal{N}(y)\, \Delta_x f(x-y)\, \mathrm{d}y \\
&:= I_\varepsilon + J_\varepsilon.
\end{aligned}$$

I_ε can be easily estimated by the local integrability of $\mathcal{N}$. Indeed, we have

$$|I_\varepsilon| \le \left\|\nabla^2 f\right\|_{L^\infty(\mathbb{R}^n)} \int_{B(0,\varepsilon)} |\mathcal{N}(y)|\, \mathrm{d}y \le C\varepsilon^2.$$

Thus,

$$I_\varepsilon \to 0 \qquad \text{as } \varepsilon \to 0.$$

For computing J_ε, note that $\mathcal{N}$ is smooth away from the origin and thus, integrating by parts twice, we obtain

$$\begin{aligned}
J_\varepsilon &= \int_{\mathbb{R}^n \setminus B(0,\varepsilon)} \mathcal{N}(y)\, \Delta_x f(x-y)\, \mathrm{d}y \\
&= \int_{\mathbb{R}^n \setminus B(0,\varepsilon)} \mathcal{N}(y)\, (-1)^2\, \Delta_y f(x-y)\, \mathrm{d}y \\
&= \int_{\mathbb{R}^n \setminus B(0,\varepsilon)} \mathcal{N}(y)\, \Delta_y f(x-y)\, \mathrm{d}y \\
&= \int_{\mathbb{R}^n \setminus B(0,\varepsilon)} \Delta\mathcal{N}(y)\, f(x-y)\, \mathrm{d}y + \int_{\partial B(0,\varepsilon)} \mathcal{N}(y)\, \frac{\partial f}{\partial \nu}(x-y)\, \mathrm{d}\Sigma_y \\
&\qquad - \int_{\partial B(0,\varepsilon)} \frac{\partial \mathcal{N}}{\partial \nu}(y)\, f(x-y)\, \mathrm{d}\Sigma_y,
\end{aligned}$$

where ν denotes the **inward** normal on $\partial B(0,\varepsilon)$, since that is the *outward* normal from the side of $\mathbb{R}^n \setminus B(0,\varepsilon)$. Now it is easy to check that $\Delta\mathcal{N} = 0$ in $\mathbb{R}^n \setminus B(0,\varepsilon)$ and thus we have

$$\begin{aligned}
J_\varepsilon &= \int_{\partial B(0,\varepsilon)} \mathcal{N}(y)\, \frac{\partial f}{\partial \nu}(x-y)\, \mathrm{d}\Sigma_y - \int_{\partial B(0,\varepsilon)} \frac{\partial \mathcal{N}}{\partial \nu}(y)\, f(x-y)\, \mathrm{d}\Sigma_y \\
&:= J_\varepsilon^1 + J_\varepsilon^2.
\end{aligned}$$

The estimate of J_ε^1 is similar to I_ε. We have

$$\left|J_\varepsilon^1\right| \le \|\nabla f\|_{L^\infty(\mathbb{R}^n)} \int_{\partial B(0,\varepsilon)} |\mathcal{N}(y)|\, \mathrm{d}\Sigma_y \le C\varepsilon.$$

Thus,

$$J_\varepsilon^1 \to 0 \qquad \text{as } \varepsilon \to 0.$$

So we have shown

$$
\begin{aligned}
\mathcal{N} * \Delta f &= \lim_{\varepsilon \to 0} J_{\varepsilon}^{2} \\
&= - \lim_{\varepsilon \to 0} \int_{\partial B(0,\varepsilon)} \frac{\partial \mathcal{N}}{\partial \nu} (y) \, f(x-y) \, \mathrm{d}\Sigma_y \\
&= - \lim_{\varepsilon \to 0} \int_{\partial B(0,\varepsilon)} \langle \nu(y), \nabla \mathcal{N}(y) \rangle f(x-y) \, \mathrm{d}\Sigma_y \\
&= - \lim_{\varepsilon \to 0} \int_{\partial B(0,\varepsilon)} \left\langle -\frac{y}{|y|}, -\frac{1}{|\mathbb{S}^{n-1}|} \frac{y}{|y|^{n}} \right\rangle f(x-y) \, \mathrm{d}\Sigma_y \\
&= -\frac{1}{|\mathbb{S}^{n-1}|} \lim_{\varepsilon \to 0} \int_{\partial B(0,\varepsilon)} \frac{1}{|y|^{n-1}} f(x-y) \, \mathrm{d}\Sigma_y \\
&= - \lim_{\varepsilon \to 0} \left(\frac{1}{|\mathbb{S}^{n-1}| \, \varepsilon^{n-1}} \int_{\partial B(0,\varepsilon)} f(x-y) \, \mathrm{d}\Sigma_y \right) \\
&= - \lim_{\varepsilon \to 0} ⨍_{\partial B(0,\varepsilon)} f(x-y) \, \mathrm{d}\Sigma_y \\
&= -f(x).
\end{aligned}
$$

This completes the proof. □

From what we have proved above, clearly, we have

$$
\frac{\partial^2 u}{\partial x_j \partial x_k} = \frac{\partial^2}{\partial x_j \partial x_k} [\mathcal{N} * f] = \frac{\partial^2 \mathcal{N}}{\partial x_j \partial x_k} * f,
$$

where the last equality is only formal so far. Thus, taking our clue from this, our guess for the kernel K_{jk} for the operator T_{jk} is given by

$$
K_{jk}(x) := -\frac{1}{|\mathbb{S}^{n-1}|} \left[\frac{\delta_{jk}}{|x|^{n}} - \frac{n x_j x_k}{|x|^{n+2}} \right].
$$

Thus, we can now try to 'define' our operator T_{jk} as

$$
T_{jk} f := K_{jk} * f.
$$

More explicitly,

$$
\begin{aligned}
T_{jk} f(x) &= \left[\frac{\partial^2 \mathcal{N}}{\partial x_j \partial x_k} * f \right] (x) \\
&= \int_{\mathbb{R}^n} \frac{\partial^2 \mathcal{N}}{\partial x_j \partial x_k} (x-y) \, f(y) \, \mathrm{d}y \\
&= -\frac{1}{|\mathbb{S}^{n-1}|} \int_{\mathbb{R}^n} \left[\frac{\delta_{jk}}{|x-y|^{n}} - \frac{n (x_j - y_j)(x_k - y_k)}{|x-y|^{n+2}} \right] f(y) \, \mathrm{d}y.
\end{aligned}
$$

To get an idea of the trouble, suppose for the moment that $K_{jk} \in L^1(\mathbb{R}^n)$. Then using Young's inequality for convolutions, we would have

$$\|T_{jk}f\|_{L^p(\mathbb{R}^n)} = \|K_{jk} * f\|_{L^p(\mathbb{R}^n)} \leq \|K_{jk}\|_{L^1(\mathbb{R}^n)} \|f\|_{L^p(\mathbb{R}^n)}$$

for **every** $1 \leq p \leq \infty$, i.e. **including** $p = 1$ **and** $p = \infty$. However, since K_{jk} is **not integrable and not even locally integrable around the origin**. Again, taking our clue from the computations for the Newtonian potential, we want to cut out the singularity and analyze operators given by

$$T_\varepsilon f(x) = \int_{\mathbb{R}^n \setminus B_\varepsilon(x)} K(x-y) f(y) \, dy, \tag{5}$$

where the kernel K has the form

$$K(x) = \frac{\theta(x)}{|x|^n}, \tag{6}$$

where $\theta \in L^\infty(\mathbb{R}^n)$ is a **bounded measurable homogeneous function of degree** 0. Note that in the case of the kernels K_{jk} above, we have

$$\theta(x) = -\frac{1}{|\mathbb{S}^{n-1}|}\left[\delta_{jk} - \frac{n x_j x_k}{|x|^2}\right]. \tag{7}$$

Note that $T_\varepsilon f$ is a nice convolution operator with an integrable kernel for every $\varepsilon > 0$. So we can hope to define the operator

$$Tf := \lim_{\varepsilon \to 0} T_\varepsilon f,$$

in the sense of *Cauchy principal value*. However, this simply is false without further assumptions, as the following simple result shows.

Proposition 4. *Let* $f = \mathbb{1}_{[-1,1]}$ *and for every* $\varepsilon > 0$, *consider*

$$T_\varepsilon f(x) := \int_{|x-t|>\varepsilon} \frac{f(t)}{|x-t|} \, dt.$$

Then we have

$$\lim_{\varepsilon \to 0} T_\varepsilon f(x) = +\infty \qquad \text{for every } x \in [-1,1].$$

The reason for this difficulty is that

$$\lim_{\varepsilon \to 0} \int_{B_1(0)\setminus B_\varepsilon(0)} K(x) \, dx = \lim_{\varepsilon \to 0} \int_{B_1(0)\setminus B_\varepsilon(0)} \frac{\theta(x)}{|x|^n} \, dx$$

need not exist. Recalling complex analysis, one might be tempted to think that we should cut out the singularity by some other way and not balls. First let us prove that this is not the case.

Proposition 5. *Let $\Omega \subset \mathbb{R}^n$ be any open set such that $0 \in \Omega$. Let U, V be any two open neighborhoods of 0 in $\mathbb{R}^n$. Then*

$$\lim_{\varepsilon \to 0} \int_{\Omega \setminus \varepsilon U} \frac{\theta(x)}{|x|^n}\, \mathrm{d}x \quad \textit{exists iff} \quad \lim_{\varepsilon \to 0} \int_{\Omega \setminus \varepsilon V} \frac{\theta(x)}{|x|^n}\, \mathrm{d}x \quad \textit{exists.}$$

Proof. The proof is elementary. For $\varepsilon > 0$ small enough, we have $\varepsilon U, \varepsilon V \subset \Omega$. Now we have

$$\begin{aligned} \int_{\Omega \setminus \varepsilon U} \frac{\theta(x)}{|x|^n}\, \mathrm{d}x - \int_{\Omega \setminus \varepsilon V} \frac{\theta(x)}{|x|^n}\, \mathrm{d}x &= \int_{\varepsilon(V \setminus U)} \frac{\theta(x)}{|x|^n}\, \mathrm{d}x - \int_{\varepsilon(U \setminus V)} \frac{\theta(x)}{|x|^n}\, \mathrm{d}x \\ &= \int_{V \setminus U} \frac{\theta(\varepsilon x)}{|\varepsilon x|^n} \varepsilon^n\, \mathrm{d}x - \int_{U \setminus V} \frac{\theta(\varepsilon x)}{|\varepsilon x|^n} \varepsilon^n\, \mathrm{d}x \\ &= \int_{V \setminus U} \frac{\theta(\varepsilon x)}{|x|^n}\, \mathrm{d}x - \int_{U \setminus V} \frac{\theta(\varepsilon x)}{|x|^n}\, \mathrm{d}x \\ &= \int_{V \setminus U} \frac{\theta(x)}{|x|^n}\, \mathrm{d}x - \int_{U \setminus V} \frac{\theta(x)}{|x|^n}\, \mathrm{d}x. \end{aligned}$$

Since the domains of integration on the right does not contain the singularity, the result follows. ☐

1.3 Cancellation property and L^2 estimate

As we have seen above, we can not expect to make sense of the singular integrals in the principal value sense without further hypothesis. However, for our kernel K_{jk}, we have already proved the L^2 estimates. So now we investigate under what additional assumptions L^2 estimates would hold. First, observe that our kernel satisfies a remarkable property, which will turn out to be essential to what we would be doing.

Proposition 6. *For any $1 \leq j, k \leq n$, if θ is given by (7), then we have*

$$\int_{\mathbb{S}^{n-1}} \theta(y)\, \mathrm{d}\Sigma_y = 0. \tag{8}$$

The proof is left as an exercise. An immediate, but very useful consequence of this observation is the following.

Proposition 7. *Let K be given (6), where θ satisfies (8). Then for any $0 < R_1 < R_2 < \infty$, we have*

$$\int_{R_1 < |x| < R_2} K(x)\, \mathrm{d}x = 0.$$

Is this property important? Indeed it is. This is in fact equivalent to the existence of integral of the kernel around the singularity in the principal value sense, as we now show.

Theorem 8. *The limit*

$$\lim_{\varepsilon\to 0}\int_{B_1(0)\setminus B_\varepsilon(0)} \frac{\theta(x)}{|x|^n}\ \mathrm{d}x$$

exists if and only if θ satisfies (8).

Proof. We have

$$\begin{aligned}\int_{B_1(0)\setminus B_\varepsilon(0)} \frac{\theta(x)}{|x|^n}\ \mathrm{d}x &= \int_\varepsilon^1 \rho^{n-1}\left(\int_{\mathbb{S}^{n-1}} \frac{1}{\rho^n}\theta(\rho\zeta)\ \mathrm{d}\Sigma_\zeta\right)\ \mathrm{d}\rho \\ &= \int_\varepsilon^1 \frac{1}{\rho}\ \mathrm{d}\rho \int_{\mathbb{S}^{n-1}} \theta(\zeta)\ \mathrm{d}\Sigma_\zeta \\ &= -\log\varepsilon \int_{\mathbb{S}^{n-1}} \theta(\zeta)\ \mathrm{d}\Sigma_\zeta = \log\left(\frac{1}{\varepsilon}\right)\int_{\mathbb{S}^{n-1}} \theta(\zeta)\ \mathrm{d}\Sigma_\zeta.\end{aligned}$$

This clearly blows up as $\varepsilon\to 0$ if and only if (8) is violated. □

Before proceeding, we first set some terminology.

Definition 9. *A function $K:\mathbb{R}^n\setminus\{0\}\to\mathbb{R}$ is called a* ***Calderon-Zygmund kernel*** *or* ***CZ kernel*** *if*

(a) K is positively homogeneous of degree $-n$, i.e.

$$K(x) = \frac{\theta(x/|x|)}{|x|^n} \qquad \text{for all } x\in\mathbb{R}^n\setminus\{0\},$$

for some measurable function $\theta:\mathbb{S}^{n-1}\to\mathbb{R}$,

(b) $\theta\in L^\infty\left(\mathbb{S}^{n-1}\right)$ and

(c) θ enjoys the cancellation property (8).

The cancellation property (and that $\theta\in L^\infty$) already implies the L^2 boundedness (see Theorem 7.20 in [3] for a proof). However, we would prove the result under additional regularity assumptions on K. There are several reasons for this. Firstly, this would make our life a lot simpler. Secondly, the additional assumption would anyway be needed to prove the L^p boundedness for $p\neq 2$ and finally, our kernel K_{jk} satisfies an even stronger property. Namely, our kernel K_{jk} enjoys good regularity properties away from zero. More precisely, $K_{jk}\in C^\infty(\mathbb{R}^n\setminus\{0\})$ and its derivative decays fast enough away from the origin.

Proposition 10. *$K_{jk}\in C^\infty(\mathbb{R}^n\setminus\{0\})$ and there exists a constant $B>0$, depending only on n, such that*

$$|\nabla K_{jk}(x)| \leq \frac{B}{|x|^{n+1}} \qquad \text{for } x\in\mathbb{R}^n\setminus\{0\}.$$

Proof. This is just easy computation. We can directly show that we have

$$\frac{\partial^3 \mathcal{N}}{\partial x_j \partial x_k \partial x_l}(x) = -\frac{1}{|\mathbb{S}^{n-1}|}\left[\frac{n(n+2)\, x_j x_k x_l}{|x|^{n+4}} - \frac{n\delta_{jk} x_l}{|x|^{n+2}} - \frac{n\delta_{jl} x_k}{|x|^{n+2}} - \frac{n\delta_{kl} x_j}{|x|^{n+2}}\right].$$

The estimate is obvious now. □

As a consequence of this regularity, we have the following.

Proposition 11. *There exists a constant $C > 0$ such that for every $1 \leq j, k \leq n$, we have*

$$\sup_{y \neq 0} \int_{|x|>2|y|} |K_{jk}(x-y) - K_{jk}(x)| \, \mathrm{d}x \leq C.$$

Proof. We have

$$|K_{jk}(x-y) - K_{jk}(x)| \leq \sup_{t \in [0,1]} |\nabla K_{jk}(x-ty)|\, |y| \leq B \sup_{t \in [0,1]} \frac{|y|}{|x-ty|^{n+1}}.$$

Now if $|x| > 2|y|$, then for any $t \in [0,1]$, we have

$$|x| \leq |x-ty| + |ty| \leq |x-ty| + |y| \leq |x-ty| + \frac{1}{2}|x|.$$

This implies $|x-ty| \geq |x|/2$ and thus, we deuce

$$|K_{jk}(x-y) - K_{jk}(x)| \leq B \sup_{t \in [0,1]} \frac{|y|}{|x-ty|^{n+1}} \leq 2^{n+1} B \frac{|y|}{|x|^{n+1}}.$$

Integrating, we have

$$\int_{|x|>2|y|} |K_{jk}(x-y) - K_{jk}(x)| \, \mathrm{d}x \leq C|y| \int_{|x|>2|y|} \frac{1}{|x|^{n+1}} \, \mathrm{d}x = C.$$

This completes the proof. □

Now we need another definition.

Definition 12 (Hörmander condition)**.** *A CZ kernel K is said to satisfy the **Hörmander condition** if there exists a constant $C > 0$ such that*

$$\sup_{y \neq 0} \int_{|x|>2|y|} |K(x-y) - K(x)| \, \mathrm{d}x \leq C.$$

Now we prove

Theorem 13 (L^2 estimate)**.** *Let K be CZ kernel satisfying the Hörmander condition. Let $f \in L^2(\mathbb{R}^n)$. For any $\varepsilon > 0$, define the operators*

$$T_\varepsilon f(x) = \int_{\mathbb{R}^n \setminus B_\varepsilon(x)} K(x-y) f(y) \, \mathrm{d}y \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

Then we have

(i) *For every* $\varepsilon > 0$, $T_\varepsilon f \in L^2(\mathbb{R}^n)$ *and there exists a constant* $A_2 > 0$, *independent of* f *and* $\varepsilon > 0$, *such that we have the estimates*

$$\|T_\varepsilon f\|_{L^2(\mathbb{R}^n)} \leq A_2 \|f\|_{L^2(\mathbb{R}^n)}.$$

(ii) $T_\varepsilon f$ *converges to a limit, denoted by* Tf *in* $L^2(\mathbb{R}^n)$ *as* $\varepsilon \to 0$ *and the map* $f \mapsto Tf$ *defines a bounded linear operator from* $L^2(\mathbb{R}^n)$ *to itself and satisfies*

$$\|Tf\|_{L^2(\mathbb{R}^n)} \leq A_2 \|f\|_{L^2(\mathbb{R}^n)}.$$

Proof. We first prove (i). The kernel for the operator T_ε is

$$K_\varepsilon(x) = \begin{cases} K(x) & \text{if } |x| > \varepsilon, \\ 0 & \text{if } |x| \leq \varepsilon. \end{cases}$$

Clearly, K_ε is a CZ kernel for every $\varepsilon > 0$ and satifies the Hörmander condition with a constant that depends on the Hörmander condition constant of K and the dimension n, but is independent of ε. Note that $K_\varepsilon \in L^2(\mathbb{R}^n)$ for every $\varepsilon > 0$. We want to show that $\widehat{K_\varepsilon}$ is in $L^\infty(\mathbb{R}^n)$ and the L^∞ norm is bounded independently of $\varepsilon > 0$. This would prove the uniform bound in (i).

We begin by showing that for $\varepsilon = 1$, the Fourier transform is a bounded function, i.e. we show $\widehat{K_1}$ is in $L^\infty(\mathbb{R}^n)$. We have

$$\begin{aligned} \widehat{K_1}(\xi) &= \int_{\mathbb{R}^n} e^{-i\langle \xi, x\rangle} K_1(x) \, dx \\ &= \int\limits_{|x| < \frac{2\pi}{|\xi|}} e^{-i\langle \xi, x\rangle} K_1(x) \, dx + \int\limits_{|x| \geq \frac{2\pi}{|\xi|}} e^{-i\langle \xi, x\rangle} K_1(x) \, dx \\ &= \int\limits_{1 < |x| < \frac{2\pi}{|\xi|}} e^{-i\langle \xi, x\rangle} K_1(x) \, dx + \int\limits_{|x| \geq \frac{2\pi}{|\xi|}} e^{-i\langle \xi, x\rangle} K_1(x) \, dx \\ &:= I_1 + I_2. \end{aligned}$$

To estimate I_1, we use the cancellation property. We have

$$\begin{aligned} I_1 &= \int\limits_{1 < |x| < \frac{2\pi}{|\xi|}} e^{-i\langle \xi, x\rangle} K_1(x) \, dx \\ &= \int\limits_{1 < |x| < \frac{2\pi}{|\xi|}} \left[e^{-i\langle \xi, x\rangle} - 1\right] K_1(x) \, dx. \end{aligned}$$

Now we use the inequality

$$\left|e^{it} - 1\right| \leq |t| \qquad \text{for } t \in \mathbb{R}.$$

This and the Cauchy-Schwarz inequality yields the estimate

$$\begin{aligned}
|I_1| &\leq \int\limits_{1<|x|<\frac{2\pi}{|\xi|}} \left|e^{-i\langle \xi, x\rangle} - 1\right| |K_1(x)| \, \mathrm{d}x \\
&\leq \int\limits_{1<|x|<\frac{2\pi}{|\xi|}} |\langle \xi, x\rangle| \, |K_1(x)| \, \mathrm{d}x \\
&\leq |\xi| \int\limits_{1<|x|<\frac{2\pi}{|\xi|}} |x| \, |K_1(x)| \, \mathrm{d}x \\
&\leq C |\xi| \int\limits_{1<|x|<\frac{2\pi}{|\xi|}} |x| \frac{1}{|x|^n} \, \mathrm{d}x \\
&\leq C |\xi| \int\limits_{0<|x|<\frac{2\pi}{|\xi|}} \frac{1}{|x|^{n-1}} \, \mathrm{d}x = 2\pi C.
\end{aligned}$$

For I_2, we plan to use the Hörmander condition. We set $z = \pi \frac{\xi}{|\xi|^2}$. The choice is dictated by the fact that for this z, we have

$$e^{i\langle \xi, z\rangle} = e^{i\pi} = -1.$$

Now we write

$$\begin{aligned}
&\int_{\mathbb{R}^n} e^{-i\langle \xi, x\rangle} K_1(x) \, \mathrm{d}x \\
&\quad = \frac{1}{2} \int_{\mathbb{R}^n} e^{-i\langle \xi, x\rangle} K_1(x) \, \mathrm{d}x + \frac{1}{2} \int_{\mathbb{R}^n} e^{-i\langle \xi, x\rangle} K_1(x) \, \mathrm{d}x \\
&\quad = \frac{1}{2} \int_{\mathbb{R}^n} e^{-i\langle \xi, x\rangle} K_1(x) \, \mathrm{d}x + \frac{1}{2} \int_{\mathbb{R}^n} e^{-i\langle \xi, y-z\rangle} K_1(y-z) \, \mathrm{d}y \\
&\quad = \frac{1}{2} \int_{\mathbb{R}^n} e^{-i\langle \xi, x\rangle} K_1(x) \, \mathrm{d}x + \frac{1}{2} \int_{\mathbb{R}^n} e^{-i\langle \xi, y\rangle} e^{i\langle \xi, z\rangle} K_1(y-z) \, \mathrm{d}y \\
&\quad = \frac{1}{2} \int_{\mathbb{R}^n} e^{-i\langle \xi, x\rangle} K_1(x) \, \mathrm{d}x - \frac{1}{2} \int_{\mathbb{R}^n} e^{-i\langle \xi, y\rangle} K_1(y-z) \, \mathrm{d}y \\
&\quad = \frac{1}{2} \int_{\mathbb{R}^n} e^{-i\langle \xi, x\rangle} \left[K_1(x) - K_1(x-z)\right] \, \mathrm{d}x
\end{aligned}$$

Thus, we have

$$
\begin{aligned}
&\int\limits_{|x|\geq\frac{2\pi}{|\xi|}} e^{-i\langle\xi,x\rangle}K_1(x)\ \mathrm{d}x\\
&=\frac{1}{2}\int\limits_{|x|\geq\frac{2\pi}{|\xi|}} e^{-i\langle\xi,x\rangle}\left[K_1(x)-K_1(x-z)\right]\ \mathrm{d}x\\
&\quad+\frac{1}{2}\int\limits_{|x|<\frac{2\pi}{|\xi|}} e^{-i\langle\xi,x\rangle}\left[K_1(x)-K_1(x-z)\right]\ \mathrm{d}x-\int\limits_{|x|<\frac{2\pi}{|\xi|}} e^{-i\langle\xi,x\rangle}K_1(x)\ \mathrm{d}x\\
&=\frac{1}{2}\int\limits_{|x|\geq\frac{2\pi}{|\xi|}} e^{-i\langle\xi,x\rangle}\left[K_1(x)-K_1(x-z)\right]\ \mathrm{d}x\\
&\quad-\frac{1}{2}\int\limits_{|x|<\frac{2\pi}{|\xi|}} e^{-i\langle\xi,x\rangle}K_1(x-z)\ \mathrm{d}x-\frac{1}{2}\int\limits_{|x|<\frac{2\pi}{|\xi|}} e^{-i\langle\xi,x\rangle}K_1(x)\ \mathrm{d}x\\
&=\frac{1}{2}\int\limits_{|x|\geq\frac{2\pi}{|\xi|}} e^{-i\langle\xi,x\rangle}\left[K_1(x)-K_1(x-z)\right]\ \mathrm{d}x\\
&\quad+\frac{1}{2}\int\limits_{|y+z|<\frac{2\pi}{|\xi|}} e^{-i\langle\xi,y\rangle}K_1(y)\ \mathrm{d}y-\frac{1}{2}\int\limits_{|x|<\frac{2\pi}{|\xi|}} e^{-i\langle\xi,x\rangle}K_1(x)\ \mathrm{d}x\\
&=\frac{1}{2}\int\limits_{|x|\geq\frac{2\pi}{|\xi|}} e^{-i\langle\xi,x\rangle}\left[K_1(x)-K_1(x-z)\right]\ \mathrm{d}x+\frac{1}{2}\int\limits_{\substack{|x|<\frac{2\pi}{|\xi|},\\ |x+z|<\frac{2\pi}{|\xi|}}} e^{-i\langle\xi,x\rangle}K_1(x)\ \mathrm{d}x\\
&:=J_1+J_2.
\end{aligned}
$$

Now note if $|x+z|<2\pi/|\xi|$, then we have

$$
|x|\leq|x+z|+|z|<\frac{2\pi}{|\xi|}+\frac{\pi}{|\xi|}=\frac{3\pi}{|\xi|}.
$$

Hence we have

$$
\begin{aligned}
|J_2| &\leq \frac{1}{2}\left|\int\limits_{\substack{|x|<\frac{2\pi}{|\xi|},\\ |x+z|<\frac{2\pi}{|\xi|}}} e^{-i\langle \xi,x\rangle} K_1(x)\ \mathrm{d}x\right| \\
&\leq \frac{1}{2}\int\limits_{\substack{|x|<\frac{2\pi}{|\xi|},\\ |x+z|<\frac{2\pi}{|\xi|}}} \left|e^{-i\langle \xi,x\rangle}\right| |K_1(x)|\ \mathrm{d}x \\
&\leq \frac{C}{2}\int_{\frac{2\pi}{|\xi|}}^{\frac{3\pi}{|\xi|}} r^{n-1}\frac{1}{r^n}\ \mathrm{d}r \\
&= \frac{C}{2}\log\left(\frac{3\pi/|\xi|}{2\pi/|\xi|}\right) = \frac{C}{2}\log\left(\frac{3}{2}\right).
\end{aligned}
$$

For J_1, we use the Hörmander condition to deduce

$$
\begin{aligned}
|J_1| &\leq \frac{1}{2}\int\limits_{|x|\geq\frac{2\pi}{|\xi|}} \left|e^{-i\langle \xi,x\rangle}\right| |K_1(x)-K_1(x-z)|\ \mathrm{d}x \\
&\leq \frac{1}{2}\int\limits_{|x|\geq\frac{2\pi}{|\xi|}} |K_1(x)-K_1(x-z)|\ \mathrm{d}x \\
&= \frac{1}{2}\int\limits_{|x|\geq 2|z|} |K_1(x)-K_1(x-z)|\ \mathrm{d}x \leq C,
\end{aligned}
$$

where in the last line, we have used the fact that $|z| = \pi/|\xi|$.

This settles the case $\varepsilon = 1$. For the general case, fixe $\varepsilon > 0$ and note that if we define the kernel

$$K'(x) := \varepsilon^n K(\varepsilon x),$$

then it is easy to check that K' also satisfies the hypotheses of the result with the **same constants as** K. Thus, our previous result for applied to K' implies that if

$$K_1'(x) := \begin{cases} K'(x) & \text{if } |x| > 1, \\ 0 & \text{if } |x| \leq 1, \end{cases}$$

then we have

$$\left|\widehat{K_1'}(\xi)\right| \leq C, \qquad \text{for a.e. } \xi \in \mathbb{R}^n.$$

But now it is easy to check that we have

$$\widehat{K_\varepsilon}(\xi) = \widehat{K_1'}(\varepsilon\xi).$$

Hence we have proved that $\widehat{K_\varepsilon} \in L^\infty(\mathbb{R}^n)$ and we have

$$\left\|\widehat{K_\varepsilon}\right\|_{L^\infty(\mathbb{R}^n)} \leq C,$$

where C is a constant independent of $\varepsilon > 0$. This completes the proof of (i).

Now we prove (ii). By (i), we have shown that for any $f \in L^2(\mathbb{R}^n)$, the sequence $\{T_\varepsilon f\}_{\varepsilon>0}$ is uniformly bounded in $L^2(\mathbb{R}^n)$. Since $L^2(\mathbb{R}^n)$ is a Banach space, to prove (ii), it is enough to prove that the sequence $\{T_\varepsilon f\}_{\varepsilon>0}$ is Cauchy. To this end, assume $0 < \delta < \varepsilon$ and fix $\eta > 0$. Since $C_c^\infty(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, we can find $g \in C_c^\infty(\mathbb{R}^n)$ such that we have

$$\|f - g\|_{L^2(\mathbb{R}^n)} < \eta.$$

Now we have

$$\begin{aligned}\|T_\varepsilon f - T_\delta f\|_{L^2(\mathbb{R}^n)} \leq \|T_\varepsilon g - T_\delta g\|_{L^2(\mathbb{R}^n)} + \|T_\varepsilon (f-g)\|_{L^2(\mathbb{R}^n)} \\ + \|T_\delta (f-g)\|_{L^2(\mathbb{R}^n)}.\end{aligned}$$

By the uniform bound, this implies

$$\begin{aligned}\|T_\varepsilon f - T_\delta f\|_{L^2(\mathbb{R}^n)} &\leq \|T_\varepsilon g - T_\delta g\|_{L^2(\mathbb{R}^n)} + 2A_2 \|f-g\|_{L^2(\mathbb{R}^n)} \\ &\leq \|T_\varepsilon g - T_\delta g\|_{L^2(\mathbb{R}^n)} + 2A_2\eta.\end{aligned}$$

Now we claim that we have

$$\lim_{\delta,\varepsilon\to 0} \|T_\varepsilon g - T_\delta g\|_{L^2(\mathbb{R}^n)} = 0$$

for any $g \in C_c^\infty(\mathbb{R}^n)$. The claim implies the result. Clearly, assuming the claim, we would have

$$\lim_{\delta,\varepsilon\to 0} \|T_\varepsilon f - T_\delta f\|_{L^2(\mathbb{R}^n)} \leq 2A_2\eta.$$

Since $\eta > 0$ is arbitrary, this means $\{T_\varepsilon f\}_{\varepsilon>0}$ is Cauchy in $L^2(\mathbb{R}^n)$. Thus, it only remains to show the claim. We have

$$\begin{aligned}T_\varepsilon g(x) - T_\delta g(x) &= \int_{|y|\geq\varepsilon} K(y)\, g(x-y)\ \mathrm{d}y - \int_{|y|\geq\delta} K(y)\, g(x-y)\ \mathrm{d}y \\ &= -\int_{\delta<|y|<\varepsilon} K(y)\, g(x-y)\ \mathrm{d}y \\ &= -\int_{\delta<|y|<\varepsilon} K(y)\,[g(x-y) - g(x)]\ \mathrm{d}y, \qquad (9)\end{aligned}$$

where we have used the cancellation property in the last line. Thus, we deduce

$$\begin{aligned}\left|T_{\varepsilon}g\left(x\right)-T_{\delta}g\left(x\right)\right| &\leq \int_{\delta<|y|<\varepsilon}\left|K\left(y\right)\right|\left|g\left(x-y\right)-g\left(x\right)\right|\,\mathrm{d}y\\ &\leq C\left\|\nabla g\right\|_{L^{\infty}(\mathbb{R}^{n})}\int_{\delta<|y|<\varepsilon}\frac{1}{\left|y\right|^{n}}\left|y\right|\,\mathrm{d}y\\ &\leq C\int_{\delta}^{\varepsilon}r^{n-1}\frac{1}{r^{n-1}}\,\mathrm{d}r\\ &\leq C\int_{0}^{\varepsilon}\,\mathrm{d}r\leq C\varepsilon\to 0 \qquad \text{as } \varepsilon\to 0.\end{aligned}$$

Thus, we have

$$\left\|T_{\varepsilon}g-T_{\delta}g\right\|_{L^{\infty}(\mathbb{R}^{n})}\to 0 \qquad \text{as } \varepsilon,\delta\to 0.$$

But now observe that the expression in (9) makes it clear that $T_{\varepsilon}g-T_{\delta}g$ is compactly supported, as g has compact support and y varies in the spherical shell $\delta<|y|<\varepsilon$.[1] Since the support is compact and thus have finite measure, we have

$$\left\|T_{\varepsilon}g-T_{\delta}g\right\|_{L^{2}(\mathbb{R}^{n})}\leq\left\|T_{\varepsilon}g-T_{\delta}g\right\|_{L^{\infty}(\mathbb{R}^{n})}\left|\operatorname{supp}\left(T_{\varepsilon}g-T_{\delta}g\right)\right|\to 0.$$

This completes the proof of the fact that $\left\{T_{\varepsilon}f\right\}_{\varepsilon>0}$ is Cauchy in $L^{2}\left(\mathbb{R}^{n}\right)$. Thus, $\left\{T_{\varepsilon}f\right\}_{\varepsilon>0}$ is convergent in $L^{2}\left(\mathbb{R}^{n}\right)$ and converges to some $h\in L^{2}\left(\mathbb{R}^{n}\right)$. We now set

$$Tf:=h \qquad \text{for all } f\in L^{2}\left(\mathbb{R}^{n}\right),$$

where h is the unique limit

$$T_{\varepsilon}f\to h \qquad \text{strongly in } L^{2}\left(\mathbb{R}^{n}\right).$$

Thus, we immediately deduce the estimate

$$\left\|Tf\right\|_{L^{2}(\mathbb{R}^{n})}=\lim_{\varepsilon\to 0}\left\|T_{\varepsilon}f\right\|_{L^{2}(\mathbb{R}^{n})}\leq A_{2}\left\|f\right\|_{L^{2}(\mathbb{R}^{n})}.$$

This completes the proof. □

2 Real analysis tools

2.1 Covering lemmas

We first start with a very simple, but still immensely useful result.

[1] Note that neither $T_{\varepsilon}g$ nor $T_{\delta}g$ have compact support, but only their difference must be compactly supported.

Lemma 14 (Vitali covering lemma (simplified version))**.** *Suppose we have a finite family of balls* $\{B_{r_i}(x_i)\}_i$*, then there exists a sub family of disjoint balls* $\{B_{r_k}(x_k)\}_k$ *such that*

$$\bigcup_i B_{r_i}(x_i) \subset \bigcup_k B_{3r_k}(x_k)$$

Proof. First we arrange the balls in the descending order of their radii, i.e. the largest ball (or one of the largest) as the first ball. We then add the first ball to our subcollection. Now, if the second ball is disjoint from the first ball, we add it to the subcollection. Then we select the ball with the largest radius (or one of them if there are are more than one with the same largest radius) which does not intersect the first ball as our second ball. Then we pick the ball the ball with the largest radius (or one of them if there are are more than one with the same largest radius) which does not intersect the either of the two balls we have chosen as our third ball. We continue in this fashion. The process would stop after finitely many steps, as we started with finitely many balls. Then all the remaining balls intersect at least one of the balls in our collection. Now note that if any two balls intersect and one of them has radius less than or equal to the other, then the ball with the smaller or equal radius would be completely contained inside the ball of radius three times that of the larger (or equal) radius. This **engulfing property** ensures the result. □

2.2 Distribution function and weak L^p

We first introduce a tool to study the behavior of L^p functions. Roughly, if a function is in L^p, then although the values of the function can be large, but the measure of the set where this happens has to be correspondingly small enough. Equivalently, the measure of its *super level sets* must decay in a certain manner as the level rises. This is best expressed by the follong function.

Definition 15 (Distribution function)**.** *Let* $(\Omega, \mathcal{F}, \mu)$ *be measure space and* $f : \Omega \to [0, \infty]$ *be a nonnegative measurable function. Given a nonnegative real number* $t \geq 0$*, we define the distribution function of* f*,* $\alpha_f : [0, \infty) \to [0, \infty]$ *by*

$$\alpha_f(t) := \mu\left(\{x \in \Omega : |f(x)| > t\}\right)$$

Remark 16. *When* μ *is the Lebesgue measure in* $\mathbb{R}^n$*, we would just write*

$$\alpha_f(t) := \left|\{x \in \Omega : |f(x)| > t\}\right|.$$

We now state a formula commonly known as the **Layer Cake formula**. In the same setting as above, we have,

Proposition 17. *For all* $1 \leq p < \infty$*,*

$$\int_\Omega (f(x))^p \, d\mu = p \int_0^\infty t^{p-1} \alpha_f(t) dt$$

Proof. Rewriting the LHS, we have

$$\begin{aligned}\int_\Omega (f(x))^p \, d\mu &= \int_\Omega p \int_0^{f(x)} t^{p-1} dt \, d\mu \\ &= \int_\Omega \int_0^\infty p t^{p-1} \chi(x)_{\{f(x)>t\}} dt \, d\mu \\ &= p \int_0^\infty t^{p-1} \alpha_f(t) dt\end{aligned}$$

This proves the result. □

Proposition 18. *Let $\Phi : [0, \infty] \to [0, \infty]$ be a C^1, nondecreasing function such that $\Phi(0) = 0$. Then*

$$\int_{\mathbb{R}^n} \Phi(|f(x)|) \, \mathrm{d}x = \int_0^\infty \Phi'(t) \, \alpha_f(t) \, \mathrm{d}t.$$

Proof is left as an exercise.

Theorem 19 (Chebyshev's inequality)**.** *Let $f \in L^p(\Omega)$ for some $1 \le p < \infty$ and $\Omega \subset \mathbb{R}^n$ measurable. Then we have*

$$\alpha_f(t) := |\{x \in \Omega : |f(x)| > t\}| \le \frac{1}{t^p} \|f\|^p_{L^p(\Omega)}.$$

Proof. We have that

$$s \mathbb{1}_{\{g(x) \ge s\}} \le g(x)$$

Integrating this on Ω, we deduce

$$s |\{x \in \Omega : |f(x)| > s\}| \le \int_\Omega g(x) \mathrm{d}x.$$

With $s = t^p$ and $g = |f|^p$, and noting that $\{x \in \Omega : |f(x)|^p > t\} = \{x \in \Omega : |f(x)|^p > t^p\}$ we obtain the theorem. □

A simple consequence of Chebyshev's inequality is that if $f \in L^p(\Omega)$, then

$$\sup_{t>0} t^p |\{x \in \Omega : |f| > t\}| \le \|f\|^p_{L^p(\Omega)} < \infty.$$

The answer to the natural converse question is false as seen by the function $f(x) = 1/x$ on the interval $[0, 1]$ with $p = 1$. However, the polynomial decay of the distribution function is important enough to merit a definition.

Definition 20. *(Weak L^p or Marcinkiewicz space) For $1 \le p < \infty$, we define*

$$L^p_w(\Omega) := \left\{ f : \Omega \to \mathbb{R}, f \text{ measurable } : \sup_{t>0} t^p \alpha_f(t) < \infty \right\}$$

In general,

$$L^p(\Omega) \subsetneq L^p_w(\Omega),$$

as the following example shows

$$f(x) := \frac{1}{|x|^{\frac{n}{p}}}.$$

Since $1/|x|^n$ is not locally integrable around the origin (check via polar coordinates) in $\mathbb{R}^n$, clearly

$$f \in L^p_w(B^n_1(0)), \quad \text{but} \quad f \notin L^p(B^n_1(0)).$$

Remark 21. *The space L^p_w is the Lorentz space $L^{(p,\infty)}$ and is often denoted this way. The expression*

$$\|f\|_{L^{(p,\infty)}} := \sup_{t>0} t^p \alpha_f(t)$$

does not define a norm, as the triangle inequality fails in general. However, as we have

$$\{x \in \Omega : |f(x) + g(x)| > t\} \subset \{x \in \Omega : |f(x)| > t/2\} \cup \{x \in \Omega : |g(x)| > t/2\},$$

it is easy to that

$$\|f + g\|_{L^{(p,\infty)}} \leq 2\left(\|f\|_{L^{(p,\infty)}} + \|g\|_{L^{(p,\infty)}}\right).$$

Thus, it is a ***quasinorm****, not a norm. $L^{(p,\infty)}$ is a quasi-Banach space under this quasinorm. When $p \neq 1$, the quasinorm however is equivalent to a norm, but this is not so for $p = 1$. A lot of harmonic analysis (if not most, or all) would be trivial if weak L^1 would have been a normed space.*

2.3 Maximal functions

Definition 22 (Hardy-Littlewood Maximal function)**.** *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Define,*

$$Mf(x) := \sup_{\substack{Q_r(x)\\ r>0}} \frac{1}{|Q_r(x)|} \int_{Q_r(x)} |f(y)| dy$$

where $Q_r(x)$ is a cube of side length r centered at x with sides parallel to the axes.

This is the **centered maximal function**. One can also define the **uncentered** one by only requiring $x \in Q$, not necessarily the center of Q. We can also replace cubes with balls of radius r centered around x in the definition (or the uncentered ball version by using balls containing x). For all these versions, their general behavior, for our purposes, would not differ much at all. By the Lebesgue Differentiation theorem, we have that $Mf \geq |f|$. It is not difficult to show that Mf is **never** in $L^1(\mathbb{R}^n)$ unless if $f \equiv 0$.

Definition 23. *The map*

$$f \mapsto Mf$$

is called the ***maximal operator****.*

Theorem 24 (Hardy-Littlewood-Wiener maximal theorem)**.** *Let $f \in L^1_{loc}(\mathbb{R}^n)$.*

(i) If $f \in L^p(\mathbb{R}^n)$ for some $1 \leq p \leq \infty$, then Mf is finite for a.e. $x \in \mathbb{R}^n$.

(ii) If $f \in L^\infty(\mathbb{R}^n)$, then $Mf \in L^\infty(\mathbb{R}^n)$ and we have

$$\|Mf\|_{L^\infty(\mathbb{R}^n)} \leq \|f\|_{L^\infty(\mathbb{R}^n)}.$$

(iii) If $f \in L^1(\mathbb{R}^n)$, then $Mf \in L^1_w(\mathbb{R}^n)$ and there exists a constant $A > 0$, depending only on the dimension n, such that

$$\sup_{t>0} t\ |\{x \in \mathbb{R}^n : |Mf(x)| > t\}| \leq A\|f\|_{L^1(\mathbb{R}^n)}.$$

(iv) If $f \in L^p(\mathbb{R}^n)$ for $1 < p < \infty$, then $Mf \in L^p(\mathbb{R}^n)$ and there exists a constant $A_p > 0$, depending only on the dimension n and p, such that

$$\|Mf\|_{L^p(\mathbb{R}^n)} \leq A_p\|f\|_{L^p(\mathbb{R}^n)}.$$

Proof. (ii) is completely obvious. We prove (iii) and (iv) and leave (i) as an exercise. We first prove (iii). Let $t > 0$ and let K be a compact set such that $K \subset \{|Mf| > t\}$. Then for any $x \in K$, there exists $r(x) > 0$ such that,

$$\int_{Q_{r(x)}(x)} |f(y)|dy > t\,(r(x))^n.$$

Now the collection of cubes $\cup_{x\in K} Q_{r(x)}(x)$ defines an open cover of K. By compactness of K, there exists a finite subcover $\left\{Q_{r_j}(x_j)\right\}$. Using Vitali's Covering lemma, we may obtain a finite disjoint subfamily $\{Q_{r_i}(x_i)\}_{i=1}^m$ such that

$$K \subset \bigcup_{i=1}^m Q_{3r_i}(x_i)$$

Thus, we deduce

$$\begin{aligned} |K| &\leq \sum_{i=1}^m (3r_i(x_i))^n \\ &\leq 3^n \sum_{i=1}^m (r_i(x_i))^n \leq \frac{3^n}{t} \sum_{i=1}^m \int_{Q_{r_i(x_i)}(x_i)} |f(y)|dy \leq \frac{3^n}{t} \int_{\mathbb{R}^n} |f|. \end{aligned}$$

Since this is true for every compact set sitting inside $\{|Mf| > t\}$, by inner regularity of Lebesgue measure, we have

$$|\{|Mf| > t\}| \leq \frac{3^n}{t} \|f\|_{L^1(\mathbb{R}^n)}.$$

This proves (iii). For (iv), we define $f_1 = f.\mathbb{1}_{\{|f(x)|>t/2\}}$. Then clearly, $|f(x)| \leq |f_1(x)| + t/2$ and consequently $Mf \leq Mf_1 + t/2$. Thus, we have

$$|\{|Mf| > t\}| \leq \left|\left\{|Mf_1| > \frac{t}{2}\right\}\right| \leq \frac{2.3^n}{t} \int_{\mathbb{R}^n} |f_1| \leq \frac{2.3^n}{t} \int_{\{|f(x)|>t/2\}} |f|$$

Now, using this in the layer cake formula and employing Fubini, we deduce

$$\begin{aligned}
\int_{\mathbb{R}^n} (Mf)^p &= p \int_0^\infty t^{p-1} |\{x \in \mathbb{R}^n : Mf > t\}| \, \mathrm{d}t \\
&\leq 2.3^n p \int_0^\infty t^{p-1} \frac{1}{t} \int_{\{|f(x)|>t/2\}} |f| \, \mathrm{d}x \, \mathrm{d}t \\
&\leq 2.3^n p \int_{\mathbb{R}^n} |f(x)| \int_0^{2|f(x)|} t^{p-2} \, \mathrm{d}t \, \mathrm{d}x \\
&\leq \frac{3^n 2^p p}{p-1} \|f\|_{L^p}^p.
\end{aligned}$$

This completes the proof. □

Remark 25. *Note that that the constant blows up as $p \to 1$.*

As a consequence, we can prove the Lebesgue differentiation theorem.

Corollary 26 (Lebesgue differentiation theorem)**.** *If $f \in L^1_{loc}(\mathbb{R}^n)$, then*

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y) dy = f(x) \qquad \textit{for a.e. } x \in \mathbb{R}^n.$$

Proof. We may assume $f \in L^1(\mathbb{R}^n)$. Define

$$A_r f(x) = \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y) \mathrm{d}y.$$

Suppose $g \in C_c(\mathbb{R}^n)$, we have that

$$\lim_{r\to 0} A_r g(x) = g(x) \qquad \text{for all } x \in \mathbb{R}^n.$$

We now have

$$A_r f - f = A_r (f - g) + A_r g - g + g - f.$$

Noting that $\limsup_{r\to 0} |A_r(f-g)| \leq M(f-g)$ we have,

$$\limsup_{r\to 0} |A_r f - f| \leq |M(f-g)| + |f-g|.$$

So we have for any $\varepsilon > 0$

$$\left|\left\{\limsup_{r\to 0} |A_r f - f| > \varepsilon\right\}\right| \leq \left|\left\{|M(f-g)| > \frac{\varepsilon}{2}\right\}\right| + \left|\left\{|f-g| > \frac{\varepsilon}{2}\right\}\right|$$

Now, we have by the weak $(1,1)$ estimate for the maximal function,

$$\left|\left\{|M(f-g)| > \frac{\varepsilon}{2}\right\}\right| \leq \frac{2C}{\varepsilon}\|f-g\|_{L^1(\mathbb{R}^n)}.$$

By Chebyshev's inequality, we have

$$\left|\left\{|f-g| > \frac{\varepsilon}{2}\right\}\right| \leq \frac{2}{\varepsilon}\|f-g\|_{L^1(\mathbb{R}^n)}.$$

Thus, by the density of $C_c(\mathbb{R}^n)$ in $L^1(\mathbb{R}^n)$, the RHS can be made arbitrarily small. This completes the proof. □

Remark 27. *One can actually prove the stronger statement*

$$\lim_{r\to 0} \frac{1}{|B_r(x)|}\int_{B_r(x)} |f(y) - f(x)|\, dy = 0 \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

We leave it as an exercise.

2.4 Marcinkiewicz interpolation theorem

We now want to prove an interpolation theorem. Before this, we need a few notions.

Definition 28. *Let $\Omega \subset \mathbb{R}^n$ be open. For any $1 \leq p < q \leq \infty$, the space $L^p(\Omega) + L^q(\Omega)$ is defined as the set of measurable functions $f : \Omega \to \mathbb{R}$ such that there exists $f_1 \in L^p(\Omega)$ and $f_2 \in L^q(\Omega)$ and we can write*

$$f = f_1 + f_2.$$

Remark 29. *Note that such a decomposition of f is far from unique.*

Obviously, by taking $f_1 = 0$ or $f_2 = 0$, it is easy to see that $L^p(\Omega), L^q(\Omega) \subset L^p(\Omega) + L^q(\Omega)$. But even more is true.

Proposition 30. *For every $p \leq r \leq q$, we have*

$$L^r(\Omega) \subset L^p(\Omega) + L^q(\Omega).$$

Proof. For any $\gamma > 0$, we write

$$f = f\mathbb{1}_{\{|f|>\gamma\}} + f\mathbb{1}_{\{|f|\leq\gamma\}} := f_1 + f_2.$$

Clearly,

$$\int_\Omega |f_1|^p \leq \gamma^{r-p}\int_\Omega |f|^r.$$

On the other hand, clearly $f_2 \in L^\infty(\Omega)$ by construction and if $q \neq \infty$, we have

$$\int_\Omega |f_2|^q \leq \gamma^{q-r} \int_\Omega |f|^r .$$

□

Definition 31. *Let $\Omega_1, \Omega_2 \subset \mathbb{R}^n$ be two open subsets. Let $\mathcal{M}(\Omega_2)$ denote the space of measurable functions over Ω_2. We say that a map $T : L^p(\Omega_1) + L^q(\Omega_1) \to \mathcal{M}(\Omega_2)$ is **Q-subadditive** if there exists a constant $Q > 0$ such that*

$$|T(f+g)| \leq Q(|Tf| + |Tg|) \qquad \text{for all } f, g \in L^p(\Omega_1) + L^q(\Omega_1).$$

Remark 32. *Note that T need not be linear, even if T is 1-subadditive. Every linear map is of course 1-subadditive, but the maximal operator is 1-subadditive, but not linear.*

Definition 33. *Let $T : L^p(\Omega_1) + L^q(\Omega_1) \to \mathcal{M}(\Omega_2)$ be a **Q-subadditive** map. T is said to be of **weak type** (p,p) if T maps $L^p(\Omega_1)$ into $L^p_w(\Omega_2)$ and there exists a constant C such that,*

$$\sup_{t>0} t^p \left|\{x \in \Omega_2 : |Tf| > t\}\right| \leq C\|f\|^p_{L^p(\Omega_1)} \qquad \text{for all } f \in L^p(\Omega_1).$$

*We say T is of **strong type** (p,p) if T maps $L^p(\Omega_1)$ into $L^p(\Omega_2)$ and there exists a constant C such that*

$$\|Tf\|_{L^p(\Omega_2)} \leq C\|f\|_{L^p(\Omega_1)} \qquad \text{for all } f \in L^p(\Omega_1).$$

We define weak type (∞,∞) to be the same as strong type (∞,∞).

Now we are in aposition to state the interpolation theorem.

Theorem 34 (Marcinkiewicz's interpolation theorem)**.** *Let $\Omega_1, \Omega_2 \subset \mathbb{R}^n$ be two open subsets. Let $1 \leq p < q \leq \infty$. Let $T : L^p(\Omega_1) + L^q(\Omega_1) \to \mathcal{M}(\Omega_2)$ be a Q-subadditive map which is of weak type (p,p) and weak type (q,q). Then T is of strong type (r,r) for every $p < r < q$.*

Proof. First, we prove for $1 \leq p < q < \infty$. Let $f \in L^r(\Omega_1)$. For a $s > 0$, let

$$f_1 = f\mathbb{1}_{\{|f|>s\}} \qquad f_2 = f\mathbb{1}_{\{|f|\leq s\}}$$

We have

$$f = f_1 + f_2$$

The idea will be to let this splitting of f vary by letting s vary. We have that

$$|Tf| \leq Q(|Tf_1| + |Tf_2|)$$

Let A_p and A_q be the constants of the weak type (p,p) and weak type (q,q) inequalities respectively.

We now have that

$$\{|Tf| > t\} \subset \left\{|Tf_1| > \frac{t}{2Q}\right\} \cup \left\{|Tf_2| > \frac{t}{2Q}\right\}$$

So we have

$$\begin{aligned}\alpha_{Tf}(t) &\le \alpha_{Tf_1}\left(\frac{t}{2Q}\right) + \alpha_{Tf_2}\left(\frac{t}{2Q}\right)\\ &\le \frac{A_p}{(t/2Q)^p}\int_{\Omega_1}|f_1|^p + \frac{A_q}{(t/2Q)^q}\int_{\Omega_1}|f_2|^q.\end{aligned}$$

Now we have,

$$\begin{aligned}\int_{\Omega_2}|Tf(x)|^r dx &= r\int_0^\infty t^{r-1}\alpha_{Tf}(t)dt\\ &\le r\int_0^\infty t^{r-1}\left\{\frac{A_p}{(t/2Q)^p}\int_{\Omega_1}|f_1|^p + \frac{A_q}{(t/2Q)^q}\int_{\Omega_1}|f_2|^q\right\}dt\\ &= A_p 2^p Q^p r\int_0^\infty \left(\int_{|f|>s}|f_1|^p\right)t^{r-1-p}dt\\ &\qquad + A_q 2^q Q^q r\int_0^\infty\left(\int_{|f|\le s}|f_2|^q\right)t^{r-1-q}dt\end{aligned}$$

Now, the choice of s was arbitrary. In particular, we may let it vary. Setting $s = t$, we get

$$\begin{aligned}&\int_{\Omega_2}|Tf(x)|^r dx\\ &\le A_p 2^p Q^p r\int_0^\infty\left(\int_{|f|>t}|f_1|^p\right)t^{r-1-p}dt\\ &\qquad + A_q 2^q Q^q r\int_0^\infty\left(\int_{|f|\le t}|f_2|^q\right)t^{r-1-q}dt\\ &\le A_p 2^p Q^p r\int_0^\infty\left(\int_{|f|>t}|f|^p\right)t^{r-1-p}dt\\ &\qquad + A_q 2^q Q^q r\int_0^\infty\left(\int_{|f|\le t}|f|^q\right)t^{r-1-q}dt\\ &= A_p 2^p Q^p r\int_{\Omega_1}|f|^p dx\int_0^{|f|}t^{r-1-p}dt\\ &\qquad + A_q 2^q Q^q r\int_{\Omega_1}|f|dx\int_{|f|}^\infty t^{r-1-q}dt\\ &= \left\{A_p 2^p Q^p r\frac{1}{r-p} + A_q 2^q Q^q r\frac{1}{q-r}\right\}\int_{\Omega_1}|f|^r.\end{aligned}$$

This proves the desired inequality. When $q = \infty$, Take any $f \in L^r(\Omega_1)$ for $r \in (p, \infty)$. As before define f_1 and f_2

$$f_1 = f\mathbb{1}_{\{|f|>s\}} \qquad f_2 = f\mathbb{1}_{\{|f|\leq s\}}.$$

We have that $f_2 \in L^\infty(\Omega_1)$ and $f_1 \in L^p(\Omega_1)$. Let A_p and A_∞ be the constants of the weak (p,p) and weak (∞,∞) inequalities respectively. So we have,

$$\begin{aligned} |Tf| &\leq Q(|Tf_1| + |Tf_2| \\ &\leq Q|Tf_1| + QA_\infty \|f_2\|_{L^\infty(\Omega_1)} \\ &\leq Q|Tf_1| + QA_\infty s \end{aligned}$$

Now we pick s such that $QA\infty s = \frac{t}{2}$. So we have

$$|Tf| \leq Q|Tf_1| + \frac{t}{2}$$

This gives us,

$$\{|Tf| > t\} \subset \left\{|Tf_1| > \frac{t}{2Q}\right\}.$$

Thus, we deduce

$$\begin{aligned} |\{x \in \Omega_2 : |Tf| > t\}) &\leq \left|\left\{x \in \Omega_2 : |Tf_1| > \frac{t}{2Q}\right\}\right| \\ &\leq \frac{A_p}{(t/2Q)^p} \|f_1\|_{L^p(\Omega_1)} \end{aligned}$$

So we have,

$$\begin{aligned} \int_{\Omega_2} |Tf(x)|^r dx &= r\int_0^\infty t^{r-1}\alpha_{Tf}(t)dt \\ &\leq r\int_0^\infty t^{r-1}\frac{A_p 2^p Q^p}{t^p}\int_{|f|>\frac{t}{2QA_\infty}} |f|^p dtdx \\ &\leq A_p 2^p Q^p r\int_{\Omega_1} |f|^p dx \int_0^{2QA_\infty|f|} t^{r-1-p}dt \\ &\leq A_p 2^p Q^p r\frac{(2QA_\infty)^{r-p}}{r-p}\int_{\Omega_1} |f|^r \end{aligned}$$

which proves the required inequality. □

Remark 35. *Note that the Marcinkiewicz interpolation theorem can be used to prove part* (*iv*) *of the Hardy-Littlewood-Wiener maximal theorem from part* (*ii*) *and* (*iii*). *Indeed, the maximal operator is* 1*-subadditive and by* (*iii*), *is of weak type* $(1,1)$ *and of strong type* (∞,∞) *by* (*ii*). *Hence, maximal operator is of strong* (p,p) *for every* $1 < p < \infty$, *by the Marcinkiewicz interpolation theorem.*

2.5 Calderon-Zygmund decomposition

We have already seen one way of splitting a function f. Now we want to split a functio in two parts, with more refined control over the pieces. A basic tool for this is the following decomposition, known as the **Calderon-Zygmund decomposition**. This simple device is an extremely robust, flexible and potent tool.

Theorem 36 (CZ decomposition in a cube)**.** *Let $Q \subset \mathbb{R}^n$ be an open cube and let $f \in L^1(Q)$. Let $\alpha > 0$ be a real number such that*

$$\frac{1}{|Q|}\int_Q |f| \le \alpha.$$

Then there exists a countable family of open subcubes $\{Q_i\}_{i=1}^{\infty}$, with sides parallel to the original cube Q and with pairwise mutually disjoint interiors such that

(a) For every i, we have

$$\alpha < \frac{1}{|Q_i|}\int_{Q_i} |f| \le 2^n \alpha,$$

(b) and we have

$$|f| \le \alpha \qquad \text{for a.e. } x \text{ on } Q \setminus \bigcup_{i=1}^{\infty} Q_i.$$

Proof. If $|f| \le \alpha$ *a.e.* on Q, then we are done. If not, bisect each side of Q to obtain 2^n congruent subcubes. In each of those subcubes Q', exactly one of the two possibilities can occur.

- **Case 1:**

$$⨍_{Q'} |f| \le \alpha.$$

- **Case 2:**

$$⨍_{Q'} |f| > \alpha.$$

Add those subcubes where the second case occurs to our subcollection $\{Q_i\}$. Where Case 1 occurs, we again bisect the sides of those subcubes and continue the process. Clearly, the process can go on only countably many times and we end up with a countable collection of subcubes $\{Q_i\}_{i\in\mathbb{N}}$. Now, for any of these subcubes, note that their immediate 'parent' cube (denoted by $\tilde{Q}_i$) was not selected, otherwise we would not even bisect the sides of $\tilde{Q}_i$. Hence we have

$$\alpha < \frac{1}{|Q_i|}\int_{Q_i} |f| \le \frac{1}{|Q_i|}\int_{\tilde{Q}_i} |f| \le \frac{2^n}{|\tilde{Q}_i|}\int_{\tilde{Q}_i} |f| \le 2^n \alpha.$$

Now for any point in the complement of this collection is contained in a sequence of cubes where Case 1 occured. Thus, we have a sequence of cubes of shrinking side length $\{C_i(x)\}_{i\in\mathbb{N}}$ such that $x \in C_i(x)$ for each $i \in \mathbb{N}$. Thus, by the Lebesgue Differentiation theorem (uncentered version), we deduce,

$$f(x) = \lim_{\operatorname{diam} C_i(x)\to 0} \frac{1}{|C_i(x)|} \int_{C_i(x)} |f| \leq \alpha.$$

This completes the proof. □

As a simple corollary, we have the Calderon-Zygmund decomposition on all of $\mathbb{R}^n$ for any $\alpha > 0$.

Theorem 37 (CZ decomposition in $\mathbb{R}^n$)**.** *Let $f \in L^1(\mathbb{R}^n)$. Let $\alpha > 0$ be a real number. Then there exists a countable family of open cubes $\{Q_i\}_{i=1}^{\infty}$, with sides parallel to the coordinate axes and with pairwise mutually disjoint interiors such that*

(a) For every i, we have

$$\alpha < \frac{1}{|Q_i|} \int_{Q_i} |f| \leq 2^n \alpha,$$

(b) and we have

$$|f| \leq \alpha \qquad \textit{for a.e. } x \textit{ on } Q \setminus \bigcup_{i=1}^{\infty} Q_i.$$

Proof. Divide $\mathbb{R}^n$ into countable number of congruent cubes with sides parallel to the coordinate axes with side length L. Since $f \in L^1(\mathbb{R}^n)$, choose $L > 0$ large enough such that

$$\alpha L^n \geq \|f\|_{L^1(\mathbb{R}^n)}.$$

Now we apply the CZ decomposition to each of these cubes. □

Remark 38. *We often say CZ decomposition **of** f **at level** α.*

As a consequence, we can split a function into two parts.

Theorem 39 (CZ decomposition of functions in $\mathbb{R}^n$)**.** *Let $f \in L^1(\mathbb{R}^n)$. Let $\alpha > 0$ be a real number. Then there exist a bounded function g and a countable family of L^1 functions $\{b_i\}_{i\in\mathbb{N}}$ and a countable collection of open cubes $\{Q_i\}_{i=1}^{\infty}$, with sides parallel to the coordinate axes and with pairwise mutually disjoint interiors such that*

(i) We have

$$f = g + b := g + \sum_{i=1}^{\infty} b_i.$$

(ii) $f = g$ for a.e. $x \in \mathcal{G}$, where $\mathcal{G} := \mathbb{R}^n \setminus \bigcup_{i=1}^{\infty} Q_i$.

(iii) We have the estimates

$$\|g\|_{L^{\infty}(\mathbb{R}^n)} \leq 2^n \alpha \qquad \textit{and} \qquad \|g\|_{L^1(\mathbb{R}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)} .$$

(iv) For every i, we have $b_i \equiv 0$ outside $\overline{Q_i}$ and we have

$$⨍_{Q_i} b_i (y) \, \mathrm{d}y = 0.$$

(v) We have

$$\sum_{i=1}^{\infty} \|b_i\|_{L^1(\mathbb{R}^n)} \leq 2 \|f\|_{L^1(\mathbb{R}^n)} .$$

(vi) For the set $\mathcal{F} := \bigcup\limits_{i=1}^{\infty} Q_i$, we have

$$|\mathcal{F}| \leq \frac{1}{\alpha} \|f\|_{L^1(\mathbb{R}^n)} .$$

Remark 40. *The notation g and b stands for the 'good part' and 'bad part' of the function f respectively and the notation $\mathcal{G}$ denotes the 'good set' and $\mathcal{F}$ is called the 'bad set'.*

Proof. Apply the Calderon-Zygmund decomposition to f at level $\alpha > 0$ to obtain a countable collection of cubes $\{Q_i\}_{i=1}^{\infty}$ and define

$$g(x) := \begin{cases} f(x) & \text{in } \mathbb{R}^n \setminus \bigcup\limits_{i=1}^{\infty} Q_i, \\ ⨍_{Q_i} f(y) \, \mathrm{d}y & \text{in } Q_i. \end{cases}$$

Set $b = f - g$ and $b_i = b \mathbb{1}_{Q_i}$ for each $i \in \mathbb{N}$. Verification of the claimed properties is left as an easy exercise. □

3 Singular integrals

3.1 Weak $(1,1)$ estimate

We are now ready to prove our main estimate, which is the key step towards the Calderon-Zygmund theorem we are planning to prove.

Theorem 41. *Let K be CZ kernel satisfying the Hörmander condition. Let $f \in L^1(\mathbb{R}^n)$. For any $\varepsilon > 0$, define the operators*

$$T_{\varepsilon} f(x) = \int_{\mathbb{R}^n \setminus B_{\varepsilon}(x)} K(x-y) f(y) \, \mathrm{d}y \qquad \textit{for a.e. } x \in \mathbb{R}^n.$$

Then for every $\varepsilon > 0$, $T_\varepsilon f \in L^1_w(\mathbb{R}^n)$ and there exists a constant $A_1 > 0$, independent of f and $\varepsilon > 0$, such that we have the estimates

$$|\{x \in \mathbb{R}^n : |T_\varepsilon f(x)| > t\}| \leq \frac{A_1}{t} \|f\|_{L^1(\mathbb{R}^n)} \qquad \text{for all } t > 0.$$

Proof. Fix $t > 0$. Apply the CZ decomposition to f at level t to obtain a bounded function g and a countable family of L^1 functions $\{b_i\}_{i\in\mathbb{N}}$ and a countable collection of open cubes $\{Q_i\}_{i=1}^{\infty}$, with sides parallel to the coordinate axes and with pairwise mutually disjoint interiors, as in Theorem 39. Thus, we have

$$T_\varepsilon f = T_\varepsilon g + T_\varepsilon b := T_\varepsilon g + \sum_{i=1}^{\infty} T_\varepsilon b_i.$$

Thus, we deduce

$$\{x : |T_\varepsilon f(x)| > t\} \subset \{x : |T_\varepsilon g(x)| > t/2\} \cup \{x : |T_\varepsilon b(x)| > t/2\}.$$

This implies

$$|\{x \in \mathbb{R}^n : |T_\varepsilon f(x)| > t\}| \leq |\{x \in \mathbb{R}^n : |T_\varepsilon g(x)| > t/2\}| + |\{x \in \mathbb{R}^n : |T_\varepsilon b(x)| > t/2\}|. \quad (10)$$

Note that $g \in L^2(\mathbb{R}^n)$ and we have the estimate

$$\|g\|^2_{L^2(\mathbb{R}^n)} \leq \|g\|_{L^\infty(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)} \leq 2^n t \|g\|_{L^1(\mathbb{R}^n)} \leq 2^n t \|f\|_{L^1(\mathbb{R}^n)}.$$

Combining this with Chebyshev's inequality and Theorem 13, we have

$$|\{x \in \mathbb{R}^n : |T_\varepsilon g(x)| > t/2\}| \leq \frac{4}{t^2} \|T_\varepsilon g\|^2_{L^2(\mathbb{R}^n)} \leq \frac{4A_2}{t^2} \|g\|^2_{L^2(\mathbb{R}^n)} \leq \frac{2^{n+2}A_2}{t} \|f\|_{L^1(\mathbb{R}^n)}. \quad (11)$$

Now for each $i \in \mathbb{N}$, let Q_i^* denote that cube with the same center and parallel sides as Q_i, but with side length $2\sqrt{n} l_i$, where l_i is the side length of the cube Q_i. We set

$$\mathcal{F}^* := \bigcup_{i=1}^{\infty} Q_i^* \qquad \text{and} \qquad \mathcal{G}^* = \mathbb{R}^n \setminus \mathcal{F}^*.$$

Clearly, we have

$$\{x \in \mathbb{R}^n : |T_\varepsilon b(x)| > t/2\} \subset \{x \in \mathcal{G}^* : |T_\varepsilon b(x)| > t/2\} \cup \{x \in \mathcal{F}^* : |T_\varepsilon b(x)| > t/2\}.$$

Thus, we deduce

$$|\{x \in \mathbb{R}^n : |T_\varepsilon b(x)| > t/2\}| \leq |\{x \in \mathcal{G}^* : |T_\varepsilon b(x)| > t/2\}| + |\{x \in \mathcal{F}^* : |T_\varepsilon b(x)| > t/2\}|. \quad (12)$$

But we have

$$\begin{aligned}\left|\{x\in\mathcal{F}^*:\left|T_\varepsilon b\,(x)\right|>t/2\}\right| &\leq \left|\mathcal{F}^*\right| \\ &\leq \sum_{i=1}^{\infty}\left|Q_i^*\right| \\ &= \left(2\sqrt{n}\right)^n\sum_{i=1}^{\infty}\left|Q_i\right| \\ &= \left(2\sqrt{n}\right)^n\left|\mathcal{F}\right| \leq \frac{\left(2\sqrt{n}\right)^n}{t}\left\|f\right\|_{L^1(\mathbb{R}^n)}. \end{aligned} \tag{13}$$

Now we estimate $\left|\{x\in\mathcal{G}^*:\left|T_\varepsilon b\,(x)\right|>t/2\}\right|$. Note that since $⨍_{Q_i} b_i = 0$ for each $i\in\mathbb{N}$, we can write

$$T_\varepsilon b_i\,(x) = \int_{Q_i} K_\varepsilon\,(x-y)\,b_i\,(y)\;\mathrm{d}y = \int_{Q_i}\left[K_\varepsilon\,(x-y)-K_\varepsilon\,(x-y_i)\right]b_i\,(y)\;\mathrm{d}y,$$

where y_i denotes the center of the cube Q_i. Hence we have

$$T_\varepsilon b\,(x) = \sum_{i=1}^{\infty}\int_{Q_i}\left[K_\varepsilon\,(x-y)-K_\varepsilon\,(x-y_i)\right]b_i\,(y)\;\mathrm{d}y.$$

Using this and Fubini, we deduce

$$\begin{aligned}\int_{\mathcal{G}^*}\left|T_\varepsilon b\,(x)\right|\;\mathrm{d}x &\leq \sum_{i=1}^{\infty}\int_{x\notin Q_i^*}\left(\int_{Q_i}\left|K_\varepsilon\,(x-y)-K_\varepsilon\,(x-y_i)\right|\left|b_i\,(y)\right|\;\mathrm{d}y\right)\;\mathrm{d}x \\ &= \sum_{i=1}^{\infty}\int_{Q_i}\left|b_i\,(y)\right|\left(\int_{x\notin Q_i^*}\left|K_\varepsilon\,(x-y)-K_\varepsilon\,(x-y_i)\right|\;\mathrm{d}x\right)\;\mathrm{d}y,\end{aligned} \tag{14}$$

as long as we can show that the integral on the right is finite. Now observe that since y_i is the center of the cube Q_i with side length l_i, we have

$$\left|y-y_i\right| \leq \frac{\sqrt{n}}{2}l_i.$$

On the other hand, since Q_i^* has side length $2\sqrt{n}l_i$ and center y_i, for any $x\notin Q_i^*$, we have

$$\left|x-y_i\right| > \frac{2\sqrt{n}}{2}l_i \geq 2\left|y-y_i\right|.$$

Thus, setting $x' = x-y_i$ and $y' = y-y_i$ and changing variables, the integral inside the parentheses can be written as

$$\int_{x\notin Q_i^*}\left|K_\varepsilon\,(x-y)-K_\varepsilon\,(x-y_i)\right|\;\mathrm{d}x = \int_{|x'|>2|y'|}\left|K_\varepsilon\,(x'-y')-K_\varepsilon\,(x')\right|\;\mathrm{d}x'.$$

Since this is bounded uniformly w.r.t y' by the Hörmander condition, we finally arrive at

$$\int_{\mathcal{G}^*} |T_\varepsilon b(x)| \, \mathrm{d}x \leq \sum_{i=1}^{\infty} \int_{Q_i} |b_i(y)| \left(\int_{x \notin Q_i^*} |K_\varepsilon(x-y) - K_\varepsilon(x-y_i)| \, \mathrm{d}x \right) \mathrm{d}y$$
$$\leq C \sum_{i=1}^{\infty} \int_{Q_i} |b_i(y)| \, \mathrm{d}y = C \sum_{i=1}^{\infty} \|b_i\|_{L^1(\mathbb{R}^n)} \leq 2C \|f\|_{L^1(\mathbb{R}^n)}.$$

Hence, by Chebyshev's inequality, we deduce

$$|\{x \in \mathcal{G}^* : |T_\varepsilon b(x)| > t/2\}| \leq \frac{2}{t} \int_{\mathcal{G}^*} |T_\varepsilon b(x)| \, \mathrm{d}x \leq \frac{4C}{t} \|f\|_{L^1(\mathbb{R}^n)}. \tag{15}$$

Finally, in view of (10) and (12), combining (11), (13) and (15), we have

$$|\{x \in \mathbb{R}^n : |T_\varepsilon f(x)| > t\}| \leq \frac{A_1}{t} \|f\|_{L^1(\mathbb{R}^n)},$$

where

$$A_1 := 2^{n+2} A_2 + 2^n n^{\frac{n}{2}} + 4C,$$

where C is the constant in the Hörmander condition. This completes the proof. □

3.2 Calderon-Zygmund theorem

Now we prove our main result, often called the Calderon-Zygmund theorem or the Calderon-Zygmund inequality.

Theorem 42. *Let K be CZ kernel satisfying the Hörmander condition. Let $1 < p < \infty$ and let $f \in L^p(\mathbb{R}^n)$. For any $\varepsilon > 0$, define the operators*

$$T_\varepsilon f(x) = \int_{\mathbb{R}^n \setminus B_\varepsilon(x)} K(x-y) f(y) \, \mathrm{d}y \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

Then we have

(i) For every $\varepsilon > 0$, $T_\varepsilon f \in L^p(\mathbb{R}^n)$ and there exists a constant $A_p > 0$, independent of f and $\varepsilon > 0$, such that we have the estimates

$$\|T_\varepsilon f\|_{L^p(\mathbb{R}^n)} \leq A_p \|f\|_{L^p(\mathbb{R}^n)}.$$

(ii) $T_\varepsilon f$ converges to a limit, denoted by Tf in $L^p(\mathbb{R}^n)$ as $\varepsilon \to 0$ and the map $f \mapsto Tf$ defines a bounded linear operator from $L^p(\mathbb{R}^n)$ to itself and satisfies

$$\|Tf\|_{L^p(\mathbb{R}^n)} \leq A_p \|f\|_{L^p(\mathbb{R}^n)}.$$

Remark 43. *As the proof will reveal, we also have $A_p = A_{p'}$.*

Proof. We first prove (i). The operators T_ε are of strong type $(2,2)$, by Theorem 13 and also of weak type $(1,1)$, by Theorem 41, uniformly w.r.t. $\varepsilon > 0$. Thus, by the Marcinkiewicz interpolation theorem, The operators T_ε are also of strong type (p,p) for any $1 < p \leq 2$, again uniformly w.r.t. $\varepsilon > 0$. This proves (i) if $1 < p \leq 2$.

If $2 < p < \infty$, we proceed via a duality argument. Define the kernel

$$\widetilde{K}(x) := K(-x) \qquad \text{for all } x \in \mathbb{R}^n \setminus \{0\}.$$

Define the operators

$$\widetilde{T}_\varepsilon f(x) = \int_{\mathbb{R}^n \setminus B_\varepsilon(x)} \widetilde{K}(x-y) f(y)\ \mathrm{d}y \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

Note that $\widetilde{K}$ is a CZ kernel and satisfies the Hörmander condition with the same constants as K. Now since $2 < p < \infty$, we have $1 < p' < 2$. Thus, by the arguments above, the operators $\widetilde{T}_\varepsilon$ are of strong type (p', p'), uniformly w.r.t. $\varepsilon > 0$. More precisely, we have the estimates

$$\left\| \widetilde{T}_\varepsilon \phi \right\|_{L^{p'}(\mathbb{R}^n)} \leq A_{p'} \|\phi\|_{L^{p'}(\mathbb{R}^n)},$$

for all $\varepsilon > 0$ and any $\phi \in L^{p'}(\mathbb{R}^n)$. Note that the constant $A_{p'}$ for $\widetilde{T}_\varepsilon$ is the same for T_ε. Now for any $f, \phi \in C_c^\infty(\mathbb{R}^n)$, we have

$$\begin{aligned}
\int_{\mathbb{R}^n} T_\varepsilon f(x)\, \phi(x)\ \mathrm{d}x &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n \setminus B_\varepsilon(x)} K(x-y) f(y)\ \mathrm{d}y \right] \phi(x)\ \mathrm{d}x \\
&\overset{\text{Fubini}}{=} \int_{\mathbb{R}^n} f(y) \left[\int_{\mathbb{R}^n \setminus B_\varepsilon(y)} K(x-y)\, \phi(x)\ \mathrm{d}x \right] \mathrm{d}y \\
&= \int_{\mathbb{R}^n} f(y) \left[\int_{\mathbb{R}^n \setminus B_\varepsilon(y)} \widetilde{K}(y-x)\, \phi(x)\ \mathrm{d}x \right] \mathrm{d}y \\
&= \int_{\mathbb{R}^n} f(y)\, \widetilde{T}_\varepsilon \phi(y)\ \mathrm{d}y.
\end{aligned}$$

By the dual characterization of L^p norms and the density of $C_c^\infty(\mathbb{R}^n)$, we deduce

$$\begin{aligned}
\|T_\varepsilon f\|_{L^p(\mathbb{R}^n)} &= \sup_{\substack{\phi \in C_c^\infty(\mathbb{R}^n), \\ \|\phi\|_{L^{p'}(\mathbb{R}^n)} \leq 1}} \left| \int_{\mathbb{R}^n} T_\varepsilon f(x)\, \phi(x)\ \mathrm{d}x \right| \\
&= \sup_{\substack{\phi \in C_c^\infty(\mathbb{R}^n), \\ \|\phi\|_{L^{p'}(\mathbb{R}^n)} \leq 1}} \left| \int_{\mathbb{R}^n} f(y)\, \widetilde{T}_\varepsilon \phi(y)\ \mathrm{d}y \right| \\
&\overset{\text{Hölder}}{\leq} \sup_{\substack{\phi \in C_c^\infty(\mathbb{R}^n), \\ \|\phi\|_{L^{p'}(\mathbb{R}^n)} \leq 1}} \|f\|_{L^p(\mathbb{R}^n)} \left\| \widetilde{T}_\varepsilon \phi \right\|_{L^{p'}(\mathbb{R}^n)} \leq A_{p'} \|f\|_{L^p(\mathbb{R}^n)},
\end{aligned}$$

for any $f \in C_c^\infty(\mathbb{R}^n)$. By density of $C_c^\infty(\mathbb{R}^n)$ in $L^p(\mathbb{R}^n)$, this estimate holds for any $f \in L^p(\mathbb{R}^n)$. This completes the proof of (i).

The proof of (ii) can now be derived from the uniform bound in (i) by arguing exactly as was done in the proof of (ii) of Theorem 13. □

3.3 L^p estimate for Newtonian potential

Now we are almost ready to use the Calderon-Zygmund theorem to prove L^p estimates for the Newton's potential. First we need a notation.

Notation 44. *For any $1 \leq p \leq \infty$, we define the space*

$$L_c^p(\mathbb{R}^n) := \{f \in L^p(\mathbb{R}^n) : \operatorname{supp} f \text{ is a compact set in } \mathbb{R}^n\}.$$

Theorem 45. *Let $f \in L_c^p(\mathbb{R}^n)$ for some $1 < p < \infty$. Then there exists a $w \in W_{loc}^{2,p}(\mathbb{R}^n)$ with $\nabla^2 w \in L^p(\mathbb{R}^n)$ and satisfies*

$$-\Delta w = f \qquad \text{in } \mathbb{R}^n,$$

in the sense of distributions and there exists a constant $C_p = C_p(p,n) > 0$ such that we have the estimate

$$\left\|\nabla^2 w\right\|_{L^p(\mathbb{R}^n)} \leq C_p \|f\|_{L^p(\mathbb{R}^n)}.$$

Remark 46. *As the proof will show, if $f \in C_c^\infty(\mathbb{R}^n)$, and Nf denotes the Newtonian potential of f, defined by*

$$Nf := \mathcal{N} * f,$$

then

$$w \equiv Nf \qquad \text{in } \mathbb{R}^n.$$

For this reason, we would just call w as ***the*** *Newtonian potential of f, when $f \in L^p(\mathbb{R}^n)$ and would henceforth denote w simply by the notation Nf or $\mathcal{N} * f$. Note that as the Newtonian kernel and its first derivative is only locally integrable, but not in $L^1(\mathbb{R}^n)$, Nf is in general, never in $W^{2,p}(\mathbb{R}^n)$. Thus, there is no easy way to make sense of the convolution $\mathcal{N} * f$ directly when $f \in L^p(\mathbb{R}^n)$.*

Proof. First assume $f \in C_c^\infty(\mathbb{R}^n)$ and set $\mathcal{C} = \operatorname{supp} f$. Then Nf is well defined and is in fact a smooth function and satisfies the PDE in the pointwise sense. Clearly, the map

$$f \mapsto \nabla^2 Nf$$

is a CZ operator which satisfies the Hörmander condition. Hence, by the Calderon-Zygmund inequality, we have

$$\left\|\nabla^2 Nf\right\|_{L^p(\mathbb{R}^n)} \leq C_p \|f\|_{L^p(\mathbb{R}^n)}.$$

Now let $K \subset \mathbb{R}^n$ be any compact subset. We plan to estimate

$$\|Nf\|_{L^p(K)} \qquad \text{and} \qquad \|\nabla Nf\|_{L^p(K)} .$$

Now let $\phi \in C_c^\infty(\mathbb{R}^n)$ such that $\phi \equiv 1$ in an open neighborhood of the compact subset $K - \mathcal{C}$, defined as

$$K - \mathcal{C} := \{x - y : x \in K, y \in \mathcal{C}\} .$$

Now, for any $x \in K$, we have

$$\begin{aligned} Nf(x) &= \int_{\mathbb{R}^n} \mathcal{N}(x-y) f(y) \, \mathrm{d}y \\ &= \int_{\mathbb{R}^n} (\phi\mathcal{N})(x-y) f(y) \, \mathrm{d}y = [\phi\mathcal{N} * f](x) . \end{aligned}$$

Now, it is easy to show that $\mathcal{N}$ is locally integrable around the origin in $\mathbb{R}^n$. Since $\mathcal{N}$ is a smooth function away from the origin anyway, we deduce that $\mathcal{N} \in L^1_{loc}(\mathbb{R}^n)$. Since $\phi \in C_c^\infty(\mathbb{R}^n)$, we have $\phi\mathcal{N} \in L^1(\mathbb{R}^n)$. Thus, by the Young's inequality for convolutions, we deduce

$$\|Nf\|_{L^p(K)} = \|\phi\mathcal{N} * f\|_{L^p(K)} \leq \|\phi\mathcal{N} * f\|_{L^p(\mathbb{R}^n)} \leq \|\phi\mathcal{N}\|_{L^1(\mathbb{R}^n)} \|f\|_{L^p(\mathbb{R}^n)} .$$

Exactly similar arguments prove

$$\begin{aligned} \|\nabla Nf\|_{L^p(K)} = \|\phi\nabla\mathcal{N} * f\|_{L^p(K)} &\leq \|\phi\nabla\mathcal{N} * f\|_{L^p(\mathbb{R}^n)} \\ &\leq \|\phi\nabla\mathcal{N}\|_{L^1(\mathbb{R}^n)} \|f\|_{L^p(\mathbb{R}^n)} . \end{aligned}$$

Note that the numbers

$$\|\phi\mathcal{N}\|_{L^1(\mathbb{R}^n)} \qquad \text{and} \qquad \|\phi\nabla\mathcal{N}\|_{L^1(\mathbb{R}^n)}$$

depend on the compact sets K and $\mathcal{C}$ and the dimension n. Thus, we can write the estimates

$$\|Nf\|_{L^p(K)} + \|\nabla Nf\|_{L^p(K)} \leq C(n, K, \mathcal{C}) \|f\|_{L^p(\mathbb{R}^n)} .$$

This proves our result when $f \in C_c^\infty(\mathbb{R}^n)$.

For the general case, let $f \in L^p_c(\mathbb{R}^n)$ and let $\mathcal{C} = \operatorname{supp} f$. By approximation, we can find a sequence $\{f_s\}_{s\in\mathbb{N}} \subset C_c^\infty(\mathbb{R}^n)$ such that

$$f_s \to f \qquad \text{strongly in } L^p(\mathbb{R}^n)$$

and

$$\operatorname{supp} f_s \subset \mathcal{C} \qquad \text{for every } s \in \mathbb{N}.$$

Thus, by our arguments in the previous case, we have

$$\begin{aligned} \left\|\nabla^2 Nf_{s_1} - \nabla^2 Nf_{s_2}\right\|_{L^p(\mathbb{R}^n)} &= \left\|\nabla^2 N(f_{s_1} - f_{s_2})\right\|_{L^p(\mathbb{R}^n)} \\ &\leq C_p \|f_{s_1} - f_{s_2}\|_{L^p(\mathbb{R}^n)} \to 0 \qquad \text{as } s_1, s_2 \to \infty. \end{aligned}$$

Also, for any compact set K, we have

$$\|Nf_{s_1} - Nf_{s_2}\|_{L^p(K)} + \|\nabla Nf_{s_1} - \nabla Nf_{s_2}\|_{L^p(K)} \\ \leq C(n, K, \mathcal{C}) \|f_{s_1} - f_{s_2}\|_{L^p(\mathbb{R}^n)} \to 0 \qquad \text{as } s_1, s_2 \to \infty.$$

This shows that for any compact set K, the sequence of smooth functions $\{Nf_s\} \subset C^\infty(\mathbb{R}^n)$, restricted to K, defines a Cauchy sequence is $W^{2,p}(K)$ and thus converges in $W^{2,p}(K)$ to a limit, which we denote by h_K. By the strong convergence in $L^p(K)$, we also have

$$Nf_s(x) \text{ is convergent for a.e. } x \in K.$$

Since K is arbitrary, we have

$$Nf_s(x) \text{ is convergent for a.e. } x \in \mathbb{R}^n.$$

Set

$$w(x) = \lim_{s\to\infty} Nf_s(x) \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

It is now easy to check that w must agree a.e. with h_K in K. Again, since $K \subset \mathbb{R}^n$ is an arbitary compact subset, we have $w \in W^{2,p}_{loc}(\mathbb{R}^n)$. Moreover, we have the estimates

$$\left\|\nabla^2 w\right\|_{L^p(\mathbb{R}^n)} \leq \liminf_{s\to\infty} \left\|\nabla^2 Nf_s\right\|_{L^p(\mathbb{R}^n)} \\ \leq C_p \liminf_{s\to\infty} \|f_s\|_{L^p(\mathbb{R}^n)} = C_p \|f\|_{L^p(\mathbb{R}^n)}.$$

To see that w satisfies the PDE in the sense of distributions, pick $\phi \in C_c^\infty(\mathbb{R}^n)$ and set $K := \operatorname{supp}\phi$. Since

$$Nf_s \to w \qquad\qquad \text{strongly in } L^p(K),$$

we deduce

$$\begin{aligned} -\int_{\mathbb{R}^n} w\Delta\phi &= -\int_K w\Delta\phi \\ &= -\lim_{s\to\infty} \int_K Nf_s \Delta\phi \\ &= -\lim_{s\to\infty} \int_{\mathbb{R}^n} Nf_s \Delta\phi \\ &= -\lim_{s\to\infty} \int_{\mathbb{R}^n} \phi \Delta Nf_s = \lim_{s\to\infty} \int_{\mathbb{R}^n} \phi f_s = \int_{\mathbb{R}^n} \phi f. \end{aligned}$$

This completes the proof. □

4 L^p estimates for elliptic equations via singular integrals

4.1 Interior L^p estimates for the Laplacian

Now we can prove the interior $W^{2,p}$ estimates for the Laplacian.

Theorem 47. *Let $n \geq 3$ and $\Omega \subset \mathbb{R}^n$ be a bounded open set. Let $f \in L^p(\Omega)$ for some $1 < p < \infty$ and let $u \in L^p(\Omega)$ satisfy*

$$-\Delta u = f \qquad\qquad \text{in } \Omega,$$

in the sense of distributions. Then $u \in W^{2,p}_{loc}(\Omega)$ and for any $\Omega_1 \subset\subset \Omega$, there exists a constant $C = C(n, p, \Omega, \Omega_1) > 0$ such that we have the estimate

$$\|u\|_{W^{2,p}(\Omega_1)} \leq C\left(\|u\|_{L^p(\Omega)} + \|f\|_{L^p(\Omega)}\right).$$

Proof. Define

$$\tilde{f} := \begin{cases} f & \text{in } \Omega, \\ 0 & \text{in } \mathbb{R}^n \setminus \Omega. \end{cases}$$

Then clearly, $\tilde{f} \in L^p_c(\mathbb{R}^n)$. Let $w = N\tilde{f}$, i.e. the function w given by Theorem 45. Set $v := u - w$. Then clearly

$$-\Delta v = 0 \qquad\qquad \text{in } \Omega,$$

in the sense of distributions. By Weyl's lemma (see Theorem 4.7, Page 118 in [1]), the distribution (actually an L^p function here) v is a smooth harmonic function in Ω. Now, the standard derivative estimate for harmonic functions (see Theorem 7, page 29 of [2]) implies that we have

$$\sup_{x\in\overline{\Omega_1}} |v|, \sup_{x\in\overline{\Omega_1}} |\nabla v|, \sup_{x\in\overline{\Omega_1}} \left|\nabla^2 v\right| \leq C(\Omega, \Omega_1, n)\|v\|_{L^1(\Omega)}.$$

Hence we deduce

$$\|v\|_{W^{2,p}(\Omega_1)} \leq C(\Omega, \Omega_1, n, p)\|v\|_{L^1(\Omega)} \overset{\text{Hölder}}{\leq} C(\Omega, \Omega_1, n, p)\|v\|_{L^p(\Omega)}.$$

Hence, we have

$$\begin{aligned}\|u\|_{W^{2,p}(\Omega_1)} &\leq \|v\|_{W^{2,p}(\Omega_1)} + \|w\|_{W^{2,p}(\Omega_1)} \\ &\leq C(\Omega, \Omega_1, n, p)\|v\|_{L^p(\Omega)} + C(\Omega, \Omega_1, n, p)\left\|\tilde{f}\right\|_{L^p(\mathbb{R}^n)}.\end{aligned}$$

But we also have

$$\begin{aligned}\|v\|_{L^p(\Omega)} &= \|u - w\|_{L^p(\Omega)} \\ &\leq \|u\|_{L^p(\Omega)} + \|w\|_{L^p(\Omega)} \\ &\leq \|u\|_{L^p(\Omega)} + C(\Omega, n, p)\left\|\tilde{f}\right\|_{L^p(\mathbb{R}^n)}.\end{aligned}$$

Thus, we arrive at

$$\begin{aligned}\|u\|_{W^{2,p}(\Omega_1)} &\leq C(\Omega,\Omega_1,n,p)\left(\|u\|_{L^p(\Omega)}+\left\|\tilde{f}\right\|_{L^p(\mathbb{R}^n)}\right)\\ &= C(\Omega,\Omega_1,n,p)\left(\|u\|_{L^p(\Omega)}+\|f\|_{L^p(\Omega)}\right).\end{aligned}$$

This completes the proof. □

Note that if $g \in L^p_c(\mathbb{R}^n)$, then for any $1 \leq i, j \leq n$, we have

$$\frac{\partial}{\partial x_j}\left[\mathcal{N} * \left(\frac{\partial g}{\partial x_i}\right)\right] \simeq \frac{\partial^2 \mathcal{N}}{\partial x_i \partial x_j} * g.$$

Hence, for any $F \in L^p_c(\mathbb{R}^n;\mathbb{R}^n)$, each component of the map

$$F \mapsto \nabla^2 N F \simeq \nabla N(\operatorname{div} F)$$

is also a CZ operator satisfying the Hörmander conditions. Thus, exactly the same arguments as above proves the following result about gradient L^p estimates.

Theorem 48. *Let $n \geq 3$ and $\Omega \subset \mathbb{R}^n$ be a bounded open set. Let $F \in L^p(\Omega;\mathbb{R}^n)$ for some $1 < p < \infty$ and let $u \in L^p(\Omega)$ satisfy*

$$-\Delta u = \operatorname{div} F \qquad \text{in } \Omega,$$

in the sense of distributions. Then $u \in W^{1,p}_{loc}(\Omega)$ and for any $\Omega_1 \subset\subset \Omega$, there exists a constant $C = C(n,p,\Omega,\Omega_1) > 0$ such that we have the estimate

$$\|u\|_{W^{1,p}(\Omega_1)} \leq C\left(\|u\|_{L^p(\Omega)} + \|F\|_{L^p(\Omega;\mathbb{R}^n)}\right).$$

If $1 < p < n$, similar arguments coupled with Sobolev embedding proves the following.

Theorem 49. *Let $n \geq 3$ and $\Omega \subset \mathbb{R}^n$ be a bounded open set. Let $f \in L^{\frac{np}{n+p}}(\Omega)$ and $F \in L^p(\Omega;\mathbb{R}^n)$ for some $1 < p < n$. Let $u \in L^p(\Omega)$ satisfy*

$$-\Delta u = f + \operatorname{div} F \qquad \text{in } \Omega,$$

in the sense of distributions. Then $u \in W^{1,p}_{loc}(\Omega)$ and for any $\Omega_1 \subset\subset \Omega$, there exists a constant $C = C(n,p,\Omega,\Omega_1) > 0$ such that we have the estimate

$$\|u\|_{W^{1,p}(\Omega_1)} \leq C\left(\|u\|_{L^p(\Omega)} + \|f\|_{L^{\frac{np}{n+p}}(\Omega)} + \|F\|_{L^p(\Omega;\mathbb{R}^n)}\right).$$

Remark 50. *Note that if we are trying to solve the Dirichlet boundary value problem*

$$\begin{cases} -\Delta u = f & \text{in } \Omega,\\ u = 0 & \text{on } \partial\Omega,\end{cases}$$

with $F \in L^p(\Omega;\mathbb{R}^n)$ and obtain interior L^pestimates for a solution, the ***existence of solutions*** *is* ***not*** *assured if $1 < p < 2$, as the Newtonian potential solution would not satisfy the boundary condition. For $p \geq 2$, one can use standard variational methods (or Lax-Milgram argument) to obtain existence of a solution in $W^{1,2}(\Omega)$ and argue by bootstrapping. For $1 < p < 2$, even existence of a solution to the homogeneous Dirichlet BVP requires* ***global estimates*** *and* ***uniqueness arguments*** *coupled with an* ***approximation procedure*** *to establish the existence of solutions. We leave it to the interested reader to carefully work out the argument in detail.*

4.2 Interior L^p estimates for constant coefficients

Consider the equation

$$-\operatorname{div}(A\nabla u) = f \qquad \text{in } \Omega,$$

where $\Omega \subset \mathbb{R}^n$ is a bounded open set, $A \in \mathrm{Symm}_{n\times n}$ is a symmetric $n \times n$ matrix which is uniformly elliptic, i.e. there exists a constant $\lambda > 0$ such that

$$\langle A\xi, \xi\rangle \geq \lambda |\xi|^2 \qquad \text{for every } \xi \in \mathbb{R}^n.$$

We are interested in deriving the interior $W^{2,p}$ estimates. There is a simple trick to reduce this question to the interior $W^{2,p}$ estimate for the Laplacian, which we now describe.

Since A is symmetric and uniformly elliptic, all its eigenvalues are positive and A is diagonalizable. Thus, there exists a matrix $P \in \mathbb{SO}(n)$ and a diagonal matrix

$$D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of the matrix A such that

$$A = P^{\intercal} A P.$$

Denote

$$\sqrt{D} := \operatorname{diag}\left(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}\right).$$

Now we set

$$\tilde{\Omega} := \left\{x \in \mathbb{R}^n : \sqrt{D}Px \in \Omega\right\}$$

and

$$v(x) := u\left(\sqrt{D}Px\right) \qquad \text{for all } x \in \tilde{\Omega}.$$

Since the map $x \mapsto \sqrt{D}Px$ defines a smooth affine diffeomorphism of $\mathbb{R}^n$ to itself, $\tilde{\Omega}$ is also open and bounded. Now it is easy to verify by direct calculation that we have

$$-\Delta v(x) = -\operatorname{div}(A\nabla u)\left(\sqrt{D}Px\right) \qquad \text{for all } x \in \tilde{\Omega},$$

if $u \in C^2(\Omega)$. But these also holds in the weak sense for $W^{2,p}$ functions and thus, proving $W^{2,p}$ estimate for u is reduced to proving $W^{2,p}$ estimate for v. Analogous considerations hold for $W^{1,p}$ estimates as well.

4.3 Interior L^p estimates for Lipschitz coefficients

Now we consider the equation

$$-\operatorname{div}\left(A\left(x\right)\nabla u\right) = \operatorname{div} F \qquad \text{in } \Omega,$$

where $\Omega \subset \mathbb{R}^n$ is a bounded open set, $A \in \operatorname{Lip}\left(\overline{\Omega}; \operatorname{Symm}_{n\times n}\right)$ is a symmetric $n \times n$ matrix field which is uniformly elliptic, i.e. there exists a constant $\lambda > 0$ such that

$$\langle A\left(x\right)\xi, \xi\rangle \geq \lambda\left|\xi\right|^2 \qquad \text{for every } \xi \in \mathbb{R}^n \text{ and every } x \in \overline{\Omega}.$$

We are interested in deriving the interior $W^{1,p}$ estimates. For the result, the assumption of Lipschitz continuity for the coefficient is not sharp. But this makes our life quite a bit simpler, so we would prove this under this assumption. The plan is to use the **Korn's freezing trick** and write

$$-\operatorname{div}\left(A\left(x_0\right)\nabla u\right) = \operatorname{div}\left(\left[A\left(x\right) - A\left(x_0\right)\right]\nabla u\right) + \operatorname{div} F \qquad \text{in } \Omega,$$

for some $x_0 \in \Omega$. Now for any radius $R > 0$ such that $B_{2R}\left(x_0\right) \subset\subset \Omega$, assuming $u \in W^{1,p}\left(B_{2R}\left(x_0\right)\right)$, we would have the estimate

$$\begin{aligned}
&\left\|u\right\|_{W^{1,p}(B_R(x_0))} \\
&\leq C\left(\left\|u\right\|_{L^p(\Omega)} + \left\|F\right\|_{L^p(\Omega;\mathbb{R}^n)} + + \left\|\left[A\left(x\right) - A\left(x_0\right)\right]\nabla u\right\|_{L^p(B_{2R}(x_0);\mathbb{R}^n)}\right) \\
&\leq C\left(\left\|u\right\|_{L^p(\Omega)} + \left\|F\right\|_{L^p(\Omega;\mathbb{R}^n)} + 2R\operatorname{Lip}\left(A\right)\left\|\nabla u\right\|_{L^p(B_{2R}(x_0);\mathbb{R}^n)}\right).
\end{aligned}$$

We plan to absorb the last term on the RHS in the LHS by choosing $R > 0$ small enough. Unfortunately, there are two issues with this approach. The first is that we do not know $u \in W^{1,p}\left(B_{2R}\left(x_0\right)\right)$ to begin with and the second, somewhat more serious issue is that the last term on the right has L^p norm of ∇u on a ball of radius $2R$, whereas on the left we have the $W^{1,p}$ norm of u on a ball of radius R, i.e. the norms are on the **different sets**. We would address both these difficulties by a **localization**, **approximation** and a **covering** argument.

Localization: Fix some $\Omega_1 \subset\subset \Omega$. Choose Ω_2 such that we have

$$\Omega_1 \subset\subset \Omega_2 \subset\subset \Omega.$$

Now choose $\eta \in C_c^\infty\left(\Omega_2\right)$ such that $\eta \equiv 1$ on Ω_1. Set

$$v := \eta u \qquad \text{in } \mathbb{R}^n.$$

Approximation: Now choose a standard mollifying kernel $\psi \in C_c^\infty\left(B_1\left(0\right)\right)$ and for $\delta > 0$, set

$$v_\delta := v * \psi_\delta,$$

where

$$\psi_\delta(x) := \frac{1}{\delta^n}\psi\left(\frac{x}{\delta}\right) \qquad \text{for all } x \in \mathbb{R}^n.$$

Since $\operatorname{supp} v \subset \Omega_2 \subset\subset \Omega$, there exists $\delta_0 > 0$ small enough such that $\operatorname{supp} v_\delta \subset\subset \Omega_{\delta_0}$ for all $0 < \delta < \delta_0$, where

$$\Omega_{\delta_0} := \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \delta_0\}.$$

We plan to derive the equation satisfied by v_δ. By a straight forward calculation, we have

$$-\operatorname{div}(A(x)\nabla v_\delta) = H(u, \eta, \delta) \qquad \text{in } \Omega_{\delta_0},$$

where

$$H(u, \eta, \delta) := G(u, \eta) * \psi_\delta - \operatorname{div}[A(x)(\nabla v * \psi_\delta) - (A(x)\nabla v) * \psi_\delta]$$

and

$$\begin{aligned} G(u, \eta) &:= \eta \operatorname{div} F - \langle \nabla\eta, A(x)(\nabla u)\rangle - \operatorname{div}(uA(x)(\nabla\eta)) \\ &= -\langle \nabla\eta, F + A(x)(\nabla u)\rangle - \operatorname{div}(\eta F + uA(x)(\nabla\eta)) \\ &:= -G_1(u, \eta) - \operatorname{div} G_2(u, \eta). \end{aligned}$$

Now we first claim if $u \in W^{1,\frac{np}{n+p}}(\Omega)$, then

$$H_3 := A(x)(\nabla v * \psi_\delta) - (A(x)\nabla v) * \psi_\delta \in L^p_c(\mathbb{R}^n; \mathbb{R}^n),$$

with estimates which are independent of $0 < \delta < \delta_0$. To see this, we write

$$\begin{aligned} |A(x)(\nabla v * \psi_\delta)(x) &- [(A(x)\nabla v) * \psi_\delta](x)| \\ &\le \int_{B(0,\delta)} |A(x) - A(x-z)| |\nabla v(x-z)| \psi_\delta(z)\ \mathrm{d}z \\ &\le \operatorname{Lip}(A) \int_{B(0,\delta)} |z| |\nabla v(x-z)| \psi_\delta(z)\ \mathrm{d}z \\ &\le \delta \operatorname{Lip}(A) \int_{B(0,\delta)} |\nabla v(x-z)| \psi_\delta(z)\ \mathrm{d}z \\ &= \delta \operatorname{Lip}(A) [|\nabla v| * \psi_\delta](x). \end{aligned}$$

Thus, by Young's inequality for convolutions, we deduce

$$\begin{aligned} \|A(\nabla v * \psi_\delta) &- (A(x)\nabla v) * \psi_\delta\|_{L^p(\mathbb{R}^n)} \\ &\le \delta \operatorname{Lip}(A) \||\nabla v| * \psi_\delta\|_{L^p(\mathbb{R}^n)} \\ &\le \delta \operatorname{Lip}(A) \|\psi_\delta\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \|\nabla v\|_{L^{\frac{np}{n+p}}(\mathbb{R}^n)}. \end{aligned}$$

Now, since $v = \eta u$, we clearly have

$$\|\nabla v\|_{L^{\frac{np}{n+p}}(\mathbb{R}^n)} \leq C(\eta) \|u\|_{W^{1,\frac{np}{n+p}}(\Omega)}.$$

Also, a direct computation yields

$$\begin{aligned}\|\psi_\delta\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} &= \frac{1}{\delta^n}\left(\int_{\mathbb{R}^n}\left[\psi\left(\frac{x}{\delta}\right)\right]^{\frac{n}{n-1}}\ \mathrm{d}x\right)^{\frac{n-1}{n}} \\ &= \frac{1}{\delta^n}\left(\delta^n\int_{\mathbb{R}^n}\left[\psi(z)\right]^{\frac{n}{n-1}}\ \mathrm{d}z\right)^{\frac{n-1}{n}} = \frac{1}{\delta}\|\psi\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)}.\end{aligned}$$

Combining these last three estimates, we arrive at

$$\begin{aligned}&\|A(\nabla v * \psi_\delta) - (A(x)\nabla v) * \psi_\delta\|_{L^p(\mathbb{R}^n)} \\ &\qquad\leq C(\eta)\operatorname{Lip}(A)\|\psi\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)}\|u\|_{W^{1,\frac{np}{n+p}}(\Omega)}.\end{aligned}$$

Now note that we have

$$\operatorname{div} G_2(u,\eta) * \psi_\delta = \operatorname{div}\left[G_2(u,\eta) * \psi_\delta\right] := \operatorname{div} H_2.$$

Moreover, we have the estimate

$$\begin{aligned}\|G_2(u,\eta) * \psi_\delta\|_{L^p(\mathbb{R}^n)} &\leq \|\psi_\delta\|_{L^1(\mathbb{R}^n)}\|G_2(u,\eta)\|_{L^p(\mathbb{R}^n)} \\ &= \|\psi\|_{L^1(\mathbb{R}^n)}\|G_2(u,\eta)\|_{L^p(\mathbb{R}^n)} \\ &= \|\psi\|_{L^1(\mathbb{R}^n)}\|\eta F + uA(\nabla\eta)\|_{L^p(\mathbb{R}^n)} \\ &\leq C\left(\eta, \|A\|_{L^\infty(\Omega)}\right)\|\psi\|_{L^1(\mathbb{R}^n)}\left(\|F\|_{L^p(\Omega)} + \|u\|_{L^p(\Omega)}\right).\end{aligned}$$

By Sobolev embedding, this implies

$$\|G_2(u,\eta) * \psi_\delta\|_{L^p(\mathbb{R}^n)} \leq C\left(\|F\|_{L^p(\Omega)} + \|u\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right).$$

Now we show that if $u \in W^{1,\frac{np}{n+p}}(\Omega)$, then

$$H_1 := G_1(u,\eta) * \psi_\delta \in L^{\frac{np}{n+p}}(\mathbb{R}^n).$$

Indeed, we have

$$\begin{aligned}&\|G_1(u,\eta) * \psi_\delta\|_{L^{\frac{np}{n+p}}(\mathbb{R}^n)} \\ &\leq \|\psi_\delta\|_{L^1(\mathbb{R}^n)}\|G_1(u,\eta)\|_{L^{\frac{np}{n+p}}(\mathbb{R}^n)} \\ &= \|\psi\|_{L^1(\mathbb{R}^n)}\|G_1(u,\eta)\|_{L^{\frac{np}{n+p}}(\mathbb{R}^n)} \\ &= \|\psi\|_{L^1(\mathbb{R}^n)}\|\langle\nabla\eta, F + A(x)(\nabla u)\rangle\|_{L^{\frac{np}{n+p}}(\mathbb{R}^n)} \\ &\leq C\left(\eta, \|A\|_{L^\infty(\Omega)}\right)\|\psi\|_{L^1(\mathbb{R}^n)}\left(\|F\|_{L^{\frac{np}{n+p}}(\Omega)} + \|\nabla u\|_{L^{\frac{np}{n+p}}(\Omega)}\right) \\ &\leq C\left(\eta, \|A\|_{L^\infty(\Omega)}, \Omega\right)\|\psi\|_{L^1(\mathbb{R}^n)}\left(\|F\|_{L^p(\Omega)} + \|u\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right).\end{aligned}$$

To summarize, we have so far shown that v_δ satisfies the PDE

$$-\operatorname{div}\left(A\left(x\right)\nabla v_\delta\right) = -H_1 - \operatorname{div} H_2 - \operatorname{div} H_3 \qquad \text{in } \Omega_{\delta_0},$$

where $H_1 \in L^{\frac{np}{n+p}}\left(\mathbb{R}^n\right)$ and $H_2, H_3 \in L^p\left(\mathbb{R}^n;\mathbb{R}^n\right)$ if $u \in W^{1,\frac{np}{n+p}}\left(\Omega\right)$ and we have the estimates

$$\begin{aligned}
\left\|H_1\right\|_{L^{\frac{np}{n+p}}\left(\mathbb{R}^n\right)} &\leq C\left(\left\|F\right\|_{L^p(\Omega)} + \left\|u\right\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right), \\
\left\|H_2\right\|_{L^p\left(\mathbb{R}^n\right)} &\leq C\left(\left\|F\right\|_{L^p(\Omega)} + \left\|u\right\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right) \\
\left\|H_3\right\|_{L^p\left(\mathbb{R}^n\right)} &\leq C\left\|u\right\|_{W^{1,\frac{np}{n+p}}(\Omega)}.
\end{aligned}$$

Now we fix $\varepsilon > 0$ and by the uniform continuity of A, choose a radius $0 < R < \delta_0/4$ such that

$$\sup_{x\in\Omega_{\delta_0}} \left\|A - A\left(x\right)\right\|_{L^\infty(B_{2R}(x))} < \varepsilon.$$

Now since Ω_{δ_0} is precompact, we can cover Ω_{δ_0} by finitely many balls of radius R such that the number of overlapping is bounded above by a constant that depends only on the dimension n and the set Ω_{δ_0}, but not on R or δ. Thus, there are finitely many balls with centers $x_1, \ldots, x_N$ such that

$$\Omega_{\delta_0} \subset\subset \bigcup_{i=1}^{N} B_R\left(x_i\right).$$

Now for each $1 \leq i \leq N$, we write the PDE as

$$-\operatorname{div}\left(A\left(x_i\right)\nabla v_\delta\right) = -H_1 - \operatorname{div} H_2 - \operatorname{div} H_3 - \operatorname{div}\left(\left[A\left(x_i\right) - A\left(x\right)\right]\nabla v_\delta\right)$$

in $B_{2R}\left(x_i\right)$. Thus, for each $1 \leq i \leq N$, we have

$$\begin{aligned}
\left\|\nabla v_\delta\right\|_{L^p(B_R(x_i))} &\leq C\left(\left\|F\right\|_{L^p(\Omega)} + \left\|u\right\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right) \\
&\qquad + C\left\|A - A\left(x\right)\right\|_{L^\infty(B_{2R}(x_i))}\left\|\nabla v_\delta\right\|_{L^p(B_{2R}(x_i))} \\
&\leq C\left(\left\|F\right\|_{L^p(\Omega)} + \left\|u\right\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right) + C\varepsilon\left\|\nabla v_\delta\right\|_{L^p(B_{2R}(x_i))}.
\end{aligned}$$

Since the number of overlapping balls is bounded by a constant C_{overlap}, which is independent of R and δ, summing the estimates, we deduce

$$\begin{aligned}
\left\|\nabla v_\delta\right\|_{L^p\left(\bigcup_{i=1}^N B_R(x_i)\right)} &\leq C\left(\left\|F\right\|_{L^p(\Omega)} + \left\|u\right\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right) \\
&\qquad + CC_{\text{overlap}}\varepsilon\left\|\nabla v_\delta\right\|_{L^p\left(\bigcup_{i=1}^N B_{2R}(x_i)\right)}.
\end{aligned}$$

Thus, choosing $\varepsilon > 0$ small enough to absorb the last term in the left, we have

$$\left\|\nabla v_\delta\right\|_{L^p(\Omega)} \leq C\left(\left\|F\right\|_{L^p(\Omega)} + \left\|u\right\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right).$$

Now since v_δ has compact support in Ω, by Poincaré inequality, we have

$$\|v_\delta\|_{W^{1,p}(\Omega)} \leq C\left(\|F\|_{L^p(\Omega)} + \|u\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right).$$

Thus, $\{v_\delta\}_{\delta>0}$ is uniformly bounded in $W^{1,p}(\Omega)$ and thus, up to the extraction of a subsequence, converges weakly in $W^{1,p}(\Omega)$. But since the weak limit can only be v, we have

$$v_\delta \rightharpoonup v \qquad \text{weakly in } W^{1,p}(\Omega).$$

Thus, by weak lower semicontinuity of the norm, we have

$$\|v\|_{W^{1,p}(\Omega)} \leq \liminf_{\delta\to 0} \|v_\delta\|_{W^{1,p}(\Omega)} \leq C\left(\|F\|_{L^p(\Omega)} + \|u\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right).$$

But since $\eta \equiv 1$ on Ω_1, we deduce

$$\|u\|_{W^{1,p}(\Omega_1)} \leq \|v\|_{W^{1,p}(\Omega)} \leq C\left(\|F\|_{L^p(\Omega)} + \|u\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right).$$

Thus, we have proved that the following.

Theorem 51. *Let $\Omega \subset \mathbb{R}^n$ is a bounded open set. $A \in \operatorname{Lip}\left(\overline{\Omega}; \operatorname{Symm}_{n\times n}\right)$ is a symmetric $n\times n$ matrix field which is Lipschitz and uniformly elliptic, i.e. there exists a constant $\lambda > 0$ such that*

$$\langle A(x)\xi, \xi\rangle \geq \lambda |\xi|^2 \qquad \text{for every } \xi \in \mathbb{R}^n \text{ and every } x \in \overline{\Omega}.$$

Let $1 < p < \infty$ and let $F \in L^p(\Omega; \mathbb{R}^n)$. If $u \in W^{1,\frac{np}{n+p}}(\Omega)$ is a distributional solution of

$$-\operatorname{div}(A(x)\nabla u) = \operatorname{div} F \qquad \text{in } \Omega,$$

then $u \in W^{1,p}_{loc}(\Omega)$ and for any $\Omega_1 \subset\subset \Omega$, there exists a constant

$$C = C\left(n, p, \Omega_1, \Omega, \lambda, \operatorname{Lip} A, \|A\|_{L^\infty(\Omega)}\right) > 0$$

such that we have the estimate

$$\|u\|_{W^{1,p}(\Omega_1)} \leq C\left(\|F\|_{L^p(\Omega)} + \|u\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right).$$

By bootstrapping, one can actually have the following.

Theorem 52. *Let $\Omega \subset \mathbb{R}^n$ is a bounded open set. $A \in \operatorname{Lip}\left(\overline{\Omega}; \operatorname{Symm}_{n\times n}\right)$ is a symmetric $n\times n$ matrix field which is Lipschitz and uniformly elliptic, i.e. there exists a constant $\lambda > 0$ such that*

$$\langle A(x)\xi, \xi\rangle \geq \lambda |\xi|^2 \qquad \text{for every } \xi \in \mathbb{R}^n \text{ and every } x \in \overline{\Omega}.$$

Let $1 < q < p < \infty$ and let $F \in L^p(\Omega;\mathbb{R}^n)$. If $u \in W^{1,q}(\Omega)$ is a distributional solution of

$$-\operatorname{div}(A(x)\nabla u) = \operatorname{div} F \qquad \text{in } \Omega,$$

then $u \in W^{1,p}_{loc}(\Omega)$ and for any $\Omega_1 \subset\subset \Omega$, there exists a constant

$$C = C\left(n, p, \Omega_1, \Omega, \lambda, \operatorname{Lip} A, \|A\|_{L^\infty(\Omega)}\right) > 0$$

such that we have the estimate

$$\|u\|_{W^{1,p}(\Omega_1)} \leq C\left(\|F\|_{L^p(\Omega)} + \|u\|_{W^{1,q}(\Omega)}\right).$$

The proof is easy. Without loss of generality, we can assume $1 < q < n$. Now if $nq/(n-q) < p$, then applying the previous theorem, we have $u \in W^{1,\frac{nq}{n-q}}_{loc}$. Now if $nq/(n-q) \geq n$, then u is also in $W^{1,n-\varepsilon}_{loc}$ for any $\varepsilon > 0$ and we can choose $\varepsilon > 0$ such that $n(n-\varepsilon)/\varepsilon = p$. If $nq/(n-q) < n$ and $nq/(n-2q) < p$, we can apply the last result again to conclude that $u \in W^{1,\frac{nq}{n-2q}}_{loc}$. We can continue this process until we reach n, in which case we reach p in the next step, or till we reach p.

Remark 53. *Note that Theorem 52 is not always useful. However, for $2 < p < \infty$, this immediately establishes the desired interior L^p estimate, as existence theory gives the existence of a $W^{1,2}$ weak solution and we can apply the result with $q = 2$.*

4.4 Boundary estimates

We are not going to prove the boundary L^p estimates. We would only sketch the basic arguments. By localizing and flattening the boundary, the boundary estimates reduce to deriving the L^p estimates for solutions in a half ball. We would just show how a reflection can argument can reduce the L^p estimates in a half ball to interior estimates in a ball.

Proposition 54. *Let $1 < p < \infty$ and let $R > 0$. Let $F \in L^p\left(B_R^+;\mathbb{R}^n\right)$ and let $u \in C^\infty\left(B_R^+\right)$ be such that $u \equiv 0$ on $\partial B_R^+ \cap \{x \in \mathbb{R}^n : x_n = 0\}$ and u vanishes in a neighborhood of the curved boundary of B_R^+. Let u satisfy*

$$-\operatorname{div}(A\nabla u) = \operatorname{div} F \qquad \text{in } B_R^+,$$

where A is a constant symmetric $n \times n$ matrix which is uniformly elliptic. Set

$$\tilde{u}(x) := \begin{cases} u(x) & \text{if } x_n > 0, \\ -u(x_1, \ldots, x_{n-1}, -x_n) & \text{if } x_n < 0. \end{cases}$$

Then $\tilde{u} \in C_c^\infty(B_R)$ satisfies the equation

$$-\operatorname{div}\left(\tilde{A}\nabla\tilde{u}\right) = \operatorname{div}\tilde{F} \qquad \text{in } B_R,$$

where $\tilde{A}$ is also a symmetric uniformly elliptic matrix with the same ellipticity constant as A and $\tilde{F} \in L^p(B_R)$ with an estimate

$$\left\|\tilde{F}\right\|_{L^p(B_R;\mathbb{R}^n)} \leq c\,\|F\|_{L^p\left(B_R^+;\mathbb{R}^n\right)},$$

The proof is a straight forward calculation and is skipped.

4.5 Failure of the estimates at endpoints

Now we give examples to show that the L^p estimates does not extend to borderline cases, i.e. $p = 1$ and $p = \infty$.

Example 55 (Failure of L^1 estimate)**.** *Let $\mathbb{D} \subset \mathbb{R}^2$ be the open unit disk. Define*

$$u(x) = \log\log\frac{e}{|x|} \qquad \text{for a.e. } x \in \mathbb{D}.$$

Then $u \in W_0^{1,2}(\mathbb{D})$, $\Delta u \in L^1(\mathbb{D})$, but $u \notin W^{2,1}(\mathbb{D})$.

Example 56 (Failure of L^∞ estimate)**.** *Let $\mathbb{D} \subset \mathbb{R}^2$ be the open unit disk. Define, in polar coordinates,*

$$u(r,\theta) = r^2\log r\cos 2\theta \qquad \text{a.e. in } \mathbb{D}.$$

Then $u \in W_0^{1,2}(\mathbb{D})$, $\Delta u \in L^\infty(\mathbb{D})$, but $u \notin W^{2,\infty}(\mathbb{D})$.

5 BMO and interpolation

So far, we have derived the L^p estimates using the singular integral estimates. Recall that we have interpolated between a weak $(1,1)$ estimate and a strong $(2,2)$ estimate to obtain the result for $1 < p \leq 2$ and obtain the case $2 < p < \infty$ by duality. It is possible to somewhat reverse the manner of doing things. Roughly speaking, instead of interpolating between 'almost L^1 estimate and L^2 estimate, we can also interpolate between 'almost L^∞' estimate and L^2 estimate. One can also avoid singular integrals altogether and instead directly establish estimates for energy-weak solutions.

5.1 BMO and the John-Nirenberg estimate

We now define the BMO space, which is going to serve as our substitute for L^∞.

Definition 57. *Let Q be a n-dimensional cube in $\mathbb{R}^n$. We define the space of functions of bounded mean oscillation $BMO(Q)$ as:*

$$BMO(Q) := \left\{u \in L^1(Q) : \sup_{Q' \subset Q}\frac{1}{|Q'|}\int_{Q'}|u - (u)_{Q'}| < \infty\right\}$$

where the supremum is taken over all n-dimensional subcubes Q' of Q.

The function

$$[u]_{BMO(Q)} := \sup_{Q' \subset Q} \int_{Q'} |u - (u)_{Q'}|$$

forms a seminorm on $BMO(Q)$.

By $BMO(Q)\backslash\{0\}$ we will mean BMO with the equivalence class of 0 removed, where the equivalence relation is $u \sim v$ *if and only if* $u - v = constant$.

One of the important properties of BMO functions is the following estimate, known as the John-Nirenberg inequality.

Theorem 58 (John-Nirenberg lemma)**.** *Let Q_0 be a n-dimensional cube in $\mathbb{R}^n$. There are constants $c_1, c_2 > 0$ depending only on n such that*

$$\mu\left(\{|u - (u)_Q| > t\}\right) \leq c_1 e^{\frac{-c_2 t}{[u]_{BMO(Q)}}} |Q|$$

for all cubes with sides parallel to the axes $Q \subset Q_0$ and all $t > 0$.

Proof. We may assume without loss of generality that, $[u]_{BMO(Q)} = 1$ since

$$\{|u - (u)_Q| > t\} = \left\{\left|\frac{u}{[u]_{BMO(Q)}} - \frac{(u)_Q}{[u]_{BMO(Q)}}\right| > \frac{t}{[u]_{BMO(Q)}}\right\}$$

It also suffices to prove for Q_0, as the $BMO(Q) \leq BMO(Q_0)$ for any subcube $Q \subset Q_0$.

Now, choose an

$$\alpha > 1 \geq \frac{1}{|Q_0|}\int_{Q_0} |u - (u)_{Q_0}|$$

Now, apply the Calderon-Zymund decomposition to $|u - (u)_{Q_0}|$ with α to obtain a collection of non-overlapping cubes $\{Q_i^1\}$ such that

$$\alpha \leq \frac{1}{|Q_i^1|}\int_{Q_i^1} |u - (u)_{Q_0}| \leq 2^n \alpha$$

$$|u - (u)_{Q_0}| \leq \alpha \quad a.e.\ x \in Q_0 \backslash \bigcup_i Q_i^1$$

So we have

$$\sum_i |Q_i^1| \leq \frac{1}{\alpha}\int_{Q_0} |u - (u)_{Q_0}| \leq \frac{1}{\alpha}|Q_0|$$

and

$$|(u)_{Q_i^1} - (u)_{Q_0}| \leq \left|\frac{1}{|Q_i^1|}\int_{Q_i^1} u - \frac{1}{|Q_i^1|}\int_{Q_i^1} (u)_{Q_0}\right| \leq \frac{1}{|Q_i^1|}\int_{Q_i^1} |u - (u)_{Q_0}| \leq 2^n \alpha$$

Now, the idea is to iterate the CZ decomposition. The definition of the BMO seminorm gives us that

$$\frac{1}{|Q_i|}\int_{Q_i} |u - (u)_{Q_i}| \leq 1 < \alpha$$

Now, we apply the decomposition to $|u-(u)_{Q_i}|$ on each of the Q_i^1 to obtain a collection of non-overlapping subcubes $\{Q_j^2\}$ so that on each subcube Q_j^2 (which sits inside Q_i^1 say), we have

$$\alpha \leq \frac{1}{|Q_j^2|}\int_{Q_j^2} |u-(u)_{Q_i^1}| \leq 2^n\alpha$$

$$|u-(u)_{Q_i^1}| \leq \alpha \quad a.e.\ x \in Q_i^1 \backslash \bigcup_i Q_j^2$$

We have for the whole collection $\{Q_j^2\}$

$$\sum_j |Q_j^2| \leq \sum_i \frac{1}{\alpha}\int_{Q_i^1} |u-(u)_{Q_i^1}| \leq \frac{1}{\alpha}\sum_i |Q_i| \leq \frac{1}{\alpha^2}|Q_0|$$

So we have,

$$|u-(u)_{Q_0}| \leq 2.2^n\alpha \quad a.e.\ x \in Q_0 \backslash \bigcup_j Q_j^2$$

This is clear, since if $x \in Q_0 \backslash \bigcup_j Q_j^2$ and not in any of the $Q_i^1 s$, $|u-(u)_{Q_0}| \leq \alpha$ and if x is in some Q_i^1, then we may use triangle inequality as

$$\begin{aligned} |u-(u)_{Q_0}| &\leq |u-(u)_{Q_i^1}| + |(u)_{Q_0} - (u)_{Q_i^1}| \\ &\leq \alpha + 2^n\alpha \leq 2.2^n\alpha \end{aligned}$$

Continuing this process, for each integer $k \geq 1$, we have a collection of non-overlapping cubes $\{Q_i^k\}$ such that,

$$\sum_i |Q_i^k| \leq \frac{1}{\alpha^k}|Q_0|$$

and

$$|u-(u)_{Q_0}| \leq k.2^n\alpha \quad a.e.\ x \in Q_0 \backslash \bigcup_i Q_i^k$$

Thus,

$$|\{x \in Q_0 : |u(x)-(u)_{Q_0}| > k.2^n\alpha\}| \leq \sum_i |Q_i^k| \leq \frac{1}{\alpha^k}|Q_0|$$

For any t, there exists a k so that $t \in [k.2^n\alpha, (k+1).2^n\alpha)$. We have

$$\alpha^{-k} = \alpha.\alpha^{-(k+1)} = \alpha.e^{-(k+1)\log\alpha} \leq \alpha.e^{-\frac{\log\alpha}{2^n\alpha}t}$$

So we have

$$|\{x \in Q_0 : |u(x)-(u)_{Q_0}| > t\}| \leq \alpha e^{-\frac{\log\alpha}{2^n\alpha}t}|Q_0|.$$

This completes the proof. $\square$

This would imply a result that is going to be crucial for us. But first let us define the Campanato spaces. Let $\Omega \subset \mathbb{R}^n$ be Lipschitz domain. We will denote $B_\rho(x_0) \cap \Omega$ by $\Omega(x_0, \rho)$.

Definition 59. *We define the Campanato space $\mathcal{L}^{p,\lambda}(\Omega)$ for every $1 \le p \le \infty$ and $\lambda \ge 0$ as the collection of all $f \in L^p(\Omega)$ such that*

$$[u]^p_{\mathcal{L}^{p,\lambda}(\Omega)} := \sup_{x_0 \in \Omega;\, \rho > 0} \frac{1}{\rho^\lambda} \int_{\Omega(x_0,\rho)} |u - (u)_{x_0,\rho}|^p < +\infty$$

Remark 60. *Note that $\mathcal{L}^{1,n}$ is, by definition, BMO, where we have used balls instead of cubes, which changes nothing.*

Corollary 61. *For every $1 \le p < \infty$, the Campanato space $\mathcal{L}^{p,n}(Q_0)$ is isomorphic to $BMO(Q_0)$*

Proof. If we have a $u \in BMO(Q_0)$, then we have

$$\begin{aligned}\int_Q |u - (u)_Q|^p &= p \int_0^\infty t^{p-1} \left|\{x \in Q : |u(x) - (u)_Q| > t\}\right| dt \\ &\le p.c_1 \int_0^\infty t^{p-1} e^{-\frac{c_2 t}{[u]_{BMO(Q_0)}}} |Q| dt\end{aligned}$$

Making a substitution $\frac{c_2 t}{[u]_{BMO(Q_0)}} = s$, we have

$$\begin{aligned}&\le pc_1 \left(\frac{[u]_{BMO(Q_0)}}{c_2}\right)^p \int_0^\infty s^{p-1} e^{-s} ds \\ &\le C(n,p)[u]^p_{BMO(Q_0)}|Q|\end{aligned}$$

Dividing by $|Q|$ and letting side length of Q be ρ and taking supremum over $\rho > 0$, we get that

$$[u]_{\mathcal{L}^{p,n}(\Omega)} \le C(n,p)[u]_{BMO(Q_0)}$$

The converse directly follows from Jensen's inequality. If $u \in \mathcal{L}^{p,\lambda}(\Omega)$, then we have

$$\left(\frac{1}{|Q|}\int_Q |u - (u)_Q|\right)^p \le \frac{1}{|Q|}\int_Q |u - (u)_Q|^p$$

So taking supremum over $Q \subset Q_0$, we have

$$[u]^p_{BMO(Q_0)} \le [u]^p_{\mathcal{L}^{p,n}(Q_0)}$$

□

It also follows from this that if $u \in BMO(Q_0)$ then $u \in L^p(Q_0)$ for all $1 \le p < \infty$.

We state a theorem that gives equivalent statements to the John Nirenberg Lemma

Theorem 62. *The following are equivalent:*

1. $u \in BMO(Q_0)$
2. *There exist* $c_1, c_2 > 0$ *such that for* $Q \subset Q_0$ *we have,*

$$\mu\left(\{|u-(u)_Q| > t\}\right) \leq c_1 e^{\frac{-c_2 t}{[u]_{BMO(Q)}}} |Q|$$

3. *There exists* $c_3, c_4 > 0$ *so that for* $Q \subset Q_0$ *we have,*

$$\fint_Q e^{c_3|u-(u)_Q|} - 1 < c_4$$

4. *There exist* $c_5, c_6 > 0$ *so that,*

$$\left(\fint_Q e^{c_6 u}\right)\left(\fint_Q e^{-c_6} u\right) \leq c_5$$

5.2 Sharp maximal function and Fefferman-Stein inequality

Definition 63. *The sharp function of* $u \in L^1(Q_0)$ *as*

$$u^\sharp(x) = \sup_{x \in Q \subset Q_0} \frac{1}{|Q|} \int_Q |u(y) - (u)_Q| dy$$

We define the centered sharp function as:

$$\tilde{u}(x) := \sup_{Q(x) \subset Q} \frac{1}{|Q(x)|} \int_{Q(x)} |u(y) - (u)_{Q(x)}| dy$$

where the supremum is taken over cubes with center x. We have

$$\tilde{u}(x) \leq u^\sharp(x) \leq 2^n \tilde{u}(x), \quad [u]_{BMO(Q_0)} = \|\tilde{u}\|_{L^\infty(Q_0)}$$

We further have,

$$\begin{aligned}
u^\sharp(x) \leq 2^n \tilde{u}(x) &\leq 2^n \sup_{Q(x) \subset Q_0} \frac{1}{|Q(x)|} \int_{Q(x)} |u(y) - (u)_Q| \\
&\leq 2^n \sup_{Q(x) \subset Q_0} \frac{1}{|Q(x)|} \int_{Q(x)} |u(y)| dy + (u)_Q \\
&\leq 2^n.2. \sup_{Q(x) \subset Q_0} \frac{1}{|Q(x)|} \int_{Q(x)} |u(y)| dy \\
&\leq 2^{n+1} Mu(x)
\end{aligned}$$

Hence, if $u \in L^p(Q_0)$ for $1 < p \leq \infty$, then $u^\sharp \in L^p(Q_0)$. The converse is the following theorem important result.

Theorem 64 (Fefferman-Stein)**.** *Consider $u \in L^1(Q_0)$, and suppose that $u^\sharp \in L^p(Q_0)$ for some $p > 1$. Then $u \in L^p(Q_0)$ and*

$$\|u\|_{L^p(Q_0)} \leq c(n,p) \left\{ \|u^\sharp\|_{L^p(Q_0)} + |Q_0|^{1/p} \frac{1}{|Q_0|} \int_{Q_0} |u| \right\}$$

We begin with a proposition.

Proposition 65. *Set*

$$\mu(t) = \sum_i |Q_i^t|$$

where $\{Q_i^t\}$ is the Calderon-Zygmund family of cubes for $|u|$ at level t. Then we have

$$\mu((2^n+1)t) \leq \left|\left\{x \in Q_0 : u^\sharp > \beta t\right\}\right| + \beta\mu(t)$$

for any $\beta \in (0,1)$ and any t with

$$\frac{1}{|Q_0|} \int_{Q_0} |u| < t$$

Proof. Set $s = (2^n+1)t$. Let $\{Q_j^t\}$ and $\{Q_i^s\}$ be the Calderon-Zygmund family of cubes corresponding to t and s respectively. We have

$$\mu(s) = \sum_j \sum_{i : Q_i^s \subset Q_j^t} |Q_i^s|$$

For a fixed j, we have two possibilities:
Case 1: $Q_j^t \subset \{x \in Q_0 : u^\sharp > \beta t\}$

$$\sum_{i : Q_i^s \subset Q_j^t} |Q_i^s| \leq \left|\left\{x \in Q_0 : u^\sharp > \beta t\right\}\right|$$

Case 2: There is a $y \in Q_j^t$ so that $u^\sharp \leq \beta t$. In this case we have,

$$\frac{1}{|Q_j^t|} \int_{Q_j^t} |u - (u)_{Q_j^t}| \leq \beta t$$

So,

$$\begin{aligned}\frac{1}{|Q_i^s|} \int_{Q_i^s} |u - (u)_{Q_j^t}| &\geq \frac{1}{|Q_i^s|} \int_{Q_i^s} |u| - \frac{1}{|Q_i^s|} \int_{Q_i^s} |u| - (u)_{Q_j^t} \\ &\geq \frac{1}{|Q_i^s|} \int_{Q_i^s} |u| - \frac{1}{|Q_j^t|} \int_{Q_j^t} |u| \geq s - 2^n t = t\end{aligned}$$

So we have,

$$t \sum_{i : Q_i^s \subset Q_j^t} |Q_i^s| \leq \int_{Q_i^s} |u - (u)_{Q_j^t}| \leq \int_{Q_j^t} |u - (u)_{Q_j^t}| \leq \beta t |Q_j^t|$$

which gives,

$$\sum_{i:Q_i^s \subset Q_j^t} \leq \beta |Q_j^t|$$

In both cases, summing over j, we get

$$\mu(s) \leq \left|\left\{x \in Q_0 : u^\sharp > \beta t\right\}\right| + \beta\mu(t)$$

□

Now, we give a proof of Theorem (64).

Proof. We have by the above proposition,

$$\mu(t) \leq \left|\left\{x \in Q_0 : u^\sharp > \beta(2^n+1)^{-1}t\right\}\right| + \beta\mu\left((2^n+1)^{-1}t\right)$$

for

$$\frac{1}{|Q_0|}\int_{Q_0}|u| < \frac{t}{(2^n+1)}$$

Define

$$t_0 := (2^n+1)\frac{1}{|Q_0|}\int_{Q_0}|u|$$

We have,

$$F_u(t) \leq \mu(t)$$

Now,

$$\begin{aligned}\|u\|_{L^p(Q_0)}^p &= p\int_0^\infty t^{p-1}F_u(t)\\ &\leq p\int_0^\infty t^{p-1}\mu(t)\end{aligned}$$

For a $s > t_0$, we define

$$I_s := p\int_0^s t^{p-1}\mu(t)dt$$

And we have,

$$\begin{aligned}I_s &\leq p\int_0^{t_0} t^{p-1}\mu(t) + p\int_{t_0}^s t^{p-1}\mu(t)\\ &\leq (I) + (II)\end{aligned}$$

We have,

$$\begin{aligned}(I) &= p\int_0^{t_0} t^{p-1}\mu(t)\\ &\leq \int_0^{t_0} t^{p-1}\frac{1}{t}\int_{Q_0}|u|\\ &\leq c(n,p)\frac{1}{|Q_0|}\left(\int_{Q_0}|u|\right)^p\end{aligned}$$

and for (II), making use of the above proposition we get,

$$\begin{aligned}(II) &= p\int_{t_0}^{s} t^{p-1}\mu(t)\\ &\leq p\int_{t_0}^{s} t^{p-1}\left|\left\{u^\sharp > \beta(2^n+1)^{-1}t\right\}\right| + \beta p\int_{t_0}^{s} t^{p-1}\mu\left((2^n+1)^{-1}t\right)\\ &\leq (i)+(ii)\end{aligned}$$

Now, by a change of variable, we have

$$\begin{aligned}(i) &= p\int_{t_0}^{s} t^{p-1}\left|\left\{u^\sharp > \beta(2^n+1)^{-1}t\right\}\right| dt\\ &\leq p\left(\frac{2^n+1}{\beta}\right)^p\int_0^\infty s^{p-1}\left|\left\{u^\sharp > s\right\}\right| ds\\ &\leq \left(\frac{2^n+1}{\beta}\right)^p\int_{Q_0}|u^\sharp|^p\end{aligned}$$

and for (ii), we have again by change of variable,

$$\begin{aligned}(ii) &= \beta p\int_{t_0}^{s} t^{p-1}\mu\left((2^n+1)^{-1}t\right) dt\\ &\leq \beta p\,(2^n+1)^p\int_{(2^n+1)^{-1}t_0}^{(2^n+1)^{-1}s} s^{p-1}\mu(s)ds\end{aligned}$$

Noting that $s > (2^n+1)^{-1}s$ we have,

$$\begin{aligned}&\leq \beta p(2^n+1)^p\int_0^s s^{p-1}\mu(s)ds\\ &\leq \beta(2^n+1)^p I_s\end{aligned}$$

Choosing $\beta = \frac{1}{2}(2^n+1)^{-p}$, and combining everything till now, we have,

$$\frac{1}{2}I_s \leq c(n,p)\frac{1}{|Q_0|}\left(\int_{Q_0}|u|\right)^p + c(n,p)\int_{Q_0}|u^\sharp|^p$$

Since this is true for all s, we have

$$\int_{Q_0}|u|^p \leq c(n,p)\left\{\frac{1}{|Q_0|}\left(\int_{Q_0}|u|\right)^p + \int_{Q_0}|u^\sharp|^p\right\}$$

The stated result now follows with an application of Jensen's inequality. □

5.3 Stampacchia interpolation theorem

As a consequence, we can prove the Stampacchia interpolation theorem, which allows us to interpolate between L^p and BMO.

Theorem 66. *Let $1 \le p < \infty$ and let T be a linear operator of strong type (p,p) and bounded from L^∞ into BMO, i.e.,*

$$\|Tu\|_{L^p(Q_0)} \le C_1 \|u\|_{L^p(Q_0)} \quad \forall\, u \in L^p(Q_0)$$

and

$$[Tu]_{BMO(Q_0)} \le C_2 \|u\|_{L^\infty(Q_0)} \quad \forall\, u \in L^\infty(Q_0)$$

Then, T is of strong type (r,r) for all $r \in (p,\infty)$

Proof. Consider the map $\mathcal{T}$ defined by

$$\mathcal{T}(u) := (Tu)^\sharp$$

We have that, $\mathcal{T}$ is sublinear and
1. is of type (p,p) if $p > 1$ since

$$\begin{aligned}\|\mathcal{T}u\|_{L^p(Q_0)} &\le c(n)\|M(Tu)\|_{L^pQ_0}\\ &\le c(n,p)\|Tu\|_{L^p(Q_0)}\\ &\le c(n,p)c_1\|u\|_{L^p(Q_0)}\end{aligned}$$

2. is of weak type $(1,1)$ since

$$\begin{aligned}\left|\left\{(Tu)^\sharp > t\right\}\right| &\le |\{M(Tu) > t/c(n)\}|\\ &\le C(n)\frac{1}{t}\|Tu\|_{L^1(Q_0)}\\ &\le C(n)c_1\frac{1}{t}\|u\|_{L^p(Q_0)}\end{aligned}$$

3. is of strong type (∞,∞) since

$$\|\mathcal{T}(u)\|_{L^\infty(Q_0)} \le 2^n[Tu]_{BMO(Q_0)} \le 2^n c_2\|u\|_{L^\infty(Q_0)}$$

So by the Marcinkiewicz interpolation theorem, we have $\mathcal{T}$ is of strong type (r,r) for all $r \in (q,\infty)$. So, we have for all $r \in (q,\infty)$,

$$\|(Tu)^\sharp\|_{L^r(Q_0)} \le c\|u\|_{L^r(Q_0)}$$

If $p = 1$, we already have $\|Tu\|_{L^1(Q_0)} \le c_1\|u\|_{L^1(Q_0)}$. So by Hölder's we will have,

$$\|Tu\|_{L^1(Q_0)} \le c_3\|u\|_{L^r(Q_0)}$$

If $p > 1$, by Hölder's inequality and the strong (p,p) estimate, we have

$$\|Tu\|_{L^1(Q_0)} \le c_4\|Tu\|_{L^p(Q_0)} \le c_5\|u\|_{L^p(Q_0)} \le c_6\|u\|_{L^r(Q_0)}$$

Now, using Fefferman-Stein, we have for all $r \in (p,\infty)$

$$\begin{aligned}\|Tu\|_{L^r(Q_0)} &\le c(n,p)\left\{\|(Tu)^\sharp\|_{L^r(Q_0)} + |Q_0|^{1/r}\frac{1}{|Q_0|}\int_{Q_0}|Tu|\right\}\\ &\le c(n,p)\left\{c\|u\|_{L^r(Q_0)} + |Q_0|^{(1/r)-1}\|Tu\|_{L^1(Q_0)}\right\}\\ &\le c(n,p)\left\{c\|u\|_{L^r(Q_0)} + |Q_0|^{(1/r)-1}c_6\|u\|_{L^r(Q_0)}\right\}\\ &\le c_7\|u\|_{L^r(Q_0)}\end{aligned}$$

□

6 L^p estimates using Campanato method

6.1 Global L^p estimates for constant coefficients for $p \geq 2$

We now prove the L^p estimates using the Campanato-Stampacchia method.

Consider the problem

$$\begin{cases} -\operatorname{div}(A\nabla u) = \operatorname{div} F & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Let $u \in W_0^{1,2}(\Omega)$ be the unique solution. Define the operator T by $F \mapsto \nabla u$. We have that T is strong type $(2,2)$ as, by the weak formulation with the test function as u itself, we have that

$$\Lambda \int_\Omega |\nabla u|^2 \leq \int_\Omega \langle A\nabla u, \nabla u\rangle = \int_\Omega \langle F, \nabla u\rangle \leq \frac{1}{\varepsilon}\int_\Omega |F|^2 + \varepsilon \int_\Omega |\nabla u|^2$$

Choosing small enough ε, we have that

$$\|\nabla u\|_{L^2(\Omega)} \leq C\|F\|_{L^2(\Omega)}$$

The Camapanto estimates (including boundary estimates) tells us that T maps maps continuously $L^\infty(\Omega)$ into BMO. Indeed, by the Campanato estimate (see Chapter 5 of [3]) if $F \in \mathcal{L}^{2,n}(\Omega;\mathbb{R}^n)$, we have

$$\|\nabla u\|_{\mathcal{L}^{2,n}(\Omega)} \leq c\left(\|\nabla u\|_{L^2(\Omega)} + [F]_{\mathcal{L}^{2,n}(\Omega)}\right)$$

where we can use $\|\nabla u\|_{L^2(\Omega)} \leq C\|F\|_{L^2(\Omega)}$ to get

$$\|\nabla u\|_{\mathcal{L}^{2,n}(\Omega)} \leq c(\|F\|_{\mathcal{L}^{2,n}(\Omega)})$$

So noting that

$$[F]_{\mathcal{L}^{2,n}(\Omega)} \leq c\|F\|_{L^\infty(\Omega)}$$

and using the fact that $\mathcal{L}^{2,n}$ is isomorphic to BMO with equivalent seminorms, we have

$$[\nabla u]_{BMO(\Omega)} \leq c\|F\|_{L^\infty(\Omega)}$$

Stampacchia's interpolation theorem immediately gives us that, T is of strong type (r,r) for all $r \in (2,\infty)$. So we have the estimate

$$\|\nabla u\|_{L^r(\Omega)} \leq c\|F\|_{L^r(\Omega)}$$

For the general problem

$$\begin{cases} -\operatorname{div}(A\nabla u) = \operatorname{div} F & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases}$$

Noting that $v := u - g \in W_0^{1,2}(\Omega)$ solves the homogeneous boundary value problem, we can write

$$\begin{cases} -\operatorname{div}(A\nabla v) = \operatorname{div}(F + A\nabla g) & \text{in } \Omega, \\ v = 0 & \text{on } \partial\Omega. \end{cases}$$

Thus, we have the following theorem

Theorem 67. *Let $u \in W^{1,2}(\Omega)$ solve*

$$\begin{cases} -\operatorname{div}(A\nabla u) = \operatorname{div} F & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases}$$

where A satisfies the uniform ellipticity condition, $F \in L^p(\Omega; \mathbb{R}^n)$ and $g \in W^{1,p}(\Omega)$ for some $2 \leq p < \infty$. Then $\nabla u \in L^p$ and we have the estimate,

$$\|\nabla u\|_{L^p(\Omega)} \leq c\left(\|\nabla g\|_{L^p(\Omega)} + \|F\|_{L^p(\Omega)}\right).$$

6.2 Global L^p estimates for constant coefficients for $1 < p < 2$

For the case $1 < p < 2$, we use a duality argument, along with uniqueness and approximation.

Theorem 68. *Let $F \in L^p(\Omega; \mathbb{R}^n)$ and $g \in W^{1,p}(\Omega)$ for some $1 < p < 2$.. Let A be a symmetric $n \times n$ matrix which is uniformly elliptic. Then there exists unique $u \in W^{1,p}(\Omega)$ which solves*

$$\begin{cases} -\operatorname{div}(A\nabla u) = \operatorname{div} F & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases}$$

Moreover, we have the estimate,

$$\|\nabla u\|_{L^p(\Omega)} \leq c\left(\|\nabla g\|_{L^p(\Omega)} + \|F\|_{L^p(\Omega)}\right).$$

Proof. Assume $g = 0$, as we can reduce to this case as before. Let $\{F_\varepsilon\}_{\varepsilon > 0} \subset C_c^\infty(\Omega; \mathbb{R}^n)$ be a sequence such that

$$F_\varepsilon \to F \qquad \text{strongly in } L^p(\Omega).$$

Since $F_\varepsilon \in L^2(\Omega; \mathbb{R}^n)$ for every $\varepsilon > 0$, by Lax-Milgram or variational method, we can find $u_\varepsilon \in W_0^{1,2}(\Omega)$ solving

$$\begin{cases} -\operatorname{div}(A\nabla u_\varepsilon) = \operatorname{div} F_\varepsilon & \text{in } \Omega, \\ u_\varepsilon = 0 & \text{on } \partial\Omega. \end{cases}$$

Note that this implies

$$\int_\Omega \langle A\nabla u_\varepsilon, \nabla\varphi\rangle = -\int_\Omega \langle F_\varepsilon, \nabla\varphi\rangle \qquad \text{for any } \varphi \in C_c^\infty(\Omega).$$

Consequently, by density of $C_c^\infty(\Omega)$ in $W_0^{1,p'}(\Omega)$, we have

$$\int_\Omega \langle A\nabla u_\varepsilon, \nabla\psi\rangle = -\int_\Omega \langle F_\varepsilon, \nabla\psi\rangle \qquad \text{for any } \psi \in W_0^{1,p'}(\Omega). \tag{16}$$

Now by the dual characterization of norm and the density of $C_c^\infty(\Omega;\mathbb{R}^n)$ in $L^{p'}(\Omega;\mathbb{R}^n)$, we have

$$\|\nabla u_\varepsilon\|_{L^p(\Omega)} = \sup_{\substack{\zeta\in C_c^\infty(\Omega;\mathbb{R}^n),\\ \|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)}\leq 1}} \left|\int_\Omega \langle \nabla u_\varepsilon, \zeta\rangle\right| = \sup_{\substack{\zeta\in C_c^\infty(\Omega;\mathbb{R}^n),\\ \|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)}\leq 1}} \left|\int_\Omega \langle u_\varepsilon, \operatorname{div}\zeta\rangle\right|.$$

Now, since $p' > 2$, we can find $\psi \in W_0^{1,p'}(\Omega)$ solving

$$\begin{cases} -\operatorname{div}(A^\intercal\nabla\psi) = \operatorname{div}\zeta & \text{in } \Omega, \\ \psi = 0 & \text{on } \partial\Omega. \end{cases}$$

By Theorem 67, we have the estimate

$$\|\nabla\psi\|_{L^{p'}(\Omega;;\mathbb{R}^n)} \leq c\|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)}. \tag{17}$$

Now we have,

$$\begin{aligned}
\sup_{\substack{\zeta\in C_c^\infty(\Omega;\mathbb{R}^n),\\ \|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)}\leq 1}} \left|\int_\Omega \langle u_\varepsilon, \operatorname{div}\zeta\rangle\right| &= \sup_{\substack{\zeta\in C_c^\infty(\Omega;\mathbb{R}^n),\\ \|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)}\leq 1}} \left|\int_\Omega \langle u_\varepsilon, \operatorname{div}(A^\intercal\nabla\psi)\rangle\right| \\
&= \sup_{\substack{\zeta\in C_c^\infty(\Omega;\mathbb{R}^n),\\ \|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)}\leq 1}} \left|\int_\Omega \langle \nabla u_\varepsilon, A^\intercal\nabla\psi\rangle\right| \\
&= \sup_{\substack{\zeta\in C_c^\infty(\Omega;\mathbb{R}^n),\\ \|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)}\leq 1}} \left|\int_\Omega \langle A\nabla u_\varepsilon, \nabla\psi\rangle\right| \\
&\overset{(16)}{=} \sup_{\substack{\zeta\in C_c^\infty(\Omega;\mathbb{R}^n),\\ \|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)}\leq 1}} \left|\int_\Omega \langle F_\varepsilon, \nabla\psi\rangle\right| \\
&\leq \sup_{\substack{\zeta\in C_c^\infty(\Omega;\mathbb{R}^n),\\ \|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)}\leq 1}} \|F_\varepsilon\|_{L^p(\Omega;\mathbb{R}^n)}\|\nabla\psi\|_{L^{p'}(\Omega;\mathbb{R}^n)} \\
&\overset{(17)}{\leq} \sup_{\substack{\zeta\in C_c^\infty(\Omega;\mathbb{R}^n),\\ \|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)}\leq 1}} \|F_\varepsilon\|_{L^p(\Omega;\mathbb{R}^n)}\|\zeta\|_{L^{p'}(\Omega;\mathbb{R}^n)} \\
&\leq c\|F_\varepsilon\|_{L^p(\Omega;\mathbb{R}^n)}.
\end{aligned}$$

Hence, by Poincaré inequality, we have

$$\|\nabla u_\varepsilon\|_{W^{1,p}(\Omega)} \leq C\,\|\nabla u_\varepsilon\|_{L^p(\Omega;\mathbb{R}^n)} \leq C\|F_\varepsilon\|_{L^p(\Omega;\mathbb{R}^n)},$$

where the constant in independent of $\varepsilon > 0$. Now, by the strong convergence in L^p, $\{F_\varepsilon\}_{\varepsilon>0}$ is uniformly bounded in $L^p(\Omega;\mathbb{R}^n)$ and thus, $\{u_\varepsilon\}_{\varepsilon>0}$ is uniformly bounded in $W_0^{1,p}(\Omega)$. Hence, up to the extraction of a subsequence which we do not relabel, we have

$$u_\varepsilon \to u \qquad\qquad \text{weakly in } W^{1,p}(\Omega)$$

for some $u \in W_0^{1,p}(\Omega)$. It is now easy to check that $u \in W_0^{1,p}(\Omega)$ solves

$$\begin{cases} -\operatorname{div}(A\nabla u) = \operatorname{div} F & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

in the sense of distributions. Uniqueness follows from Lemma 69. This completes the proof. ☐

6.3 Uniqueness of solution in the nonvariational setting

In the last result, we have used the following uniqueness result, which establishes uniqueness of solutions to the Dirichlet problem in cases where the solution is not expected to have 'finite energy', i.e. solutions do not belong to $W^{1,2}(\Omega)$.

Lemma 69. *Let $F \in L^p(\Omega;\mathbb{R}^n)$ and $g \in W^{1,p}(\Omega)$ for some $1 < p < 2$. Let A be a symmetric $n \times n$ matrix which is uniformly elliptic. Then there exists **at most one** $u \in W^{1,p}(\Omega)$ which solves*

$$\begin{cases} -\operatorname{div}(A\nabla u) = \operatorname{div} F & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases}$$

Proof. By linearity of the equation, we only need to show that if $v \in W_0^{1,p}(\Omega)$ solves

$$\begin{cases} -\operatorname{div}(A\nabla v) = 0 & \text{in } \Omega, \\ v = 0 & \text{on } \partial\Omega, \end{cases} \tag{18}$$

then $v = 0$. This would be a semi-trivial integration by parts argument if $p \geq 2$. In our case, however, it is somewhat more delicate. Since v solves (18), we have

$$\int_\Omega \langle A\nabla v, \nabla\varphi\rangle = 0 \qquad \text{for any } \varphi \in C_c^\infty(\Omega).$$

Consequently, by density of $C_c^\infty(\Omega)$ in $W_0^{1,p'}(\Omega)$, we have

$$\int_\Omega \langle A\nabla v, \nabla\psi\rangle = 0 \qquad \text{for any } \psi \in W_0^{1,p'}(\Omega). \tag{19}$$

Now set

$$G := |\nabla v|^{p-2}\nabla v.$$

Since $v \in W_0^{1,p}(\Omega)$, we have $G \in L^{p'}(\Omega;\mathbb{R}^n)$. Since $p' > 2$, by Lax-Milgram or variational arguments, we can find $\psi \in W_0^{1,p'}(\Omega)$ which solves

$$\begin{cases} -\operatorname{div}(A^{\intercal}\nabla\psi) = \operatorname{div} G & \text{in } \Omega, \\ \psi = 0 & \text{on } \partial\Omega. \end{cases}$$

Using this, we deduce

$$\begin{aligned} \int_\Omega |\nabla v|^p &= \int_\Omega \langle \nabla v, G\rangle \\ &= -\int_\Omega \langle v, \operatorname{div} G\rangle = \int_\Omega \langle v, \operatorname{div}(A^{\intercal}\nabla\psi)\rangle = -\int_\Omega \langle A\nabla v, \nabla\psi\rangle \overset{(19)}{=} 0. \end{aligned}$$

This proves $v = 0$ and completes the proof. □

6.4 Interior L^p estimates for continuous coefficients

We now establish the interior regularity result in the case where the coefficients of A are continuous.

Theorem 70. *Let $\Omega \subset \mathbb{R}^n$ is a bounded open set. $A \in C\left(\overline{\Omega};\mathrm{Symm}_{n\times n}\right)$ is a symmetric $n \times n$ matrix field which is uniformly continuous and uniformly elliptic, i.e. there exists a constant $\lambda > 0$ such that*

$$\langle A(x)\,\xi, \xi\rangle \geq \lambda\,|\xi|^2 \qquad \text{for every } \xi \in \mathbb{R}^n \text{ and every } x \in \overline{\Omega}.$$

Let $1 < p < \infty$ and let $F \in L^p(\Omega;\mathbb{R}^n)$. If $u \in W^{1,\frac{np}{n+p}}(\Omega)$ is a distributional solution of

$$-\operatorname{div}(A(x)\nabla u) = \operatorname{div} F \qquad \text{in } \Omega,$$

then $u \in W^{1,p}_{loc}(\Omega)$ and for any $\Omega_1 \subset\subset \Omega$, there exists a constant

$$C = C\left(n, p, \Omega_1, \Omega, \lambda, \omega_A, \|A\|_{L^\infty(\Omega)}\right) > 0$$

such that we have the estimate

$$\|u\|_{W^{1,p}(\Omega_1)} \leq C\left(\|F\|_{L^p(\Omega)} + \|u\|_{W^{1,\frac{np}{n+p}}(\Omega)}\right).$$

Proof. By a covering argument, it suffices to prove the estimate for balls. For a point x_0, take a ball $B_R(x_0) \subset\subset \Omega$ and a function $\eta \in C_c^\infty(B_R(x_0))$ with $\eta = 1$ on $B_{R/2}(x_0)$. We compute, for any $\psi \in C_c^\infty(\Omega)$,

$$\begin{aligned} \int_\Omega \langle A(x)\nabla(\eta u), \nabla\psi\rangle &= \int_\Omega \langle A(x)\eta\nabla u, \nabla\psi\rangle + \langle A(x)u\nabla\eta, \nabla\psi\rangle \\ &= \int_\Omega \langle \eta F, \nabla\psi\rangle + \langle A(x)u\nabla\eta, \nabla\psi\rangle. \end{aligned}$$

Now, we perform the freezing trick to get,

$$\int_\Omega \langle (A(x) - A(x_0) + A(x_0)\nabla(\eta u), \nabla\psi\rangle = \int_\Omega \langle \eta F, \nabla\psi\rangle + \langle A(x)u\nabla\eta, \nabla\psi\rangle$$

which gives

$$\int_\Omega \langle A(x_0)\nabla(\eta u), \nabla\psi\rangle$$
$$= \int_\Omega \langle A(x_0) - A(x)\nabla(\eta u), \nabla\psi\rangle + \langle \eta F, \nabla\psi\rangle + \langle A(x)u\nabla\eta, \nabla\psi\rangle$$

for any $\psi \in C_c^\infty(\Omega)$. Thus, $v := \eta u$ solves the PDE

$$\begin{aligned} -\operatorname{div}\left(A\left(x_0\right)\nabla v\right) &= \operatorname{div}\left((A(x_0) - A(x))\nabla v\right) \\ &\quad - \operatorname{div}(\eta F) - \operatorname{div}(A(x)u\nabla\eta) \qquad \text{in } B_R(x_0). \end{aligned} \tag{20}$$

Note that if $u \in W^{1,\frac{np}{n+p}}(\Omega)$, then by the Sobolev inequality, we have $u \in L^p(\Omega)$ and we have the estimate

$$\|u\|_{L^p(\Omega)} \leq C\,\|u\|_{W^{1,\frac{np}{n+p}}(\Omega)}.$$

Thus, by either Lax-Milgram and Theorem 67 or Theorem 68, we can find $w \in W_0^{1,p}(B_R(x_0))$ such that

$$-\Delta w = -\operatorname{div}(\underbrace{\eta F}_{\in L^p}) - \operatorname{div}(\underbrace{A(x)u\nabla\eta}_{\in L^p}) \qquad \text{in } B_R(x_0).$$

By the L^p estimate for constant coefficient operators, we deduce $\nabla w \in L^p(B_R(x_0);\mathbb{R}^n)$. Now, for a $\zeta \in W_0^{1,p}(B_R(x_0))$, let $\theta \in W_0^{1,p}(B_R(x_0))$ be the weak solution of

$$-\operatorname{div}(A(x_0)\nabla\theta) = -\operatorname{div}\left((A(x_0) - A(x))\nabla v + \nabla w\right)$$

Again, by the L^p estimates for constant coefficient operators, we have the estimate

$$\|\nabla\theta\|_{L^p(B_R(x_0))} \leq c\|A(x_0) - A(x)\|_{L^\infty(B_R(x_0))}\|\nabla v\|_{L^p(B_R(x_0))} + c\|\nabla w\|_{L^p(B_R(x_0))}$$

Now, consider the map $T : W_0^{1,p}(B_R(x_0)) \to W_0^{1,p}(B_R(x_0))$ given by

$$Tv = \theta$$

Note that by Poincaré, L^p norm of the gradient is an equivalent norm on $W_0^{1,p}$. Now, choosing $R > 0$ small enough, we can ensure that this map is a contraction. We have

$$\|\theta_1 - \theta_2\|_{L^p(B_R(x_0))} \leq c\omega(R)\|\nabla v_1 - \nabla v_2\|_{L^p(B_R(x_0))}$$

So by Banach's fixed point theorem, we have a unique fixed point for T, say ψ. Now, $\psi \in W_0^{1,p}(B_R(x_0))$ satisfies

$$-\operatorname{div}(A(x_0)\nabla\psi) = -\operatorname{div}\left((A(x_0) - A(x))\nabla\psi + \nabla w\right) \qquad \text{in } B_R(x_0).$$

So by uniqueness of $W_0^{1,p}\left(B_R(x_0)\right)$ solutions of the PDE (20), given by Lemma 69, ψ must coincide with ηu on $B_R(x_0)$. Now, since $\eta = 1$ on $B_{R/2}(x_0)$,, we have $u = \psi$ on $B_{R/2}(x_0)$ and consequently, $u \in W^{1,p}\left(B_{R/2}(x_0)\right)$. This completes the proof. □

As before, by bootstrapping, this implies the following result.

Theorem 71. *Let $\Omega \subset \mathbb{R}^n$ is a bounded open set. $A \in C\left(\overline{\Omega}; \mathrm{Symm}_{n\times n}\right)$ is a symmetric $n \times n$ matrix field which is uniformly continuous and uniformly elliptic, i.e. there exists a constant $\lambda > 0$ such that*

$$\langle A(x)\xi, \xi\rangle \geq \lambda |\xi|^2 \qquad \text{for every } \xi \in \mathbb{R}^n \text{ and every } x \in \overline{\Omega}.$$

Let $1 < q < p < \infty$ and let $F \in L^p(\Omega; \mathbb{R}^n)$. If $u \in W^{1,q}(\Omega)$ is a distributional solution of

$$-\operatorname{div}\left(A(x)\nabla u\right) = \operatorname{div} F \qquad\qquad \text{in } \Omega,$$

then $u \in W^{1,p}_{loc}(\Omega)$ and for any $\Omega_1 \subset\subset \Omega$, there exists a constant

$$C = C\left(n, p, \Omega_1, \Omega, \lambda, \omega_A, \|A\|_{L^\infty(\Omega)}\right) > 0$$

such that we have the estimate

$$\|u\|_{W^{1,p}(\Omega_1)} \leq C\left(\|F\|_{L^p(\Omega)} + \|u\|_{W^{1,q}(\Omega)}\right).$$

Appendix A Recap: Basic properties of Fourier transform

A.1 Fourier transform in $L^1(\mathbb{R}^n)$

Definition 72 (Fourier transform in L^1)**.** *Let $u \in L^1(\mathbb{R}^n)$. We define **the Fourier transform of** u**,** denoted $\hat{u}$, as*

$$\hat{u}(\xi) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-i\langle \xi, x\rangle} u(x)\ \mathrm{d}x \qquad \text{for all } \xi \in \mathbb{R}^n.$$

Definition 73 (Inverse Fourier transform in L^1)**.** *Let $u \in L^1(\mathbb{R}^n)$. We define **the inverse Fourier transform of** u**,** denoted $\check{u}$, as*

$$\check{u}(x) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{i\langle \xi, x\rangle} u(\xi)\ \mathrm{d}\xi \qquad \text{for all } x \in \mathbb{R}^n.$$

It is quite easy to see that the Fourier transform actually is finite for a.e. $\xi \in \mathbb{R}^n$. But we have something more.

Proposition 74 (Plancherel-Parseval identity)**.** *Let $u \in L^1(\mathbb{R}^n)$. Then $\hat{u} \in L^\infty(\mathbb{R}^n)$ and for all $u, v \in L^1(\mathbb{R}^n)$, we have the identity, sometimes called the* ***Plancherel-Parseval identity****,*

$$\int_{\mathbb{R}^n} u\hat{v} = \int_{\mathbb{R}^n} \hat{u}v.$$

Definition 75 (Gaussian)**.** *For any point $p \in \mathbb{R}^n$ and any two real numbers $a, \sigma > 0$, the* ***Gaussian*** *with* ***mean*** *p and* ***variance*** *σ^2 is a real-valued function $f : \mathbb{R}^n \to \mathbb{R}$ defined by*

$$f(x) = ae^{-\frac{|x-p|^2}{2\sigma^2}} \qquad \text{for all } x \in \mathbb{R}^n.$$

A Gaussian is called ***normalized Gaussian*** *if we have*

$$a = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}}.$$

Proposition 76 (FT of Gaussian is another Gaussian)**.** *For any $\varepsilon > 0$, we have*

$$\int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle - \varepsilon|x|^2}\, \mathrm{d}x = \left(\frac{\pi}{\varepsilon}\right)^{\frac{n}{2}} e^{-\frac{|\xi|^2}{4\varepsilon}}.$$

A.2 Fourier transform in $L^2(\mathbb{R}^n)$

Theorem 77 (Plancherel theorem)**.** *Let $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Then $\hat{u}, \check{u} \in L^2(\mathbb{R}^n)$ and we have*

$$\|u\|_{L^2(\mathbb{R}^n)} = \|\hat{u}\|_{L^2(\mathbb{R}^n)} = \|\check{u}\|_{L^2(\mathbb{R}^n)}.$$

We finish off our discussion of Fourier transform in L^1 and L^2 by a simple result, which records how Fourier transform behaves with respect to affine change of variables.

Theorem 78. *Let $u \in L^1(\mathbb{R}^n)$. Then the following holds.*

(i) ***Translation:*** *For any $a \in \mathbb{R}^n$, set $\tau_a u := u(x+a)$. Then*

$$(\tau_a u)\hat{}\,(\xi) = e^{i\langle \xi, a\rangle}\hat{u}(\xi) \qquad \text{for } \xi \in \mathbb{R}^n.$$

(ii) ***Change of Variable:*** *Let $T \in \mathbb{GL}(n, \mathbb{R})$. Then*

$$(u \circ T)\hat{} = |\mathrm{Det}\, T|^{-1}\, \hat{u} \circ \left(T^{-1}\right)^T.$$

In particular, we have

(a) ***Dilation:*** *Let $\lambda \neq 0$ be a real number and let $u_\lambda(x) := u(\lambda x)$. Then*

$$\hat{u}_\lambda(\xi) = \frac{1}{|\lambda|^n}\hat{u}\left(\frac{1}{\lambda}\xi\right) \qquad \text{for } \xi \in \mathbb{R}^n.$$

(b) ***Orthogonal transformations:*** *Let $R \in \mathbb{O}(n, \mathbb{R})$. Then*

$$(u \circ R)\hat{} = \hat{u} \circ R.$$

A.3 Schwartz space $\mathcal{S}(\mathbb{R}^n)$

Definition 79 (Schwartz space)**.** *The* ***space of rapidly decaying functions*** *on $\mathbb{R}^n$ or the* ***Schwartz space*** *on $\mathbb{R}^n$, denoted by $\mathcal{S}(\mathbb{R}^n)$, is defined as*

$$\mathcal{S}(\mathbb{R}^n) := \left\{ u \in C^\infty(\mathbb{R}^n) : \sup_{x\in\mathbb{R}^n} \left| x^\beta D^\alpha u(x) \right| < \infty \text{ for all multiindices } \alpha, \beta \right\}$$

It is easy to see that we have

$$C_c^\infty(\mathbb{R}^n) \subsetneq \mathcal{S}(\mathbb{R}^n) \subsetneq L^p(\mathbb{R}^n) \cap C(\mathbb{R}^n),$$

for any $1 \leq p \leq \infty$ The following result is an important feature of the Schwartz space.

Theorem 80. *Let $u \in \mathcal{S}(\mathbb{R}^n)$ and let $P(x)$ be a polynomial in $x \in \mathbb{R}^n$.*

(i) Define the function $g : \mathbb{R}^n \to \mathbb{R}$ by

$$g(x) = P(x)\, u(x) \qquad \text{for all } x \in \mathbb{R}^n.$$

Then $g \in \mathcal{S}(\mathbb{R}^n)$. Moreover, for every fixed polynomial $P(x)$, the map

$$u \mapsto P(x)\, u$$

is a linear continuous map from $\mathcal{S}(\mathbb{R}^n)$ to itself.

(ii) For each multiindex α, we have $D^\alpha u \in \mathcal{S}(\mathbb{R}^n)$. Moreover, for every fixed multiindex α, the map

$$u \mapsto D^\alpha u$$

is a linear continuous map from $\mathcal{S}(\mathbb{R}^n)$ to itself.

Theorem 81. *Let $u \in \mathcal{S}(\mathbb{R}^n)$. Then*

(i) ***Derivatives to multiplication by ξ:*** *For each multiindex α, we have*

$$\left(D^\alpha u\right)^{\wedge}(\xi) = (i\xi)^\alpha\, \hat{u}(\xi) \qquad \text{for every } \xi \in \mathbb{R}^n.$$

(ii) ***Multiplication by x to derivatives:*** *For each multiindex α, we have*

$$D^\alpha \hat{u} = \left[(-ix)^\alpha u\right]^{\wedge}.$$

(iii) $\hat{u}, \check{u} \in \mathcal{S}(\mathbb{R}^n)$. Moreover, the maps $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ given by

$$u \mapsto \hat{u}$$

and $\mathcal{F}^{-1} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ given by

$$u \mapsto \check{u}$$

are both linear and continuous as maps from $\mathcal{S}(\mathbb{R}^n)$ to itself.

Theorem 82 (Fourier inversion formula)**.** *Let $u \in \mathcal{S}(\mathbb{R}^n)$. Then*

$$u = (\hat{u})^{\vee} = (\check{u})^{\wedge}.$$

Theorem 83. *Let $u, v \in \mathcal{S}(\mathbb{R}^n)$. Then*

$$(u * v)^{\wedge} = (2\pi)^{\frac{n}{2}}\, \hat{u}\hat{v} \qquad \text{and} \qquad (uv)^{\wedge} = (2\pi)^{-\frac{n}{2}}\, \hat{u} * \hat{v}.$$

A.4 Tempered distributions

Definition 84. *A* ***tempered distribution*** *on* $\mathbb{R}^n$ *is a continuous linear functional on* $\mathcal{S}(\mathbb{R}^n)$.

Remark 85. *The definition says that a linear map* $T : \mathcal{S}(\mathbb{R}^n) \to \mathbb{R}$ *is called a tempered distribution if* T *is continuous. But since* $\mathcal{S}(\mathbb{R}^n)$ *is a metric space,* T *is continuous if and only if it is* ***sequentially continuous****, i.e. for every sequence* $\phi_s \to \phi$ *in* $\mathcal{S}(\mathbb{R}^n)$, *we must have*

$$T(\phi_s) \to T(\phi).$$

Definition 86. *Let* $T \in \mathcal{S}'(\mathbb{R}^n)$. *Then the Fourier transform of* T, *denoted* $\hat{T}$, *is another tempered distribution which is defined by the action*

$$\hat{T}(\phi) = T\left(\hat{\phi}\right) \qquad \text{for every } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Definition 87. *Let* $T \in \mathcal{S}'(\mathbb{R}^n)$. *Then for any multiindex* α, *the* ***distributional derivative*** *of* T, *denoted* $D^\alpha T$, *is another tempered distribution which is defined by the action*

$$D^\alpha T(\phi) = (-1)^{|\alpha|} T(D^\alpha \phi) \qquad \text{for every } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Remark 88. *The reason for the somewhat strange sign is the integration by parts formula.*

Definition 89. *Let* $T \in \mathcal{S}'(\mathbb{R}^n)$ *and let* $P(x)$ *be a polynomial in* $\mathbb{R}^n$. *Then the multiplication of* T *by* P, *is another tempered distribution which is denoted by* $P(x)T$ *and is defined by the action*

$$P(x)T(\phi) = T(P(x)\phi) \qquad \text{for every } \phi \in \mathcal{S}(\mathbb{R}^n).$$

References

[1] Dacorogna, B. *Introduction to the calculus of variations*, third ed. Imperial College Press, London, 2015.

[2] Evans, L. C. *Partial differential equations*, vol. 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[3] Giaquinta, M., and Martinazzi, L. *An introduction to the regularity theory for elliptic systems, harmonic maps and minimal graphs*, second ed., vol. 11 of *Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]*. Edizioni della Normale, Pisa, 2012.