Lecture notes, Singular Integrals

Joaquim Bruna, UAB

April 28, 2016

Chapter 1

The Hardy-Littlewood maximal function

In this chapter we will deal with the Hardy-Littlewood maximal function in a rather general context and show its role in some important aspects of Harmonic Analysis. We begin with some definitions.

1.1 Some definitions

We will denote by $(X, \mu), (Y, \nu), \ldots$ measure spaces, although most of the time $X = \mathbf{R^d}$ equipped with Lebesgue measure. $L^p(X)$ denotes the usual Lebesgue spaces, and $L^0(X)$ the space of measurable functions equipped with the notion of convergence in measure.

We denote by λ_f the distribution function of f,

$$\lambda_f(s) = \mu\{|f| > s\}.$$

If φ is increasing then

$$\int_X \varphi(|f|)d\mu = \int_X \int_0^{|f(x)|} \varphi'(s)dsd\mu(x) = \int_0^{+\infty} \varphi'(s)\lambda_f(s)ds.$$

In particular, $\|f\|_p^p = p\int_0^\infty s^{p-1}\lambda_f(s)\,ds$.

The Tchebichev's inequality: if if $f \in L^p(X), p < +\infty$, and $A_s = \{x : |f(x)| > s\}$, then

$$s^p\mu(A_s) \leq \int_{A_s} |f^p|d\mu \leq \int_X |f^p|d\mu = \|f\|_p^p.$$

The functions f such that $\mu(A_s) = O(s^{-p}), p < \infty$, are said to be *weak* L^p-functions, a space that we denote by $L^{p,\infty}(X)$. A typical example of a weak L^1 function which is not in L^1 is $|x|^{-d}$ is $\mathbf{R^d}$.

If a mapping $T : L^p(X) \to L^0(Y)$, $1 \leq p, q \leq +\infty$, satisfies

$$\|Tf\|_q \leq C\|f\|_p$$

we say that it is of *strong type* (p,q). If T is linear, this amounts to T being continuous from $L^p(X)$ to $L^q(Y)$. If T is of strong type $(p,q), q < +\infty$ then

$$s^q \nu\{|Tf| > s\} \leq \|Tf\|_q^q \leq (C\|f\|_p)^q.$$

A mapping such that

$$\nu\{|Tf| > s\} \leq (\frac{C\|f\|_p}{s})^q.$$

and thus mapping L^p to $L^{q,\infty}$ is said to be of *weak type* $(p,q), q < +\infty$. By definition, when $q = +\infty$ weak type (p,∞) will de identified with strong type, $\|Tf\|_\infty \leq C\|f\|_p$

1.2 How maximal functions arise

Suppose we have a family of linear operators $T_t : L^p(X) \to L^0(Y)$ depending on $t > 0$ and we are interested in the statement "the limit $\lim_{t\to 0} T_t f(x)$ exists a.e. on X"; or that $X = Y$ and we are interested in the statement " $\lim_t T_t f(x) = f(x)$ for a.e. x". Often in analysis we know that this is the case if f belongs to a certain class of functions E, and often this class E is dense in L^p. For instance, consider the model case in which $X = \mathbf{R^d}$ and

$$T_t f(x) = \frac{1}{|B(x,t)|} \int_{B(x,t)} f(y)dy,$$

is the average of f in the ball $B(x,t)$; when f is say continuous with compact support (a dense subspace of the $L^p(X)$, then trivially $T_t f(x) \to f(x)$ as $t \to 0$.

Coming back to the general situation, assume $Y = X$ and that we are interested in the statement $\lim_t T_t f(x) = f(x) a.e. x$. We consider the maximal operator

$$T^* f(x) = \sup_t |T_t f(x)| \leq +\infty.$$

Note that T^* is not linear but sub-additive: $T^*(f+g) \leq T^* f + T^* g$.

Theorem 1. If T^* is of weak type (p,q), then the class

$$E = \{f : \lim_t T_t f(x) = f(x) a.e. x\}$$

is closed in $L^p(X)$.

Proof. Assume that $f_n \in E$ approach f. To show that $f \in E$, we will show that for fixed $s > 0$, the set $\{x : \limsup_t |T_t f(x) - f(x)| > s\}$ has zero measure. Now, from

$$|T_t f(x) - f(x)| \leq |T_t f(x) - T_t f_n(x)| + |T_t f_n(x) - f_n(x)| + |f_n(x) - f(x)|,$$

and the fact that $f_n \in E$, we see that

$$\limsup_t |T_t f(x) - f(x)| \leq \limsup_t |T_t f(x) - T_t f_n(x)| + |f_n(x) - f(x)| \leq$$
$$\leq T^*(f - f_n)(x) + |f_n(x) - f(x)|,$$

hence

$$\{x : \limsup_t |T_t f(x) - f(x)| > s\} \subset \{x : T^*(f - f_n)(x) > \frac{s}{2}\} \cup \{|f_n(x) - f(x)| > \frac{s}{2}\}$$

Then the hypothesis and Tchebychev's inequality imply

$$\mu\{x : \limsup_t |T_t f(x) - f(x)| > s\} \leq (\frac{C}{s}\|f - f_n\|_p)^q + (\frac{C}{s}\|f - f_n\|_p)^p,$$

and so making $n \to +\infty$ we are done. □

Note that we can deal in an analogous way with the statement "the limit $\lim_{t\to 0} T_t f(x)$ exists a.e. on X", for which we would work with $\limsup_t T_t f(x) - \liminf_t T_t f(x)$, bounded by $2T^* f$.

All this gives the basic recipe to prove a.e. convergence results: find a good dense class and prove a weak estimate for the maximal associated function. This is somehow similar to the situation in the uniform boundedness principle for families of operators defined in Banach spaces.

1.3 The general setting for the Hardy-Littlewood maximal function

The standard setting is $\mathbf{R^d}$ where we consider the means

$$T_t f(x) = \frac{1}{|B(x,t)|} \int_{B(x,t)} f(y)dy,$$

that make sense for $f \in L^1_{loc}(\mathbf{R^d})$, and ask about their a.e. limit.

Having this in mind we will consider however a more general situation of potential interest. We suppose that we are still in $X = \mathbf{R^d}$ and we are given the balls $B(x, \delta)$ in some way, satisfying

- They increase in δ.
- For fixed x, their intersection is x and their union is $\mathbf{R^d}$.

- There exists $c_1 > 1$ such that whenever $B(x,\delta) \cap B(y,\delta) \neq \emptyset$, then $B(y,\delta) \subset B(x, c_1\delta)$ (and so also $B(x,\delta) \subset B(y, c_1\delta)$, engulfing property). We denote by B^* the ball obtained from B dilating by c_1 its radious. In particular $x \in B(y,\delta)$ does not imply $y \in B(x,\delta)$, but it does imply $y \in B(x, c_1\delta)$.

and that the regular Borel measure satisfies

- The map $x \mapsto \mu(U \cap B(x,\delta))$ is continuous in x. This implies that the means of $f \in L^1_{loc}$ are continuous in x too.
- There exists $c_2 > 1$ such that $\mu(B(x, c_1\delta)) \leq c_2 \mu(B(x,\delta)$ (doubling condition)

Note that we do not require that the measures of balls just depend on the radious, they might depend on x too.

We then define the *centered Hardy-Littlewood maximal function* by

$$Mf(x) = \sup_\delta \frac{1}{\mu(B(x,\delta))} \int_{B(x,\delta)} |f(y)| d\mu(y). \tag{1.1}$$

The uncentered maximal function $\overline{M}f$ is defined using all balls containing x, so trivially $Mf \leq \overline{M}f$. But if $x \in B(y,\delta)$, then $B(y,\delta)$ is included in $B(x, c_1\delta)$, and $B(x,\delta)$ in $B(y, c_1\delta)$, so that all these balls have the same measure up to constants; it then follows that $\overline{M}f \leq CMf$ for some constant C.

The following are examples of this general setting, besides the standard Euclidean balls.

- Fix a star-shaped domain U with respect the origin in $\mathbf{R^d}$, for instance a convex set containing the origin, and put $B(x,\delta) = x + \delta U$, so that U is the "unit ball".
- A quasi-distance $\rho(x,y)$ satisfies the three basic properties:
  1. $\rho(x,y) = 0$ only if $x = y$.
  2. $\rho(x,y) \leq c\rho(y,x)$
  3. $\rho(x,y) \leq C(\rho(x,z) + \rho(z,y))$

  If we then set $B(x,\delta) = \{y : \rho(x,y) < \delta\}$ we get a system of balls with the above requirements. Conversely we can define a quasidistance from the balls by defining $\rho(x,y) = \inf\{\delta : y \in B(x,\delta)\}$.
- These quasi-distances may be non isotropic. For instance, fixed positive exponents $\alpha_1, \ldots, \alpha_d$ set $\rho(x,y) = \max |x_i - y_i|^{1/\alpha_i}$.

- An interesting case arises in connection with the following fact in Riemanian geometry. Suppose we are given a family of vector fields $D_1, \cdots, D_k, k < d$ in $\mathbf{R^d}$. If at each point these vectors, together with their commutators, span the whole of $\mathbf{R^d}$ it can be proved that for sufficiently close points x, y there exists a curve $\gamma : [0,1] \to \mathbf{R^d}$ joining these two points, $\gamma(0) = x, \gamma(1) = y$ and such that at each point $\gamma(t)$ the tangent vector lies in the span of $D_1, \cdots, D_k$. It makes sense then to define $\rho(x, y)$ as the infimum of the lengths of all such curves. This is a quasidistance. This situation appears in particular in function theory in several complex variables.

1.4 Estimates for the Hardy-Littlewood maximal function

Our purpose is to prove

Theorem 2. In the setting described above, with Mf defined as in (1.1), one has

1. For $f \in L^p(\mu), 1 \leq p \leq +\infty$, Mf is finite μ- almost everywhere.
2. For $f \in L^1(\mu)$, Mf satisfies a weak L^1-estimate
$$\mu\{Mf > s\} \leq \frac{C}{s} \|f\|_1$$
3. For $f \in L^p(\mu), 1 < p \leq +\infty$, Mf is also in L^p and $\|Mf\|_p \leq A_p \|f\|_p$.

Note that the first statement follow from the others, than the last statement is trivial for $p = +\infty$ and that in the standard case, euclidian balls and Lebesgue measure, then the first statement holds for all $f \in L^1_{loc}(\mathbf{R^d})$.

We need the following Vitali-type covering lemma.

Lemma 1. Assume that E is the union of a finite number of balls $B_1, \ldots, B_N$. Then we can extract a subcollection $\hat{B}_1, \cdots, \hat{B}_k$ of mutually disjoint balls such that $\mu(\cup_{j=1}^k \hat{B}_j) = \sum_j \mu(\hat{B}_j) \geq c\mu(E)$, the constant c depending just on c_1, c_2.

Proof. We first chose $\hat{B}_1$ the ball with biggest radious, then $\hat{B}_2$ is chosen as the biggest among the remaining not meeting $\hat{B}_1$, and so on till we have no more balls. If B_i has not been selected it must meet some of the $\hat{B}_j$. If $\hat{B}_j$ is the first one meeting B_i, since when choosing the j-th ball it has not been selected, B_i has radious not bigger than that of $\hat{B}_j$, whence $B_i \subset \hat{B}_j^*$. Thus $E \subset \cup_j \hat{B}_j^*$, therefore $\mu(E) \leq \sum_j \mu(\hat{B}_j^*) \leq C \sum_j \mu(\hat{B}_j)$. □

We give now the proof of the theorem. We work with the equivalent uncentered version $\overline{M}f$ and consider $A_s = \{\overline{M}f > s\}$, which is open. Every $x \in A_s$ is the center of some ball B_x such that

$$\frac{1}{\mu(B_x)} \int_{B_x} |f(y)| d\mu(y) > s.$$

A fixed compact set K of A_s will be covered by a finite number of these balls to which we apply the lemma, to get disjoint balls B_j such that

$$\mu(K) \leq C \sum_j \mu(B_j) \leq \frac{C}{s} \sum_j \int_{B_j} |f(y)| d\mu(y) = \frac{C}{s} \int_{\cup B_j} |f(y)| d\mu(y) \leq \frac{C}{s} \|f\|_1.$$

Since K is arbitrary this proves the second statement. Now assume that $f \in L^p(\mu), 1 < p$ and consider again A_s. Let $f_1(x)$ be equal to $f(x)$ if $|f(x)| > \frac{s}{2}$ and zero otherwise (so that $f_1 \in L^1(\mu)$), and put $f = f_1 + f_2$, with f_2 bounded by $\frac{s}{2}$. Then $\overline{M}f \leq \overline{M}f_1 + \overline{M}f_2 \leq \overline{M}f_1 + \frac{s}{2}$, whence $A_s \subset \{\overline{M}f_1 > \frac{s}{2}\}$. By the estimate we just proved for L^1,

$$\mu(A_s) \leq \frac{C}{s} \|f_1\|_1 = \frac{C}{s} \int_{|f| > \frac{s}{2}} |f(y)| d\mu(y).$$

Then, by Fubini's theorem

$$\|\overline{M}f\|_p^p = \int_0^{+\infty} p s^{p-1} \mu(A_s) ds \leq C \int_0^{+\infty} s^{p-2} \int_{|f(y)| > \frac{s}{2}} |f(y)| d\mu(y) =$$

$$= C \int_{\mathbf{R^d}} |f(y)| \int_0^{2|f(y)|} s^{p-2} ds = A_p \int_{\mathbf{R^d}} |f(y)|^p d\mu(y),$$

as claimed.

One may check that the constant A_p behaves like $\frac{1}{p-1}$ and explodes at 1. In fact, in the standard case, unless $f = 0$, Mf is never integrable. Indeed, consider a ball B where $|f|$ has a non-zero mass m; for big enough $|x|$ the ball $B(x, k|x|)$ will include B and so $Mf(x) \geq c|x|^{-d}$, which is not integrable at infinity.

However, if f satisfies an additional condition, Mf is integrable over sets of finite measure:

Proposition 1. If A has finite measure, then

$$\int_A Mf d\mu \leq C\mu(A) + \int_{\mathbf{R^d}} |f(x)| \log^+ |f(x)| d\mu(x).$$

This follows from the same proof as above, where now

$$\int_A |\overline{M}f| d\mu = \int_0^{+\infty} \mu\{x \in A, |\overline{M}f(x)| > s\}.$$

For $s < 1$ we bound simply by $\mu(A)$, while for $s > 1$ we use the weak inequality as before.

1.5 Some consequences

Of course the first consequence is Lebesgue's differentiation theorem

Theorem 3. For $f \in L^1_{loc}(\mathbf{R^d})$, one has

$$\lim_{\delta \to 0} \frac{1}{|B(x,\delta)|} \int_{B(x,\delta)} f(y)dy = f(x), a.e.x.$$

In fact,

$$\lim_{\delta \to 0} \frac{1}{|B(x,\delta)|} \int_{B(x,\delta)} |f(y) - f(x)|dy = 0, a.e.x.$$

We already proved the first statement; now this gives trivially that $|f| \leq Mf$ whence the maximal function associated to the second statement is bounded by $2Mf$; since this second statement is also trivially true for continuous functions, we are done. The points above are called the *Lebesgue points* of f.

In dimension $d = 1$, the Lebesgue theorem becomes the fundamental theorem of calculus: for $f \in L^1(\mathbf{R})$, the indefinite integral

$$F(x) = \int_0^x f(t)\,dt,$$

has a.e. a derivative equal to $f(x)$. As it is well known, the functions F of this form are exactly the absolutely continuous functions, meaning that for each $\varepsilon > 0$ there is $\delta > 0$ such that the variation $\sum_j |F(b_j) - F(a_j)|$ is less than ε for all consecutive intervals (a_j, b_j) with total length $\sum_j (b_j - a_j) < \delta$.

We now turn to another application of Mf in the standard setting, regarding *approximations of the identity.* Recall that an approximation of the identity is a family (φ_t) of the form $\varphi_t(x) = t^{-d}\varphi(x/t)$, with a fixed $\varphi \in L^1(\mathbf{R^d})$ with integral one. The term "approximation of the identity" comes from the fact that in a certain sense, one has that $\lim_t \varphi_t = \delta_0$, the unit for convolution. The precise statement is that

- If $f \in L^p(\mathbf{R^d}), \mathbf{1 \leq p} < +\infty$, then $\varphi_t * f \to f$ in $L^p(\mathbf{R^d})$
- If f is continuous at x_0 then $\varphi_t * f(x_0) \to f(x_0)$.

Approximate identities are related with summation methods for the Fourier integral and with classical PDE's as follows. For $f \in L^1(\mathbf{R^d})$ or $f \in L^2(\mathbf{R^d})$ the inverse Fourier transform

$$f(x) = \int_{\mathbf{R^d}} \hat{f}(\xi) e^{2\pi i x \cdot \xi}\, d\xi,$$

does not makes sense as an absolutely convergent integral and so alternative summation methods are proposed. A general scheme is as follows: we consider a continuous integrable function h such that $h(0) = 1$ and the mean

$$\int_{\mathbf{R^d}} \hat{f}(\xi) e^{2\pi i x\cdot\xi} h(t\xi) d\xi.$$

Now, Fubini's theorem implies that

$$\int_{\mathbf{R^d}} \hat{f}(\xi) g(\xi) d\xi = \int_{\mathbf{R^d}} f(y)\hat{g}(y) dy, f, g \in L^1(\mathbf{R^d}).$$

If $\hat{h} = \Phi$, the Fourier transform of $h(t\xi)$ is $t^{-d}\Phi(\frac{y}{t}) = \Phi_t(y)$, whence the Fourier transform of $e^{2\pi i x\cdot\xi}h(t\xi)$ is $\Phi_t(y-x)$ and so

$$\int_{\mathbf{R^d}} \hat{f}(\xi) e^{2\pi i x\cdot\xi} h(t\xi) d\xi = (f * \Phi_t)(x).$$

The most simple choice is $h(x) = e^{-\pi|x|^2}$ (Gauss means), for which it is easy to compute that $\Phi(\xi) = e^{-\pi|\xi|^2}$, that is $\hat{h} = h$. For this choice we then have

$$\int_{\mathbf{R^d}} \hat{f}(\xi) e^{-t\pi|\xi|^2} e^{2\pi i x\cdot\xi}\, d\xi = (f * \Phi_{\sqrt{t}})(x).$$

The Gauss means are connected with the heath diffusion problem: indeed one can check that $u(t,x) = (f * \Phi_{\sqrt{t}})(x)$ is a solution of the heath equation

$$\frac{\partial u}{\partial t} = \frac{1}{4}\Delta_x u(x,t).$$

Another choice is $h(x) = e^{-2\pi|x|}$ leading to the Abel means. One can check that in this case

$$\Phi(\xi) = c_d \frac{1}{(1+|\xi|^2)^{(d+1)/2}}, c_d = \frac{\Gamma[\frac{d+1}{2}]}{\pi^{(d+1)/2}}.$$

and that in this case $u(t,x) = f * \Phi_t(x)$ satisfies

$$\frac{\partial^2 u}{\partial t^2} + \Delta_x u(t,x) = 0.$$

that is, it is harmonic in the half-space. This function Φ is called the Poisson kernel P.

In both cases, the pointwise limit $\lim_t \Phi_t * f(x)$ becomes the vertical limit $\lim_t u(x,t)$. We expect that in both cases the initial boundary condition at $t = 0$ is f, meaning $\lim_t u(x,t) = f(x)$. We know that this limit takes place in $L^p(\mathbf{R^d})$ if $f \in L^p(\mathbf{R^d})$, and now we would like to know about the pointwise behaviour. So we are led to consider the maximal function

$$(\Phi)^*(f)(x) = \sup_t |(\Phi_t * f)(x)|.$$

We will see next that in most cases this maximal function is dominated by Mf.

Lemma 2. If $\Psi \in L^1(\mathbf{R^d})$ is positive and radial then $|(\Psi * f)(x)| \leq \|\Psi\|_1 Mf(x)$.

Proof. It is enough to prove it for Ψ simple, that is, of the form $\Psi = \sum_j a_j \chi_{B_j}$ where χ_{B_j} denotes the characteristic function of a ball B_j centered at zero and $a_j \geq 0$. In this case, noticing that $\frac{1}{|B_j|}\chi_{B_j} * f(x)$ is the mean of f over B_j, we have

$$|(\Psi * f)(x)| \leq \sum_j a_j |(\chi_{B_j} * f)(x)| \leq \sum_j a_j |B_j| Mf(x) =$$

$$= Mf(x) \sum_j a_j |B_j| = \|\Psi\|_1 Mf(x).$$

□

It follows from the lemma that if Φ_t is an approximation of the identity with Φ radial then $(\Phi)^*(f)(x) \leq Mf(x)$ and so it satisfies a weak estimate implying that $\lim_t (f * \Phi_t)(x) = f(x) a.e.$. This will hold as well if the *least radially decreasing majorant* of Φ,

$$\Psi(x) = \sup_{|y| \geq |x|} |\Phi(x)|,$$

is in $L^1(\mathbf{R^d})$.

This fact applies to both Gauss and Abel means and hence we get that the above solutions of the heath and Laplace equation have vertical boundary values equal to f a.e. It is not hard to see, for the Poisson kernel P, that not only the vertical maximal function

$$\sup_t |u(x,t)| = \sup_t |(f * P_t)(x)|,$$

is controlled by Mf, but also the *non-tangentail maximal function.* This is defined by associating to each point $x \in \mathbf{R^d}$ the cone in the upper half-space

$$\Gamma(x) = \{(y,t) : |x - y| \leq ct\}.$$

The corresponding maximal function

$$N^* f(x) = \sup_{(y,t) \in \Gamma(x)} |u(y,t)|,$$

is called the non-tangential maximal function. It can be proved by easy modification of the argument above that also it is controlled by Mf, that is, $N^* f(x) \leq CMf(x)$ for some constant C, and therefore it satisfies as well a weak L^1- estimate. Since obviously pointwise non-tangential convergence occurs for continuous functions with compact support one gets

Theorem 4. (Fatou's theorem). If $f \in L^p(\mathbf{R^d}), \mathbf{1 \leq p \leq +\infty}$, the harmonic function in the upper half-space $u(x,t) = (f * P_t)(x)$ has non-tangential limit $f(x)$ at almost all x.