Function Spaces Notes

1 Preliminaries

1.1 Outer Measure

A function $\mu^* : 2^X \to [0, \infty]$ is an outer measure if

- $\mu^*(\emptyset) = 0$
- $\mu^*(\bigcup A_k) \leq \sum \mu^*(A_k)$ for any collection of subsets A_k.

If $X = \mathbb{R}$ we define the Lebesgue outer measure m^* by

$$m^*(E) = \inf\{\sum |I_j| : \text{for all intervals } I_k \text{ such that} \bigcup I_k \supset E\}$$

We say that a set X is μ^*-measurable, or measurable with respect to μ^* if $\forall B \subset X$ we have that $\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \setminus A)$. If a set E is μ^* measurable, we denote its measure by $\mu(E)$.[1] Note that if $\mu^*(A)$ is zero, then A is measurable.

If E_k are μ^*-measurable sets then:

- $\cap E_k$ and $\cup E_k$ are measurable
- If E_k are disjoint then $\mu\left(\bigcup E_k\right) = \sum_k \mu(E_k)$.
- If E_k is an increasing sequence of sets then $\lim \mu(E_k) = \mu(\bigcup E_k)$.
- If E_k is a decreasing sequence of sets and $\mu(E_1) < \infty$ then $\lim \mu(E_k) = \mu(\bigcap E_k)$.

The collection of measurable sets is a $\sigma-$**algebra**, namely, it contains the empty set, is closed under countable unions and is closed under complements. We will write $\mathcal{B}$ for the Borel σ-algebra, the smallest σ algebra containing all open sets. We can define this in two ways, either by considering the intersection of all σ-algebras containing the open sets, or considering the $\sigma-$algebra generated by the open sets. Our attention will mainly be restricted to $\mathbb{R}^n$.

Let μ^* be an outer measure. We say that μ^* is regular if $\forall E \subset X$ $\exists$ a measurable B with $B \supset E$ and $\mu^*(B) = \mu^*(E)$. We say that μ^* on $\mathbb{R}^n$ is Borel if every Borel subset is μ^* measurable. We say μ^* is Borel regular if μ^* is Borel and $\forall A \subset \mathbb{R}^n$ $\exists B \in \mathcal{B}$ such that $A \subset B$ and $\mu^*(B) = \mu^*(A)$. We say that μ^* is Radon if it is Borel regular and every compact subset has finite measure.

Note that if E_k increases and μ^* is regular then regardless of measurability we have that $\lim \mu^*(E_k) = \mu^*(\bigcup E_k)$.

We will write $(\mu^* \lfloor A)$ for the restriction of μ^* to the set A, namely, $(\mu^* \lfloor A)(E) = \mu^*(A \cap E)$.

Lemma 1.1.

- *Let $A \subset \mathbb{R}^n$ be Borel. Then $(\mu^* \lfloor A)$ is a Borel regular measure.*
- *If μ^* is Borel regular, $A \subset \mathbb{R}^n$ is measurable and $\mu^*(A) < \infty$ then $(\mu^* \lfloor A)$ is Radon.*
- *If μ^* is a Borel measure $B \in \mathcal{B}$ of finite μ^* outer measure then $\forall \epsilon > 0$ $\exists K, G$ with K open and G closed such that $K \subset B \subset G$ and $\mu^*(B \setminus K) < \epsilon$ and $\mu^*(G \setminus B) < \epsilon$.*
- *The previous statement holds true for $\epsilon = 0$ under the weaker condition that K is F^σ and G is G^δ.*

[1] This means that for a measurable set $\mu^*(E) = \mu(E)$, but this is neither here nor there.

- *If B is Borel, μ^* Randon on $\mathbb{R}^n$ then $\forall \epsilon > 0$ $\exists O$ where O is open such that $\mu(O \setminus B) < \epsilon$), ie, we do not need to have that $\mu^*(B)$ finite.*

We also have Caratheodory's criterion:
Lemma 1.2. *If μ^* is an outer measure on $\mathbb{R}^n$ and if $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$ for all sets A, B with $d(A, B) > 0$ then μ^* is a Borel measure.*

1.1.1 Some Finer Properties of Measure

There are two equivalent definitions for Lebesgue measure: the first is the one we gave above, namely $\inf\{\sum |I_j| : \cup I_j \supset a\}$, and the second is $\inf\{\sum \text{diam } C_j : C_j \subset \mathbb{R}, \cup C_j \supset A\}$. The second has the advantage that it's defined for all subsets of $\mathbb{R}$ c_j. However, we can argue that these are the same (one way is to note that the Lebesgue measure is regular, so for any C_j we can cover it by intervals that are only ϵ too much, thus the first is at most ϵ above, therefore it's less than or equal to the second, but the second is obviously greater than or equal.)

We make some statements about the definitions of measure and specifically $\mathbb{R}^n$ to clean up some of the details:
Lemma 1.3. *In the definition of outer measure, we must take countably many sets to have countable additivity.*

This can be observed (rather simply) by noting that if we took uncountably many $\mu(A) = \mu(\bigcup_{a \in A} a) \leq \sum \mu(a) = 0$, so everything would have 0 size.
Claim 1.4. *For $n \geq 1$ there does not equal a countably additive set function on $\mathbb{R}^n$ that is invariant under isometries (eg, non-measurable sets aren't a bug, they must always exist). Moreover, for $n \geq 3$ there does not even exist a finitely additive set measure.*

The culmination of the non existence of a finitely additive set measure in $n = 3$ is the Banach Tarski paradox, namely, we can decompose the unit sphere into 2 copies of itself.

The existence of a finitely additive set measure for $n = 1, 2$ was proven by Banach.

1.2 Integrability

Take $(X, \mathcal{M}, \mu)$ a measure space. A function $f : X \to \overline{R}$ is measurable if for all open sets O we have $f^{-1}(O) \in \mathcal{M}$ (ie, preimages of open sets are measurable.) We call a function **simple** if we can write $s(x) = \sum \alpha_j 1_{E_j}$ for a finite sequence of measurable sets E_j.
Theorem 1.5. *If $f : X \to [0, \infty]$ is measurable then there exists a sequence s_n of simple functions s_n such that s_n converges to f pointwise, and $s_0 \leq s_1 \leq \dots$.*

We are now in a position to define the **Lebesgue integral**: for a simple function $s_n = \sum_{i=1}^k \alpha_i 1_{E_i}$ we define the integral as $\int f \, d\mu = \sum_{i=1}^k \alpha_i \mu(E_i)$.[2] For positive measurable functions we take $\int f \, d\mu = \sup\{\int s \, d\mu : s \leq f \text{ for } s \text{ simple }\}$. Then for functions that are not necessarily positive, we define $\int f \, d\mu$ only in the case $\int |f| \, d\mu$ exists, and when it does we define it as $\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu$, where $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$.

The following theorems are known:
Theorem 1.6.
- *If f_n increases pointwise to f for f_n, f both positive, then $\int_X f_n \, d\mu$ increases to $\int f \, d\mu$.*

[2] Here, as in most of measure theory, we make the convention $0 \cdot \infty = 0$, otherwise everything breaks.

- *If $f_n : X \to [0,\infty]$ then $\int \sum f_n \, d\mu = \sum \int f_n \, d\mu$.*
- *If $f_n \geq 0$ then $\int \liminf f_n \, d\mu \leq \liminf \int f_n \, d\mu$.*
- *If $f \geq 0$ and $\nu(E) = \int_E f \, d\mu$ for $E \in \mathcal{M}$, then $\nu : \mathcal{M} \to [0,\infty]$ is a measure on $\mathcal{M}$. Moreover, $\int_E g d\nu = \int_E gf \, d\mu$, ie, $f \, d\mu = d\nu$.*
- *If f_n is measurable, and $\lim f_n(x) = f(x)$, and we have a **dominating function** $\phi(x)$ for which $|f_n(x)| < \phi(x)$, and $\int \phi(x) \, dx < \infty$, then $\lim \int f_n(x) \, d\mu = \int f(x) \, d\mu$.*

1.3 L^p spaces

We are now in a position to define the L^p spaces. We define

$$L^p(\mu) = \{f : X \to \overline{R} : \int_X |f(x)|^p \, d\mu < \infty\}$$

for $1 \leq p < \infty$. For $p = \infty$ we define $L^\infty(\mu) = \{f : X \to \mathbb{R} : \|f\|_\infty < \infty\}$, where $\|f\|_\infty$ is the essential supremum of f, defined as $\inf\{y > 0 : \mu(x : |f(x)| > y)\} = 0$. We turn each of these into normed spaces, defining $\|f\|_p = \left(\int_X |f(x)|^p \, d\mu\right)^{1/p}$ (and leaving $\|.\|_\infty$ as it is.)

We recall three key inequalities: **Minkowski's inequality** states that $\|f + g\|_p \leq \|f\|_p + \|g\|_p$, **Holder's inequality** states that $\|fg\|_1 \leq \|f\|_p \|g\|_q$ whenever $\frac{1}{p} + \frac{1}{q} = 1$ and **Jensen's inequality** which states that if $\mu(X) = 1$, and φ is a convex function, then $\varphi(\int_X g \, d\mu) \leq \int_X \varphi \circ g \, d\mu$.

We now give various properties of the L^p spaces:

- $(L^p, \|.\|_p)$ is a (semi-)normed space.[3]
- $(L^p, \|.\|_p)$ are Banach spaces, this is the Riesz Fischer theorem.
- If $f_n \to f$ in L^p then there is a subsequence f_{n_j} that converges to f almost everywhere.

We note that we can not strengthen the last claim to $f_n \to f$ almost everywhere. Indeed, consider $f_{jn} = 1_{\frac{j-1}{n}, \frac{j}{n}}$ and consider them in the order $(1,1), (1,2), (2,2), (1,3), \ldots$, denoting by φ_k the k'th term in this series. We note that $\|\varphi_k\| \to 0$, but for any fixed x we have that $\varphi_k(x)$ is not convergent.

We know that the L_p closure of the smooth functions of compact support is equal to L^p: this is generally argued using a convolution with a bump function. We may take a more direct proof: this is Lusin's theorem.

If we have $\mathbb{R}^n$ we define the measure of the set $I_1 \times \cdots \times I_n$ to be $\prod_{i=1}^n |I_i|$, and we define m_n^* to be an outer measure defined by $m^*(X) = \inf\{$sums of measures of product sets that cover $X\}$. The theory of L^p spaces generalizes readily.

1.3.1 Computation of the p-norm

The following section gives us the majority of the calculation to show that L^p has dual L^q where $p^{-1} + q^{-1} = 1$.[4]

[3] It's only a semi-norm because of functions of measure zero. Often we just quotient out by the equivalence relation $f \sim g$ iff $f(x) = g(x)$ almost everywhere.

[4] Well: it shows that the map $U : L_q \to L_p^*$ is an into isometry. To show that it is onto is a pain: the proof I found online involves arguing that L_p is uniformly convex and (thus) to each point of the dual ball there is a norming

We know by Holders inequality that $\int |fg|\,dx \leq \|f\|_p \|g\|_q$, so if $\|g\|_q = 1$ then $\int |fg|\,dx \leq \|f\|_p$. However, the following computation shows that there exists a g for which equality holds: set $g = C|f|^{p-1}\text{sign}f$. To ensure that the q norm of g equals 1, take $C = \frac{1}{\|f\|^{p/q}}$. It can be seen easily that for such a function g we have that $\|fg\|_1 = \|f\|_p$.

If $f \in L_\infty$ a similar result holds true (when we impose the condition of σ finiteness), namely, $\|f\|_{L_\infty} = \sup \int fg\,dx$ where we take the supremum over functions whose 1-norm is equal to 1. (This is essentially due to the fact that the unit ball of L_1 is weak star dense in the unit ball of L_∞, however we give a proof here nonetheless). We assume that $X = \bigcup A_j$ with A_j increasing and $\mu(A_j) < \infty$. Take any $0 < \lambda < \|f\|_{L_\infty}$ and consider the set E_λ, where f is above λ in absolute value. Then consider $g = \frac{1_F \text{sign}f}{\mu(F)}$, where $F = A_{j_0} \cap E_\lambda$. This suffices.

Claim 1.7. *If $\int |fg|\,d\mu < \infty$ for all $g \in L_q$ then $\exists 0 < k < \infty$ such that $|\int fg\,d\mu| \leq k$ for each $\|g\|_q = 1$.*

This is the uniform boundedness principle in surprise.

Proof. Assume not: take functions $\|g_n\|_q = 1$ but $\int |fg|\,d\mu = 3^n$. Consider $\sum \frac{1}{2^n} g_n$. This is such that $\int |fg|\,d\mu \leq \sum_n \int |fg_n|\,d\mu = \sum \frac{3}{2}^n = \infty$, a contradiction. $\square$

2 Decompositions, Maximal Functions and Coverings

2.1 Rearrangement of Functions

Let $\varphi : (0,\infty) \to [0,\infty)$ be a non-decreasing right continuous function that vanishes at infinity. Let $t = \varphi(\tau)$. We define the inverse by $\varphi^*(t) = \inf\{\tau : \varphi(\tau) \leq t\}$ (this is legitimately *an* inverse of φ, the only issue is that we have to look at what happens on the 'bad' sets where φ^* could take multiple values. This is simply one such choice.) We note the following properties:

- φ^* is non increasing.
- $\int_0^\infty \varphi^*(t)\,dt = \int_0^\infty \varphi(t)\,dt$, which is evident by considering the area under the curve.
- $t \geq \varphi(\varphi^*(t))$ and $\tau \geq \varphi * (\varphi(t))$: these two follow (essentially) since φ^* is the least possible inverse.
- $|\{t : \varphi^*(t) > \alpha\}| = \varphi(\alpha)$: We have that $\{t : \varphi^*(t) > \alpha\} = (0, \varphi(\alpha))$ for $t > 0$: if $t \notin (0, \varphi(\alpha))$, which means $t \geq \varphi(\alpha)$. We thus have that $\varphi^*(t) \leq \varphi^*(\varphi(\alpha)) \leq \alpha$, by the third property. Conversely, $t \notin \{t : \varphi^*(t) > \alpha\}$, means that $\varphi^*(t) \leq \alpha$ implies that $\varphi(\alpha) \leq t$, applying the third property again.
- φ^* is right continuous, again, this is because φ^* is the least possible inverse.

2.1.1 Property †

We say that a function satisfies property † if the sets where the function is big vanish at infinity, namely, given $(X, \mathcal{M}, \mu)$ and $f : X \to \mathbb{R}$ is measurable, we say that f satisfies property † if

point.(Also it's worth noting (as per normal) that the dual of L^∞ is not L^1 for any non trivial case. Suppose that there are countably many disjoint sets with measure non-zero, and this gives us an isometry to ℓ_∞. Now ℓ_∞ does not have dual ℓ_1, eg, take any non principal ultrafilter and associate a sequence with its limit along the ultrafilter.)

$\mu(x : f(x) > y) \to 0$ as $y \to \infty$, and this measure is finite for all y.

Claim 2.1. *If $f \in L^p$ we have that property † is true for f.*

Proof. It is evidently true for L^∞. We know that for $1 \leq p < \infty$ we know that $E_y = \{x : |f(x)| > y\}$ is such that $\mu(E_y) = \int_{E_y} d\mu \leq \frac{1}{y^p} \int_{E_y} |f(x)|^p \, d\mu \leq \frac{1}{y^p} \int_X |f(x)|^p \, d\mu = \frac{\|f\|_p}{y^p}$. Hence $\mu(E_y) = \mathcal{O}y^{-p}$, and thus it tends to zero. □

We know that not all functions that satisfy † are in L^p: take $f(x) = \frac{1}{x}$. We know that $|\{x : \frac{1}{x} > y\}| = \frac{1}{y}$, but $f(x)$ is not in any L^p.

2.1.2 Distribution function of f

Let $\lambda_f(y) = \mu\{x : |f(x)| > y\}$, for $y > 0$.We know that λ_f is non-increasing, right continuous and behaves similarly to φ (of these only right continuity is difficult to see: if y_n is a sequence that decreases to y then $\{x : |f(x)| > y_1\} \supset \{x : |f(x)| > y_2\} \supset \ldots$, and thus $\bigcup_n \{x : |f(x)| > y_n\} = \{x : |f(x)| > y\}$, and the measure of the left hand side converges to the measure of the right hand side.

We define the **equi-distributed rearrangement of** f by $f^*(t) = \inf\{y : \lambda_f(y) \leq t\}$ for $t > 0$. Note that f^* maps $(0, \infty)$ to $\mathbb{R}$. f^* is monotone decreasing, and has the following properties:

- $|\{t : f^*(t) > \alpha\}| = \mu\{x : |f(x)| > \alpha\}$
- For $1 \leq p < \infty$ we have that $p \int_0^\infty y^{p-1} \lambda_f(y) \, dy = \int_X |f(x)|^p \, d\mu$, to see this note that $\int_X |f(x)|^p \, d\mu = \int_X \int_0^\infty 1_{f(x)^p > t} dt dx = \int_0^\infty |\{f(x)^p > t\}| dt$. Substituting, we get that $\int_0^\infty p y^{p-1} \lambda_f(y) \, dy = \|f\|_p$. Since f^* is equimeasurable with f this gives us that we preserve the L^p norm by rearranging.

2.2 Hardy Littlewood Principles

We first fix some notation. Let $Q = Q(x, h)$ be a cube in $\mathbb{R}^n$ centred at x of side length h. We write $\alpha Q = Q(x, \alpha h)$, for $\alpha > 0$.

We say that $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ if $f : \mathbb{R}^n \to \bar{R}$ is measurable and $f 1_K \in L^1$ for any compact K. We note that $L^1_{\text{loc}} \supset L^1$ and, eg, $C(\mathbb{R}) \subset L^1_{\text{loc}}$. L^1_{loc} is not a normed space a priori, in fact, it is a locally convex topological vector space with semi norms given by $\|f 1_K\|_1$, where we take this over every compact K.

There are two forms of the Hardy Littlewood principle in terms of the following two operators:

- The operator M_1 defined as $(M_1 f)(x) = \sup_h \frac{1}{|Q(x,h)|} \int_{Q(x,h)} |f(x)| \, dx$, the so-called **centred maximal operator**.
- The operator M_2 defined as $(M_2 f)(x) = \sup_Q \frac{1}{|Q|} \int_Q |f(x)| \, dx$, where we take the supremum over all cubes Q containing x, the so-called **uncentred maximal operator**.

It is worth noting a couple of things about the Maximal operators (in both forms):

- If f is L^1 and non zero, then Mf is positive for all x. Take the set E on which f is non negative, and by assumption $|E| > 0$. So, taking any Q that intersects E we have that $\int_Q |f| \, dt$ is strictly positive.

- If we have φ a non-increasing function (as per section 2.1) on $(0, \infty)$, then $M\varphi(x) = \frac{1}{x}\int_0^x \varphi(t)\,dt$. It is clear that $M\varphi$ is bigger than this thing, and we can show that this is non-increasing: for $0 < x < y$ we have that $\frac{1}{x}\int_0^x \varphi(t)\,dt = \int_0^1 \varphi(xu)\,du \geq \int_0^1 \varphi(yu)\,du$ by the fact φ is non increasing, and thus $\geq \frac{1}{y}\int_0^y \varphi(t)\,dt$.
- Mf can be $\gg f$. For an example, consider the function $f(x) = \frac{1}{x \log^2 x}$ for $x < \frac{1}{e}$. By the above computation, we get that $Mf(x) = \frac{1}{x \log x}$, which is much bigger.

Claim 2.2. *The two versions are linearly equivalent, eg, there exist constants c, C independent of f and x such that $cM_1f(x) \leq M_2f(x) \leq CM_1f(x)$.*

Proof. First of all note that $c = 1$ suffices: we don't have to worry about this.

Note that for any Q containing the point x we have can construct some $\tilde{Q}$ such that $Q \subset \tilde{Q} \subset 3Q$, and x is the centre of $\tilde{Q}$. We now perform the computation: $\frac{1}{|Q|}\int_Q |f|dx \leq \frac{1}{|Q|}\int_{\tilde{Q}} |f|dx \leq \frac{|\tilde{Q}|}{|Q|}\frac{1}{|\tilde{Q}|}\int_{\tilde{Q}} |f(x)|dx \leq \frac{|3Q|}{|Q|}M_1f(x) = 3^n M_1 f(x)$. Taking supremum over all cubes Q gives the result. □

Note that $M_1 f$ and $M_2 f$ are measurable functions (simply because the sup of a family of functions is measurable). Also note that for $g \in C_c(\mathbb{R}^n)$ we have that $|\frac{1}{|Q|}\int_Q g(t)\,dt - g(x)| \to 0$ as $Q \to x$ (this is some variant of the integral mean value theorem.)[5]

We now make a series of statements that aim to show that the previous result is true for functions in L_1. The only weakening we get is that the result is only true for almost all x.

Claim 2.3. *Suppose that there exists some constant $c > 0$ (depending only on n) such that for all $f \in L_1(\mathbb{R}^n)$ the set $\lambda_{Mf}(y) = |\{x : Mf(x) > y\}| \leq \frac{c}{y}\|f\|_1$ for all $y > 0$.*

Then for any $f \in L_1(\mathbb{R}^n)$ we have that $\frac{1}{|Q|}\int_Q |f(t) - f(x)|\,dt \to 0$ as $|Q| \to 0$ for almost every x.

The idea of the proof is that we can approximate f by a continuously compactly supported g, and then look at where $f - g$ is 'big'.

Proof. Let E denote the set where $\{x : \limsup_{Q \to x} \frac{1}{|Q|}\int_Q |f(t) - f(x)|\,dt > 0\}$. Our goal is to show that the measure of E is zero. Let E_j denote the set where the lim sup is bigger than $\frac{1}{j}$. It is evident that $E = \cup E_j$, so if we can show that the measure of E_j is zero we are done.

Fix some $g \in C_c(\mathbb{R}^n)$ (which we will later take to approximate f very well in the L_1 norm.) Since $g \in C_c(\mathbb{R}^n)$ from the previous result we get that the set $E_j = \{x : \limsup_{Q \to x} \frac{1}{|Q|}\int_Q |(f - g)(t) - (f - g)(x)|\,dt > \frac{1}{j}$. Thus, $E_j \subset \{x : \limsup_{Q \to x} \frac{1}{|Q|}\int_Q |(f - g)(t)|\,dt + |(f - g)(x)| > \frac{1}{j}\}$. We can split this into two sets, namely, the set where the first one is $> \frac{1}{2j}$ and the one where the second is $> \frac{1}{2j}$.

The first of these is $\leq 2jc\|f - g\|_1$, by the hypothesis. The second of these is less than $2j\|f - g\|_1$ by the formula on the decrease in L_1 norm. So, the sum of these is bounded by $2j(c + 1)\|f - g\|_1$. But then, since $C_c(\mathbb{R}^n)$ is dense in $L_1(\mathbb{R}^n)$ we have that $|E_j| = 0$, eg, $|E| = 0$. □

We now prove a sort of converse to this,

[5] Apply it in each direction at a time to get that thre is some point r in the cube such that the integral equals $|Q|f(r)$, and then apply continuity.

Claim 2.4. *There exists some c (again, depending only on n) such that if* $\frac{1}{|Q|}\int_Q |f_0(t) - f_0(x)|dt \to 0$ *as* $Q \to x$ *for almost all x, then* $|\{x : (Mf_0)(x) > y\}| \leq \frac{c}{y}\|f_0\|_1$.

Proof. Fix $y > 0$, and let $s \geq \frac{\|f_0\|_1}{y}$. Partition $\mathbb{R}^n$ into a collection of disjoint cubes $\{Q\}$ such that $|Q| = s$. Note that $\frac{1}{|Q|}\int_Q |f_0|\,dx \leq \frac{1}{s}\|f_0\|_1 \leq y$ for each $Q \in \{Q\}$.

We will exhibit a sequence of cubes Q_j that are countable and disjoint such that

1. $y < \frac{1}{|Q_j|}\int_{Q_j} |f_0|\,dx \leq 2^n y$
2. For almost all x not in any Q_j, $|f_0(x)| \leq y$.

To exhibit such a collection, we subdivide each Q in $\{Q\}$ into 2^n disjoint congruent cubes, $\tilde{Q}_j$. If $\frac{1}{|\tilde{Q}_j|}\int_{\tilde{Q}_j} |f_0|\,dt > y$ then put $\tilde{Q}_j$ into the collection. If not, we have that $frac1|\tilde{Q}_j|\int_{\tilde{Q}_j} |f_0|\,dt \leq y$. We then repeat this subdivision process on each $\tilde{Q}_j$.

Evidently the collection of cubes we generate is countable. Let $Q' \supset Q$ be the *parent* cube of Q, namely, the cube from which we subdivide. Note that the collection Q_j is distinct from the collection $\{Q\}$, so every cube has a parent such that $\frac{1}{|Q|}\int_Q |f_0|\,dx \leq y$.

For any cube Q_j we observe that $y < \frac{1}{|Q_j|}\int_{Q_j} |f_0|\,dt \leq \frac{|Q_j'|}{|Q_j|}\frac{1}{|Q_j|}\int_{Q_j'} |f_0|dt \leq 2^n y$, since Q_j' is emphatically not in the collection Q.

To show the second property, if $x \notin \bigcup Q_j$ then there exists $\{Q_k\}_{k\geq 1}$ such that $Q_k \to x$ and $\frac{1}{|Q_k|}\int |f_0|\,dt \leq y$. But, by hypothesis, as $Q_k \to x$ we have that $\frac{1}{|Q_k|}\int |f_0|\,dt \to f_0(x)$ for almost all x, so $|f_0(x)| \leq y$.

We note that $|\bigcup Q_j| = \sum |Q_j| \leq \frac{1}{y}\sum\int_Q |f_0|\,dt \leq \frac{\|f_0\|_{L_1}}{y}$.

We now claim that there exists a $c \in (0, \infty)$ (depending only on n) such that if $x \notin \bigcup 2Q_j$ then $Mf(x) \leq cy$.

Fix some $x \notin \bigcup 2Q_j$, and $x \in Q$. Then we estimate the size of Mf:

$$\frac{1}{|Q|}\int_Q |f_0|\,dt = \frac{1}{|Q|}\int_{Q\setminus\cup Q_j} |f_0|\,dt + \frac{1}{|Q|}\int_{Q\cap\cup Q_j} |f_0|\,dt$$

The first of these is $\leq y$. To estimate the second of these we need to work a little harder: the second term $= \frac{1}{|Q|}\sum_j \int_{Q\cap Q_j}\int_{Q\cap Q_j} |f_0|\,dt$, as the Q_j's are disjoint, $\leq \frac{1}{|Q|}\sum |Q_j|\frac{1}{|Q_j|}\int_{Q_j\cap Q}\int |f_0|\,dt \leq 2^n y\frac{1}{|Q|}\sum |Q_j| \leq 2^n y\frac{|3Q|}{Q} = 6^n y$.

So we have that $\frac{1}{|Q|}\int_Q |f_0|\,dt \leq (1 + 6^n)y$, eg, taking sup over all cubes Q we get that $Mf(x) \leq cy$.

Thus, $|\{x : Mf(x) > cy\}| \leq |\bigcup 2Q_j| \leq \sum |2Q_j| \leq 2^n \sum |Q_j| \leq \frac{2^n}{y}\|f_0\|_1$. □

We have a basic corollary: $\{x : (Mg)(x) > y\} \leq \frac{c}{y}\|g\|_1$ for all g that are C_c, evidently the hypotheses are satisfied for this all such g.

Claim 2.5. *The previous corollary extends by density to all* L_1 *functions, namely,* $\exists 0 < c < \infty$ *independent of both* f, y *such that* $\lambda_{Mf}(y) = |\{x : Mfx > y\}| \leq \frac{c}{y}\|f\|_1$.

Proof. Let g_j be a sequence in $C_c(\mathbb{R}^n)$ such that $\|f-g_j\|_1 \to 0$.

Fix y and consider $E = \{x : Mf(x) > y\}$. Set $E_j = \{x : Mg_j(x) > y\}$. By the corollary proved above, there is some c such that $|E_j| \le \frac{c}{y}\|g_j\|_1$.

For a fixed $x \in E$ we have that $Mf(x) > y$, eg, there is some cube Q such that $\frac{1}{Q}\int_Q |f(t)|\,dt > y$. Thus, for sufficiently large j we have that $\frac{1}{Q}\int_Q |g_j(t)|\,dt > y$ as well, eg, $x \in E_j$ for each j sufficiently large. Thus we have that $E \subset \bigcup_{j=1}^{\infty}\bigcap_{k\ge j} E_k$. Writing $A_k = \bigcap_{j\ge k} E_j$ we have that A_k is a decreasing sequence, eg, $|A_k| \to |\bigcup A_k| \ge |E|$. However, $A_k \subset E_k$ and so we have that $|E| \le \liminf |E_k| \le \frac{c}{y}\|f\|_1$. □

We are now done with a proof of the **Lebesgue differentiation theorem**, which states the following:

Theorem 2.6. *For $f \in L_1(\mathbb{R}^n)$ we have that $\frac{1}{|Q|}\int |f(t)-f(x)|\,dt \to 0$ as $Q \to x$ for almost all x.*

Proof. Claim 2.5 followed by Claim 2.3 gives the result. □

2.3 Strong and Weak (p,q) operators

We will generalize results of the form that we proved in the previous section. Consider an operator T from the set of measurable functions to the set of measurable functions such that $|T(f_1+f_2)(x)| \le |Tf_1(x)| + |Tf_2(x)|$ (ie, T is sublinear). We say that such an operator is **strong** (p,q) if T is bounded as a map from L_q to L_p, namely, if $\|Tf\|_p \le c_{p,q}\|f\|_q$.

We say that such an operator is **weak** (p,q) if $\lambda_{Tf}(y) = |\{x : (Tf)(x) > y\}| \le c\left(\frac{\|f\|_q}{y}\right)^p$.

It is evident that strong (p,q) implies weak (p,q), eg, $|\{x : |Tf(x) > y|\}| \le \frac{1}{y^p}\|Tf\|_p^p \le c\left(\frac{\|f\|_q}{y}\right)^p$. The converse is not true: namely, it is not true that $\|Mf\|_1 \le c\|f\|_1$.[6]

Theorem 2.7. *The maximal operator Mf is weak (1,1).*

Proof. This is the contents of claim 2.5. □

We have shown that Mf is weak (1,1) and it is not strong (1,1). However, we will now show that it is actually strong (p,p) for all $p \ne 0$. To achieve this we will need a lemma:

Lemma 2.8. *For $f \in L_1$ we have the following inequality:*

$$|\{x : Mf(x) > y\}| \le \frac{c}{y}\int\limits_{\{x:|f(x)|>y/2\}} |f(t)|\,dt$$

Proof. For $f \in L_1$, fix $y > 0$ and split f into $f_y + f^y$, the first term where we restrict f to where it's smaller than $y/2$ in modulus, and the second term where we restrict to where it is larger than $y/2$ in modulus. We note that $Mf(x) \le Mf^y + Mf_y$, and hence $\lambda_{Mf}(y) \le \lambda_{Mf^y}(y/2) + \lambda_{Mf_y}(y/2)$.

[6] In fact, Mf is not in L_1 whenever $f \in L_1$ isn't identically zero. Assume that there is some set on which $|f| \ge 0$ and that we can take a compact K on which $|f| > \gamma > 0$. Then $Mf > M\gamma 1_K \ge \frac{\gamma|K|}{x}$, which is certainly not L_1.

Note $Mf_y(x) = \sup_Q \frac{1}{|Q|}\int_Q |f_y|\,dt \leq \frac{y}{2}$, eg, $\{x : Mf_y(y) > \frac{y}{2}\} = \emptyset$. So, by the inequality at the end of the previous paragraph, the left hand side is $\leq \lambda_{Mf^y}(y/2) \leq \frac{c}{y/2}\|f^y\|_c \leq \frac{c}{y/2}\int\limits_{\{x:|f(x)|>y/2\}} |f(t)|\,dt$. □

We can now prove the **Hardy-Littlewood theorem**:

Theorem 2.9. *There are constants c_p for $1 < p \leq \infty$ such that $\|Mf\|_p \leq c_p\|f\|_p$.*

Proof. For $p = \infty$ this result is trivial, $\frac{1}{|Q|}\int_Q |f|\,dt < \|f\|_\infty$, so restrict to the case $1 < p < \infty$.

We have that $\|Mf\|_p^p = p\int_0^\infty \lambda_{Mf}(y)\,dy$ by the formula proven in section 2.1.2. By the previous lemma, this is $\leq cp\int_0^\infty y^{p-2}\int\limits_{\{x:|f(x)|>y/2\}} |f(x)|\,dxdy$. Interchanging our integrals by Fubini's theorem we get $= cp\int_{\mathbb{R}^n} |f(x)|\int_0^{2|f(x)|} y^{p-2}dydx$. Performing the integral, we get that this $= \frac{cp2^{p-1}}{p-1}\|f\|_p^p$, ie, the theorem holds with $c_p = \frac{cp2^{p-1}}{p-1}^{1/p}$. We note that as $p \to 1$ c_p blows up. □

The constant here has an implicit dependence upon the dimension: in the previous lemma c depended upon the dimension due to the fact that we were using the weak (1,1) inequality. Smarter methods give c_p independent of the dimension.

We can now show that, although Mf is much bigger than f for some functions, for some other functions we have that $Mf \sim f$.

Claim 2.10. *There are functions f for which Mf is linearly equivalent to f, eg, there exists constant C such that $Mf(x) < Cf(x)$.*

Proof. Pick some $f \in L^p$ and consider the fact that there exists a constant c_p such that $\|Mf\|_p \leq c_p\|f\|_p$. Setting $M^n f(x) = M^{n-1}Mf(x)$ for $n \geq 2$ we see that $\|M^n f\|_p \leq c_p^n\|f\|_p$. Set $\varphi(x) = \sum \frac{1}{2^j c_p^j} M^j f$. Then $\|\varphi\|_p \leq \sum \frac{1}{2^j c_p^j} c_p^j \|f\|_p$, eg, $\varphi \in L_p$.

Moreover, $M\varphi(x) \leq \sum \frac{1}{2^j c_p^j} M^{j+1}\varphi(x) \leq 2c_p\varphi(x)$, eg, φ is linearly equivalent to the maximum function of φ. □

We note that in the above, we can generalize all the results to L^1_{loc}: all we do is intersect with a compact set K and zero f outside of it and the proofs go through.

We say that for some function $u \in L^1_{\text{loc}}$, with $u > 0$ everywhere that $u \in \mathcal{A}_1(\mathbb{R}^n)$ if $Mu \leq cu$ for some constant u (almost everywhere).

For E in $\mathbb{R}^n$ we will define (for the purposes of this proof) that $u(E) = \int_E u\,dx$, and $\|f\|_{p,u} = \left(\int_{\mathbb{R}^n} |f|^p u\,du\right)^{1/p}$.

Theorem 2.11. *$u \in \mathcal{A}_1$ if and only if $u(\{x : Mf(x) > y\}) \leq \frac{c\|f\|}{y}$.*

Proof. First we show the $\Longleftarrow$ direction.

We know that for any fixed x_0 we have that $\frac{1}{|Q|}\int_Q |u|\,dt \to u(x_0)$ as $Q \to x_0$. We know that $\frac{1}{|Q|}\int_Q u\,dx = \frac{u(Q)}{|Q|}$. Let $x_0 \in Q' \subset Q$.

Take $f = 1_{Q'}$. We claim that $Q \subset \{x : Mf(x) \geq \frac{|Q'|}{|Q|}$.

For $x \in Q$ we have that $Mf(x) \geq \frac{1}{|Q|}\int_Q |f|\, dt = \frac{|Q'|}{|Q|}$, and the hypothesis shows that $u(Q) \leq \frac{c}{\frac{|Q'|}{|Q|}} u(Q')$, eg $\frac{u(Q)}{Q} \leq c\frac{u(Q')}{|Q'|} \leq cu(x_0)$.

Now we show the $\implies$ direction.

We first establish the result for $g \in C_c$, and then extend the result by a density argument (since u is finite on compact sets, $u\, dx$ is regular, and this suffices to show that C_c is dense in L^1.)

Fix some $g \in C_c$. By the proof of Claim 2.4 there exists some sequence of cubes Q_j such that $y < \frac{1}{|Q_j|}\int_{Q_j} |g| \leq 2^n y$ and $|g(x)| \leq y$ for $x \notin \cup Q_j$, and there exists some c such that $\{x : Mg(x) > cy\} \subset \cup 2Q_j$, and $u(\{x : Mg(x) \geq cy\} \leq \sum u(2Q_j)$.

We claim that $u(2Q) \subset cu(Q)$. This is true as $\frac{u(2Q)}{|2Q|} \leq 2\inf_{x\in 2Q} u(x) \leq c\inf_{x\in Q} u(x) \leq \frac{c}{|Q|}\int_Q u\, dt$. (This is the statement that u is a doubling measure.)

Then $u\{x : Mg(x) > cy\} \leq c\frac{\sum u(Q_j)}{|Q_j|}|Q_j| \leq \frac{c}{y}\sum \frac{u(Q_j)}{|Q_j|}\int_{Q_j} |g| u u^{=1} dx$. We note that $\frac{u(Q_j}{|Q_j|} \leq cu(t)$ for all $t \in Q_j$ implies that $u^{-1}(t) \leq \frac{c|Q_j|}{u(Q_j)}$.

Putting all of this together gives that we have $\leq \frac{c}{y}\|g\|_{1,u}$. □

2.4 Marcinkiewicz Interpolation Theorem

We now show a result on operators from the L^p spaces, more specifically, if T is an operator whose domain contains L^{p_1} and L^{p_2} with $p_1 < p_2$. In this case we wish to argue that some properties of T on the L^{p_1} and L^{p_2} gives some properties on the intermediary values of p.

To do this more precisely, we define $T : L^{p_1} \oplus L^{p_2} \to \{\text{measurable functions}\}$. Our first observation is that $L^{p_1} \oplus L^{p_2}$ contains L^p for all intermediary p: eg, write $f = f1_{|f|>1} + f1_{|f|\leq 1}$. We have growth for the large scale behaviour in the p_1 stuff and in the small scale behaviour by the p_2 stuff, and thus we're fine.

Theorem 2.12. *Suppose that such a mapping T is sublinear, and that T is weak (p_1, p_1) and weak (p_2, p_2). Then T is strong (p, p) for each p between p_1 and p_2.*

We note that we can not strengthen this for p at the end points: weak does not imply strong.

Proof. Take $p_2 < \infty$ first of all, and $p_1 < p < p_2$.

We know that $\|Tf\|_p^p = p\int_0^\infty y^{p-1}\lambda_{Tf}(y)\, dy$. Fix some $y > 0$, and write $f = f^y + f_y$, where we take $f_y = f$ restricted to where $|f| > y$ (and f^y to be the other thing.) By sublinearity, we have that $\lambda_{Tf}(y) \leq \lambda_{Tf_y}(y/2) + \lambda_{Tf^y}(y/2)$, so we can split $\|Tf\|_p^p$ into two parts, the one where we focus on f_y and one where we focus on f^y, which we denote I_1 and I_2 separately.

We will perform the computation for I_1, the computation for I_2 is similar:

$$I_1 = p\int_0^\infty y^{p-1}\lambda_{Tf_y}(y/2)\, dy \lesssim \int_0^\infty y^{p-1-p_1}\|f_y\|_{p_1}^{p_1}\, dy$$

where this inequality follows by the definition of weak (p, p).[7] This

$$= \int_0^\infty y^{p-p_1-1}\int_{\{x:|f(x)|>y\}} |f(x)|^p\, dxdy = \int_{\mathbb{R}^n} |f(x)|^{p_1}\int_0^{|f(x)|} y^{p-p_1-1}\, dydx = \|f\|_p^p$$

[7] Note that $f \lesssim g$ means that $f \leq cg$ for some constant c.

The first equality comes from Fubini's theorem and rearranging the integrals, and the second comes from performing them. Thus we get that T is strong (p,p).

If $p_2 = \infty$ then $\|Tf\|_\infty \leq c_\infty \|f\|_\infty$. When we now perform the split, we split it into f_y where $|f| > \frac{y}{2c_\infty}$. The second term in the split then vanishes, so we can continue as before. $\square$

2.5 Approximate Continuity

For a measurable set $E \subset \mathbb{R}^n$, we say that x_0 is **a point of density** for E if $\frac{|Q \cap E|}{|Q|} \to 1$ as $Q \to x$, and is **a point of dispersion** if the limit is 0.

Claim 2.13. *If E is measurable, almost every point in E is a point of density, and almost every point in the complement is a point of dispersion*

Proof. Set $f = 1_E$. This is measurable, and by the Lesbegue differentiation theorem we have that $\frac{1}{|Q|}\int_Q |f|\, dt \to f(x)$ almost everywhere. But this is exactly $\frac{|Q \cap E|}{|Q|}$. $\square$

If we have a measurable function f we say that x_0 is a **point of approximate continuity** if $\forall \epsilon > 0$ the set $\{x : |f(x) - f(x_0)| < \epsilon\}$ has x_0 as a point of density.

It's worth noting that approximate continuity is genuinely different from continuity: consider the intervals $I_n = [1/n, 1/n + 1/2^n]$, and $E = \cup I_n$. These intervals are such that 0 is a point of density for E^c (almost obviously) so any function f which is zero outside E and anything inside E is approximately continuous at 0. Not many of these functions are continuous.

We now push this to the extreme:

Theorem 2.14. *Let f be a measurable function. Then f is approximately continuous for almost all $x \in \mathbb{R}^n$.*

Proof. Let r_1, r_2 be rationals and let $E_{r_1,r_2} = \{x : r_1 < f(x) < r_2\}$. We know that there is a null set N_{r_1,r_2} such that $x \in E_{r_1,r_2} \setminus N_{r_1,r_2}$ is a point of density of E_{r_1,r_2}. Set $\bigcup N_{r_1,r_2} = N$ (where we take the union over all pairs of rationals $r_1 < r_2$) and we claim that if $x \notin N$ then f is approximately continuous at x.

Fix some $x_0 \notin N$, and some $\epsilon > 0$. Consider the set $E_\epsilon = \{x : f(x_0) - \epsilon < f(x) < f(x_0) + \epsilon\}$. There are rationals r_1, r_2 such that $f(x_0) - \epsilon \leq r_1 < f(x) < r_2 \leq f(x_0) + \epsilon$, eg, $E_\epsilon \supset E_{r_1,r_2}$, and thus x is a point of density. This means that x is a point of approximate continuity. $\square$

2.6 Generalizations of Maximum functions

We now wish to attempt to generalize the results of sections 2.2 - 2.4 in some ways. We wish to generalize it to measures other than the Lebesgue measure. The key property of Lebesgue measure we used was that $\mu(2Q) \leq c\mu(Q)$ for some constant c independent of Q. We call a measure that obeys a linear relation like this a **doubling measure**.

We write $M_\mu f(x)$ for the obvious generalization, eg, $M_\mu f(x) = \sup_Q \frac{1}{\mu(Q)} \int_Q |f(t)|\, dt$, where the supremum is over cubes Q containing x. We need the convention that if $\mu(Q) = 0$ we set it equal to zero, and if $\mu(Q) = \infty$ we set it equal to zero.

Theorem 2.15. *If μ is a doubling measure, then*

- $\mu\{x : M_\mu f(x) > y\} \leq \frac{c}{y}\|f\|_1$.

- *For $1 < p < \infty$ we have that $\|M_\mu f\|_p \le c_p \|f\|_p$.*

Proof. The proof follows identically as for Lebesgue measure. □

We now give an example to show that the fact that we have a doubling measure is essential. Consider $x_0 = 0$ and x_n being a sequence of points lying on the line $y = 1 - x$ converging monotonically to the point $x = (0.5, 0.5)$ (say). Let $\mu(E)$ denote the measure that counts the number of x_j inside the set E. μ is clearly Borel.

Consider the function $f(x) = 1_{x_0}(x)$. Take cubes whose axes are parallel to the co-ordinate axes and whose top left corner hits the point x_j and are big enough to contain x_0. Then $M_\mu f(x_j) > \frac{1}{|Q|}\int_Q f(x)\,d\mu > \frac{1}{2}$. So, $\mu\{x : M_\mu f(x) > \frac{1}{3}\} = \infty$, which is not less than any constant, so μ is not weak $(1,1)$ (or strong (p,p), as this example also shows.)

2.6.1 Dyadic Cubes

We say that a subset E of $\mathbb{R}^n$ is a **dyadic cube** if we can write it in the form $[\frac{l_1}{2^k}, \frac{l_1+1}{2^k}) \times \cdots \times [\frac{l_n}{2^k}, \frac{l_n+1}{2^k})$. We denote the set of all dyadic cubes by Δ. The prototypical dyadic cube is $Q_0 = [0,1)^n$.

The key property of dyadic cubes is that if we have $Q, Q' \in \Delta$ then either $Q \cap Q' = \emptyset$, or one of the cubes is contained in the other. Moreover, if we have any cube Q' in $\mathbb{R}^n$ then there is some cube $Q \in \Delta$ such that $Q \subset Q' \subset 5Q$.

Consider, for a doubling measure μ, the maximal operator $M^*_\mu(f) = \sup \frac{1}{\mu(Q)}\int_Q f(t)\,d\mu$, where we take the supremum over Q containing x and $Q \in \Delta$. The way that the dyadic cubes nest give us a much stronger maximal principal for this restricted class of maximal operators:
Theorem 2.16. *For $f \in L^1(d\mu)$ we have that*

- *$\mu\{x : M^*_\mu f(x) > y\} \le \frac{1}{y}\|f\|_{L^1(d\mu)}$ (eg, weak (1,1))*
- *$\|M^*_\mu(f)\|_{L^p(d\mu)} \le c_p \|f\|_{L^p(d\mu)}$ (eg, strong (p,p).)*

Proof. Define $M^*_{\mu,r} f(x) = \sup \frac{1}{\mu(Q)}\int_Q |f|\,d\mu$, where we take the supremum over all dyadic cubes Q containing x whose diameter is less than r.

Take $E_{r,y} = \{x : M^*_{\mu,r} f(x) > y\}$. As $r \to \infty$ we have that $E_{r,y}$ increases to $E_y = \{x : M^*_\mu f(x) > y\}$. If $x \in E_{r,y}$ there exists some $Q_x \in \Delta$ such that $x \in Q_x$ and the diameter of $Q_x < r$. Let Q_x be the maximal such cube. We know that $E_{r,y} = \cup Q_x = \cup Q_j$ for disjoint Q_j (by the property that dyadic cubes have empty intersection or are the same) and thus $\mu(E_{r,y}) = \sum \mu(Q_j) \le \frac{1}{y}\|f\|_{L^1(d\mu)}$. Taking $r \to \infty$ we have that $\mu(E_y) \le \frac{1}{y}\|f\|_{L^1(d\mu)}$.

Using the Marcinkiewicz interpolation theorem the second result follows. □

2.6.2 Approximate Identities

This section wasn't lectured here (in fact, it was lectured after Exponential L^p spaces) but it fits in here best of all I think.

For $k : \mathbb{R}^n \to [0,\infty)$ we define $k_\epsilon = \frac{1}{\epsilon^n}K(\frac{x}{\epsilon})$ and $Kf(x) = \sup_{\epsilon>0} k_\epsilon * f$. Let $Mf(x) = \sup_{r>0}\frac{1}{|B(x;r)|}\int_{B(x,r)}|f(t)|\,dt$ be the maximal function on balls (not cubes).

Theorem 2.17. *Suppose that $\varphi : \mathbb{R}^n \to \mathbb{R}^+$ is radially non-increasing and radially symmetric. Then for $f \in L^1_{loc}(\mathbb{R}^n)$ we have that $\sup_\epsilon |\varphi_\epsilon * f(x)| \leq \|\varphi\|_{L^1} Mf(x)$. Thus the operator $M_\varphi f(x) = \sup_\epsilon |\varphi_\epsilon * f(x)|$ is weak(1,1) and strong p,p for $1 < p \leq \infty$.*

Proof. Consider $E \subset \mathbb{R}^n \times [0,\infty] = \{(x,y) : \varphi(x) > y\}$. We know that $\varphi(x) = \int_0^\infty 1_E(x,y)\,dy = \int_0^\varphi(x) 1\,dy$.

Then $\varphi * |f|(x) = \int_{\mathbb{R}^n}|f(x-t)|\varphi(t)\,dt = \int_{\mathbb{R}^n}\int_0^\infty |f(x-t)|1_E(t,y)\,dtdy$. We rewrite this integral in a horrid way using Fubini: $\int_0^\infty \int_{\{t:\varphi(t)>y\}=B(0,r_y)}|f(x-t)|\,dt$. Introducing the measure of the ball of radius r_y we get this equals $\int_0^\infty |B(0;r_y)|\frac{1}{|B(0;r_y)|}\int_{B(0;r_y)}|f(x-t)|\,dt$ by the definition of the maximal function this is $\leq Mf(x)\int_0^\infty |B(0;r_y)|\,dy = Mf(x)\|\varphi\|_{L^1}$, by the rearrangement formulae.

Repeating this argument for φ_ϵ the result follows. □

We prove a strengthening of this result.

Theorem 2.18. *Suppose that K is radially symmetric $\|K\|_{L^1} = 1$ and there exists $\alpha > 0$ such that $|y|^{-\alpha}K(y)$ is radially non-increasing. Then the conclusions of the previous theorem hold.*

Sketch of proof. Set $E = \{(y,t) : k(y)|y|^{-\alpha >}t\}$. Define $E^t = \{y : k(y)|y|^{-\alpha} > t\} = B(0,r_t)$. If $f \geq 0$ we have that $\int_{\mathbb{R}^n} f(x-y)k(y)\,dy = \int_{\mathbb{R}^n} f(x-y)|y|^\alpha k(y)|y|^{-\alpha}\,dy$. This is equal to $\int_0^\infty\int_{B(0,r_t)} f(x-y)|y|^\alpha\,dydt$ by the same considerations as the previous proof. This is $\leq cMf(x)\int_0^\infty dt\int_{B(0;r_t)}|y|^\alpha\,dy$[8].

So we have that $|k*f| \leq cMf(x)\int_0^\infty dt\int_{B(0;r_t)}|y|^\alpha\,dy = \|k\|_{L^1}$.[9] Therefore, as per the proof of the previous theorem, we have that the maximal function in terms of k is weak (1,1) and strong (p,p). □

In the non radially symmetric case there are two results that we quote. The first is due to Coifman.

Theorem 2.19. *Suppose $k : \mathbb{R}^n \to \mathbb{R}$ is radially non increasing (eg, there is some α for which $|u|^\alpha k(ut)$ is non increasing in u for every fixed t) then $Kf(x) = \sup_\epsilon k_\epsilon \star f(x)$ is strong (p,p) for $p \in (1,\infty)$ (not $\infty]$ like the conclusion in the previous case.)*

The second is due to Zygmund and more general:

Theorem 2.20. *Suppose that $k_\alpha : \mathbb{R}^n \to \mathbb{R}$ is defined for $\alpha \in \Gamma$ such that:*

- *$\exists c_1$ such that $\|k_\alpha\|_{L^1} \leq c_1$*
- *$\exists c_2$ such that $\int_{|x|>4y}\sup_\alpha |k_\alpha(x-y) - k_\alpha(x)|\,dx < c_2$ for c_2 independent of y.*

Then the conclusion is as in theorem 2.17.

[8] This step was given as an exercise.
[9] As was this.

2.7 Besicovitch Covering Theorem

The Besicovitch covering theorem asserts that if we cover a set E with uncountably many balls, one centred at each $x \in E$ we can, in fact, toss away all but countably many of them and retain a cover. Moreover, we can do this in a way such that each point in the set is only in finitely many of the balls.

If E is a bounded subset of $\mathbb{R}^n$ we consider a side length function $h : E \to (0, \infty)$ and observe that $E \subset \bigcup_{x \in E} Q(x, h(x))$, and we denote $Q(x, h(x)) = Q_x$.
Theorem 2.21. *Suppose E comes equipped with some h as detailed above. Then there exists countable* x_j *such that* $E \subset \bigcup Q_{x_j}$ *and* $\sum 1_{Q_j}(x) \leq c_n$*, where* $c_n = 4^n$.

We first note that we can not remove the hypothesis that E is bounded: suppose we have E as an unbounded set. Then $E \subset \bigcup Q(x, 2|x| + 1)$. This is such that every cube contains the point 0, so any infinite collection of them does not have the bounded overlap property.

To prove this, the majority of work goes into proving the bounded overlap lemma:
Lemma 2.22. *If* $\{Q_j\}$ *are parallel cubes defined by* $Q_j = Q(x_j, h_j)$ *are such that:*

- *For* $j > k$ *we have that* $x_j \notin Q(x_k, h_k)$
- $|Q_j| < 2|Q_k|$

Then $\exists 0 < c_n < \infty$ *(equalling* 4^n*) such that* $\sum 1_{Q_j}(x) < c_n$ *for all* $x \in \mathbb{R}^n$*, eg, x is in at most* 4^n *cubes.*

We will assume that the sides are parallel to the axes and that the cubes are centred. The fact that the cubes are centred is essential. In what follows we do the $n = 2$ case, the same proof generalizes.

Proof. Our goal is to show that there are at most four possible squares in each quadrant that intersect 0: there are four quadrants and thus we will be done for 0. Translation then suffices to finish the proof. We will call the first quadrant I.

Let $j_1 < j_2 < \dots$ be all of the indices such that $x_j \in I$ and $x_{j_i} \in Q_{x_{j_i}}$. Consider the square $[0, h_1] \times [0, h_1]$ and consider splitting it into four subsquares, A, B, C, D labelled counter clockwise.

We know that $A \subset Q_{x_{j_1}}$ (the same holds for B, C, D). We claim that for each $i > 1$ we have that x_{j_i} is either in A, B, C, D.

By hypothesis we know that $|Q_{j_i}| \leq 2|Q_{j_1}|$, eg, $h_{j_i}^2 \leq 2h_{j_1}^2$, in particular, $h_{j_i} \leq \sqrt{2} h_{j_1}$. Since Q_{j_i} contains 0 by assumption, if $x_{j_i} \notin [0, h_{j_1}] \times [0, h_{j_1}]$ we have that $h_{j_i} \geq 2h_{j_1}$, a contradiction.

There is thus at most one x_{j_i} in A. Similarly, there is at most one in C, D, eg, 4 in this quadrant in total. $\square$

We are now in a position to prove Besicovitch.

Proof of theorem 2.17. Let $\alpha_1 = \sup\{|Q(x, h(x))| : x \in E = E_1\}$. If $\alpha_1 = \infty$ we are done, there would be some sufficiently large cube containing all of E. If $\alpha_1 < \infty$ we know that there exists $x \in E_1$ with $|Q_{x_1}| > \frac{\alpha_1}{2}$.

Inductively construct cubes Q_j centred at x_j of size $> \alpha_j/2$, where $\alpha_j = \sup\{|Q_x| : x \in E_j = E_{j-1}/Q_{j-1}\}$. We have that α_j is a decreasing sequence, and (since it is bounded below by zero) it converges to some constant δ. The claim is now that Q_{x_j} has bounded overlap, eg, that the hypothesis of the bounded overlap lemma hold.

It is evident for $j > k$ that $x_j \notin Q_{x_k}$ and also that $|Q_{x_j}| \leq \alpha_j \alpha \alpha_k \leq 2|Q_{x_k}|$, eg, the hypotheses of the bounded overlap lemma hold and we have that the cubes have bounded overlap.

We just need to show now that $E \subset Q(x_j, h(x_j))$: first we show that $\alpha_j \to 0$. Suppose it doesn't, ie, $\alpha_j > \delta$ for each δ. But then, since $x_j \notin Q_{x_k}$ we have that $\|x_j - x_k\| \leq \delta$ for each pair of indices (j, k). A standard volumetric argument shows that the number of x_j is thus bounded, which is a contradiction.

Now if $x \in E$ and $x \notin \bigcup Q_{x_j}$ then $|Q_x| \leq \alpha_j$ for each j, eg, $|Q_x| = 0$. Thus $h(x) = 0$, a contradiction, so we're done. □

We are now in a position to prove a generalization of Theorem 2.7, eg, the fact that the operator Mf is weak $(1,1)$.

For two Borel measures μ, ν for which μ is Radon (and positive), consider the function $M_{\mu,\nu}(x)$ $= \sup_Q \frac{\nu(Q(x,h))}{\mu(Q(x,h))}$. (For $d\nu = |f| d\, mu$ this recovers the definition of the centralized maximal operator we had before.) This function need not be measurable in x (with regards to either measure.)

Claim 2.23. *There is an equivalent* $(1,1)$ *principle for the function* $M_{\mu,\nu}$*, namely,* $\mu^*(\{x : M_{\mu,\nu}(x) > y\}) \leq \frac{c}{y}\nu(\mathbb{R}^n)$.

Proof. Fix some $y > 0$ and let $E_y = \{x : M_{\mu,\nu}(x) > y\}$, and $E_{y,r} = E_y \cap B(x; r)$ (to use Besicovitch we need a bounded set.)

If $x \in E_{y,r}$ there is some $h(x) > 0$ such that $\frac{\nu(Q(x,h(x)))}{\mu(Q(x,h(x)))} > y$, eg, $E_{y,r} \subset \bigcup Q(x, h(x))$, where we take the union over all $x \in E_{y,r}$. Write $Q_x = Q(x, h(x))$.

By Besicovitch there is a countable collection Q_j that cover $E_{y,r}$ and $\sum 1_{Q_j(x)} < c_n$ for all $x \in \mathbb{R}^n$. Thus $\mu^*(E_{y,r}) \leq \sum \mu(Q_j) \leq \frac{1}{y} \sum \nu(Q_j) \leq \frac{1}{y} \int_{\mathbb{R}^n} \sum 1_{Q_j}(x)\, dx$ where we have used the monotone convergence theorem to interchange summation and integration and thus this is $< \frac{c_n}{y}\nu(\mathbb{R}^n)$. □

We give a (somewhat) trivial example to show that it is necessary to have the cubes centred. Consider the line $x + y = 2$ and consider a sequence of points x_n on the line converging to some $(1,1)$ 'from above' in the first quadrant, and set $x_0 = 0$. Consider cubes Q_n which contain 0 and have the top left vertex at x_n. Set $E = \{x_n\}_{n=0}^{\infty}$. By construction, the only possible subcovering of E is the trivial one, namely, every cube. But this gives $\sum 1_{Q_j}(0) = \infty$, thus we can not continue.

The issue here is that we have no form of bounded overlap: one can prove a form of bounded overlap if the centres of the cubes are all contained in the interior in some *strict* way, eg, if all of them lie within the middle third (in all axes).

2.8 Vitali Covering Lemma

There is a substantial weakening of the Besicovitch covering theorem which we prove now, the hypotheses involved are much stronger, and we get out a result which is correspondingly a little stronger (namely, the cubes involved are disjoint.)

Lemma 2.24. *If $E \subset \mathbb{R}^n$ is a set of bounded Lebesgue outer measure (ie $m^*(E) < \infty$) (and of bounded diameter) which is covered by cubes Q_α (for α in some indexing set Γ) then there is a countable subcollection such that:*

- *Q_i are disjoint cubes*
- *$E \subset \bigcup 5Q_j$*
- *$\exists c_n$ (independent of E and only depending on n such that $\sum |Q_j| \geq c_n m^*(E)$ (where $c_n = 5^{-n}$).*

We note that we can remove the hypotheses of finite outer measure (intersecting with a finite set and taking the union outwards) but the proof is really really tied up with the fact that we have the Lebesgue measure, removing this is impossible.

Proof. As in the proof of Besicovitch let $\alpha_1 = \sup\{|Q_\alpha| : \alpha \in \Gamma\}$. If α_1 is infinite we are done: some cube is big enough to cover the set. If not then take some Q_1 such that $|Q_1| > \alpha_1/2$.

Inductively continue, picking $\alpha_n = \sup\{|Q_\alpha| : \alpha \in \Gamma, Q_\alpha \cap Q_j = 0 \,\forall j < n\}$ and Q_n such that $|Q_n| \geq \alpha_n/2$. We note that α_n is a decreasing sequence. If $\sum |Q_n| = \infty$ we're done, if not $|Q_j| \to 0$, eg, $\alpha_j \to 0$.

We claim that $E \subset \bigcup 5Q_j$. So, fix any $\alpha \in \Gamma$. We show that $Q_\alpha \in \bigcup 5Q_j$, moreover, there is some Q_j such that $Q_\alpha \subset Q_j$. Suppose $Q_\alpha \not\subset \bigcup 5Q_j$. Then $Q_\alpha \cap \cup Q_j \neq \emptyset$, as otherwise $|Q_\alpha| < \alpha_j$ for each j, but then $|Q_\alpha| = 0$, a contradiction. Take j_0 to be the first integer such that $Q_\alpha \cap Q_j \neq \emptyset$. We have that $Q_\alpha \cap (Q_1 \cup \cdots \cup Q_{j_0-1}) = \emptyset$, eg, $|Q_\alpha| < \alpha_{j_0}$. But then $Q_\alpha \subset 5Q_{j_0}$, a contradiction, eg, $Q_\alpha \subset \bigcup 5Q_j$.

Finally, $m^*(E) \leq \sum |5Q_j| = 5^n \sum |Q_j|$. □

This proof can be modified for doubling measures, however we shall not pursue this here.

We can give a modification of Vitali's covering theorem: this concerns fine coverings. We say that a collection of compact cubes $\{Q\}$ is a **fine cover** if $\forall x \in E$ there is a subcollection Q_j such that $Q_j \to x$ and $|Q_j| \to 0$.

Theorem 2.25. *Suppose that Q is a fine covering of a set E of finite Lebesgue measure. Then there exists a disjoint collection $Q_j \in \{Q\}$ such that $m^*(E \setminus \cup Q_j) = 0$, eg, the set $E \setminus Q_j$ is null.*

Proof. First of all, there is some open set G containing E that has finite Lebesgue measure. We can assume for all $Q \in \{Q\}$ that $Q \subset G$, as for all $x \in E$ we can simply ignore cubes that are big enough to be outside the set G.

Take $\{Q_j\}$ according to the construction in the previous theorem, we are now guaranteed that $\alpha_1 < \infty$.

For N a positive integer, we claim that $E / \cup_{j=1}^N Q_j \subset \bigcup_{j \geq N+1} 5Q_j$. We know that Q_j are closed and compact, so for each x not in $\cup_{j=1}^N Q_j$ there exists some Q_x in the collection $\{Q\}$ such that Q_x intersects none of the cubes. We know, however, that $Q \cap \cup Q_j$ is non empty, so (as before) there would be some j_0 for which $Q \subset 5Q_{j_0}$.

But then $\sum |Q_j| < \infty$ gives us that $\sum |5Q_j| < \infty$, eg, $|5Q_j| \to 0$ as $j \to \infty$, and thus $m^*(E \setminus \cup Q_j) \leq \sum_N^\infty |5Q_j|$, which tends to zero. □

3 Differentiation

3.1 Upper and Lower Derivatives

We now move on to looking at differentiability of functions, and sets where they are (or are not) differentiable. Given some function $f : [a, b] \to \mathbb{R}$ we define the **upper derivative** of f at a point x by $\overline{D}f(x) = \limsup_{h\to 0} \frac{f(x+h)-f(x)}{h}$ and the **lower derivative** of f at a point x by $\underline{D}f(x) = \liminf_h \frac{f(x+h)-f(x)}{h}$. Evidently a function is differentiable at a point x if $\overline{D}f(x) = \underline{D}f(x)$.

Claim 3.1. *For $f : I \to \mathbb{R}$, the upper and lower derivatives are measurable.*

Proof. Set $E = \{x : \overline{D}f(x) > a\}$. If we show that the set E is measurable, then $\overline{D}f(x)$ is a measurable function, and the same proof works in the case of $\underline{D}f$.

For h, k consider $E_{h,k} = \bigcup_i \{I : |I| \leq \frac{1}{k}, I = [\alpha, \beta], \frac{f(\beta)-f(\alpha)}{\beta-\alpha} > a + \frac{1}{h}\}$. We have that $E = \bigcup_h \bigcap_k E_{h,k}$, but $E_{h,k}$ is measurable as it is the union of intervals, and thus E is measurable. □

As an application of Vitali's covering theorem we can give a proof of Lebesgue's fundamental differentiation theorem:

Theorem 3.2. *Let $f : [a, b] \to \mathbb{R}$ be a non-decreasing function. Then $f'(x)$ exists almost everywhere, is non-negative wherever it exists and $\int_a^b f'(x)\,dx \leq f(b) - f(a)$.*

This is some form of the fundamental theorem of calculus. As an example where we get strict inequality take the distribution function of the variable which is chosen uniformly at random from the Cantor set. Clearly such a function has derivative zero everywhere outside the Cantor set, and the Cantor set has measure zero, eg, the integral of $f'(x)$ is zero, but $f(1) - f(0) = 1$.

Proof. We clearly have that $0 \leq \underline{D}f(x) \leq \overline{D}f(x) \leq \infty$. Our first goal is to establish that $\overline{D}f(x) < \infty$ almost everywhere.

Let $A = \{x : \overline{D}f(x) = \infty\}$. Fix some $N > 0$. If $x \in A$ then, clearly, $\overline{D}f(x) > N$, eg, there exists a sequence $x_j \to x$ such that $\frac{f(x_j)-f(x)}{x_j - x} > n$ for each j. Define I_j to be the interval between x and x_j. The collection $\{I_j(x)\}$ provides a fine covering of A, and thus there is a subcollection $I_j = [a_j, b_j]$ of disjoint intervals such that $|A \setminus \cup I_j| = 0$.

Now $|A| = |A \setminus \cup I_j| + |A \cap \cup I_j| \leq \sum |I_j|$. But $|I_j| \leq \frac{1}{N}(f(b_j) - f(a_j))$, and since f is decreasing, when we sum over them we get that this is $\leq \frac{1}{n}(f(b) - f(a))$, which converges to zero.

Our second step is to show that $\underline{D}f(x) = \overline{D}f(x)$ almost everywhere. Again this proceeds via an application of Vitali's theorem. Take $E = \{x : \underline{D}f(x) < \overline{D}f(x)\}$, and as per normal, consider $E_{s,t} = \{x : \underline{D}f(x) < s < t < \overline{D}f(x)\}$. We note that $E = \bigcup_{s<t\in\mathbb{Q}} E_{s,t}$, a countable union. So if we can show the measure of $E_{s,t}$ is zero for each s, t we will be done. Let $r = E_{s,t}$ and for some $\epsilon > 0$ pick some open set G that contains $E_{s,t}$ and $|G| < r + \epsilon$.

Since $\underline{D}f(x) = \liminf \frac{f(x+h)-f(x)}{h}$, for any $x \in E_{s,t}$ there is some sequence $x_j \to x$ such that $\frac{f(x_j)-f(x)}{x_j-x} < s$. As per the first part define intervals I_j that are the intervals between x_j and x. Since $E_{s,t} \subset G$ we can (without loss of generality) assume that each $I_j \subset G$. The collection $\{I_j(x)\}$ provides a fine cover of $E_{s,t}$, thus we can take some collection $I_j = [a_j, b_j]$ that are disjoint and $|E_{s,t} / \bigcup I_j| < \epsilon$.

We have that $0 < f(b_j) - f(a_j) \leq s|x_j - x| = s|I_j|$. Therefore $\sum f(b_j) - f(a_j) < s|\cup I_j| \leq s(r+\epsilon)$, as each $I_j \subset G$.

Let $A = E_{s,t} \cap \bigcup I_j$, eg, $|A| = r = |E_{s,t}|$. We now perform a similar argument using the upper derivative: as before there exists a sequence x_j such that $\frac{f(x_j)-f(x)}{x_j - x} > t$. As before there is a sequence $J_i(x)$ of intervals with endpoints x_i, x and this forms a fine cover of A. Pass to a sequence $J_i = [\alpha_i, \beta_i]$, and this is such that $|A / \cup J_i| = 0$ and $t|J_i| \leq f(\beta_i) - f(\alpha_i)$.

We note now that $tr = t|A| \leq t\sum |J_i| \leq \sum_i f(\beta_i) - f(\alpha_i)$ We rewrite this sum as first a sum over j and then a sum over i such that $I_j \supset J_i$.

This gives $tr \leq \sum_j \sum_{i: I_j \supset J_i} f(\beta_i) - f(\alpha_i) \leq \sum_j f(b_j) - f(a_j) \leq s(r+\epsilon)$. But since ϵ was arbitrary we get that $tr \leq sr$, and thus (since $t > s$) we get that $r = 0$. Thus f is differentiable almost everywhere.

Finally we need to show that $\int_a^b f'(x)\,dx \leq f(b) - f(a)$: we extend f so that $f(x) = f(b)$ for all $x > b$.

Let $f_n(x) = nf(x+1/n) - nf(x)$. We know that the limit almost everywhere of $f_n(x) = f'(x)$, and we have by Fatou's Lemma that $\int_a^b \liminf f_n \leq \liminf \int_a^b f_n$. The first of these is equal to $\int_a^b f'(x)$.

To bound the second of these, note that this equals

$$n\int_a^b f(x+1/n)\,dx - n\int_a^b f(x)\,dx = n\int_{a+1/n}^{b+1/n} f(x)\,dx - \int_a^b f(x)\,dx$$

by substitution. This is equal to

$$n\int_b^{b+1/n} f(x)\,dx - \int_a^{a+1/n} f(x)\,dx$$

The first integral here is a constant $f(b)$ and the second of these is bounded from above by $f(a)$ (as f is non decreasing), thus we have that $\liminf \int_a^b f_n(x)\,dx \leq f(b) - f(a)$, as required. □

3.2 The Fundamental Theorem of Calculus

The above shows us that it is non trivial to state precisely when the fundamental theorem of calculus holds: there is in general the inequality, but we need more hypotheses to get equality.

We say that a function $f : [a,b] \to \mathbb{R}$ is **absolutely continuous** (or $f \in AC[I]$) on $[a,b]$ if for all $\epsilon > 0$ there exists a δ such that for all partitions I_j such that $\sum |I_j| < \delta$ we have that $\sum |f(b_j) - f(a_j)| < \epsilon$. We note that this is a much stronger condition than continuity, and a function is absolutely continuous on an *interval* not at a *point*.

Claim 3.3. *If $f \in L^1[a,b]$ then $\int_a^x f(t)\,dt$ is absolutely continuous.*

Proof. For $f \in L^\infty[a,b]$ the result is obvious, namely, if $\|f\|_\infty < K$ and $|E| < \delta$ then $\int_E f\,dx < K\delta$.

For $f \notin L^\infty[a,b]$ we can use the fact that L^∞ is dense in L^1, namely, if we consider the functions $f_n(x) = f(x)$ for $|f(x)| \leq n$ and zero otherwise, then there is some n_0 such that $\|f - f_{n_0}\|_1 \leq \epsilon/2$. Then $\int_E f \leq \int |f - f_{n_0}| + \int_E f_{n_0}$, whence the result. □

This gives us a technique for showing that a function if of bounded variation: express $f(x) = \int_a^x F(t)\,dt$ for some function F.

We say that a function $f : [a,b] \to \mathbb{R}$ is of **bounded variation** on $[a,b]$ (or $f \in BV[I]$) if the function $V(f,[a,b]) = \sup_{a=x_0<x_1<\cdots<x_n=b} \sum |f(x_{i+1}) - f(x_i)|$ is bounded above by some absolute constant. A function is of bounded variation if and only if it can be written as the difference of two monotonic functions: evidently any monotonic function is of bounded variation, and if we have a function of bounded variation on $[a,b]$ then $f = V(f,[a,x]) - (V(f,[a,x]) - f)$ is such a decomposition.

Claim 3.4. *Any function that is absolutely continuous is of bounded variation.*

Proof. Take the δ corresponding to $\epsilon = 1$ in the definition. The total variation of f on any interval of length δ is at most 1, but then there are at most $\frac{b-a}{\delta} + 1$ such sub intervals, and thus we're done. □

Corollary 3.5. *Suppose $f \in AC[I]$ or $f \in BV[I]$. Then f' exists almost everywhere and $f' \in L^1$.*

Proof. f is the difference of two monotonic functions, say $m - n$. Differentiation is linear, and theorem 3.2 asserts that m' and n' exists almost everywhere, eg, f' exists almost everywhere, and is L^1 (as $\int |f'| \le \int |m'| - \int |n'|$). □

We are now able to state the conditions for the fundamental theorem of calculus, and give a proof:

Theorem 3.6. *If we have $f : [a,b] \to \mathbb{R}$ such that f' exists almost everywhere with $f' \in L^1$ then $f(x) = f(a) + \int_a^x f'(t)\,dt$ if and only if $f \in AC[I]$.*

Proof. $\implies$ is clear, f would necessarily be absolutely continuous by claim 3.3.

$\impliedby$ is harder: define $\varphi(x) = \int_a^x f'(t)\,dt - f(x)$. It is clear with what we have done so far that $\varphi' = 0$ almost everywhere, so the proof comes down to the following lemma, of interest in its own right. □

Lemma 3.7. *Suppose that φ is a function for which $\varphi' = 0$ almost everywhere. Then φ is constant.*

Proof. It is enough to show that $\varphi(a) - \varphi(b) = 0$ and then apply the result to any interval $[a,x]$ to get the result. Take $\epsilon > 0$. By the argument in theorem 3.2, for any $\epsilon > 0$ there are sets $I_j = [a_j, b_j]$ that are disjoint and $|\varphi(a_j) - \varphi(b_j)| < \epsilon |I_j|$, and $|I \setminus \cup I_j| = 0$, eg, $|I| = \sum |I_j|$. There exists a δ (corresponding to absolute continuity) such that $|I \setminus \bigcup_{j=1}^n I_j| < \delta(\epsilon)$. We have that $\varphi(a) - \varphi(b) \le \sum_j |\varphi(a_j) - \varphi(b_j)| \le \epsilon \sum |I_j|$, where the first inequality follows as a_j, b_j forms a partition of $[a,b]$. Taking ϵ to zero gives the result. □

We can also show the following slightly neat fact:

Claim 3.8. *Suppose that N is a null set. Then there is an absolutely continuous function f such that N is a subset of the points for which f' does not exist.*

Proof. Since N is null there is a sequence of open sets G_j such that $|G_j| \le \frac{1}{2^j}$ and $\cap G_j \supset N$. Set $f(x) = \sum |G_n \cap [a,x]| = \int_a^x \sum 1_{G_n}(t)\,dt$. Such a function is absolutely continuous, and evidently not differentiable for any $x \in N$: $\frac{f(x+h)-f(x)}{h} \ge N_h$ if $[x, x+h] \subset G_{N_h}$. Such N_h can be chosen to $\to \infty$ as $h \to 0$ (by simply taking the largest such N_h. □

3.3 Decomposition Theorems

The following section gives some examples of decomposition theorems.

Theorem 3.9. *Suppose that $f \in BV(I)$. Then there are functions g, h such that $g \in AC(I)$ and $h' = 0$ almost everywhere such that $g + h = f$. Moreover this decomposition is unique up to the addition of constants.*

Proof. To note the uniqueness: simply assume that there were two, $f = g_1 + h_1 = g_2 + h_2$. Then $g_1 - g_2 = h_2 - h_1$. Differentiating we get $g_1' - g_2' = h_2' - h_1' = 0$, as both h_1, h_2 have zero derivative. But this gives the result, by Lemma 3.7.

For existence: we know that for $f \in BV(I)$ we have that $f = m_1 - m_2$ with m_i monotone, eg, its derivative exists almost everywhere and is in L_1. Then we have, if $g = \int_a^x f'(t)\, dt$ that g is absolutely continuous and $(f - g)' = 0$ almost everywhere. □

Theorem 3.10. *For $f \in AC(I)$ we have that if $V(x) = V(f, [a, x])$ is absolutely continuous. Moreover, $V(x) = \int_a^x |f'|\, dt$.*

Proof. We have that

$$0 \leq V(\beta) - V(\alpha) = V(f, [\alpha, \beta])$$

by definition of the variation,

$$= \sup \sum |f(x_i) - f(x_{i-1})| = \sup \sum |\int_{x_{i-1}}^{x_i} f'(x)\, dx| \leq \sup \sum \int_{x_{i-1}}^{x_i} |f'(x)|\, dx = \int_\alpha^\beta |f'(x)|\, dx$$

Take some $[\alpha_j, \beta_j]$ a subpartition. Then $\sum |V(\beta_j - V(\alpha_j)| \leq \sum \int_{\alpha_j}^{\beta_j} |f'(x)|\, dx \leq \int_{\cup[\alpha_j,\beta_j]} |f'(x)|\, dx$. Since f' is in L^1 we have that for this union sufficiently small in measure, this is less than ϵ, thus recovering that $V(x)$ is in $AC(I)$.

We have also shown that $V' \leq |f'(x)|$. For the converse, note that $\frac{V(x+h)-V(x)}{h} = \frac{V(f,[x,x+h])}{h} \geq \frac{|f(x+h)-f(x)|}{h}$. Taking the limit gives us that $|V'(x)| \geq f'(x)$, as required. □

We now give the statement (and one proof) of some of the more important statements about decompositions of measures, starting with the famed Radon Nikodym theorem.

We say that a measure ν is **absolutely continuous** with respect to a measure μ if $\nu(A) = 0$ whenever $\mu(A) = 0$. We write $\mu \ll \nu$. We say that a measure ν is **singular** with respect to μ, written $\nu \perp \mu$, if $X = A \cup B$ with A, B disjoint and $\mu(A) = \nu(B) = 0$.

Theorem 3.11. *Suppose X is a σ finite measure space equipped with measures μ, ν such that ν is absolutely continuous with respect to μ. Then there exists some function $f \in L^1(\mu)$ such that $\nu(A) = \int_A f\, d\mu$.*

We will use this statement without proof to show the following result:

Theorem 3.12. *Suppose we have σ finite measures μ and ν. Then there exist unique measures ν_0 and ν_1 such that $\nu = \nu_0 + \nu_1$, $\nu_0 \ll \mu$ and $\nu_1 \perp \mu$.*

Proof. The uniqueness of the measures is obvious: they obviously live on disjoint subsets, and they can only be one thing on these disjoint subsets.

Let $\lambda = \mu + \nu$. We have that $\mu \ll \lambda$, and thus there exists some $f \in L^1$ such that $\mu(E) = \int_E f\, d\lambda$, where $f \geq 0$ everywhere. Let $A = \{f > 0\}$ and $B = \{f = 0\}$.

Set $\nu_0(E) = \nu(A \cap E)$ and $\nu_1(E) = \nu(B \cap E)$. We claim that this decomposition works: note that $\nu_1(A) = \nu(A \cap B) = 0$, eg, $\nu_1 \perp \mu$.

Now, if $\mu(E) = 0$ then $\lambda(E \cap A) = 0$, namely, $\nu_0(E) = 0$ as well. □

4 Sobolev Spaces

Our next topic to look at is something completely different: we look at Sobolev spaces. Given $f \in L^1_{loc}(\mathbb{R}^n)$ we wish to define some notion of a derivative on an expanded class of functions that coincides with the normal derivative where it exists. Moreover, we would like nice things to be true for such a function (eg, f can be recovered from its derivative.)

We fix some notation: set $\alpha = (\alpha_1, \dots, \alpha_n)$ to be a multi-index (eg, $\alpha_j \in \mathbb{N} \cup \{0\}$ and set $|\alpha| = \|\alpha\|_1$. Define $D^\alpha \varphi(x) = \frac{\partial^{|\alpha|}}{\partial^{\alpha_1}_{x_1} \dots \partial^{\alpha_n}_{x_n}} \varphi$.

We say that g is the α **weak derivative** of f (written $g = D^\alpha f$) if for all compactly supported smooth $\varphi(x)$ we have that $\int_{\mathbb{R}^n} (D^\alpha f) \varphi \, dx = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f(D^\alpha \varphi) \, dx$. Evidently these two notions coincide if the function is classically differentiable.

We note that this notion is independent of taking $\mathbb{R}^n$ as our source space, if we have any open set Ω we can instead look at weak derivatives on Ω, simply by taking $\varphi(x)$ compactly supported and smooth within Ω.

Claim 4.1. *Weak derivatives are unique.*

Proof. The main thing to check is that if $\int g(x)\varphi(x) \, dx = 0$ for all functions $\varphi(x)$ then g is zero. To do this we consider mollifications, take $\varphi_0 \in C_c^\infty$ such that the support of φ is in the unit ball, $\varphi \geq 0$ and $\int \varphi = 1$ (such functions exist.) Define $\varphi_\epsilon(x) = \epsilon^{-n}\varphi(x/\epsilon)$. These functions are chosen such that φ_ϵ is an approximation to the identity, ie, $\int_{\mathbb{R}}^n g(t)\varphi_\epsilon(x-t) \, dt \to g(x)$ almost everywhere as $\epsilon \to 0$. But this is as required. □

We define the **Sobolev Space** $W^{k,p}$ as the functions $f \in L^p$ for which $D^\alpha f$ exists (and is in L^p) for all $|\alpha| \leq k$. We equip this space with the norm $\|f\|_{k,p} = \sum_{|\alpha| \leq k} \|D^\alpha f\|_p$.

We note that the Sobolev spaces are Banach spaces, any Cauchy sequence in $W_{k,p}$ is such that each derivative is a Cauchy sequence in L^p. Take the limit and note that the definition of the weak derivative is continuous, so we are done.

We have some more basic facts about Sobolev spaces: The spaces $W^{k,p}$ are separable for $p \neq \infty$, simply due to the fact that they are subsets of $L^p(\mathbb{R}^n)$ which is separable (eg, because they are.)

If f has compact support and is absolutely continuous, then f' (which a priori we know exists and is L^1) is equal to the weak derivative.

If we have $f, g : \mathbb{R} \to \mathbb{R}$ functions of compact support with g being the weak derivative of f, then there is some $\tilde{f} \in AC$ which is absolutely continuous. The idea is that $\tilde{f} = \int_{-\infty}^x g(t) \, dt$, however we just need to check that this works by mollifying: set f_ϵ and g_ϵ to be convolutions of f, g with φ_ϵ. We can take a subsequence converging almost everywhere and in L^1. Then f_ϵ and g_ϵ are both C_c^∞. Our claim is now that $g_\epsilon = f'_\epsilon$: but $f'_\epsilon = \int_{-\infty}^\infty f(t)\varphi'_\epsilon(x-t)dt = \int_{-\infty}^\infty g(t)\varphi_\epsilon(x-t) \, dt = g_\epsilon(t)$. Now, by the Fundamental Theorem of Calculus, we have that for any two points x_1, x_2 that $f_\epsilon(x_2) - f_\epsilon(x_1) = \int_{x_1}^{x_2} g_\epsilon(t) \, dt$. Letting $\epsilon \to 0$ we get $f(x_2) = f(x_1) + \int_{x_1}^{x_2} g(t) \, dt$

wherever x_2 and x_1 are points where the convergence was almost everywhere. But, as we claimed above, this is absolutely continuous.

4.1 Poincare Inequalities and Imbedding Theorems

Consider a function $f : \mathbb{R}^n \to \mathbb{R}$ that is smooth. We wish to be able to recover f from its derivative, eg, similar to the case of $\mathbb{R}$ where $f(x) = \int_{-\infty}^{x} f'(t)\, dt$.

We write ∇f to be the vector $(\partial_{x_1} f, \ldots, \partial_{x_n} f)$.

Theorem 4.2. *Suppose we have a smooth function $f : \mathbb{R}^n \to \mathbb{R}$. Then $f(x) = c_n \int_{\mathbb{R}^n} \frac{\nabla f(x-y)\cdot y}{|y|^n}\, dy$, where c_n is the reciprocal of the surface measure of the sphere.*

Proof. We aim to prove the result by induction: starting at $n = 1$. For $n = 1$ note that $f(x) = \int_{-\infty}^{x} f'(u)\, du = \int_0^\infty f'(x-t)\, dt$. For t positive we have that $1 = \frac{t}{|t|}$, so we have that $f(x) = \int_0^\infty f'(x-t)\frac{t}{|t|}\, dt$.

Similarly, $f(x) = -\int_x^\infty f'(u)\, du = -\int_{-\infty}^0 f'(x-t)dt = \int_{-\infty}^0 f'(x-t)\frac{t}{|t|}\, dt$. Adding these together gives $f(t) = \frac{1}{2}\int_{-\infty}^\infty f'(t)\frac{t}{|t|}\, dt$ as required.

For the general case $n = k$, consider for some fixed x and some fixed $\xi \in S^n$ the function $g(u) = f(x + t\xi)$. This is a function from $\mathbb{R}$ to $\mathbb{R}$ so we have that $f(x) = g(0) = \int_0^\infty g'(-t)\, dt = \int_0^\infty \nabla f(x - t\xi)\cdot \xi\, dt$. Integrating this expression in ξ over the unit sphere and making the switch to polar co-ordinates we get the result. □

We can prove another result which gives us a bound on the norm of the function in terms of bounds on the derivatives:

Theorem 4.3. *If $f \in C_c^\infty(\mathbb{R}^n)$ and $\frac{1}{q} = 1 - \frac{1}{n}$ then we have that $\|f\|_{L^q} \le \left(\prod_{j=1}^n \|\partial_{x_i} f\|_{L^1}\right)^{1/n} \le \frac{1}{n}\sum \|\partial_{x_i} f\|_{L^1}$.*

Proof. We proceed by induction: we know that $f(x) = \int_{-\infty}^x f'(t)\, dt$, eg, $\|f\|_\infty \le \|f'\|_1$.

Write $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$ and decompose $x \in \mathbb{R}^n$ into (x', x_n). Define $I_j(x) = \int_{\mathbb{R}^{n-1}} |\partial_{x_j}(x', x)|\, dx'$ and $I_n(x') = \int_{\mathbb{R}} |\partial_{x_n} f(x', x)| dx$.

The inductive hypothesis gives that $\left(\int_{\mathbb{R}^{n-1}} |f(x', x_n)|^q\, dx'\right)^{1/q'} \le (\prod_{j=1}^{n-1} \|\partial_{x_i} f|)^{1/n-1}$ where $\frac{1}{q'} = 1 - \frac{1}{n-1}$. By our definitions we have the second of these $= \prod_{j=1}^{n-1} I_j(x_n)^{1/n-1}$.

We have that $\int_{\mathbb{R}^{n-1}} |f(x', x_n)|^q dx' = \int_{\mathbb{R}^{n-1}} |f(x', x_n)|^{q-1} |f(x', x_n)|\, dx$. We apply Holders inequality, and use the fact that $|f(x', x_n)| \le I_n(x')$ to get

$$\le \left(\int_{\mathbb{R}^{n-1}} I_n(x_n) dx'\right)^{\frac{1}{n-1}} \left(\int_{\mathbb{R}^{n-1}} f(x', x_n)^{q'} dx'\right)^{\frac{1}{q'}}$$

But now integrating with respect to x_n gives the result. □

The previous inequality is known as a Poincare inequality or a Sobolev inequality. The general form of the Sobolev inequality is known as the Sobolev-Gagliardo-Nirenberg inequality:

Claim 4.4. *There is a constant c such that the mapping from $W^{1,p}(\mathbb{R}^n) \to L^{p^*}(\mathbb{R}^n)$ is a continuous (but not compact) embedding, where $p^* = \frac{n-p}{np}$.*

We will not give a proof of this here: the proof goes by density, which will take us a while to develop.

Lemma 4.5. *For $1 \leq p < \infty$ we have that $f \in W^{k,p}(\mathbb{R}^n)$ if and only if there is a sequence $f_n \in C_c^\infty$ such that $f_n \to f$ in L^p and $D^\alpha f_n$ is Cauchy in L^p.*

Proof. For $\Longleftarrow$ note that f would then be in $W^{k,p}$ by completeness.

For $\Longrightarrow$ we have to work harder. Take φ_ϵ an approximation to the identity, and set $f_\epsilon = f * \varphi_\epsilon$. We have that $f_\epsilon \in C^\infty \cap L^p$ and that $f_\epsilon \to f$ in L^p.

We claim that $D^\alpha f_\epsilon = D^\alpha f * \varphi_\epsilon$, but this is true because: $D^\alpha f_\epsilon = f * D^\alpha \varphi_\epsilon = \int f(t) D^\alpha \varphi_\epsilon(x - t)\,dt = (-1)^{|\alpha|} \int f(t)(D^\alpha \varphi_\epsilon)(x-t)\,dt = (-1)^{2|\alpha|} \int D^\alpha f(t) \varphi_\epsilon(x-t)\,dt$ as required. So, $D^\alpha f_\epsilon \to D^\alpha f$ as $\epsilon \to 0$.

Let $\theta(x)$ be in $C_c^\infty(\mathbb{R}^n)$ with $\theta(0) = 1$. Let $f_{\epsilon,\delta}(x) = \theta(\delta x) f_\epsilon(x)$. We now claim that there exists some ϵ_j and δ_j as per the theorem.

Note that for any δ there is some ϵ_j such that $\|D^\alpha f_{\delta,\epsilon_j} - D^\alpha f\| \leq 1/j$. Moreover, there is some δ_j such that $\|D^\alpha f - D^\alpha(f\theta(\delta x))\|_p \leq 1/j$, as this equals $\int |D^\alpha f|^p |1 - \theta(\delta x)|^p\,dx$, but the term $|D^\alpha f| M$ where $M = \sup |1 - \theta(\delta x)|^p$ is a dominating term in L^1, so by LDCT we are done. □

Our goal now is to prove a theorem involving oscillation, which we define as follows: the p-**oscillation** of a function is defined as $\omega_p(f,t) = \left(\int |f(x+t) - f(x)|^p\,dx\right)^{1/p}$, this is some measure of continuity.

Theorem 4.6. *Suppose that $1 < p < \infty$ and $\omega_p(f,t) = O(|t|)$ (eg, $\frac{\omega_p(f,t)}{t} \leq C$ for some absolute constant C, for all t sufficiently small). Then the first order weak derivatives of f exist and are p integrable.*

We need to mention two facts about weak convergence. We say that a sequence $f_n \to f$ weakly (written $f_n \rightharpoonup f$) in a Banach space X if $f(x_n) \to f(x)$ for any $f \in X^*$.

In L^p note that the dual is L^q so we need that $\int f_n g\,dx \to \int f g\,dx$ for all $g \in L^q$.

Lemma 4.7. *If f_j is a sequence in L^p such that $\|f_j\| \leq K$ then there is a subsequence f_{n_j} such that $f_{n_j} \rightharpoonup f$ for some $f \in L^p$ (ie, there is a weakly convergent subsequence.)*

Proof. L^p is reflexive, therefore the unit ball is weakly compact. Weak compactness implies weak sequential compactness by the Eberlein Smulian theorem, thus there is a weak convergent subsequence. □

Lemma 4.8. *If $f_n \to f$ in L^p then $f_n \rightharpoonup f$.*

Proof. This is true in any Banach space: suppose $x_n \to x$. Then $|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\| \|x_n - x\| \to 0$. □

Proof of Theorem 4.6. We need to show that there are some f_j such that for each $\varphi \in C_c^\infty(\mathbb{R}^n)$ we have that $\int f \partial_{x_j} \varphi\,dx = -\int f_j \varphi\,dx$.

Let e_j be the j'th canonical basis vector of $\mathbb{R}^n$ and take $F_m(x) = m(f(x + e_j/m) - f(x))$. We know that $\|F_m\|_p$ is bounded (by hypothesis) and thus there exists a weakly convergent subsequence. Let f_j denote the limit of the weakly convergent subsequence.

Now $\int m(f(x+e_j/m)-f(x))\varphi(x)\,dx = \int f(x)m(\varphi(x-e_j/m)-\varphi(x))\,dx$. Taking the limit using DCT we get that this equals $-\int f(x)\partial_{x_j}\varphi(x)\,dx$. Taking the limit in the first expression, using the definition of weak limits, gives $\int f_j\varphi\,dx$, as required. □

4.1.1 Embedding Theorems

Theorem 4.9. *Let $k \in \mathbb{N}$, and look at $W^{k,p}(\mathbb{R}^n)$. There are three different behaviours depending upon the value of p:*

- *If $1 \leq p < n/k$, and $1/q = 1/p - n/k$ then $W^{k,p} \hookrightarrow L^q(\mathbb{R}^n)$.*
- *If $p = n/k$ then for any compact set k, $f|_K \in \bigcap r \geq nL^r(\mathbb{R}^n)$.*
- *If $p > n/k$ then there is a continuous function $\tilde{f}$ such that $f = \tilde{f}$ almost everywhere.*

The value $p = n/k$ is variously called things such as the critical exponent. The value of q in the first part is called the Sobolev exponent.

Here (I think) is the first time we use the following notation: $A \subset\subset B$. This means that A is compact and $A \subset B$, we say A is compactly supported in B.

Proof. We first show the results for $f \in W^{1,p}(\mathbb{R}^n)$. Once we have shown it for this, the other cases will follow inductively (essentially, $f \in W^{k,p} \implies f, \partial_{x_j}f \in W^{k-1,p}$.)

Case 1: We show that for $f \in W^{1,p}(\mathbb{R}^n)$ that there is some constant c such that $\|f\|_{L^q} \leq c\|\nabla f\|_{L^p}$.

We showed that for $g \in C_c^\infty(\mathbb{R}^n)$ that $|f(x)| \leq c\int_{\mathbb{R}^n} |\nabla f(x-y)|\frac{1}{|y|^{n-1}}\,dy$. We also showed that if $q = \frac{n}{n-1}$ then $\|f\|_{L^q} \leq \frac{1}{n}\sum \|\partial_{x_j}f\|_{L^1} \leq c\|\nabla f\|_{L^1}$.

We know (by density) that there is some sequence of functions $f_m \in C_c^\infty(\mathbb{R}^n)$ such that $f_m \to f$ in $W^{1,p}$, so it suffices to show that $\|f_m - f\| \to 0$ in L^q, then the results shown will suffice.

We show that it is Cauchy, $\|f_k - f_m\| \leq c\|\nabla(f_k - f_m)\|_{L^q} \to 0$ by the first part. The derivatives converge in L^p so are also Cauchy, eg, the sequence is Cauchy. Thus $\|f_m\|_q \leq c\|\nabla f_m\|_p$, taking the limit on both sides yields the result.

Case 2: Set $p = n$ and we need to show that for K compact and $n \leq r < \infty$ that $f \in W^{1,n}(\mathbb{R}^n)$ implies that $f \in L^r$.

Take $\eta \in C_c^\infty(\mathbb{R}^n)$ with $\eta(x) = 1$ on K. We show that $\eta \cdot f \in L^r$. Take $f_m \in C_c^\infty$ such that $f_m \to f$ in $W^{1,p}$. Then $\eta f_m \to \eta f$ in $W^{1,p}$.[10]

Let $K_1 \subset\subset \operatorname{supp}\eta$. We have $r_0 > 0$ such that $K_1 - K \subset r_0 B_{\mathbb{R}^n}$.

For some $x \in K$ we have that $|\eta f_m(x)| \leq c\int_K |\nabla(\eta f_m)(x-y)|\frac{dy}{y^{n-1}}$. Apply Young's inequality, $\frac{1}{r} = \frac{1}{s} + \frac{1}{n} - 1$ and thus $\frac{1}{s} = 1 + \frac{1}{r} - \frac{1}{n} > 1 - \frac{1}{n} = \frac{n-1}{n}$. So we have that $s < \frac{n}{n-1}$. Also $s > 1$, eg, $1 \leq s < \frac{n}{n-1}$.

Therefore, $\|\eta f_m\|_{L^r} \leq c\|\nabla(\eta f_m)\|_{L^n}\left(\int_{|y|\leq r_0} \frac{dy}{|y|^{s(n-1)}}\,dy\right)^{1/s}$ $(\star)$. The integral is bounded by some constant.

[10] This follows from either the product rule or some variant of the product rule, which works here because η is smooth.

Note that ηf_m is Cauchy in L^r by the same argument, applying the inequality $\star$ to $(\eta(f_m - f_k))$. Thus, ηf_m (which is certainly in L^r) is Cauchy in L^r, thus $\eta f_m \in L^r$.

Case 3: Suppose that $p > n/k$. Fix K compact $\subset \mathbb{R}^n$. Then for $x \in K$ we have that $|\eta f_m(x)| \leq \int_{|y| \leq y_0} |\nabla \eta f_m(x-y)| \frac{dy}{y^{n-1}}$. By Holder this is $\leq \left(\int_{|y| \leq y_0} \frac{dy}{|y|^{p'(n-1)}}\right)^{1/p} \|\nabla(\eta f_m)\|_{L^p}$. But this is uniformly bounded, eg, the sequence ηf_m is Cauchy in the uniform norm, thus $\eta f_m \to \eta f$ uniformly, eg, the limit function (which is guaranteed to be continuous) is equal to ηf almost everywhere.

Inductive step: Assume the result is true for $1, 2, \ldots, k-1$ and $f \in W^{k,p}$. Then $f, \partial_{x_i} f$ are all in $W^{k-1,p}$.

Thus for the case 1 we have that $1 \leq p \leq \frac{n}{k} < \frac{n}{k-1}$, eg, $f, \partial_{x_i} f \in L^{q'}(\mathbb{R}^n)$, where $\frac{1}{q'} = \frac{1}{p} - \frac{k-1}{n}$ is the Sobolev exponent for $k-1$. But then $f \in W^{1,q'}(\mathbb{R}^n)$, thus $f \in L^q(\mathbb{R}^n)$ by the case one.

Similar considerations suffice for cases two and three. $\square$

We have an important result here that is necessary to use at some points.

Lemma 4.10. *Suppose $f \in L^p \cap L^q$ where $p < q$ (on any domain). Then $f \in L^r$ for $p \leq r \leq q$ inclusive.*

Proof. By Holder. $\square$

We can, in fact, get stronger results in case 3 above. Case 3 gives us that there are continuous representatives. The following theorem tells us that these representatives are Holder continuous for a certain coefficient:

Theorem 4.11. *If $f \in W^{1,p}(\mathbb{R}^n)$ for $p > n$ then $\exists u$ such that $u = f$ almost everywhere such that there exists some constant c with $|u(x) - u(y)| \leq c\|x - y\|^\lambda$ for $\lambda = 1 - \frac{n}{p}$.*

If $f \in W^{k,p}$ we can apply this process to the k'th derivatives of f to get us that the k'th derivatives are λ Holder continuous, where $\lambda = k - [k - \frac{n}{p}]$ where [] denotes the greatest integer less than. This follows inductively in a similar way to the final part of Proof 4.9.

Proof. Assume that $f \in C_c^\infty(\mathbb{R}^n)$. We will show that there is some universal constant c such that $|f(x) - f(y)| \leq c|x - y|^\lambda \|f\|_{W^{1,p}}$.

Suppose that $|x - y| \leq \sigma$ and let Q_σ denote a cube of side length σ such that $x, y \in Q_\sigma$. Take any $z \in Q_\sigma$.

Our main goal in this proof will be to bound the quantity that we will call $(\dagger) = |f(x) - \frac{1}{\sigma^n}\int_{Q_\sigma} f(z)\,dz| \leq \frac{1}{\sigma^n}\int_{Q_\sigma} |f(x) - f(z)|\,dz$. We bound the inner term now.

We have that $f(x) = f(z) - \int_0^1 \frac{d}{dt}(f(x + t(z-x)))dt = \int_0^1 \nabla f(x + t(z-x)) \cdot (z-x)\,dt$.

Then $|f(x) - f(z)| \leq |z - x| \int_0^1 \nabla f(x + t(z-x))\,dt \leq \sqrt{n}\sigma \int_0^1 \nabla f(x + t(z-x))\,dt$.

Thus we have that $(\dagger) \leq \frac{\sqrt{n}}{\sigma^{n-1}} \int_{Q_\sigma}\int_0^1 |\nabla f(x + t(z-x))|\,dtdz$. Change the order of integration by Fubini, and introduce a new variable, to get that this is $= \frac{\sqrt{n}}{\sigma^{n-1}} \int_0^1 t^{-n} \int_{Q_{t\sigma}} 1 \cdot |\nabla f(v)|\,dv\,dt$.

But this is bounded by $\frac{\sqrt{n}}{\sigma^{n-1}} \int_0^1 t^{-n} \|f\|_{W^{1,p}} |Q_{t\sigma}|^{1/p'} dt$ where we have used Holder on 1 and $|\nabla f(v)|$. But now we can perform this integral to get that this is bounded above by (where C is some absolute constant depending on n) $C\sqrt{n}\sigma^{1-n/p}\|f\|_{W^{1,p}}$.

But then $|f(x) - f(y)| \le 2C\sqrt{n}\sigma^{1-n/p}\|f\|_{W^{1,p}}$, by the triangle inequality. □

4.2 Exponential L^p spaces

We noted in Theorem 4.9 that there were functions which belong to every L^p space for sufficiently large p. However, we should observe that such functions needn't be in L^∞: consider $f(x) = \log(1/x)$. This is evidently unbounded on $(0, 1/2)$ but we have that $\int_0^{1/2} f^n(x)\,dx$ exists for all large enough n (eg, it exists for $n = 2$, in this case we can just do the integral.)

We are thus interested in some finer distinction between functions with these properties.

To this end we define exponential L^p spaces as follows: for a compact set K we say that $f \in \exp(L^s(K))$ if there is some $\alpha > 0$ such that $\int_K \exp(\alpha|f(x)|^s)\,dx < \infty$ for some α. These spaces aren't really nice spaces, there's no obvious norm for them.

However, if we unpack the definition we get that $\sum \frac{\alpha^n}{n!}\|f1_K\|_{sn}^{sn} < \infty$, eg, we are in every L^{sn} space, and whats more, we do it in a bounded way. This means that we're in every L^r space (by interpolation) for sufficiently large r.

This is what we mean by 'finer distinction': we are lying between functions that lie in every L^n space and L^∞.

We also have that $\exp(L^s(K))$ is not necessarily those functions that are in every L^p space. This should be evident, we have that the sum needs to *converge* and not simply the norms exist. We have to come up with a function for which the sn norm grows faster than $\frac{n!}{\alpha^n}$ for any α. The function $(\log(1/x))^2$ on $(0, 1/2)$ works.

So now the question is what conditions do we need so that the norms existing for all r give us that we're in the exponential space. The next lemma gives a sufficient condition:

Lemma 4.12. *Let $f : \mathbb{R}^n \to \mathbb{R}$ such that $\|f1_K\|_{L^r} = O(r^{1/s})$. Then $f \in \exp(L^s(K))$.*

Proof. We need to exhibit an α such that $\sum \frac{\alpha^n}{n!}\|f1_k\|_{sn}^{sn} < \infty$. We know that there is some constant M such that $\|f1_K\|_{sn}^{sn} \le M(sn)^n$ for n sufficiently large.

Thus we get that the sum $\sum \frac{\alpha^n}{n!}\|f1_K\|_{sn}^{sn} \le c\sum \frac{(\alpha M^s)^n s^n n^n}{n!}$. By Stirlings formula this bounded by $\le c\sum \frac{(\alpha M^s se)^n}{\sqrt{2\pi n}}$ (perhaps up to a modification in c.) Thus if α is sufficiently small this converges. □

We can also state that Sobolev functions are in these spaces for some value of s:

Theorem 4.13. *If $f \in W^{k,n/k}(\mathbb{R}^n)$ for K compact then $f \in \exp(L^{n/(n-1)}(K))$.*

Proof. For $n \le r < \infty$ we have that $\|\eta f_m\|_{L^r} \le c\|\nabla\eta f_m\|_{L^r}\left(\int_0^{r_0} \frac{\rho^{n-1}}{\rho^{(n-1)s}}\,d\rho\right)^{1/s}$. We compute the integral to get the decay listed in the previous lemma. □

4.3 Sobolev Spaces on open subsets of $\mathbb{R}^n$

We define $W^{k,p}(\Omega) = \{f \in L^p(\Omega)$:all weak derivatives up to order k exist and are $\in L^p(\Omega)\}$.

These spaces are separable for $1 \le p < \infty$, reflexive for $1 < p < \infty$ and Hilbert if $p = 2$ with inner product $\int fg dx + \sum\int \partial_i f \partial_i g\,dx$

There is an equivalent definition of $W^{1,p}(\Omega)$ in terms of the distributions, we say $f \in W^{1,p}(\Omega)$ if $f \in L^p(\Omega)$ and the distributional derivatives are also in $L^p(\Omega)$.

The spaces $W^{1,p}(\Omega)$ also play nicely with regular derivatives, eg, $C_1(\bar{\Omega}) \subset W^{1,p}(\Omega)$ and if f is both $W^{1,p}(\Omega)$ and $C(\Omega)$ with $\partial_i f \in C(\Omega)$ then $f \in C(\Omega)$. Classical derivatives agree with weak derivatives if they exist.

We now embark upon stating and proving density theorems, a task which will take a while.
Theorem 4.14. *Suppose that $\Omega \subset \mathbb{R}^n$ is an open set. Let $f \in W^{1,p}(\Omega)$, for $1 \leq p < \infty$. Then there is a sequence of $u_n \in C_c^\infty(\mathbb{R}^n)$ such that $u_n|_\Omega \to f \in L^p(\Omega)$ and $\nabla u_n|_\Omega \to \nabla f$ for each $\omega \subset\subset \Omega$.*

Note that $u_n \in C_c^\infty(\mathbb{R}^n)$ and not $C_c^\infty(\Omega)$. In general we can not demand this.

Proof. Extend f by zero outside of Ω and denote the extension by $\bar{f}$. Let $f_\epsilon = \eta_\epsilon * \bar{f}$, where η_ϵ is a mollifier. Then $f_\epsilon \to \bar{f}$ in $L^p(\mathbb{R}^n)$. We need to show that for every $\omega \subset\subset \Omega$ that $\nabla f_\epsilon|_\omega \to \nabla f|_\omega$ in $L^p(\omega)$.

Take $\zeta \in C_c^1(\Omega)$ such that $\zeta(z) = 1$ on some ϵ' neighbourhood of ω, and consider $\overline{\zeta f}$. Then $\eta_\epsilon * \overline{\zeta f} = \eta_\epsilon * f$ on ω for sufficiently small ϵ.

Now $\partial_{x_i}(\eta_\epsilon * \overline{\zeta f}) = \eta_\epsilon * \overline{\zeta \partial_{x_i} f + f \partial_{x_i} \zeta}$ by the product rule. This converges as $\epsilon \to 0$ to $\zeta \partial_{x_i} f + f \partial_{x_i} \zeta$ which converges to $\partial_{x_i} f$ in $L^p(\omega)$. Now take some sequence ω_n that exhausts Ω. This suffices. $\square$

There are two other density results. We will not prove one of them here, and the other will follow from considerations later on:
Theorem 4.15. *For $\Omega \subset \mathbb{R}^n$ open, and $f \in W^{1,p}(\mathbb{R}^n)$ then there exist $u_n \in C^\infty(\Omega) \cap W^{1,p}(\Omega)$ such that $u_n \to f$ in $W^{1,p}(\Omega)$.*

This is close to best possible, in the sense that generally without some restriction on $\partial\Omega$ we can not have approximations in $C_c^1(\mathbb{R}^n)$. We detail these restrictions now.
Theorem 4.16. *If Ω is open and bounded, and $\partial\Omega$ is of class C^1 then $\forall f \in W^{1,p}$ for $1 \leq p < \infty$ there exists $u_n \in C_c^\infty(\mathbb{R}^n)$ such that $u_n \to f$ when restricted to Ω.*

We will now detail exactly what this theorem means.

4.4 Class C^1

Let Ω be an open subset of $\mathbb{R}^n$. We say Ω is of class C^1 if for all $x \in \partial\Omega$ there is an open neighbourhood of x denoted N_x and a map $H : Q \to N_x$ (where $Q =$ the solid cylinder, the set of $x = (x', x_n)$ of $\mathbb{R}^n$ for which $\|x'\| \leq 1$ and $|x_n| \leq 1$) for which:

- H is a diffeomorphism, eg, $H \in C^1(\Omega)$ and $H^{-1} \in C^1(\Omega)$ (H thus is bijective)
- $H(Q \cap \mathbb{R}^{n,+}) = \Omega \cap N_x$ (eg, it maps one side of the cylinder to one side of Ω.
- $H(Q \cap \{x_n = 0\}) = N_x \cap \partial\Omega$ (eg, it maps an $n-1$ plane to the boundary.)

We could also demand (in what follows) that $\partial\Omega$ is Lipschitz, ie for all $x \in \partial\Omega$ there is h and $H : \mathbb{R}^{n-1} \to \mathbb{R}$ such that (up to isometries) we have that $Q(x,h) \cap \Omega = \{y : H(y_1, \ldots, y_{n-1}) < y_n\} \cap Q(x,h)$. We are not going to use this definition again.

4.5 Characterizations of Sobolev spaces

We now give a characterization of $W^{1,p}(\Omega)$:

Theorem 4.17. *Consider the following hypotheses:*

1. $f \in W^{1,p}(\Omega)$
2. *There exists a c such that for all* $\varphi \in C_c^\infty$ *we have* $\int f \partial_{x_i} \varphi \, dx \leq c\|\varphi\|_{L^q}$ *where q is the Holder dual exponent.*
3. *There exists a c such that* $\forall \omega \subset\subset \Omega$ *and* $|h| < d(\omega, \partial\Omega)$ *we have that* $\|\tau_h f - f\|_{L^1(\omega)} \leq c|h|$.

If $p = 1$ *then* $1 \implies 2 \iff 3$. *If* $1 < p < \infty$ *then* $1 \iff 2 \iff 3$.

Proof. One implies two is clear.

For two implies one: consider the map $p_i : C_c^\infty \to \mathbb{R}$ given by $\varphi \mapsto \int f \partial_{x_i} \varphi \, dx$. We know that this is bounded, so by Hahn-Banach we can extend the map to all of L^q. But then since the dual of L^p is L^q in this range we have that the extension can be written in the form $\int g_i \varphi \, dx$. But this says that $\int g_i \varphi \, dx = \int f \partial_{x_i} \varphi \, dx$. This is exactly the definition of the weak derivative.

For one implies three: take $f \in C_c^\infty(\mathbb{R}^n)$ and consider for any h the function $v(t) = f(x + th)$. We have that $f(x + h) - f(x) = v(1) - v(0) = \int_0^1 v'(s) \, ds = \int_0^1 h \cdot \nabla f(x + th) \, dt$. Thus $|\tau_h f - f|^p \leq \int_0^1 |h|^p |\nabla f(x + th)|^p \, dt$ by Holders inequality (with one function being chosen to be 1, which integrates to 1.)

Integrating this expression over ω we get that $\|\tau_h f - f\|_{L^p}^p \leq \int_\omega \int_0^1 |h|^p |\nabla f(x + th)|^p \, dt \, dx$. We then do our regular trick of interchanging integration to get $= |h|^p \int_0^1 \int_{\omega+th} |\nabla f|^p \, dt$.

We now appeal to density, there exists some $f_n \in C_c^\infty(\mathbb{R}^n)$ such that $f_n \to f$ in L^p and $\nabla f_n \to \nabla f$ in $\omega \subset\subset \Omega$. Therefore $\|\tau_h f - f\|_{L^p(\omega)}^p \leq |h|^p \int_{\omega^prime} |\nabla f|^p \, dt$ where ω' chosen such that $\omega' \subset\subset \Omega$ with $\omega' + th$ still inside Ω. Taking the limit as $n \to \infty$ we get 1 implies 3.

Now we show 3 implies 2. Take $\varphi \in C_c^\infty$. We know that $|\int (\tau_h f - f)\varphi \, dx| \leq c\|h\|\|\varphi\|_{L^q}$ by Holders inequality, eg, $\int (f(x + h) - f(x))\varphi(x) \, dx = \int f(y)(\varphi(y - h) - \varphi(y)) \, dy$. Taking absolute value and putting this inside the integral and taking the limit (setting $h = te_i$) gives us the result (we end up differentiating φ.)

Finally, 2 implies 3. This is already done for $p \neq 1$, so the only thing we need to do is test it in the case $p = 1$.

Fix ω, h appropriately. There exists some $\varphi \in C_c^\infty(\omega)$ such that $\|(f(x + h) - f(x))(1 - \omega)\|_1 < \epsilon$ with $\|\varphi - 1_\omega\|_1 < \epsilon$. Then $|\int_\omega (f(x + h) - f(x))\varphi \, dx| = |\int_{\omega-h} f(x)(\varphi(x - h) - \varphi(x)) \, dx| = |\int_{\omega-h} f(x) \int_0^1 \nabla\varphi(x - th) \cdot h dt \, dx| \leq |h| \int_0^1 \int_{\omega-h} |f(x)||\nabla\varphi(x - th)| \leq c|h||\nabla\varphi|$. Fixing $|\nabla\varphi| \leq 1$ we are done. ☐

We can get a characterization of $W^{1,\infty}$.

Corollary 4.18. *If* $f \in L^\infty(I)$ *then* $f \in W^{1,\infty}(I)$ *if and only if there exists a c such that* $|f(x + h) - f(x)| \leq c|h|$ *for almost all* x, y.

For any n if $\Omega \subset \mathbb{R}^n$ *then* $f \in W^{1,p}(\Omega)$ *if and only if* $|f(x) - f(y)| < cd(x, y)$, *where* $d(x, y)$ *is the geodesic distance in* Ω.

But this says exactly that such functions have continuous representatives.

For any $p \in [1, \infty]$ we have that if $\nabla f = 0$ almost everywhere on Ω then f is constant almost everywhere on connected components (as we would expect).

4.6 Compactness Results

We give a (brief) explanation of compactness results, giving the statement of the L^p case and then giving as a corollary the statement of the result in $W^{k,p}$.

The following is known as the **Riesz Frechet Kolmogorov** theorem:

Theorem 4.19. *Suppose A is a set in $L^p(\mathbb{R}^n)$ which is uniformly bounded in norm. Then A is relatively compact if and only if $\int_{|x|>r} |f|^p \, dx \to 0$ uniformly in A and $\lim_{a \to 0} \|\tau_a f - f\| = 0$ uniformly in A.*

Corollary 4.20. *If $\Omega \subset \mathbb{R}^n$ is open and B is the unit ball of $W^{1,p}(\Omega)$ for $1 \leq p \leq \infty$ then B is relatively compact in $L^p(\omega)$ for every $\omega \subset\subset \Omega$.*

Proof. This follows simply from Riesz Frechet Kolmogorov, noting that the in B the second term tends to zero because $\|\nabla f\| \leq 1$. □

This can be strengthened in the case that $\partial\Omega$ is of class C^1 and bounded, then B is precompact in $L^p(\Omega)$. This is false if Ω has unbounded boundary or is not regular enough.

4.7 Extension Theorems

If Ω is of C^1 and $\partial\Omega$ is bounded, then there exists a bounded linear operator $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ that extends f and $\|Ef\| \leq K\|f\|$.

If Ω is also bounded, we can prescribe the support on any $\Omega' \supset\supset \Omega$.

We give the main lemma that proves the result here but don't give the actual result, due to it being very fiddly and involving partitions of unity more than anything else.

So, we give the result in the case of a function $f \in W^{1,p}$ defined on Q^+, the set of points $x = (x', x_n)$ where $|x'| \leq 1$ and $0 < x_n < 1$.

Once we have sketched the proof of this lemma, all that is left to do is apply it at every 'local cylinder' that exists because the domain is C^1. So, after this we just glue them together using C^∞ partitions of unity.

In our case we consider the function $f^*(x) = f(x', x_n)$ if x_n is positive and $f^*(x', -x_n)$ if x_n is negative. We need to check that such an extension works, eg, $\|f^*\|_{W^{1,p}(Q} \leq 2\|f\|_{W^{1,p}(Q^+)}$. Consider a function $\theta(t) \in C^\infty$ given by $\theta(t) = 0$ for $t < 1/2$ and 1 for $t > 1$. Then, given a test function $\varphi(t) \in C_c^1(Q)$ we consider the function $\psi(x', x_n) = \varphi(x', x_n) - \varphi(x', -x_n)$. Then we look at $\theta(kt)\psi(x)$. This function has the property that we've 'smoothed out' the bad plane of the function f^*. But, when we take the limit, we recover the integral.

So then $\int_Q f^* \frac{\partial \varphi}{\partial x_i} \, dx = \int_{Q^+} f \frac{\partial \psi}{\partial x_i} \, dx$, eg, we get that (for $1 \leq i \leq n-1$) $\left(\frac{\partial f^*}{\partial x_i}\right) = \left(\frac{\partial f}{\partial x_i}\right)^*$. Similar arguments prove that $\frac{\partial f^*}{\partial x_n} = \pm \frac{\partial f}{\partial x_n}$, where the $\pm$ is whether x_n is above zero or below zero. This gives us that $\|f^*\|_{W^{1,p}(Q)} \leq 2\|f\|_{W^{1,p}(Q^+)}$.

This (finally) gives us the result of Theorem 4.16: given any Ω which is open, bounded and of class C^1 we can extend functions on it to $\mathbb{R}^n$. But then by Theorem 4.14 we can use that there exist functions $v_k \to f$ in $W^{1,p}(\omega)$ for all $\omega \subset\subset \mathbb{R}^n$. But then $\Omega \subset \Omega' \subset\subset \mathbb{R}^n$, so we're done.

We can, in fact, remove the hypotheses of boundedness by taking $\zeta = 1$ on B_X and 0 outside of $2B_X$. Then the extension operator applied to $\zeta(x/n)f$ works: we map into a bounded domain first.

4.7.1 Chain Rule

For $f \in W^{1,p}(\Omega)$, Ω as in the extension theorem, $g \in C_1(\mathbb{R})$ g' bounded we have that $\frac{\partial}{\partial x_i}(g \circ f) = g' \cdot \frac{\partial f}{\partial x_i}$ where we evaluate g' at the place we'd expect it (namely $f(x)$).

For $f \in W^{1,p}$ we thus have $f^\pm$ and $|f| \in W^{1,p}(\Omega)$, where $\nabla f^+ = \nabla f$ ae on $\{f \geq 0\}$ and 0 on $\{f \leq 0\}$ (same for ∇f^-.)

We also have a Jacobian rule: if $\Omega, \Omega' \subset \mathbb{R}^n$, H is a diffeomorphism between Ω' and Ω, $f \in W^{1,p}(\Omega)$ and the Jacobian of both H, H^{-1} are bounded, then $f \circ H \in W^{1,p}(\Omega')$ with $\frac{\partial}{\partial y_i}(f \circ H)(y) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}|_{H(y)} \frac{\partial H_i}{\partial y_j}$.

4.7.2 Imbedding Theorem

Because we now have the results of extension, we can formulate an imbedding theorem for $\Omega \subset \mathbb{R}^n$ (which we did not have before):

Corollary 4.21. *Suppose we have Ω that satisfies the hypotheses of the extension theorem. Then:*

- *For $p < n$ we have that $W^{1,p}(\Omega) \subset L^{p^*}(\Omega)$ for $p^* = \frac{np}{n-p}$ where the embedding is continuous.*
- *For $p = n$, $W^{1,p}(\Omega) \subset L^q(\Omega)$ for all $p < q < \infty$.*
- *For $p > n$ we have that $W^{1,p}(\Omega) \subset L^\infty(\Omega)$ and for each f there is a Holder continuous representative with exponent $1 - \frac{n}{p}$, where $|f(x) - f(y)| \leq c\|f\|_{W^{1,p}(\Omega)}|x-y|^\alpha$.*

Proof. Extend and then apply the imbedding theorem in $\mathbb{R}^n$. □

4.8 Poincare Inequalities and Traces

For $1 \leq p < \infty$ and Ω an open subset of $\mathbb{R}^n$ we want to look at functions that vanish on $\partial\Omega$. However, most generally, $\partial\Omega$ is a set of measure zero, so there's no way of defining this. We define $W_0^{1,p}(\Omega) = \overline{C_c^1(\Omega)}^{W^{1,p}(\Omega)}$, in some sense, these are functions that vanish on the boundary (or, at least, they can be well approximated by things that vanish on the boundary.)

Evidently, if $\Omega = \mathbb{R}^n$ we have that $W^{1,p}(\mathbb{R}^n) = W_0^{1,p}(\mathbb{R}^n)$. However, in other cases this is not true.

This space is a separable Banach space (for $1 \leq p < \infty$) when we equip it with the subspace norm. We write H_0^1 to be the space $W_0^{1,2}$, a Hilbert space.

We list some basic facts about functions of this type:

Theorem 4.22. *If $f \in W^{1,p}(\Omega)$ and $1 \leq p < \infty$ with the support of f compact, then $f \in W^{1,p}(\Omega)$.*

If Ω is of class C^1 and $f \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$ then $f = 0$ on $\partial\Omega \iff f \in W_0^{1,p}(\Omega)$.

The following fact generalizes Theorem 4.17:

Theorem 4.23. *If Ω is of class C^1 for $1 < p < \infty$, and $f \in L^p(\Omega)$ then $f \in W_0^{1,p}(\Omega) \iff$ there exists a constant c such that $|\int_\Omega f \frac{\partial \varphi}{\partial x_i} \, dx| \leq c\|\varphi\|_{L^q}$ where q is the conjugate index to p, $\iff$ $F(x) = f(x)$ for $x \in \Omega$ and 0 for $x \notin \Omega$ is in $W^{1,p}(\mathbb{R}^n)$.*

We denote the conditions I, II and III respectively.

Proof. $I \implies II$: There exist functions $f_n \in C_c^1(\mathbb{R}^n)$ with $f_n \to f$ in $W^{1,p}(\mathbb{R}^n)$. Then $|\int f_n \partial_{x_i} \varphi \, dx| = |\int \partial_{x_i} f_n \varphi \, dx| \leq \|\nabla f_n\|_p \|\varphi\|_q$. Taking the limit gives the result.

$II \implies III$: For $\varphi \in C_c^1(\mathbb{R}^n)$ we have that $\int_{\mathbb{R}^n} F \partial_{x_i} \varphi| = |\int_\Omega f \partial_{x_i} \varphi| \leq c\|\varphi\|_{L^q(\Omega)} \leq c\|\varphi\|_{L^q(\mathbb{R}^n)}$, so theorem 4.16 takes it.

$III \implies I$: we truncate, as in the proof of density 4. □

The proof of Sobolev-Gagliardo-Nirenberg thus generalizes to $W_0^{1,p}$.

We now prove a Poincare inequality:

Theorem 4.24. *If Ω is bounded and open, $f \in W_0^{1,p}(\Omega)$ then $\|f\|_{L^p(\Omega)} \leq c\|\nabla f\|_{L^p(\Omega)}$.*

Proof. This follows as a corollary of the previous theorem. □

4.8.1 Duality Results

There is no simple characterization of $W^{1,p}(\Omega)^*$: we do not get anything like $W^{k,p'}$ which we've come to expect (for some different k perhaps.)

However, if we have $W_0^{1,p}(\Omega)$ we can take the dual to get $W^{-1,p'}(\Omega)$. We have not looked at negative index sobolev spaces yet, however, we can think of them in two ways, either generalizing the Fourier transform relationship or to think of it in the following way: $u \in \mathcal{D}'(\Omega)$ such that $u = \sum_{|\alpha| \leq 1} \partial_i^\alpha u_\alpha$ where $u_\alpha \in L^p$, eg, we take things in D' (the dual of distributions) which are at most 1 derivative of L^p functions.

So we can view any $T \in W_0^{1,p}(\Omega)^*$ as the extension to $W^{1,p}(\Omega)$ of a distribution $\tau \in \mathcal{D}'(\Omega)$. An extension to $W^{1,p})_0(\Omega)$ is unique, due to our definition of $W_0^{1,p}(\Omega)$ we have a version of density. Without density there need not be a unique extension, this (essentially) comes down to the fact that for functionals in $W^{1,p}(\Omega)$ they can take both the values of the function in the interior of Ω (which is what the distribution behaviour comes down to) and the values of the function on the boundary, which is the concept developed in the next section.

4.8.2 Imbeddings of $W_0^{1,p}(\Omega)$

When Ω is bounded and of class C^1 we, in fact, get stronger embeddability results. These are mainly related to the compactness of the embeddings, namely, the results of corollary 4.23 hold, except the embeddings are compact (in certain cases). More accurately:

Theorem 4.25.
- *For $p < n$ we have that $W^{1,p}(\Omega) \subset L^{p^*}(\Omega)$ for $p^* = \frac{np}{n-p}$ where the embedding is continuous. Moreover, for $L^p(\Omega)$, for $p < p^*$ we have that the embedding is compact.*
- *For $p = n$, $W^{1,p}(\Omega) \subset L^q(\Omega)$ for all $p < q < \infty$, and this embedding is compact.*
- *For $p > n$ we have that $W^{1,p}(\Omega) \subset L^\infty(\Omega)$ and for each f there is a Holder continuous representative with exponent $1 - \frac{n}{p}$, where $|f(x) - f(y)| \leq c\|f\|_{W^{1,p}(\Omega)}|x-y|^\alpha$ and this embedding is compact.*

4.8.3 Trace

Trace gives us a way of defining Sobolev functions on the boundary. We will do the case of $\mathbb{R}^n_+$ which extends to C^1 functions through the use of local charts.

The operator $T : C^1_c(\mathbb{R}^n_+) \to C^1_c(\mathbb{R}^{n-1})$ is defined by $T(f) = f|_{\{x_n=0\}}$ can be extended by density to a continuous bounded linear operator from $W^{1,p}(\mathbb{R}^n_+)$ to $L^p(\mathbb{R}^{n-1})$. This is called the **trace** operator, and for $\Gamma = \partial\Omega$ it defines $f|_\Gamma$ in a natural way. For $f \in L^p(\Omega)$ there's no way of doing this, we don't have enough information to define f on the boundary: essentially this follows because the boundary is a null set.

We have some properties of the Trace operator that we will not prove:

- If $f \in W^{1,p}(\Omega)$ then $f|_\Gamma \in W^{1-1/p,p}(\Omega)$, where we take the fractional Sobolev index here. Moreover, this embedding is continuous, namely, $\|f\|_{W^{1-1/p,p}(\Gamma)} \leq c\|f\|_{W^{1,p}(\Omega)}$.[11]
- We have that $W^{1,p}_0(\Omega) = \{f \in W^{1,p}(\Omega) : f|_\Gamma = 0\}$.
- We get the more general form of integration by parts this way: $\int_\Omega \partial_i f g\, dx = -\int_\Omega f\partial_i g + \int_\Gamma fg(n \cdot e_i)\, dx$ (where n is the surface normal.)

4.8.4 Hausdorff Measure

The Hausdorff measure gives us a way of understanding trace more generally.

The $s-$dimensional **Hausdorff Measure** $\mathcal{H}^s$ is defined on any set A to be

$$\mathcal{H}^s(A) = \liminf_{\delta \to 0} \left[\sum \alpha(s) \left(\frac{\text{diam}(C_j)}{2} \right)^s |A \subset \cup_{j=1}^\infty C_j, \text{diam}(C_j) < \delta \right]$$

We denote the thing we are taking the limit of as $\mathcal{H}^s_\delta(A)$.

In particular, $\mathcal{H}^0$ is the counting measure. It is also evident that $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}(A)$, and that $\mathcal{H}^s(OA) = \mathcal{H}^s(A)$ where O is any isometry of $\mathbb{R}^n$.

We now make various (unproven) statements about the Hausdorff measure:

- $\mathcal{H}^s$ is Borel, regular, not Radon, and for any $s < n$ we have that $(\mathbb{R}^n, \mathcal{H}^s)$ is not σ finite.
- $\mathcal{H}^n$ is Lebesgue measure on $\mathbb{R}^n$: the case of $\mathbb{R}$ is easier.
- $\mathcal{H}^s(A) = 0$ for any $s > n$ and any $A \subset \mathbb{R}^n$.

[11] The fractional Sobolev spaces are naturally defined by the Fourier Transform.

- The behaviour of $\mathcal{H}^s$ has a phase transition: it's infinite for all sufficiently small s and zero for all sufficiently large s. Moreover, if we are finite at one point, we are zero for every point after, and infinite for every point before. For a given set A the critical value of s is called the Hausdorff dimension of A.

Now, for a function $f \in W^{1,p}(\Omega)$ we have that the trace operator maps as follows: $T : W^{1,p}(\Omega) \to L^p(\partial\Omega, \mathcal{H}^{n-1})$. Moreover, this mapping is continuous.