3. Poincaré's inequality

A continuously differentiable function is approximated by an affine function up to a quantifiable error. This is the definition and also the content of Taylor's theorem. Sobolev functions do not admit such a simple pointwise approximability, but they are still approximated by polynomials at average. The quantitative statement of this property is Poincaré's inequality, which we can also use to characterize weak differentiability. Given $k \geq 0$, let $\mathcal{P}_k$ be the family of all polynomials of degree at most k, k included.

Theorem 3.1 (Poincaré inequality). *Let $k \geq 0$ and $p \geq 1$. Then there is a constant C such that for all $f \in W^{k,p}_{loc}(\Omega)$ and all balls $B(x, 2r) \subset \Omega$*

$$\inf_{\pi \in \mathcal{P}_{k-1}} \left(⨍_{B(x,r)} |f(y) - \pi(y)|^p \, dy \right)^{1/p} \leq C r^k \left(⨍_{B(x,r)} |\nabla^k f(y)|^p \, dy \right)^{1/p}.$$

The claim also holds with the Euclidean balls $B(x, r)$ replaced by the cubes $Q(x, r)$ centered at x and having side length r.

If f were a $C^k_{loc}(-1, 1)$ function, the claim would follow immediately by testing the left hand side with an order $k - 1$ Taylor polynomial:

$$f(x) - \left(f(0) + f'(0)x + \frac{1}{2} f''(0)x^2 \right) = \int_0^x \int_0^{s_3} \int_0^{s_2} f^{(3)}(s_1) \, ds_1 ds_2 ds_3$$

for $k = 3$ and similarly for higher orders. The case of a Sobolev function follows by approximation. One just has to pick a Taylor polynomial of a suitably good approximant and show that it will do for the final inequality. The multidimensional case could follow the same reasoning, but passing from line integrals to a solid integral on the right hand side is more involved. However, we can circumvent all this by the following (weird) trick. Chapter 3 in [2] shows a few more of its applications.

Proof. Without loss of generality, we may assume $r = 1$ and $x = 0$. Once we prove the claim for one ball, the inequality for other balls follows by translation and dilation invariance of the inequality. Suppose the claim were false. Then there existed a sequence of functions f_j such that

$$\inf_{\pi \in \mathcal{P}_{k-1}} \left(⨍_{B(0,1)} |f_j(x) - \pi(x)|^p \, dx \right)^{1/p} > j \|f_j\|_{\dot{W}^{k,p}(B(0,1))}.$$

For each j, let $\pi_j \in \mathcal{P}_{k-1}$ be a polynomial attaining the minimal value on the left hand side such that

$$\left(⨍_{B(0,1)} |f_j(x) - \pi_j(x)|^p \, dx \right)^{1/p} \geq j \|f_j\|_{\dot{W}^{k,p}(B(0,1))}.$$

Let $g_j = \lambda_j (f_j - \pi_j)$ where $\lambda_j > 0$ is chosen so that $\|g_j\|_{L^p(B(0,1))} = 1$ for all $j \geq 0$. By construction

$$1 = \|g_j\|_{L^p(B(0,1))} \geq j \|g_j\|_{\dot{W}^{k,p}(B(0,1))}.$$

By weak sequential compactness of the closed unit ball of $L^p(B(0, 1))$, we can find $g \in L^p(B(0, 1))$ so that passing to a subsequence, $g_j \to g$ weakly in $L^p(B(0, 1))$. In particular, we know that $\partial^\alpha g$ with $|\alpha| = k$ exist as distributions and $\partial^\alpha g_j \to \partial^\alpha g$ in the sense of distributions. As $j \|g_j\|_{\dot{W}^{k,p}(B(0,1))}$ stays bounded while $j \to \infty$, we see that

$$\lim_{j \to \infty} \|g_j\|_{\dot{W}^{k,p}(B(0,1))} = 0.$$

Hence $\partial^\alpha g = 0$ for all multi-indices α with $|\alpha| = k$. It follows that g is a polynomial of degree at most $k-1$. By repeated application of Lemma 1.11, we see that g_j has a subsequence that converges in $\dot{W}^{k-1,p}(B)$ norm and further a subsequence that converges in $L^p(B)$ norm. The construction then implies

$$1 = \|g\|_{L^p(B(0,1))} = \lim_{i \to \infty} \|g_i\|_{L^p(B(0,1))} = \lim_{i \to \infty} \inf_{\pi \in \mathcal{P}_{k-1}} \|g_i - \pi\|_{L^p(B(0,1))} \leq \lim_{i \to \infty} \|g_i - g\|_{L^p(B(0,1))} = 0,$$

a contradiction. □

When r is small, the factor r^k on the right hand side is very small and when r is large, the estimate is not so favorable. We can describe this alternatively in terms of the left hand side belonging to $L^q(\Omega)$ with $q > p$. By testing such an inequality with functions $f(\delta x)$ and sending $\delta \to 0$ and $\delta \to \infty$, one sees that the condition

$$\frac{k}{n} = \frac{1}{p} - \frac{1}{q}$$

is necessary for the inequality. Such an estimate with $L^q(\Omega)$ and $L^p(\Omega)$ norms is the content of a Sobolev–Poincaré inequality.

We introduce a bit of notation before proving the Sobolev–Poincaré inequality. We start with the dyadic cubes. Consider a fixed cube Q_0, which we may assume to be the unit cube centered at the origin for simplicity. We let $\mathcal{D}_j(Q_0)$ with $j \geq 0$ be the 2^{jn} congruent cubes with side length 2^{-j} that cover Q_0. We let $\mathcal{D}(Q_0) = \bigcup_{j=0}^{\infty} \mathcal{D}_j$. This is the collection of dyadic cubes. If two cubes $Q, Q' \in \mathcal{D}$ meet in a set of positive n dimensional Lebesgue measure, then they are equal or one is contained in the other.

We also need a few properties of the polynomials. Consider the inner product

$$(f,g) \mapsto \int_{Q_0} f(x)g(x)\,dx$$

on $\mathcal{P}_{k-1}(Q_0)$, the polynomials of degree at most $k-1$ restricted to Q_0. We can apply the Gram–Schmidt algorithm on the set of monomials of total degree at most $k-1$ to form a finite and orthonormal basis $\{\phi_\nu\}_\nu$. We define P_{Q_0} as a projection $L^1(Q_0) \to \mathcal{P}_{k-1}(Q_0)$ using the inner product as before and this basis by setting

$$P_{Q_0} f(x) = \sum_\nu \langle f, \phi_\nu \rangle \phi_\nu(x).$$

We define the projections relative to other cubes $L^1(Q) \to \mathcal{P}_{k-1}(Q)$ by the same way but now using the differently normalized inner product

$$(f,g) \mapsto \frac{1}{|Q|} \int_Q f(x)g(x)\,dx.$$

Define

$$M_P f(x) = \sup_{Q \in \mathcal{D}} 1_Q(x) |P_Q f(x)|.$$

The case $k = 1$ gives the mean value, which is a good example to keep in mind. The projections leave polynomials in $\mathcal{P}_{k-1}(Q_0)$ invariant. Inspecting the explicit formula of the projection, one sees that

$$|P_Q f(x)| \leq C \fint_Q |f(y)|\ dy, \quad M_P f(x) \leq C M_\Omega f(x). \tag{3.1}$$

In particular, the Lebesgue differentiation theorem applies to polynomials. The convergence

$$\lim_{|Q| \to 0, Q \ni x} P_Q f(x) = f(x)$$

takes place for almost every $x \in Q_0$.

Theorem 3.2 (Sobolev–Poincaré inequality)**.** *Let $0 < k < n/p$ and $q = pn/(n-kp)$. Then there is a constant C such that for each $f \in W^{k,p}_{loc}(\Omega)$ and each cube $Q(x, 2r) \subset \Omega$ there is a polynomial $\pi \in \mathcal{P}_{k-1}$ with*

$$\left(\fint_{Q(x,r)} |f(y) - \pi(y)|^q\,dy \right)^{1/q} \leq C r^k \left(\fint_{Q(x,r)} |\nabla^k f(y)|^p\,dy \right)^{1/p}.$$

Proof. We first prove

$$\|f - \pi\|_{L^{q,\infty}(Q(x,r))} \leq C \|\nabla^k f\|_{L^p(Q(x,r))}. \tag{3.2}$$

By translation and dilation invariance of the local inequality, we may assume $x = 0$ and $r = 1$. Let π be a polynomial as on the left hand side of the inequality in Theorem 3.1 with $p = 1$ and set

$$E_\lambda = \{x \in Q_0 : |f(y) - \pi(y)| > \lambda\}.$$

Let $\varepsilon > 0$ and $\lambda > 0$. Define

$$E_{\lambda,-} = \left\{x \in Q_0 : M^k_\Omega |\nabla^k f|(x) \leq \varepsilon\lambda\right\},$$
$$E_{\lambda,+} = \left\{x \in Q_0 : M^k_\Omega |\nabla^k f|(x) > \varepsilon\lambda\right\}.$$

The inclusion up to a set of measure zero defines a partial order on dyadic cubes. Let $\mathcal{M}$ be the family of dyadic cubes Q maximal with

$$\inf_{y \in Q} |P_Q(f - \pi)(y)| = \inf_{y \in Q} |P_Q f(y) - \pi(y)| > \lambda.$$

The infimum is zero for Q_0, and hence the Q_0 cannot be maximal. The maximal cubes are necessarily pairwise disjoint, and by the Lebesgue differentiation theorem $\mathcal{M}$ covers $E_\lambda \supset E_{2\lambda}$. Let $\mathcal{M}_-$ be the subfamily of cubes in $\mathcal{M}$ that meet $E_{\lambda,-}$, and let $\mathcal{M}_+ = \mathcal{M} \setminus \mathcal{M}_-$. By Theorem 2.4, it clearly holds

$$\sum_{Q \in \mathcal{M}_+} |Q| \leq |\{x \in Q_0 : M^k_\Omega(\nabla^k f)(x) > \varepsilon\lambda\}| \leq (\varepsilon\lambda)^{-q} \|M^k_\Omega(1_{Q_0}\nabla^k f)\|^q_{L^{q,\infty}(\Omega)} \leq C(\varepsilon\lambda)^{-q} \|\nabla^k f\|^q_{L^p(Q_0)}. \tag{3.3}$$

Consider then $Q \in \mathcal{M}_-$. Let $\hat{Q}$ be the unique dyadic cube with $\hat{Q} \supset Q$ and $|\hat{Q}| = 2^n|Q|$. By definition, for all $y \in Q \cap E_{2\lambda}$

$$2\lambda < |P_Q(f-\pi)(y)| \le |P_Q(f - P_{\hat{Q}}f)(y)| + |P_Q(P_{\hat{Q}}f(y) - \pi)(y)| \le |P_Q(f - P_{\hat{Q}}f)(y)| + \lambda$$

so that by equation (3.1) and Theorem 2.3

$$|Q \cap E_{2\lambda}| \le |\{x \in Q : M_P(1_Q(f - P_{\hat{Q}}f))(x) > \lambda\}| \le \frac{C}{\lambda}\int_Q |f(x) - P_{\hat{Q}}f(x)|\,dx. \tag{3.4}$$

By Theorem 3.1, the definition of the fractional maximal function and the fact that $Q \in \mathcal{M}_-$

$$\frac{C}{\lambda}\int_Q |f(x) - P_{\hat{Q}}f(x)|\,dx \le \frac{C}{\lambda}\cdot \ell(Q)^k \int_{\hat{Q}} |\nabla^k f(x)|\,dx \le \frac{C|Q|}{\lambda}\cdot \inf_{y\in Q} M^{k,d}(\nabla^k f)(y) \le C\varepsilon|Q|. \tag{3.5}$$

Summing equation (3.4) over $Q \in \mathcal{M}_-$, using equation (3.5) and combining with (3.3), we see that

$$\sum_{Q\in\mathcal{M}} |Q| \le C(\varepsilon\lambda)^{-q}\|\nabla^k f\|_{L^p(Q_0)} + C\varepsilon \sum_{Q\in\mathcal{M}_-} |Q|$$

Choosing ε small enough so as to move the right most sum to the left, we obtain

$$|\{x \in Q_0 : |f(y) - \pi(y)| > 2\lambda\}| \le \sum_{Q\in\mathcal{M}} |Q| \le C\lambda^{-q}\|\nabla^k f\|_{L^p(Q_0)}$$

whence the claimed inequality (3.2) follows by multiplying by λ^q and taking supremum over $\lambda > 0$.

To conclude the proof, we notice that f and $g = |\nabla^k f|$ satisfy the assumptions of Proposition 2.5. Hence the weak Sobolev–Poincaré inequality implies a strong Sobolev–Poincaré inequality. □

Now we can read the global Sobolev embedding as a corollary.

Corollary 3.3 (Sobolev embedding)**.** *Let $1 \le p < \infty$ and $k \ge 0$. Then there is a constant C such that if $f \in W^{k,p}(\mathbb{R}^n)$ and $0 \le k < n/p$, then*

$$\|f\|_{L^q(\mathbb{R}^n)} \le C\|f\|_{\dot{W}^{k,p}(\mathbb{R}^n)}, \quad \frac{k}{n} = \frac{1}{p} - \frac{1}{q}.$$

Proof. This follows by taking $r \to \infty$ in Theorem 3.2. For each $r > 0$ there is an almost optimal polynomial π_r with

$$\left(\int_{Q(x,r)} |f(y) - \pi_r(y)|^q\,dy\right)^{1/q} \le C\left(\int_{\mathbb{R}^n} |\nabla^k f(y)|^p\,dy\right)^{1/p}.$$

Take a sequence $r_i \to \infty$ and let $R > 0$. Then $1_{B(0,R)}\pi_{r_i}$ is bounded in $L^q(\mathbb{R}^n)$, and it has a weakly convergent subsequence. By a diagonal argument, π_{r_i} has a weakly convergent subsequence. But a sequence of degree $k-1$ polynomials can only converge if all its coefficients converge. Then the limit is another polynomial and the convergence takes place almost everywhere. The only polynomial in $L^q(\mathbb{R}^n)$ is the zero polynomial, and the claim then follows by Fatou's lemma. □

Next we discuss the converse of Poincaré's inequality. If a function satisfies a Poincaré inequality with locally integrable right hand side, then we can conclude it is weakly differentiable.

Theorem 3.4. *Let $D \ge 1$ and $k \ge 1$. There exists a constant C such that if $f \in L^1_{loc}(\Omega)$, $g \in L^1_{loc}(\Omega)$ is non-negative, and*

$$\inf_{\pi\in\mathcal{P}_{k-1}} \fint_{B(x,r)} |f(y) - \pi(y)|\,dy \le r^k \fint_{B(x,Dr)} g(y)\,dy$$

holds for all $B(x, 2Dr) \subset \Omega$, then f is k times weakly differentiable and $|\nabla^k f(x)| \le Cg(x)$ for almost every $x \in \Omega$.

Proof. As f is a distribution, it suffices to show its distributional derivatives are locally integrable functions. Let η be a smooth and non-negative function of total mass one. Assume that $\operatorname{supp}\eta \subset B(0,1)$. For $\varepsilon > 0$, let $\eta_\varepsilon(x) = \varepsilon^{-n}\eta(\varepsilon^{-1}x)$. For $j \ge 1$, let $\Omega_j = \{x \in \Omega : \operatorname{dist}(x, \Omega^c) > 1/j\}$. Then $\eta_\varepsilon * f \to f$ in the sense of $C_c^\infty(\Omega_j)$ for any j.

Let $\alpha \in \mathbb{N}^n$ be a multi-index with $|\alpha| = k$. Then by partial integration

$$\int \partial^\alpha \eta_\varepsilon(x)\cdot \pi(x)\,dx = 0$$

for all polynomials $\pi \in \mathcal{P}_{k-1}$. For any such polynomial π

$$|\eta_\varepsilon * \partial^\alpha f(x)| = |\partial^\alpha \eta_\varepsilon * f(x)| = \frac{1}{\varepsilon^{n+1}}\left|\int_{B(0,\varepsilon)} (\partial^\alpha\eta)_\varepsilon(x-y)(f(y) - \pi(y))\,dy\right| \le \frac{C}{\varepsilon^{n+1}}\int_{B(x,\varepsilon)} |f(y) - \pi(y)|\,dy.$$

Taking infimum over $\pi \in \mathcal{P}_{k-1}$ and applying the assumption, we obtain

$$|\eta_\varepsilon * \partial^\alpha f(x)| \le C \fint_{B(x,D\varepsilon)} |g(y)|\, dy.$$

Let $K \subset \Omega$ be compact and let $\varphi \in C_c^\infty(K)$. Then

$$|\langle \partial^\alpha f, \varphi\rangle| = \lim_{\varepsilon\to 0} |\langle \eta_\varepsilon * \partial^\alpha f, \varphi\rangle| \le C \lim_{\varepsilon \to 0} \int_K \fint_{B(0,D\varepsilon)} |g(y-x)\varphi(x)|\, dydx \le C \left(\int_K f(x)\, dx\right) \|\varphi\|_\infty$$

and hence the linear functional $\partial^\alpha f$ is bounded in supremum norm. As the smooth functions are dense in continuous functions, we conclude that $\partial^\alpha f$ extends to an element of $C_c(K)^*$ for all K, that is, $\partial^\alpha f$ is a signed Radon measure. By the control through g, the Lebesgue differentiation theorem and the Radon–Nikodym theorem, $\partial^\alpha f$ is a measurable and further locally integrable function subject to the claimed bound. ☐

We can go one step further and give a pointwise characterization of weak differentiability. This is a direct consequence of the Poincaré inequality. The main advantage of the pointwise characterization is that it can be used to extend the notion of Sobolev space to low-dimensional, possibly fractal subsets of $\mathbb{R}^n$. First we state the condition that follows if weak differentiability is assumed.

Proposition 3.5. *Let $\alpha \in [0,1)$. There is a constant C depending on the dimension and α such that if $f \in W^{1,1}_{loc}(\Omega)$, then for almost every $(x,y) \in \Omega^2$ with $B(x,2|x-y|) \Subset \Omega$*

$$|f(x) - f(y)| \le C|x-y|^{1-\alpha}(M_c^\alpha(1_{B(x,2|x-y|)}|\nabla f|)(x) + M_c^\alpha(1_{B(y,|x-y|)}|\nabla f|)(y)).$$

Proof. Fix the points x and y and denote $R = |x-y|$. By Lebesgue differentiation theorem and Poincaré inequality for $z \in \{x,y\}$

$$\begin{aligned}|f(z) - f_{B(z,R)}| &\le \sum_{k=1}^\infty |f_{B(z,2^{-k}R)} - f_{B(z,2^{-k+1}R)}| \le \sum_{k=1}^\infty \fint_{B(z,2^{-k}R)} |f(\zeta) - f_{B(z,2^{-k+1}R)}|\, d\zeta \\ &\le C\sum_{k=1}^\infty 2^{-k+1}R \fint_{B(z,2^{-k+1}R)} |\nabla f(\zeta)|\, d\zeta \le C\left(\sum_{k=1}^\infty (2^{-k+1}R)^{1-\alpha}\right) M_c^\alpha(1_{B(z,R)}|\nabla f|)(z) \\ &= CR^{1-\alpha} M_c^\alpha(1_{B(z,R)}|\nabla f|)(z).\end{aligned}$$

Using the fact $B(x,R) \cup B(y,R) \subset B(x,2R)$ and Poincaré's inequality, we see

$$\begin{aligned}|f_{B(x,R)} - f_{B(y,R)}| &\le |f_{B(x,R)} - f_{B(x,2R)}| + |f_{B(x,2R)} - f_{B(y,R)}| \le 2^{n+1} \fint_{B(x,2R)} |f(\zeta) - f_{B(x,2R)}|\, d\zeta \\ &\le CR \fint_{B(x,2R)} |\nabla f(\zeta)|\ d\zeta \le CR^{1-\alpha} M_c^\alpha(1_{B(x,2R)}|\nabla f|)(x)\end{aligned}$$

so that

$$\begin{aligned}|f(x) - f(y)| &\le |f(x) - f_{B(x,R)}| + |f_{B(x,R)} - f_{B(y,R)}| + |f_{B(x,R)} - f(y)| \\ &\le CR^{1-\alpha}(M_c^\alpha(1_{B(x,2R)}|\nabla f|)(x) + M_c^\alpha(1_{B(y,R)}|\nabla f|)(y)).\end{aligned}$$

This concludes the proof. ☐

As a corollary, we can deduce Morrey's inequality.

Corollary 3.6. *Let $p > n$ and $f \in W^{1,p}_{loc}(\Omega)$. Then (up to a choice of representative) $f \in C^{1-n/p}_{loc}(\Omega)$.*

More precisely, for every open $U \Subset \Omega$ let $d := \operatorname{dist}(U,\Omega^c)/6$. Then there is a constant with

$$\|f\|_{C^{1-n/p}(K)} \le C(p,n,\operatorname{diam}(U),d)\|f\|_{W^{1,p}(U+B(0,5d))}.$$

Proof. Let $U \Subset \Omega$ be open and $K = \overline{U}$. There is $r > 0$ and a finite set of points $\{x_i\}_{i=1}^N$ such that the balls $B(x_i,r)$ with $B(x_i,6r) \Subset \Omega$ cover K. Fix one i. Apply Proposition 3.5 with $\alpha = n/p$ to notice that for almost every pair $x,y \in B(x_i,r)$

$$|f(x) - f(y)| \le C|x-y|^{1-n/p}(M_c^{n/p}(1_{B(x,2|x-y|)}|\nabla f|)(x) + M_c^{n/p}(1_{B(y,|x-y|)}|\nabla f|)(y))$$

with C only depending on dimension and p. By Hölder's inequality

$$\begin{aligned}M_c^{n/p}(1_{B(x,2|x-y|)}|\nabla f|)(x) + M_c^{n/p}(1_{B(y,|x-y|)}|\nabla f|)(y) &\le C\left(\int_{B(x,2|x-y|)\cup B(y,|x-y|)} |\nabla f(z)|^p\ dz\right)^{1/p} \\ &\le C\left(\int_{B(x_i,5r)} |\nabla f(z)|^p\ dz\right)^{1/p}.\end{aligned}$$

Hence for almost every $x, y \in K$

$$|f(x) - f(y)| \le NC|x-y|^{1-n/p} \left(\int_{K+B(0,5r)} |\nabla f(z)|^p \ dz \right)^{1/p}.$$

It follows that the limit

$$\lim_{\delta \to 0} \fint_{B(x,\delta)} f(z)\, dz$$

exists everywhere in K and defines a Hölder continuous function that coincides with the original f almost everywhere. We identify f with this limit from now on. Then also for all $x \in K$

$$\begin{aligned} |f(x)| &\le \fint_{B(x,r)} |f(z)|\, dz + \fint_{B(x,r)} |f(z) - f(x)|\, dz \\ &\le \fint_{B(x,r)} |f(z)|\, dz + NCr^{1-n/p} \left(\int_{K+B(0,5r)} |\nabla f(z)|^p \ dz \right)^{1/p} \end{aligned}$$

so that

$$\|f\|_{C^{1-n/p}(K)} \le \frac{1}{|B(0,r)|} \int_{K+B(0,5r)} |f(z)|\, dz + NCr^{1-n/p} \left(\int_{K+B(0,5r)} |\nabla f(z)|^p \ dz \right)^{1/p},$$

and the claimed inequality is proved to hold. □

As another corollary, we notice that an inequality similar to that in Proposition 3.5 follows for higher order differences. Given $h \in \mathbb{R}^n$, define

$$\Delta_h f(x) = f(x+h) - f(x), \quad \Delta_h^k f(x) = \Delta_h^{k-1} \Delta_h f(x).$$

Corollary 3.7. *Let $k \ge 1$. There is a constant C such that if $f \in W^{k,1}_{loc}(\Omega)$, then for almost every $x \in \Omega$ and almost every $B(0, \operatorname{dist}(x, \Omega^c)/(4k))$*

$$|\Delta_h^k f(x)| \le C|h|^k \sum_{j=0}^{k} M_c^k |\nabla^k f|(x + jh).$$

Proof. We can apply Proposition 3.5 with $\alpha = 0$ to obtain

$$|\Delta_h^k f(x)| \le C|h| (M_c(1_{B(x,2|h|)} |\nabla \Delta_h^{k-1} f|)(x) + M_c(1_{B(x+h,|h|)} |\nabla \Delta_h^{k-1} f|)(x+h)).$$

Iterating the argument with ∇f, $\nabla^2 f$, and so forth in place of f we conclude the proof. □

The converse of Proposition 3.5 is an immediate consequence of Theorem 3.4. For future reference, we first define a new type of spaces, the Hajłasz–Sobolev spaces, and we then show that these spaces coincide with the first order Sobolev spaces in $\mathbb{R}^n$.

Definition 3.8 (Hajłasz–Sobolev space)**.** Let (X, d, μ) be a metric measure space. Let $E \subset X$ be a measurable set. The Hajłasz–Sobolev space $M^{1,p}(E)$ is defined as the family of all measurable functions $f : E \to \mathbb{R}$ for which there exists non-negative $g \in L^p(E, \mu)$ such that for $\mu \otimes \mu$ almost every $(x, y) \in E^2$

$$|f(x) - f(y)| \le d(x,y) \left(g(x) + g(y) \right).$$

We norm this space with

$$\|f\|_{M^{1,p}(E)} := \|f\|_{L^p(E,\mu)} + \inf_g \|g\|_{L^p(E,\mu)}$$

where the infimum is over all g admissible as above. The functions g are called Hajłasz gradients of f.

Proposition 3.9. *For $(X, d, \mu) = (\mathbb{R}^n, |\cdot|, \mathcal{L}^n)$, Ω open and $1 \le p < \infty$, the continuous embedding $M^{1,p}(\Omega) \subset W^{1,p}(\Omega)$ holds.*

Proof. Given $f \in M^{1,p}(\Omega)$ and a Hajłasz gradient g, we see that for all balls $B \Subset \Omega$

$$\fint_B |f(x) - f_B|\, dx \le \fint_B \fint_B |f(x) - f(y)|\, dx dy \le \fint_B \fint_B |x-y| (g(x) + g(y))\, dx dy \le 2r(B) \fint_B g(x)\, dx.$$

By Theorem 3.4 f is weakly differentiable and $|\nabla f(x)| \le Cg(x)$ for almost every $x \in \Omega$. It then holds $\|f\|_{W^{1,p}} \le C\|f\|_{M^{1,p}(\Omega)}$. □

The higher order finite differences cannot be defined in the generality of the above proposition. However, we can give a converse to Corollary 3.7 in domains $\Omega \subset \mathbb{R}^n$.

Proposition 3.10. *Let $k \geq 1$ and $p > 1$. Let $f \in L^1_{loc}(\Omega)$. If there is a constant $C \geq 1$ and a non-negative function $g \in L^p_{loc}(\Omega)$ such that for almost every x and $h \in B(0, \operatorname{dist}(x, \Omega^c)/(2k))$*

$$|\Delta_h^k f(x)| \leq C|h|^k \sum_{j=0}^{k} g(x + jh),$$

then f is k times weakly differentiable and there is a constant C' only depending on k, p, n and C with $|\nabla^k f(y)| \leq C' g(y)$ for almost every $y \in \Omega$.

Proof. This is left as an exercise. Do not try to apply Theorem 3.4 as a black box but try to modify its proof. □

We saw in Proposition 3.5 that weakly differentiable functions satisfy a local version of Definition 3.8 when $1 < p < \infty$. The actual global condition is satisfied under additional assumptions on the domain that we do not want to discuss too much in detail here. For instance, if $\Omega = \mathbb{R}^n$ or Ω is a bounded domain whose boundary is Lipschitz regular, then $M^{1,p}(\Omega) = W^{1,p}(\Omega)$ for $1 < p < \infty$. The case $p = 1$ is interestingly different already for a local condition. For simplicity, consider $\Omega = \mathbb{R}^n$. Then $M^{1,1}(\mathbb{R}^n)$ turns out to be the Hardy–Sobolev space $H^{1,1}(\mathbb{R}^n)$, the space of all functions whose first distributional derivatives are in the real variable Hardy space $H^1(\mathbb{R}^n)$. The space $H^1(\mathbb{R}^n)$ consists of those locally integrable functions that the smooth maximal operator sends to $L^1(\mathbb{R}^n)$: Let $\varphi \in C^\infty(\mathbb{R}^n)$, $\operatorname{supp} \varphi \subset B(0,1)$, $1_{B(0,1/2)} \leq \varphi$. The smooth maximal function is defined as

$$\mathcal{M}_\varphi f(x) = \sup_{t>0} |\varphi_t * f(x)|.$$

While M_c maps no non-trivial function to $L^1(\mathbb{R}^n)$ (any positive mass causes a tail of order $\sim |x|^{-n} \notin L^1(\mathbb{R}^n)$), the fact that $\mathcal{M}_\varphi$ respects the signs (no absolute value inside the convolution) and the fact that the convolution kernel is smooth (unlike $1_{B(0,1)}$ for Hardy–Littlewood) change the picture completely for functions with mean value zero, a property that is very common for derivatives. We cannot delve any deeper into the theory for Hardy spaces here (for time constraint), but in the case of interest, the basic theory of Hardy spaces can be found in Section 6.4 of [4], the connection of Hardy–Sobolev theory to Calderón maximal functions is discussed in [7], and the pointwise characterization of Hardy–Sobolev functions can be found in [5].

The pointwise formula is very easy to generalize to fractional orders of smoothness, very differently from the original definition of Sobolev space $W^{1,p}(\mathbb{R}^n)$. However, the role of the resulting functions space in the general theory is currently unclear. In the next lecture, we introduce another and more commonly used candidate to be called fractional Sobolev space, the so-called Sobolev–Slobodeckij space. The Sobolev–Slobodeckij space does not coincide with the traditional Sobolev space in any nice way when talking about integer order smoothness, but it turns out to be the right space of fractional smoothness when talking about traces.