TOPICS 2011 - WHITNEY EXTENSION THEOREM

YOAV BENYAMINI

1. The plan

In §2 we introduce partitions of unity and apply them to construct linear extension operators on spaces of continuous functions.
In §3 we introduce the Whitney cover of open sets in $\mathbb{R}^n$ and apply it to construct linear extension operators for Lipschitz functions on closed subsets of $\mathbb{R}^n$.
In §4 we prove Whitney's characterization of the restrictions of jet of C^k functions to closed subsets of $\mathbb{R}^n$ and construct linear extension operators for C^k jets on closed subsets of $\mathbb{R}^n$.
In §5 we prove Sard's theorem and present Whitney's proof that weaker smoothness is not sufficient.

2. Partitions of unity

Definition 2.1. Let K be a Hausdorff space and let $\mathcal{U}$ be an open cover of K.

(i) A cover $\mathcal{V}$ of K is called a refinement of $\mathcal{U}$ if for every $V \in \mathcal{V}$ there is a $U \in \mathcal{U}$ s.t. $V \subset U$.

(ii) The cover is said to be locally finite if every $x \in K$ has a nbhd which intersects only a finite number of the sets in $\mathcal{U}$.

(iii) K is called paracompact if every open cover of K has a locally finite refinement.

(iv) A partition of unity, subordinated to an open cover $\mathcal{U}$ is a family of continuous functions $f_\alpha : K \to [0,1]$ s.t. $\{\psi_\alpha > 0\}$ is a locally finite refinement of $\mathcal{U}$ and s.t. $\sum \psi_\alpha(x) = 1$ for every $x \in K$. (Note that this is a locally finite sum.)

Partitions of unity will serve to "glue" local constructions to a global one. For example, assume $\{\psi_\alpha\}$ is a partition of unity and for each α the function f_α is defined on some set containing $\{\psi_\alpha > 0\}$ and has a certain property $\mathcal{P}$. Then the function $f = \sum \psi_\alpha f_\alpha$ is well defined on all of K and is a locally a finite convex combination of the f_α's. For certain properties $\mathcal{P}$ (e.g., continuity, nonnegativity, boundedness by a uniform constant) this suffices to ensure that also f has $\mathcal{P}$.

Theorem 2.2. *(i) A paracompact space is normal.*

(ii) A compact Hausdorff space is paracompact.

(iii) A metric space is paracompact.

(iv) In a paracompact space any open cover admits a partition of unity subordinated to it.

Proof. We shall not prove the easy part (i) (hint: prove first that it is regular), the trivial part (ii), and the hard part (iii) (the original proof of Stone (1948) was quite involved, but M. E. Rudin [6] gave a short, simple and ingenious proof).

To prove (iv) we may assume that the given cover, $\mathcal{U}$, is locally finite, and we first find, by transfinite induction, for each $U \in \mathcal{U}$ an open set V_U s.t. $\overline{V_U} \subset U$ and s.t. $\{V_U\}$ is still a cover of K.

To this end index the cover as $\{U_\alpha\}_{\alpha<\alpha_0}$ and assume the V_α's have been found for all $\alpha < \beta$. Replacing, for $\alpha < \beta$, the U_α's by the V_α's the problem reduces to finding a $V \subset W'$ for some fixed W' in a cover $\mathcal{W}$. But by (i) K is normal, hence there are disjoint open sets separating the two closed and disjoint sets

$$K \setminus W' \quad \text{and } K \setminus \bigcup\{W : W \in \mathcal{W} \ \& \ W \neq W'\}$$

and if O is the open set containing $K \setminus W'$, take $V_{W'}$ to be the interior of $K \setminus O$.

Now, by the normality of K and Tietze's theorem, find for each $U \in \mathcal{U}$ a function $\varphi_U : K \to [0,1]$ s.t. $f(x) \equiv 1$ on V_U and $\{f_U > 0\} \subset U$. Then put

$$\psi_U = \frac{\varphi_U}{\sum_{\mathcal{U}} \varphi_U} .$$

□

As an application we discuss linear extension operators for continuous functions. Tietze's theorem allows, under certain circumstances, to extend functions from a closed subset H of a topological space K to the whole space. Can we find a continuous linear operator that does the extension simultaneously for all the continuous functions on the subset? The answer is trivially yes if H is a retract of $K-$ if $r : K \to H$ is the retraction, then $Tf(x) = f(r(x))$ is such an operator. But, of course, retracts are quite rare.

The following theorem gives conditions under which a linear extension operator exists. Example 2.4 shows that this is not always possible.

Theorem 2.3 (Borsuk - Kakutani)**.** *Let K be a metric space and let $H \subset K$ be a closed subset. Then there is a linear extension operator $T : C_B(H) \to C_B(K)$ satisfying $\|T\| = 1$ and $T1 = 1$ (so, in particular, $Tf \geq 0$ whenever $f \geq 0$).*

Proof. Denote the metric by d. The balls $B_z = B(z, dist(z,H)/2)$ for $z \notin H$ cover $K \setminus H$, and note that $d(z,H) \leq 2d(x,H)$ for every $x \in B_z$.

Let ψ_α be a partition of unity on $K \setminus H$ subordinated to its cover $\{B_z\}$ and choose, for each α, a point z_α s.t. $\{\psi_\alpha \neq 0\} \subset B_{z_\alpha}$. Then choose $y_\alpha \in H$ s.t. $d(z_\alpha, y_\alpha) < 2d(z_\alpha, H)$. It follows that if $\psi_\alpha(x) \neq 0$ then

$$d(x,y_\alpha) \leq d(x,z_\alpha) + d(z_\alpha,y_\alpha) < \frac{1}{2}d(z_\alpha,H) + 2d(z_\alpha,H) = \frac{5d(z_\alpha,H)}{2} \leq 5d(x,H).$$

Define the linear operator T by

$$Tf(x) = \begin{cases} \sum \psi_\alpha(x) f(y_\alpha) & x \notin H \\ f(x) & x \in H \end{cases}$$

for $f \in C_B(H)$. The local finiteness of $\{\psi_\alpha \neq 0\}$ implies that each $x \notin H$ has a nbhd W where the sum is actually a finite convex combination of values of f. Thus $\|Tf\| \leq \|f\|$ and $T1 = 1$. Also Tf is continuous separately on H and on $K \setminus H$.

To show that Tf is actually continuous, fix $y \in H$ and we need to estimate $f(y) - Tf(x)$ for $x \notin H$. Thus fix x and let $A_x = \{\alpha : \psi_\alpha(x) \neq 0\}$. The set A_x is finite and $Tf(x) = \sum_{A_x} \psi_\alpha(x) f(y_\alpha)$, and for $\alpha \in A_x$

$$d(y,y_\alpha) \leq d(y,x) + d(x,y_\alpha) \leq d(x,y) + 5d(x,H) \leq 6d(x,y).$$

Fix now $\varepsilon > 0$ and choose δ s.t. $t \in H$ and $d(y,t) < \delta$ imply that $|f(t)-f(y)| < \varepsilon$. If x is such that $d(y,x) < \delta/6$ then $d(y,y_\alpha) < \delta$ for every $\alpha \in A_x$. Hence

$$|f(y) - Tf(x)| = \left| \sum_{A_x} \psi_\alpha(x)\big(f(y) - f(y_\alpha)\big) \right| \le \sum_{A_x} \psi_\alpha(x)\varepsilon = \varepsilon.$$

□

Example 2.4. The Borsuk - Kakutani theorem holds, more generally, when K is paracompact and H is compact and metrizable. But continuous linear extension operators do not necessarily exist otherwise. Before we give a concrete example, we first show that if K is a compact Hausdorff space which contains a countable dense sequence $\{x_j\}$ and $H \subset K$ is a closed subset which contains an uncountable family of disjoint open sets U_t, then there is no *bounded* linear extension $T : C(H) \to C(K)$.

Assume for contradiction that such an operator T exists.

Choose, for each t, a function $0 \le g_t \in C(H)$ supported in U_t s.t. $\|g_t\| = 1$. It follows that $\|Tg_t\| \ge 1$ (because Tg_t is an extension of g_t). Thus each of the uncountably many nonempty open sets $V_t = \{Tg_t > 1/2\}$ must contain one of the x_j's – and there are only countably many of them. Thus there is a point $x \in K$ which belongs to uncountably many sets V_t. Fix $n > 2\|T\|$ and choose $t_1, \dots, t_n$ s.t. $x \in V_{t_j}$. Put $g = \sum_{j \le n} g_{t_j}$. As the g_t's are disjointly supported it follows that $\|g\| = 1$. But this leads to a contradiction because

$$\|T\| \ge \|Tg\| \ge \sum_{j \le n} Tg_{t_j}(x) > n/2 > \|T\|.$$

It remains to give an example of such K and H.

Let $I = [0,1]$. Then $K = I^I$ with the product topology is compact and separable. Indeed, let $\mathcal{P}$ denote the (countable) set of finite partitions $P = \{I_j\}$ of I by disjoint intervals with rational endpoints. It then follows from the definition of the product topology that the countable set

$$\Big\{ f_{P;r_j} : f_{\big|I_j} \equiv r_j \,;\, r_j \in \mathbb{Q} \,;\, P \in \mathcal{P} \Big\}$$

is dense in K. For the set $H \subset K$ we take

$$H = \Big\{ f \in K : \sum_{t \in I} f(t) \le 1 \Big\} = \bigcap_{t_1, \dots, t_n} \Big\{ f \in K : \sum_{1}^{n} f(t_i) \le 1 \Big\}.$$

As an intersection of closed sets H is closed, and the uncountably many open sets $U_t = \{f \in H : f(t) > 1/2\}$ for $t \in I$ are, indeed, pairwise disjoint.

3. The Whitney cover

To extend continuous functions it was enough that the partition of unity was locally finite. For other classes of functions, for example, to extend Lipschitz functions to Lipschitz functions or for functions with a certain number of continuous derivatives, this is not enough. In addition to conditions on the smoothness of the ψ's we shall need a uniform bound on the maximal number of intersecting sets in the cover, as well as a bound on the distance of V_U from U^c. In this section we describe the Whitney covers of open subsets of $\mathbb{R}^n$, which give such uniform bounds.

Such covers will be applied to construct "Whitney type" partitions of unity and extensions.

By a cube in $\mathbb{R}^n$ we shall mean a closed cube with sides parallel to the axes. Two cubes are said to be disjoint if their interiors are disjoint. The distance of a cube Q from a set F is denoted by $dist(Q, F)$.

Theorem 3.1 (Whitney). *Let $F \subset \mathbb{R}^n$ be closed. Then its complement $F^c = \mathbb{R}^n \setminus F$ admits a cover $\mathcal{Q} = \{Q_k\}$ by closed cubes with disjoint interiors s.t.*

(i) $diam(Q) \le dist(Q, F) \le 4diam(Q)$ for every $Q \in \mathcal{Q}$.

(ii) If $Q_1, Q_2 \in \mathcal{Q}$ touch, then $diam(Q_1) \le 4diam(Q_2)$.

(iii) A cube $Q \in \mathcal{Q}$ touches at most $N = N(n) = (12)^n$ other cubes in $\mathcal{Q}$.

(iv) Fix $0 < \varepsilon < 1/4$ and let $\mathcal{Q}^*$ be the collection of cubes Q^* obtain by expanding each $Q \in \mathcal{Q}$ by a factor of $1 + \varepsilon$ around its center. Then $\mathcal{Q}^*$ is a cover of F^c s.t. each point in F^c has a nbhd which intersects at most N cubes $Q^* \in \mathcal{Q}^*$.

Proof. Let $\mathcal{M}_0$ be the collection of all cubes with edges of length one and integer vertices, and for $k \in \mathbb{Z}$ let $\mathcal{M}_k$ be $2^{-k}\mathcal{M}_0$. A cube in $\mathcal{M}_k$ has diameter $2^{-k}\sqrt{n}$.

The "layers"

$$\Omega_k = \Big\{x \in \mathbb{R}^n : 2\sqrt{n}2^{-k} < d(x, F) \le 2\sqrt{n}2^{-k+1}\Big\}$$

are disjoint from F and cover F^c, and we define a collection $\mathcal{Q}_0$ of (not necessarily disjoint) cubes by

$$\mathcal{Q}_0 = \bigcup_k \Big\{Q \in \mathcal{M}_k : Q \cap \Omega_k \ne \emptyset\Big\}.$$

We first check that every $Q \in \mathcal{Q}_0$ satisfies

$$diam(Q) \le dist(Q, F) \le 4diam(Q).$$

Indeed, fix k and $Q \in \mathcal{M}_k$ s.t. $Q \cap \Omega_k \ne \emptyset$, and choose $x \in Q \cap \Omega_k$. Then

$$diam(Q) = 2^{-k}\sqrt{n} = 2\sqrt{n}2^{-k} - diam(Q) < d(x, F) - diam(Q) \le dist(Q, F)$$

and

$$dist(Q, F) \le d(x, F) \le 2\sqrt{n}2^{-k+1} = 4diam(Q).$$

The cubes in $\mathcal{Q}_0$ are disjoint from F and cover F^c, but they are not disjoint, so we now pass to a subcollection $\mathcal{Q} \subset \mathcal{Q}_0$ of disjoint cubes which still covers F^c. The point is that if $Q_1, Q_2 \in \cup\mathcal{M}_k$ are not disjoint, then one of them contains the other. Moreover, every $Q \in \mathcal{Q}_0$ is contained in a (unique!) maximal $Q' \in \mathcal{Q}_0$ because the diameters of the cubes containing Q are bounded by $d(Q, F)$.

Thus define $\mathcal{Q}$ to be the set of these maximal cubes, and (i) is already proved.

Part (ii) follows from (i). Indeed, the fact the cubes touch and (i) imply that

$$diam(Q_1) \le dist(Q_1, F) \le dist(Q_2, F) + diam(Q_2) \le 5diam(Q_2, F)$$

but the diameters are powers of 2, so necessarily $diam(Q_1) \le 4diam(Q_2, F)$.

To prove (iii) assume that $Q \in \mathcal{M}_k$. By (ii) the touching cubes can only be in $\mathcal{M}_{k-1}$, $\mathcal{M}_k$ or $\mathcal{M}_{k+1}$. But Q touches exactly 3^n cubes in $\mathcal{M}_k$ (including Q itself), and each cube in $\mathcal{M}_k$ is split into 4^n sub-cubes in $\mathcal{M}_{k+1}$, so the maximal numbers of cubes that can touch Q is $3^n 4^n = (12)^n$.

Finally, to prove (iv) we note that by (i) each Q^* is contained in F^c. Also Q^* intersects Q' iff Q touches Q'. Indeed, by (ii) the diameters of all the cubes touching

Q are at least $diam(Q)/4$ so its expansion Q^* is contained in their union (including Q itself), and by (iii) there are at most N such cubes. And, of course, every $x \in F^c$ is contained in some $Q' \in \mathcal{Q}$. As ε is strictly smaller than $1/4$, this also holds for points in some nbhd of x even when $x \in \partial Q$. □

The parameter $\varepsilon < 1/4$ will be fixed from here on. We shall not specify its value nor shall we note the dependence on ε of constants that depend on it.

The Whitney cover will be applied in the next section to extend smooth functions from closed subsets of $\mathbb{R}^n$. We finish this section with a "baby case" of such a theorem.

Let X be a metric space and let $f : X \to \mathbb{R}$ be a Lipschitz function. We denote its Lipschitz constant sup $\frac{f(x)-f(y)}{d(x,y)}$ by $Lip(f)$ and consider the Banach space $Lip(X)$ of Lipschitz functions on X with the norm $\|f\|_{Lip} = \max\{\|f\|_\infty, Lip(f)\}$.

Theorem 3.2. *Let $F \subset \mathbb{R}^n$ be a closed set. Then there is a continuous linear extension operator $T : Lip(F) \to Lip(\mathbb{R}^n)$. The norm of T depends only on n and not on F.*

Proof. In what follows the letter C will stand for a constant $C > 0$ which may depend only on n, and which could have different values at different places. We shall also write $a \sim b$ when they are equivalent up to a constant that may depend only on n.

Thus, let $\mathcal{Q} = \{Q_k\}$ be the Whitney cover of F^c, and we summarize Theorem 3.1 by saying that $d(x, F) \sim diam(Q^*) \sim dist(Q^*, F)$ for every x in any of the $N = N(n)$ cubes Q_k^* that intersect Q^*.

For each k let $p_k \in F$ be a closest point to Q_k and let $\varphi_k(x) = dist(x, Q_k^*)$. These functions are 1-Lipschitz and $\varphi_k(x) \le Cd(x, F)$. We consider the partition of unity on F^c, subordinated to the cover $\mathcal{Q}^*$, given by $\psi_k(x) = \frac{\varphi_k}{\sum \varphi_j}$ and define

$$Tf(x) = \begin{cases} \sum \psi_k(x) f(p_k) & x \notin F \\ f(x) & x \in F. \end{cases}$$

As in Theorem 2.3 Tf is continuous and bounded on $\mathbb{R}^n$ by the same bound as f, and we need to estimate its Lipschitz constant.

We first estimate the Lipschitz constants $Lip(\psi_k)$. Put $g = \sum \varphi_j$ and note that $C_1 d(x, F) \le g(x) \le C_2 d(x, F)$. Indeed, if $x \in Q_k$ then the k'th term, namely $\varphi_k(x)$ is already bigger than $\frac{\varepsilon diam(Q_k)}{(1+\varepsilon)} \sim d(x, F)$. Also, $\varphi_j(x) \ne 0$ for at most $N(n)$ terms in the sum and each of them is bounded above by $Cd(x, F)$. Similarly $Lip(\varphi_k) = 1$ and $Lip(g) \sim 1$.

Fix now $x, y \in Q_m$, then

$$\begin{aligned} \psi_m(x) - \psi_m(y) &= \frac{\varphi_m(y)}{g(y)} - \frac{\varphi_m(x)}{g(x)} \\ &= \frac{g(x)(\varphi_m(y) - \varphi_m(x)) + \varphi_m(x)(g(x) - g(y))}{g(y)g(x)} \\ &\le C\frac{d(x, F)d(x, y) + d(x, F)d(x, y)}{d(x, F)d(y, F)} \sim \frac{d(x, y)}{d(Q_m, F)}. \end{aligned}$$

Assume now that $Lip(f)=1$ and write, using $\sum(\psi_k(y)-\psi_k(x))=1-2=0$

$$\begin{aligned}|Tf(y)-Tf(x)| &= \left|\sum(\psi_k(y)-\psi_k(x))f(p_k)\right| \\ &= \left|\sum(\psi_k(y)-\psi_k(x))(f(p_k)-f(p_m))\right| \\ &\leq C\sum\frac{d(x,y)}{d(Q_m,F)}d(p_k,p_m)\leq Cd(x,y)\end{aligned}$$

because the sum has at most N nonzero terms and

$$\begin{aligned}d(p_k,p_m) &\leq dist(F,Q_k)+diam(Q_k)+diam(Q_m)+dist(F,Q_m) \\ &\sim diam(Q_m)\sim diam(Q_k)\end{aligned}$$

for each of them.

We are now ready to finish the proof.

Fix $x,y\in\mathbb{R}^n$ and we consider two cases. The first is that the interval $[x,y]$ does not contain any point in F. In this case we can find successive points $x=z_0,z_1,\ldots,z_k=y$ in the interval s.t. each pair z_m,z_{m+1} belongs to some cube Q_m. It follows that

$$|Tf(y)-Tf(x)|\leq\sum|Tf(z_{m+1})-Tf(z_m)|\leq C\sum d(z_m,z_{m+1})=Cd(x,y).$$

If $[x,y]$ contains points in F let $p,q\in F$ be the points in the interval nearest to x and y respectively. As

$$|Tf(y)-Tf(x)|\leq|Tf(y)-Tf(q)|+|Tf(q)-Tf(p)|+|Tf(p)-Tf(x)|$$

* and

$$d(y,q)+d(q,p)+d(p,x)=d(y,x)$$

and since $|Tf(q)-Tf(p)|=|f(q)-f(p)|\leq d(p,q)$ it is enough to prove that $|Tf(y)-Tf(q)|\leq Cd(y,q)$. But this follows from the first case and the continuity of Tf: for each $z\in(q,y)$ we already know that $|Tf(y)-Tf(z)|\leq Cd(y,z)$, and we let $z\to q$. □

4. Extensions of smooth functions

In this section we shall extend C^k functions on F to C^k functions on all of $\mathbb{R}^n$. We shall need, of course, a partition of unity with C^k functions and with estimates on their derivatives – and this is simple and will be done in Lemma 4.1. The main problem, however, is what do we mean by C^k functions on a general closed set $F\subset\mathbb{R}^n$. Whitney gave a good definition of this notion and we shall explain it and then prove the extension theorem.

Lemma 4.1. *Let $F\subset\mathbb{R}^n$ be closed and let $\mathcal{Q}$ be a Whitney cover for F^c. Then there is a C^∞ partition of unity $\{\psi_k\}$ subordinated by $\mathcal{Q}^*$ s.t.*

$$|D^\alpha\psi_k(x)|\leq\frac{C_\alpha}{d^{|\alpha|}(x,F)}$$

for every multi-index α and $x\in F^c$.

Proof. Let $0 \le h \le 1$ be a C^∞ function on $\mathbb{R}$ supported in $[-\varepsilon, 1+\varepsilon]$ s.t. $h \equiv 1$ on $[0,1]$. Then $\varphi(x) = \prod h(x_j)$ is a C^∞ function on $\mathbb{R}^n$ supported in $[-\varepsilon, 1+\varepsilon]^n$, and we put $\varphi_k(x) = g((x - z_k)/l_k)$ where z_k is the center and l_k is the side length of Q_k^*. As the φ's are obtained by rescaling a fixed function h and the scaling factor l_k is proportional to $diam(Q_k) \sim d(x, F)$ whenever $x \in Q_k^*$, it is clear that $|D^\alpha \varphi_k(x)| \le \frac{C_\alpha}{d^{|\alpha|}(x,F)}$.

We now define $\psi_k = \frac{\varphi_k}{\sum \varphi_j} = \frac{\varphi_k}{g}$ and we prove the estimate for the derivatives of the ψ's. Indeed, writing $\varphi_k = \psi_k g$, the estimates hold for the φ_k's and g (as a sum of N φ's). Differentiating the product, and then isolating the highest derivative of ψ_k gives that

$$D^\alpha \psi_k(x) = D^\alpha \varphi_k(x) - \sum_{\beta<\alpha} a_{\alpha,\beta} D^\beta \psi_k(x) D^{\alpha-\beta} g(x)$$

and the estimate now follows by induction on $|\alpha|$. □

To have an idea how to define C^k functions on closed sets, we first reformulate the condition that f is C^k on $\mathbb{R}^n$. Let

$$P(x) = P_y(x, f) = f(y) + \sum_{1\le|\alpha|\le k} \frac{D^\alpha f(y)}{\alpha!}(x-y)^\alpha$$

be the k-degree Taylor polynomial of f around the point y, evaluated at x, and let $R(x) = f(x) - P(x)$ be the remainder.

A k-jet is a collection of functions $(u^\alpha)_{|\alpha|\le k}$. The k-jet of $f \in C^k$ is the collection $(D^\alpha f)_{|\alpha|\le k}$ of its derivatives.

Note that the $(k - |\alpha|)$ Taylor polynomial of $D^\alpha f$ is just $D^\alpha P$.

Proposition 4.2. *A k-jet (f^α) on $\mathbb{R}^n$ is the k-jet of a C^k function $f = f^0$ iff for every $y \in \mathbb{R}^n$ there is a k-degree polynomial P_y s.t. for every compact set A and for every $|\alpha| \le k$*

$$f^\alpha(x) - D^\alpha P_y(x) = o(\|y - x\|^{k-|\alpha|})$$

uniformly for $y \in A$.

Proof. We give the proof for $n = 1$. This also proves the general case by restricting the functions to lines and noting that the estimate do not depend on the choice of the line but only on the moduli of continuity or the quantitative formulation of the o-condition.

Assume $f \in C^k$ and take P_y to be the k'th Taylor polynomial of f at y. For simplicity, we just treat the case of $k = 2$, the proof of the general case is similar. Write

$$\begin{aligned} R_2(x) &= f(x) - f(y) - f'(y)(x-y) - \frac{f''(\theta_x)}{2}(x-y)^2 \\ &= \int_y^x f'(t)dt - f'(y)(x-y) - \frac{f''(\theta_x)}{2}(x-y)^2 \\ &= \int_y^x (f'(t) - f'(y))dt - \frac{f''(\theta_x)}{2}(x-y)^2 \\ &= \int_y^x \int_y^t (f''(s) - f''(\theta_x))dsdt \le \max_{y\le s,\theta\le x} |f''(s) - f''(\theta)| \frac{(x-y)^2}{2} \end{aligned}$$

and the result follows from the continuity of f''.

The converse direction is immediate. □

This observation led Whitney to the following definition of C^k functions on general closed sets

Definition 4.3. Let $F \subset \mathbb{R}^n$ be a closed set. A k-jet (f^α) on F is said to be the jet of the C^k function $f = f^0$ on F if for every $z \in F$ there is a k-degree polynomial P_z s.t. for every compact set $A \subset F$ and for every $|\alpha| \le k$

$$f^\alpha(y) - D^\alpha P_z(y) = o(\|z - y\|^{k-|\alpha|})$$

uniformly for $y, z \in A$.

Remark. Note that the definition requires as data not only the function $f = f^0$ but the whole k-jet (f^α). We shall comment on this at the end of the chapter.

For a C^k function on $\mathbb{R}^n$ the Taylor polynomials $P_y(x)$ depend continuously on y (and x) with an $o\big((\|y - z\| + \|x - y\|)^k\big)$ estimate – and similarly for $D^\alpha P_y$. This is also true for the P_z's in the definition of C^k k-jets on F.

Lemma 4.4. *Let $F \subset \mathbb{R}^n$ be a closed set and let (f^α) be a C^k jet on F. For $z \in F$ let P_z be the polynomials in the definition of a C^k jet. Then for every compact set $A \subset F$ and for every $|\alpha| \le k$*

$$D^\alpha P_y(x) - D^\alpha P_z(x) = o\big((\|y - z\| + \|z - x\|)^{k-|\alpha|}\big)$$

uniformly for $y, z \in A$.

Proof. It is enough to estimate $P_y - P_z$ (i.e., when $\alpha = 0$) because $D^\alpha P_y$ is just the polynomial associated with the $(k - |\alpha|)$-jet $(f^{\alpha+\gamma})_{\gamma \le k-|\alpha|}$.

Thus, expand the polynomial $P_y(x) - P_z(x) = g(x) = \sum \frac{D^\alpha g(z)}{\alpha!}(x - z)^\alpha$ around the point z, and estimate $D^\alpha g(z)$ as

$$D^\alpha(P_y(x) - P_z(x))\big|_{x=z} = D^\alpha P_y(z) - f^\alpha(z) = o(\|y - z\|^{k-|\alpha|}).$$

It follows that

$$\begin{aligned} P_y(x) - P_z(x) &= \sum \frac{D^\alpha g(z)}{\alpha!}(x - z)^\alpha = \sum o(\|y - z\|^{k-|\alpha|}) \cdot \|z - x\|^{|\alpha|} \\ &= o\big((\|y - z\| + \|x - z\|)^k\big). \end{aligned}$$

□

Theorem 4.5. *Let $F \subset \mathbb{R}^n$ be closed and let (f^α) be a C^k jet on F. Then $f = f^0$ can be extended to a C^k function f on all of $\mathbb{R}^n$ with $D^\alpha f = f^\alpha$ for all $|\alpha| \le k$.*

Proof. Fix a Whitney cover of F^c and choose, for each j a point $y_j \in F$ nearest to Q_j. Let ψ_j be the partition of unity of lemma 4.1, and define

$$Tf(x) = \begin{cases} f(x) & x \in F \\ \sum \psi_j(x) P_{y_j}(x) & x \notin F. \end{cases}$$

We need to show that for every $z \in \mathbb{R}^n$ there is a polynomial P_z so that $D^\alpha P_z(x) - D^\alpha Tf(x) = o(\|z - x\|^{k-|\alpha|})$. As in previous extension theorems this is obvious for $x, z \in F$ (by the definition of a C^k jet) and when $z \notin F$ (because Tf is C^∞ in a nbhd of such a point). Thus we only need to treat $z \in F$ and $x \notin F$.

We start with $\alpha = 0$, i.e., $P_z(x) - Tf(x) = o(\|z-x\|^k)$. Indeed,

$$Tf(x) - P_z(x) = \sum \psi_j(x)(P_{y_j}(x) - P_z(x))$$

but

$$P_{y_j}(x) - P_z(x) = o\big((\|y_j - z\| + \|x-z\|)^k\big) = o(\|z-x\|^k)$$

by Lemma 4.4 and because $\|y_j - z\| \le C\|x-z\|$ for the j's for which $\psi_j(x) \neq 0$.

We now pass to the first derivative, and denote by D the derivative with respect to x_1, say. We need to check that $DP_z(x) - DTf(x) = o(\|z-x\|^{k-1})$ and write

$$DTf(x) = \sum \psi_j(x) DP_{y_j}(x) + \sum D\psi_j(x) P_{y_j}(x).$$

As for $\alpha = 0$, the first sum is $DP_z(x)$ up to $o(\|z-x\|^{k-1})$. To estimate the second sum let $y \in F$ be nearest to x and, as $\sum D\psi_j(x) \equiv 0$, write it as

$$\sum D\psi_j(x)(P_{y_j}(x) - P_y(x)) = O(1/d(x,F)) \cdot o(\|y-x\|^k) = o(\|z-x\|^{k-1}).$$

The proof for general α is similar. □

5. Sard's theorem

The following theorem of Sard plays an important role in the study of smooth functions between differentiable manifolds. (For a nice application see Hirsch's proof of the non-retraction and Brouwer's fixed point theorems in [5], page 14.) The case of mappings between manifolds reduces trivially to mapping between Euclidean spaces - and this is how we formulate it.

Definition 5.1. Let $U \subset \mathbb{R}^s, V \subset \mathbb{R}^t$ be open and let $f : U \to V$ be C^1. A point $x \in U$ is called a critical point of f if the rank of the Jacobi matrix $Jf(x)$ is smaller than its maximal possible value, namely, $\min(s,t)$. A point $y \in V$ is a critical value of f if $y = f(x)$ for some critical point x.

We denote the set of critical points by $\mathcal{C}$, and then the set of critical values is $f(\mathcal{C})$.

Theorem 5.2. *With the notation above, if f is C^k and $k \ge s/t$ then $f(\mathcal{C})$ has measure zero.*

Proof. We shall only discuss the case $s = 2, t = 1;\ k \ge 2$ of the theorem. This case essentially contains (in a simple form) all the ingredients of the general case. For the proof of the general case see [5], [4] or [2]. We shall then apply Whitney's extension theorem to show that in this case $k = 1$ is not enough.

The theorem is of a local nature (it is enough to show that every $x \in U$ has a nbhd W s.t. $f(\mathcal{C} \cap W)$ has measure zero). We may thus assume that U is the close unit cube.

As a preliminary step we need to treat the case $s = 1$, but the proof for $s = t$ is the same for any s (and, of course, f only needs to be C^1). Given $\varepsilon > 0$ there is a $\delta > 0$ s.t. the $|det(Jf)| < \varepsilon$ in any subcube Q of side δ which intersects $\mathcal{C}$. It follows that the volume of $f(Q)$ of such a cube Q is smaller than $\varepsilon\delta^s$. Now cover U with δ^{-s} cubes of side δ, then the total measure of those $f(Q)$'s s.t. $Q \cap \mathcal{C} \neq \emptyset$ is smaller than $\varepsilon\delta^s \times \delta^{-s} = \varepsilon$.

Turning to $s = 2$, let $\mathcal{C}_1 \subset \mathcal{C}$ be the set of points where also all the second derivatives are zero. Cover U with ε^{-2} squares of side ε. If Q is one of these cubes which intersects $\mathcal{C}_1$ at a point a, say, then by Taylor's formula

$$f(x) - f(a) = o(\|x - a\|^2).$$

It follows that the length of the interval $f(Q)$ for such a Q is $o(\varepsilon^2)$. Thus the total measure of these $f(Q)$'s is $\varepsilon^{-2} \times o(\varepsilon^2) = o(1)$.

Assume now that $a \in \mathcal{C} \setminus \mathcal{C}_1$ and we shall find a nbhd W of a s.t. $f(\mathcal{C} \cap W)$ has measure zero.

Assume, for example, that $f''_{x,x}(a) \neq 0$. It follows that the Jacobi matrix of the mapping $g(x, y) = (f'_x(x, y), y)$ is non-singular at a. Thus it is a diffeomorphism of some nbhd W of a onto an open set $W' \subset \mathbb{R}^2$, and the set $\mathcal{C}'$ of critical points of $f \circ g^{-1}$ in W' is just $g(W \cap \mathcal{C})$. But if (x, y) is a critical point of f then $g(x, y) = (0, y)$, i.e., $\mathcal{C}'$ is contained in the line $x = 0$. By the case $s = t = 1$ it follows that the measure of $f(W \cap \mathcal{C}) = (f \circ g^{-1})(\mathcal{C}')$ is zero. □

A C^1 function $f : \mathbb{R} \to \mathbb{R}$ is determined (up to a constant) by its derivative. Let Γ be a simple curve in $\mathbb{R}^2$ and let $f : \mathbb{R}^2 \to \mathbb{R}$ be C^1. Is $f_{|\Gamma}$ determined by $\nabla f_{|\Gamma}$? Equivalently, if $\nabla f_{|\Gamma} \equiv 0$ is f constant on Γ?

If f is C^2 Sard's theorem implies that the answer is yes. Indeed, if $\nabla f_{|\Gamma} \equiv 0$ then the interval $f(\Gamma)$ is contained in the set of critical values of f, which, by Sard's theorem has measure zero – so it reduces to a single point.

Whitney [9] constructed an example which shows that, at least for $s = 2, t = 1$ the restriction $k \geq 2$ cannot be reduced to $k = 1$. To this end he constructed a curve $\Gamma \subset \mathbb{R}^2$ and a Cantor set in Γ and then considered the associated Cantor function f on the curve. The construction is such that f satisfies $|s - s_0| = o(\|f(s) - f(s_0)\|)$ on the curve hence, by Whitney's extension theorem it extends to a C^1 function f on all of $\mathbb{R}^n$. But then $f(\mathcal{C}) \supset f(\Gamma)$ which is the whole unit interval.

Instead of Whitney's construction we shall follow Glaeser [3] (pages 57-58) who noted that the classical Koch curve gives a similar example.

Example 5.3. We recall the construction of the Koch curve. Replace the middle third of $K_0 = [0, 1]$ by the two other edges of the isosceles triangle whose base is this middle third. The resulting figure is a curve, $K_1 = K_1(t)$, which maps the four intervals $I_j = [\frac{j-1}{4}, \frac{j}{4}]$ isometrically onto the four intervals of the image. The curve $K_2 = K_2(t)$ is obtained by repeating the same construction on each of the intervals I_j. Continue inductively, and the Koch curve $K = K(t)$ is the limit curve of the $K_n(t)$'s.

Define $f(s) = K^{-1}(t)$, and we want to prove that $f(s) - f(s_0) = o(\|s - s_0\|)$ for every $s_0 \in K$. Indeed, let n be the minimal index s.t. s and s_0 are separated by an edge in K_n. It follows that $\|s - s_0\| \geq \frac{1}{2 \cdot 3^n}$. On the other hand

$$|f(s) - f(s_0)| = |t - t_0| \leq K \sum_{j \geq n} \frac{1}{4^j} = O(1/4^n)$$

hence $\frac{|f(s) - f(s_0)|}{\|s - s_0\|} = O\left(\frac{3^n}{4^n}\right) \to 0$. (This is actually the same proof as Koch's proof that K is not differentiable anywhere.)

Remark. Consider $[0, 1]$ equipped with the metric $d(s, t) = |s - t|^\alpha$ for some $0 < \alpha < 1$. If $\varphi : [0, 1] \to \mathbb{R}^2$ is a Lipschitz embedding, i.e., φ and its inverse are

Lipschitz, then it is clear that $f = \varphi^{-1}$ satisfied $f(s) - f(s_0) = o(\|s - s_0\|)$, and is an example of the same type.

The computations for the Koch curve show that, in fact, $K(t)$ is such a Lipschitz embedding for $\alpha = \frac{\log 3}{\log 4}$.

Assouad [1] investigates the relations between n, k and α so that the cube $[0,1]^k$ with the metric $d(s,t) = \|s - t\|^\alpha$ is Lipschitz embeddable in $\mathbb{R}^n$. In particular, he proves that for $k = 1$ and $n = 2$ this is true iff $1/2 < \alpha \leq 1$.

References

[1] P. Assouad, Plongements Lipschtziens dans $\mathbb{R}^n$, Bull. Soc. Math. France, 111 (1983), 429-448.
[2] S. G. Krantz and H. R. Parks, The geometry of domains in space.
[3] G. Glaeser, Etude de quelques algebres Tayloriennes, J. d'Analyse Mth. 6 (1958), 1-125.
[4] V. G. Guillemin and A. Pollack, Differential Topology.
[5] J. Milnor, Topology from a Differential Point of View.
[6] M. E. Rudin, A new proof that metric spaces are paracompact, Proc. Amer. Math. Soc, 20 (1969), 603.
[7] E. M. Stein, Singular Integrals and Differentiability properties of functions.
[8] H. Whitney, Analytic extensions of differentiable functions defined in closed sets, Trans. Amer. Math. Soc, 36 (1934), 63-89.
[9] H. Whitney, A function not constant on a connected set of critical points, Duke Math. J., 4 (1935), 514-517