Math Colloquium: Professor Fefferman

Math Club

July 3, 2026

Topic: Lipschitz selection

Abstract: Suppose that to each point x in a metric space X we associate a convex subset $K(x)$ in $\mathbb{R}^D$. We'd like to find a Lipschitz map $F : X \to \mathbb{R}^D$ such that $F(x)$ lies in $K(x)$ for each x in X. Such an F may or may not exist. That's one version of the problem of finding smooth functions that agree approximately with data.

Observe some data. x axis, y axis. (Draw x, y plane. Draw at each of some x's, values of y. Draw in interpolated function. Could also be very squiggly function!) Don't know if function is smooth, but tempting to guess that. If it comes from real world, don't care if exactly passes through data points, as they are measured! We can add error bars. Want to find as smooth function as possible that approximately agrees with data points.

One flavor of this problem: Lipschitz selection. $F : X \to \mathbb{R}^D$ Lipschitz means $|F(x) - F(y)| \leqslant Cd(x, y)$, where C is Lipschitz constant/norm of F.

Problem of Lipschitz selection: (draw MS (X, d). For each point x, have convex set $K(x)$ in $\mathbb{R}^d$. Say compact and convex. Draw map from points of X to rectangles $K(x)$ in $\mathbb{R}^d$. Think of $K(x)$ as a target, i.e. want to find F s.t. $F(x)$ lies somewhere in $K(x)$, but don't care where.

Problem: Find $F : X \to \mathbb{R}^D$ s.t.

- $F(x) \in K(x)$ for all x;
- C should be more or less as small as possible, subject to i.

More or less? Suppose F_1, F_2 both satisfy i. Look at Lipschitz constants $\|F_1\|_{Lip}, \|F_2\|_{Lip}$. Say F_1 does at least as well as F_2 if $\|F_1\|_{Lip} \leqslant C \|F_2\|_{Lip}$. (Why?) Not just about minimizing Lipschitz constant. Going from squiggly $\to$ smooth is big change. But smooth $\to$ super smooth is not big change, since noisy anyway. C only depends on D (?).

Note: Lipschitz norm is infimum of all possible constants. Lipschitz norm can be 0 for nonzero functions, so it's a seminorm. There are many other kinds of smoothness, like maximum of size of fourth derivatives. But only on a metric space means this is simplest to state.

Examples: suppose finite MS and associate line segment in plane. (Add drawing from ellipsoid MS to line segments.) $x \to K(x), y \to K(y)$. If x lies close to y, then F must place $F(x), F(y)$ close to each other. That may not be possible. Ex: if the line segments $K(x), K(y)$ very far apart! Lipschitz constant will be enormous. A simple way to guess how well we can do (i.e. how small make Lipschitz constants) is look at $d(x, y)$ vs $\min |K(x) - K(y)|$ (distance between sets, exists by compact). That can allow us to guess how well we can solve the Lipschitz selection problem.

Meh. Too optimistic! Take 3 closeby points in equilateral triangle. Draw 3 sides of big equilateral triangle, with $K(x), K(y), K(z)$. x close to y means put both $F(x), F(y)$ close to vertex. But when try to put $F(z)$ close... too far! Out of luck in trying to put everything close. Any 2 point subproblem is easy, but all 3 is much worse.

Fancier version of this with 4 points. Can make s.t. any 3 can be solved with small Lipschitz constant, but all 4 requires big Lipschitz constant.

But it turns out 4 point subproblems are good enough! If every 4 point subproblem can be solved with small Lipschitz constant, then the whole big problem can be solved with a small Lipschitz constant. This is a special case of a theorem. We will discuss 1 of the ideas of the proof of the theorem.

Theorem. *Given D, $\exists C(D)$ s.t. following holds:*

Let (X, d) be MS.

For each $x \in X$, associate $K(x)$ be compact convex subset of $\mathbb{R}^D$.

Suppose each $K(x)$ has dimension at most k. (i.e. a segment has dimension 1, a disc has dimension 2, an ellipsoid has dimension 3... A dimension of a set is the dimension of the smallest affine space that contains it.)

Assume that for every $k^\#$ point subset $S \subset X$ there exists solution to subproblem for S (i.e. $F^S : S \to \mathbb{R}^D$ s.t. $F(x) \in K(x) \forall x \in S$ and $\left\|F^S\right\|_{Lip} \leqslant 1$).

(Think of $k^\#$ as 4...)

Then there exists a nice solution for the problem on the whole MS, i.e. $F : X \to \mathbb{R}^D$ s.t. $F(x) \in K(x) \forall x \in X$ and $\|F\|_{Lip} \leqslant C(D)$.

Wait, what is $k^\#$? $= \min(2^D, 2^{k+1})$.

For ex: lines in the plane. $D = 2$, $k = 1$, $k^\# = 2^2 = 4$. If you can solve every 4 point subproblem with norm $\leqslant 1$, can solve whole problem with norm $\leqslant C(D)$.

This problem has been open a long time!

The question of fitting functions to data is an old and venerable subject. This line of inquiry goes back to beautiful papers by Whitney in 1934. In late 1970s and 80s, two Soviets Rutney and Pavel Shartsman were told the problem (the precise version, different?). They were told Whitney did important work, but weren't told what work. This predates the (widespread) internet, former Soviet Union. Couldn't get the 1934 paper! Instead guessed what Whitney had done, guessed

wrong, and came up with novel ideas! Lipschitz selection started.

Solved variations: affine space instead of convex compact. Professor Fefferman also worked in this field for 20 years. Turns out Schartsman and Professor Fefferman proved the theorem together! How did they do it, what got put together? What was the fact that was known for a long time by experts on metric spaces, but solved it?

Shartsman reduced the problem to MS (X, d) to special case of metric tree. (To be defined...) Professor Fefferman and his coworkers (Arie Israel and Kevin Luli, him) proved the theorem when (X, d) is Euclidean. Metric trees are very different from $\mathbb{R}^n$, but they had 1 crucial similarly... a hidden resemblance between $\mathbb{R}^n$ and metric trees. Let us explain the connection.

Graph = nodes with edges (draw example of tree, with no cycles). If you added an edge, it'd create a cycle. Tree = connected graph with no cycles. Given x and y (draw), there is 1 and only 1 nonredundant path from x to y. Label each edge with a positive number $(3.14159, 2.718, 17, 144, 53, 2021, 7)$. Think of these numbers as the length of the edges. The distance between x, y is the sum of the lengths along the path from x to y. This makes the nodes a MS!

Draw hydra graph. Start with 3 coming out of center, and keep drawing 2 heads coming out, so on forever. If you want to fit into sheet of paper, the edges need to get shorter and shorter. This is a discrete model of the unit disc with its non-Euclidean Poincare metric. (See screenshot.) Can declare that every edge has length 1. That means the geometry of this tree at infinity is radically different from $\mathbb{R}^n$. Imagine large ball of radius R in $\mathbb{R}^n$. Then volume is CR^n. If you do the same thing for metric tree, as $R \to \infty$, the number of points inside grows exponentially with R! Exponential in R. Compare R vs $R + 5$ ball in $\mathbb{R}^n$; negligible difference. But in this metric, a clear majority of the points lie outside the ball of radius R! The geometry of such a graph is very interesting and completely different from that of Euclidean space.

Nevertheless, this metric tree has a property common with Euclidean space, just what was needed for Professor Fefferman to adapt the proof to the case of metric trees.

Smidge of proof in Euclidean case: imagine trying to prove the result on big cube. (If we can, we can let size go to infinity and use compactness argument.) It turns out it's possible to measure the difficulty of a Lipschitz selection problem. Assign label between 0 and D – the bigger the label, the harder the problem. 0 means $K(x)$ are so huge that we can take F constant! When label D, studying general case.

Proof will proceed on induction on label. Base case trivial. Fix number between 0 and D. Assume can solve all labels less than that number. Calderon-Zygmund decomposition. Have big cube. Don't like it. Bisect it and produce 4 subcubes. Top left happy face, rest unhappy face. Keep upper left, bisect other 3. Keep going... Bisect one more time after that and happy with all! (Calling cubes, because this works in $\mathbb{R}^n$.) Such a procedure is Calderon-Zygmund decomposition. Need rule to say happy/unhappy: happy if know how to solve subproblem from that cube! Use nice cutoff function/partition of unity (that is Lipschitz) to put together local solutions and solve on entire cube. That's the idea for Euclidean proof!

... Does the metric tree look like cutting up in cubes? Turns out there is something like it, something like a decomposition into cubes that's true for a metric tree.

Nagata dimension: (X, d) MS is Nagata (D, c) ($D \in \mathbb{Z}$, $c > 0$) iff $\forall s$ (length scale, sidelengths of cubes) $\exists$ covering of X by subsets $\{X_\nu\}$ with 2 props:

i) Each X_ν has diameter $\leqslant s$. (Diameter = sup of distances.)

ii) Any ball of radius $c \cdot s$ meets at most $D + 1$ of the subsets $\{X_\nu\}$.

Think for a moment that X_ν is a bunch of cubes covering the plane with checkerboard of squares of diameter s, but if you take any circle of radius $s/10$, it meets at most $D + 1$ of the subsets.

Turns out Nagata dimension $D = n$ for $\mathbb{R}^n$. These X_ν are enough like a grid of cubes to adapt the proof to any MS that's Nagata (D, c).

How to study any metric tree?

Proposition (surprising): every metric tree is Nagata $(1, 1/16)$.

(Much more subtle proof: any metric graph embeds in the plane is dimension $\leqslant 4096$.)

Proof sketch (for hydra): How to make appropriate covering for any length scale s? Serious issue arises when s is large. If $s < 16$, then ball of radius $c \cdot s = 1$ only hits 1 thing, trivial. Fix s. Suppose $s = 2^k$. Draw slightly different graph, from left to right. Left to right binary tree starting from origin. For each x, find distance $x \to 0$. If $d(x, 0) \in [s \cdot \ell, s \cdot (\ell + 1)]$, say d is in Layer(ℓ).

Ancestor of x: pick out the single y on obvious path to origin from 0 to x s.t. $d(y, 0) = (\ell - 1) \cdot s$. Stop a little short.

Issue: if $x \in$ Layer(0), just say ancestor is origin.

Ready to construct X_ν. Let $X(z, \ell) = \{x \in \text{Layer}(\ell) | z = \text{Anc}(x)\}$.

First of all, its a covering. Every point x belongs to $X(\text{its ancestor}, \text{layer})$.

Next claim $\operatorname{diam} X(z, \ell) =\leqslant 4s$.

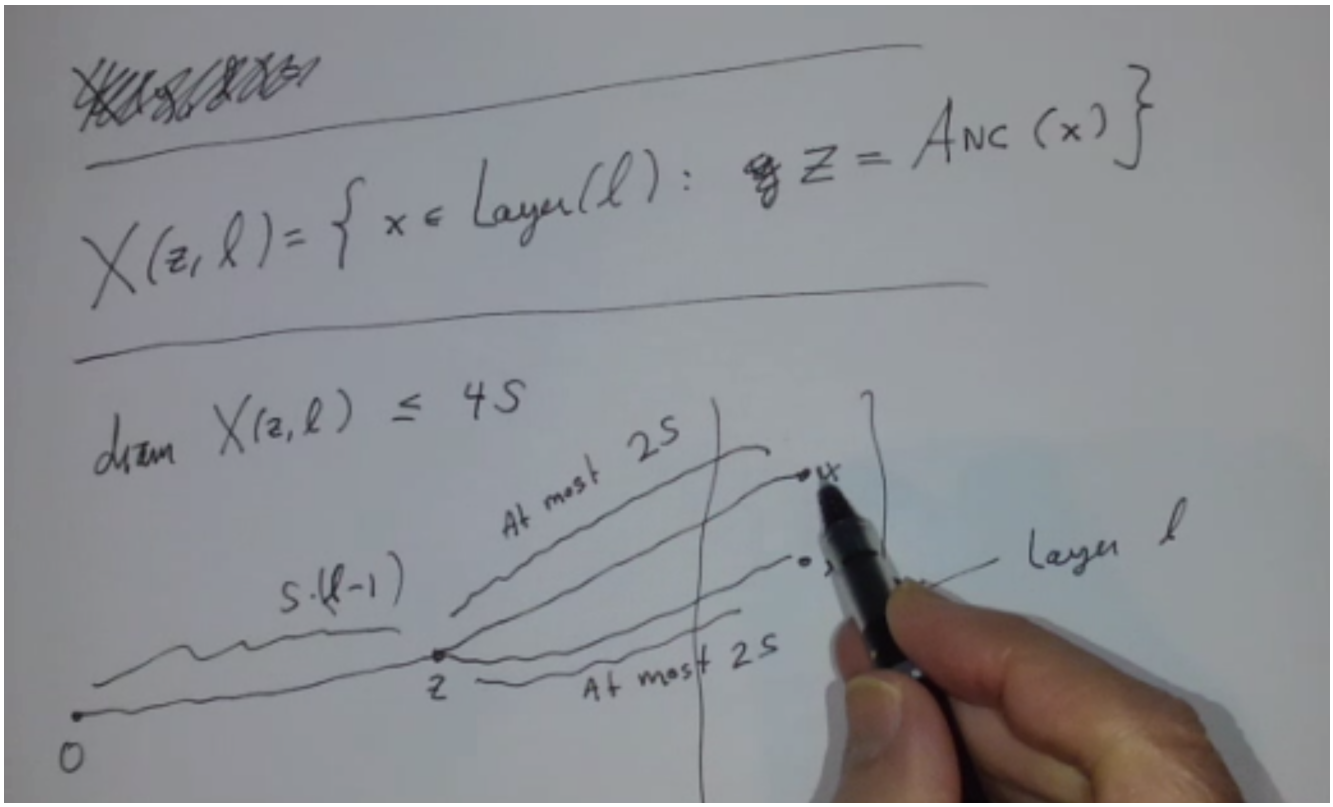
Draw layer 4... (missing?)

Finally we show any ball of radius $r \leqslant s/4$ meets at most 2 $X(z, \ell)$'s. Specifically, we should at most 1 for ℓ even and at most 1 for ℓ odd.

Just do even case. Claim only 1. FTSOC assume there are two, $X(z, \ell), X(z', \ell')$, with points x, x'. We will force x, x' to be far apart $\geqslant s$, hence no small ball covers both.

Case 1: $\ell = \ell'$. Then $d(x, 0) \in [s\ell, s(\ell + 1)]$ and $d(x', 0) \in [s\ell', s(\ell' + 1)]$. If $\ell \neq \ell'$, WLOG $\ell' \geqslant \ell + 2$. Thus $d(x, 0) \leqslant s\ell + 1$ and $d(x', 0) \geqslant s\ell'$, hence $d(x, x') \geqslant s$, as desired.

Case 2 (more interesting): $\ell = \ell'$. See drawing below. Consider $x \in X(z, \ell), x' \in X(z', \ell)$, $z \neq z'$. Shorest path $x \to x'$ transits via nearest common ancestor. How far do we travel? Go from x to z to divergence point of z, z', turn around and go to z', and continue to x'. As $d(z, x) \geqslant s$ and $d(z', x') \geqslant s$, that's $\geqslant 2s$, as desired.



Thus metric tree can be paved by bricks, in a sense.

1 Ideas for further reading

Whitney Extension Problem + 1934 papers

Machine learning

Topological data analysis

Full proofs of results:

https://www.ams.org/journals/jams/2014-27-01/S0894-0347-2013-00763-8/,

https://link.springer.com/content/pdf/10.1007/BF02921632.pdf

Metric trees

Metric graphs