Bi-Lipschitz embedding in Banach spaces

Joint work with Jeff Cheeger

Def. A map $f : X \to Y$ between metric spaces is L-biLipschitz if

$$\frac{1}{L}\, d(x, x') \leq d(f(x), f(x')) \;\leq L\, d(x, x')$$

for every pair of points x, $x' \in X$.

Problem. *Given metric spaces X and Y, is there a biLipschitz embedding of X into Y? If so, what is the optimal biLipschitz constant?*

In this lecture, Y will be a Banach space, e.g. L^p.

The embedding problems arising from Computer Science usually involve mapping finite metric spaces into Banach spaces, e.g. L^2 or L^1.

In many interesting cases, one can relate the discrete and continuous embedding problems.

Example. Suppose X is a compact metric space, and $\{X_i\}$ is a sequence of finite subsets of X which converges to X. Then any L-biLipschitz embedding

$$f : X \longrightarrow V$$

yields a sequence

$$\{f_i : X_i \to V\}$$

of L-biLipschitz embeddings.

For certain Banach spaces V, the converse holds: the existence of a sequence $\{f_i : X_i \to V\}$ of L-biLipschitz embeddings implies the existence of an L-biLipschitz embedding $X \to V$.

Note: This implication fails for ℓ^1, as we will see later.

Embedding results

(Kuratowski) Every metric space embeds in L^∞.

Embedding in $\mathbb{R}^n$. In order to biLipschitz embed in $\mathbb{R}^n$, a metric must have Hausdorff dimension at most n. In fact every R-ball must be coverable by at most

$$C\left(\frac{R}{r}\right)^n$$

r-balls, for all $r \leq R$.

Def. A metric space is **doubling** if there is a constant N such that any ball can be covered by at most N balls of half the radius.

Thm (Assouad). If (X, d) is a doubling metric space and $0 < \alpha < 1$, then there is a biLipschitz embedding

$$f : (X, d^\alpha) \to \mathbb{R}^n.$$

Here the dimension of the target and biLipschitz constant can be bounded in terms of the doubling constant of X and α.

Note: (X, d^α) contains no rectifiable curves.

Theorem. (Rademacher) Every Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable almost everywhere.

Corollary. If $(\mathbb{R}^k, \|\cdot\|_1)$ and $(\mathbb{R}^l, \|\cdot\|_2)$ are finite dimensional normed spaces and there is an L-biLipschitz embedding

$$(\mathbb{R}^k, \|\cdot\|_1) \to (\mathbb{R}^l, \|\cdot\|_2),$$

then there is also an L-biLipschitz linear embedding.

Theorem (Pansu) The Heisenberg group equipped with the Carnot-Caratheodory distance does not biLipschitz embed into $\mathbb{R}^n$, for any n.

The Heisenberg group is the 3-dimensional Lie group

$$H := \left\{ \begin{bmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \mid a,\, b,\, c \in \mathbb{R} \right\}.$$

Let X, Y, Z be the standard basis for the Lie algebra:

$$X := \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad Y := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$Z := \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Let g be the left invariant Riemannian metric making X, Y, and Z orthonormal.

Let ξ be the left invariant subbundle of the tangent bundle TH whose fiber at the identity is the 2-plane spanned by X and Y. This is a left-invariant contact structure on H.

Def. The **Carnot-Caratheodory distance** between two points p and q in H is the infimum of the lengths of curves

$$\gamma : [0, 1] \to H,$$

such that γ joins p to q and γ is everywhere tangent to ξ.

In coordinates, one may view ξ as the kernel of the 1-form

$$\theta = dz - xdy,$$

where the orthonormal basis of ξ at (x, y, z) is

$$\frac{\partial}{\partial x}, \quad \frac{\partial}{\partial y} + x\frac{\partial}{\partial z}\,.$$

$$d_{\mathrm{CC}}\,(\,(x, y, z_1), (x, y, z_2)\,) \;=\; C\,|z_1 - z_2|^{\frac{1}{2}}.$$

The Carnot distance is homogeneous, and for every $\lambda \in (0, \infty)$ there is a self-similarity $S_\lambda : H \to H$ of scale factor λ:

$$S_\lambda\left(\begin{bmatrix}1 & a & c\\ 0 & 1 & b\\ 0 & 0 & 1\end{bmatrix}\right) = \begin{bmatrix}1 & \lambda a & \lambda^2 c\\ 0 & 1 & \lambda b\\ 0 & 0 & 1\end{bmatrix}$$

Theorem (Pansu) Every Lipschitz map $f : H \to \mathbb{R}^n$ is (Carnot) differentiable almost everywhere.

The map is **differentiable** at $p \in H$ if the maps $f_\lambda : H \to \mathbb{R}^n$ given by

$$f_\lambda(g) = \frac{1}{\lambda}\left(f(p\, S_\lambda g) - f(p) \right)$$

converge as $\lambda \to 0$ to a Lie group homomorphism $Df : H \to \mathbb{R}^n$.

Problem. Characterize (meaningfully) metric spaces which biLipschitz embed in L^2 (or in $\mathbb{R}^n$ for some n).

Generalizations of Pansu's theorem

Def. A Banach space V has the **Radon-Nikodym property** if every Lipschitz map $\mathbb{R} \to V$ is differentiable almost everywhere.

Observations:

- V has the Radon-Nikodym property iff every Lipschitz map $\mathbb{R}^k \to V$ is differentiable almost everywhere.

- If V has the Radon-Nikodym property, then Pansu's differentiation theorem and the non-embedding corollary apply to Lipschitz maps $f : H \to V$.

Examples.

- Reflexive spaces: L^p, $1 < p < \infty$.

- Separable dual spaces: V is separable and $V = W^*$ for some Banach space W. The space of summable sequences, ℓ^1.

Nonexamples.

- L^∞.

- L^1.

Consider the "moving characteristic function", $f : [0, 1] \to L^1([0, 1])$,

$$f(t) = \chi_{[0,t]}.$$

Differentiation for more general domains

Theorem (Cheeger) Rademacher's theorem extends to Lipschitz functions

$$f : X \to \mathbb{R}$$

where (X, μ) is a doubling metric measure space satisfying a Poincare inequality.

Examples.

- (H, d_{CC}).

- Limits of sequences of Riemannian n-manifolds with a uniform lower bound on Ricci curvature.

- Fractal examples of Bourdon-Pajot, and Laakso.

A remarkable feature of Cheeger's work is that he is able to make sense of the idea of differentiation even though the domain has no affine structure or group structure.

He was able to use his differentiation theorem to give a general condition which implies biLipschitz nonembeddability into Euclidean space.

Theorem. (Cheeger-K.) The same differentiation theory applies to Lipschitz maps $f : X \to V$, where (X, μ) is a doubling metric measure space satisfying a Poincare inequality, and V is either a reflexive space or a separable dual space.

Cor. The Laakso and Bourdon-Pajot examples do not biLipschitz embed in any separable dual space. In particular they do not biLipschitz embed in the space of summable sequences, ℓ^1.

Recall that differentiation fails for Lipschitz maps $\mathbb{R} \to L^1$.

Theorem. (Cheeger-K.) The Laakso example admits a biLipschitz embedding into L^1.

Thus the Laakso space biLipschitz embeds in L^1 but not ℓ^1.

Conjecture. (Lee-Naor) The Heisenberg group does not biLipschitz embed in L^1.

Theorem If $f : H \to L^1$ is a Lipschitz map, then for almost every point $p \in H$, the map f collapses in the direction of the center:

$$\frac{d(f(p\ \mathrm{exp}(tZ), f(p))}{d(p\ \mathrm{exp}(tZ), p)} \to 0$$

as $t \to 0$. Here $Z \in L(H)$ generates the center of H.

Let $G \subset H$ be the integral Heisenberg group.

Cor. G does not biLipschitz embed into L^1, and the optimal biLipschitz constant for embeddings

$$B(e, r) \cap G \longrightarrow L^1$$

tends to infinity as $r \to \infty$.

Note: G may be viewed as the vertex set of a Cayley graph Γ, which is a homogeneous, bounded valence graph.

$$|B(e, r) \cap G| \leq C\, r^4.$$

Theorem (Lee-Naor) There is a metric ρ on H which is biLipschitz equivalent to d_{CC} such that

$$(H, \rho^{\frac{1}{2}})$$

isometrically embeds in L^2.

Cor. The Heisenberg group (H, d_{CC}) is a counterexample to the Goemans-Linial conjecture.

The conjecture had been disproved earlier by an example of Khot-Vishnoi.

The geometry of L^1

Def. Let S be a set. A **cut** of S is a decomposition

$$S = A \sqcup B .$$

To each such cut, one may associate a cut (pseudo)metric d_A:

$$d_A(x,y) = |\chi_A(x) - \chi_A(y)|;$$

i.e the distance between x and y is zero if they lie in the same subset of the cut, and 1 otherwise.

Lemma. Let d be a metric on a finite set S. Then (S,d) isometrically embeds into L^1 iff d is a nonnegative linear combination of cut metrics:

$$d = \sum_i a_i \, d_{A_i},$$

where $a_i \geq 0$ and $A_i \subset S$.

Generalization to infinite sets

Suppose (X, μ) is a measure space,

$$f : (X, \mu) \to L^1$$

is an L^1-mapping, and let

$$d_f : X \times X \to \mathbb{R}$$

be the pullback of the distance function on L^1:

$$d_f(x_1, x_2) \;=\; d(f(x_1), f(x_2)).$$

Then d_f has a representation as a superposition of cut metrics.

For simplicity, assume that $f(x) \in L^1$ is a nonnegative function for μ a.e. $x \in X$.

There is a Borel measure Σ on Meas(X) such that

$$d_f(x, x') = \int_{\text{Meas}(X)} d_E(x, x')\, d\Sigma(E),$$

and

$$(*) \quad \int_{\text{Meas}(X)} \mu(E)\, d\Sigma(E) < \infty.$$

Here Meas(X) denotes the collection of (equivalence classes of) finite measure subsets of X.

We refer to any Borel measure on Meas(X) satisfying $(*)$ as an L^1 cut measure.

Now suppose X is a ball in $\mathbb{R}^n$ or in the Heisenberg group.

Prop. If $f : X \to L^1$ is a Lipschitz map, then the associated cut measure Σ_f has finite perimeter:

$$\mathrm{PER}(\Sigma_f) := \int_{\mathrm{Meas}(X)} \mathrm{PER}(E)\, d\Sigma(E) < \infty.$$

Def. The **perimeter** of $E \subset X$ is the total variation of its characteristic function χ_E:

$$\mathrm{PER}(E) := \liminf_{f \to \chi_E} \int_X |\nabla f|\, d\mu,$$

where f is C^1 and the convergence is in L^1.

When $X = H$, the notation ∇f refers to the horizontal gradient.

Other ingredients in the proof

Def. A **half-space** in H is the inverse image of a half-plane under the epimorphism $H \to \mathbb{R}^2$, in other words a subset of the form

$$\left\{ \begin{bmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \;\middle|\; \alpha a + \beta b \geq \gamma \right\} \subset H.$$

Theorem (Ambrosio, Franchi-Serapioni-Serra-Cassano) If $E \subset H$ is a set of finite perimeter, then for

$$\text{per}(E) \quad a.e. \quad x \in H,$$

blow-ups of E centered at x converge to a half-space Y:

$$S_\lambda(\, x^{-1}H\,) \xrightarrow{L^1_{\text{loc}}} Y \quad \text{as} \quad \lambda \to 0.$$

The corresponding theorem for sets of finite perimeter in $\mathbb{R}^n$ is due to De Giorgi, circa 1955.

The main part of the argument involves producing a quantitative version of the AF-SSC regularity result for families of sets of finite perimeter.

Let d_f denote the pullback distance for a Lipschitz map $f : H \to L^1$ as before, and Σ denotes the corresponding cut measure. The families version of AFSSC implies that when one blows-up the cut measure Σ, it converges in a natural sense to a cut measure supported on half-spaces.

The final step is to observe that the distance function associated with a cut measure supported on half-spaces collapses in the direction of the center.