LECTURE NOTES FOR MATH 8250
RANDOM METRIC SPACES
SPRING 2019.

LECTURERS: MATTHEW KAHLE AND FACUNDO MÉMOLI

Contents

Date: June 6, 2019.

1. Lecture 1. Jan 7th. Scribe Katie

1.1. Metric Spaces: Definitions, Examples, and Invariants.

Definition 1.1 (metric space). A *metric space* is a pair (X, d_X) where X is a set and $d_X : X \times X \to \mathbb{R}_+$ such that

(1) $d_X(x, x') = d_X(x', x) \;\forall\; x, x' \in X$.

(2) $d_X(x, x') \geqslant 0$, with $d_X(x, x') = 0 \iff x = x'$.

(3) $d_X(x, x') + d_X(x', x'') \geqslant d_X(x, x'') \;\forall\; x, x', x'' \in X$. (triangle inequality, denoted Δ_1 in this lecture)

Example 1.2. Examples of Metric Spaces:

(1) $(\mathbb{R}, |\cdot|)$, $(\mathbb{R}^n, \|\cdot\|)$

(2) ultrametric spaces (UMS)

(3) tree metric spaces (TMS)

Definition 1.3 (ultrametric space). (X, u_X) metric is called *ultrametric* if it satisfies a strong triangle inequality (denoted by Δ_∞ in this lecture):

$$\max(u_X(x, x'), u_X(x', x'')) \geqslant u_X(x, x'') \text{ for all } x, x', x'' \in X.$$

Exercise 1.4. $\Delta_\infty \implies \Delta_1$.

Definition 1.5 (tree metric space). Finite metric space (X, d_X) is a *tree metric space* $\iff$ it satisfies the 4-point condition:

$$\max(d_X(x_1, x_2) + d_X(x_3, x_4), d_X(x_2, x_3) + d_X(x_1, x_4)) \geqslant d_X(x_1, x_3) + d_X(x_2, x_4).$$

Exercise 1.6. Finite (X, u_X) is a UMS $\implies$ it is a TMS.

Intuition: TMS are "simple" in the sense that there is a "TREE" underneath $\implies$ can draw them.

Goal: Measure failure to be a UMS/TMS.

Relaxation of being ultrametric $\to$ ultrametricity.

Definition 1.7 (ultrametricity). Let (X, d_X) be any metric space. Define the *ultrametricity* of X to be $\mathrm{ult}(X) = \inf\{\delta > 0 | \forall x, x', x'' \in X, \delta + \max(d_X(x, x'), d_X(x', x'')) \geqslant d_X(x, x'')\}$.

Question: Suppose that metric space (X, d_X) has $\mathrm{ult}(X) = \delta < \infty$. Does there exist an ultrametric u_X on X "close to d_X"?

Theorem 1.8 (Gromov, 1980s). *Given (X, d_X) finite m.s. with $\mathrm{ult}(X) = \delta$, there exists u_X ultrametric on X such that $\|d_X - u_X\|_\infty < c \cdot \delta \cdot \log(|X|)$ for some constant c.*

Definition 1.9 (single linkage map). Define the single linkage map $\mathcal{H}$ from a finite metric space to a finite ultrametric space as follows:
For a finite metric space (X, d_X), write $\mathcal{H}(X, d_X) = (X, u_X)$,
where $u_X(x, x') = \min\{\max_i d_X(x_i, x_{i+1}),\ \text{all } \underset{\substack{\| \\ x}}{x_0}, x_1, ..., \underset{\substack{\| \\ x'}}{x_n}\}$ for $x, x' \in X$.

Claim 1.10. u_X is a legit ultrametric on X.

Proof of Claim: Need to prove that for every $x, x', x'' \in X$,

$\max(u_X(x,x'), u_X(x',x'')) \geqslant u_X(x,x'')$. Take sequences

$\underset{\substack{\| \\ x}}{x_0}, x_1, ..., \underset{\substack{\| \\ x'}}{x_n}$ & $\max_i d_X(x_i, x_{i+1}) = u_X(x,x')$, $\underset{\substack{\| \\ x'}}{\bar{x}_0}, \bar{x}_1, ..., \underset{\substack{\| \\ x''}}{\bar{x}_m}$ & $\max_i d_X(\bar{x}_i, \bar{x}_{i+1}) = u_X(x', x'')$.

Let $\underset{\substack{\| \\ x}}{z_0}, z_1, ..., \underset{\substack{\| \\ x''}}{z_N}$ be the concatenation of both sequences. Then $\max_j d_X(z_j, z_{j+1}) \geqslant u_X(x, x'')$ by the definition of $u_X(x,x'')$.

By construction, $\max_j d_X(z_j, z_{j+1}) \leqslant \max(u_X(x,x'), u_X(x', x''))$, which proves the claim.

$u_X^* \leqslant d_X$.

Proposition 1.11. *u_X^* is the maximal sub-dominant ultrametric on (X, d_X).*

Let $\mathcal{U}^{\leqslant}(X) = \{u_X \text{ u.m. on } X \text{ such that } u_X \leqslant d_X\}$.

Then $u_X^*(x,x') = \sup\{u_X(x,x'), u_X \in \mathcal{U}(X)\}$, where $\mathcal{U}(X)$ denotes the collection of ultrametrics on X.

Exercise 1.12. Prove the previous proposition.

Idea of Proof of Gromov's Result. Consider $u_X = u_X^*$. Since $u_X^* \leqslant d_X$, we want to prove that $d_X \leqslant c \cdot \log(|X|) \cdot \mathrm{ult}(X) + u_X$.

Proof is by induction. For any points $x, x', x'' \in X$,

$$\delta + \max(u_X(x,x'), u_X(x',x'')) \geqslant u_X(x,x''), \quad (*)$$

where $\delta = \mathrm{ult}(X)$.

Claim is true when $|X| = 3$. For $|X| = 5$, consider points $\underset{\substack{\| \\ x}}{x_1}, x_2, x_3, x_4, \underset{\substack{\| \\ x'}}{x_5}$. Want to find an upper bound for $d_X(x,x') - u_X(x,x')$.

By (*),

$\max(d_X(x_1,x_2), d_X(x_2,x_3)) + \delta \geqslant d_X(x_1,x_3)$,

$\max(d_X(x_3,x_4), d_X(x_4,x_5)) + \delta \geqslant d_X(x_3,x_5)$.

Thus, $\max_i (d_X(x_i,x_{i+1})) + \delta \geqslant \max(d_X(x_1,x_3), d_X(x_3,x_5)) \geqslant d_X(x_1,x_5) - \delta$.

Therefore, $2\delta + u_X^*(x,x') \geqslant d_X(x,x')$. Can do this argument more generally to complete the proof. $\square$

Consider the map $\mathcal{H} : \mathcal{M}(X) \to \mathcal{U}(X)$, where X is a finite set,
$\mathcal{M}(X) = \{d : X \times X \to \mathbb{R}_+ \text{ such that } d \text{ is a metric on } X\}$, and
$\mathcal{U}(X) = \{u : X \times X \to \mathbb{R}_+ \text{ such that } u \text{ is an ultrametric on } X\}$.
Then $\mathcal{U} \subseteq \mathcal{M}$.

Theorem 1.13. *(Stability) $\mathcal{M}(X) \to \mathcal{U}(X)$ is 1-Lipschitz under $\|\cdot\|_\infty \implies$ for every $d, d' \in \mathcal{M}(X)$ with $d \mapsto u$ and $d' \mapsto u'$, $\|d - d'\|_\infty \geqslant \|u - u'\|_\infty$.*

Proof: Exercise.

Relaxation of being tree metric → hyperbolicity.

Definition 1.14 (hyperbolicity)**.** Let (X, d_X) be a compact metric space. Define the *hyperbolocity* as follows:
$\mathrm{hyp}(X) = \inf\{\delta > 0 | \forall\ x_1, x_2, x_3, x_4, \delta + \max(d_X(x_1, x_2) + d_X(x_3, x_4), d_X(x_1, x_3) + d_X(x_2, x_4)) \geqslant d_X(x_1, x_4) + d_X(x_2, x_3)\}$.

Theorem 1.15 (Gromov)**.** *For any finite metric space (X, d_X), there exists a tree metric t_X on X such that $\|d_X - t_X\|_\infty \leqslant c \cdot \log(|X|) \cdot \mathrm{hyp}(X)$ for some constant c.*

Proof: See Gromov's "Hyperbolic Groups" in book.

Metric Invariants:

Definition 1.16 (diameter, separation maps)**.**

Diameter map: $X \mapsto \mathrm{diam}(X) := \max_{x,x' \in X} d_X(x, x')$.
Separation map: $X \mapsto \mathrm{sep}(X) := \inf_{x \neq x'} d_X(x, x')$.

Definition 1.17 (distance preserving, isometry)**.**

(1). A map $f : (X, d_X) \to (Y, d_Y)$ is called *distance preserving* iff for all $x, x' \in X$, $d_X(x, x') = d_Y(f(x), f(x'))$. Any such map is often called an isometric embedding.
(2). f is an *isometry* between X and Y iff f is distance preserving and surjective.

(1). $f : X \to Y$ distance preserving $\implies f$ must be injective. Otherwise, there exists $x, x' \in X$, $x \neq x'$, such that $f(x) = f(x')$. But, this implies that $0 < d_X(x, x') = d_Y(f(x), f(x')) = 0$, a contradiction.
(2). By (1), an isometry $f : X \to Y$ is actually bijective. We say that $X \cong Y$.
(3). A map $\iota : \mathcal{M} \to \mathbb{R}$ is invariant if and only if for every $(X, d_X), (Y, d_Y) \in \mathcal{M}$ with $X \cong Y$, $\iota(X) = \iota(Y)$.

Quantum Mechanics Question: Can you identify a family $\{\iota_\alpha : \mathcal{M} \to \mathbb{R}\}_{\alpha \in A}$ of invariants such that if $\iota_\alpha(X) = \iota_\alpha(Y)$ for every $\alpha \in A$, then $X \cong Y$?

Yes, but not directly for $\mathbb{R}$; look at target space $\mathcal{T}_\alpha$ that depends on α. Let $A = \mathbb{N}, \alpha = n \in \mathbb{N}$, and $\mathcal{T}_n = \mathrm{pow}(\mathbb{R}_+^{n \times n})$, the collection of $n \times n$ square matrices. The map $\iota_n : \mathcal{M} \to \mathcal{T}_n$ where $(X, d_X) \mapsto \iota_n(X)$ is "the collection of all" $n \times n$ distance matrices induced by X.

Definition 1.18. Given $n \in N, (X, d_X)$, define the map $\Psi_X^{(n)} : \underbrace{X \times ... \times X}_{n \text{ times}} \to \mathbb{R}_+^{n \times n}$ by

$(x_1, ..., x_n) \mapsto \big((d_X(x_i, x_j)\big)_{i,j=1}^n$.

Example 1.19. When $n = 1$, $\Psi_X^{(1)}(x_1) = \big((0)\big)$.

When $n = 2$, $\Psi_X^{(2)}(x_1, x_2) = \begin{pmatrix} 0 & d_X(x_1, x_2) \\ d_X(x_1, x_2) & 0 \end{pmatrix}$.

When $n = 3$, $\Psi_X^{(3)}(x_1, x_2, x_3) = \begin{bmatrix} 0 & d_X(x_1, x_2) & d_X(x_1, x_3) \\ d_X(x_1, x_2) & 0 & d_X(x_2, x_3) \\ d_X(x_1, x_3) & d(x_2, x_3) & 0 \end{bmatrix}$.

Definition 1.20 (Curvature Sets)**.** $\kappa_n(X) = \mathrm{im}(\Psi_X^{(n)})$

Example 1.21. $\kappa_1(X) = \{(0)\}$.

$$\kappa_2(X) = \left\{ \begin{pmatrix} 0 & \delta \\ \delta & 0 \end{pmatrix}, \delta = d_X(x, x'),\ \text{for } x, x' \in X \right\}.$$

Theorem 1.22 (Gromov's Metric Space Reconstruction Theorem)**.** *Let $X, Y, \in \mathcal{M}$. If $\kappa_n(X) = \kappa_n(Y)$ for every $n \in \mathbb{N}$, then $X \cong Y$.*

Exercise 1.23. Prove Gromov's Metric Space Reconstruction Theorem.

Example 1.24. Let $X = S^1$ with angular metric.

$\kappa_1(X) = \{(0)\}$.

$$\kappa_2(X) = \left\{ \begin{pmatrix} 0 & \delta \\ \delta & 0 \end{pmatrix}, \delta \in [0, \pi] \right\}.$$

For $\kappa_3(X)$, think of the possible configurations of 3 points, x_1, x_2, x_3 on S^1. Let α, β, γ represent the angular distances between x_1 and x_2, x_2 and x_3, and x_1 and x_3, respectively. There are two cases:

Case 1: All three points lie on the same side of the origin, as illustrated in Figure 1 below.

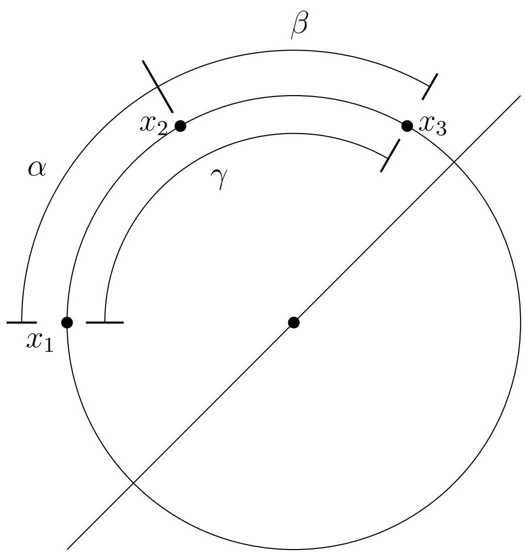
Then $\alpha + \beta = \gamma$. We can permute the three points to get $\alpha + \gamma = \beta$ and $\gamma + \beta = \alpha$ as well.

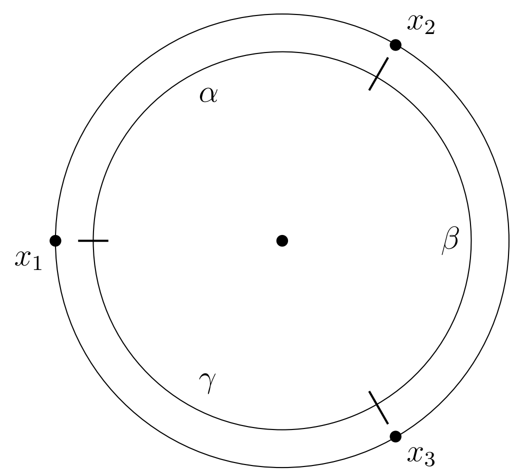
Case 2: Points x_1, x_2, x_3 do not all lie on the same side of the origin, as illustrated in Figure 2 below.

Then $\alpha + \beta + \gamma = 2\pi$.

Figure 1:

Figure 2:





Thus, any matrix $M \in \kappa_3(S^1)$ is of the form $\begin{bmatrix} 0 & \alpha & \beta \\ & 0 & \gamma \\ & & 0 \end{bmatrix} \to \mathbb{R}^3$.

$\kappa_3(S^1)$ is the tetrahedron with vertices $(0,0,0), (\pi, 0, \pi), (\pi, \pi, 0), (0, \pi, \pi)$. $\kappa_3(S^1) \cong S^2$.

Research Question: Compute $\kappa_n(S^1)$ for every n.

2. Lecture 2. Date Jan 10th. Scribe Austin

Definition 2.1. (X, d_X) metric space, $\epsilon \geqslant 0$.

(1) An ϵ-net for X is any subset $N \subseteq X$ such that $\bigcup_{x \in N} B_\epsilon(x) = X$.

(2) An ϵ-separated set in X is any $S \subseteq X$ such that $d_X(s, s') \geqslant \epsilon$ for all $s \neq s' \in S$.

A metric space (X, d_X) is totally bounded iff for all $\epsilon > 0$, it has a finite ϵ-net.

Exercise 2.2. Prove that

(1) if there exists $\frac{\epsilon}{3}$-net for X with cardinality n, then any ϵ-separated set in X has cardinality at most n

(2) A maximal ϵ-separated set in X is also an ϵ-net for X

2.1. **Quantification of Totally Boundedness.**

Definition 2.3. The covering function of a metric space (X, d_X) is given by

$$\begin{aligned} \mathrm{cov}_X : \mathbb{R}_+ &\to \mathbb{N} \\ \epsilon &\mapsto \mathrm{cov}_X(\epsilon) \end{aligned}$$

where $\mathrm{cov}_X(\epsilon) := \inf\{n \in \mathbb{N} : \exists \epsilon\text{-nets of } X \text{ with at most } n \text{ elements}\}$.

Definition 2.4. the packing function of X is given by

$$\begin{aligned} \operatorname{pack}_X : \mathbb{R}_+ &\to \mathbb{N} \\ \epsilon &\mapsto \operatorname{pack}_X(\epsilon) \end{aligned}$$

where $\operatorname{pack}_X(\epsilon) := \sup\{|S| : S \subseteq X \text{ is } \epsilon\text{-separated }\}$.

These functions are invariant under isometries.

Recall:

Theorem 2.5. *(X, d_X) m.s. is compact iff X is complete and totally bounded.*

Exercise 2.6. X compact implies both pack_X and cov_X are finite.

Recall:

- Notation: $\mathcal{M}$ is the collection of all metric spaces
- isometric embeddings ($\varphi : X \to Y$ is an isometric embedding iff $d_X(x, x') = d_Y(\varphi(x), \varphi(x'))$ for all $x, x' \in X$) and isometries (surjective isometric embedding)

Proposition 2.7. *If $f : X \to X$ distance preserving, and X is compact, then f is surjective (and therefore an isometry).*

Proof. Assume $p \in X \backslash f(X)$. f continuous implies $f(X)$ is compact, therefore closed. Thus there exists $\epsilon > 0$ so that $B_\epsilon(p) \cap f(X) = \varnothing$. Take $S \subseteq X$ maximally ϵ-separated. Let $n = |S|$ (exists by exercise, as X is compact and hence packing number is finite). Then $f(S)$ is also ϵ-separated because f is distance-preserving. Let $s \in S$. Then

$$d_X(p, f(s)) \geqslant \min_{x \in X} d_X(p, f(x)) \geqslant \epsilon$$

So $\{p\} \cup f(S)$ is also ϵ-separated, and has cardinality $n + 1$ contradicting maximality of S.

Hence f is surjective. □

Definition 2.8.
- $f : X \to Y$ is non-expanding iff $d_X(x, x') \geqslant d_Y(f(x), f(x'))$ for all $x, x' \in X$
- $f : X \to Y$ is non-contracting iff $d_X(x, x') \leqslant d_Y(f(x), f(x'))$ for all $x, x' \in X$

Theorem 2.9. *X compact metric space.*

(1) if $f : X \to X$ is non-expanding and surjective, then f is distance-preserving (hence isometry)

(2) if $f : X \to X$ is non-contracting, then f is distance-preserving

Proof. Exercise or see BBI. □

2.2. **Enlarging or Extending Metric Spaces.**

Enlarging metric spaces.

Example 2.10. Hausdorff "functor": take (X, d_X) compact metric space, and consider $C(X) = \{A \subseteq X : A \text{ closed}\}$. The Hausdorff distance associated to X is the function

$$\begin{aligned} d_H^X : C(X) \times C(X) &\to \mathbb{R}_+ \\ (A, B) &\mapsto \inf\{\epsilon > 0 : A \subseteq B^\epsilon \text{ and } B \subseteq A^\epsilon\} \end{aligned}$$

where

$$A^\epsilon := \{x \in X : \exists a \in A \text{ with } d_X(a, x) \leqslant \epsilon\}$$

Denote $H(X, d_X) = (C(X), d_H^X)$.

Theorem 2.11. *$(C(X), d_H^X)$ is a metric space. Furthermore, if X is compact then so is $(C(X), d_H^X)$ (variation of Prokhorov theorem).*

Finally, the function

$$\begin{aligned} j_X : X &\to C(X) \\ x &\mapsto \{x\} \end{aligned}$$

is an isometric embedding.

Example 2.12. Kuratowski embedding: $K(X) := (L^\infty(X), \|\cdot\|_\infty)$ where $L^\infty(X)$ is the space of bounded functions $f : X \to \mathbb{R}$.

Definition 2.13. The Kuratowski embedding is given by

$$\begin{aligned} k_X : X &\to L^\infty(X) \\ x &\mapsto d_X(x, \cdot) \end{aligned}$$

Proposition 2.14. *$k_X : X \to L^\infty(X)$ is distance-preserving.*

Proof. We have:

$$\mathrm{dist}(k_X(x), k_X(x')) = \|k_X(x) - k_X(x')\|_\infty = \|d_X(x, \cdot), d_X(x', \cdot)\| = \sup_{p \in X} |d_X(x, p) - d_X(x', p)|$$

By the triangle inequality for d_X,

$$d_X(x, x') + d_X(x', p) \geqslant d_X(x, p) \Rightarrow d_X(x, x') \geqslant d_X(x, p) - d_X(x', p)$$

Swapping x, x' implies

$$d_X(x, x') \geqslant |d_X(x, p) - d_X(x', p)|$$

Hence the above supremum is bounded by $d_X(x, x')$; as the bound is attained by $p = x$ or $p = x'$, it follows that $\mathrm{dist}(k_X(x), k_X(x')) = d_X(x, x')$ as desired. □

Definition 2.15. Filling radius of (M, g^M) an orientable Riemannian manifold of dimension n. Observe that $k_X(M) \subseteq L^\infty(M)$, so we can consider the thickening $(k_X(M))^\epsilon \subseteq L^\infty(M)$.

Consider $\ell_\epsilon : k_X(M) \to (k_X(M))^\epsilon$. This induces a map on homology: $(\ell_\epsilon)_\# : H_n(M) \to H_n(M^\epsilon)$. Then

$$\mathrm{FillRad}(M) := \inf\{\epsilon > 0 : \ell_\epsilon([M]) = 0.\}$$

The FillRad was defined by Gromov in the study of systolic inequalities.

Extensions of metric spaces. In what follows, let $(X, d_X) \in \mathcal{M}$ be compact.

(1) How does one add a point to X?

(2) How much freedom is there in that process?

Given a distance matrix on $X = \{x_1, \dots, x_n\}$:

$$d_X = \begin{array}{c|c} & \begin{array}{cccc} x_1 & x_2 & \dots & x_n \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array} & \begin{pmatrix} 0 & d_{12} & \dots & d_{1n} \\ d_{12} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ d_{1n} & \dots & \dots & 0 \end{pmatrix} \end{array}$$

We wish to add a new point x^*, and consider possible distance matrices on $X^* = X \cup \{x^*\}$:

$$d_{X^*} = \begin{array}{c|c} & \begin{array}{ccccc} x_1 & x_2 & \dots & x_n & x^* \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \\ x^* \end{array} & \begin{pmatrix} 0 & d_{12} & \dots & d_{1n} & \cdot \\ d_{12} & 0 & \ddots & \vdots & \cdot \\ \vdots & \ddots & \ddots & \vdots & f \\ d_{1n} & \dots & \dots & 0 & \cdot \\ \cdot & \cdot & f & \cdot & \cdot \end{pmatrix} \end{array}$$

How to choose f such that d_{X^*} is a legitimate metric on X^*? (only interesting restriction is triangle inequality). This leads to Katetov-functions on X:

$$\Delta_1(X) = \{f : X \to \mathbb{R}^+ : f(x) + f(x') \geqslant d_X(x, x') \geqslant |f(x) - f(x')|\}$$

Proposition 2.16. *d_{X^*} satisfies the triangle inequality iff $f \in \Delta_1(X)$.*

Proof. Exercise. (Write out triangle inequality for d_{X^*} for p, x, x', with $x, x' \in X$. □

2.3. **Universal Metric Spaces.** For now we know how to create various 'thickenings' $T(X)$ of a given metric space X. SO, $T(X)$ depends on X. Can we construct a metric space that fits all metric spaces at once?

Definition 2.17. A metric space (U, d_U) is Urysohn Universal if

(1) It is separable and complete (Polish)

(2) For any finite subset $X \subset U$, the following holds:

- Make X into a metric space by restricting d_U, i.e., $(X, d_U|_{X \times X} = d_X)$.
- Consider any one point extension (in the sense of Katetov) of (X, d_X), call it (X^*, d_{X^*}).
- Then there exists $u^* \in U$ so that $d_U(u^*, x) = f(x) = d_{X^*}(x^*, x)$ for all $x \in X$.

Theorem 2.18. *(Pavel Urysohn, 1920s) There exists at least one Urysohn universal space, U. Furthermore, any two Urysohn universal spaces are isometric.*

Proposition 2.19. *Any Polish space admits an isometric embedding into U.*

Proof. Likely presentations. □

Theorem 2.20. *(Vershik, 2004) (X_n) converges to an Urysohn Universal metric space.*

[This will be one of the topics for presentation as well.]

3. Lecture 3. Jan 14th. Scribe Jimin

Recall that given $X, Y \in \mathcal{M}$, an isometry between X and Y is any map $\phi : X \to Y$ such that ϕ is distance preserving and surjective.

ϕ being distance preserving means $d_X(x, x') = d_Y(\phi(x), \phi(x'))$ for all $x, x' \in X$.

Question: How do we relax this?

Definition 3.1 (Distortion)**.** For any isometry $\phi : X \to Y$, the distortion of ϕ is defined by

$$\mathrm{dis}(\phi) := \sup_{x,x' \in X} |d_X(x, x') - d_Y(\phi(x), \phi(x'))|$$

.

An idea is to let $\phi(X)$ be an ε-net for Y.

$$\phi(X) = Y$$

Definition 3.2. Given $\varepsilon > 0$, we say that X is ε-equivalent to Y, denoted by $X \cong_\varepsilon Y$, if there exists $\phi : X \to Y$ such that $\mathrm{dis}(\phi) < \varepsilon$ and $\phi(X)$ is an ε net for Y.

Now we consider the following definition.

Definition 3.3. $\hat{d}(X, Y) := \inf\{\varepsilon | X \cong_\varepsilon Y\}$

Exercise 3.4. Try to see that $\hat{d}$ does not satisfy the triangle inequality and also fails the symmetry.

Let's go back to the definition of an isometry. We define an isomery to be a map that is distance preserving and surjective. A map being distance preserving implies that it is injective. So we have a bijection.

Let's try to relax "there exists a bijection $\phi : X \to Y$ preserving distance exactly". If there exists a bijection $\phi : X \to Y$ then there exists $\phi^{-1} : Y \to X$ so that $\phi \circ \phi^{-1}$ & $\phi^{-1} \circ \phi$ are identity maps. How we make it preserve the distance exactly is what introduces the notion of distortion.

"Relaxation" is to find $\phi : X \to Y$ & $\psi : Y \to X$ such that $\phi \circ \psi$ & $\psi \circ \phi$ are ε - close to the identity respectively.

Definition 3.5. Given $X, Y \in \mathcal{M}$, $\phi : X \to Y$, and $\psi : Y \to X$,

$$\mathrm{codis}(\phi, \psi) := \sup_{\substack{x \in X \\ y \in Y}} |d_X(x, \psi(y)) - d_Y(\phi(x), y)|.$$

Remark 3.6. Assume that $\operatorname{codis}(\phi, \psi) < \varepsilon$. Then

$$d_X(x, \psi(y)) < \varepsilon \text{ for all } x \in X$$
$$d_Y(\phi(x), y) < \varepsilon \text{ for all } y \in Y$$

Proof. The assumption means that for all x in X and y in Y, the following holds.

$$|d_X(x, \psi(y)) - d_Y(\phi(x), y)| < \varepsilon.$$

Take $y = \phi(x)$. Then we have $|d_X(x, \psi(y))| < \varepsilon$. □

Definition 3.7 (Gromov-Hausdorff distance on $\mathcal{M}$)**.**

$$\operatorname{dist}(X, Y) = \frac{1}{2} \inf_{\phi,\psi} \max\{\operatorname{dis}(\phi), \operatorname{dis}(\psi), \operatorname{codis}(\phi, \psi) \text{ for } X, Y \in M\}$$

Question: Is this finite? Do we get finite number?

Take any point $x_0 \in X$, and $y_0 \in Y$. Suppose we have

$$\phi \text{ maps everything } \mapsto y_0,$$
$$\psi \text{ maps everything } \mapsto x_0.$$

Then

$$\operatorname{dis}(\phi) \leqslant \operatorname{diam}(X),$$
$$\operatorname{dis}(\psi) \leqslant \operatorname{diam}(Y),$$
$$\operatorname{codis}(\phi, \psi) \leqslant \operatorname{diam}(X) + \operatorname{diam}(Y).$$

Thus,

$$\operatorname{dist}(X, Y) \leqslant \frac{1}{2}(\operatorname{diam}(X) + \operatorname{diam}(Y)).$$

Exercise 3.8. Prove that $\operatorname{dist}(X, Y) \leqslant \frac{1}{2} \max(\operatorname{diam}(X), \operatorname{diam}(Y))$.

Comments: This is not the original definition given by Gromov in the 1980's. This is actually given by Kalton-Ostrovskii in 2000's.

Definition 3.9 (The original definition)**.** Given $X, Y \subset \mathcal{M}$, assume there exists (Z, d_Z) a sufficiently large/rich space such that X is isometrically embedded into Z by ι_X and Y is isometrically embedded into Z by ι_Y.

$$X \overset{\text{iso}}{\underset{\iota_X}{\hookrightarrow}} Z$$

$$Y \overset{\text{iso}}{\underset{\iota_Y}{\hookrightarrow}} Z$$

Consider

$$\inf\{d_H^Z(\iota_x(X), \iota_Y(Y)), \text{ all } (Z, \iota_X, \iota_Y)\} =: \operatorname{dist}(X, Y)$$

Remark 3.10. The two definitions agree.

Remark 3.11. Using Gromov's definition, we can prove **Exercise 5.8**.

Consider $Z = X \amalg Y$.

$$\begin{array}{cc} & \begin{array}{cc} X & Y \end{array} \\ \begin{array}{c} X \\ Y \end{array} & \begin{pmatrix} d_X & * \\ * & d_Y \end{pmatrix} \end{array} = d_Z$$

Choose $d_{(X,Y)} = \frac{1}{2}\max(\operatorname{diam}(X), \operatorname{diam}(Y))$. Then it's clear that $X \subset Y^\varepsilon$ for all ε. The real question is whether d_Z satisfies the triangle inequality on $X \amalg Y$.

Need to prove: $d_Z(x, x') \leqslant d_Z(x, y) + d_Z(y', x') = \max(\operatorname{diam}(X), \operatorname{diam}(Y))$ for all $x, x' \in X$.

Exercise 3.12. Prove the triangle inequality for dist using dis and codis definitions.

Some interpretation of Gromov's definition of d_{GH}

Definition 3.13 (3rd definition)**.** Given sets X and Y in $\mathcal{M}$ a correspondence between them is any subset $R \subset X \times Y$ such that

$$\pi_x(R) = X \text{ and } \pi_Y(R) = Y.$$

If X and Y are in compact metric spaces, then the distortion of a correspondence R between X and Y is defined by

$$\operatorname{dis}(R) := \sup_{(x,y),(x',y') \in R} |d_X(x, x') - d_Y(y, y')|.$$

$\Rightarrow$ $\mathrm{d}_{\mathrm{GH}}(X, Y) = \frac{1}{2}\inf_R \operatorname{dis}(R)$.

Exercise 3.14.
1) Suppose that you have $\phi : X \to Y$ and $\psi Y \to X$. Induce

$$\begin{aligned} R(\phi, \psi) &= \{(x, \phi(x)) : x \in X\} \cup \{(\psi(y), y) : y \in Y\} && (1) \\ &\subset X \times Y. && (2) \end{aligned}$$

Claim: $R(\phi, \psi)$ is a correspondence between X and Y.

2) $\operatorname{dis}(R(\phi, \psi)) = \max\{\operatorname{dis}(\phi), \operatorname{dis}(\psi), \operatorname{codis}(\phi, \psi)\}$.

Theorem 3.15 (Kalton-Ostrovskii, 2000s)**.** *All three definitions agree.*

Theorem 3.16 (Gromov, 1980s)**.** *$d_{\mathrm{GH}} : M \times M \to R_+$ satisfies*

(1) symmetry

(2) triangle inequality

(3) $d_{\mathrm{GH}}(X, Y) = 0$ iff $X \cong^{iso} Y$ for $X, Y \in \mathcal{M}$.

$(\mathcal{M}/\cong, d_{\mathrm{GH}})$ is a metric space .

Theorem 3.17 (P.Petersen, 2000s)**.** *d_{GH} is a complete metric on $\mathcal{M}/\cong$.*

Theorem 3.18 (Ivanov, 2015)**.** *d_{GH} is a geodesic metric.*

This means that for $X, Y \in \mathcal{M}$, there exists $\gamma : [0,1] \to \mathcal{M}$ such that

$$\gamma(0) \cong^{iso} X, \gamma(1) \cong^{iso} Y, \text{ and}$$
$$d_{\mathrm{GH}}(\gamma(s), \gamma(t)) = |t-s| d_{\mathrm{GH}}(X,Y) \text{ for all } t, s \in [0,1].$$

3.1. **Geodesics on $\mathcal{M}$. Claim**: If $X, Y \in \mathcal{M}$, then there exists an optimal correspondence R, a closed subset of $X \times Y$. i.e. There exists $R \subset X \times Y$ closed, a correspondence with $\operatorname{dis}(R) = 2d_{\mathrm{GH}}(X,Y)$.

Construction: Define $R^{opt}(X,Y) = \{R : \text{ optimal correspondence between } X \text{ and } Y\}$.

Given any $R \in R^{opt}(X,Y)$, we can construct γ_R, a geodesic between X and Y by

$$\gamma_R(t) = (R, d_t), \text{ for } t \in [0,1] \text{ where } d_t : R \times R \to \mathbb{R}_+.$$

Theorem 3.19 (Gromov's precompactness theorem)**.** *Let $N : \mathbb{R}_+ \to \mathbb{N}$ be given $D > 0$. Consider the class, $\mathcal{F}(N,D) \subset \mathcal{M}$,*

$$\mathcal{F}(N,D) := \{x \in M | \operatorname{diam}(X) \leqslant D, N_X(\varepsilon) \leqslant N(\varepsilon), \varepsilon > 0\}.$$

Then $\mathcal{F}(N,D)$ is totally bounded as a subset of $(\mathcal{M}, d_{\mathrm{GH}})$.

Applications to Riemannian geometry Groomv's precompactness theorem interacts well with Lower bounds on Ricci Curvature.

For $C \in \mathbb{R}$, a natural number m, and $D > 0$, let $\mathcal{R}(m,D,C)$ be the collection of all compact Riemannian manifolds M such that

$\dim(M) = m$
$\operatorname{diam}(M) \leqslant D$
$\text{Ricci} \geqslant C (C \in \mathbb{R})$.

Then $\mathcal{R}(m,D,C)$ is totally bounded in GH sense.

Question: How do we inject some randomness into these ideas? $\Rightarrow$ metric measure spaces

4. Lecture 4. Jan 16th. Scribe Woojin

Facundo Mémoli's lecture:

Recall the space $(\mathcal{M}, d_{\mathrm{GH}})$ of compact metric spaces equipped with the *Gromov-Hausdorff distance.*

4.1. Motivation of the Gromov-Wasserstein distance: we want a metric which is relevant/sensitive to statistical measurements.

Question 4.1. *Consider any two finite metric spaces (X, d_X) and (Y, d_Y). Take any (not necessarily different) two points x, x' in X in the uniformly random way and observe how often $d_X(x,x')$ is 1. If $d_{\mathrm{GH}}(X,Y)$ is small, can we say that the same experiment with (Y, d_Y) results in a similar result?*

The answer to the above question is *NO.*

Example 4.2. Let $\Delta_n(1) := (\{0,\ldots,n\}, d_{\Delta_n})$, where $d_{\Delta_n}(i,j) = 1 - \delta_{ij}$. For $\varepsilon > 0$, let $\Delta_n(\varepsilon) := (\Delta_n, \varepsilon \cdot d_{\Delta_n})$. As $\varepsilon \searrow 0$ and $n \nearrow \infty$,

- one can check that $d_{\mathrm{GH}}(\Delta_2(\varepsilon), \Delta_n(\varepsilon)) \searrow 0$.
- However, the probability of the event $d_{\Delta_2(\varepsilon)}(\cdot,\cdot) = \varepsilon$ is $1/2$, while the probability of the event $d_{\Delta_n(\varepsilon)}(\cdot,\cdot) = \varepsilon$ approaches to 1.

Motivated by the above example, we change our approach and consider the space $(\mathcal{M}^{\mathrm{w}}, d_{\mathrm{GW,p}})$ of *metric measure spaces* (mm-spaces) equipped with the *Gromov-Wasserstein distance*, which will be defined below.

4.2. mm-spaces and coupling measures.

Definition 4.3 (mm-spaces)**.** Let $\mathcal{X} = (X, d_X, \mu_X)$ be a compact metric space equipped with a Borel probability measure μ_X on (X, d_X). We call $\mathcal{X}$ a *metric measure space* (mm-space).

Notation 4.4 (Collection of mm-spaces)**.** By $\mathcal{M}^{\mathrm{w}}$, we denote the collection of all mm-spaces *with full supports.*

Definition 4.5 (Isomorphisms between mm-spaces)**.** Consider any two mm-spaces $\mathcal{X} = (X, d_X, \mu_X)$ and $\mathcal{Y} = (Y, d_Y, \mu_Y)$. We say that $\mathcal{X}$ and $\mathcal{Y}$ are *isomorphic* if there exists an isometry $\phi : (X, d_X) \to (Y, d_Y)$ with $\phi_\# \mu_X = \mu_Y$, i.e. the measure μ_Y is the push-forward measure of μ_X via ϕ. We write $\mathcal{X} \cong \mathcal{Y}$ in this case.

In order to introduce the *Gromov-Wasserstein distance between mm-spaces*, we first introduce the notion of *coupling measures (between measures)*, which is analogous to the notion of correspondences (between sets).

Definition 4.6 (Coupling measures)**.** Given any two probability spaces (X, μ_X) and (Y, μ_Y), let μ be a probability measure on $X \times Y$. We say that μ is a *coupling between μ_X and μ_Y* if $(\pi_X)_\# \mu = \mu_X$ and $(\pi_Y)_\# \mu = \mu_Y$.

Note that one typical example of coupling measure between μ_X and μ_Y is the product measure $\mu_X \otimes \mu_Y$.

Notation 4.7 (Collection of couplings)**.** By $\mathcal{U}(\mu_X, \mu_Y)$, we denote the collection of all coupling measures between μ_X and μ_Y.

4.3. The Gromov-Wasserstein distance.

We shall define the Gromov-Wasserstein distance between mm-spaces in an analogous way to the Gromov-Hausdorff distance between metric spaces. Let (X, d_X) and (Y, d_Y) be any two compact metric spaces. Recall that the 3rd version of the definition of the Gromov-Hausdorff distance $d_{\mathrm{GH}}((X, d_X), (Y, d_Y)$ was defined as $d_{\mathrm{GH}}((X, d_X), (Y, d_Y) = \frac{1}{2} \inf_R \mathrm{dis}(R)$ where the infimum is taken over all correspondences between X and Y, and

$$\mathrm{dis}(R) := \sup_{\substack{(x,y)\in R \\ (x',y')\in R}} |d_X(x,x') - d_Y(y,y')|.$$

In particular, let us write

$$S_{X,Y}((x,y),(x',y')) := |d_X(x,x') - d_Y(y,y')|. \tag{3}$$

Now we define the Gromov-Wasserstein distance. Let $\mathcal{X} = (X,, d_X, \mu_X)$ and $\mathcal{Y} = (Y, d_Y, \mu_Y)$ be any two mm-spaces. Take any $\mu \in \mathcal{U}(\mu_X, \mu_Y)$ (Definition 4.6). Note that the product measure $\mu \otimes \mu$ is a measure on $(X \times Y) \times (X \times Y)$.

Definition 4.8 (The p-th distortion of a coupling)**.** Let $\mathcal{X} = (X,, d_X, \mu_X)$ and $\mathcal{Y} = (Y, d_Y, \mu_Y)$ be any two mm-spaces. Pick any $\mu \in \mathcal{U}(\mu_X, \mu_Y)$. For $p \in [1, \infty)$, let us define the *p-th distortion of* μ as

$$\mathrm{dis}_p(\mu) := \left(\iint\limits_{(X \times Y) \times (X \times Y)} S_{X,Y}((x,y),(x',y'))^p \; d(\mu \otimes \mu) \right)^{\frac{1}{p}}$$

Remark 4.9 (Exercise)**.** One can check that

$$\lim_{p \to \infty} \mathrm{dis}_p(\mu) = \mathrm{dis}(\mathrm{supp}[\mu]),$$

where $\mathrm{supp}[\mu]$ is the support of μ.

Definition 4.10 (The Gromov-Wasserstein distance)**.** Let $\mathcal{X} = (X,, d_X, \mu_X)$ and $\mathcal{Y} = (Y, d_Y, \mu_Y)$ be any two mm-spaces. For $p \in [1, \infty)$, their p-th Gromov-Wasserstein distance is defined as

$$d_{\mathrm{GW,p}}(\mathcal{X}, \mathcal{Y}) := \frac{1}{2} \inf_{\mu \in \mathcal{U}(\mu_X, \mu_Y)} \mathrm{dis}_p(\mu).$$

Theorem 4.11. *$d_{\mathrm{GW,p}}$ is a legitimate metric on the quotient space $\mathcal{M}^{\mathrm{w}}/\cong$ (Notation 4.4, Definition 4.5).*

Remark 4.12 (Exercise: The metric space $(\mathcal{M}^{\mathrm{w}}, d_{\mathrm{GW,p}})$ is *not* complete[1])**.** For each $n \in \mathbb{N}$, consider Δ_n (Example 4.2) equipped with the uniform probability measure. It is known that for $m, n \in \mathbb{N}$ with $m \geqslant n$

$$d_{\mathrm{GW,1}}(\Delta_m, \Delta_n) \approx \frac{1}{2n}.$$

Think about what is the potential limit of the sequence $\{\Delta_n\}_{n \in \mathbb{N}}$ in $\mathcal{M}^{\mathrm{w}}$ and conclude that $(\mathcal{M}^{\mathrm{w}}, d_{\mathrm{GW,1}})$ is not complete.

Exercise 4.13 (Estimation of $d_{\mathrm{GW,p}}$)**.** Let $\mathbb{S}^n$ be the n-th sphere of radius 1 equipped with the geodesic distance and the normalized volume measure. How can we estimate $d_{\mathrm{GW,1}}(\mathbb{S}^m, \mathbb{S}^n)$? This problem can be a project problem. Also, see Remark 4.15 below.

Definition 4.14 (Covering number function)**.** Let (X, d_X) be a compact metric space. The covering number function $N_{(X,d_X)} : [0, \infty) \to \mathbb{N}$ of (X, d_X) is defined as

$$N_{(X,d_X)}(\varepsilon) := \inf\{n \in \mathbb{N} : X \text{ can be covered by } n \text{ open balls of radius } \varepsilon\}.$$

Remark 4.15 (Estimation of d_{GH})**.** By utilizing the *covering number function* of spheres, one can estimate the Gromov-Hausdorff distance $d_{\mathrm{GH}}(\mathbb{S}^m, \mathbb{S}^n)$ from below. Namely, we have the inequality:

$$2 \cdot d_{\mathrm{GH}}(\mathbb{S}^m, \mathbb{S}^n) \geqslant d_{\mathrm{I}}(N_{\mathbb{S}^n}, N_{\mathbb{S}^m}),$$

[1]This fact is in contract with the fact that $(\mathcal{M}, d_{\mathrm{GH}})$ is complete (proved by P.Peterson).

where d_{I} is the so-called *interleaving distance* (we do not deal with its precise definition for now). Also, we know

$$d_{\mathrm{GH}}(\mathbb{S}^m, \mathbb{S}^n) \leqslant \frac{1}{2} \max(\mathrm{diam}(\mathbb{S}^m), \mathrm{diam}(\mathbb{S}^n)) = \frac{\pi}{2}.$$

In general, the value $\frac{\pi}{2}$ is not identical to $d_{\mathrm{GH}}(\mathbb{S}^m, \mathbb{S}^n)$: It is known that $d_{\mathrm{GH}}(\mathbb{S}^1, \mathbb{S}^2) = \frac{\pi}{3}$.

Matthew Kahle's lecture:

4.4. **Expectation.** Let X be a random variable. We define the expectation $\mathbb{E}[X]$ of X.

- Let X be a nonnegative integer random variable. Then,
$$\mathbb{E}[X] = \sum_{i=0}^{\infty} i \cdot \mathbb{P}[X = i].$$
- Let X be a real-valued random variable with a density function f, i.e. $\mathbb{P}(X \in U) = \int_U f(x)\, dx$. Then,
$$E[X] = \int_{-\infty}^{\infty} f(x)\, dx.$$

Example 4.16 (Not every random variable has a mean)**.** Consider the following random variables:

- Let X be a positive integer random variable with the probability distribution
$$\mathbb{P}[X = i] = \frac{6}{\pi^2} \cdot \frac{1}{i^2},\ i \in \mathbb{N}.$$
- Let Y be a real-valued random variable with the density function $f : \mathbb{R} \to \mathbb{R}$ defined as
$$f(x) = \frac{\pi^{-1}}{x^2 + 1} \text{ (Cauchy's distribution).}$$

Proposition 4.17 (Linearity of expectation)**.** *For any two random variables A and B, and for any constant $c \in \mathbb{R}$, we have*

$$\mathbb{E}[A + B] = \mathbb{E}[A] + \mathbb{E}[B], \qquad \mathbb{E}[cA] = c \cdot \mathbb{E}[A].$$

Theorem 4.18 (Markov's inequality)**.** *If X is a nonnegative random variable, then for any $a > 0$,*

$$\mathbb{P}[X \geqslant a] \leqslant \frac{\mathbb{E}[X]}{a}.$$

Proof.

$$\begin{aligned} a \cdot \mathbb{P}[X \geqslant a] &\leqslant \sum_{i \geqslant a} i \cdot \mathbb{P}[X = i] \\ &\leqslant \sum_{i \geqslant 0} i \cdot \mathbb{P}[X = i] \\ &= \mathbb{E}[X]. \end{aligned}$$

∎

Exercise 4.19. Prove the Markov's inequality for real-valued random variables.

We introduce the notion of random graphs. See the following references: *Random graphs* by Bollobás, Janson, Riordan and *The probabilistic method* by Alon and Spencer (4th edition).

Definition 4.20 (Erdös-Rényi model (edge-independent model))**.** Let $n \in \mathbb{N}$ and $p \in [0, \infty)$. By $G(n,p)$, we mean a random graph with the vertex set $[n] := \{1, \dots, n\}$, where each edge appears with probability p independently.

Equivalently, the random graph $G(n,p)$ can be described as follows: if H is a certain graph on $[n]$ with exactly m edges,

$$\mathbb{P}[G(n,p) = H] = p^m \cdot (1-p)^{\binom{n}{2}-m}.$$

Example 4.21. Given a random graph $G(n,p)$, let X be the number of K_4 subgraphs in $G(n,p)$, where K_4 is the complete graph on 4 vertices. Then,

$$\mathbb{E}[X] = \binom{n}{4} p^6,$$

by the following argument: Let us index each 4-subset of $[n]$ by $i = 1, \dots, \binom{n}{4}$. Let X_i be the indicator random variable defined as

$$X_i = \begin{cases} 1, & \text{if 4-subset indexed by } i \text{ spans a } K_4 \text{ subgraph in } G(n,p), \\ 0, & \text{otherwise.} \end{cases}$$

Write X as a sum of indicator's random variables:

$$X = X_1 + \dots + X_{\binom{n}{4}}.$$

By using the linearity of expectation (Proposition 4.17), we can obtain the claim.

Example 4.22. Given a random graph $G(n,p)$, let Y be the number of induced copies of C_4, where C_4 is the cycle graph on 4 vertices. One can check that

$$\mathbb{E}[Y] = 3 \cdot \binom{n}{4} \cdot p^4 (1-p)^2.$$

In particular, the factor 3 above comes out of the fact that there are 3 different graphs on 4 vertices which are isomorphic to C_4.

Let us go back to Example 4.21. Assuming that $p << n^{-\frac{2}{3}}$, we can induce that as $n \to \infty$,

$$\mathbb{E}[\# \text{ of } K_4 \text{ subgraphs in } G(n,p)] \to 0$$

since $\binom{n}{4} p^6 << n^4 p^6 << 1$. Hence, the Markov's inequality tells us that if $p << n^{-\frac{2}{3}}$, then[2]

$$\mathbb{P}[\text{there exists } K_4 \text{ in } G(n,p)] \to 0 \text{ as } n \to \infty$$

because

$$\mathbb{P}[\text{there exists } K_4 \text{ in } G(n,p)] = \mathbb{P}[\# \text{ of } K_4 \text{ subgraphs} \geqslant 1] \leqslant \mathbb{E}[\# \text{ of } K_4 \text{ subgraphs}].$$

[2]Given any $f, g : \mathbb{N} \to \mathbb{R}$, we write $f << g$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.

What if $p >> n^{-\frac{2}{3}}$? Then, $\mathbb{E}[\# \text{ of } K_4 \text{ subgraphs}] \to \infty$ as $n \to \infty$. Does this imply that $\mathbb{P}[\text{there exists } K_4 \text{ in } G(n,p)] \to 1$? The answer is *NO*.[3]

Exercise 4.23. Give an example of a sequence $\{X_i\}_{i\in\mathbb{N}}$ of random variables such that $E[X_n] \to \infty$ but $\mathbb{P}[X_n = 0] \to 1$ as $n \to \infty$.

4.5. Variance.

Given a random variable X, we define the *variance* of X

$$\mathbf{Var}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2,$$

where the second equality is left as an exercise. Note that $\mathbf{Var}[X] \geqslant 0$ by definition. Also, we remark that some random variables do not admit its variance. We make use of σ to denote $\sqrt{\mathbf{Var}[X]}$, which is called the *standard deviation* of X.

Theorem 4.24 (Chebyshev's inequality). *Let X be any random variable with $\mathbb{E}[X] = \mu$ and $\mathbf{Var}[X] = \sigma^2$. Then,*

$$\mathbb{P}[|X - \mu| \geqslant \lambda\sigma] \leqslant \frac{1}{\lambda^2}.$$

Note that when $\lambda \leqslant 1$, the above theorem says nothing.

Exercise 4.25. Show that Chebyshev's inequality is best possible without more information about X.

5. Lecture 5. Jan 23rd. Scribe Gustavo

5.1. Second Moment Method.

Notation 5.1. We will use μ for $\mathbb{E}[X]$ and σ^2 for $\mathbf{Var}[E]$.

Proposition 5.2. *Suppose that $\{X_n\}$ is a sequence of random variables such that*

(1) $\mathbb{E}[X_n] \to \infty$, *and*

(2) $\mathbf{Var}[X_n] \ll \mathbb{E}[X_n]^2$,

then $\mathbb{P}[X_n > 0] \to 1$.

For a proof of Proposition 5.2, see Chapter 4 of *The Probabilistic Method*.

We may return now to Example 4.21 and conclude that

$$\mathbb{P}[\text{there exists } K_4 \text{ in } G(n,p)] \to 1$$

by checking condition (2) in Proposition 5.2. This is left as an Exercise.

Question 5.3. *In Example 4.21 and the previous comments, we covered two cases concerning the growth of p, namely, when $p \gg n^{-2/3}$ and when $p \ll n^{-2/3}$. What if $p = c\, n^{-2/3}$ for some constant $c > 0$?*

5.2. Three important distributions.

[3] In fact, it is true that $\mathbb{P}[\text{there exists } K_4 \text{ in } G(n,p)] \to 1$ as $n \to \infty$, but this is not induced from the Markov's inequality.

Poisson distribution. For any $\mu > 0$, a random variable X is Poisson with mean μ if

$$\mathbb{P}[X = k] = e^{-\mu}\,\frac{\mu^k}{k!}$$

for any $k \geqslant 0$. We denote this distribution by $\mathrm{Pois}(\mu)$.

Exercise 5.4. Check the following:

(1) $\sum_{k\geqslant 0} \mathbb{P}[X = k] = 1$.

(2) $\mathbb{E}[X] = \sum_{k\geqslant 0} k e^{-\mu}\,\frac{\mu^k}{k!} = \mu$.

Theorem 5.5 (Brun's Sieve)**.** *Suppose that $\{X_n\}$ is a sequence of random variables such that $\mathbb{E}[X_n] \to \mu$ and $\mathbb{E}[\binom{X_n}{r}] \to \frac{\mu^r}{r!}$ for every $r > 1$ as $n \to \infty$. Then $X_n \to \mathrm{Pois}(\mu)$ in distribution, that is, for every $k \geqslant 0$, we have $\mathbb{P}[X_n = k] \to e^{-\mu}\frac{\mu^k}{k!}$.*

Example 5.6. Let σ be a uniform random permutation in the symmetric group Σ_n on $[n]$, i.e., each permutation has probability $\frac{1}{n!}$. Define X_n as the number of fixed points. Then $\mathbb{E}[X_n] = 1$ for $n \geqslant 1$ (use linearity of expectation, note that $\mathbb{E}[\text{“1” is a fixed point}] = \frac{1}{n}$). Note also that for $n \geqslant r$,

$$\mathbb{E}\left[\binom{X_n}{r}\right] = \binom{X_n}{r}\,\frac{(n-r)!}{n!} = \frac{1}{r!}.$$

Hence $\mathbb{E}\left[\binom{X_n}{r}\right] \to \frac{\mu^r}{r!}$ for $\mu = 1$. Then, by Brun's Sieve, $X_n \to \mathrm{Pois}(1)$. As a corollary, $\mathbb{P}[\text{no fixed points}] = e^{-1}$.

Binomial distribution. For any $n \geqslant 1$ and $0 \leqslant p \leqslant 1$, we define the binomial distribution $\mathrm{Bin}(n,p)$ as the number of successes in n independent trials, where p is the probability of success in any given trial. We can easily check that

$$\mathbb{P}[\mathrm{Bin}(n,p) = k] = \binom{n}{k} p^k (1-p)^{n-k}$$

for any $0 \leqslant k \leqslant n$, and that $\mathbb{E}[\mathrm{Bin}(n,p)] = np$ (linearity of expectation).

Now consider $p = c/n$ for some c fixed and $n \to \infty$. Then $\mathbb{E}[\mathrm{Bin}(n,p)] = c$ for $n \geqslant 1$. Fix k. Letting $n \to \infty$ and since $1 - p \approx e^{-p}$ when $p \to 0$, we note that

$$\mathbb{P}[\mathrm{Bin}(n,p) = k] = \binom{n}{k} p^k (1-p)^{n-k} \approx \frac{n^k}{k!}\left(\frac{c}{n}\right)^k e^{-\frac{c}{n}(n-k)} \approx \frac{c^k}{k!} e^{-c},$$

which is precisely the Poisson distribution.

Exercise 5.7. Consider $G(n,p)$ where $p = c\, n^{-2/3}$ for some $c > 0$ fixed. Then

$$\mathbb{E}[\#K_4 \text{ subgraphs}] = \binom{n}{4} p^6 \to \frac{c^6}{24}$$

as $n \to \infty$. Show that $\#K_4$ subgraphs $\to \mathrm{Pois}(c^6/24)$. In particular,

$$\mathbb{P}[\text{no } K_4 \text{ subgraphs}] \to e^{-c^6/24}.$$

Normal distribution ("Gaussian"). The normal distribution $\mathcal{N}(0,1)$ (with mean 0 and variance 1) is a probability distribution on $\mathbf{R}$. Its probability density function is given by $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. In particular, for $a < b$,

$$\mathbb{P}[a \leqslant X \leqslant b] = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx.$$

Definition 5.8. Let $\{X_n\}$ be a sequence of random variables. We say that $\{X_n\}$ obeys a Central Limit Theorem (CLT) if

$$\frac{X_n - \mathbb{E}[X_n]}{\sqrt{\mathbf{Var}[X_n]}} \to \mathcal{N}(0,1)$$

in distribution. In other words, if for every $a < b$,

$$\mathbb{P}\left[a \leqslant \frac{X_n - \mathbb{E}[X_n]}{\sqrt{\mathbf{Var}[X_n]}} \leqslant b\right] \to \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$$

as $n \to \infty$.

Example 5.9.

- Let $Y_1, Y_2, Y_3, \ldots$ be independent identically distributed random variables with finite mean and variance. Let $X_n = \sum_{i=1}^n Y_i$. Then

  $$\frac{X_n - \mathbb{E}[X_n]}{\sqrt{\mathbf{Var}[X_n]}} \to \mathcal{N}(0,1)$$

  in distribution.
- Let $\mu_1, \mu_2, \ldots$ be any sequence of real numbers tending to ∞. Let $X_n = \mathrm{Pois}(\mu_n)$. Then $\{X_n\}$ obeys a CLT.
- Let $c_1, c_2, \ldots$ be a sequence of real numbers tending to ∞ (and such that $c_n \leqslant 1$). Let $X_n = \mathrm{Bin}(n,p)$ with $p = c_n/n$. Then $\{X_n\}$ obeys a CLT.

6. Lecture 6. Jan 28th. Ling Zhou

6.1. **Some exercises on Poisson distribution.**

Exercise 6.1. Suppose $X_1 = \mathrm{Pois}(\mu_1)$ and $X_2 = \mathrm{Pois}(\mu_2)$ are two independent random variables. Show that $X_1 + X_2 = \mathrm{Pois}(\mu_1 + \mu_2)$.

Exercise 6.2. Consider $G(n,p)$ and $p = \dfrac{\log n + C}{n}$ with a fixed constant $C \in \mathbb{R}$. Let X be the number of isolated vertices, where a vertex is isolated if it has degree zero. Observe that $\mathbb{E}(X) = n(1-p)^{n-1}$, which follows from that fact the expectation for the $i-$th vertex to be isolated is $(1-p)^{n-1}$. Show that $X \to \mathrm{Pois}(e^{-C})$, i.e. $\mathbb{P}(X=0) \to e^{-e^{-C}}$ as $n \to \infty$.

6.2. **Random Geometric Graphs.** The study of random geometric graphs (r.g.g.) is motivated by the followings:

- its nice application in statistics,
- it is a more 'realistic' model than $G(n, p)$ for many situations. For example, in social networks, we denote $A \sim B$ When two persons A, B are friends. Given $A \sim B$ and $B \sim C$, $\mathbb{P}[A \sim C]$ is high.

The basic idea of random geometric graphs is that the vertices are obtained by taking random points in a space. Our reference book is *Random Geometric Graphs* by Penrose.

Example 6.3. Choose n points independently identically distributed (i.d.d.) uniformly in $[0, 1]^d$, and connect every pair of points within distance r. In other words, $p \sim q \iff d(p, q) \leqslant r$. Here are a few comments for this example:

- usually $n \to \infty$, $r = r(n)$ depends on n,
- more general setting: consider a distribution on $\mathbb{R}^d$ with bounded measurable density functions,
- usually $d \geqslant 2$ is fixed,
- philosophical comment: $G(n, p)$ looks like an r.g.g. when $d \to \infty$ quickly.

Notation 6.4. A random geometric graph is denoted by $G(n, r)$ with n the number of vertices and r the distance of adjacency.

Proposition 6.5. *Let $d = 2$.*

- *if $r << n^{-3/4}$, then a.s.s. there are no K_3 subgraphs;*
- *if $r >> n^{-3/4}$, then a.s.s. there exists a K_3 subgraphs.*

Here a.s.s. means asymptotically almost surely, i.e. the probability goes to 1 as $n \to \infty$.

Proof. Claim $\mathbb{E}[\#K_3 \text{ subgraphs}] = cn^3r^4$ for some constant $c > 0$. Then the first statement follows from the claim and Markov's inequality, and the second statement follows from the claim and the second movement method.

Now we prove the claim. Let x, y, z be three vertices. Notice that $\mathbb{P}[y \sim x] = r^2$, $\mathbb{P}[z \sim x] = r^2$ and $\mathbb{P}[y \sim z | x \sim z, y \sim x] \geqslant \epsilon$ for some constant $\epsilon > 0$. Then

$$\mathbb{P}[x \sim y, y \sim z, x \sim z] = \mathbb{P}[y \sim z | x \sim z, y \sim x]\mathbb{P}[x \sim z]\mathbb{P}[y \sim x] \approx cr^2r^2 = cr^4.$$

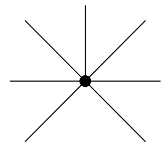
It follows that $\mathbb{E}[\#K_3 \text{ subgraphs}] \approx \binom{n}{3} cr^4 \approx \dfrac{c}{6} n^3r^4$. □

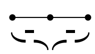
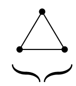
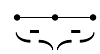
Proposition 6.6. *Let $d = 2$.*

- *if $r << n^{-4/6}$, then a.s.s. there are no K_4 subgraphs;*
- *if $r >> n^{-4/6}$, then a.s.s. there exists a K_4 subgraphs.*

Proposition 6.7. *Let H be a geometrically feasible induced subgraph (i.e. possible as an induced subgraph) in $\mathbb{R}^d$ with k vertices. Then $\mathbb{E}[\#$ induced subgraphs isomorphic to $H] \approx c_H n^k r^{d(k-1)}$.*

Example 6.8. $K_{1,7}$ is not feasible in $\mathbb{R}^2$:



Example 6.9 (subgraph counts)**.** #  $= 3 \cdot \#$  $+1 \cdot \#$ 

subgraphs induced graph induced graph

The formula in Proposition (6.7) suggests that something interesting may happen when $r \approx n^{-1/d}$.

"Percolation": $d \geqslant 2$.

Proposition 6.10. *There exists a constant λ_d such that*

- *if $r < (\lambda_d - \epsilon)n^{-1/d}$, then a.s.s. all components are of order $O(\log n)$;*
- *if $r > (\lambda_d + \epsilon)n^{-1/d}$, then a.s.s. there exists a unique giant component with $\Omega(n)$ vertices.*

Fact 6.11. If $r \geqslant (\frac{c_d \log n}{n})^{1/d}$, then a.s.s. $G(n,r)$ is connected.

sketch of proof for $d = 2$. Recall that the n vertices of $G(n,r)$ are chosen i.i.d. uniformly randomly in $[0,1]^2$. First, we divide $[0,1]^2$ into k^2 congruent squares, with $1/k$ on a side. The strategy is that if there is at least on point in every square and $r > 3k$, then the graph is connected.

Set $k = \sqrt{\frac{n}{c \log n}}$ with $c > 0$ to be determined. The area of a square is $\frac{1}{k^2} = \frac{c \log n}{n}$. Then

$$\mathbb{P}[(\text{no points in a given square})] = (1 - \frac{c \log n}{n})^n < e^{-\frac{c \log n}{n} n} = n^{-c}.$$

Choose $c > 1$. Then we apply union bound (i.e. $\mathbb{P}(A \text{ or } B) \leqslant \mathbb{P}(A) + \mathbb{P}(B)$) to get

$$\mathbb{P}[(\text{some square does not get any points})] \leqslant k^2 n^{-c} < nn^{-c} < n^{1-c} \to 0,$$

as $n \to \infty$. Thus, a.s.s. all squares contain a point. $\square$

Example 6.12 ("Coupon Collector" Problem)**.** Roll a fair 6−sided die. What is $\mathbb{E}[\#$ rolls before hitting every number at least once$]$? (Hint: the expected waiting time in a Bernoulli process is $1/p$.)

7. Lecture 7, February 4, Kritika Singhal

7.1. **Poisson Point Processes.** Let $d \geqslant 1$. A *Poisson point process* is a way of choosing random points in $\mathbb{R}^d$ with some "nice" properties.

Example 7.1. A uniform Poisson point process on $[0,1]^d$ of intensity $\lambda > 0$ has the following properties:

(1) *Property* 1: The total number of points is a Poisson random variable with mean λ. This implies that $\mathbb{E}[\text{total number of points}] = \lambda$, and $\mathbb{P}[\text{number of points} = k] = \frac{e^{-\lambda} \cdot \lambda^k}{k!}$.

(2) *Property* 2: If $U \subseteq [0,1]^d$ is a measurable subset of d-dimensional volume V, then the number of points in U is a Poisson random variable with mean $\lambda \cdot V$. This implies that $\mathbb{E}[\text{number of points in } U] = \lambda \cdot V$.

(3) *Property* 3: If $U, U' \subseteq [0,1]^d$ are disjoint, then the number of points in U and the number of points in U' are independent random variables. This property is referred to as *spatial independence.*

The three properties described above characterize a Poisson point process. We note that the property of spatial independence does not exist for random geometric graphs.

We now construct a Poisson point process on $[0,1]$ of intensity $\lambda > 0$ using the following steps:

(1) We first choose $n \sim \text{Pois}(\lambda)$, i.e. n is a Poisson random variable with mean λ.

(2) We then drop n points into $[0,1]$ uniformly, randomly and independently.

We check that the above Poisson point process has the desired properties.

(1) Property 1 is true by definition.

(2) For Property 2, we choose $U = [a,b]$, with $0 \leqslant a < b \leqslant 1$. We want to show that the number of points in U is a Poisson random variable with mean $\lambda(b-a)$. This is equivalent to showing that $\mathbb{P}[\text{number of points in } U = k] = \frac{e^{-\lambda(b-a)} \cdot (\lambda(b-a))^k}{k!}$ for every $k \geqslant 0$. Let $n \sim \text{Pois}(\lambda)$ be the total number of points in $[0,1]$. Then

$$\begin{aligned}\mathbb{P}[\text{number of points in } U = k] &= \sum_{i=0}^{\infty} \mathbb{P}[n = i] \cdot \mathbb{P}[\text{number of points in } U = k \mid n = i] \\ &= \sum_{i=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{i!} \binom{i}{k} (b-a)^k (1-(b-a))^{i-k} \\ &= e^{-\lambda(b-a)} \frac{(\lambda(b-a))^k}{k!}.\end{aligned}$$

The last equality is left as an exercise. The proof of Property 3 is the following exercise.

Exercise 7.2. Check if $0 \leqslant a < b < c < d \leqslant 1$, then the number of points in $[a,b]$ is independent of the number of points in $[c,d]$.

Hint: Check that $\mathbb{P}[\text{number of points in } [a,b] = k \text{ and number of points in } [c,d] = l] = \mathbb{P}[\text{number of points in } [a,b] = k] \cdot \mathbb{P}[\text{number of points in } [c,d] = l]$.

We now have two different models for obtaining random geometric graphs. In the first model, we take n points i.i.d uniformly in $[0,1]^d$ and connect two of them if they are close. In the second model, we take a uniform Poisson point process on $[0,1]^d$ of intensity n. We let $N \sim \text{Pois}(n)$ and take N i.i.d points uniformly randomly in $[0,1]^d$. In the second model, we have spatial independence which we do not have in the first model. Interestingly, we have that these two models are the same as $n \to \infty$. A reference for this is the section on Poissonization and de-Poissonization in Matthew Penrose's book "Random Geometric Graphs".

There are ways of going between these two models. A useful fact in this regard is that there are tail bounds or "Chernoff-type" bounds for $\mathrm{Pois}(n)$. An easy to use bound is the following: if $X \sim \mathrm{Pois}(n)$, then $\mathbb{P}[X > (1+\varepsilon)n] \leqslant e^{-\varepsilon^2 \cdot n}$. A sharper bound is shown in the following theorem. A reference for this is the book "Concentration of measures for the analysis of randomized algorithms" by Dubhashi and Panconesi.

Theorem 7.3. *Let $h : (-1,\infty) \to \mathbb{R}$ be defined as $h(u) = \frac{2(1+u)-\log(1+u)-u}{u^2}$. Let $\lambda > 0, x > 0$ and $X \sim \mathrm{Pois}(\lambda)$. Then, $\mathbb{P}[X \geqslant \lambda + x] \leqslant e^{\frac{-x^2}{2\lambda}\cdot h\left(\frac{x}{\lambda}\right)}$ and $\mathbb{P}[X \leqslant \lambda - x] \leqslant e^{\frac{-x^2}{2\lambda}\cdot h\left(\frac{-x}{\lambda}\right)}$.*

Another commonly studied object is the homogeneous Poisson point process on $\mathbb{R}^d$. This process has same properties as a Poisson point process, except property 1. This process gives an infinite random geometric graph, and is an object of study in continuum percolation or Gilbert disc model.

7.2. Brownian motion and scaling limits.

Brownian motion is a natural phenomenon, that is named after botanist Robert Brown. The corresponding mathematical object is a stochastic process called *Wiener process*.

Definition 7.4 (Wiener process)**.** A Wiener process is a random function from $\mathbb{R}_{\geqslant 0}$ to $\mathbb{R}$, that assigns to every time t, a W_t, satisfying the following properties:

1. $W_0 = 0$ almost surely (with probability 1).
2. W has independent increments, i.e. for $t > 0, u \geqslant 0$, $W_{t+u} - W_t$ is independent of past values of W_s for $s < t$.
3. The increments follow a Gaussian distribution, i.e. $W_{t+u} - W_u$ is normally distributed with mean 0 and variance u.
4. W is continuous with probability 1 (i.e. W_t is continuous in t).

One way to construct a Wiener process is the following: let $z_1, z_2, z_3, \ldots$ be i.i.d normally distributed random variables with mean 0 and variance 1. Then, for $0 \leqslant t \leqslant 1$

$$W_t = \sqrt{2}\sum_{n=1}^{\infty} z_n \cdot \frac{\sin\left(n-\frac{1}{2}\right)\pi t}{\left(n-\frac{1}{2}\right)\pi t}.$$

Alternately, we may write, for $0 \leqslant t \leqslant 1$,

$$W_t = z_0 t + \sqrt{2}\sum_{n=1}^{\infty} z_n \cdot \frac{\sin \pi n t}{\pi n}.$$

Another property satisfied by a Wiener process is that if $W : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ is a Wiener process, then for every $c > 0$, $V_t = \left(\frac{1}{\sqrt{c}}\right) W_{ct}$ is also a Wiener process. This shows that a Wiener process has some kind of fractal structure.

We now construct our own random walk. Let $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ be i.i.d Bernoulli random variables. We set, for $0 \leqslant t \leqslant 1$,

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{1 \leqslant k \leqslant [nt]} \varepsilon_k.$$

We have that W_n is a random function from $[0,1]$ to $\mathbb{R}$. We want to say that this function converges to a Brownian motion. We have that $W_n(0) = 0$ with probability 1. By central limit theorem, we have that $W_n(t) - W_n(s) \sim \mathcal{N}(0, t-s)$. Similarly, we also have independent increments. We invoke Donster's theorem that says that $W_n(t)_{[0,1]} \to W_{[0,1]}$ in a suitable function space (called "Skorokhod space"). Such a limit is called a *scaling limit*. We note that none of $W_n(t)$ are continuous, but they converge to a continuous function.

Another example of a scaling limit is the following: SLE (Schramm-Loewner evolution) is another random curve in plane, besides Brownian motion. $SLE(\kappa)$ is a fractal with dimension $1 + \frac{\kappa}{8}$. We consider loop erased random walk (random walk with all loops removed) in $\mathbb{Z}^2$. It was shown by Loewner and Schramm that this random walk has scaling limit $SLE(2)$ with dimension $\frac{5}{4}$.

We end with an open problem. Consider self-avoiding random walk in $\mathbb{Z}^2$. There is uniform measure on all simple paths of length n. It is conjectured that such a random walk has scaling limit $SLE\left(\frac{8}{3}\right)$.

8. Lecture 8. February 11. Jason Bello

8.1. Metric Measure Spaces.

Recall the following notation:

- if $\mathcal{X} = (X, d_X, \mu_X)$ is a metric measure space (m.m. space), then (X, d_X) is a compact metric space and μ_x is a Borel probability measure on X such that $\operatorname{supp}[\mu_X] = X$.
- $\mathcal{M}^w$ is the collection of all m.m. spaces.
- Isomorphism: $\mathcal{X} \cong^w \mathcal{Y}$ iff there exists and isometry $\phi : X \to Y$ such that $\phi_\# \mu_X = \mu_Y$.

Question 8.1. *What if we try to study/characterize a m.m. space $\mathcal{X}$ by taking statistical measurements for $\mathcal{X}$?*

Definition 8.2. Given $(X, d_x) \in \mathcal{M}$, $n \in \mathbb{N}$, consider the map $\Psi_X^{(n)} : X \times \cdots \times X \to \mathbb{R}^{n \times n}$ such that $(x_1, \ldots, x_n) \mapsto (d_x(x_i, x_j))_{i,j=1}^n$.

Remark 8.3. Curvature sets: $K_n(X) = \operatorname{Im}(\Psi_X^{(n)})$.

Imagine sampling n i.i.d. random points on X, $\bar{x}_1, \ldots, \bar{x}_n$ and then inducing the random variable $\Psi_X^{(n)}(\bar{x}_1, \ldots, \bar{x}_n)$. We want to understand the distribution of the induced random variable.

Definition 8.4. Let $\mathcal{X} = (X, d_X, \mu_x) \in \mathcal{M}^w$, $n \in \mathbb{N}$, and define the curvature measure as $U_X^{(n)} = (\Psi_X^{(n)})_\# \mu_X \otimes \cdots \otimes \mu_X \in \mathcal{P}_1(\mathbb{R}_+^{n \times n})$

Exercise 8.5. Prove $\operatorname{supp}[U_X^{(n)}] = K_n(X)$.

Recall that inside of $\mathcal{M}$, $(K_n(X))_{n\in\mathbb{N}}$ characterizes (X, d_X) up to isometry. Similarly, we have the following theorem.

Theorem 8.6 (M.m. Reconstruction Theorem). $U_{\mathcal{X}}^{(n)} = U_{\mathcal{Y}}^{(n)}$ *for all* $n \in \mathbb{N} \iff \mathcal{X} \cong^w \mathcal{Y}$.

Question 8.7. *Is this map* $U_{\bullet}^{(n)}$ *stable? (i.e. Lipschitz)*

How much information about $\mathcal{X}$ can be extracted from $U_{\mathcal{X}}^{(2)}$? from $U_{\mathcal{X}}^{(3)}$?

8.2. Curvature Measure for $n = 2$, $U_{\mathcal{X}}^{(2)}$:

$$\Psi_X(x_1, x_2) = \begin{pmatrix} 0 & d_X(x_1, x_2) \\ d_X(x_1, x_2) & 0 \end{pmatrix} \tag{4}$$

To understand $U_{\mathcal{X}}^{(2)}$ look at $dH_{\mathcal{X}} := (d_X)_\# \mu_X \otimes \mu_X \in \mathcal{P}_1(\mathbb{R}_+^{2\times 2})$. This is called the global distribution of distances on $\mathcal{X}$. So for $t \geqslant 0$,

$$dH_{\mathcal{X}}([0,t]) = ((d_X)_\# \mu_X \otimes \mu_X)([0,t]) = \mu_X \otimes \mu_X(D_X(t)) =: H_{\mathcal{X}}(t) \tag{5}$$

where $D_X(t) = \{(x, x') \in X \times X | d_X(x, x') \leqslant t\}$.

8.3. Examples. :

- Let $\Delta_2 = \left\{ \{p, q\}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, (1/2, 1/2) \right\}$, then

$$D_X(t) = \begin{cases} \{(p,p), (q,q)\} & \text{if } 0 \leqslant t < 1 \\ \{(p,p), (p,q), (q,p)(q,q)\} & \text{if } t \geqslant 1 \end{cases}$$

and $\mu_X \otimes \mu_X(\{(p,p),(q,q)\}) = (\mu_X(p))^2 + (\mu_X(q))^2 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. Thus,

$$H_{\Delta_2}(t) = \begin{cases} \frac{1}{2} & \text{if } 0 \leqslant t < 1 \\ 0 & \text{if } t \geqslant 1 \end{cases}.$$

- Let $\Delta_n = \left(\{p_1, \ldots, p_n\}, \begin{pmatrix} 0 & & 1 \\ & \ddots & \\ 1 & & 0 \end{pmatrix}, (\frac{1}{n}, \ldots, \frac{1}{n}) \right)$, then

$$D_X(t) = \begin{cases} \{(p_i, p_i) : i = 1, \ldots, n\} & \text{for } 0 \leqslant t < 1 \\ X \times X & \text{for } t > 1 \end{cases}$$

and so for $t \in [0,1)$, $\mu_{\Delta_n}^{\otimes 2}(D_{\Delta_n}(t)) = \sum_{i=1}^n \mu_{\Delta_n}^{\otimes 2}(\{(p_i,p_i)\}) = n \cdot \frac{1}{n^2} = \frac{1}{n}$ and for $t \geqslant 1$, $\mu_{\Delta_n}^{\otimes 2}(D_{\Delta_n}(t)) = 1$.

8.4. **What about S^n spheres?** Consider $(S^n, d_{S^n}, \mu_{S^n})$ where d_{S^n} is geodesic distance and μ_{S^n} is normalized volume measure. Then for $t \in [0, \pi]$

$$H_{S^1}(t) = \mu_{S^1} \otimes \mu_{S^1}(\{(x, x') \in S^1 \times S^1 | d_{S^1}(x, x') \leqslant t\}) \tag{6}$$

and for fixed $x_0 \in S^1$,

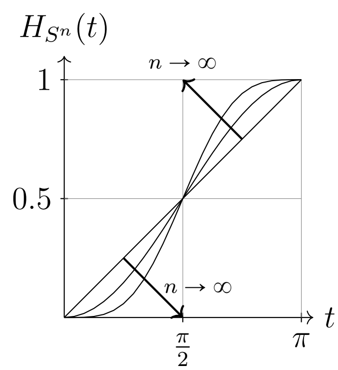
$$= \mu_{S^1}(\{x \in S^1 | d_{S^1}(x_0, x) \leqslant t\}) = \mu_{S^1}(\bar{B}_t(x_0)) = \frac{2t}{2\pi} = \frac{t}{\pi}. \tag{7}$$

Exercise 8.8. Prove $H_{S^2}(t) = \frac{1-\cos t}{2}$ for $t \in [0, \pi]$.

Proposition 8.9. *For $n \in \mathbb{N}$, $H_{S^n}(t) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi}} \int_0^t (\sin r)^{n-1} dr$ where Γ is the gamma function.(Disclaimer: may have forgotten some normalization constant.)*

Fact 8.10. $\mu_{S^n}(\bar{B}_t(x_0))$ does not depend on x_0 and

$$\mu_{S^n}(\{(x, x') \in S^n \times S^n | d_{S^n}(x, x') \leqslant t\}) = \mu_{S^n}(\bar{B}_t(x_0)) = H_{S^n}(t).$$



8.5. **Concentration of measure on spheres.** : Most of the mass of S^n lies in $(S^n_+)^\varepsilon$. In other words, any 2 random points on a high dimensional sphere are approximately orthogonal.

p-Diameters:

Definition 8.11. Let $p \in [1, \infty]$ then the p-diameter of $\mathcal{X}$ is

$$\operatorname{diam}_p(\mathcal{X}) = (\text{“}p\text{-moments of } dH_{\mathcal{X}}\text{”})^{1/p} = \left(\int_{\mathbb{R}_+} t^p dH_{\mathcal{X}}(dt)\right)^{1/p}$$

where $dH_{\mathcal{X}}(dt) = (d_X)_{\#}\mu_X \otimes \mu_X$.

Exercise 8.12. Show $\operatorname{diam}_p(\mathcal{X}) = (\iint_{X \times X} (d_X(x, x'))^p \mu_X(dx)\mu_X(dx'))^{1/p}$

Conclusion: $dH_\bullet$ discriminates between spheres of different dimensions. Thus, $U_\bullet^{(2)}$ discriminates between spheres of different dimensions.

Question 8.13. *How much information from $\mathcal{X}$ can we extract from:*

(1) $diam_{p_0}(\mathcal{X})$ for fixed $p_0 \in [1, \infty)$,

(2) $(diam_p(\mathcal{X}))_{p \geqslant 1}$?

The following proposition answers number 2.

Proposition 8.14. *s* $(diam_p(\mathcal{X}))_{p\in[1,\infty]}$ *determines* $dH_{\mathcal{X}}$.

<u>$\mathrm{diam}_p(S^n)$:</u>

- For $p = \infty$, $\mathrm{diam}_\infty(S^n) = \pi$ for all n.
- For $p = 1$, $\mathrm{diam}_1(S^n) = \pi/2$ for all n.

Proof. Let $a : S^n \to S^n$ be the antipodal map, then for all $p \in S^n$,

$$d_{S^n}(x,p) + d_{S^n}(a(x),p) = \pi = d_{S^n}(x,a(x)).$$

Integrating over p, we get

$$\int d_{S^n}(x,p)\mu_{S^n}(dp) + \int d_{S^n}(a(x),p)\mu_{S^n}(dp) = \pi$$

but since x was arbitrary,

$$\int d_{S^n}(x,p)\mu_{S^n}(dp) = \pi/2.$$

Now, integrating over x,

$$\mathrm{diam}_1(S^n) = \iint d_{S^n}(x,p)\mu_{S^n}(dp)\mu_{S^n}(dx) = \pi/2.$$

□

Exercise 8.15. $\mathrm{diam}_2(S^1) = \pi/\sqrt{3}$ and $\mathrm{diam}_2(S^2) = \sqrt{\frac{\pi^2}{2} - 1}$.

Global Distributions / Geometric Information that they carry:

Proposition 8.16. *Let* (M, g^M) *be an* m*-dimensional compact Riemannian manifold* $(\partial M = \emptyset)$ *where* $M = (M, d_M, \mu_M)$ *and* μ_M *is normalized volume measure. For* $0 < t \ll 1$*, we have the following Taylor expansion*

$$H_M(t) = \frac{w_m t^m}{Vol(M)}\left(1 - \frac{\int_M S_M(x)\mu_M(dx)}{6(m+2)}t^2 + O(t^4)\right)$$

where $Vol(M)$ *is the volume of a ball of radius* t *in* $\mathbb{R}^m$ *and* $S_M(x)$ *is scalar curvature. When* $m = 2$, $S_M(x) = Gaus(x)$ *and so the coefficient of* $t^m \cdot t^2$ *is* $\int_\mu Gaus(x)(dx)$ = *topological invariant by Gauss-Bonnet theorem.*

Proposition 8.17 (The case of smooth, planar, simple, closed)**.** *Let*

$$\mathcal{C} = (trace(\mathcal{C}), \|\cdot\|, \frac{length(\cdot)}{L})$$

where $\|\cdot\|$ *is Euclidean distance and* $L = length(\mathcal{C})$*. Then for* $0 < t \ll 1$,

$$H_{\mathcal{C}}(t) = \frac{2t}{L} + \frac{1}{12L^2}(\int_{\mathcal{C}} K^2(s)dst^3 + O(t^5)$$

where s *is arc-length.*[4]

[4]This expnasion can be found in http://www-users.math.umn.edu/~olver/vi_/hist.pdf.

Consider S^1 with Euclidean distance, $S^1 \subset \mathbb{R}$, then one can calculate by hand that $H_{S^1}(2) = \frac{2}{\pi}\arcsin(t/2)$. Thus, for $t > 0$ small

$$H_{S^1}(t) = \frac{t}{\pi} + \frac{t^3}{24\pi} + O(t^5)$$

by expanding arcsine since $\frac{\text{length(arc)}}{2\pi} = \frac{2\theta_t}{2\pi} = \frac{\theta_t}{\pi} = \frac{2}{\pi}\arcsin(t/2)$.

The following corollary is from https://arxiv.org/abs/1810.09646.

Corollary 8.18. *If $\mathcal{C}$ satisfies $H_{\mathcal{C}} = H_{S^1}$, then $\mathcal{C}$ is isometric to S^1.*

Proof. Since $H_{\mathcal{C}} = H_{S^1}$, their Taylor expansions must match. The coefficients of t must match, so $\frac{2}{L} = \frac{1}{\pi}$. Thus, $L = 2\pi$. Similarly, inequality of t^3 coefficients gives the following equality

$$\frac{1}{12L^2}\int_{\mathcal{C}} K^2(s)ds = \frac{1}{24\pi}$$

and so $\int_{\mathcal{C}} K^2(s)ds = 2\pi$. Recall the standard fact that for closed simple planar curves, $\int_{\mathcal{C}} K(s)\,ds = 2\pi$. Then, by Cauchy-Schwarz,

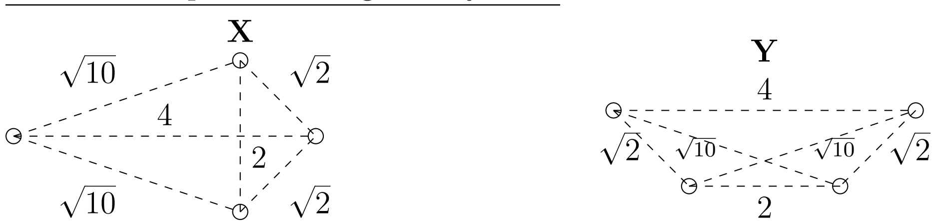
$$(2\pi)^2 = \left(\int_{\mathcal{C}} K(s)ds\right)^2 \leqslant \int_{\mathcal{C}} K^2(s)ds \cdot \int_{\mathcal{C}} ds = (2\pi)^2.$$

By the Cauchy-Schwarz conditions of equality, we have that $K \propto 1$ and thus K is constant. Hence, $\mathcal{C}$ is a circle of length 2π and so $C \cong S^1$. $\square$

Conjecture:
Simple, closed, planar curves are characterized by their global distance distributions.

Stronger Conjecture:
Bounded, closed subsets of $\mathbb{R}^2$ are characterized by their global distance distributions.

Counterexample for Stronger Conjecture:



The counterexample is due to Boutin and Kemper https://arxiv.org/abs/math/0311004.

Exercise 8.19. Show that $H_X = H_Y$.

9. Lecture 9. February 13th. Paul Duncan

Recall the example from last time of two different bounded, closed subsets of $\mathbb{R}^2$ with the same global distributions. Brinkmann and Olver observed experimentally that the curve versions are told apart by $H_\bullet$.

Conjecture: Planar simple closed curves are discriminated by $H_\bullet$.

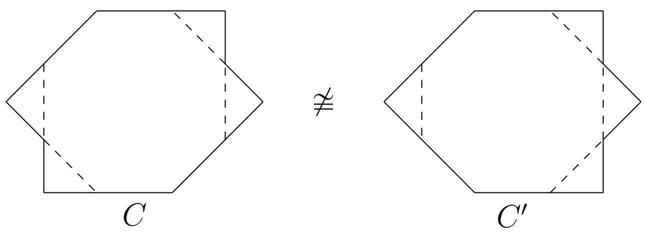
Proposition 9.1. *$S^1 \hookrightarrow \mathbb{R}^2$ is discriminated by $H_\bullet$.*

The conjecture is false.

Counterexample: $S^1 \subset \mathbb{R}^2$

$\mathcal{C} = \{$all planar simple closed curves $\subseteq \mathcal{M}$.

For every $\epsilon > 0, \exists C, C' \in B_\epsilon^{(\mathcal{M}, d_{GH})}(S^1 \subset \mathbb{R}^2)$ which satisfy $H_C = H_{C'}$, yet $C \ncong C'$, see https://arxiv.org/abs/1810.09646.



9.1. Concentration of Measure.

Recall, $\forall x_0 \in S^n, t \in [0, \pi]$,

$$H_{S^n(t)} = \mu_{S^n}(\overline{B_t(x_0)}) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})} \int_0^t (\sin r)^{n-1} dr$$

$(S^n_+)^\epsilon$ contains most of the mass for large n. This is a geometric manifestation of the so-called concentration of measure phenomenon.

Definition 9.2 (Concentration function of an mm-space)**.** Let $\mathcal{X} = (X, d_x, \mu_x) \in \mathcal{M}^W$. Let

$$\begin{aligned} \alpha_X(\epsilon) &= 1 - \inf\{\mu_X(A^\epsilon), A \subseteq X \text{ with } \mu_X(A) \geqslant 1/2\} \\ &= \sup\{\mu_X((A^\epsilon)^C), A \subseteq X \text{ with } \mu_X(A) \geqslant 1/2\} \end{aligned}$$

Theorem 9.3 (Paul Lévy, 1900s)**.** *For all $\epsilon > 0$,*

$$\alpha_{S^n}(\epsilon) \leqslant \sqrt{\frac{\pi}{8}} \exp(-\frac{n-1}{2}\epsilon^2).$$

Remark 9.4. $\mu_{S^n}((S^n_+)^\epsilon) \geqslant 1 - \sqrt{\frac{\pi}{8}} \exp(-\frac{n-1}{2}\epsilon^2)$

But the theorem guarantees something like this for any $A \subset S^n$ with $\mu_{S^n}(A) \geqslant \frac{1}{2}$.

9.2. Concentration of Measure - Gromov's point of view.

$\mathcal{X} = (X, d_X, \mu_X) \in \mathcal{M}^W$

"Quantum physics point of view"

(1) (X, d_X) "system"

(2) μ_X "state of the system"

(3) $f \in \mathrm{Lip}_1(\mathcal{X}, \mathcal{Y})$ "measurements"

Studying different $\mathcal{S}$ is interesting, but $\mathcal{S} = \mathbb{R}$ for us.

Idea: For every given measurement f, consider/compute the diameter of $f_{\#}\mu_X$, after discarding "noise."

Definition 9.5 (Partial Diameter)**.** Given $\mathcal{X} \in \mathcal{M}^W, \alpha \in [0,1]$, define

$$\mathrm{PartDiam}_\alpha(X) := \inf\{\mathrm{diam}(A), A \subseteq X, \mu_X(A) \geqslant \alpha\}.$$

Definition 9.6 (Observable diameter of an mm-space)**.** Given $\mathcal{X} \in \mathcal{M}^W, \kappa \in (0,1)$, define

$$\mathrm{ObsDiam}_\kappa(\mathcal{X}) := \sup\{\mathrm{PartDiam}_{1-\kappa}((\mathbb{R}, |-|, f_{\#}\mu_X)) | f \in \mathrm{Lip}_1(X,\mathbb{R})\}.$$

Definition 9.7. A sequence $(X_n)_{n\geqslant 1} \subseteq \mathcal{M}^W$ is called a Lévy-family iff for some $\kappa \in (0,1)$,

$$\mathrm{ObsDiam}_\kappa(X_n) \xrightarrow{n\to\infty} 0.$$

Pleas use $\mathbb{S}^n$ for spheres — not S^n

Theorem 9.8 (Gromov, Shioya, Funano,...)**.** *S^n is a Lévy family.*

Examples/Remarks

(1) $\mathrm{PartDiam}_\kappa(S^n) \overset{n\to\infty}{\not\to} 0$.

(2) Is $(\Delta_n)_{n\geqslant 1}$ a Lévy family?

(3) What about $\left(\{0,n\}, \begin{bmatrix} 0 & n \\ n & 0 \end{bmatrix}, \delta_0(1-1/n) + \delta_n(1/n)\right)$?

(1) Let $\rho = \mathrm{PartDiam}_\alpha(S^n) = \inf\{\mathrm{diam}(A), A \subseteq X, \mu_{S^n}(A) \geqslant \alpha\}$. Given $\epsilon > 0, \exists A_\epsilon \subseteq X$ with $\mu_{S^n}(A) \geqslant \alpha$ and $\mathrm{diam}(A_\epsilon) \leqslant \rho + \epsilon$. Then

$$\begin{aligned} \mathrm{diam}_1(S^n) &= \int_{S^n \times S^n} d_{S^n}(x,x')\mu_{S^n}(dx)\mu_{S^n}(dx') = \frac{\pi}{2} \\ &= \int_{A_\epsilon \times A_\epsilon} + 2\int_{A_\epsilon \times (S^n \setminus A_\epsilon)} + \int_{(S^n\setminus A_\epsilon)\times(S^n\setminus A_\epsilon)}. \end{aligned}$$

- $\int_{A_\epsilon \times A_\epsilon} d_{S^n}(x,x')\mu_{S^n}(dx)\mu_{S^n}(dx') \leqslant (\mu_{S^n}(A_\epsilon))^2(\rho+\epsilon) \leqslant \rho + \epsilon$
- $\int_{A_\epsilon \times (S^n\setminus A_\epsilon)} d_{S^n}(x,x')\mu_{S^n}(dx)\mu_{S^n}(dx') \leqslant \pi\mu_{S^n}(A_\epsilon)\mu_{S^n}(S^n\setminus A_\epsilon) \leqslant (1-\alpha)\pi$
- $\int_{(S^n\setminus A_\epsilon)\times(S^n\setminus A_\epsilon)} d_{S^n}(x,x')\mu_{S^n}(dx)\mu_{S^n}(dx') \leqslant \pi(1-\alpha)^2$

$\implies \frac{\pi}{2} \leqslant \rho + \epsilon + 2(1-\alpha)\pi(1+(1-\alpha))$

$\implies \rho \geqslant \frac{\pi}{2} - 2(1-\alpha)\pi(2-\alpha)$. Now, for $\alpha = 1-$ we obtain

$$\mathrm{PartDiam}_{1-\kappa}(\mathbb{S}^n) \geqslant \frac{\pi}{2} - 2\pi\kappa(1+\kappa) > 0 \text{ for } \kappa > 0 \text{ small.}$$

(2) To prove that $(\Delta_n)_{n\geqslant 1}$ is not a Lévy family, we need to convince ourselves that $\forall n \exists f_n \in \mathrm{Lip}_1(\Delta_n, \mathbb{R})$ such that $\mathrm{PartDiam}_{1-\kappa}(f_{n\#}\mu_n)$ is large.

Idea: Partition $\{1,...,n\} = A \sqcup B$ with $|A| \sim |B| \sim \frac{n}{2}$. Then consider f_n mapping A to a and B to b.

Claim: f_n is 1-lipschitz.

For this choice of f_n, $(f_n)_{\#}\mu_n$ has large $\mathrm{PartDiam}_{1-\kappa}$ for small κ.

(3) This is clearly a Lévy family.

Some properties:

Definition 9.9. (A poset structure on $\mathcal{M}^W$) We say that $\mathcal{X} \geqslant \mathcal{Y}$ in $\mathcal{M}^W$ iff $\exists \phi : X \to Y$ surjective, 1-lipschitz such that $\varphi_{\#}\mu_X = \mu_Y$.

Exercise 9.10. Prove that this is a poset structure (The fun part is showing that if $\mathcal{X} \leqslant \mathcal{Y}, \mathcal{Y} \leqslant \mathcal{X}$, then $X \cong^W Y$).

Proposition 9.11. *Let $\mathcal{X}, \mathcal{Y}$ in $\mathcal{M}^W, \alpha \in [0,1], \kappa \in (0,1)$. Then*

(1) $\mathrm{PartDiam}_\alpha(\mathcal{X}) \geqslant \mathrm{PartDiam}_\alpha(\mathcal{Y})$

(2) $\mathrm{ObsDiam}_\kappa(\mathcal{X})(\geqslant, \leqslant)\mathrm{ObsDiam}_\kappa(\mathcal{Y})$

(3) $\mathrm{ObsDiam}_\kappa(\mathcal{X}) \geqslant \mathrm{PartDiam}_{1-\kappa}(\mathcal{Y})$

(4) $\forall t > 0, \mathrm{ObsDiam}_\kappa(t\mathcal{X}) = t\mathrm{ObsDiam}_\kappa(\mathcal{X})$, *where* $t\mathcal{X} = (X, td_X, \mu_X)$.

Proof. Proof of (1):

Let $A_\varphi \subseteq X$ such that $\mu_X(A_\varphi) \geqslant \alpha$. Then it is enough to show that $\mathrm{PartDiam}_\alpha(Y) \leqslant \mathrm{diam}(A_\varphi)$. We have $\varphi(A_\varphi) \subseteq Y$ and

$$\mu_Y(\varphi(A_\varphi)) = \mu_X(\varphi^{-1}(\varphi(A_\varphi))) \geqslant \mu_X(A_\varphi) \geqslant \alpha.$$

Then, since $\mathrm{diam}(\varphi(A_\varphi)) \leqslant \mathrm{diam}(A_\varphi)$, we are done.

The rest of the proofs (including the direction of the inequality in (2)) are exercises.

□

Theorem 9.12 (Shioya)**.** *For $\kappa \in (0,1)$,*

(1) $\lim_{n\to\infty} \mathrm{ObsDiam}_\kappa(\sqrt{n}S_n) = \mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, |-|, \gamma^1)$, *where γ^1 is 1-dimensional Gaussian measure.*

(2) $\mathrm{ObsDiam}_\kappa(S^n) = O(\frac{1}{\sqrt{n}})$.

Proof. (1) This is one of the projects.

(2) Apply (1) and part (4) of the previous proposition.

□

Application: Let $(r_n) \subset \mathbb{R}_+$ and conisder (r_nS^n).

Corollary 9.13 (Shioya)**.** *$(r_nS^n)_{n\geqslant 1}$ is a Lévy family iff* $\frac{r_n}{\sqrt{n}} \xrightarrow{n\to\infty} 0$.

9.3. **Comparison Geometry for** $\mathrm{ObsDiam}_\kappa(-)$ **:.**

Theorem 9.14 (Bonnet-Myers)**.** *Let M be a compact Riemannian manifold of dimension n with $\partial M = \varnothing$ such that* $\mathrm{Ric}_M \geqslant (m-1)g^M$ *= the Ricci tensor on S^M. Then*

$$\mathrm{diam}(M) \leqslant \pi = \mathrm{diam}(S^n).$$

Theorem 9.15. *Let M be a m-dimensional Riemannian manifold with $\partial M = \varnothing$ and $\mathrm{Ric}_M \geqslant (m-1)g^M$. Then for any $\kappa \in (0,1)$,*

$$\mathrm{ObsDiam}_\kappa(M) \leqslant \mathrm{ObsDiam}_\kappa(S^m).$$

10. Lecture 10. February 18. Sam Mossing

10.1. **Recap: Concentration of Measure.** Let $\mathcal{X} = (X, d_X, \mu_X) \in \mathcal{M}^w$. Recall the following definitions:

- Concentration function:
$$\alpha_X(\epsilon) = 1 - \inf\{\mu_X(A^\epsilon), A \subset X \text{ measurable with } \mu_X(A) \geqslant 1/2\}$$
- Partial Diameter, ($\alpha \in (0,1)$):
$$\mathrm{PartDiam}_\alpha(\mathcal{X}) = \inf\{\mathrm{diam}(A) | A \subset X, \mu_X(A) \geqslant \alpha\}$$
- Observable Diameter, ($\kappa \in (0,1)$:
$$\mathrm{ObsDiam}_\kappa(\mathcal{X}) = \sup_{f \in \mathrm{Lip}_1(X,\mathcal{R})} \{\mathrm{PartDiam}_{1-\kappa}(f_\# \mu_X)\}$$
- Levy Family: $(\mathcal{X}_n)_n \subset \mathcal{M}^w$ is a Levy family if and only if $\mathrm{ObsDiam}_\kappa(X_n) \to 0$ for all $k \in (0,1)$

10.2. **Tensorization.** (Possible Paper: Estimate of observable diameter of ℓ_p product spaces: https://link.springer.com/content/pdf/10.1007/s00229-015-0730-1.pdf)

Let $\mathcal{X} = (X, d_X, \mu_X) \in \mathcal{M}^w$, $p \in [1, \infty)$. For all $n \in \mathbb{N}$ define $\mathcal{X}_{n,p} := (X^n, d_{X_{n,p}}, \mu_X^{\otimes n})$. Given $(x_1, ..., x_n), (x_1', ..., x_n') \in X^n$, define the metric $d_{X_{n,p}}((x_1, ..., x_n), (x_1', ..., x_n')) = (\sum_{i=1}^n d_X(x_i, x_i')^p)^{1/p}$

- Example: $X = \mathbb{S}^1$ (Tori)
- Example: $X = \{0, 1\}$ (Hamming cubes)

Theorem 10.1 (Ozawa-Shioya c. 2016)**.** *Let $\mathcal{X} \in \mathcal{M}^w$, $\kappa \in (0,1)$, $p \in [1,\infty)$. Then:*

$$\mathrm{ObsDiam}_\kappa(\mathcal{X}_{n,p}) \leqslant C_{\kappa,p}\, \mathrm{diam}(\mathcal{X})\, n^{\frac{1}{2p}}$$

where $C_{\kappa,1} = 4\sqrt{2\log(2/\kappa)}$ and $C_{\kappa,p} = 4 + 4\sqrt{2\log(2/\kappa)}$ for $p > 1$.

Example 10.2. Let $H_n = (\{0,1\}^n, d_{H_n}, \mathrm{unif})$ be the Hamming cube with uniform measure and distance metric

$$d_{H_n}((x_1, ..., x_n), (x_1', ..., x_n')) := \frac{\#\{i \in \{1, ..., n\} | x_i \neq x_i'\}}{n}.$$

For each n we can use the previous theorem with $X = \{0,1\}$ and $d_X(0,1) = diam(X) = \frac{1}{n}$ to see that $\mathrm{ObsDiam}_\kappa(H_n) \leqslant O(n^{-1/2})$. So by definition H_n is a levy family.

10.3. Relating $\alpha_X(\epsilon)$ to $\mathrm{ObsDiam}_\kappa(X)$.

Theorem 10.3. *For any $\mathcal{X} \in \mathcal{M}^w$, $\epsilon > 0$, $\kappa \in (0, 1/\epsilon)$:*

(1) $\mathrm{ObsDiam}_\kappa(X) \leqslant 2\alpha_X^{-1}(\kappa/2)$ *where, for $\nu \in (0,1)$, $\alpha_X^{-1}(\nu) = \inf\{\epsilon > 0 | \alpha_X(\epsilon) \leqslant \nu\}$ is the generalized inverse.*

(2) $\alpha_X(\epsilon) \leqslant \sup\{\kappa > 0 | \mathrm{ObsDiam}_\kappa(X) \geqslant \epsilon\} \leqslant \inf\{\kappa > 0 | \mathrm{ObsDiam}_\kappa(X) \leqslant \epsilon\}$

Proof. Postponed until later in lecture. □

Example 10.4. Applying item (2) of this theorem along with Levy's theorem (from the previous lecture) we see that $\mathrm{ObsDiam}_\kappa(\mathbb{S}^n) \leqslant \frac{4}{\sqrt{n-1}}\sqrt{\log(\frac{\sqrt{\pi/2}}{\kappa})} \to 0$. This gives us a proof showing that $\mathbb{S}^n$ is a Levy family.

Corollary 10.5. *A sequence $(\mathcal{X}_n)_n \subset \mathcal{M}^w$ is a Levy family if and only if $\alpha_{X_n}(\epsilon) \to 0$ for all $\epsilon > 0$.*

Proof. ($\Longrightarrow$) If$(\mathcal{X}_n)_n$ is a Levy family then for all $\kappa_0 \in (0,1)$, $\mathrm{ObsDiam}_{\kappa_0}(\mathcal{X}_n) \to 0$. So for every $\epsilon > 0$ there exists $N = N(\epsilon, \kappa_0) \in \mathbb{N}$ such that $\mathrm{ObsDiam}_{\kappa_0}(\mathcal{X}_n) \leqslant \epsilon$ for all $n \geqslant N$. Therefore by (2) we see that:

$$\alpha_{X_n}(\epsilon) \leqslant \inf\{\kappa > 0 | \mathrm{ObsDiam}_\kappa(\mathcal{X}_n) \leqslant \epsilon\} \leqslant \kappa_0 \text{ for } n \text{ large enough.}$$

Since κ_0 is arbitrary, this implies that $\alpha_{X_n}(\epsilon) \to 0$.

The other direction is similar and omitted. □

Example 10.6. Take the sequence $(\Delta_{2n})_n \subset \mathcal{M}^w$. Now we have the tools for another proof (using the concentration function α) that this is not a Levy family. Let $\epsilon \in (0,1)$. Since every distance between points $\{1, ..., 2n\}$ is either 0 or 1, we see for any subset $A \subset \{1, .., 2n\}$ that $A^\epsilon = A$. Thus:

$$\alpha_{\Delta_{2n}}(\epsilon) = 1 - \inf\{\mu_{\Delta_{2n}}(A^\epsilon) | A \subset \{1, ..., 2n\}, \mu_{\Delta_{2n}}(A) \geqslant \frac{1}{2}\} = 1/2 \nrightarrow 0.$$

10.4. Levy Radius.

There exists another invariant called Levy Radius that we can use to mediate between the observable diameter and the concentration function. The goal of this invariant is to detect how close any given 1-Lipschitz function is to being constant.

Definition 10.7. Given $\mathcal{X} = (X, d_X, \mu_X) \in \mathcal{M}^w$ and $f \in \mathrm{Lip}_1(X, \mathbb{R})$, we say that $a_0 \in \mathbb{R}$ is a pre-Levy mean of f if $(f_\# \mu_X)(-\infty, a_0] = (f_\# \mu_X)[a_0, \infty) \geqslant \frac{1}{2}$. Two remarks: pre-Levy means always exist but are not necessarily unique, and these are sometimes called medians.

Exercise 10.8. The set of all pre-Levy means of f is a bounded, closed interval, say $A_0(f)$.

Definition 10.9. The Levy mean of f, denoted m_f is:

$$m_f = \frac{\min(A_0(f)) + \max(A_0(f))}{2}.$$

Definition 10.10. The Levy Radius of $\mathcal{X}$ is defined as follows. Fix $\kappa \in (0,1)$. Given $\rho > 0$ and $f \in \mathrm{Lip}_1(X, \mathbb{R})$,

say property $C_\kappa^f(\rho)$ is true if $\mu_X(\{x \in \mathcal{X} : |f(x) - m_f| \geqslant \rho\}) \leqslant \kappa$.

Define:

$$\mathrm{LevyRad}_\kappa(\mathcal{X}) = \inf\{\rho > 0 : \forall f \in \mathrm{Lip}_1(X,\mathbb{R}), C_\kappa^f(\rho) \text{ is true}\}$$

10.5. **Relating Levy Radius and Observable Diameter.**

Lemma 10.11. *For all* $\kappa \in (0,1)$,

$$\mathrm{ObsDiam}_\kappa(\mathcal{X}) \leqslant 2\,\mathrm{LevyRad}_\kappa(\mathcal{X}).$$

Proof. Let $\rho > \mathrm{LevyRad}_\kappa(\mathcal{X})$. So for $f \in \mathrm{Lip}_1(X,\mathbb{R})$, $\mu_X(\{x \in X : |f(x) - m_f| \geqslant \rho\}) \leqslant \kappa$. Thus $(f_\# \mu_X)([m_f - \rho, m_f + \rho]) \geqslant 1 - \kappa$. So $\mathrm{PartDiam}_{1-\kappa}(f_\#\mu_X) \leqslant 2\rho$, and thus

$$\mathrm{ObsDiam}_\kappa(\mathcal{X}) = \sup_{f \in \mathrm{Lip}_1(X,\mathbb{R}} \mathrm{PartDiam}_{1-\kappa}(f_\#\mu_X) \leqslant 2\rho.$$

Since $\rho > \mathrm{LevyRad}_\kappa(\mathcal{X})$ is arbitrary, we take the infimum over all such ρ to conclude the result. ◻

Lemma 10.12. *For all* $\kappa \in (0, \frac{1}{2})$,

$$\mathrm{LevyRad}_\kappa(\mathcal{X}) \leqslant \mathrm{ObsDiam}_\kappa(\mathcal{X}).$$

Proof. Let $a = \mathrm{ObsDiam}_\kappa(\mathcal{X})$. Pick any $f \in \mathrm{Lip}_1(X,\mathbb{R})$, then $\mathrm{PartDiam}_{1-\kappa}(f_\#\mu_X) \leqslant a$. So by definition of partial diameter, for any $\epsilon > 0$ there exists $A_\epsilon \subset \mathbb{R}$ measurable with the following properties: $diam(A_\epsilon) \leqslant a+\epsilon$ and $f_\#\mu_X(A_\epsilon) \geqslant 1-\kappa$. Now we smooth out the set A_ϵ by taking $\ell_\epsilon = \inf A_\epsilon$, $r_\epsilon = \sup A_\epsilon$, $I_\epsilon = [\ell_\epsilon, r_\epsilon]$. Observe that $I_\epsilon \supset A_\epsilon$ with $(f_\#\mu_X)(I_\epsilon) \geqslant 1-\kappa$ and $diam(I_\epsilon) = diam(A_\epsilon) \leqslant a + \epsilon$. Next we claim $m_f \in I_\epsilon$.

Assume the claim for now. So $I_\epsilon \subset [m_f - (a+\epsilon), m_f + (a+\epsilon)]$. Thus

$$\mu_X(\{x \in X : |f(x) - m_f| \leqslant a+\epsilon\}) = (f_\#\mu_X)([m_f - (a+\epsilon), m_f + (a+\epsilon)]) \geqslant (f_\#\mu_X)(I_\epsilon) \geqslant 1-\kappa$$

Thus $\mu_X(\{x \in X : |f(x) - m_f| \geqslant a+\epsilon\}) \leqslant \kappa$, and so by definition of the Levy Radius we have $\mathrm{LevyRad}_\kappa(\mathcal{X}) \leqslant a + \epsilon$. Since $\epsilon > 0$ is arbitrary we can take the infimum over ϵ to conclude:

$$\mathrm{LevyRad}_\kappa(\mathcal{X}) \leqslant a = \mathrm{ObsDiam}_\kappa(\mathcal{X}).$$

Proof of claim: assume for contradiction that $m_f \notin I_\epsilon$. There are two cases, $m_f < \ell_\epsilon$ or $m_f > r_\epsilon$. In the second case $I_\epsilon \subset (-\infty, m_f)$. Also, recall that $\kappa \in (0, 1/2)$ by assumption, and so $1 - \kappa > 1/2$. So using this along with the definition of m_f we see that:

$$\frac{1}{2} < 1 - \kappa \leqslant (f_\#\mu_X)(I_\epsilon) \leqslant (f_\#\mu_X)((-\infty, m_f)) = \frac{1}{2}.$$

This gives us a contradiction. The proof for the first case is similar and omitted. ◻

Corollary 10.13. $(\mathcal{X}_n)_n$ *is a Levy family if and only if* $\mathrm{LevyRad}_\kappa(\mathcal{X}_n) \to 0$ *for all* $\kappa \in (0,1)$.

Remark 10.14. The moral of this section is that Levy families can be detected by checking that 1-Lipschitz functions are almost constant.

10.6. **Relating Levy Radius and Concentration Function.**

Lemma 10.15. *For all* $\epsilon > 0$, $\kappa \in (0, 1/2)$,

(1) $\mathrm{LevyRad}_{2\alpha_X(\epsilon)}(\mathcal{X}) \leqslant \epsilon$

(2) $\alpha_X\big(\mathrm{LevyRad}_\kappa(X)\big) \leqslant \kappa$

Proof. (Exercise) □

We can use this lemma to provide a proof for part (1) of Theorem 11.3 from the start of lecture.

Proof. Our goal is to show $\mathrm{ObsDiam}_\kappa(\mathcal{X}) \leqslant 2\alpha_X^{-1}(\kappa/2)$. Fix $\epsilon > 0$. By lemmas 11.12 and 11.15 respectively,

$$\mathrm{ObsDiam}_{2\alpha_X(\epsilon)}(\mathcal{X}) \leqslant 2\,\mathrm{LevyRad}_{2\alpha_X(\epsilon)}(\mathcal{X}) \leqslant 2\epsilon$$

Now fix $\kappa \in (0, 1/2)$, and so for all $\epsilon > 0$ with $2\alpha_X(\epsilon) \leqslant \kappa$, we have:

$$\mathrm{ObsDiam}_\kappa(\mathcal{X}) \leqslant \mathrm{ObsDiam}_{2\alpha_X(\epsilon)}(\mathcal{X}) \leqslant 2\epsilon.$$

Taking the infimum over these ϵ we conclude:

$$\mathrm{ObsDiam}_\kappa(\mathcal{X}) \leqslant 2\inf\{\epsilon > 0 | \alpha_X(\epsilon) \leqslant \kappa/2\} = 2\alpha_X^{-1}(\kappa/2).$$

□

10.7. **Observable Distance on mm-spaces.** We want a metric on $\mathcal{M}^w$ to detect Levy families. That is, we want a metric d such that $(\mathcal{X}_n)_n$ is a Levy family if and only if $d(\mathcal{X}_n, *) \to 0$.

Remark 10.16. Observe that the Gromov-Wasserstein distance $d_{GW,1}$ does not detect Levy families. Recall that $\mathbb{S}^n$ is a Levy family and that

$$d_{GW,1}(\mathcal{X}, *) = \frac{1}{2}\mathrm{diam}_1(\mathcal{X}) = \int\int d_X \mu_X \otimes \mu_X.$$

Since $\mathrm{diam}_1(\mathbb{S}^n) = \frac{\pi}{2}$ for all n, this shows that $d_{GW,1}(\mathbb{S}^n) = \frac{\pi}{4} \nrightarrow 0$ even though it is a Levy family.

Remark 10.17. Let $I = [0, 1)$ with 1-dimensional Lebesgue measure $\mathcal{L}^1$. It is a classical fact that given $\mathcal{X} \in \mathcal{M}^w$, there exists $\varphi_X : I \to X$ such that $\varphi_{X\#}\mathcal{L}^1 = \mu_X$. Any map like that is called a *parametrization* of $\mathcal{X}$.

Also given $\varphi_X : I \to X$, we can define the pullback map $\varphi_X^* : \mathrm{Lip}_1(X, \mathbb{R}) \to \mathcal{F}(I, \mathbb{R})$ by $\varphi_X^* f = f \circ \varphi_X$.

Let $\mathcal{F}(I, \mathbb{R})$ denote the set of all measurable functions $f : I \to \mathbb{R}$.

Definition 10.18. The Ky Fan Metric on $\mathcal{F}(I, \mathbb{R})$ is

$$d_{\mathrm{KF}}(f, g) = \inf\{\rho > 0 | \mu_X(\{x \in X : |f(x) - g(x)| \geqslant \rho\}) \leqslant \rho\}.$$

Definition 10.19. Let d_H denote Hausdorff distance. Define the observable distance , denoted d_{conc} on $\mathcal{M}^w$ by:

$$d_{conc}(\mathcal{X}, \mathcal{Y}) = \inf_{\varphi_X, \varphi_Y} d_H^{\mathrm{KF}}\Big(\varphi_X^* \mathrm{Lip}_1(X, \mathbb{R}), \varphi_Y^* \mathrm{Lip}_1(Y, \mathbb{R}\Big)$$

where we view $\varphi_X^* \mathrm{Lip}_1(X, \mathbb{R})$ and $\varphi_Y^* \mathrm{Lip}_1(Y, \mathbb{R})$ as subsets of $(\mathcal{F}(I, \mathbb{R}), d_{\mathrm{KF}})$.

Proposition 10.20. *$d_{conc}(\mathcal{X}, *)$ and $\mathrm{ObsDiam}_\kappa(\mathcal{X})$ are within a factor of 2 of eachother.*

Theorem 10.21. *$(\mathcal{X}_n)_n$ is Levy if and only if $d_{conc}(\mathcal{X}_n, *) \to 0$.*

11. Lecture 11. February 20. Sunhyuk Lim

11.1. **Urysohn Universal Space.** Presentation about Urysohn Universal Space by Samir Chowdhury.

Fréchet-defined metric space∼1905. Urysohn∼1924, Hausdorff∼1924 and Katětov∼1986 developed the following theory.

Definition 11.1 (1 point extension)**.** A metric space (X, d_X) is given. Then, a metric space (Y, d_Y) is said to be one point extension of X if $Y = X \sqcup \{y\}$ and $d_Y|_{X \times X} = d_X$.

Definition 11.2 (1 point extension property [1EP])**.** We say a metric space (U, d_U) has one point extension property if $\forall$ finite subset X of U and any one point extension (Y, d_Y) of X such that $Y = X \sqcup \{y\}$, $\exists u \in U$ such that $X \cup \{u\}$ is isometric to Y.

Theorem 11.3 (Ultrahomogeneity)**.** *Let X, Y be serabale and complete metric spaces with 1EP. A is a finite subset of X, B is a finite subset of Y, and $\phi : A \to B$ is an isometry between A and B. Then, $\exists$ isometry $\Phi : X \to Y$ extending ϕ.*

Proof. Since both X and Y are separable, there are countable dense subset $S_X \subseteq X$ and countable dense subset $S_Y \subseteq Y$. Let

$$S_X = \{x_1, x_2, \dots\}$$

and

$$S_Y = \{y_1, y_2, \dots\}.$$

$\phi : A \to B$ is given. Observe that since $A \cup \{x_1\}$ is still a finite metric space, because of the 1EP of Y, there exist $v_1 \in Y$ and isometry $f_1 : A \cup \{x_1\} \to B \cup \{v_1\}$. Also, since $B \cup \{v_1\} \cup \{y_1\}$ is a finite metric space, because of the 1EP of X, there exist $u_1 \in X$ and isometry $\Phi_1 : A \cup \{x_1\} \cup \{u_1\} \to B \cup \{v_1\} \cup \{y_1\}$. Repeat this process inductively. Then, we have sequence of isometries Φ_n for each $n \in \mathbb{Z}_{>0}$ where each Φ_n extends Φ_{n-1}. Finally, define

$$\Phi := \bigcup_n \Phi_n : A \cup S_X \to Y.$$

Obseve that $S_Y \subseteq \mathrm{Im}(\Phi)$. Then, one can extend the domain of Φ by using the completeness. □

Corollary 11.4. *Suppose X and Y are separable, complete, and 1EP. Then X and Y are isometric.*

Proof. Take $A = B = \varnothing$ in the previous proof. □

Definition 11.5 (Urysohn universal space)**.** A metric space (U, d_U) is said to be *Urysohn universal* if it is separable, complete and has one point extension property.

Remark 11.6. Any separable metric space can be isometrically embedded in Urysohn universal space (U, d_U) (take a countable dense subset $S = \{s_1, s_2, \dots\}$ and use 1EP+completeness).

Theorem 11.7. *A Urysohn universal space (U, d_U) is a geodesic space. It means that, for any two points $a, b \in U$, there exist continuous $\gamma : I = [l, r] \to U$ such that $\gamma(r) = a, \gamma(l) = b$, and $d_U(\gamma(t), \gamma(s)) = |t - s|$ for any $t, s \in I$.*

Proof. Consider closed interval $I := [0, d_U(a, b)] \subset \mathbb{R}$. Since I is separable, there exists isometric embedding $\phi : I \hookrightarrow U$. Let $l = \phi(0)$ and $r = \phi(d_U(a, b))$. Let $f : \{l, r\} \to \{a, b\}$ be the map such that $f(l) = a$ and $f(r) = b$. This f is isometry so it can be extended to global isometry $\Phi : U \to U$. Then, $\Phi(\phi(I))$ is geodesic from a to b. □

Remark 11.8. Do we have an example of complete, separable but not geodesic space (question by Prof. Matt Kahle)? Yes, sphere with euclidean metric (answer by Osman).

Remark 11.9. In fact, Urysohn universal metric space (U, d_U) has uncountably many geodesics between any two points (& branching).

Now, we will construct Urysohn universal space.

Definition 11.10 (Katětov function)**.** Let (X, d_X) be a metric space. A function $f : X \to \mathbb{R}$ is said to be *Katětov function* if it satisfies the following inequality:

$$f(x) - f(y) \leqslant d_X(x, y) \leqslant f(x) + f(y)$$

for any $x, y \in X$.

Remark 11.11. In the above definition, observe that the first inequality actually implies $|f(x) - f(y)| \leqslant d_X(x, y)$ so that f is 1-Lipschitz. Also, the second inequality means f is nonnegative.

We denote $\widetilde{E}(X) := \{\text{All Katětov functions on } X\}$.

Remark 11.12.

$$(\text{one point extension of } X) \Longleftrightarrow \widetilde{E}(X)_+$$

Let $Y = X \cup \{y\}$ be one point extension of X. Let $f : X \to \mathbb{R}$ such that $f(x) = d_Y(x, y)$. Verify $f \in \widetilde{E}(X)_+$. Conversely, given $f \in \widetilde{E}(X)_+$, Write $Y = X \cup \{y\}$ and

$$d_Y(a, b) := \begin{cases} d_X(a, b) \ \textit{if } a, b \in X \\ 0 \ \textit{if } a = y = b \\ f(a) \ \textit{if } a \in X, b = y \end{cases}$$

In other words, Katětov functions encode distance to a abstract point.

One can give metric structure on $\widetilde{E}(X)$ in the following way: $d_{\widetilde{E}(X)}(f, g) := \sup_{x \in X} |f(x) - g(x)|$.

Definition 11.13 (Katětov extension)**.** (Y, d_Y) is a metric space. X is a finite subset of Y. $f \in \widetilde{E}(X)$. We define

$$\begin{aligned} k_f : &Y \longrightarrow \mathbb{R} \\ &y \longmapsto \inf_{x \in X}(d_Y(x,y) + f(x)) \end{aligned}$$

This k_f is the Katětov extension of f.

Remark 11.14. Verify the following properties.

(1) $k_f = f$ on X.

(2) $k_f \in \widetilde{E}(Y)$.

Definition 11.15. We say $f \in \widetilde{E}(Y)$ is supported on $X \subseteq Y$ if we have $f(y) = \inf_{x \in X}(d_Y(x,y) + f(x))\ \forall y \in Y$.

We denote

$$E(Y) := \{f \in \widetilde{E}(Y) : f \text{ is finitely supported}\}.$$

Remark 11.16. (1) One can isometrically embed Y in $E(Y)$ by using the Kuratowski embedding in the follwing way:

$$\begin{aligned} Y &\hookrightarrow E(Y) \\ y &\mapsto f_y \end{aligned}$$

where $f_y : Y \to \mathbb{R}$ is the map satisfying $f(z) = d_Y(z,y)$. the support of f_y is $\{y\}$.

(2) Let X be a finite subset of Y and $f \in E(X) = \widetilde{E}(X)$. Then, $k_f \in E(Y)$.

(3) $d_{E(Y)}(f, f_y) = f(y)$. Here is the proof.

$$|f(z) - f_y(z)| = |f(z) - d_Y(z,y))| \leqslant f(y)$$

for arbitrary $z, y \in Y$ so that we have $\sup_{z \in Y} |f(z) - f_y(z)| \leqslant f(y)$. Also,

$$f(y) = |f(y) - f_y(y)| \leqslant \sup_{z \in Y} |f(z) - f_y(z)|.$$

(4) Let X be a finite subset of Y and $f \in E(X)$. Then,

$$d_{E(Y)}(k_f, f_x) = k_f(x) = f(x).$$

(5) Let X be a finite subset of Y. Any one point extension of X embeds isometrically in $E(X)$ and hence in $E(Y)$.

Proposition 11.17. *If Y is separable, then $E(Y)$ is separable ($\widetilde{E}(Y)$ may not be).*

Proof. Start by showing that for finite subset $X \subseteq Y$, $E(X)$ is separable. Let $f \in E(X)$, then one can view $f = \sum_{x \in X} c_x 1_x$. Take functions which assume rational values for each c_x, then we havbe countable dense subset of $E(X)$.

Now, let $E_n(Y) := \{f \in E(Y) : |\text{supp}(f)| \leqslant n\}$ for each $n \in \mathbb{Z}_{>0}$. Then $E(Y) = \bigcup_n E_n(Y)$, so it is also separable. ∎

Finally, we will construct Urysohn universal space in the following way:

Let Y be an arbitrary separable metric space. Define $Y_0 := Y$, $Y_1 := E(Y)$, $Y_2 := E(Y_1), \ldots,$ and $Y_\infty := \bigcup_n Y_n$. Take the metric completion $\overline{Y_\infty}$ of Y_∞. Then, $\overline{Y_\infty}$ is separable by the previous proposition. So, to show that $\overline{Y_\infty}$ is the Urysohn universal space containing Y, it is enough to show $\overline{Y_\infty}$ has 1EP.

Take d to be the supremum metric on $\overline{Y_\infty}$. For any $f, g \in \overline{Y_\infty}$, take n large enough so that $f, g \in Y_n$ and $d(f,g) = d_{Y_n}(f,g)$.

Let $X = \{x_1, x_2, \ldots, x_n\} \subseteq \overline{Y_\infty}$ and $Z = X \cup \{z\}$ such that $d_Z|_{X\times X} = d|_{X\times X}$. Define

$$\begin{aligned} f :& X \to \mathbb{R} \\ & x \mapsto d_Z(x,z) \end{aligned}$$

$f \in E(X)$. Take katětov extension $k_f : \overline{Y_\infty} \to \mathbb{R}$. Fix arbitrary $\varepsilon > 0$. Pick $y_1, \ldots, y_n \in Y_{m(\varepsilon)}$ such that $d_(x_i, y_i) < \varepsilon$ for each i. Define

$$\begin{aligned} f_\varepsilon :& \{y_1, \ldots, y_n\} \longrightarrow \mathbb{R} \\ & \qquad y_i \longmapsto k_f(y_i) \end{aligned}$$

Extend to get $k_{f_\varepsilon} : E(Y_{m(\varepsilon)}) \to \mathbb{R}$. So $k_{f_\varepsilon} \in Y_{m(\varepsilon)+1}$. We want to show $\{k_{f_\varepsilon}\}_\varepsilon$ has Cauchy property.

Let $\varepsilon, \eta > 0$. Then we have $f_\varepsilon : \{y_1, \ldots, y_n\} \longrightarrow \mathbb{R}$ with $d(x_i, y_i) < \varepsilon$ and $f_\eta : \{z_1, \ldots, z_n\} \longrightarrow \mathbb{R}$ with $d(x_i, z_i) < \eta$. Let $M := \max(m(\varepsilon), m(\eta))$. Then, $A \cup B \subset Y_M$. Now we want to bound $|k_{f_\varepsilon} - k_{f_\eta}|$ for all $y \in Y_M$. We have,

$$k_{f_\varepsilon}(y) = \inf_{a\in A}(d(a,y) + f_\varepsilon(a))$$

and

$$k_{f_\eta}(y) = \inf_{b\in B}(d(b,y) + f_\eta(b)).$$

Then.

$$|k_{f_\varepsilon}(y) - k_{f_\eta}(y)| = |\inf_{a\in A}(d(a,y) + f_\varepsilon(a)) - \inf_{b\in B}(d(b,y) + f_\eta(b))|$$

Let $\alpha = \inf_{a\in A}(d(a,y) + f_\varepsilon(a))$ and $\beta = \inf_{b\in B}(d(b,y) + f_\eta(b))$. without loss of generality, one can assume $\alpha \geqslant \beta$. Let i be such that $d(z_i, y) + k_f(z_i) = k_{f_\varepsilon}(y)$. Then,

$$\begin{aligned} |k_{f_\varepsilon}(y) - k_{f_\eta}(y)| &\leqslant |d(y_i, y) + k_f(y_i) - d(z_i, y) - k_f(z_i)| \\ &\leqslant |d(y_i, y) - d(z_i, y)| + |k_f(y_i) - k_f(z_i)| \\ &\leqslant 2(\varepsilon + \eta) \end{aligned}$$

since $vertd(y_i, y) - d(z_i, y)| \leqslant d(y_i, z_i)$ by the triangle inequality and $|k_f(y_i) - k_f(z_i)| \leqslant d(y_i, z_i)$ from the definition of Katětov.

Thus, $\{k_{f_\varepsilon}\}_\varepsilon$ is cauchy. hence the limt exists in $\overline{Y_\infty}$. Call it $\overline{y}$.

$$|d(k_{f_\varepsilon}.f_{x_i}) - d(k_{f_\varepsilon}.f_{y_i})| \leqslant |k_{f_\varepsilon}(x_i) - k_{f_\varepsilon}(y_i)| \leqslant d(x_i, y_i) < \varepsilon.$$

So, as ε goes to 0,

$$\begin{aligned}
\lim_{\varepsilon\to 0}(k_{f_\varepsilon}.f_{x_i}) &= \lim_{\varepsilon\to 0}(k_{f_\varepsilon}.f_{y_i}) \\
&= \lim_{\varepsilon\to 0} k_{f_\varepsilon}(y_i) \\
&= \lim_{\varepsilon\to 0} f_\varepsilon(y_i) \\
&= \lim_{\varepsilon\to 0} k_f(y_i^{(\varepsilon)}) = k_f(x_i) = f(x_i).
\end{aligned}$$

Thus $d(\overline{y}, f_{x_i}) = f(x_i) = d_Z(z, x_i)$.

Addendum from Samir Chowdhury:

An application:

Sunhyuk asked after class if I knew any applications of the Urysohn universal space. Here is one that I know.

We've been talking about the collection of compact metric spaces equipped with with dGH. Question: is the collection of all compact metric spaces a set? One proof providing a positive answer to this question is obtained via the Urysohn space U: simply embed every compact metric space into U. Then the collection of all compact metric spaces is a subset of the metric space U, and is therefore a set.

Non-unique/branching geodesics:

Here is a quick proof (taken from a Melleray paper) of the fact that there are uncountably many geodesics between any two points in U:

Let $a \neq b \in U$, and let $\ell = d_U(a, b)$. Let γ be a geodesic in U from a to b, and let m be its midpoint. Then γ has length ℓ.

Let $\epsilon > 0$.

Next define a map $f_\epsilon : \{a, b, m\} \to$ as follows: $a \mapsto \ell/2$, $b \mapsto \ell/2$, and $m \mapsto \epsilon$.

Then f_ϵ is Katetov. We have:

$$|f_\epsilon(a) - f_\epsilon(b)| = |\ell/2 - \ell/2| = 0 \leqslant d_U(a, b) = \ell \leqslant f_\epsilon(a) + f_\epsilon(b),$$

and also

$$|f_\epsilon(a) - f_\epsilon(m)| = |\ell/2 - \epsilon| < d_U(a, m) < \ell/2 + \epsilon = f_\epsilon(a) + f_\epsilon(m).$$

Then f corresponds to a 1-point metric extension, so there exists $z \in U$ such that $d_U(a, z) = \ell/2$, $d_U(b, z) = \ell/2$, and $d_U(z, m) = \epsilon$. But then z is on a geodesic from a to b that is different from γ.

Bonus: this example can be adapted to produce branching geodesics in U. Even worse, these geodesics may be made to branch uncountably often.

12. Lecture 12. Februray 25. Mario Gomez

Definition 12.1 (Urysohn Universal Space)**.** A metric space (U, d_U) is called Urysohn Universal if it is Polish (that is, complete and separable) and satisifies:

(1) (Universality) Every separable m.s. X has an isometric embedding $X \hookrightarrow U$.

(2) (Homogeneity) Every isometry between finite sets of points extends to an isometry of the whole space.

Remark 12.2.

(1) We have previously shown that the Urysohn Universal space exists and is unique up to isometry.

(2) We need both universality and homogeneity in order to get the Urysohn space. For example, $C[0,1]$ with the supremum norm satisfies universality, but not homogeneity.

Today's goal: "Sufficiently random" finite metric spaces converge to the Urysohn Universal space with probability 1. We will define what we mean by sufficiently random and under which probability measure.

12.1. **Random Distance Matrices.**

Definition 12.3. Let

$$\mathcal{R} = \{(r_{ij})_{i,j=1}^{\infty} : r_{ii} = 0, r_{ij} \geqslant 0, r_{ij} = r_{ji}, r_{ik} + r_{kj} \geqslant r_{ij}, \forall i, j, k\}.$$

Elements of $\mathcal{R}$ are called distance matrices (dm for short). $r \in \mathcal{R}$ is proper if there are no 0's off the main diagonal.

Remark 12.4.

(1) Every dm determines a semimetric on $\mathbb{N}$. The conditions in the definition of $\mathcal{R}$ are reflexivity, nonnegativity, symmetry, and the triangle inequality, respectively. Additionally, a proper dm determines a metric.

$$\begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & \cdots \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ \vdots \end{array} & \begin{pmatrix} 0 & r_{12} & r_{13} & \cdots \\ & 0 & r_{23} & \\ & & 0 & \\ & & & \ddots \end{pmatrix} \end{array}$$

For example, r_{12} is the distance between 1 and 2.

(2) $\mathcal{R}$ is a convex cone in the vector space of infinite real matrices, so we will call it the cone of dm's.

Definition 12.5. $\mathcal{R}_n$ =dm's of order n.

As above, every $r \in \mathcal{R}_n$ determines a (semi)metric on the space X_r consisting of n points.

Let's define a function $p_{m,n} : M_m^S \to M_n^S$, where M_k^S is the set of symmetric matrices of order k. Given $r \in M_m^S$, we write $p_{m,n}(r)$ for the northwest corner of r of order n.

$$\begin{array}{cc} \begin{array}{ccccc} 1 & \cdots & n & \cdots & m \end{array} & \\ r = \left(\begin{array}{c|c} p_{m,n}(r) & \\ \hline & \end{array}\right) & \begin{array}{c} 1 \\ \vdots \\ n \\ \vdots \\ m \end{array} \end{array}$$

Notice that $p_{m,n}(\mathcal{R}_m) = \mathcal{R}_n$. We define analogously $p_n : M_{\mathbb{N}}^S \to M_n^S$, where $p_n(\mathcal{R}) = \mathcal{R}_n$.

Remark 12.6. The cones $\mathcal{R}_n$ are invariant under conjugation by elements of S_n (that is, when rows and columns are permuted simultaneously).

Example 12.7. $\mathcal{R}_1 = \{0\}$, and $\mathcal{R}_2 = \{\left(\begin{smallmatrix} 0 & r \\ r & 0 \end{smallmatrix}\right) : r \geqslant 0\} \cong \mathbb{R}_{\geqslant 0}$.

Definition 12.8 (Admissible Vectors)**.** Let $r = (r_{ij})_{i,j=1}^n \in \mathcal{R}$. A vector $\vec{a} = (a_i)_1^n \in \mathbb{R}^n$ is <u>admissible</u> if the matrix obtained by attaching $\vec{a}$ to r as the last row and column is a distance matrix of order $n+1$.

Notation 12.9.

- $A(r)$ = set of admissible vectors for a fixed dm r.
- $r^{\vec{a}}$ is the matrix obtained by attaching $\vec{a}$ as described.

$$r^{\vec{a}} = \left(\begin{array}{cccc|c} 0 & r_{12} & \cdots & r_{1n} & a_1 \\ r_{21} & 0 & & & a_2 \\ \vdots & & \ddots & \vdots & \vdots \\ r_{n1} & & \cdots & 0 & a_n \\ \hline a_1 & a_2 & \cdots & a_n & 0 \end{array}\right)$$

Remark 12.10.

(1) The projection $p_{n+1,n}$ recovers the original matrix: $p_{n+1,n}(r^{\vec{a}}) = r$.

(2) The (semi)metric space $X_{r^{\vec{a}}}$ is a one point extension of X_r.

(3) The admissibility of $\vec{a} \in \mathbb{R}^n$, for a fixed $r \in \mathcal{R}_n$, is equivalent to the system
$$|a_i - a_j| \leqslant r_{ij} \leqslant a_i + a_j, \tag{8}$$
for all $i, j = 1, \ldots, n$. Thus, $A(r) = \{(a_i)_1^n \in \mathbb{R}^n \text{ that satisfy (8)}\}$.
Notice that the system above is very similar to the Katětov condition. Indeed, if we write $X_r = \{1, \ldots, n\}$, its metric as $d_{X_r}(i,j) = r_{ij}$, and define $f(i) = a_i$, then f is a Katětov function as per Definition 11.10.

Definition 12.11 (Projections)**.** Given $r \in \mathcal{R}_n$ and $n < N$, we define the projection
$$\begin{aligned} \chi_n^r : A(r) &\twoheadrightarrow A(p_n(r)) \\ (b_i)_{i=1}^N &\longmapsto (b_1, \ldots, b_n). \end{aligned}$$

Lemma 12.12 (Ammalgamation Lemma). *Let $r \in \mathcal{R}_n$. For any two $\vec{a}, \vec{b} \in A(r)$, there exists $h \in \mathbb{R}_{\geqslant 0}$ such that $\vec{b}' = (\underbrace{b_1, \ldots, b_n}_{\vec{b}}, h) \in A(r^{\vec{a}})$.*

$$\left(\begin{array}{cccc|cc} 0 & r_{12} & \cdots & r_{1n} & a_1 & b_1 \\ r_{21} & 0 & & & a_2 & b_2 \\ \vdots & & \ddots & \vdots & \vdots & \vdots \\ r_{n1} & & \cdots & 0 & a_n & b_n \\ \hline a_1 & a_2 & \cdots & a_n & 0 & h \\ b_1 & b_2 & \cdots & b_n & h & 0 \end{array}\right)$$

In other words, it doesn't matter in what order we attach $\vec{a}, \vec{b}$ to r. We can always expand $\vec{b}$ so that $\vec{b}'$ is still admissible by $r^{\vec{a}}$.

Proof. Consider two finite metric spaces $X = (\{x_1, \ldots, x_n\}, \rho_1)$, $Y = (\{y_1, \ldots, y_n\}, \rho_2)$, and assume that the subspaces $\{x_1, \ldots, x_{n-1}\}$ and $\{y_1, \ldots, y_{n-1}\}$ are isometric (that is, $\rho_1(x_i, x_j) = \rho_2(y_i, y_j)$ for $i, j = 1, \ldots, n-1$). We claim that we can find a metric space $Z = (\{z_1, \ldots, z_{n-1}, z_n, z_{n+1}\}, \rho)$ and isometries $I_1 : X \to Z$, $I_2 : Y \to Z$ such that $I_1(x_i) = I_2(y_i) = z_i$ for $i = 1, \ldots, n-1$, $I_1(x_n) = z_n$ and $I_2(y_n) = z_{n+1}$. Z can be any set with $(n+1)$, so the real problem is choosing an appropriate metric. Notice, though, that ρ on $\{z_1, \ldots, z_{n-1}\}$ is already given by ρ_1 and ρ_2. Moreover, $\rho(z_i, z_n) = \rho_1(x_i, x_n)$ and $\rho(z_i, z_{n+1}) = \rho_2(y_i, y_n)$ because I_1, I_2 are isometries. Thus, to build ρ we only have to find a suitable $h = \rho(z_n, z_{n+1}) \geqslant 0$. For that purpose, we have the following inequality for all $i, j = 1, \ldots, n-1$:

$$\begin{aligned} \rho_1(x_i, x_n) - \rho_2(y_i, y_n) &\leqslant \rho_1(x_i, x_j) + \rho_1(x_j, x_n) - \rho_2(y_i, y_n) \\ &= \rho_1(x_j, x_n) + \rho_2(y_i, y_j) - \rho_2(y_i, y_n) \\ &\leqslant \rho_1(x_j, x_n) + \rho_2(y_j, y_n). \end{aligned}$$

The two inequalities above are a consequence of the triangle inequality on ρ_1 and ρ_2, respectively. Then:

$$M = \max_i |\rho_1(x_i, x_n) - \rho_2(y_i, y_n)| \leqslant \min_j |\rho_1(x_j, x_n) + \rho_2(y_j, y_n)| = m,$$

so that we can choose an arbitrary $h \in [M, m]$ and set $\rho(z_n, z_{n+1}) = h$. Rephrasing the above inequality in terms of elements of Z shows that ρ satisfies the triangle inequality and is, indeed, a metric. This establishes the claim.
Now, assume $r \in \mathcal{R}_{n-1}, \vec{a}, \vec{b} \in A(r)$ and write $X_{r^{\vec{a}}} = \{x_1, \ldots, x_n\}$ and $Y_{r^{\vec{b}}} = \{y_1, \ldots, y_n\}$. The claim gives a third space Z in which $h = \rho(z_n, z_{n+1}) \geqslant 0$ is the number required by the lemma. ∎

Lemma 12.13. *For $n < N$ and $r \in \mathcal{R}_N$, χ_n^r is an epimorphism $A(r) \twoheadrightarrow A(p_n(r))$. In other words, for all $\vec{a} \in A(p_n(r))$, there exists $(b_{n+1}, \ldots, b_N)$ such that $\vec{b}' = (a_1, \ldots, a_n, b_{n+1}, \ldots, b_N) \in A(r)$.*

Proof. We can represent $r \in \mathcal{R}_N$ as a sequence of admissible vectors $r(k)$ of increasing lengths k, where $r(k) \in A(p_k(r))$:

$$\begin{array}{c} \hphantom{0 \quad r_{12} \quad \cdots \quad} r(k) \hphantom{\quad \cdots} \\ \left(\begin{array}{ccc|c|c} 0 & r_{12} & \cdots & r_{1,k+1} & \cdots \\ r_{21} & 0 & & & \\ \vdots & & \ddots & \vdots & \\ r_{k1} & & & r_{k,k+1} & \\ \cline{4-4} & & & 0 & \\ & & & \vdots & \end{array}\right)\end{array}$$

The conclusion follows by repeatedly applying the amalgamation lemma (12.12) to $\vec{a}$ and $r(n), r(n+1), \ldots, r(N)$ (every $r(k)$ is admissible for the previous distance matrix $p_{k-1}(r)$). $\square$

Now, to consider probability we take as input an arbitrary probability measure γ on $\mathbb{R}_{\geqslant 0}$. The idea of the method is to construct arbitary measures on matrices. For that we determine the distribution of the first row and successively determine the conditional measures of the entries. More explicitly, we start with the 1 point metric space $\{1\}$. To add new points $n = 2, 3, \ldots$, we set the distance r_n from n to 1 to be a random variable $\mathcal{L}_n$ i.i.d. with distribution γ.

12.2. **Universal Distance Matrices.**

Definition 12.14.

(1) $r \in \mathcal{R}$ is a <u>universal</u> distance matrix if for all $\epsilon > 0$, $n \in \mathbb{N}$, and $\vec{a} \in A(p_n(r))$, there exists $m \in \mathbb{N}$ such that $\max_{i=1,\ldots,n} |r_{i_m} - a_i| < \epsilon$. That is, for every $n \in \mathbb{N}$, $\{(r_{ij})_{i=1}^n\}_{j=n+1}^\infty$ is everywhere dense in $A(p_n(r))$.

$$\left(\begin{array}{ccc|cccc} 0 & r_{12} & \cdots & a_1 & \xleftrightarrow{\text{Distance}<\epsilon} & r_{1m} & \cdots \\ r_{21} & 0 & & & & & \\ \vdots & & \ddots & \vdots & \vdots & \vdots & \\ r_{n1} & & & a_n & \xleftrightarrow{\text{Distance}<\epsilon} & r_{nm} & \cdots \\ \hline \vdots & & & 0 & \cdots & & \\ & & & \vdots & & & \end{array}\right)$$

(2) $r \in \mathcal{R}$ is <u>almost universal</u> if for all $n \in \mathbb{N}$, the set $\{(r_{i_k,i_s})_{k,s=1}^n\}$ of order n submatrices of r if dense in $\mathcal{R}_n$.

Lemma 12.15. *Universal distance matrices are Almost universal, but not conversely.*

Theorem 12.16.

(1) The completion (U_r, ρ_r) of $(\mathbb{N}, r)$ with metric determined by a universal proper distance matrix r is the Urysohn Universal Space.

(2) *For any 2 universal proper distance matrices r, r', the completions of $(\mathbb{N}, r)$ and $(\mathbb{N}, r')$ are isometric.*

12.3. Universality of Almost All Distance Matrices.

Theorem 12.17. *The measures ν_λ (constructed inductively in paper) are concentrated on the set of universal matrices. That is, almost every distance matrix (with respect to the measures ν_λ) is universal.*

This is the precise formulation of the statement at the start of the lecture: almost every random metric space converges to the Urysohn universal space with probability 1. Lastly we have:

Theorem 12.18. *The completion of the random countable metric space $\mathbb{N}$ is the Urysohn space with probability 1.*

13. Lecture 13. February 27. Scribe Francisco Martinez

13.1. The Rado Graph.

Definition 13.1 (Rado Graph)**.** The Rado Graph, denoted in this lecture as $\mathfrak{R}$, is the countably infinite graph with vertex set $\mathbb{Z}_+$ and edges $x \sim y$ if and only if, for $x < y$, the x^{th} binary digit of y is 1 (reading from right to left).

Example 13.2. We can characterize the set $\{y : x \sim y, x < y\}$ in $\mathfrak{R}$ using congruences modulo 2^x, for instance:

- $1 \sim y$ if and only if y is odd.
- $2 \sim y$ if and only if $y \equiv 2 \text{ or } 3 \mod 4$.
- $3 \sim y$ if and only if $y \equiv 4, 5, 6, \text{ or } 7 \mod 8$

Note 13.3. When consider only the first n vertices, the first few numbers are connected to roughly half of the vertices.

Properties of the Rado Graph.

(1) **Universality:** If H is a finite (or countable) graph, then there exists an embedding $H \hookrightarrow \mathfrak{R}$ as an induced subgraph.
(2) **Homogeneity:** If $f : H \to H'$ is an isomorphism between any finite induced subgraphs of $\mathfrak{R}$, then f extends to an automorphism of $\mathfrak{R}$.
(3) **One point extension:** Let U, V be finite graphs such that U is an induced subgraph of V and V has precisely one more vertex than U, i. e. vertices(V) = vertices$(U) \cup \{v\}$. If $f : U \hookrightarrow \mathfrak{R}$ is an embedding as an induced subgraph, then f extends to $\tilde{f} : V \hookrightarrow \mathfrak{R}$ as an induced subgraph.
(4) If U, V are disjoint finite vertex subsets of $\mathfrak{R}$, then there exists a vertex x such that x is adjacent to every vertex of U and is adjacent to **no** vertex of v.

Note 13.4.

- Property (4) implies (3): For U and V as described in (3), partition the vertices of U into $A = N_V(v)$ and $B = \text{vertices}(U)\backslash A$. Then define $\widetilde{f}(v)$ as the vertex x that is adjacent to all $f(A)$ and not to any vertex in $f(B)$.
- Property (3) implies (2): the extension can be built using a back and forth argument, similar to the proof of Ultrahomogeneity of the Urysohn Universal Space given in lecture 11 (Theorem 13.3).

Other constructions/definitions of $\mathfrak{R}$.

Inductive Definition of $\mathfrak{R}$

Definition 13.5 (Inductive definition of $\mathfrak{R}$)**.**

- Stage 1: $\mathfrak{R}_1$ has 1 vertex.
- Stage 2: To get $\mathfrak{R}_2$, add two independent vertices to $\mathfrak{R}_0$, connect one of them to the old vertex, and leave the other one independent.

  $\vdots$
- Sage n+1: To get $\mathfrak{R}_{n+1}$, add 2^{a_n} independent vertices to $\mathfrak{R}_n$, where a_n is the number of vertices in $\mathfrak{R}_n$. Connect each new vertex with old vertices, such that the neighborhood sets realize all possible subsets of vertices in $\mathfrak{R}_n$.

Finally $\mathfrak{R} = \bigcup_{i=0}^{\infty} \mathfrak{R}_i$.

Note 13.6.

- Property (4) follows easily from this definition: if $U, V \subset \mathfrak{R}_n$, then one of the vertices added in the next stage satisfies the property.
- As above, if a_n is the number of vertices in $\mathfrak{R}_n$, this defines an integer sequence: $a_1 = 1$ and $a_{n+1} = a_n + 2^{a_n}$ for $n \geqslant 1$. More on this sequence at The online encyclopedia of integer sequences.

Random Construction of $\mathfrak{R}$

Similar to the finite Erdős-Rényi graphs $G(n,p)$ (see Lecture 4), they also studied a random countably infinite graph. This graph has as vertex set $\mathbb{N}$, and each edge is included independently at random with probability $p = 1/2$. The following result states its relationtip with the Rado graph.

Theorem 13.7. *Almost surely, the countably infinite random graph is isomorphic to* $\mathfrak{R}$.

Note 13.8.

- In this construction, the properties of Universality and One point extension follow easily:
  - Universality: Let H be a finite graph with n vertices. The probability that H is not isomorphic to the subgraph induced by $\{1, 2, \dots, n\}$ in the Countably Infinite Random Graph is some number $q < 1$. The probability it is not the subgraph induced by $\{n+1, n+2, \dots, 2n\}$ is also q and is independent from the first. So the probability it is not isomorphic to either one reduces to q^2.

Continuing this process, we see it must be isomorphic to some subgraph $\{am+1, am+2, \ldots, (a+1)m\}$ for some $m \in \mathbb{N}$ with probability 1.

– One point extension: follows from property (4). Given $U, V \subset \mathbb{N}$ disjoint, for any point x outside U and V, the probability it is adjacent to all vertices in U and none in V is some fixed positive number. Since this value is the same for all x and independent, the probability that at least one point satisfies the property must be 1.

- Almost surely, for any choice of the edge probability $0 < p < 1$ we get the same graph (up to isomorphisms).

Metric Geometry Point of View.
Recall we can turn $\mathfrak{R}$ intro a metric space by giving it the shortest path metric. Note that the diameter of the Rado graph is 2: by property (4) for every pair of non-adjacent vertices, there exists a third vertex adjacent to both. Hence the Rado graph takes only 0, 1 and 2 distances.

Note 13.9. The Rado graph is Universal for metric spaces (finite or countable) that only realize distances 0, 1 and 2.

Exercise 13.10. Show that the clique complex of $\mathfrak{R}$ is contractible.

Note 13.11. The Rado graph has self similarity properties. For example, if we partition the vertices of $\mathfrak{R}$ into two disjoint sets U and V, then there exists an induced subgraph either in U or V, that is isomorphic to $\mathfrak{R}$.

Open Questions/Projects.

(1) Consider the filtration of $\mathfrak{R}$ given by the first definition. Take the clique complex (Vietoris-Rips complex) of this filtration. What is the persistence Homology? Does it have any fractal-like structure?

(2) Given the same order on the vertices, let $f(n)$ be the size of the longest clique within the first n vertices. Similarly, let $g(n)$ be the size of the largest independent set up to vertex n. What is the asymptotic behaviour of $f(n)$ and $g(n)$ as $n \to \infty$?

13.2. **Recap: Observable Distance.**
Recall that the metric d_{conc} defined on $\mathcal{M}^{\mathrm{w}}$ detects Levy families, because of its relation with the observable diameter. We now prove proposition 10.20:

Proposition (10.20)**.** *For every $\mathcal{X} \in \mathcal{M}^{\mathrm{w}}$,*

$$d_{conc}(\mathcal{X}, *) \leqslant \mathrm{ObsDiam}(\mathcal{X}) \leqslant 2\, d_{conc}(\mathcal{X}, *).$$

Where $\mathrm{ObsDiam}(\mathcal{X}) := \inf_{\kappa > 0} \max\{\kappa, \mathrm{ObsDiam}_\kappa(\mathcal{X})\}$

Note 13.12. Recall the definition of d_{conc}:
Given a mm-space $\mathcal{X} = (X, d_X, \mu_X)$, let $\varphi_X : I = [0,1] \to X$ be a parametrization, i.e. a map such that $\varphi_{X\#}\mathcal{L}^1 = \mu_X$. The pullback sends any 1-Lipschitz map $f \in \mathrm{Lip}_1(X, \mathbb{R})$ to a measurable function $\varphi_X^* f = f \circ \varphi_X : (I, \mathcal{L}^1) \to (X, \mu_X)$. Thus $\varphi_X^*(\mathrm{Lip}_1(X, \mathbb{R})) \subset \mathcal{F}(I, \mathbb{R})$,

where $\mathcal{F}(I,\mathbb{R})$ is the set of all measurable functions from I to $\mathbb{R}$, endowed with the Ky Fan metric.

Thus, given two mm-spaces $\mathcal{X},\mathcal{Y}\in\mathcal{M}^{\mathrm{w}}$, we defined

$$d_{conc}(\mathcal{X},\mathcal{Y})=\inf_{\varphi_X,\varphi_Y} d_H^{KF}\left(\varphi_X^*\mathrm{Lip}_1(X,\mathbb{R})),\varphi_Y^*\mathrm{Lip}_1(Y,\mathbb{R}))\right)$$

Where d_H^{Kf} is the Hausdorff Distance induced by the Ky Fan metric.

Note 13.13. Recall the Ky Fan metric:
For $f,g\in\mathcal{F}(I,\mathbb{R})$, and $\varepsilon>0$ let

$$Q(\varepsilon)=\mathcal{L}_1\left(\{t\in I:|f(t)-g(t)|\geqslant\varepsilon\}\right).$$

Then the Ky Fan metric is $d_{KF}=\inf\{\varepsilon>0:Q(\varepsilon)\leqslant\varepsilon\}$.

Proposition 10.20. We just prove the first inequality: $d_{conc}(\mathcal{X},*)\leqslant\mathrm{ObsDiam}(\mathcal{X})$.

Assume $\mathrm{ObsDiam}_\kappa(\mathcal{X})<\varepsilon$, for some $\varepsilon>0$. That means

$$\sup_{f\in\mathrm{Lip}_1(X,\mathbb{R})}\mathrm{PartDiam}_{1-\kappa}\left(f_\#\mu_X\right)<\varepsilon.$$

Fix any $f\in\mathrm{Lip}_1(X,\mathbb{R})$, so $\mathrm{PartDiam}(f_\#\mu_X)<\varepsilon$. By definition, there exists a set $A_\varepsilon\subset\mathbb{R}$ with $\mathrm{diam}(A_\varepsilon)<\varepsilon$ and $(f_\#\mu_X(A_\varepsilon)\geqslant 1-\kappa$. By letting $a=\inf(A_\varepsilon)$ and $b=\sup(A_\varepsilon)$, we get an interval such that $b-a<\varepsilon$ and $f_\#\mu_X([a,b])\geqslant 1-\kappa$.

On the other hand, note $\mathrm{Lip}_1(*,\mathbb{R})=\mathbb{R}$. We claim the following expression, which can be deduced from the definition of Hausdorff distance:

$$d_H^{KF}\left(\varphi_X^*\mathrm{Lip}_1(X,\mathbb{R}),\mathbb{R}\right)=\sup_{f\in\mathrm{Lip}_1(X,\mathbb{R})}\inf_{c\in\mathbb{R}}d_{KF}(f,c).$$

Take $c=(a+b)/2$, so $[a,b]\subset\left[c-\frac{\varepsilon}{2},c+\frac{\varepsilon}{2}\right]$. Thus $f_\#\mu_X\left(\left[c-\frac{\varepsilon}{2},c+\frac{\varepsilon}{2}\right]\right)\geqslant 1-\kappa$ and $\mathrm{diam}\left(\left[c-\frac{\varepsilon}{2},c+\frac{\varepsilon}{2}\right]\right)\leqslant\varepsilon$. Then note

$$f_\#\mu_X\left(\left[c-\frac{\varepsilon}{2},c+\frac{\varepsilon}{2}\right]\right)=\mu_X\left\{x\in X:|f(x)-c|\leqslant\frac{\varepsilon}{2}\right\}\geqslant 1-\kappa,$$

so

$$\mu_X\left\{x\in X:|f(x)-c|\geqslant\frac{\varepsilon}{2}\right\}\leqslant\kappa\leqslant\max(\kappa,\varepsilon/2)$$

$$\mu_X\left\{x\in X:|f(x)-c|\geqslant\max(\kappa,\frac{\varepsilon}{2})\right\}\leqslant\max(\kappa,\varepsilon/2)$$

So $\inf_{c\in\mathbb{R}}d_{KF}(f,c)\leqslant\max(\kappa,\varepsilon/2)$, and this same holds for every $f\in\mathrm{Lip}_1(X,\mathbb{R})$. Then, if we let $\varepsilon\to\mathrm{ObsDiam}_\kappa(X)$, we get

$$d_{conc}(X,*)\leqslant\max(\kappa,\mathrm{ObsDiam}_\kappa(X))$$

And since this inequality holds for all κ, taking the infimum proves the desired result. □

14. Lecture 14. March 4. Samir Chowdhury

Notes from a lecture given by Osman Okutan, based on results from Shioya's book.

14.1. **Main result.** Notation: $S^n(r)$ denotes the metric measure space with the following:

- standard n-dimensional sphere of radius r in $\mathbb{R}^{n+1}$
- σ^n normalized Riemannian measure
- γ^k standard k-dimensional Gaussian measure on $\mathbb{R}^k$ with density function
$$d\gamma^k = \frac{1}{(2\pi)^{k/2}} e^{-\frac{\|x\|_2^2}{2}} dx.$$
- $I(r) = \gamma^1([0,r])$.

Theorem 14.1 (Theorem 2.21 in Shioya's book)**.** *For any $0 < \kappa < 1$, we have:*

(1)
$$\lim_{n\to\infty} \mathrm{ObsDiam}_\kappa(S^n(\sqrt{n})) = \mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, \gamma^1) = 2I^{-1}(\frac{1-\kappa}{2}).$$

(2)
$$\mathrm{ObsDiam}_\kappa(S^n(1)) = O(n^{-1/2}).$$

Hence $S^n(1)$ is a Levy family.

14.2. **Review of partial and observable diameter.**

Definition 14.2. Let (X, d, μ) be an mm-space.

- For $\alpha < 1$, define
$$\mathrm{PartDiam}_\alpha(X) := \inf\{\mathrm{diam}(A) : A \subseteq X, \mu(A) \geqslant \alpha\}.$$
- For $\kappa > 0$,
$$\mathrm{ObsDiam}_\kappa(X) = \sup\{\mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, f_\#\mu) : f : X \to \mathbb{R} \text{ 1-Lipschitz}\}.$$
- $(X_n)_n$ is called a Levy family if $\mathrm{ObsDiam}_\kappa(X_n) \to 0$ as $n \to \infty$ for each $0 < \kappa < 1$.

Definition 14.3 (Lipschitz order)**.** Let X, Y be mm-spaces. We say X is dominated by Y and write $X \prec Y$ if there exists a 1-Lipschitz function $F : Y \to X$ such that $F_\#\mu_Y = \mu_X$.

Proposition 14.4 (Prop 2.18 in Shioya)**.** *Let $X \prec Y$, and let $\kappa > 0$. Then,*

(1)
$$\mathrm{PartDiam}_{1-\kappa}(X) \leqslant \mathrm{PartDiam}_{1-k}(Y).$$

(2)
$$\mathrm{ObsDiam}_\kappa(X) \leqslant \mathrm{PartDiam}_{1-\kappa}(X).$$

(3)
$$\mathrm{ObsDiam}_\kappa(X) \leqslant \mathrm{ObsDiam}_\kappa(Y).$$

Proof. Fix $F : Y \to X$ 1-Lip such that $F_\# \mu_Y = \mu_X$.

For the first assertion, take $A \subseteq Y$ measurable such that $\mu_Y(A) \geqslant 1-\kappa$. Consider $\overline{F(A)} \subseteq X$. Then $\mu_X(\overline{F(A)}) = \mu_Y(F^{-1}(\overline{F(A)})) \geqslant \mu_Y(A) = 1-\kappa$.

Also, $\mathrm{diam}(\overline{F(A)}) = \mathrm{diam}(F(A)) \leqslant \mathrm{diam}(A)$.

Thus $\mathrm{PartDiam}_{1-\kappa}(X) \leqslant \mathrm{diam}(\overline{F(A)}) \leqslant \mathrm{diam}(A)$.

Infimizing over A, we get $\mathrm{PartDiam}_{1-\kappa}(X) \leqslant \mathrm{PartDiam}_{1-\kappa}(Y)$.

For the second assertion, take $f : X \to \mathbb{R}$ 1-Lip. Then $(\mathbb{R}, f_\# d\mu_X) \prec X$. Thus we have $\mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, f_\#\mu_X) \leqslant \mathrm{PartDiam}_{1-\kappa}(X)$.

Taking supremum over f, we get $\mathrm{ObsDiam}_\kappa(X) \leqslant \mathrm{PartDiam}_{1-\kappa}(X)$.

For the third assertion, take $f : X \to \mathbb{R}$ 1-Lip and define $\widetilde{f} := f \circ F$. Then we have:

$$\mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, f_\#\mu_X) = \mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, \widetilde{f}_\#\mu_Y) \leqslant \mathrm{ObsDiam}_\kappa(Y).$$

Taking supremum over f, we get $\mathrm{ObsDiam}_\kappa(X) \leqslant \mathrm{ObsDiam}_\kappa(Y)$.

□

Proposition 14.5 (Proposition 2.19 in Shioya)**.** *Let $\mathcal{X} = (X, d, \mu)$ be an mm-space. For $t > 0$, let $t\mathcal{X} = (X, td, \mu)$. Then* $\mathrm{ObsDiam}_\kappa(t\mathcal{X}) = t\,\mathrm{ObsDiam}_\kappa(\mathcal{X})$.

14.3. **Two statements for proving the main result.** Notation: $\pi_{n,k} : S^n(\sqrt{n}) \to \mathbb{R}^k$ denotes projection onto the first k coordinates.

Definition 14.6 (Convergence of measures)**.** Let X be a metric space and let $\mu, (\mu_n)_{n\in\mathbb{N}}$ be Borel measures. We say μ_n weakly converges to μ if

$$\lim_{n\to\infty} \int_X f d\mu_n = \int_X f d\mu$$

for all continuous, bounded test functions f.

We say μ_n vaguely converges to μ if the above condition holds for all continuous functions with compact support.

Proposition 14.7 (Maxwell-Boltzmann distribution law)**.** *For $d \in \mathbb{Z}_{\geqslant 0}$, we have*

$$\lim_{n\to\infty} \frac{d(\pi_{n,k})_\#(\sigma^n)}{dx} = \frac{d\gamma^k}{dx}.$$

In particular, $(\pi_{n,k})_\#\sigma^n) \to \gamma^k$ weakly as $n \to \infty$.

Theorem 14.8 (Normal law in the way of Levy)**.** *Let $f_n = S^n(\sqrt{n}) \to \mathbb{R}$ for $n \in \mathbb{N}$ be 1-Lip. Assume that for a subsequence $\{f_{n_i}\}$, the pushforwards $(f_{n_i})_\#\sigma^{n_i}$ converge vaguely to a Borel measure σ_∞ on $\mathbb{R}$.*

If σ_∞ is not identically zero, then $(\mathbb{R}, \sigma_\infty) \prec (\mathbb{R}, \gamma^1)$. In other words, there exists $\alpha : \mathbb{R} \to \mathbb{R}$ 1-Lip such that $\alpha_\#\gamma^1 = \sigma_\infty$.

Proof of Theorem 14.1. We start with the first assertion. Consider the projection map $\pi_{n,1} : S^n(\sqrt{n}) \to \mathbb{R}$. We have:

$$\begin{aligned} \mathrm{ObsDiam}_\kappa(S^n(\sqrt{n})) &\geqslant \mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, (\pi_{n,1})_\#(\sigma^n)) \\ \liminf_{n\to\infty} \mathrm{ObsDiam}_\kappa(S^n(\sqrt{n})) &\geqslant \liminf_{n\to\infty} \mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, (\pi_{n,1})_\#(\sigma^n)) \qquad (*) \\ &= \mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, \gamma^1) \\ &= 2I^{-1}(\frac{1-\kappa}{2}) \end{aligned}$$

Here the equality following $(*)$ assumes that PartDiam is continuous with respect to weak convergence of measures.

Now take $f_n : S^n(\sqrt{n}) \to \mathbb{R}$ 1-Lip such that $\mathrm{ObsDiam}_\kappa(S^n(\sqrt{n})) \sim \mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, (f_n)_\#(\sigma^n))$. Here $\sim$ means the we take the quantities to be as close as we need.

Pick a subsequence $\{f_{n_i}\}$ such that

$$\limsup_{n\to\infty} \mathrm{ObsDiam}_\kappa(S^n(\sqrt{n})) = \lim_{i\to\infty} \mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, (f_{n_i})_\#(\sigma^n))$$

and $(f_{n_i})_\#(\sigma^n)$ weakly converges to some σ_∞ (which is nonzero by the above lim inf calculation). Then

$$\limsup_{n\to\infty} \mathrm{ObsDiam}_\kappa(S^n(\sqrt{n})) = \mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, \sigma_\infty) \leqslant \mathrm{PartDiam}_{1-\kappa}(\mathbb{R}, \gamma^1) = 2I^{-1}(\frac{1-\kappa}{2}).$$

The second part of the theorem follows from the first part and the rescaling property of ObsDiam. □

14.4. **Proof of the Maxwell-Boltzmann distribution law.** Consider the projection $\pi_{n,k}(S^n(\sqrt{n})) \to \mathbb{R}^k$. We wish to show $(\pi_{n,k})_\#(\sigma^n) \to \gamma^k$.

Now $(\pi_{n,k})^{-1}(x)$ is isometric to the $(n-k)$-dimensional sphere with radius $(n - \|x\|_2^2)^{1/2}$. Then we have:

$$\frac{d(\pi_{n,k})_\#\sigma^n}{dx} = \frac{\mathrm{vol}_{n-k}(\pi_{n,k}^{-1}(x))}{\mathrm{vol}_n(S^n(\sqrt{n}))} = \frac{(n-\|x\|_2^2)^{\frac{n-k}{2}}}{\int (n-\|x\|_2^2)^{\frac{n-k}{2}} dx}.$$

As $n \to \infty$, the latter converges to

$$\frac{e^{-\frac{\|x\|_2^2}{2}}}{\int_{\mathbb{R}^k} e^{-\frac{\|x\|_2^2}{2}} dx} = \frac{1}{(2\pi)^{k/2}} e^{-\frac{\|x\|_2^2}{2}}.$$

15. Lecture 15. March 6. Osman Okutan

Notes from a lecture given by Kritika Singhal.

15.1. **The Box Distance.**

Definition 15.1 (The Box Distance)**.** Let $\mathcal{X} = (X, d_X, \mu_X)$ and $\mathcal{Y} = (Y, d_Y, \mu_Y)$ be metric measure spaces. Let $I := [0, 1)$ with the Lebesgue measure $\mathcal{L}$. For $\lambda > 0$, the *box distance* $\square_\lambda(\mathcal{X}, \mathcal{Y})$ is defined by

$$\begin{aligned}\square_\lambda(\mathcal{X}, \mathcal{Y}) := \inf\{\epsilon > 0 :& \exists \varphi_X : I \to X,\ \varphi_Y : I \to Y \text{ such that} (\varphi_X)_\# \mathcal{L} = \mu_X,\ (\varphi_Y)_\# \mathcal{L} = \mu_Y, \\ & \exists I_\epsilon \subseteq I \text{ measurable satisfying } |\varphi_X^* d_X(t, s) - \varphi_Y^* d_Y(t, s)| < \epsilon \ \forall t, s \in I_\epsilon, \\ & \mathcal{L}(I_\epsilon^c) < \lambda\epsilon\}.\end{aligned}$$

Theorem 15.2. *$\square_1$ is a complete metric on $\mathcal{M}^\omega$, up to isomorphism of mm-spaces.*

Proposition 15.3. $\square_1 \geqslant \mathrm{d}_{conc}$.

Example 15.4. $\square_1(S^n, S^{n+1}) \to 0$ as $n \to \infty$. Note that for any $\epsilon > 0$, $\mu_{S^{n+1}}((S^n)^\epsilon) \to 1$ as $n \to \infty$.

Exercise 15.5. Estimate $\square_1(S^n, S^m)$.

15.2. **Estimates of Box Distance by Kei Funano.**

Definition 15.6 (Uniformly distributed Borel probability measure)**.** A Borel probability measure μ on a metric space X is called *uniformly distributed* if for each x and x' in X and $r > 0$,

$$\mu(B(x, r)) = \mu(B(x', r)).$$

Lemma 15.7. *Let (X, d_X, μ_X) and (Y, d_Y, μ_Y) be mm-spaces such that μ_X and μ_Y are uniformly distributed Borel probability measures. Let $\nu_X(r)$ (resp. $\nu_Y(r)$) denote the measure of a closed ball of radius $r > 0$ in X (resp. Y). If $\nu_X(a + c) \leqslant (1 - c)\nu_Y(a/2)$ for some $a, c > 0$ and $c < 1$, then*

$$\square_1(X, Y) \geqslant c.$$

Proof. Suppose $\square_1(X, Y) < c$. Then there exists a compact subset $T \subseteq [0, 1]$ and parameters $\varphi_X : [0, 1] \to X$ and $\varphi_Y : [0, 1] \to Y$ such that the following hold:

- $\mathcal{L} > 1 - c$,
- $\varphi_X|_T$ and $\varphi_Y|_T$ are continuous,
- For all s and t in T,

$$|d_X(\varphi_X(s), \varphi_X(s)) - d_Y(\varphi_Y(s), \varphi_Y(t))| \leqslant \epsilon.$$

Note that $\varphi_X(T)$ is compact.

Let

$$l := \max\{k \in \mathbb{N} : \exists (p_i)_{i=1}^k \in \varphi_Y(T) \text{ such that } B(p_i, a/2) \cap B(p_j, a/2) = \varnothing \ \forall i, j \text{ distinct}\}.$$

Let $\{p_1, \dots, p_l\} \subseteq \varphi_Y(T)$ such that for all distinct i and j in $\{1, \dots, l\}$,

$$B(p_i, a/2) \cap B(p_j, a/2) = \varnothing.$$

We have,

$$\begin{aligned}\mu_Y(\bigcup_{i=1}^{l} B_Y(p_i, a/2)) &= \sum_{i=1}^{l} \mu_Y(B(p_i, a/2)) \\ &= l\,\nu_Y(a/2) \leqslant 1.\end{aligned}$$

Therefore,

$$l \leqslant \frac{1}{\nu_Y(a/2)}.$$

We also have $\varphi_Y(T) \subseteq \cup_{i=1}^{l} B_Y(p_i, a)$.

For $i \in \{1, \ldots, l\}$, fix $t_i \in T$ such that $p_i = \varphi_Y(t_i)$. Then, we have the following:

Claim 15.8. $\varphi_X(T) \subseteq \bigcup_{i=1}^{l} B_X(\varphi_X(t_i), a + c)$.

Proof of Claim. Let $x \in \varphi_X(T)$. There exists $s_x \in T$ such that $\varphi_X(s_x) = x$. There exists $k \in 1, \ldots, l$ such that $d_Y(p_k, \varphi_Y(s_x)) \leqslant a$, in other words $d_Y(\varphi_Y(t_k), \varphi_Y(s_x)) \leqslant a$. Therefore,

$$d_X(\varphi_X(t_k), x) = d_X(\varphi_X(t_k), \varphi_X(s_x)) \leqslant a + c.$$

This proves the claim. $\square$

We obtain

$$\begin{aligned}1 &\leqslant \frac{\sum_{i=1}^{l} \mu_X(B_X(\varphi_X(t_i), a + c)}{\mu_X(\varphi_X(T))} \\ &= l\frac{\nu_X(a + c)}{\mu_X(\varphi_X(T))} \leqslant \frac{\nu_X(a + c)}{\nu_Y(a/2)\mu_X(\varphi_X(T))} \\ &\leqslant \frac{\nu_X(a + c)}{\nu_Y(a/2)\mathcal{L}(T)} < \frac{\nu_X(a + c)}{\nu_Y(a/2)(1 - c)}.\end{aligned}$$

This is a contradiction. Therefore, $\square_1(X, Y) \geqslant c$. $\square$

Theorem 15.9 (Bishop-Gromov volume comparison theorem)**.** *Let M be a complete n-dimensional Riemannian manifold such that $\mathrm{Ric}_M \geqslant (n-1)K$. Let M_K^n be the n-dimensional simply connected space of sectional curvature K. Then, for any p in M and p_K in M_K^n, we have*

$$\varphi(r) := \frac{\mathrm{vol}_M(B(p, r))}{\mathrm{vol}_{M_K^n}(B(p_K, r))}$$

is non-increasing on $(0, \infty)$ and as consequence

$$\mathrm{vol}_M(B(p, r)) \leqslant \mathrm{vol}_{M_K^n}(B(p_K, r))\ \forall r > 0.$$

Lemma 15.10. *Let M (resp. N) be an m-dimensional (resp. n-dimensional) compact Riemannian manifold having a uniformly distributed Riemannian volume measure. Assume that $\mathrm{Ric}_M \geqslant (m-1)K_1 > 0$ and $\mathrm{Ric}_N > 0$. Let $a_N := \mathrm{vol}(N)/\mathrm{vol}(S^n)$. If $0 < c < 1$ is such that*

$$c^{n-m} \leqslant \frac{(1 - c)na_N(K_1)^{m/2}\Gamma((m + 1)/2)\Gamma(n/2)}{m2^{m+1}\pi^{m-1}\Gamma(m/2)\Gamma((n + 1)/2)}$$

and $c\sqrt{K_1} \leqslant \pi$, then $\square_1(M, N) \geqslant C$.

Proof. For $r > 0$, ;et $\nu_M(r) = \mu_M(B(x,r))$ for $x \in M$ and $\nu_N(r) = \mu_N(B(y,r))$ for $y \in N$. By Bishop-Gromov volume comparison theorem, we have

$$\begin{aligned}\nu_M(c/2) &= \mu_M(B_M(x,c/2)) \\ &= \frac{\mathrm{vol}(B_M(x,c/2))}{\mathrm{vol}(M)} \\ &\geqslant \frac{\mathrm{vol}(B_{M^m_{K_1}}(x',c/2))}{\mathrm{vol}(M^m_{K_1})} = \frac{\mathrm{vol}(S^{m-1})}{\mathrm{vol}(S^m)}\int_0^{c\sqrt{K_1}/2} \sin^{m-1}\theta d\theta.\end{aligned}$$

Since $c\sqrt{K_1} \leqslant \pi$, if $\theta \in [0, c\sqrt{K_1}/2]$, then $c \in [0,\pi/2]$ and $\sin\theta \geqslant 2\theta/\pi$. Hence,

$$\nu_M(c/2) \geqslant \frac{\mathrm{vol}(S^{m-1})}{\mathrm{vol}(S^m)}\frac{c^m(K_1)^{m/2}}{2^m m}\frac{2^{m-1}}{\pi^{m-1}}.$$

Since

$$\mathrm{vol}(S^n) = \frac{2\pi^{(n+1)/2}}{\Gamma((n+1)/2)}$$

, we get

$$\nu_M(c/2) \geqslant \frac{\Gamma((m+1)/2)}{\Gamma(m/2)}\frac{c^m K_1^{m/2}}{2m\pi^{m-\frac{1}{2}}}.$$

Let $K_2 > 0$ be such that $\mathrm{Ric}_N \geqslant (n-1)K_2 > 0$. We have

$$\begin{aligned}\nu_N(2c) &= \frac{\mathrm{vol}(B_N(y,2c))}{\mathrm{vol}(N)} \leqslant \frac{\mathrm{vol}(B_{M^n_{K_2}}(y,2c))}{\mathrm{vol}(N)} \\ &= \frac{\mathrm{vol}(S^{n-1})}{a_N(K_2)^{n/2}\mathrm{vol}(S^n)}\int_0^{2c\sqrt{K_2}} \sin^{n-1}\theta d\theta \\ &\leqslant \frac{\mathrm{vol}(S^{n-1})}{\mathrm{vol}(S^n)}\frac{(2c)^n(K_2)^{n/2}}{na_N(K_2)^{n/2}} = \frac{(2c)^n\pi^{n/2}\Gamma((n+1)/2)}{na_N\pi^{(n+1)/2}\Gamma(n/2)}\end{aligned}$$

Hence,

$$\nu_N(2c) \leqslant (1-c)\nu_M(c/2).$$

By letting $a = c$ in Lemma 15.7, we get

$$\square_1(M,N) \geqslant c.$$

$\square$

Proposition 15.11. *Let $(n_k)_{k=1}^\infty$ and $(m_k)_{k=1}^\infty$ be sequences of natural numbers such that $n_k \leqslant c_1 k$, $m_k \leqslant c_2 k$ and $|n_k - m_k| \geqslant c_3 k$ for some $c_1, c_2, c_3 > 0$, for all $k \in \mathbb{N}$. Then, both $\liminf_{k\to\infty} \square_1(S^{n_k}, S^{m_k})$ and $\liminf_{k\to\infty} \square_1(\mathbb{C}P^{n_k}, \mathbb{C}P^{m_k})$ are greater than or equal to*

$$\min(2^{-c_1/c_3}\pi^{-c_2/c_3}, 2^{-c_2/c_3}\pi^{-c_1/c_3}).$$

Proof. Without loss of generality we can assume that $n_k \geqslant m_k$ for all $k \in \mathbb{N}$. Since $0 < c < 1$, $c^{n_k - m_k} \leqslant c^{c_3 k}$. Substituting $n = n_k, m = m_k$ we obtain

$$\frac{(1-c)n_k\Gamma((m_k+1)/2)\Gamma(n_k/2)}{m_k 2^{n_k+1}\pi^{m_k-1}\Gamma(m_k/2)\Gamma((n_k+1)/2)} \geqslant (1-c)2^{-c_1k-1}\pi^{-c_2k+1}\frac{\Gamma(n_k/2)}{\Gamma((n_k+1)/2)}.$$

If

$$c \leqslant \left(\frac{(1-c)n_k \Gamma(n_k/2)}{2m_k \Gamma((n_k+1)/2)} \right)^{\frac{1}{c_3 k}} 2^{-c_1/c_3} \pi^{(-c_2/c_3)+(1/k)},$$

we get $\square_1(S^{n_k}, S^{m_k}) \geqslant c$.

$$\left(\frac{1-c}{2} \frac{n_k}{2m_k} \frac{\Gamma(n_k/2)}{\Gamma((n_k+1)/2)} \right)^{\frac{1}{c_3 k}} \to 1 \text{ as } k \to \infty.$$

For k large enough, we have

$$2^{-c_1/c_3} 2^{-c_2/c_3} = c \leqslant \left(\frac{1-c}{2} \frac{n_k}{2m_k} \frac{\Gamma(n_k/2)}{\Gamma((n_k+1)/2)} \right)^{\frac{1}{c_3 k}} 2^{-c_1/c_3} \pi^{\frac{-c_2}{c_3} + \frac{1}{c_3 k}}.$$

This completes the proof for spheres. For $\mathbb{C}P^n$, use the same strategy along with the fact $\mathrm{vol}(\mathbb{C}P^n) = \pi^n/n!$. □

16. Lecture 16. March 18th. Zhengchao Wan

Notes from a lecture given by Sunhyuk Lim.

Reference.

- Metric Measure Geometry, Section 2.6 (Shioya)
- Estimates of eigenvalues of the Laplacian by a reduced number of subsets (Funano)

Definition 16.1 (Separation Distance)**.** Let $\mathcal{X} = (X, d_X, \mu_X)$ be an metric measure space. For any real numbers $s_0, s_1, \cdots, s_N > 0$ with $N \geqslant 1$, we define the Separation Distance

$$\mathrm{Sep}(\mathcal{X}; s_0, \cdots, s_N) := \sup \left\{ \min_{i \neq j} d_X(A_i, A_j) : \begin{array}{c} A_0, \cdots, A_N \text{ are Borel subsets of } X \text{ such that} \\ \mu_X(A_i) \geqslant s_i \text{ for each } i = 0, \cdots, N. \end{array} \right\},$$

where for any subsets $A, B \subset X$, $d_X(A, B) := \inf_{a \in A, b \in B} d_X(a, b)$.

Remark 16.2. $s_0 \leqslant s_0', \cdots, s_N \leqslant s_N' \Rightarrow \mathrm{Sep}(\mathcal{X}; s_0, \cdots, s_N) \geqslant \mathrm{Sep}(\mathcal{X}; s_0', \cdots, s_N')$.

Notations: denote $\mathcal{M}$ as a closed and connected Riemannian manifold.

- $(\mathcal{M}, d_{\mathcal{M}}, \mathrm{nvol}_{\mathcal{M}})$: A mm-space with $\mathrm{nvol}_{\mathcal{M}} = \frac{\mathrm{vol}_{\mathcal{M}}}{\mathrm{vol}(\mathcal{M})}$.
- $\Delta_{\mathcal{M}}$: Laplacian operator on $L^2(\mathcal{M}) \cap \mathrm{Lip}(\mathcal{M})$.
- $\lambda_k(\mathcal{M})$: k-th eigenvalue of $\Delta_{\mathcal{M}}$.

Theorem 16.3.

$$\mathrm{Sep}(\mathcal{M}; s_0, \cdots, s_k) \leqslant \frac{2}{\sqrt{\lambda_k \cdot \min_{i=0,\cdots,k} s_i}}.$$

Proposition 16.4. *For any mm-space $\mathcal{X} = (X, d_X, \mu_X)$ and any real number $s > 0$,*

$$\mathrm{ObsDiam}_{2s}(\mathcal{X}) \leqslant \mathrm{Sep}(\mathcal{X}; s, s).$$

Corollary 16.5. *Let $\mathcal{M}$ be a closed connected Riemannian manifold, then*

$$\mathrm{ObsDiam}_{2s}(\mathcal{M}) \leqslant \mathrm{Sep}(\mathcal{M}; s, s) \leqslant \frac{2}{\sqrt{\lambda_1(\mathcal{M}) \cdot s}}.$$

Remark 16.6. If a set of closed connected Riemannian manifolds $\{\mathcal{M}_n\}_{n=1}^{\infty}$ satisfies $\lambda_1(\mathcal{M}_n) \nearrow \infty$ as $n \to \infty$, then $\{\mathcal{M}_n\}_{n=1}^{\infty}$ is a Lévy family.

Fact 16.7. k-th eigenvalue of Laplacian of $\mathbb{S}^n$ is $k(k+n-1)$. This implies that $\lambda_1(\mathbb{S}^n) = n \nearrow \infty$ as $n \nearrow \infty$.

The following facts are used to prove Theorem 16.3.

Facts. Let $\mathcal{M}$ be a closed connected Riemannian manifold.

(1) $0 = \lambda_0(\mathcal{M}) < \lambda_1(\mathcal{M}) \leqslant \lambda_2(\mathcal{M}) \leqslant \cdots \nearrow \infty$.

(2) (Rayleigh Quotient) For each k,

$$\lambda_k(\mathcal{M}) = \inf_{L} \sup_{u \notin L \setminus \{0\}} R(u),$$

where L runs over all $(k+1)$-dimensional Linear subspaces of $L^2(\mathcal{M}) \cap \mathrm{Lip}(\mathcal{M})$ and

$$R(u) = \frac{\|\nabla u\|^2_{L^2(\mathrm{nvol}_{\mathcal{M}})}}{\|u\|^2_{L^2(\mathrm{nvol}_{\mathcal{M}})}} = \frac{\int_{\mathcal{M}} \langle \nabla u(x), \nabla u(x)\rangle d\mathrm{nvol}_{\mathcal{M}}(x)}{\int_{\mathcal{M}} |u(x)|^2 d\mathrm{nvol}_{\mathcal{M}}(x)},$$

where the gradient of a Lipschitz function exists almost everywhere because of Rademacher's theorem.

Proof of Theorem 16.3. Set $s = \mathrm{Sep}(\mathcal{X}; s_0, \cdots, s_k)$ for simplicity. Assume $s > 0$. Choose arbitrary r such that $0 < r < s$. This implies, $\exists$ Borel subsets $A_0, \cdots, A_k$ such that $\mathrm{nvol}_{\mathcal{M}}(A_i) \geqslant s_i$ for $i = 0, \cdots, N$, and $d_{\mathcal{M}}(A_i, A_j) > r$ for $i \neq j$. Define

$$\begin{aligned} f_i : \mathcal{M} &\to \mathbb{R} \\ x &\mapsto \max\{1 - \frac{2}{r} d_{\mathcal{M}}(x, A_i), 0\} \end{aligned}$$

for each $i = 0, \cdots, k$.

Then f_i satisfies the following properties.

(1) f_i is $\frac{2}{r}$-Lipschitz and $f_i \in L^2(\mathrm{nvol}_{\mathcal{M}})$.

(2) $\|\nabla f_i\| \leqslant \frac{2}{r}$.

(3) $\{f_i\}_{i=0}^{k}$ is orthogonal, i.e., $\langle f_i, f_j \rangle_{L^2(\mathrm{nvol}_{\mathcal{M}})} = 0$.

(4) $\|f_i\|^2 \geqslant s_i$.

Then by taking $L_0 = \mathrm{span}\{f_0, \cdots, f_k\}$, one obtains

$$\lambda_k(\mathcal{M}) = \inf_{L} \sup_{u \in L \setminus \{0\}} R(u) \leqslant \sup_{u \in L_0 \setminus \{0\}} R(u).$$

For any $u \in L_0\backslash\{0\}$, it can be written as $u = a_0 f_0 + \cdots + a_k f_k$. Hence

$$\|u\|^2 \geqslant a_0^2 s_0 + \cdots + a_k^2 s_k \geqslant (a_0^2 + \cdots + a_k^2) \min_{i=0,\cdots,k} s_i,$$

and

$$\|\nabla u\|^2 \leqslant \frac{4}{r^2}(a_0^2 + \cdots + a_k^2).$$

Therefore

$$R(u) \leqslant \frac{4}{r^2 \cdot \min_i s_i} \Rightarrow \lambda_k(\mathcal{M}) \leqslant \frac{4}{r^2 \cdot \min_i s_i}.$$

□

Theorem 16.8 (Chung-Grigor'yan-Yau)**.** *There exists a universal constant $c > 0$ satisfying the following property. Denote $(\mathcal{M}, \mu)$ as a closed connected and weighted Riemannian manifold. Then*

$$\mathrm{Sep}((\mathcal{M}, \mu); s_0, \cdots, s_k) \leqslant \frac{c}{\sqrt{\lambda_k(\mathcal{M}, \mu)}} \cdot \max_{i=0,\cdots,k} \log \frac{1}{s_i}.$$

Remark 16.9. Funano pointed out that Theorem 16.8 still holds for weighted compact connected finite dimensional Alexandrov spaces.

Theorem 16.10. $\exists$ *universal constant $c > 0$ satisfying the following property. Let (X, μ) be a weighted compact connected finite dimensional Alexandrov space satisfying* $\mathrm{CD}(0, \infty)$*. Then*

$$\mathrm{Sep}((X, \mu); s_0, \cdots, s_l) \leqslant \frac{c^{k-l+1}}{\sqrt{\lambda_k(X, \mu)}} \cdot \max_{i=0,\cdots,l} \log \frac{1}{s_i}$$

for any $l \leqslant k$.

This theorem arises in an effort to prove the following conjecture.

Conjecture 16.11. $\lambda_{k+1}(X, \mu) \leqslant C \cdot \lambda_k(X, \mu)$.

To clarify the theorem, we will recall the definition of Alexandrov space and $\mathrm{CD}(0, \infty)$ condition.

Definition 16.12 (Alexandrov space)**.** For a complete geodesic space (X, d_X) and $K \in \mathbb{R}$, any geodesic triangle $\Delta(x, y, z)$ with perimeter $\leqslant 2D_K$, where $D_K = \frac{\pi}{\sqrt{K}}$ when $K > 0$ and $D_K = \infty$ when $K \leqslant 0$, has a comparison triangle $\Delta(\bar{x}, \bar{y}, \bar{z})$ in the model space $\mathcal{M}_K$ such that $d_X(x, y) = d_{\mathcal{M}_K}(\bar{x}, \bar{y}), d_X(x, z) = d_{\mathcal{M}_K}(\bar{x}, \bar{z}), d_X(z, y) = d_{\mathcal{M}_K}(\bar{z}, \bar{y})$. We say (X, d_X) is an Alexandrov space of curvature bounded below by $K \in \mathbb{R}$, if for any geodesic triangle $\Delta(x, y, z)$ and any point $w \in [y, z]$, we have $d_X(x, w) \geqslant d_{\mathcal{M}_K}(\bar{x}, \bar{w})$, where $\bar{w}$ is a point on the side $[\bar{y}, \bar{z}]$ of the comparison triangle $\Delta(\bar{x}, \bar{y}, \bar{z})$ such that $d_X(y, w) \geqslant d_{\mathcal{M}_K}(\bar{y}, \bar{w})$.

Cheeger has generalized the Laplacian operator to Alexandrov spaces.

Definition 16.13 (Sturm)**.** Given a metric space $\mathcal{X} = (X, d_X, \mu_X)$ and $C \in \mathbb{R}$, the curvature dimension condition $\mathrm{CD}(C, \infty)$ means, for any $\nu_0, \nu_1 \in \mathcal{P}_2(X) = (\mathcal{P}_2(X), d_{W,2})$, there exists a minimal geodesic $\nu_t : [0,1] \to \mathcal{P}_2(X)$ such that

$$\mathrm{Ent}_{\mu_X}(\nu_t) \leqslant (1-t)\mathrm{Ent}_{\mu_X}(\nu_0) + t\mathrm{Ent}_{\mu_X}(\nu_1) - \frac{C}{2}(1-t)td^2_{W,2}(\nu_0,\nu_1),\ \forall t \in [0,1],$$

where

$$\mathrm{Ent}_\mu \nu := \begin{cases} \int_X \rho \log \rho d\mu & \text{if } d\nu = \rho d\mu \\ \infty & \text{otherwise} \end{cases}$$

Example 16.14. Let $\mathcal{M}$ be a Riemannian manifold. Then

$$\mathrm{Ric}_{\mathcal{M}} \geqslant C \Leftrightarrow \mathcal{M} \text{ is } \mathrm{CD}(C,\infty).$$

Example 16.15. An n-dimensional Alexandrov space of curvature bounded below by K satisfies $\mathrm{CD}((n-1)K, \infty)$.

The following is the key Lemma of proving Theorem 16.10.

Lemma 16.16. *Let $\mathcal{X}$ be a weighted finite dimensional Alexandrov space of* $\mathrm{CD}(0,\infty)$*. If (X,μ) satisfies the follwoing for any $s > 0$ and some $D > 0$*

$$\mathrm{Sep}((X,\mu);\ \underbrace{s, \cdots, s}_{(k+1)-times}\ ) \leqslant \frac{1}{D} \cdot \log \frac{1}{s}$$

then we have

$$\mathrm{Sep}((X,\mu); \underbrace{s, \cdots, s}_{k-times}) \leqslant \frac{c}{D} \cdot \log \frac{1}{s}$$

for any $s > 0$ with a universal constant $c > 0$.

17. Lecture 17. March 25th. Paul Duncan

This lecture followed the first three sections of the expository paper on SLE of Kager and Nienhuis. This paper also contains useful appendices with background material. Lawler's summer school notes are also useful, particularly for examples.

18. Lecture 20. April 8. Mario Gómez

This lecture was based on the paper "Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees)" by Greven, Pfaffelhauer and Winter (https://link.springer.com/content/pdf/10.1007%2Fs00440-008-0169-3).
A common theme of recent lectures was using a continuous object to study a discrete random space. That is how, for example, the Continuum Random Tree emerged from studying loop-erased or loop-free graphs. A technical device used in these constructions is embedding trees into ℓ_1 and studying convergence within that metric space. This idea inspires the topic of today's paper. The authors study convergence of mm-spaces without using this particular embedding.
The paper achieves this via the Gromov-Prokhorov metric (Definition 5.1). This turns out

to be a complete and separable metric that metrizes the Gromov-weak topology (Definition 2.8). An important result is Theorem 2, the characterization of pre-compactness. Given a set of mm-spaces, they informally describe pre-compactness via two conditions. First, the spaces in the set should put most of its mass in subspaces of a uniformly bounded diameter; second, the mass of the points that have small mass around them is small. They claim that the conditions involved in the Theorem are reasonably easy to calculate, and they use them in Section 4 by an example involving Λ-coalescent trees (this was not covered in lecture). Notable sections ommited from the lecture are 4, 7, 8, and the appendix. As mentioned above, Section 4 exemplifies the definitions using trees. Sections 7 and 8 are the technical details needed to prove Theorem 2. The Appendix studies other metrics that are equivalent to Gromov-Prokhorov.

19. Lecture 21. April 10. Woojin Kim

See Appendix A.10 for Woojin's handwritten notes. The main sources for Woojin's presentation are the following:

- "Introduction to Stochastic Processes", 2nd edition by Gregory F. Lawler: This book provides an easy introduction to Brownian motions in $\mathbb{R}^n$.
- (Results for $d = 2$) Cover times for Brownian motion and random walks in two dimensions http://annals.math.princeton.edu/wp-content/uploads/annals-v160-n2-p02.pdf
- (Results for $d \geqslant 3$) Brownian Motion on Compact Manifolds: Cover Time and Late Points https://projecteuclid.org/euclid.ejp/1464037588
- A Brief Introduction to Brownian Motion on a Riemannian Manifold https://www.math.kyoto-u.ac.jp/probability/sympo/PSS03abstract.pdf

20. Lecture 23. April 17. Ling Zhou and Zhengchao Wan

The lecture followed Section 9.1, 9.2, 9.3, 9.5 of *Metric Measure Geometry* by Takashi Shioya, https://arxiv.org/abs/1410.0428.

21. Lecture 24. April 22nd. Gustavo

See the Appendix A.15 for handwritten notes of the talk (including the Proof of the theorem we didn't have time to cover). Here is the list of references used:

- "A January 2005 Invitation to Random Groups" by Ollivier.
- "Asymptotic invariants of infinite groups" by Gromov.
- "Notes on word hyperbolic groups" by Alonso, Brady, Cooper, Ferlini, Lustig, Mihalik, Shapiro and Short.
- "A sharper threshold for random groups at density one-half" by Duchin, Jankiewicz, Kilmer, Lelievre, Mackay and Sánchez.

Appendix A. Handwritten notes

(1)

Randomly covering a space — especially by Brownian motion.

Outline:

- Brownian motion in $\mathbb{R}^n$
- " in a Riemannian manifold. & cover time.

We see the importance of PDE to study Brownian motion.

+ (time permit) Covering mm-space.

List of References

- Papers by Dembo, Peres, Rosen (~~keywords~~ cover time, Brownian motion)

(2).

Brownian Motion in Euclidean space.

Def) Brownian motion (with variance 1, starting at 0) in $\mathbb{R}^n$ is a random cts function $X : [0,\infty) \to \mathbb{R}^n$

(i) $X_0 = 0$

(ii) For any $s_1 \le t_1 \le s_2 \le t_2$, $X_{t_1} - X_{s_1}$ and $X_{t_2} - X_{s_2}$ are independent.

(iii). $X_t - X_s \sim N^d(0,(t-s))$.

Thus, we have the associated prob. measure $\mathbb{P}^0$ on $C([0,\infty), \mathbb{R}^n)$.

Similarly, $\mathbb{P}^x$ is defined.

Another way to define Brownian motion Let X be a Brownian motion in $\mathbb{R}^n$ with $X_0 = x$.

Let $P_t(x,\cdot)$ be the probability density function of X_t. Then, from (iii),

$$P_t(x,y) = \frac{1}{(2\pi t)^{d/2}} e^{-|y-x|^2/2t}, \quad (y \in \mathbb{R}^n, t \in [0,\infty)).$$

$\hookrightarrow$ transition density function of the Brownian motion X.

once the initial point x_0 is given, then we recover the original definition.

Remark 1 $P_t(x,y)$ is the fundamental solution of the heat equation, i.e.

$$\frac{\partial}{\partial t} P_t(x,y) = \frac{1}{2}\underbrace{\Delta_x}_{\sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}} P_t(x,y), \qquad \lim_{t \to 0} P_t(x,y) = \delta_x(y).$$

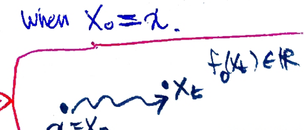
Remark 2 Let $f_0 : \mathbb{R}^n \to \mathbb{R}_+$ be an initial p.d.f on $\mathbb{R}^n$.

Let $f_t : \mathbb{R}^n \to \mathbb{R}_+$ be the distribution at $t \in [0,\infty)$ under Brownian motion.

Then,
$$f_t(x) = \int_{y \in \mathbb{R}^d} f_0(y) \cdot P_t(y,x)\, dy$$
$$= \int_{y \in \mathbb{R}^d} f_0(y) \cdot P_t(x,y)\, dy.$$

$P_t(x,\cdot) : \mathbb{R}^n \to \mathbb{R}_+$ was p.d.f of X_t when $X_0 = x$.

$$= \boxed{\mathbb{E}^x\big(f_0(X_t)\big)} \ !!$$



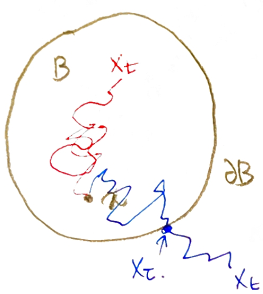
Using this, one can prove that $\frac{\partial}{\partial t} f = \frac{1}{2}\Delta f$.

So far, we did "Brownian $\longrightarrow$ Heat equation. (PDE)".

Conversely, we have the following theorem:

<u>Theorem 1</u> (Stochastic representation for solutions of the Dirichlet problem).

Let B be a bounded open in $\mathbb{R}^n$. Consider the PDE:



$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{1}{2}\Delta u & \text{on } B \\[2ex] u(0,x) = f(x) & \text{on } B \quad (I,C) \\ u(t,x) = g(x) & \text{on } \partial B \quad (B,C) \end{cases}$$

$\tau = \inf\{t \geq 0 : X_t \in B^c\}$.

Brownian perspective:

Then, $\boxed{u(t,x) = \mathbb{E}^x\left(f(X_t)\cdot \mathbb{1}\{\tau > t\} + g(X_\tau)\cdot \mathbb{1}\{\tau \leq t\}\right)}$.

Also, as $t \to \infty$, $u(t,\cdot) \to v \in C^2(B) \cap C(\overline{B})$ s.t.

(1) $\Delta v = 0$.

(2) $v = g$ on ∂B.

$\leftarrow$ this PDE has nothing to do with t.

In particular, $v = \mathbb{E}^x(\, g(X_\tau)\,)$.

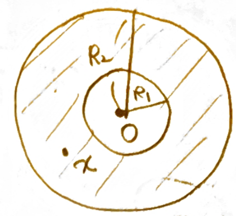
$\mathbb{P}^x$ is a measure on $C([0,\infty), \mathbb{R}^n)$.

↙ Good to see the difference of Brownian motions in different dimensions.

(5)

Recurrence and Transience. Let $0 < R_1 < R_2$.

$B = B(R_1, R_2) := \{x \in \mathbb{R}^d : R_1 < |x| < R_2\}$.



Suppose $x \in B$. Let $\tau := \tau_{\partial B} = \inf\{t \geq 0 : X_t \in \partial B\}$.

Define $g : \partial B \to \{0, 1\}$ as $y \mapsto \begin{cases} 0, & \|y\| = R_1 \\ 1 & \|y\| = R_2 \end{cases}$.

Then, the probability that $X : [0, \infty) \to \mathbb{R}^d$ $(X_0 = x)$ hits the outer sphere first is $f(x) = \mathbb{E}^x(g(X_\tau))$, which is the solution of PDE

$$(*) \quad \begin{cases} \Delta f = 0 \\ f = 1 \text{ on } \{\|y\| = R_2\} \ \& \ f = 0 \text{ on } \{\|y\| = R_1\}. \end{cases}$$

Note that f is a function that depends on $\|x\|$, i.e. $f(x) = \phi(\|x\|)$.

Write out the equation $(*)$ in spherical coordinates:

$$\Delta \phi(r) = \phi''(r) + \frac{d-1}{r}\phi'(r) = 0$$

← solve by separation of variable.

use: $\phi(R_1) = 0$ & $\phi(R_2) = 1$.

$$\Rightarrow \quad f(x) = \phi(\|x\|) = \begin{cases} \dfrac{\ln\|x\| - \ln R_1}{\ln R_2 - \ln R_1}, & d = 2 \\[2ex] \dfrac{R_1^{2-d} - \|x\|^{2-d}}{R_1^{2-d} - R_2^{2-d}}, & d \neq 2. \end{cases}$$

$$\frac{R_1 - \|x\|}{R_1 - R_2}$$

Observe ① $\|x\| = R_1 \Rightarrow f(x) = 0$.
② $\|x\| = R_2 \Rightarrow f(x) = 1$.

Interpretation

Whatever R2 is, hits R_2 first.

$$\lim_{R_1\to 0} f(x) = \begin{cases} 1, & d=2 \\ 1, & d\geq 3 \\ \frac{\|x\|}{R_2}, & d=1. \end{cases}$$

$d=2$, $d\geq 3$ $\Rightarrow$ $\mathbb{P}^x$(ever return to zero) $=0$.

$d=1$: $\Rightarrow 1-\frac{\|x\|}{R_2}>0 \Rightarrow \mathbb{P}^x$(eventually ever return to zero) $= 1$.

Skip. Exercise. Think about $\lim_{R_2\to\infty} f(x)$ and interpret it.

$$\lim_{R_2\to\infty} f(x) = \begin{cases} 0, & d=2 \\ 1-\frac{\|x\|^{2-d}}{R_1^{2-d}}, & d\geq 3 \\ 0, & d=1. \end{cases}$$

$d=2$: $\Longrightarrow$ never go to ∞, keep coming back to the circle of R_1

$d\geq 3$: $\Longrightarrow$ has a chance to go off to ∞. (eventually will).

$\square \Rightarrow$ In $\mathbb{R}^2$, the Brownian motion is neighborhood recurrent, but not point recurrent

$\bigcirc \Rightarrow$ In $\mathbb{R}^3$, " transient. (go back to an open bounded set at most finitely many times) and not point recurrent.

In $\mathbb{R}^1$, point recurrent. neighborhood.

Brownian motion on a manifold & covering time.

(7).

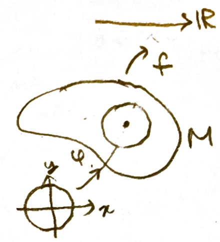
Let M be a smooth, compact, connected Riemannian manifold w.o. boundary.

Laplace-Beltrami operator $\Delta_M : C^2(M) \to C(M)$ (analogue of Δ).

On a local coordinate, $\Delta_M = \sum_{i,j} g^{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b^i \frac{\partial}{\partial x_i}$

in flat space $= \Delta$

smooth functions.



The solution of the following PDE

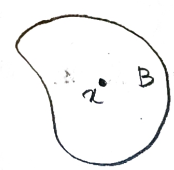
$$\frac{\partial p}{\partial t} = \frac{1}{2}\Delta_M p, \quad \lim_{t \to 0} P_t(x,y) = \delta_x(y)$$

will be the transition density function for the Brownian motion on M.

Theorem 1′ (Stochastic representation for solution). Theorem 1 still holds. Also,

$g(x) = \mathbb{E}^x(\tau_{\partial B})$ is a solution of

$$\begin{cases} \Delta_M g = -1 \\ g(x) = 0 \quad \text{when } x \in \partial B. \end{cases}$$



Let $C_\varepsilon(M)$ = time needed for the trace of a particle to become ε-cover of M.

(Wiener sausage — named after Norbert Wiener, also Wiener is German for Viennese)

• expected volume was studied in 1964.

<u>Theorem 2</u> For a Brownian motion in M^d,

- $d=2$,

$$\lim_{\varepsilon\to 0}\frac{C_\varepsilon(M)}{(\log\frac{1}{\varepsilon})^2} = \frac{2}{\pi}\cdot \mathrm{area}(M) \quad \text{a.s.}$$

- $d\geq 3$

$$\lim_{\varepsilon\to 0}\frac{C_\varepsilon(M)}{\varepsilon^{2-d}\log\frac{1}{\varepsilon}} = \boxed{d\cdot\frac{2}{(d-2)\mathrm{Vol}(S^{n-1})}\cdot \mathrm{Vol}(M)} \quad \text{a.s.}$$

Proof strategy).

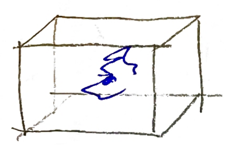
- $d=2$: Use "isothermal coordinate system".

  $\exists\, \alpha(M)>0,\ \forall x\in M,\ \Delta_M$ on $B(x,\alpha)$ is expressed as $a_x\cdot\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}\right)$

  smooth $B(x,\alpha)\to\mathbb{R}$.

  Reduce the problem to flat torus

- $d\geq 3$: no such things. Work directly on M.

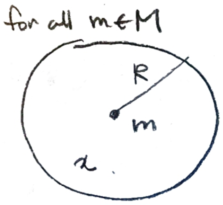


But, proof relys on similar lemmas which follow. (We focus on $d\geq 3$).

<u>Lemma 1</u> (Interior mean exit time) For sufficiently small $R>0$, for all $m\in M$

$$\mathbb{E}^x(\tau_R) = \frac{R^2-d(m,x)^2}{d} + O(R^3)$$

(first hitting time)



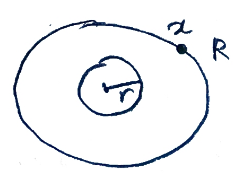
Proof) solve PDE.
(Thm 1′).

<u>Lemma 2</u> For sufficiently small $R>r>0$, for all $m\in M$ and for all $x\in\partial B(m,R)$

(a) $\alpha(R,r)\leq \mathbb{E}^x(\tau_r)\leq \beta(R,r)$.

(b) (Exterior mean exit time).

$$\|\tau_r\| := \sup_{x\in M}\mathbb{E}^x(\tau_r) \leq c\cdot r^{2-d}.$$



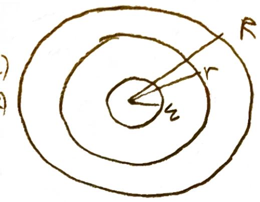
proof). Use PDE, maximal principle etc.

(9).

Lemma 3 (Hitting probabilities) For any $\delta > 0$, there exists $R_0(\delta) > 0$ s.t. if $R_0 \geq R \geq 2r \geq 4\epsilon$, then for all $m \in M$, for all $x \in \partial B(m,r)$.

$$(1-\delta)\frac{r^{2-d}-R^{2-d}}{\epsilon^{2-d}-R^{2-d}} \leq \overset{=g(x)}{\mathbb{P}^x\left(d(m, X_{\tau_{\epsilon,R}}) = \epsilon\right)} \leq (1+\delta)\frac{r^{2-d}-R^{2-d}}{\epsilon^{2-d}-R^{2-d}}$$

We computed these in $\mathbb{R}^d$



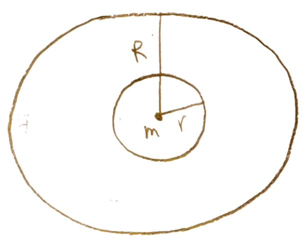
Proof. $g(x) = \mathbb{E}^x\left(h(X_{\tau_{\epsilon,R}})\right)$ where $h = 1$ on $\partial B(m,\epsilon)$, $h = 0$ on $\partial B(m,R)$

$$\Rightarrow \begin{cases} \Delta_M g = 0 & \text{on } B(m,R)\setminus B(m,\epsilon) \\ g = h & \text{on } \partial B(m,R) \cup \partial B(m,\epsilon) \end{cases}$$

Using same trick, bound $g(x)$. ⊠

Concentration of excursion times.

x °



$\tau^{(0)} := \tau_R$.

For $j \geq 1$,

$\tau^{(j)} :=$ The length of jth excursion from $\partial B(m,R)$ to itself via $\partial B(x,r)$.

(Note that $\tau^{(j)}$, $j = 1, 2, \cdots$ are i.i.d).

$\boxed{\text{FACT}}$ $\mathbb{E}^x(\tau^{(1)}) = K_M \cdot (r^{2-d} - R^{2-d})$. (go to infinite as $r \downarrow 0$).

$$\Rightarrow \mathbb{E}^x\left(\sum_{j=1}^N \tau^{(j)}\right) = N \cdot K_M \cdot (r^{2-d} - R^{2-d}) = M_{r,R,N}.$$

Lemma 4 $$\mathbb{P}^x\left(\left|\sum_{j=1}^N \tau^{(j)} - M_{r,R,N}\right| > \delta\right) \leq f(N, x).$$

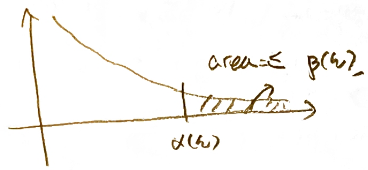
Use Chebycheff's inequality, (and difficult techniques).

Let $x \in M$. Let $T(x,\epsilon) := \tau_{\partial B(x,\epsilon)}$. (the hitting time to $\partial B(x,\epsilon)$).

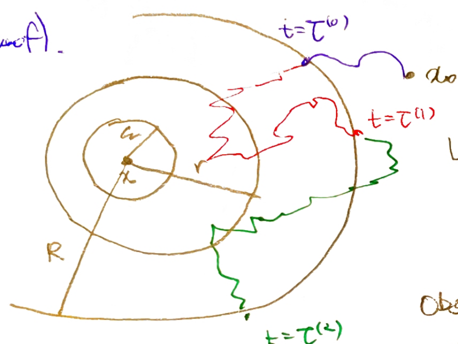
Note that $C_\epsilon(M) := \sup_{x \in M} T(x,\epsilon)$. Hence, controlling $T(x,\epsilon)$ is important.

Lemma 5 (Tail of $T(x,\epsilon)$). $\forall x, x_0 \in M$, $\exists\, \alpha, \beta : \mathbb{R}_+ \to \mathbb{R}_+$

$$\mathbb{P}^{x_0}\big(T(x,\epsilon) \geq \alpha(\epsilon)\big) \leq \beta(\epsilon), \text{ where } \lim_{\epsilon \to 0} \alpha(\epsilon) = \infty, \ \lim_{\epsilon \to 0} \beta(\epsilon) = 0.$$



proof).



Let $\tau^{(0)} := T(x,R)$

$\tau^{(j)} :=$ the j-th excursion from $\partial B(x,R)$ to itself via $\partial B(x,r)$.

Observe that, for any $N \in \mathbb{N}$, and $\alpha > 0$,

$$\mathbb{P}^{x_0}\big(T(x,\epsilon) \geq \alpha\big)$$
$$\leq \mathbb{P}^{x_0}\Big(T(x,\epsilon) \geq \sum_{j=0}^{N} \tau^{(j)}\Big) + \mathbb{P}^{x_0}\Big(\sum_{j=0}^{N} \tau^{(j)} \geq \alpha\Big).$$

Use previous lemmas to bound the RHS.

Number of boxes with side ϵ to cover E

The remaining lemmas discuss "fractal dimensions". For example,

$$\limsup_{\epsilon \to 0} \frac{\log N_E(\epsilon)}{\log(1/\epsilon)}$$

Lemma 6. Let $E \subseteq M$.

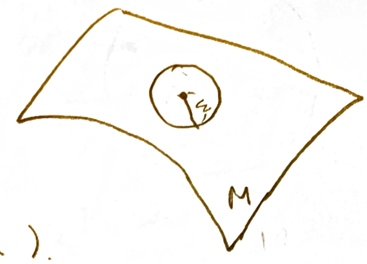
$$\limsup_{\epsilon \to 0} \frac{C_\epsilon(E)}{\epsilon^{2-d} \log\left(\frac{1}{\epsilon}\right)} = \overline{\dim}_M(E) \cdot K_M.$$

Computed in $\boxed{N_E(1/n) = c \cdot n^d}$.

(This is a partial claim of the main theorem).

(11).

Heuristic for Theorem 2 $(d=3)$



Remark. $\lambda_1 :=$ the smallest positive eigenvalue of $-\Delta_M$ on $C^2(M)$

$(\exists\, u \in C^\infty(M),\ -\Delta_M u = \lambda_1 u)$.

$\lambda_1^\varepsilon :=$ the smallest positive eigenvalue of $-\Delta_M$ on $C^2(M \setminus B(x,\varepsilon))$ with the Dirichlet condition on $\partial B(x,\varepsilon)$.

$$\left(\exists\, u,\ \begin{cases} -\Delta_M u = \lambda_1^\varepsilon u \\ u = g \text{ on } \partial B(x,\varepsilon) \end{cases} \right)$$

As $\varepsilon \searrow 0$, it is known that

$$T(x,\varepsilon) \sim \frac{1}{|\lambda_1 - \lambda_1^\varepsilon|} \sim \kappa_M \cdot \varepsilon^{2-d}.$$

Consider ε-cover $A = \{B(x_i,\varepsilon)\}_{i\in I}$, where $|I| = O(\varepsilon^{-d})$.

$$C_\varepsilon(M) \sim \max_{i\in I} T(x_i,\varepsilon)$$

$$\sim \left(d \cdot \log \frac{1}{\varepsilon}\right) \kappa_M \cdot \varepsilon^{2-d}.$$

Random Groups: Yann Ollivier. "A January 2005 invitation to random groups".

What does a "typical" group look like?

Models
- few-relator model: fixed # gen & bdd length relators
- few relator model w/ various lengths: different lengths rel.
- (Gromov's) density model: $0 \le d \le 1$ controls the # of relations

Density model: F_m = free group on $m \geq 2$ generators $a_1, \ldots, a_m$

$l \in \mathbb{Z}_{\geq 1}$. $S_l := \{\text{reduced words of length } l\} \subset F_m$.

Let $0 \le d \le 1$. A random group is $\langle a_1, \ldots, a_m \mid R \rangle$, R a $(2m-1)^{dl}$-tuple of elements of S_l

(We call R a random set of relators at density λ at length l, R is picked from S_l unif. & indep.)

Terminology: *sth* (insert property) occurs w/ overwhelming prob. at density d if its prob. of occurrence tends to 1 as $l \to \infty$ for fixed d.

Phase Transition Thm: Let G be a random group at density d.

1) • If $d < \frac{1}{2}$, w/ overwhelming prob. G is infinite, hyperbolic, torsion-free of geometric dimension 2.

2) • If $d > \frac{1}{2}$, w/ overwhelming prob. G is either $\{e\}$ or $\mathbb{Z}/2$.

Proof (2) $|S_l| = N$, $|R| = p$.

$$\boxed{1 - \frac{i}{N} \sim e^{-\frac{i}{N}}}$$

$$\Pr\left(\text{having repeated relators}\right) = 1 - \Pr\left(\text{having } m \text{ different relators}\right)$$

$$= 1 - 1 \cdot \left(1 - \frac{1}{N}\right)\left(1 - \frac{2}{N}\right)\cdots\left(1 - \frac{p-1}{N}\right)$$

$$= 1 - \prod_{i=1}^{p-1}\left(1 - \frac{i}{N}\right) \sim 1 - e^{\sum^{p-1} - \frac{i}{N}}$$

$$= 1 - e^{-\frac{(p-1)p}{2N}}.$$

For $\frac{1}{2} > \delta > 0$, $p = N^{\frac{1}{2} + \delta}$

$$= 1 - e^{-\frac{N^{2\delta}}{2} + \frac{1}{2}\frac{1}{N^{\frac{1}{2}-\delta}}} \xrightarrow{N \to \infty} 1.$$

Then very probably, we're taking the same relator twice into R.

Then, very probably, we're choosing relators $r_1, r_2 \in R$ that differ at just one position: $r_1 = w a_i^{\pm 1}$, $r_2 = w a_j^{\pm 1}$ w/ $|w| = l-1$.

$\Rightarrow$ In $G = \langle a_1, \dots, a_m \mid R \rangle$, $a_i^{\pm 1} = a_j^{\pm 1}$. Since m is fixed, this will eventually occur $\forall\, 1 \le i \le j \le m$ $\Rightarrow$ $G = \langle a \mid a^2 \rangle$ or $G = \langle a \mid a \rangle$. $\square$

Hyperbolic spaces: $X = $ (geodesic) metric space. X is δ-hyperbolic if any geodesic Δ satisfies the δ-slim condition.

$G = \langle \underset{\text{finite.}}{S} \mid R \rangle$ $\quad \overset{\Gamma}{\text{Cay}(G)}$ = Cayley graph, $V(\Gamma) = G$, $E(\Gamma) = \{ g \xrightarrow{s \in S} g \cdot s \}$

Def. G is hyperbolic if $\text{Cay}(G)$ is δ-hyperbolic for some $\delta > 0$.

Examples: Free groups. Cayley Graphs are trees, which are 0-hyp.

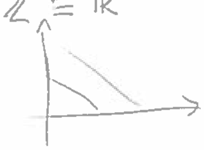
In particular, $\mathbb{Z}$.

- $\pi_1(S_g)$ $g \ge 1$ $\quad S_g$ = quotient of $\mathbb{H}^2$
  - S_g = closed surface genus g

Svarc - Milnor lemma

$G \curvearrowright Y$ on length space, properly discont. & cpt quotient $\Rightarrow \text{Cay}(G) \underset{QI}{\cong} Y$.

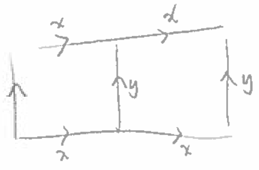
Non-examples. $\mathbb{Z}^2 \overset{QI}{\cong} \mathbb{R}^2$



$\pi_1(S^3 \setminus K)$ (K = knot) w/ $\mathbb{Z}^2$.

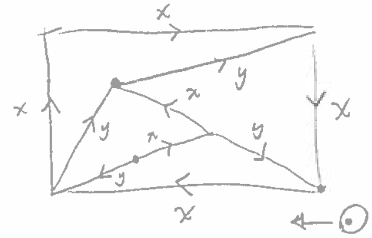
Van Kampen diagrams: A van Kampen diagram for $w \in F(S)$ over $G = \langle S \mid R \rangle$ & $w = 1$ in G is a cpt contractible subset D_w of $\mathbb{R}^2$ w/ the structure of a finite CW-complex & w/ a label of edges so $\partial C \in R$ (C a 2-cell).

$G = \langle x, y \mid xyx^{-1}y^{-1} \rangle$



$x^2 y x^{-2} y^{-1}$

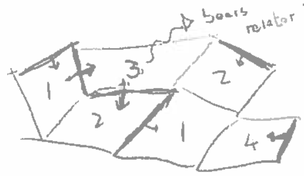
$\langle x, y \mid x^2 y^{-2}, y^{-1} x y x \rangle$



$x^4 = 1$

. A group G is hyperbolic iff $\exists\, \alpha$ s.t. $|\partial D| \geq \alpha |D|$ for any reduced van Kampen diagram for G.

<u>Abstract (van Kampen) diagram:</u> • Diagram you get after forgetting the actual relators associated to the faces but remember which faces bear same relator as each other.

D


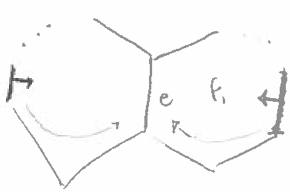
- $1 \leq n \leq |D|$, $n = \#$ distinct relators
- for each face a distinguished edge & an orientation.

An n-tuple $(w_1, \dots, w_n)$ of cyclically reduced words fulfills D if putting w_i on the faces of D bearing relator i turns D into a van Kampen diagram.

<u>Prop.</u> R = random set of relators at density d & at length l. Let D be a (reduced) abstract diagram. Let $\varepsilon > 0$ small. Then either
$$|\partial D| \geq l(1-2d-2\varepsilon)|D| \quad \text{or} \quad \Pr\left(\begin{array}{c}\exists \text{ a tuple of} \\ \text{relators in } R \\ \text{fulfilling } D\end{array}\right) \leq \frac{1}{(2m-1)^{\varepsilon l}}.$$

<u>Proof.</u> $m_i = \#$ times relator i appears in D. (Reorder so $m_1 \geq m_2 \geq \dots$)

$1 \leq i_1, i_2 \leq n$, $1 \leq k_1, k_2 \leq l$, $(i_1, k_1) > (i_2, k_2)$ means $i_1 > i_2$ or $i_1 = i_2$ & $k_1 > k_2$.



let e be an edge adjacent to two faces f_1 & f_2, so

$e = k_1$-th edge of f_1 $=$ k_2-th edge of f_2

f_1 relator i_1
f_2 relator i_2.

($(i_1, k_1) = (i_2, k_2)$ $\Rightarrow$ D is not reduced)

Then we say e belongs to f_1 if $(i_1, k_1) > (i_2, k_2)$ (or belongs to f_2 ...)

(e belongs to the second face it encounters)

$\delta(f) = \#$ edges belonging to f, $\quad \delta_i = \max\limits_{f \text{ face bearing } i} \delta(f) = \left(\begin{array}{l}\text{max. \# of} \\ \text{edges that a} \\ \text{face with relator} \\ i \text{ has}\end{array}\right)$

$$|\partial D| \geq l|D| - 2\underbrace{\sum_{f \text{ face } D} \delta(f)}_{\sum\sum \delta(f) \leq \sum\sum \delta_i} \geq l|D| - 2\overbrace{\sum_{1 \leq i \leq n} m_i \delta_i}$$

Lemma: $1 \le i \le n$. $p_i = $ prob. i randomly chosen (cycl. red.) words $w_1, \dots, w_i$ partially fulfill D ($p_0 = 1$) $\Rightarrow$ $\frac{p_i}{p_{i-1}} \le \frac{1}{(2m-1)^{\delta_i}}$ $\left(\Rightarrow \delta_i \le \log p_{i-1} - \log p_i\right)$ (I)

Proof. Sup. we have $i-1$ words $w_1, \dots, w_{i-1}$ partially fulfilling D. Let f be a face bearing relator i & realizing maximum δ_i. If the k-th edge of f belongs to f, there is another face that has that edge, so we have that prob. of being correct is $\frac{1}{2m-1}$. //

$P_i = $ Prob. $\exists$ a i-tuple of words partially fulfilling D in R.

$$P_i \le (\#R)^i p_i = (2m-1)^{id\ell} p_i \quad \left(\Rightarrow \log P_i - \log p_i \le id\ell\right) \text{(II)}$$

$$
\begin{aligned}
|\partial D| &\ge \ell |D| - 2 \sum m_i \delta_i \quad \text{(I)}\\
&\ge \ell |D| + 2 \sum m_i (\log p_i - \log p_{i-1}) \quad \text{re-index, } m_{n+1} = 0\\
&= \ell |D| + 2 \sum (m_i - m_{i+1}) \log p_i \quad \text{(II)}\\
&\ge \ell |D| + 2 \sum (m_i - m_{i+1}) (\log P_i - id\ell)\\
&= \ell |D| - 2d\ell |D| + \sum (m_i - m_{i+1}) \log P_i\\
&= (1-2d)\ell |D| + \sum (m_i - m_{i+1}) \log P_i\\
&\ge \ell |D| (1-2d) + \log P \cdot m_1\\
&\ge |D| \left(\ell(1-2d) + 2 \log P\right)
\end{aligned}
$$

$\sum (m_i - m_{i+1}) id\ell = d\ell \sum m_i = d\ell |D|$ look at partial sums

Set $P = \min_i P_i$

$\log P < 0$ & $m_1 \le |D|$.

D is fulfillable $\Leftrightarrow$ D is partially fulfillable for all $i \le n$.

$$
\begin{aligned}
\Pr(D \text{ fulfillable by relators of } R) &\le P\\
&\le (2m-1)^{\frac{1}{2}\left(|\partial D|/|D| - \ell \underbrace{(1-2d)}_{>0}\right)} \xrightarrow{\ell \to \infty} 0.
\end{aligned}
$$

Duchin, Jankiewicz, Kilmer, Lelievre, MacKay, Sánchez.

"A sharper threshold for random groups at density one-half."

Generalized density $d = \lim_{\ell \to \infty} \frac{1}{\ell} \log(|R|)$

$$|R| = (2m-1)^{\ell\left(\frac{1}{2} - f(\ell)\right)}, \quad f(\ell) = o(1)$$

Thm: $\begin{cases} G \text{ infinite hyperbolic for } f(\ell) \geq 10^5 \log^{1/3}(\ell) / \ell^{1/3}. & \left(f(\ell) \to 0 \text{ slowly} \right) \\ G \text{ is trivial for } f(\ell) \leq \log(\ell)/4\ell - \log\log(\ell)/\ell. & \left(f(\ell) \to 0 \text{ fast} \right) \end{cases}$

[Sho91a] Short et al.: Notes on word hyperbolic groups.

[Gro93] Gromov.: Asymptotic invariants of infinite groups.