1 Basic Motivations

In recent years, Optimal Mass Transport has developed as an important theory in mathematics, mainly because of it's widespread of applications. What started from a physical problem (Gaspard Monge attempting to move sand) moved into economics and logistics (Kantorovich) and was after related to physical problems such as the flow of a gas (McCann). In the last decade, researchers have noticed this theory can be used to explore General relativity. The goal of this seminar is to explain how and why we believe optimal transport will help understand and give new insights into this part of mathematical physics.

1.1 Intuition Obligation: Framework without optimal transport

Consider a manifold M with a metric g. In the context of Riemannian geometry, g is positive-definite. Meaning that, for any point p in M for tangent vector $X \in T_pM$ one has $g(X, X) \geq 0$, pseudo-Riemannian metrics are linear and symmetric maps that are not necessarily positive-definite. The signature of a metric is the set of values of evaluations of an orthonormal basis of T_pM.
A Lorentzian metric in 4 dimensions is one with signature $(-1, 1, 1, 1)$.
General relativity uses Lorentzian metrics to express the physical idea that time is not independent of space coordinates, in the sense that 'space' coordinates and time should be affected in a correlated manner in the presence of massive objects. This interplay between space and time is expressed through a Lorentzian metric. In the appropriate setting, a spacetime (M, g) is thought of as 'smooth' and 'nice' Lorentzian manifold in 4 dimensions, Einstein's equation is posed:

$$\boxed{R_{\mu\nu} + \frac{1}{2} \cdot Rg_{\mu,\nu} = \frac{1}{4\pi}T_{\mu,\nu}}$$

- $R_{\mu,\nu}$ denotes the Ricci curvature tensor
- $g_{\mu,\nu}$ denotes the metric tensor
- $T_{\mu,\nu}$ is called the Stress-Energy tensor and is derived from the matter model assumed in the spacetime.

Important Observation: To understand EE, there is a lot of geometry assumed on the reader. Further, geometry attemps to pose **coordinate-independent** equations, equations that are true for all coordinate systems but most of the insight of geometry comes from equations expressed in local coordinates.

Question 1. *Can we develop a synthethic approach to general relativity?*

Question 2. *Can we extract (enough) information of a manifold by just looking at curvature and dimension?*

Suppose both answers are yes, let's think of what that would imply. We would have more tools to understand spacetimes and maybe obtain conclusions where differential geometry is limited. Maybe spacetime is not as smooth as differential geometry needs it to be.

2 The optimal transport problem and 3 relevant definitions

Throughout this notes metric spaces are **complete and separable**

Definition 1. *(Monge's problem)*
The original problem of Monge is related to moving sand.
Given two subsets U, V of same volume $\mathbb{R}^3$ find a volume-preserving function $T : U \to V$, that minimizes the expected distance

$$\int_U |x - T(x)| dx$$

Monge's problem has been generalized and adequated to modern notation:
Let $(X, S_X, \mu), (Y, S_Y, \mu)$ be two measure spaces. Monge's problem consists in finding

$$\inf_{T_{\#}\mu=\nu} \int_X c(x, T(x)) d\mu(x)$$

Where the notation $T_{\#}\mu = \nu$ means that for every $A \in S_Y$, $\mu(T^{-1}(A)) = \nu(A)$. This condition means that T is measure-preserving.

Note that we have generalized to any cost function. It is clear that the problem won't have solutions if c is not asked to satisfy some conditions. Throughout this talk we will mainly focus on the cost being a power of **some kind** of distance. Conditions on c that ensure nice properties of the problem are lower semi-continuity, the twist condition and non-degeneracy of the Hessian.
Observe that because $\mu(X) = \nu(Y)$, there is no loss of generality by assuming μ and ν are probability measures.

Difficulty 1. *μ and ν can be supported in very strange sets. The cost function would have to be very non-linear. We know little about theoretical non-smooth non-linear optimization. This **high no linearity** of T made Monge's problem too complicated to solve. One can think of this difficulty as having TOO many things to choose from.*

Kantorovich solved this difficulty by changing the problem to a linear problem. (!!!). An amazing achievement that enhanced optimal transport theory.

Definition 2. *(Kantorovich's Problem)*
Let $(X, S_X, \mu), (Y, S_Y, \mu)$ be two probability spaces. The objective is to find a minimizer of the problem

$$\inf_{\pi \in \Gamma(\mu,\nu)} \int c(x,y) d\pi(x,y)$$

where $\Gamma(\mu,\nu)$ is the set of probability measures in $X \times Y$ having μ and ν as it's marginals. I.e. $\pi(A \times Y) = \mu(A), \pi(X \times B) = \nu(B)$

We will focus on the study of Kantorovich's problem as it can be shown to be a relaxation of the Monge's problem. If one has a solution to Monge's problem, T, set $\pi_T = (id, T)_{\#}\mu$ to be the measure in $P(X \times Y)$ completely concentrated on the graph of T, π_T is optimal in Kantorovich sense.
Recall that a function F is lower semicontinuous if $F(x) \leq \liminf_{y \to x} F(y)$

Theorem 1. *(Kantorovich's Duality) Let $(X, S_X, \mu), (Y, S_Y, \mu)$ be two measure spaces where X and Y are polish spaces, **if** c is lower semi-continuous and non-negative then*

$$\inf_{\pi \in \Gamma(\mu,\nu)} \int c(x,y) d\pi(x,y) = \sup_{(\phi,\psi) \in \Phi_c} \left\{ \int \phi(x) d\mu(x) + \int \psi d\nu(y) \right\}$$

Where $\Phi_c = \{(\phi, \psi) \in L^1(\mu) \times L^1(\nu) : \phi(x) + \psi(y) \leq c(x,y)\}$

Proof. The inequality $(\geq)$ is direct because if $(\phi, \psi) \in \Phi_c$ then

$$\int \phi(x) d\mu(x) + \int \psi(y) d\nu(y) = \int_{X \times Y} (\phi(x) + \psi(y)) d\pi \leq \int_{X \times Y} c(x,y) d\pi$$

The reverse inequality is more complicated. $(\leq)$. We can assume that X, Y are compact and proceed by exhaustion, let us proof only the compact case.
Define $T : \{\phi \in C(X \times Y) : \phi(x,y) = f(x) + g(y)\} \to \mathbb{R}$ given by

$$T(\phi) = \int_X f(x) d\mu(x) + \int_Y g(y) d\nu(y)$$

By Hahn-Banach, T can be extended to a functional $\tilde{T} : C(X \times Y) \to \mathbb{R}$. One can check that $\tilde{T}$ is continuous and $X \times Y$ is compact, hence by Riesz-Kakutani-Markov there exists a non-negative Radon measure π such that

$$\tilde{T}(\phi) = \int_{X \times Y} \phi(x,y) d\pi(x,y)$$

But if $\phi(x,y) = f(x) + g(y)$, then

$$\tilde{T}(\phi) = T(\phi) = \int_X f(x) d\mu(x) + \int_Y g(y) d\nu(y)$$

Because HB preserves the norm, the result is obtained by noting that the marginals of π are μ and ν.

Remark. *One can show, by showing compactness of $\Gamma(\mu,\nu)$ (which follows from Prokhorov's theorem (that states that compactness in the set of probability measures is characterized by tightness), from which the infimum is always attained.*

Theorem 2. *(Brenier's and McCann's)*

1. *Brenier's. In $(X, S_X, \mu) = (\mathbb{R}^n, \mathcal{B}, \mu), (Y, S_y, \nu) = (\mathbb{R}^n, \mathcal{B}, \nu)$, with $c(x,y) = \frac{1}{2}|x-y|^2$, if $\mu << \lambda^n$, and*
$$\int_X |x|^2 d\mu(x) < \infty, \int_Y |y|^2 d\nu < \infty$$
there exists a solution to Monge's problem and it is of the form:
$$T(x) = id(x) - \nabla\phi(x)$$
For some convex function ϕ.

2. *McCann's. Let (M,g) be a smooth, connected, compact Riemannian manifold, if $(X, S_X, \mu) = (M, \mathcal{B}, \mu), (Y, S_y, \nu) = (M, \mathcal{B}, \nu)$, where μ and ν have compact support, with $c(x,y) = \frac{1}{2}d(x,y)^2$, if $\mu << vol$, if $\mu << \lambda^n$, and*
$$\int_X |x|^2 d\mu(x) < \infty, \int_Y |y|^2 d\nu < \infty$$
there exists a solution to Monge's problem and it is of the form:
$$T(x) = \exp_x(\nabla\phi(x))$$
for some $d^2/2$-concave function ϕ.

Proof. (Brenier's)
From the identity $|x-y|^2 = |x|^2 + |y|^2 - 2xy$ we observe that for $\pi \in P(X \times Y)$

$$\int c(x,y)d\pi = \frac{1}{2}\left(\int_{X\times Y} |x|^2 + |y|^2 - 2xy d\pi(x,y)\right) = \frac{1}{2}\int_X |x|^2 d\mu(x) + \frac{1}{2}\int_Y |y|^2 d\nu(y) - \int_{X\times Y} xy d\pi(x,y)$$

Hence, as the first two terms are fixed (because μ and ν are given, the infimum in Kantorovich problem reduces to studying

$$\inf_{\pi\in\Gamma(\mu,\nu)} \int_{X\times Y} c(x,y)d\pi(x,y) = a - \sup_{\pi\in\Gamma(\mu,\nu)} \int xy d\pi(x,y)$$

where a is a constant.
Observe know that for every $(\phi, \psi) \in \Phi_c$,

$$\phi(x) + \psi(y) \leq \frac{1}{2}|x|^2 + \frac{1}{2}|y|^2 - xy$$

By defining $\tilde{\phi}(x) = \frac{1}{2}|x|^2 - \phi(x)$, $\tilde{\psi}(y) = \frac{1}{2}|y|^2 - \psi(y)$ we have

$$x \cdot y \leq \tilde{\phi}(x) + \tilde{\psi}(y)$$

Substituting we get

$$\inf_{\pi \in \Gamma(\mu,\nu)} \int_{X\times Y} c(x,y) d\pi(x,y) = a - \sup_{\pi \in \Gamma(\mu,\nu)} \int_{X\times Y} xy d\pi(x,y) = a - \inf \left\{ \int_X \tilde{\phi}(x) d\mu(x) + \tilde{\psi}(y) d\nu(y) \right\}$$

$$\sup_{\pi \in \Gamma(\mu,\nu)} \int_{X\times Y} xy d\pi(x,y) = \inf \left\{ \int_X \tilde{\phi}(x) d\mu(x) + \tilde{\psi}(y) d\nu(y) \right\}$$

where the last infimum is taken over all $(\tilde{\psi}, \tilde{\phi}) \in L^1(\mu) \times L^1(\nu)$ satisfying $x \cdot y \leq \tilde{\phi}(x) + \tilde{\psi}(y)$.

By the remark made after the proof of Kantorovich's Duality, we know there exists an optimal transference plan π, one can show, (see [1]) proposition 2.9 that $(\tilde{\psi}, \tilde{\phi})$ can be replaced by $(\tilde{\phi}, \tilde{\phi}^*)$ where $\tilde{\phi}^*$ denotes the Legendre transform:

$$\tilde{\phi}^*(y) = \inf_{x \in X} \{\tilde{\phi}(x) - x \cdot y\}$$

Assuming these optimizers exist,

$$\int_{X\times Y} x \cdot y d\pi = \int_X \tilde{\phi}(x) d\mu(x) + \int_Y \tilde{\phi}^*(y) d\nu(y) = \int_{X\times Y} \tilde{\phi}(x) + \int_Y \tilde{\phi}^*(y) d\pi(x,y)$$

Then,

$$\int_{X\times Y} \tilde{\phi}(x) + \int_Y \tilde{\phi}^*(y) - x \cdot y d\pi(x,y) = 0$$

But by definition of $\tilde{\phi}^*$ the quantity is always non-negative, so π-a.e. $\tilde{\phi}(x) + \phi^*(y) - x \cdot y = 0$
But this condition is well-known to be a characterization of the subdifferential $\partial \tilde{\phi}^+(y) = \{x : \forall z \ \tilde{\phi}^*(z) \geq +\tilde{\phi}^*(y) \geq \langle x, z - y \rangle\}$, vectors that lie above the tangent.
A little more argumentation gives the result but this is good enough for now.

Proof. (McCann's)
As before (it can be proven) suppose that a pair (ϕ, ϕ^c) realizes the supremum in Kantorovich duality.

Lemma 1. *(The supremum is unchanged if we consider Lipschitz maps)*

So by the Lemma ϕ is Lipschitz, so $\nabla\phi : M \to TM$ and $\exp_x(\nabla\phi)$ are measurable.
We aim to show that $\exp_x(\nabla\phi)_{\#}\mu = \nu$, so by duality we look at the integrable of continuous functions.
For $x, y \in M$ and $0 < \epsilon < 1$, and $f \in C(M)$ define

$$\phi_\epsilon = \phi + \epsilon f, \qquad \psi_\epsilon = \phi_\epsilon^*$$

Let x be a point where ϕ^* is differentiable.
All infimums in Legendre transforms are attained because ϕ, h are continuous and M is compact, though we haven't shown ϕ is continuous.

Lemma 2. *(At $\epsilon = 0$, $y = \exp_x(\nabla\phi)$ attains this minimum.) Suppose that ϕ is differentiable at x_0 we aim to show that $c(x_0, y) = \phi(x_0) + \phi^*(y)$ if and only if $y = \exp_x(\nabla\phi(x))$.*
Suppose that x is a point such that the minimum is attained, that is $c(x, y) = \phi(x) + \phi^*(y)$, let also $F(z) = c(z, y)$, let also $z = \exp_x(v)$
From the definition of the Legendre transform we know that

$$F(z) - \phi(z) - \phi^*(y) \geq 0 = F(x) - \phi(x) + \phi^*(y)$$

whence

$$F(z) \geq F(x) - \phi(x) + \phi(z)$$

No using the taylor expansion of ϕ and $\exp_x$, by the chain rule,

$$F(z) \geq F(x) - \phi(x) + \phi(x) + \langle \nabla\phi(x), v \rangle + O(|v|)$$

But this is the definition of F having subgradient $\nabla\phi$ at x. $(\nabla\phi(x) \in \partial\phi(x))$ To show that F is also superdifferentiable one can use ***Hopf-Rinow*** *theorem to define $\exp_x(\sigma(1))$. By compactness of M, one can see that the Legendre transform is achieved by some y, which by what we just shown is of the form $\exp_x(\nabla\phi(x))$.*

Now we use this lemma to prove McCann's theorem, for small enough ϵ, $y_\epsilon = \exp_x(\nabla\phi(x)) + o(\epsilon)$. Hence, by optimality,

$$c(x, \exp_x(\nabla\phi(x))) - \phi(\exp_x(\nabla\phi(x))) - \epsilon h(y_\epsilon) \leq \psi_\epsilon(x) \leq c(x, y) - \phi(y) - \epsilon h(y)$$

So after obtaining a uniform bound on the rate of convergence,

$$0 = \lim_{\epsilon \to 0} \int \frac{\psi_\epsilon(x) - \psi_0(x)}{\epsilon} d\mu(x) + \int h(x) d\nu(x) = -\int h(\exp_x(\nabla\phi(x)) d\mu(x) + \int h(x) d\nu(x)$$

Which concludes that $\exp_x(\nabla\phi)_{\#}\mu = \nu$ by duality. We do not duel with uniqueness in this work.

So we have found a solution taking points in position x to position $T(x)$. But that may not be sufficient, we may want to analyze what was the path traveled by each particle. Let us not define nor write the time-dependent version but let us focus on a possible path

Definition 3. *(McCann's interpolation)*

1. *In $\mathbb{R}^n$ given a transport map $T = Id - \nabla\phi$, pushing μ onto ν, define at time $t \in [0, 1]$:*

$$\rho_t = [t \cdot id(x) + (1-t)\nabla\phi(x)]_{\#}\mu$$

2. *In (M, g) for a transport map $T = \exp(\nabla\phi)$, for $t \in [0, 1]$ define*

$$\rho_t = \exp_x(t\nabla\phi(x))_{\#}\mu$$

Definition 4. *(Wasserstein's metric and Wasserstein's geodesics) For a metric space X we define*

$$P_p(X) = \{\mu \in P(X) : \int d(x, x_0)^p d\mu(x) < \infty \text{ (for some } x_0)\}$$

We define the Wasserstein distance in $P_p(X)$ to be the optimal transport cost in Kantorovich sense between μ and ν:

$$W_p(\mu, \nu) = \left(\inf_{\pi \in \Gamma(\mu,\nu)} \left\{ \int d(x, y)^p d\pi(x, y) \right\} \right)^{1/p}$$

Example 1. *(Transport between Dirac deltas)*

$$W_p(\delta_x, \delta_y) = d(x, y)$$

The previous example is particularly important because it shows how (X, d) is embedded in $(P_p(X), W_p)$.

Remark. *W_p is a distance for every $p \geq 1$. Further, W_1 metrizes weak convergence.*

We say that $\{\rho_t\}_{t\in[0,1]}$ is a Wasserstein p-geodesic if

$$W_p(\rho_t, \rho_s) = |t - s| W_p(\rho_0, \rho_1)$$

The idea is to use Wasserstein geodesics as we would use geodesics in a length space.

Definition 5. *(K-convexity)*
A function $f : \mathbb{R} \to \mathbb{R}$ is said to be K-convex if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{K}{2}|x-y|^2$$

Remark. *Note that if f is twice differentiable, K convexity is nothing but $f'' \leq -K$*

A functional $F : P_p \to \mathbb{R}$ is said to be K convex if for any $\rho_0, \rho_1 \in P_p(X)$ there exists a Wasserstein geodesic ρ_t satisfying

$$F(\rho_t) \leq tF(\rho_1) + (1-t)F(\rho_0) - \frac{K}{2}W(\rho_0, \rho_1)^2$$

Definition 6. *(Shannon's and Renyi's entropies)*
Let $\mu, \nu \in P_p(X)$, decompose $\mu = \rho\nu + \mu_{sing}$ by Hahn-decomposition.
We define Shannon's entropy of μ relative to ν as

$$Ent(\mu|\nu) = \int \rho \log \rho d\nu$$

We define Renyi's N-entropy of μ relative to ν as

$$Ent_R(\mu|\nu) = -\int \rho^{-1/N} d\nu$$

Remark. *Note that for any given metric space (X, d), Wasserstein geodesics in $P(X, W)$ still make sense.*

3 Curvature in metric measure spaces

Our objective now is try to understand what curvature bounds mean in contexts where there are no notions of derivatives nor coordinates.

Definition 7. *(Sturm's definition of Ricci curvature bounds)*
In a metric measured space (X, d, μ) we say that (X, d, μ) has curvature $\geq K$, if Shannon's entropy relative to μ is K-convex, that is for every $\rho_0, \rho_1 \in P_p(X)$ there exists a Wasserstein geodesic ρ_t for which

$$Ent(\rho_t|\mu) \leq tEnt(\rho_1|\mu) + (1-t)Ent(\rho_0|\mu) - \frac{K}{2}W_2(\rho_0, \rho_1)^2$$

We define the curvature of our metric measure space by taking the supremum:

$$Curv(X, d_x, \mu) = \sup\{K : (X, d_x, \mu) \text{ has curvature } \geq K\}$$

Definition 8. *Curvature-dimension condition*
Given N and K, define,

$$\tau_{K,N}^{(t)}(\theta)\begin{cases}\infty & K\theta^2 \geq \pi^2(N-1)\\ t^{1/N}\left(\frac{\sin(t\theta\sqrt{K/(N-1)})}{\sin(\theta\sqrt{K/(N-1)})}\right)^{1-1/N} & 0 < K\theta^2 < (N-1)\pi^2\\ t & K\theta^2 = 0, K\theta^2 < 0, N = 1\\ t^{1/N}\left(\frac{\sin(t\theta\sqrt{-K/(N-1)})}{\sin(\theta\sqrt{-K/(N-1)})}\right)^{1-1/N} & K\theta^2 < 0, N > 1\end{cases}$$

We say that (X, d, μ) satisfies $CD(K, N)$ if for every Wasserstein geodesic ρ_t joining $d\nu_0(x) = \rho_0(x)d\mu(x)$ and $d\nu_1(x) = \rho_1(x)d\mu(x)$, there exists a coupling q such that the Renyi N-entropy is K-convex in the following, modified sense, for all $N' \geq N$

$$Ent_R(\rho_t, \mu) \leq -\int_{X\times X} \tau_{K,N'}^{(1-t)}\rho_0^{-1/N'}(x_0) + \tau_{K,N}^{t} d(x_1)\rho_1^{-1/N'}(x_1)dq(x_0, x_1)$$

3.1 Properties

Example 2. *(Gaussian line)*
Let $(X, d, \mu) = (\mathbb{R}, |x-y|, d\mu(x) = \exp(-Kx^2/2))$ then the space has curvature $\geq K$.

Theorem 3. *(Curvature is in fact a generalization)*
Consider $(M, d, e^{-V}dvol)$ where M is a Riemannian manifold of dimension n, d is Riemannian distance then

$$Curv((M, d, e^{-V}dvol) \geq K \textit{ if and only if } Ricc(\xi,\xi) + Hess(V)(\xi,\xi) \geq K|\xi|^2$$

Proof. Suppose that $\nu_0 = \rho_0 vol$, $\nu_1 = \rho_1 vol$, assume that ρ is continuous, so by McCann's theorem there exist a Wasserstein geodesic

$$\nu_t = \exp_x(-t\nabla\phi(x))_{\#}v_0$$

for some $d^2/2-$convex function ϕ.
For notational simplicity, define $F_t(x) = \exp_x(-t\nabla\phi(x))$, so that $\nu_t = F_t(x)\#\mu$

$$Ent(\nu_t|e^{-V}vol) = \int e^{V(x)}\rho_t(x)\log\left[e^{V(x)}\rho_t(x))\right]e^{-V(x)}dvol(x) = \int \rho_t(x)\log[e^{V(x)}\rho_t(x)]dvol(x)$$

Further, by change of variable formula: $\rho_t(x) = \rho_0(F_t^{-1}(x))|dF_t^{-1}|$
Hence,

$$Ent(\nu_t|e^{-V}vol) = \int \rho_t \log\rho_t dvol + \int \rho_t \log(e^{V(x)})dvol$$

$$= \int (\rho_0(F_t^{-1}(x))|dF_t^{-1}|)\log(\rho_0(F_t^{-1}(x))|dF_t^{-1}|)dvol(x) + \int V(F_t)\rho_0 dvol$$

$$= \int \rho_0 \log \rho_0 dvol + \int \rho_0(F_t^{-1}(x))\log|dF_t^{-1}|dvol + \int V(F_t)\rho_0 dvol$$

$$= \int \rho_0 \log \rho_0 dvol + \int \rho_0(x)\log|dF_t^{-1}(F_t(x))|dvol + \int V(F_t)\rho_0 dvol$$

Now by the theorem for the derivative of the inverse function, dF_t is a linear isomorphism, hence

$$Ent(v_t|e^{-V}vol) = \int \rho_0 \log(\rho_0)dvol(x) - \int \rho_0 \cdot y_t dvol(x) + \int V(F_t)\rho_0 dvol \tag{☎}$$

where $y_t(x) = \log(|dF_t(x)|)$. Now the key geometric insight is that y_t satisfies the inequality:

$$\frac{d^2 y_t}{dt^2} \leq -\frac{1}{n}\left(\frac{dy_t}{dt}\right)^2 - Ricc\left(\frac{dF}{dt}, \frac{dF}{dt}\right)$$

Now we differentiate (☎), twice with respect to t, by the mentioned inequality,

$$\frac{\partial^2}{\partial t^2}Ent(v_t|e^{-V}vol) \geq \int \frac{1}{n}\left(\frac{dy_t}{dt}\right)^2 \rho_0 dvol + \int Ricc\left(\frac{dF}{dt}, \frac{dF}{dt}\right)\rho_0 dvol + Hess_V\left(\frac{dF}{dt}, \frac{dF}{dt}\right)\rho_0 dvol$$

Therefore,

$$\frac{\partial^2}{\partial t^2}Ent(v_t|e^{-V}vol) \geq \int \left[Ricc\left(\frac{dF}{dt}, \frac{dF}{dt}\right) + Hess_V\left(\frac{dF}{dt}, \frac{dF}{dt}\right)\right]\rho_0 dvol \geq KW_2(\nu_0, \nu_1)^2$$

Theorem 4. *(Curvature dimension is a generalization) Consider $(M, d, e^{-V}dvol)$ where M is a Riemannian manifold, d is Riemannian distance then it satisfies $CD(K, N)$ if $Ricc \geq K$ and $dim(M) \leq N$*

The proof of this theorem is very similar to the one before, one checks the same inequality as before but adds some (relatively) laborious estimates using Jacobi fields. The avid reader is referred to [4] Theorem 1.7. **Note that there is no only if**

Example 3. *(Riemannian Manifold of constant sectional curvature)*
If a Riemannian manifold (M, g) of dimension n has constant sectional curvature K then

$$Curv(M, d, vol) = K(n-1)$$

Remark. *(The power of CD conditions)*
There are many ways to define Curvature-dimension conditions: Riemannian CDK, entropic CDK, BE CDK, local CDK, CDK, lax CDK but they all agree on something. They are necessary (and sometimes sufficient) to imply global properties on a space. The usual functional inequalities from PDE theory (Sobolev,Harnack,log-sobolev,Poincaré) are implied by curvature conditions.* ***We expect*** *to be able to derive most intrinsic geometric properties from this conditions.*

Definition 9. *(Measure contraction property $MCP(K, N)$)*

Given two numbers $K, N \in \mathbb{R}$,with $N \geq 1$, we say that a metric measure space (X, d_x, μ) satisfies the measure contraction property $MCP(K, N)$ if and only if for each $0 < t < 1$ there exists a Markov kernel P_t from $X \times X$ to X such that for $\mu \times \mu$-a.e. (x, y) and for $P_t(x, y;)$-a.e. z the point z is at intermediate point of x and y, and such that for μ-a.e. x and for every measurable $A \subseteq X$

$$\int_X \xi^{(t)}_{K,N}(d(x, y))P_t(x, y; A)d\mu(y) \leq \mu(A)$$

$$\int_X \xi^{(1-t)}_{K,N}(d(x, y))P_t(x, y; A)d\mu(y) \leq \mu(A)$$

where $\xi^{(t)}_{N,K} = (\tau^{(t)}_{K,N})^N$

Theorem 5. *(Partial Converse) The converse of theorem 4 is true for the MCP. If $(M, g\mu)$ satisfies $MCP(K, N)$ then M has dimension $\leq N$*

No proof, sorry, [4] Corollary 5.5.
This formulation is equivalent to a formulation of the CD(K,N) condition when μ_0 is a Dirac mass. I haven't seen this connection. In general, $MCP(K, N)$ does not imply the curvature is bounded from below by K
MCP is also stable under HG convergence.
I haven't built a lot of intuition for this formulation, the reader is encouraged to talk to me about it or skip it and go to the TIMELIKE Measure Contraction Property.

3.2 Convergence and stability

When we defined the Kantorovich problem, we used a probability measure, π on the product space. This probability is called a transference plan or **coupling** between μ and ν.
A coupling between two metrics in different spaces, $(X, d_x, \mu), (Y, d_y, \nu)$, $X \cap Y = \varnothing$, is a pseudometric $\mathbf{d} : X \cup Y \to \mathbb{R}$ such that $d(x_1, x_2) = \mathbf{d}(x_1, x_2)$ and $d(y_1, y_2) = \mathbf{d}(y_1, y_2)$ for every $(x_1, x_2) \in supp(\mu)$ and $(y_1, y_2) \in supp(\nu)$

Definition 10. *(Hausdorff-Gromov convergence)*
Given two metric measure spaces $(X, d_x, \mu), (Y, d_y, \nu)$ we define the Hausdorff-Gromov distance between them:

$$D((X, d_x, \mu), (Y, d_y, \nu)) = \left(\inf \int \mathbf{d}^2(x, y) d\pi \right)^{1/2}$$

where the infimum is taken among all possible couplings π of μ and ν and all possible couplings $\mathbf{d}$ *of d_x and d_y*

Suppose that we have a converging sequence of metric measure spaces all satisfying a curvature bound, does the limiting metric space satisfy the same curvature bound?
One of the main results of this definition of curvature is it's stability under Hausdorff-Gromov convergence.

Theorem 6. *(Stability)*
Let (X_n, d_n, μ_n) metric measure spaces with uniformly bounded diameter.
If

$$(X_n, d_n, \mu_n) \to (X, d, \mu) \textit{ in } D-metric$$

then

$$\limsup_n Curv(X_n, d_n, \mu_n) \leq Curv(X, d, \mu)$$

Proof. The proof can be found in [4] and it is not presented here as it depends on the notion of lax curvature-dimension conditions and midpoints, which I think requires too many concepts.

Remark. *There is also a stability result for $CD(K, N)$, written similarly.*

4 Main result: Hawking's singularity theorem

4.1 Unfamiliar definitions of General Relativity

Let (M, g) be a smooth semi-Riemannian manifold, that is a manifold M together with a smooth function g, such that at every point $p \in M$ $g : T_pM \times T_pM \to \mathbb{R}$ is symmetric and bilinear:

$$g(X, Y) = g(Y, X)$$

$$g(\alpha X + \beta Y, Z) = \alpha g(X, Z) + \beta g(Y, Z)$$

Notice that different to the Riemannian setting, we do not require g to be non-degenerate. Meaning that there can exist $Z \in T_pM$ such that $g_p(Z, Z) = 0$.
A Lorentzian metric is a semi-Riemannian metric of signature $(-, +, +, ..)$

Definition 11. *1. Timelike, lightlike, spacelike We say that a vector field Z is timelike if $g(Z, Z) < 0$, lightlike if $g(Z, Z) = 0$ and spacelike if $g(Z, Z) > 0$.*

*Assume also that M is **oriented** in the sense that every timelike or lighlike vector can be classified into future oriented or past oriented. In the tangent bundle, define the Lagrangian for the metric at power q as*

$$L(v, p : q) = \begin{cases} \frac{1}{q} \sum_{a=1,b=1}^{n} (g_{a,b} v^a v^b)^{q/2} & v \text{ is future oriented} \\ \infty & \text{otherwise} \end{cases}$$

Define the action of L along a curve

$$A(\gamma) = \int_0^1 L(\gamma'(t), \gamma(t), q) dt$$

Finally take $\ell(x, y, q) = \inf\{A(\gamma, q) : \gamma$ is a continuous curve $, \gamma(0) = x, \gamma(1) = y\}$

Intuition 1. *This lagrangian is just the setup for the principle of least action. $A(\gamma)$ is to be understood as the action required to travel through γ. We add the ∞ part so that we can only travel through future oriented curves. Think of it, it is ∞-costly to go back in time.*

2. $I^+(V)$ Is the set of points achievable from V by future directed curves. $J^+(x) = \{y : \ell(x, y, q) \leq 0\}, J^-(y) = \{x : \ell(x, y, q) \leq 0\}$

Intuition 2. *We want to find the sets that we can actually get there from a specific position in spacetime. We also analyze the possible places we could come from to be at here-now.*

3. Non-branching means that if z lies in the geodesic joining x and y, the unique geodesics from x to z and z to y are the restrictions of the original curve.

Intuition 3. *Non-branching requires that there are no ramifications of geodesics. There are not concatenations of geodesics leading to the same position-time.*

4. The space is called locally causally closed if for every point $x \in X$ there exists a neiborhood N_x for which $\{(x, y) : x \leq y\}$ is closed, where $x \leq y$ means causality. Meaning there exsists a causal (non-spacelike) curve from x to y. The set of causal pairs is denoted $X_\leq$

Intuition 4. *Locally causally closed represents the idea that in a neighborhood of here-now all points that we can travel to, are achievable. In the sense that, locally, one can get to points. These points are not infinitesimally close unachievable.*

5. *We say X is globally hyperbolic if $J^+(x) \cap J^-(y)$ is compact for every $(x,y) \in X_\leq$*
   We say it is K-globally hyperbolic if it is a globally hyperbolic Lorentzian length space satisfying $I^+(x), I^-(x) \neq 0$ for every $x \in X$

   Intuition 5. *Global hyperbolicity is equivalent to the existence of a Cauchy hypersurface. A Cauchy hypersurface can be though of a spacelike part of the spacetime "at the same time". One thinks of it as the universe at an instant of time. Surprisingly, in the Optimal Transport setting we need it to achieve maximums and minimums, hence it is better to work with compact sets rather than Cauchy hypersurfaces.*

6. *V is called achronal if $x \not<< y$ for every $x, y \in V$, where $x << y$ means there exists a timelike curve between x and y.*
   A set V is called (FTC) or future timelike complete, if for every $x \in I^+(V)$, $J^-(x) \cap V$ is pre-compact in V.

   Intuition 6. *Achronal means that there is no way to travel in time from diferent points of the set.*
   FTC, intuitively, is a set on which we can travel.

7. *If V is achronal, modify the time separation function τ as follows:*

$$\tau_V(x) = \begin{cases} \sup\limits_{y \in V} \tau(y,x) & x \in I^+(V) \\ \sup\limits_{y \in V} -\tau(y,x) & x \in I^-(V) \\ 0 & otherwise \end{cases}$$

   Intuition 7. *This modification of the time separation function gives the concept of the maximal proper time (to future or past) from our position.*

8. *A Borel FTC achronal subset has curvature bounded from below by H_0 if $\mathcal{H}_0$ is a non-negative Radon-measure with $(f_0)_\# q \leq \mathcal{H}_0$ such that*

$$\limsup_{t \to 0} \frac{\mu(V_{t,\phi}) - t \int_V \phi d\mathcal{H}_0}{t^2/2} \geq H_0 \int_H \phi^2 d\mathcal{H}_0$$

   where $V_{t,\phi} = \{x \in \mathcal{T}_V : 0 < \tau_V < t\phi(\mathbf{a}(x))\}$ and $\mathbf{a} : \mathcal{T}_V \to V$ is called the initial map projection.

   Intuition 8. *I still don't have it. Sorry. I thought it was related to the fact that one can study $n-1$ dimensional Hausdorff measures as some kind of derivative of the $n$dimensional Lebesgue measure, but I am not sure what the initial map projection is actually doing.*

9.

$$D_{H_0,K,N} = \begin{cases} \frac{\pi}{2}\sqrt{N-1/K} & K>0, N>1, H_0=0 \\ \sqrt{(N-1)/K}\cot^{-1}\left(-H_0/\sqrt{K(N-1)}\right) & K>0, N>1, H_0\neq 0 \\ -\dfrac{N-1}{H_0} & K=0, N>1, H_0<0 \\ \sqrt{(N-1)/-K}\cot^{-1}\left(-H_0/\sqrt{-K(N-1)}\right) & K<0, N>1, H_0\neq 0 \end{cases}$$

Intuition 9. *This look like the distortion coefficients on Jacobi fields, but are closely related to the 'convexity-weights' we used when we defined $CD(K,N)$*

Definition 12. ***TIMELIKE Measure contraction property*** $MCP_p(K,N)$
Let $(X, d, \mu, <, <<, \tau)$ a measured Lorentzian pre-length space.
We say it satisfies a $MCP_p(K,N)$ if for any $\mu_0 \in P_c(X) \cap \{\rho : Ent(\rho|\mu) < \infty\}$ and for any $x_1 \in X$, such that $x << x_1$, μ-a.e. on x, there exists an ℓ_p-geodesic ρ_t, such that $\rho_0 = \mu, \rho_1 = \delta_{x_1}$ such that

$$U_N(\rho_t|\mu) \geq \sigma_{K,N}^{(1-t)}(||\tau(\cdot, x_1)||_{L^2(\rho_0)})U_N(\rho_0|\mu)$$

4.2 Hawking Singularity Theorem

Theorem 7. *(Hawking Singularity Theorem)*
Let $(X, d, \nu, <<, \leq, \tau)$ be a timelike non-branching, locally causally closed, K-globally hyperbolic, Lorentzian geodesic space satisfying $TMCP_P^e(K,N)$for some $p \in (0,1), K \in \mathbb{R}, N \in [1,\infty)$ and assume that the causally-reversed structure satisfies the same conditions.Let $V \subseteq X$ be a Borel achronal FTC subset having forward mean curvature bounded above by H_0 in the sense of the previous definition. If

1. $K > 0, N > 1, H_0 \in \mathbb{R}$
2. $K = 0, N > 1, H_0 < 0$
3. $K < 0, N > 1, H_0 < \sqrt{-K(N-1)}$

then for every $x \in I^+(V)$ it holds $\tau_V \leq D_{H_0,K,N}$ In particular, for every timelike geodesic γ with $\gamma(0) \in V$, the maximal (on the right) domain of definition is contained in $[0, D_{H_0,K,N}]$

Proof not provided, sketched in the talk. With just a drawing, reference [**?**]

5 Brief Summary of Talk

1. Motivation: General Relativity (and semi-Riemanian, Riemannian geometry) can be expressed in a synthetic manner. Just by analyzing the set of probabilities on top.

2. Curvature dimension conditions in Riemannian Geometry determine plenty of geometrical properties.

3. Sturm defined a curvature-dimension condition in metric measure spaces by the convexity of Entropies of measures

4. The definition of the measure contraction property can be extended to Lorentzian setting

5. MCP depends on p but it is the hypothesis needed for Hawking's theorem

References

[1] Topics in Optimal Transportation (Graduate Studies in Mathematics, Vol. 58) , Cedric Villani, Uk edition.

[2] Polar factorization of maps on Riemannian manifolds, R.J. McCann, Geometric and Functional Analysis GAFA volume 11, pages 589–608 , (2001)

[3] Optimal transport in Lorentzian synthetic spaces, synthetic timelike Ricci curvature lower bounds and applications, Fabio Cavalletti, Andrea Mondino, arXiv:2004.08934

[4] On the geometry of metric measure spaces, Karl-Theodor Sturm, Acta Math. Volume 196, Number 1 (2006), 65-131.