



A geometric approach to clustering

Identity of indiscernibles

$$d(x, y) = 0 \quad \Leftrightarrow \quad x = y$$

Non-negativity

$$d(x, y) \geq 0$$

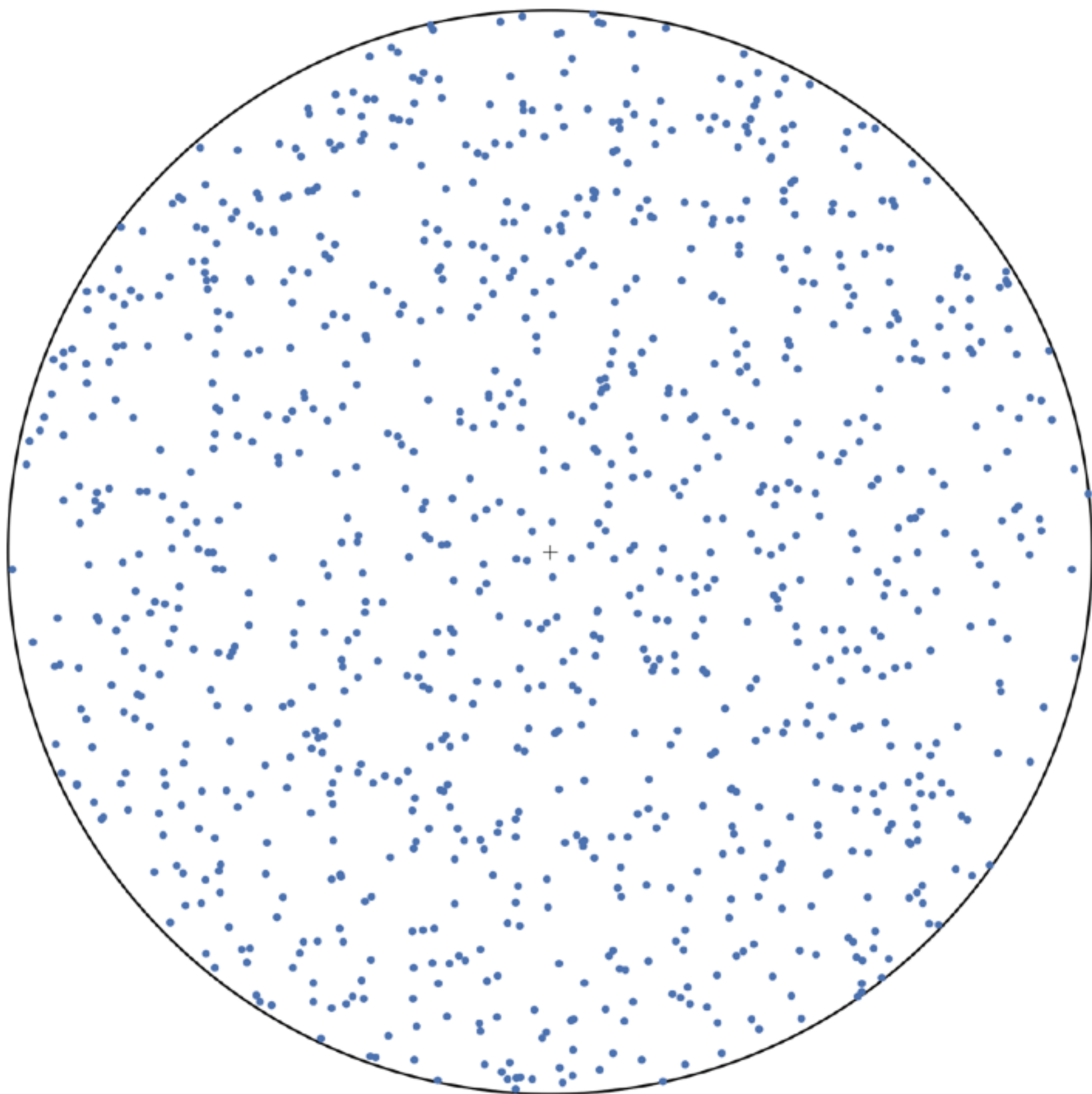
Symmetry

$$d(x, y) = d(y, x)$$

Triangle inequality

$$d(x, y) \leq d(x, z) + d(z, y)$$

Properties of any metric space



Simple random geometric graph

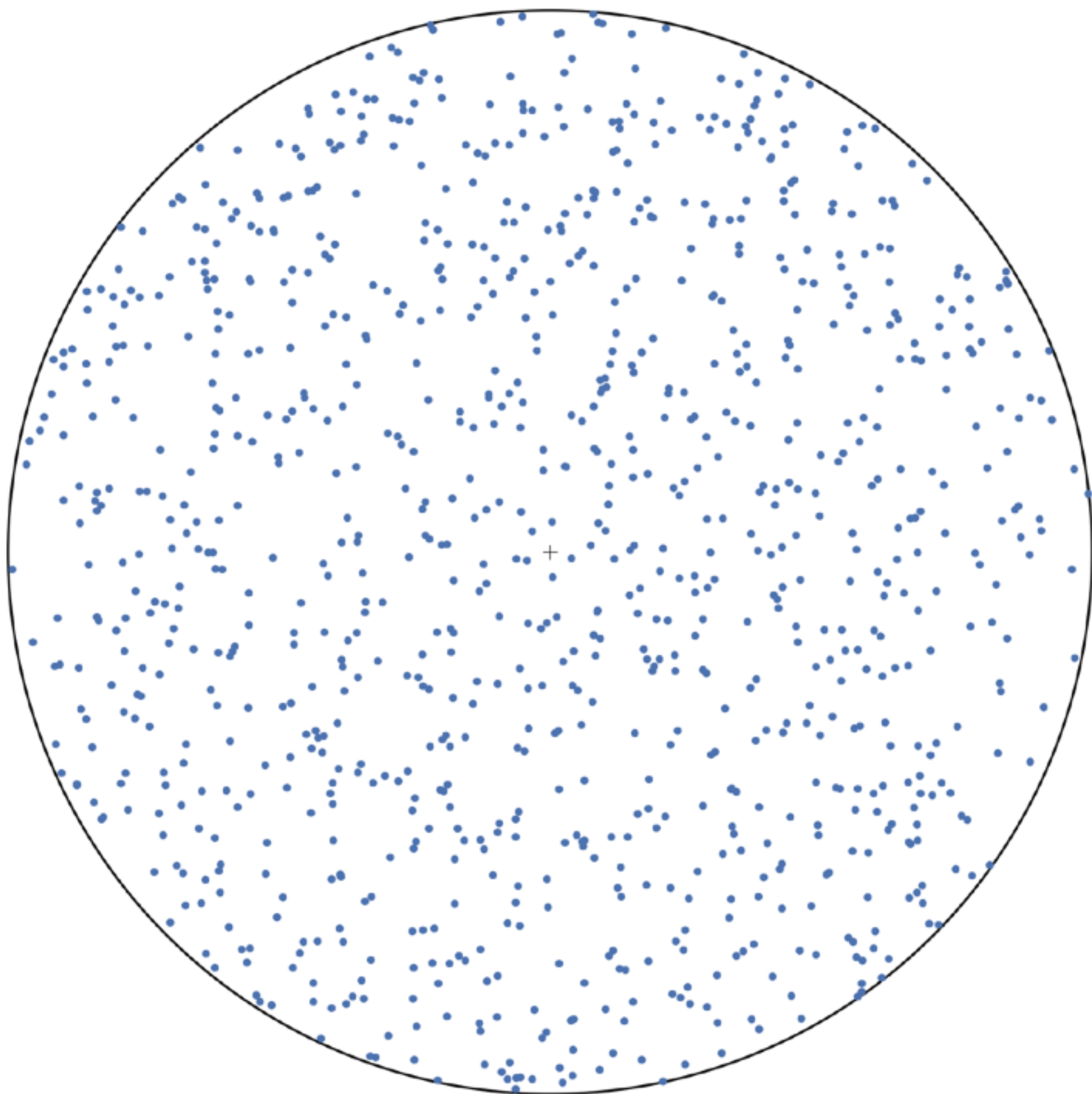
1. Sprinkle N nodes uniformly on a disk of radius R .
2. Connect any nodes separated by a distance less than r .

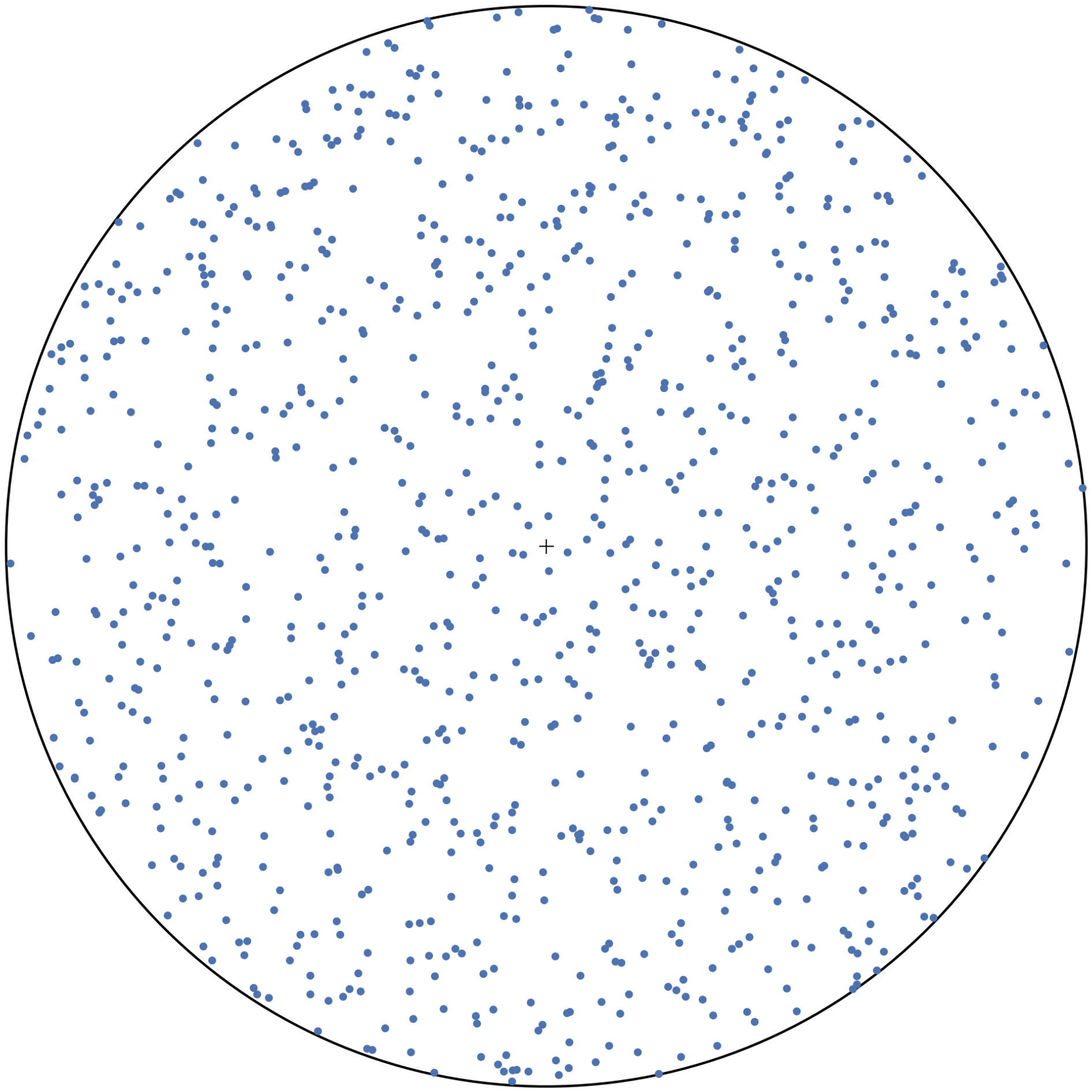
✓ high clustering

✗ binomial/Poisson degree distribution

Assume that the nodes are embedded in a metric space and that any two nodes are connected with a probability that is a decreasing function of the distance between them.

For further info, see Phys. Rep. 499, 1-101 (2011)





A geometric approach to clustering

Assume that the nodes are **embedded in a metric space** and that any two nodes are connected with a probability that is a **decreasing function of the distance** between them.

Properties of any metric space

Identity of indiscernibles $d(x, y) = 0 \Leftrightarrow x = y$

Non-negativity $d(x, y) \geq 0$

Symmetry $d(x, y) = d(y, x)$

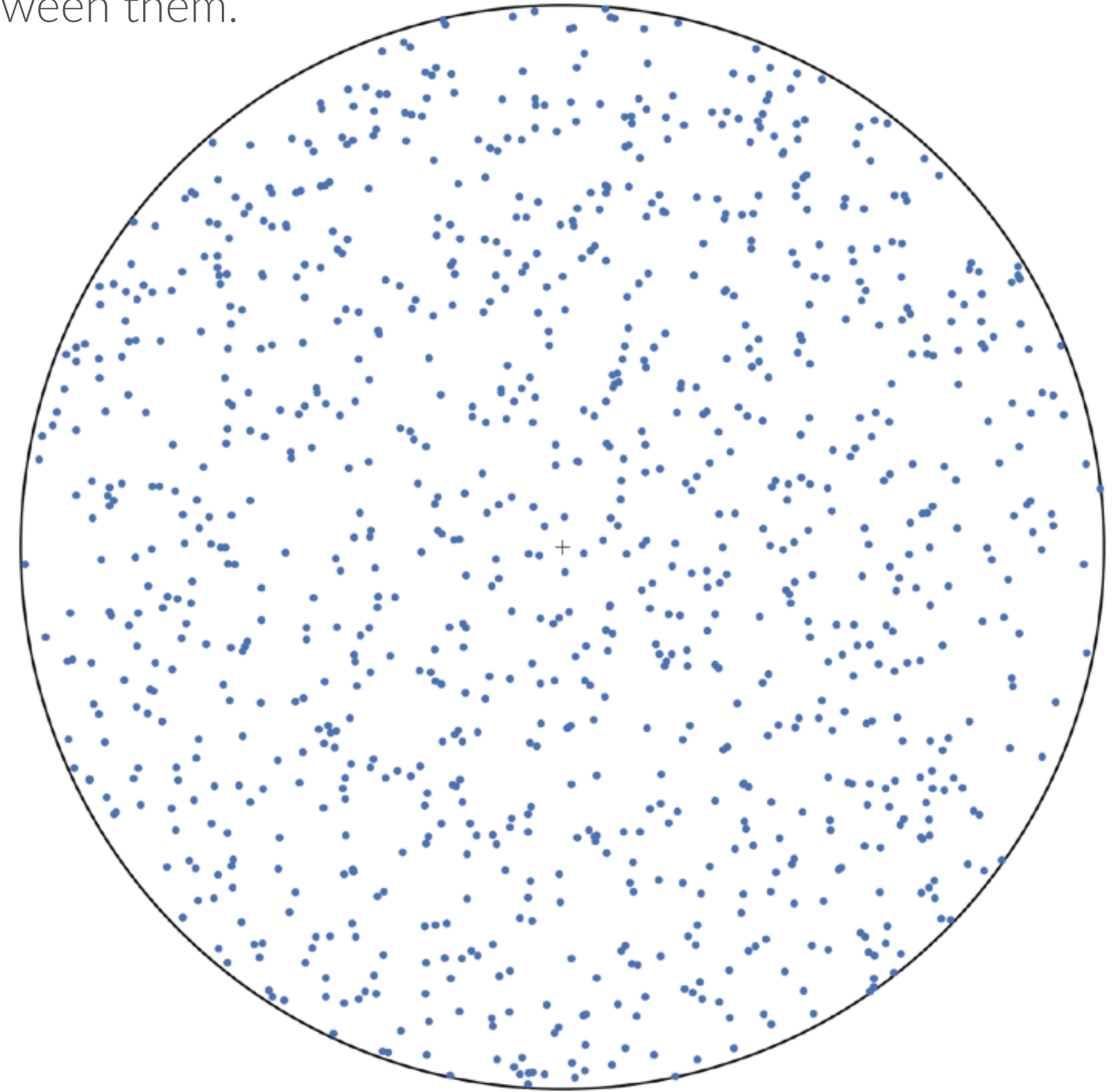
Triangle inequality $d(x, y) \leq d(x, z) + d(z, y)$

Simple random geometric graph

1. Sprinkle N nodes uniformly on a disk of radius R .
2. Connect any nodes separated by a distance less than r .

✓ high clustering

✗ binomial/Poisson degree distribution



Hyperbolic geometry

- ▷ Space of constant **negative curvature** (as opposed to flat or Euclidean space, or spherical space)
- ▷ Model for the $D = 2$ hyperbolic space : positive sheet of the **hyperboloid** defined by

$$x^2 + y^2 - z^2 = -1$$

- ▷ Distance between points (x_1, y_1, z_1) and (x_2, y_2, z_2) is

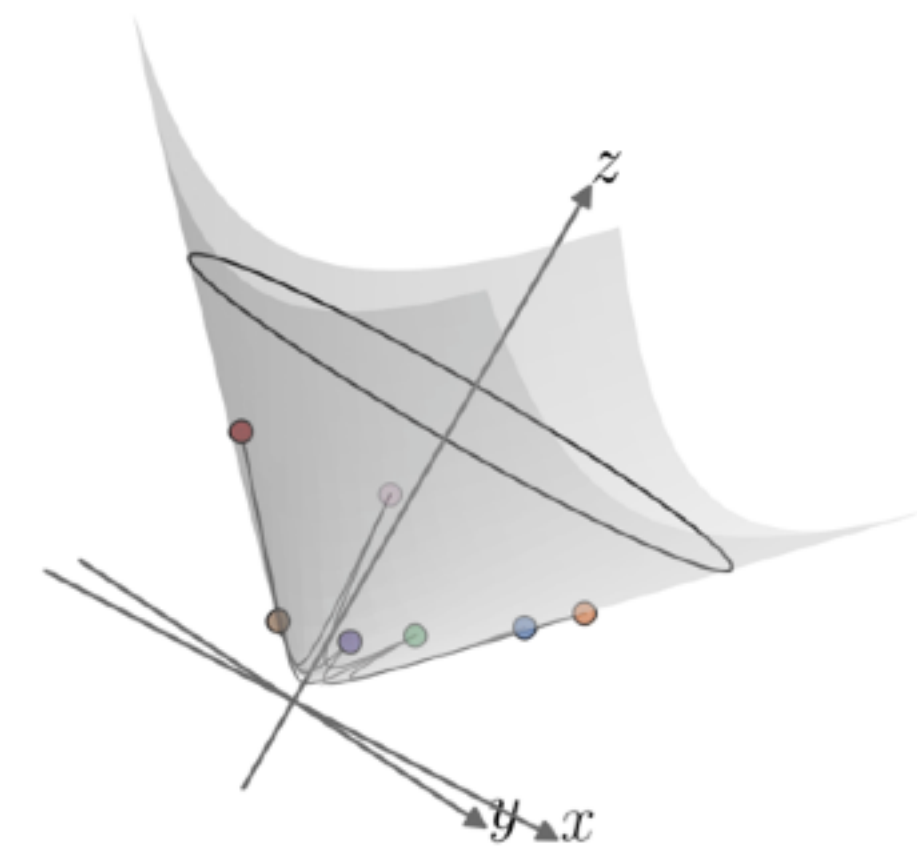
$$d(1, 2) = \operatorname{arccosh}(z_1 z_2 - x_1 x_2 - y_1 y_2)$$

- ▷ Polar coordinates

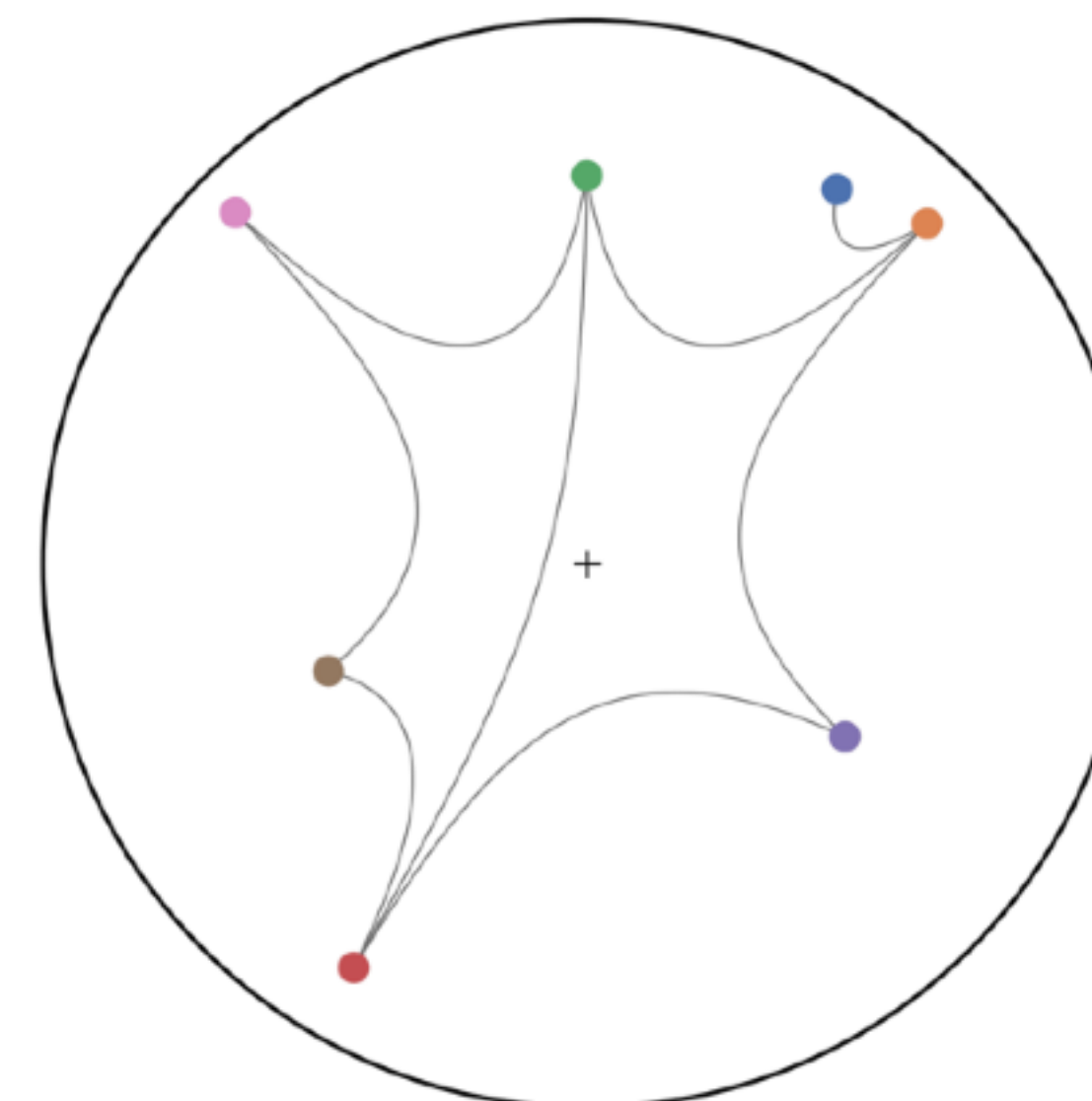
$$x = \sinh(r) \cos(\theta)$$

$$y = \sinh(r) \sin(\theta)$$

$$z = \cosh(r)$$



hyperboloid in $\mathbb{R}^{2,1}$



hyperbolic disk (r, θ)