

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

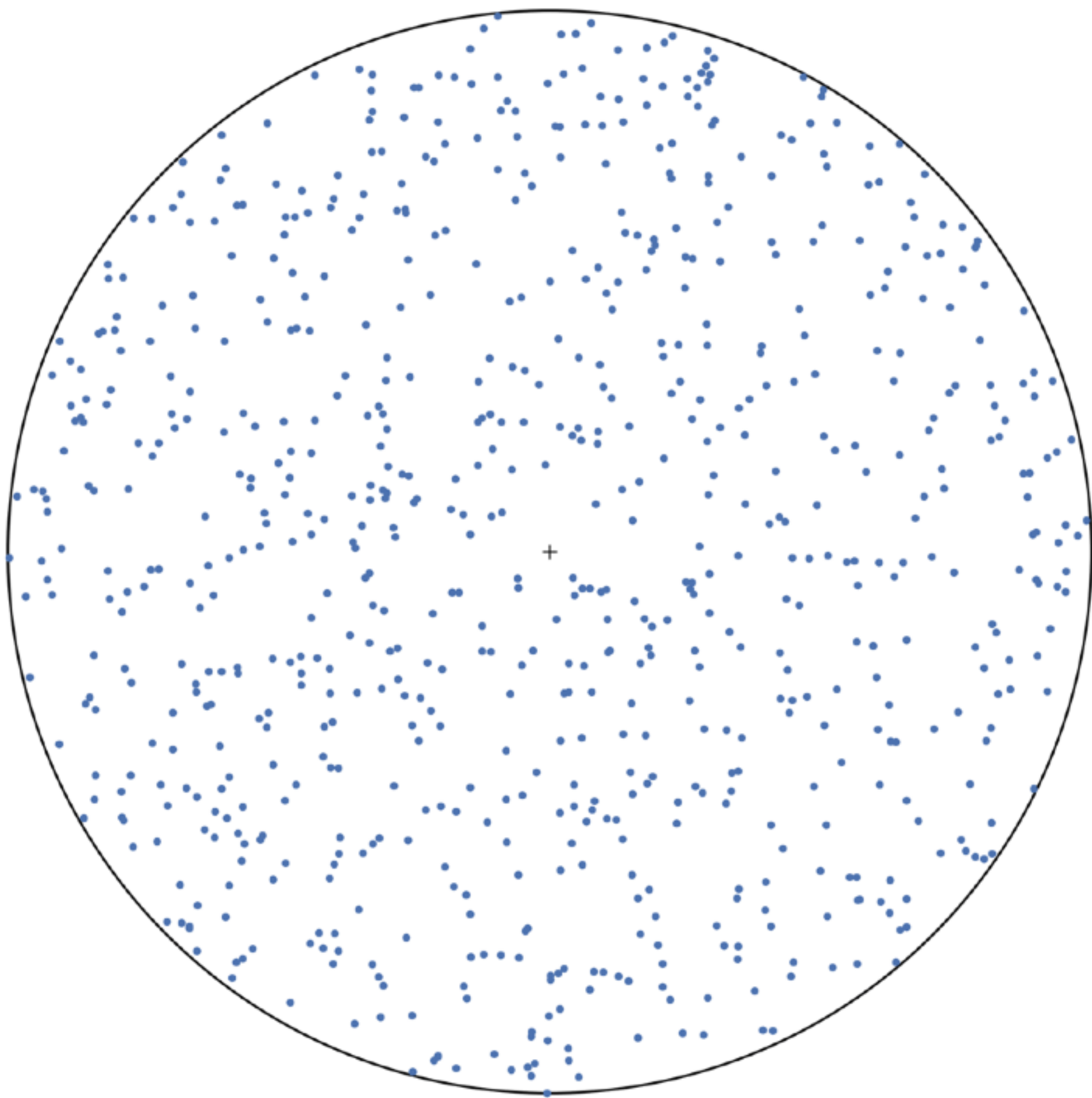
Assume that the nodes are embedded in a metric space and that any two nodes are connected with a probability that is a decreasing function of the distance between them.

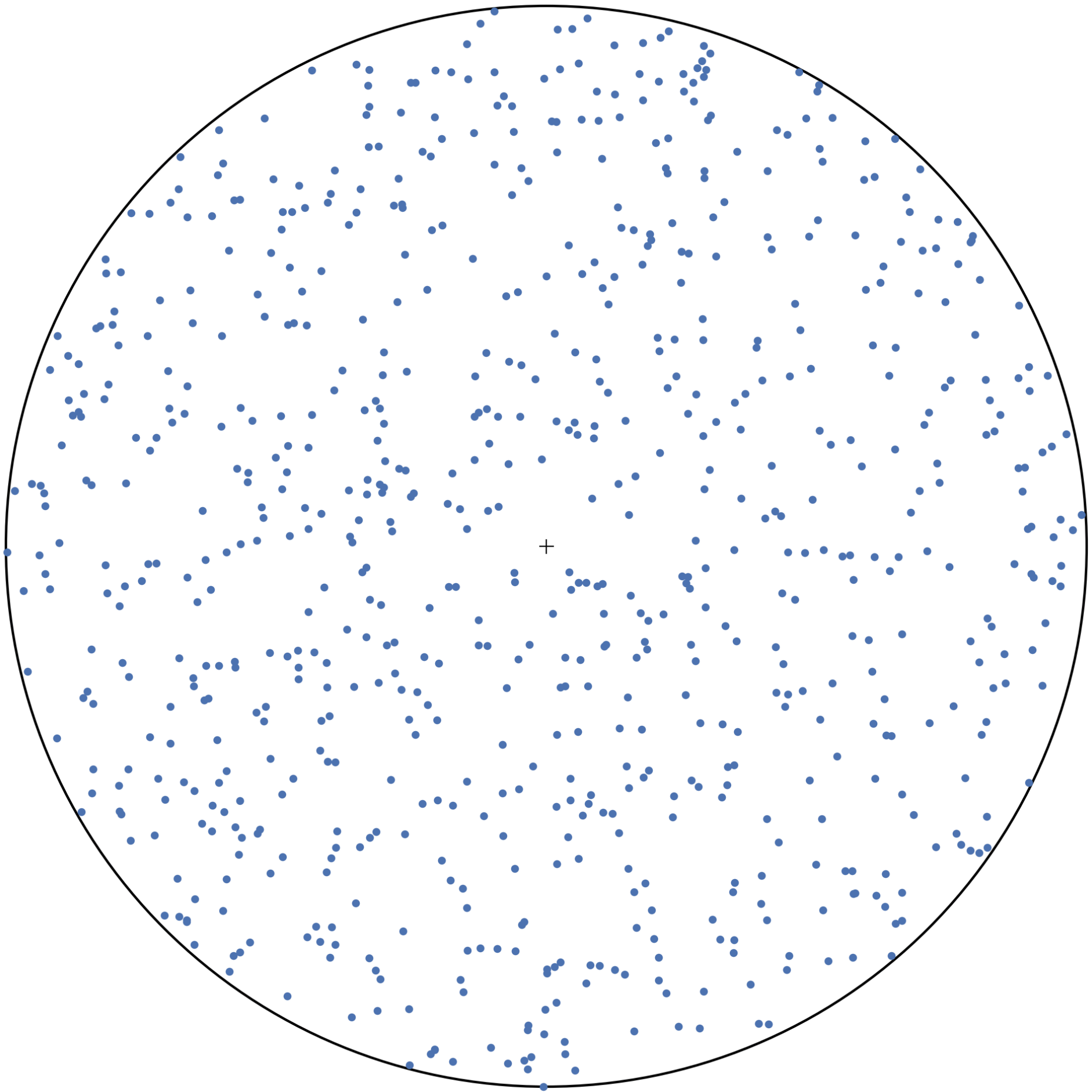
For further info, see Phys. Rep. 499, 1-101 (2011)



1

5





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

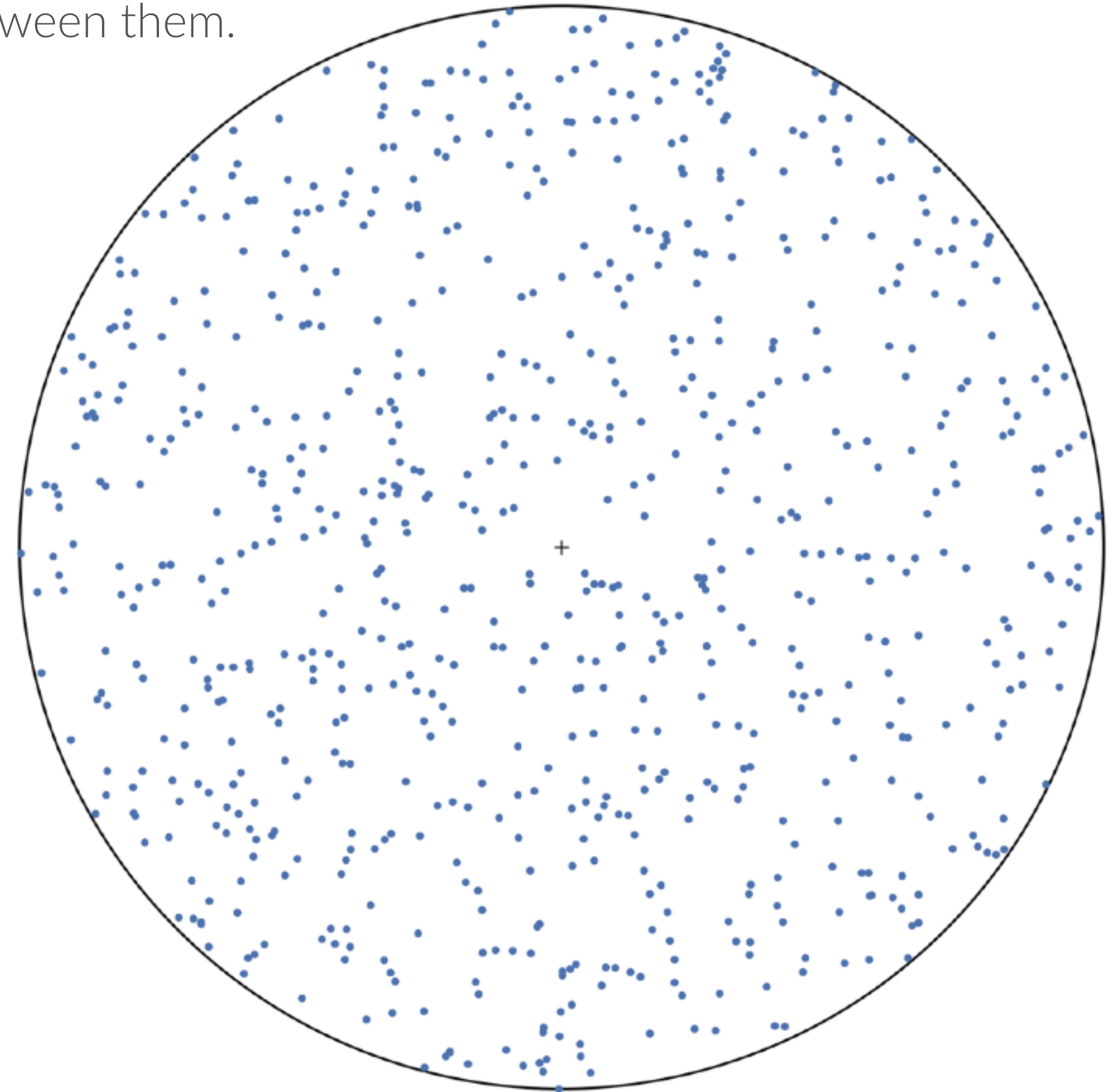
Triangle inequality  $d(x, y) \leq d(x, z) + d(z, y)$

## 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

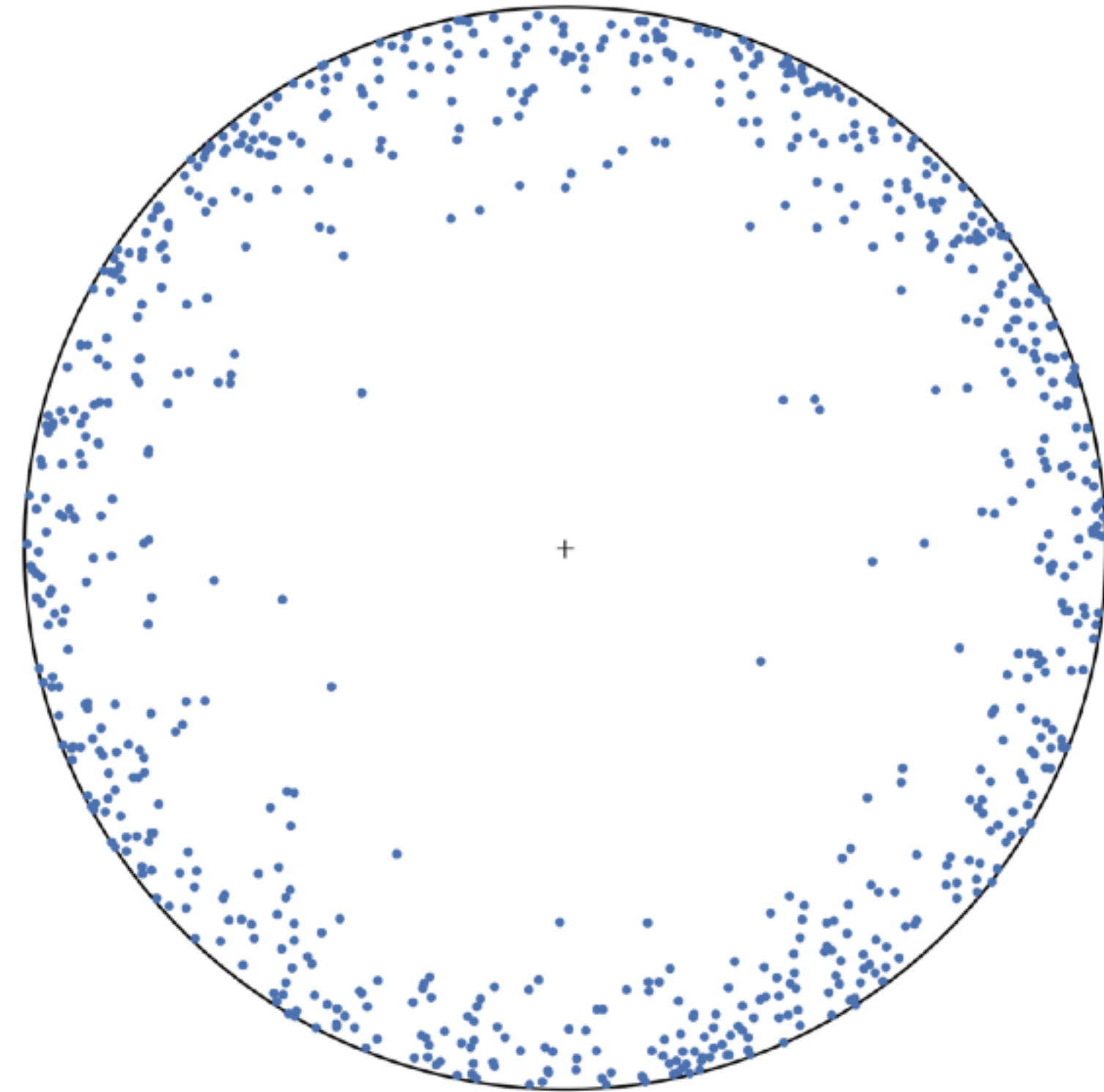
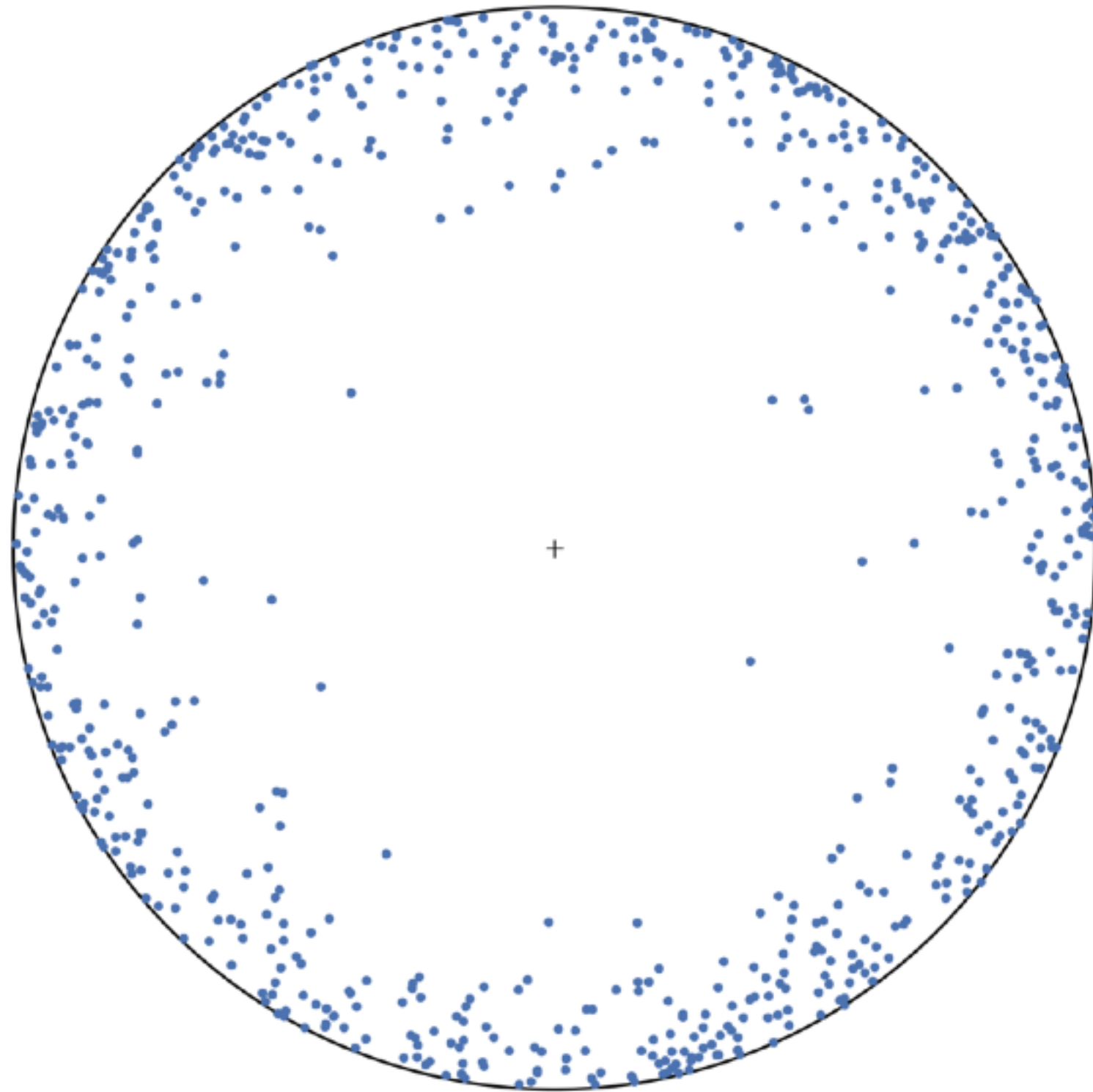
✗ binomial/Poisson degree distribution



# A geometric approach to clustering: Hyperbolic geometry

## Simple random geometric graph

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



- ✓ high clustering
- ✓ power-law degree distribution with exponent  $-3$