Let $A$ be a finite set in $\mathbb{R}^d$, $|A| = n$.

$d \geq 2$ size of $A$

For convenience, assume that $A \subseteq [0, 1]^d$, the unit cube.

We shall assume throughout that $n > \text{the amount of the national debt of } d$.

The purpose of this lecture is to study the "dimension" of $A$ using a quantity that comes from geometric measure theory:

$$I_s(A) = n^{-2} \sum_{a \neq a'} |a-a'|^{-s}$$

\(0 < s < d\) discrete energy of order \(d-s\)

\(\bigcirc\)
Let us begin by realizing that we have seen the underlying idea before.

Let \( B_d = \{ x \in \mathbb{R}^d : |x| \leq 1 \} \)

the unit ball

Then \( \int_{B_d} \int_{B_d} |x-y|^s \, dx \, dy < \infty \) \( y = (y_1, y_2, \ldots, y_d) \)

By \( B_d \)

\( \int_{B_d} \int_{B_d} \) \( s < d \)

But what if instead of the unit ball we take a lower dimensional object?

Simple example: \( d = 2 \)

\( X = \{ (z, 0) : 0 \leq z \leq 1 \} \)

What does it mean to integrate over \( X \)?

We know how to do this from multi-variable calculus:
We get
\[ \iint_{\Omega} |z - z'|^{-s} \, d\Omega d\sigma < \infty \]
\[ \text{if } s < 1, \]
so even though \( X = \{ (0,0) : t \in [0,1] \} \subset \mathbb{R}^2 \)
\[ \leq 1 \mathbb{R}^2 \]
the exponent restriction is 1-dimensional, not 2-dimensional because \( X \) is 1-dimensional.

Let's consider a less transparent example:
\( d = 2 \) \( X = \{ (\cos \theta, \sin \theta) : \theta \in [0,2\pi] \} \subset \mathbb{R}^2 \)
The corresponding "energy" quantity is:
\[ \iint \left| x(\theta) - y(\phi) \right|^{-s} \, d\theta d\phi \]
\( x(\theta) = (\cos \theta, \sin \theta) \quad y(\phi) = (\cos \phi, \sin \phi) \)
We get
\[ \frac{1}{2 \sqrt{1 - \cos(\theta - \varphi)}} \left[ 1 - \frac{s}{2} \right] \, d\theta d\varphi \]

\[ \cos(\theta - \varphi) \approx 1 - \frac{(\theta - \varphi)^2}{2} \]
so the integral above is
\[ \approx \frac{1}{2} \left[ 1 - \frac{s}{2} \right] \, d\theta d\varphi < \infty \]

iff \( s < 1 \), just like the case of the line segment.

With these calculations behind us, let's specialize to the discrete setting.

\[ A = \left\{ \left( \frac{i}{n}, \frac{i}{n} \right): 1 \leq i \leq n \right\} \]

\[ \ell = \left( \frac{1}{n}, \frac{1}{n} \right) \]
We have
\[ n - \sum_{i \neq j} \text{dist}(\frac{i}{n}, \frac{j}{n}), \frac{j}{n}) \]

\[ \approx n^{-2} \sum_{i \neq j} |i-j|^2 \cdot n^{-s} \cdot n^{1-s} = n^{1-s} \cdot n^{-s} \cdot n \quad 1 \leq i \leq n \]

Let \( u = i-j, \ v = j \)

We get
\[ n^{-2} \cdot n^{1-s} \sum_{u \neq 0} |u| \]

\[ \approx n^{-1+s} \sum_{u \sim n} |u| \cdot n \quad \text{if } s \leq 1 \]

and \[ \approx n^{-1+s} \quad \text{if } s > 1. \]

It follows that \( \mathcal{L}_s(A) \leq C \) independently.

If \( s \leq 1 \),

\[ \text{What happens if } s = 1? \]
Let us formalize things a bit.

**Definition:** Let $\{A_n\}$ be a family of subsets of $[0,1]$, $d \geq 2$, such that $|A_n| = \binom{n}{d}$, integer valued function of $n$.

and $A_n \subseteq A_{n+1}$ $\forall n$.

We say that $\{A_n\}$ is $s$-adaptable if

$$
\exists (A_n) \subseteq \mathbb{C}
$$

independently of $n$.

$$
\mathcal{S} = \{ n \}_{1 \leq k \leq d}
$$

**Exercise:** Suppose that $A_n \subseteq k$-dimensional manifold. Then $\{A_n\}$ is not $s$-adaptable if $s > k$.

But there are stranger objects in the world than manifolds!
The classical tridec Cantor set is constructed as follows:

\[ C_2 = \frac{1}{3}, \frac{2}{3} \]

\[ C_3 = \frac{7}{9}, \frac{8}{9} \]

and so on!

Let \( A_n = C_0 \times C_0 \) \[ |A_n| = 2^{n+1} \cdot 2^{n+1} = 2^{2n+2} \]

We have

\[ 2^{-4(n+1)} \sum_{a \neq a'} |a - a'| - s \]

\[ \approx 2^{-4n} \sum_{j=0}^{n} \sum_{s=1}^{3^n} 2^j \]

since it makes no sense to consider \( |a - a'| < 3^{-n} \)!!
Key question: If \( a \in \mathbb{C}_n \times \mathbb{C}_n \), how many \( a \)'s are there such that \( |a - a'| \approx 3^{-j}, \quad 0 \leq j \leq n \) ?

The answer is \( \approx \frac{2^{2n}}{2^{2j}} \) total size of \( \mathbb{C}_n \).

It follows that

\[
2^{-4n} \sum_{j=0}^{n} 3 \sum_{|a-a'| \approx 3^{-j}} 1
\]

\[
\approx 2^{-4n} \cdot 2^n \cdot 2^n \cdot \sum_{j=0}^{n} 3 \cdot 2 \cdot 2^{-2j}
\]

and the sum converges if

\[
s < 2 \cdot \frac{\log(2)}{\log(3)} = \text{Sentrical}
\]

This critical value for \( s \) gives us a notion of dimension.
Let's amplify some of the technical points that arose above. We used the fact that
\[ \sum_{u=1}^{n} u^{-s} \approx \frac{n^{1-s}}{-s+1} \]
as \( s \to 1 \), with constants independent of \( n \).

There are many ways to see this, but here is one that I find particularly instructive.
\[ \sum_{u=1}^{n} u^{-s} = \sum_{k=0}^{\log_2(n)} \sum_{u=2^k}^{2^{k+1}-1} u^{-s} \]
small amount of cheating here—please fix it!
\[ \sum_{k=0}^{\log_2(n)} 2^{-k s} 2^k \Rightarrow \text{converges if } s > 1 \]
yields \[ \sum_{u=1}^{n} u^{-s+1} \approx \log(n) \text{ if } s=1 \]
and \[ \sum_{u=1}^{n} u^{-s} \approx \frac{n^{1-s}}{-s+1} \text{ if } s < 1. \]
Let's take a more careful look at the Cantor set introduced above.

\[ 0 \quad \overset{1}{/} \quad \frac{1}{3} \quad \frac{2}{3} \]

\[ 0 \quad \overset{2}{/} \quad \frac{1}{9} \quad \frac{2}{9} \quad \frac{7}{9} \quad \frac{8}{9} \]

At the \( n \)th stage, we have \( 2^n \) intervals
of length \( 3^{-n} \), so the total length of
these intervals \( \to 0 \) as \( n \to \infty \).

There is a notion of dimension, called the
Hausdorff dimension, which can be expressed
in this context as the largest \( s \) for
\[ \int \int |x-y|^{-s} \, du(x)du(y) \]
“natural measure on the Cantor set”
Going back to families of discrete sets, we are left with some fascinating questions.

Q: Can you construct a family of sets \( \{ A_n \} \) with \( A_{n+1} \supset A_n \), so that

\[
I_s(A_n) = n^{-2} \sum_{a \neq a'} |a - a'| - s
\]

is bounded independently of \( n \).

If \( s < s_0 \) given any pre-assigned \( s_0 \in (0, d) \)?

Q: Is it possible to use \( I_s(A_n) \) in conjunction with other tools to study clustering and dimension of large point sets?