Geometry Equals Multi-Scale S.V.D.

Conclusions .

Promising directions

Nonlinear adaptations of singular value decompositions ,

·Local SVD models for data representation ,used in weather prediction as well as non parametric data modeling

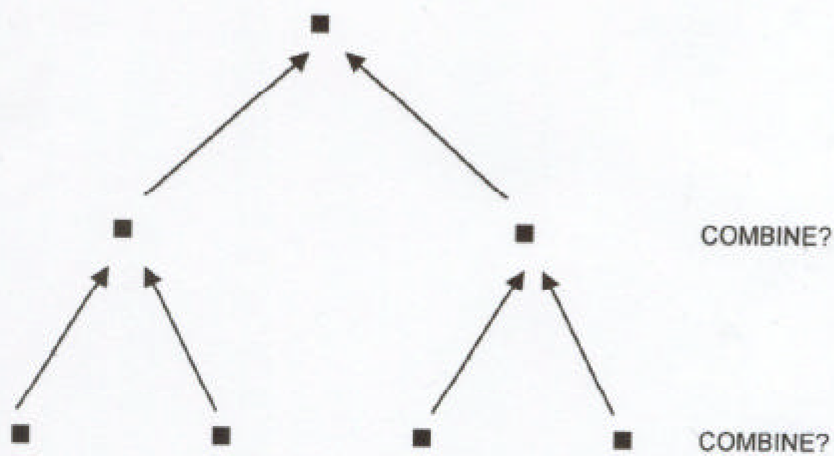
·Multilinear and separation of variable approach,for functional approximation and effective operator calculus .

• *Basis selection methods for dimensional reduction.* Independent component analysis ,sparse bases and localized adapted functional transcriptions .

Not Today!

BOTTOM-UP APPROACH

1.CHECK NEIGHBORING BOXES: IF GOOD DESCRIPTION THEN COMBINE.
2.IF COMBINE, REPEAT. IF DON'T COMBINE, STOP.
3. REPEAT UNTIL TREE EXHAUSTE



Hidden In This Lecture:

G. David - S. Semmes Theory (1990's)

See: Analysis of and on Uniformly Rectifiable Sets (A.M.S., 1993)

PHILOSOPHY

In $\mathbb{R}^n$, n large,

THERE ARE

NO FUNCTIONS

(or very few)

BUT THERE ARE

PROBABILIY MEASURES

E.G. DATA SETS

Step Zero: Study Probability Measures by Methods Analogous to Multiscale Function Theory

What is a Curve?

Yes

Yes

Yes

$$F : [0,1] \longrightarrow \mathbb{R}^n$$

continuous

We often identify F with its Image.

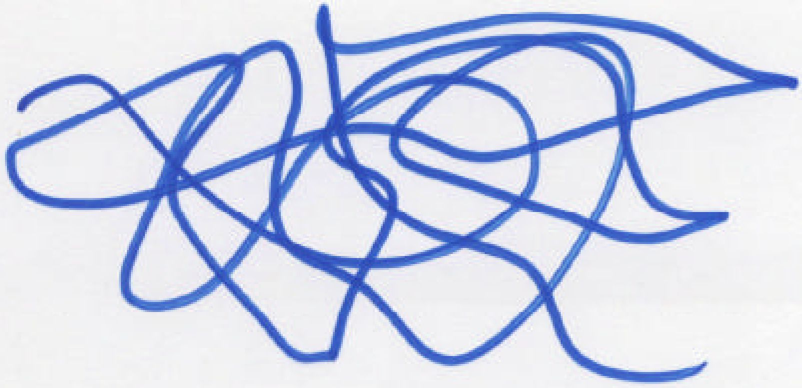
Note there are space filling curves.

Lipschitz Functions:

$$|F(x) - F(y)| \leq M\,|x-y|$$

Bi Lipschitz Functions:

$$\frac{1}{M} \leq \frac{|F(x) - F(y)|}{|x-y|} \leq M$$



K connected ↗

If K is <u>connected</u> and

length$(K) = \ell(K) = \mathcal{H}^1(K) < +\infty$

There is

$$F : [0, A] \xrightarrow{\text{onto}} K,$$

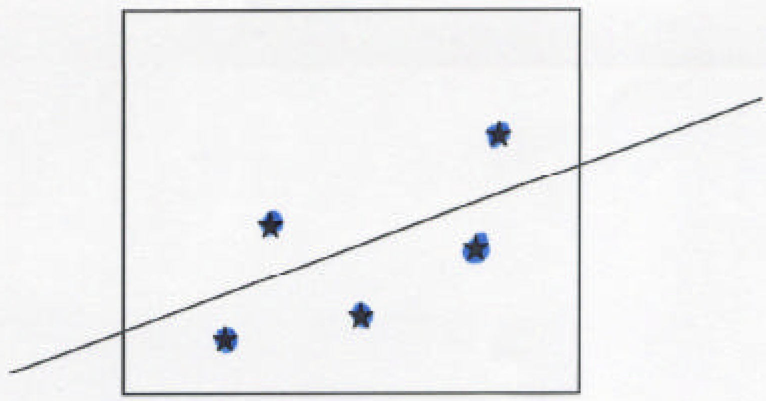
$$|F(x) - F(y)| \leq |x - y|,$$

$\ell \leq A \leq 2\ell$,

almost all points on K
are hit at most twice,

$$|F'(x)| = 1 \quad (\text{a.e. } x)$$

almost every point on K
has a "tangent"



$$\beta_\infty(Q) = \inf(\sup l(Q)^{-1} \text{distance}(z, L))$$

Where:

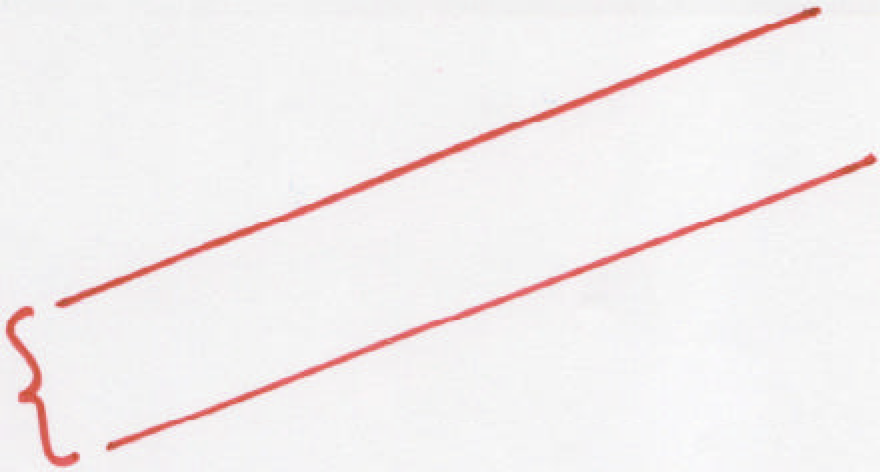
1. **K = Blue Set**
2. **Q = Cube (Or Square)**
3. **L = Red Line**
4. **inf is over all choices of lines**
5. **sup is over all z in $K \cap Q$**

SO $\beta_\infty(Q)$ is the NORMALIZED DISTANCE FROM K TO THE BEST APPROXIMATING LINE.

COMMENTS:

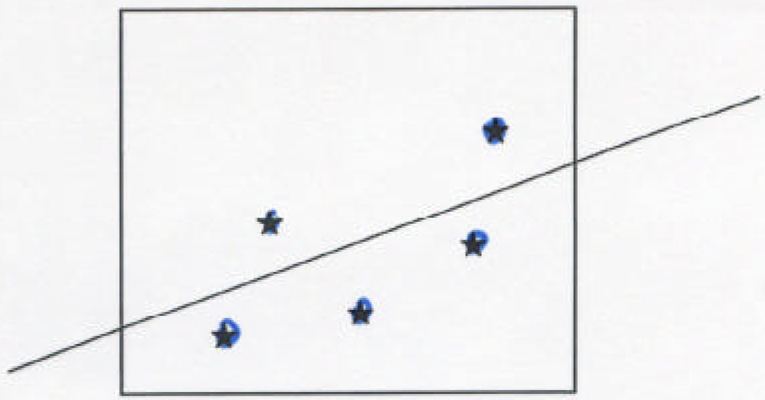
1. **$\beta_\infty(Q)$ is VERY SENSITIVE TO NOISE**
2. **THE UNITS OF $\beta_\infty(Q)$ ARE DIMENSIONLESS (INDEPENDENT OF THE LENGTH SCALE OF Q)**

width $\beta_{\infty}(Q)\,\ell(Q)$



$0 \leq \beta \leq$ "1"

Definition of β_2 for Probability distributions (μ)



$$\beta_2(Q) = \inf(\mu(Q)^{-1} \int (l(Q)^{-1} \text{distance}(z, L))^2 d\mu(z))^{1/2}$$

Where:

1. μ = Probability Distribution
2. Q = Cube (Or Square)
3. L = Red Line
4. inf is over all choices of lines
5. Integral is Over Q

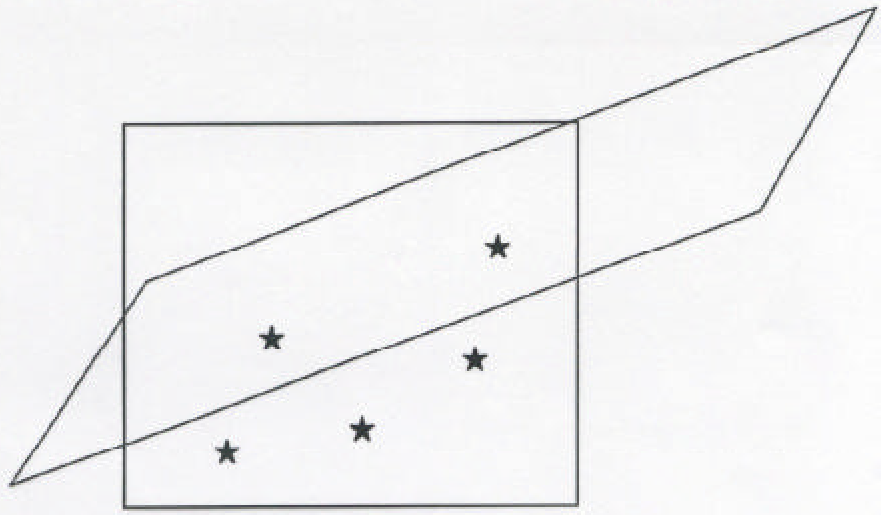
SO $\beta_2(Q)$ is the LEAST MEAN SQUARE DISTANCE TO THE BEST APPROXIMATING LINE (W.R.T. μ ON Q)

COMMENTS:

2.$\beta_2(Q)$ is NOT SENSITIVE TO NOISE

3. **THE UNITS OF $\beta_2(Q)$ ARE DIMENSIONLESS (INDEPENDENT OF THE LENGTH SCALE OF Q)**
4. **This is Classical Statistics**

Higher Dimensions: Definition of β_2 for Probability distributions (μ)



Suppose μ is on $\mathbb{R}^n$ and $d < n$ (d= Dimension to be studied)

$$\beta_{2,d}(Q) \equiv \beta_2(Q) = \inf(\mu(Q)^{-1} \int (l(Q)^{-1}\text{distance}(z, P))^2 d\mu(z))^{1/2}$$

Where:

6. μ = Probability Distribution
7. Q = Cube (Or Square)
8. P = Hyperplane of Dimension d
9. inf Is Over All Choices of Hyperplanes P
10. Integral is Over Q

SO $\beta_{2,d}(Q)$ is the LEAST MEAN SQUARE DISTANCE TO THE BEST APPROXIMATING HYPERPLANE (W.R.T. μ ON Q)
COMMENTS: As previously, this is classical statistics. The quantity $\beta_2(Q)$ is dimensionless.

β_∞ number for a cube Q:

$$\beta_{\infty Q} = \frac{\text{smallest width of a strip containing } K \cap Q}{l(Q)}$$

Theorem (P. W. Jones, 1990)

Let K be a subset of $\mathbb{R}^2$. K is contained in a curve with finite length $\iff$ diam(K) and $\sum \beta_Q^2 \cdot l(Q)$ are finite.
Moreover, the length of the shortest curve containing K is comparable to diam$(K) + \sum \beta_Q^2 \cdot l(Q)$. $(\beta = \beta_\infty)$

Theorem (C. Bishop, P. W. Jones, 1990)

If K is a bounded set in $\mathbb{R}^n$ and if $J(x) \le M$ for all $x \in K$, then K is contained in a curve of length not exceeding $c_1 e^{c_2 M}$diam(K).

Higher Dimensions, β_∞
$d = 1$!

(K. Okikiolu) The First Theorem Holds in $\mathbb{R}^n$, $n \geq 3$. (1992)

Remark: Estimates ↑ are exponential in n. (The Curse of Dimensionality.)

BUT

R. Schul: The First Theorem Holds in Hilbert Space, so "everything" can be done with estimates independent of n.

↑
In Process of Being Checked (Sept., 2004)

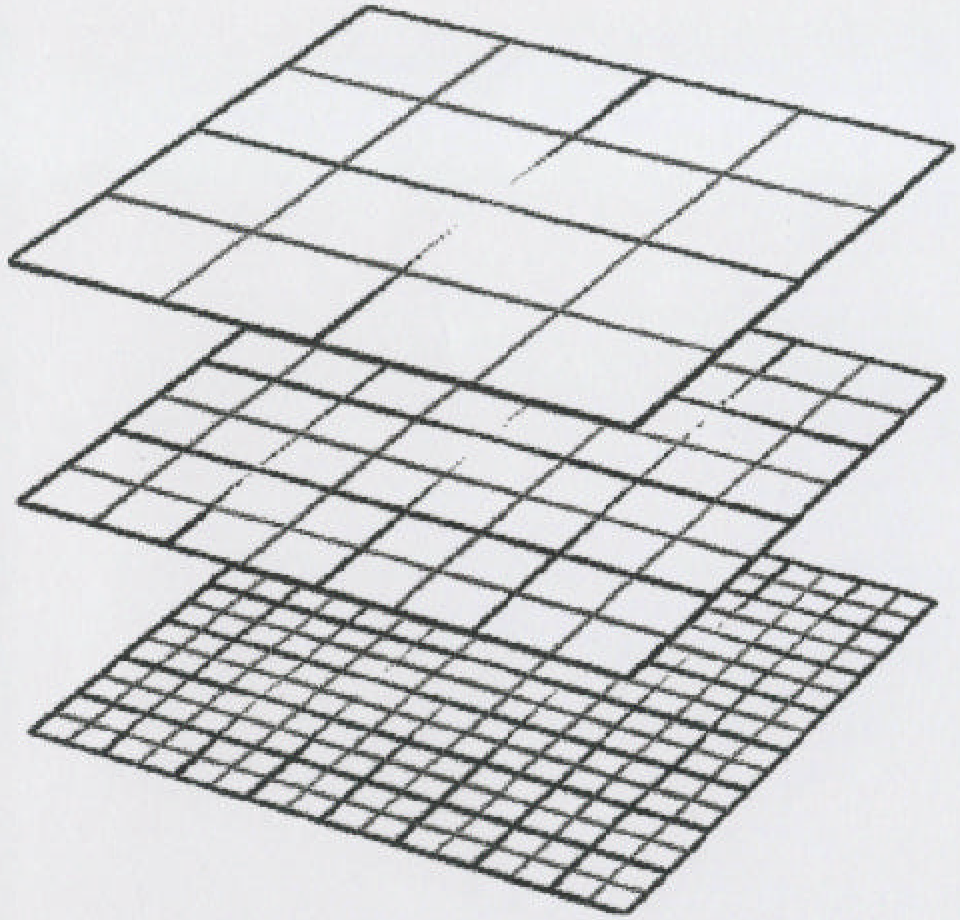


Figure 1: Wavelet Pyramid

DICTIONARY

wavelets	β numbers
$a_{j,k}$ for function f	$\beta_{Q_{j,k}}$ for set K
analysis and synthesis of the function f	analysis and synthesis of curve $\Gamma \supseteq K$
$\|f\|^2 = \sum \|a_{j,k}\|^2$	$l(\Gamma) \sim \sum \beta_Q^2 \cdot l(Q)$
square function $W_\psi(x)^2$	$J(x)$

Comments On

$$\sum \beta^2(Q)\,\ell(Q)$$

1. The sum is over <u>all</u> dyadic Q:

$$\sum = \sum_{n=-\infty}^{\infty} \sum_{\ell(Q)=2^{-n}}$$

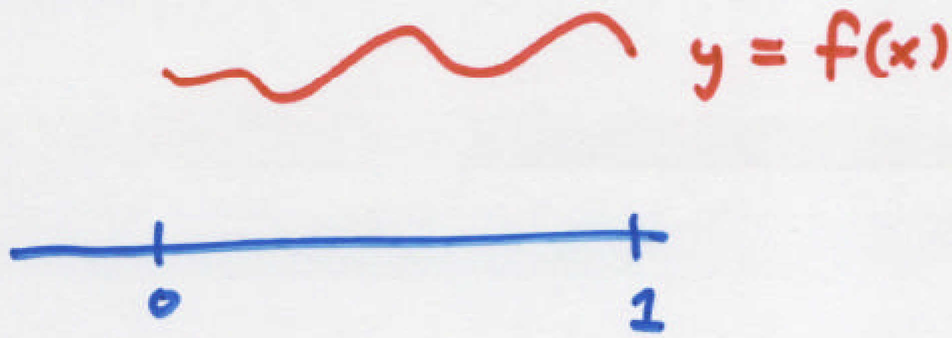
2. The sensitivity to noise is much too large for numerical versions. We therefore replace

$$\beta_\infty(Q) \text{ by } \beta_2(Q)$$

least mean squares

and develop a parallel theory.

You Have Seen TSP



$$\text{length} = \int_0^1 \sqrt{1 + (dy/dx)^2}\, dx$$

$$\sim \int_0^1 \left\{ 1 + \left(\frac{dy}{dx}\right)^2 \right\} dx$$

$\frac{1}{2}$!

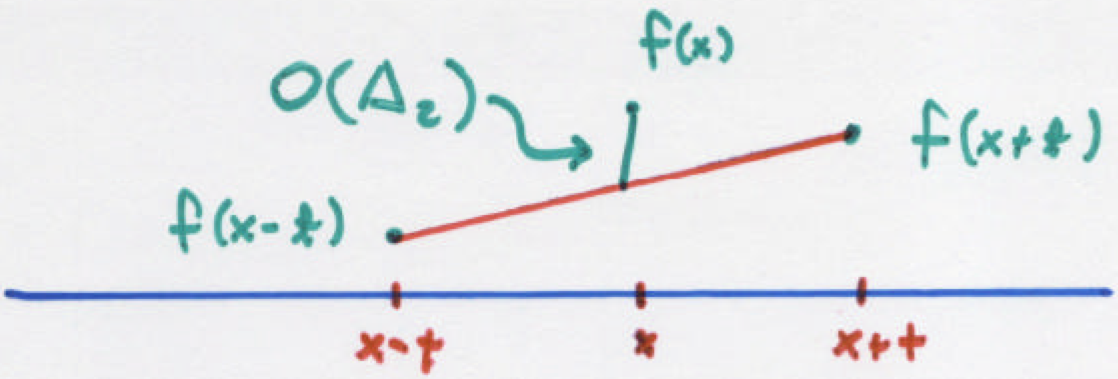
$$\cong \text{diam} + \int_0^1 \left(\frac{dy}{dx}\right)^2 dx$$

Careful Pythagoras:

$$\int_0^1 \left(\frac{dy}{dx}\right)^2 dx \sim \sum \beta^2(Q)\, \ell(Q)$$

Sobolev and Beta

$$\Delta_2 f(x,t) = f(x+t) + f(x-t) - 2f(x)$$



So $$\frac{|\Delta_2 f(x,t)|}{t} \approx \beta_\infty(\{y_1, y_2, y_3\})$$

(at scale t)

Square, Integrate dx, $\frac{dt}{t}$

$$\int_{(\mathbb{R})} \int_0^\infty \frac{(\Delta_2 f(x,t))^2}{t^3}\, dt\, dx \approx \sum \beta_\infty^2(Q)\, \ell(Q)$$

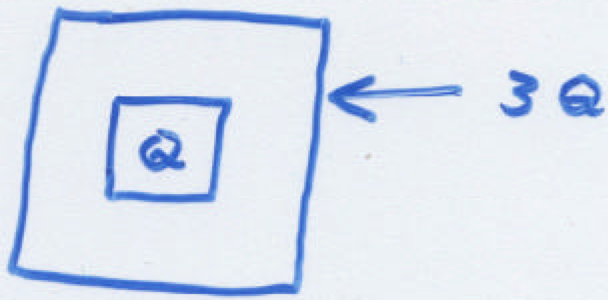
if F Lipschitz

$$= c \int_{\mathbb{R}} |f'|^2$$

Techicalities:

Q = Dyadic Cube

1. Compute $\beta(3Q)$



or

2. Use 2^n special translations of the dyadic grid

"Edge Effects"

Translated Dyadic Grids

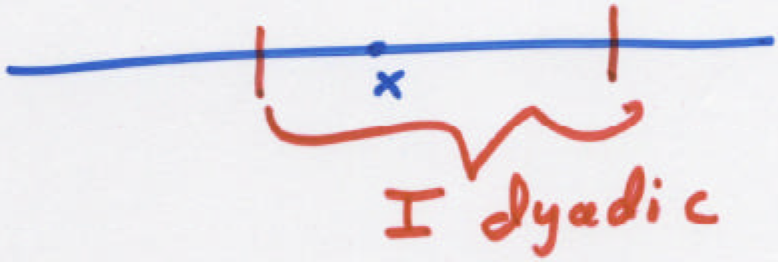
$(2,3) = 1$!

The $1/3$ Trick:

Let $x \in \mathbb{R}$, $n \geq 0$. Then either:

1. $\exists\ I$ dyadic, $\ell(I) = 2^{-n}$,

$$\operatorname{dist}(x, I^c) \geq \tfrac{1}{6}\,\ell(I)$$



x is far from the edge

or

2. $\exists\ \tilde{I} \in$ dyadic grid $+ 1/3$ (translate by $1/3$), $\ell(\tilde{I}) = 2^{-n}$

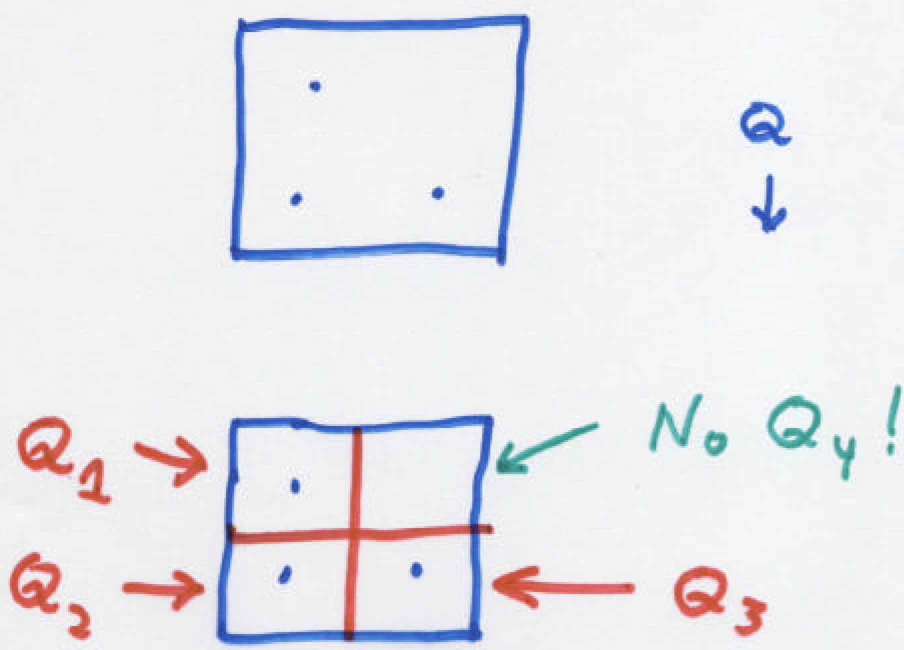
$$\operatorname{dist}(x, \tilde{I}^c) \geq \tfrac{1}{6}\,\ell(I)$$

$\mathbb{R}^n$: Same, but 2^n Grids (Homework)

Comments on C.O.D.

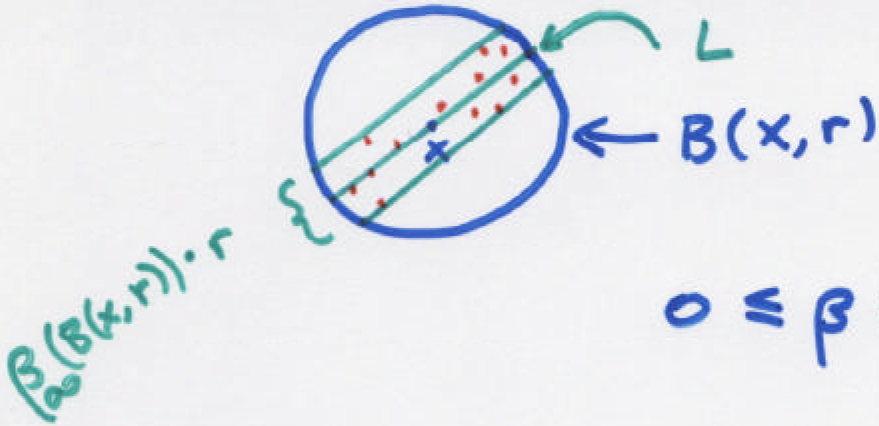
1. In $\mathbb{R}^n$ there are 2^n subcubes of $[0,1]^n$ with length $1/2$.
Too many to search! But a data set with N points occupies at most N cubes.
The binary expansion of $x \in \mathbb{R}^n$ carries a natural coding. Correctly used this organizes the cubes at each scale. AND

this is hierarchical:



Comment 2.

Balls "do not have C.O.D."
(but are difficult to implement numerically)



$0 \le \beta \le 2$

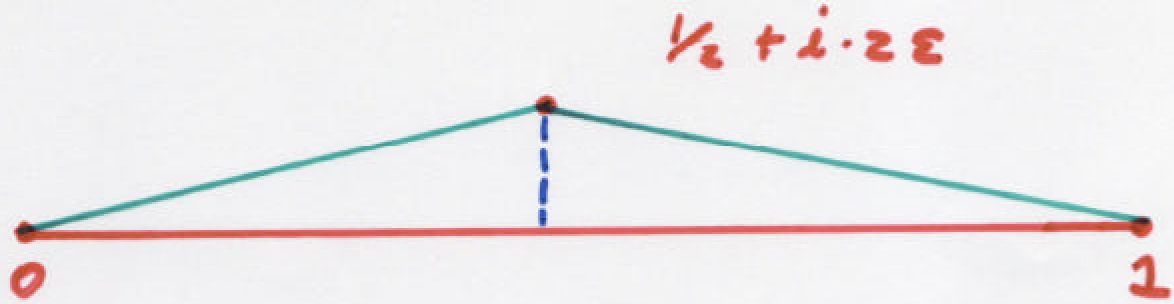
Then

$$\sum \beta^2(Q)\,\ell(Q) \sim \int_{\mathbb{R}^2} \int_0^\infty \beta^2(x,r)\,\frac{dr}{r^2}\,dx$$

$$\int_{\mathbb{R}^n} \int_0^\infty (\cdots)\,\frac{dr}{r^n}\,dx$$

Pythagorus Was Right!

TSP



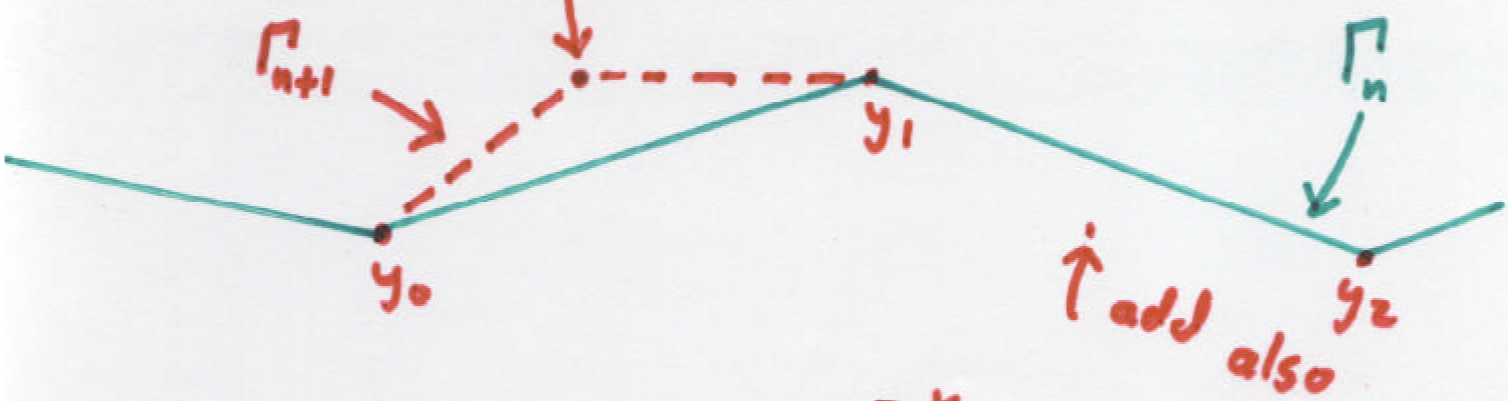
β_∞ for $\{0, 1, 1/2 + i \cdot 2\varepsilon\} = \varepsilon$

$$\ell(\text{Green}) = 2\sqrt{(1/2)^2 + (2\varepsilon)^2}$$
$$= \sqrt{1 + 16\varepsilon^2}$$
$$\sim 1 + 8\varepsilon^2$$
$$= 1 + 8\beta_\infty^2$$

Length added is <u>quadratic</u>

Construct Γ_{n+1} from Γ_n

add a vertex $x = x_{n,j}$



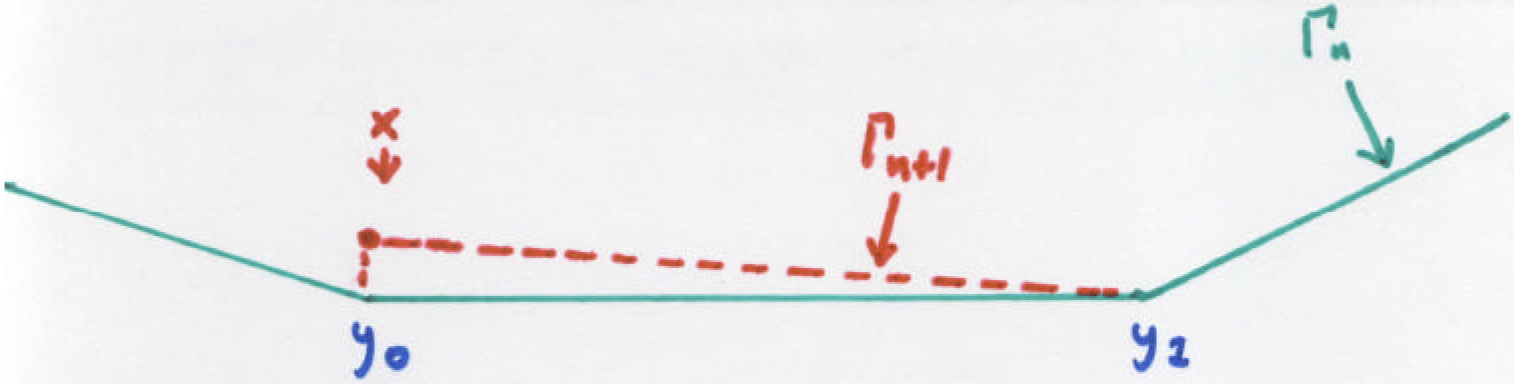
1. $|y_0 - y_1| \sim 2^{-n}$
2. "Furthest Insertion" of x
3. Compute β for $\{y_0, y_1, x\}$

   Model: $\beta_\infty(Q) \sim \beta_\infty(Q, \{\ \})$

   $|x - y_0| \sim |x - y_1| \sim 2^{-n}$

4. Pythagorus: length added is $\approx \beta_\infty^2(Q)\,\ell(Q)$ (see dotted lines)
5. Different points (vertices) added to Γ_{n+1} have "disjoint" Q, $\ell(Q) \sim 2^{-n}$

Technicalities



$|x - y_0| \sim 2^{-n}$, $|y_0 - y_1| \geq 10^3 \cdot 2^{-n}$

$\beta(Q) \approx 0$

Length Added $\sim 2^{-n} \gg \beta_\infty^2(Q) \cdot 2^{-n}$

"Free Interval"

Solution:

$$\text{Length Added} \leq \frac{1}{2}\,\ell(I_Q)$$

Free Interval

$$\ell(\Gamma) \stackrel{\sim}{\leq} \text{diam} + \sum_Q \beta_\infty^2(Q)\,\ell(Q)$$

Proof:

$$\ell(\Gamma_{n+1}) - \ell(\Gamma_n) \stackrel{\sim}{\leq} \sum_{\ell = 2^{-n}} \beta^2(Q)\,\ell(Q) + \sum_{\text{Free}_n} \tfrac{1}{2}\ell(I_Q)$$

$$\ell(\Gamma_0) \sim \text{diameter}$$

Telescope:

$$\ell(\Gamma_n) \leq c\,\text{diam} + c\sum \beta^2(Q)\,\ell(Q) + \tfrac{1}{2}\sum \ell(I_Q)$$

up to stage n

$$\leq c\,\text{diam} + c\sum \beta^2(Q)\,\ell(Q) + \tfrac{1}{2}\ell(\Gamma_n)$$

$$\Rightarrow \ell(\Gamma_n) \leq 2c\,\text{diam} + 2c\sum \beta^2(Q)\,\ell(Q)$$

Some $d = 1$ Objects

1.

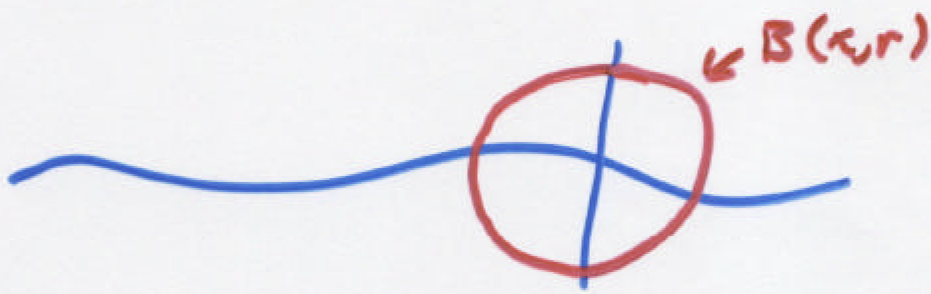
Lipschitz Graph

$$\Gamma = \{(x, A(x)) : x \in \mathbb{R} \text{ or } I\}$$

$$|A(x) - A(y)| \le M|x - y|$$

$$|A'| \le M$$

2.



Ahlfors-David Condition:

$$\ell(K \cap B(x,r)) \le Mr$$

Not too much length in any Ball. ($\longleftrightarrow$ BMO!)

3. Small Constant "Quasicircles"

K = curve = "fractal"

$\beta_\infty(Q) \leq \varepsilon_0$

$\longleftrightarrow$ K is "Riefenberg Flat"

Th. $\beta_\infty(K) \leq \varepsilon_0$ (sup over Q)

$\Longrightarrow K \subset \Gamma$ a "δ Flat Curve"

Applications to Dimensions

1. Upper Bounds: $\beta_\infty(Q) \leq \varepsilon$ for <u>all</u> Q

$$\Rightarrow \mathrm{Dim}(K) \leq 1 + c_0 \varepsilon^2$$

(Easy, available also for approximation by d planes: $\leq d + c\varepsilon^2$)

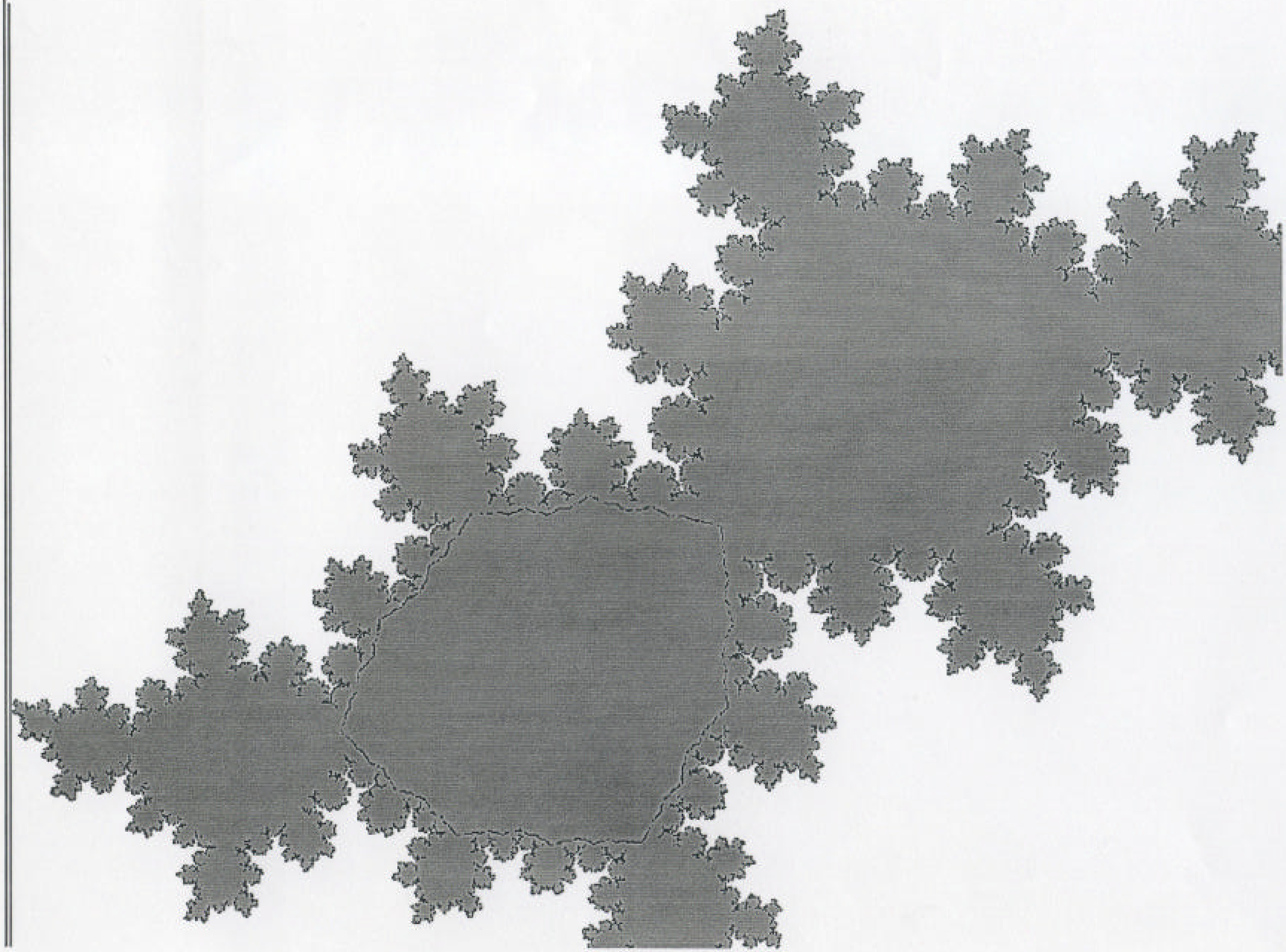
2. Lower Bounds: Here we require K is connected. Then if $K \cap Q \neq \emptyset$

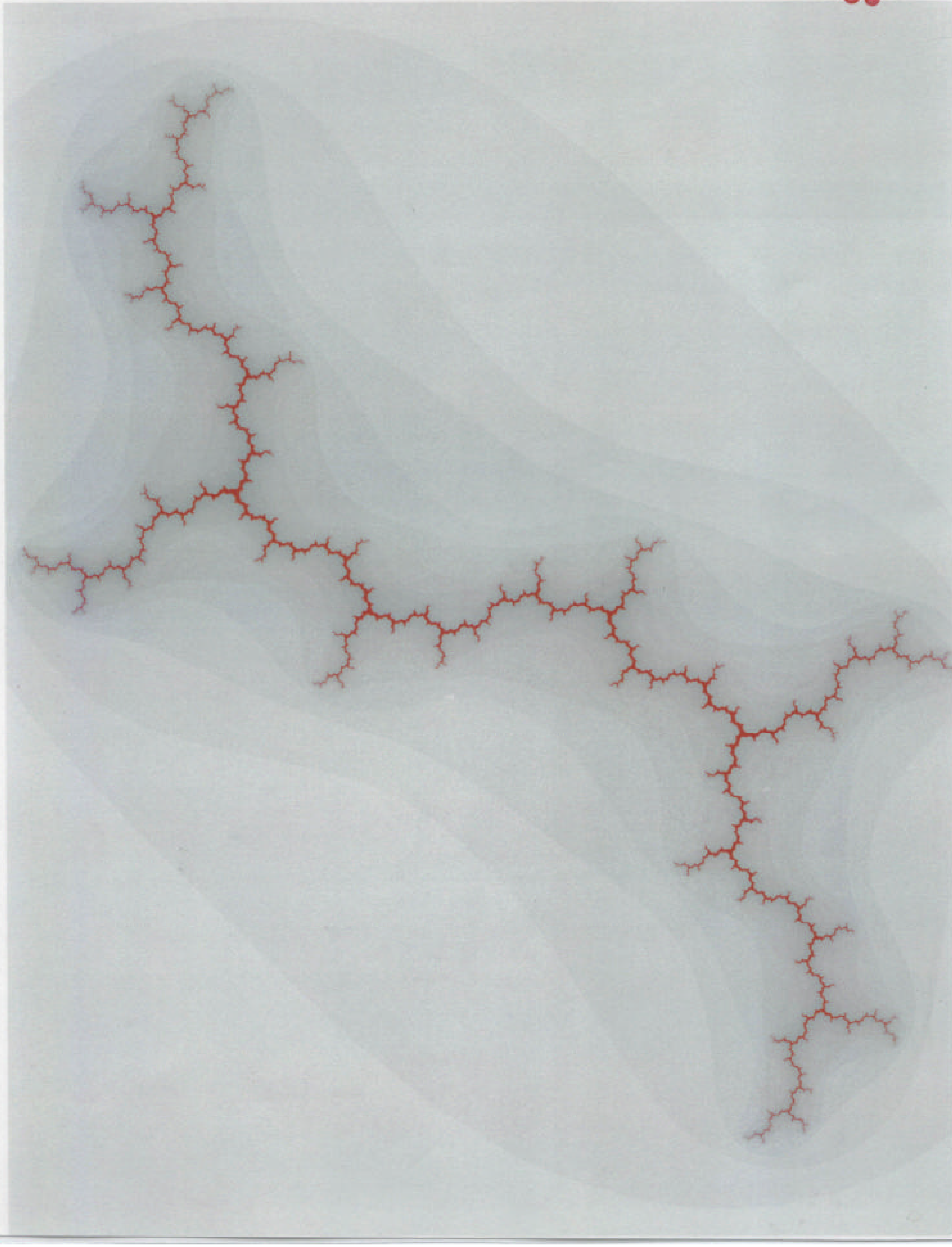
$$\Rightarrow \beta_\infty(Q) \geq \varepsilon,$$

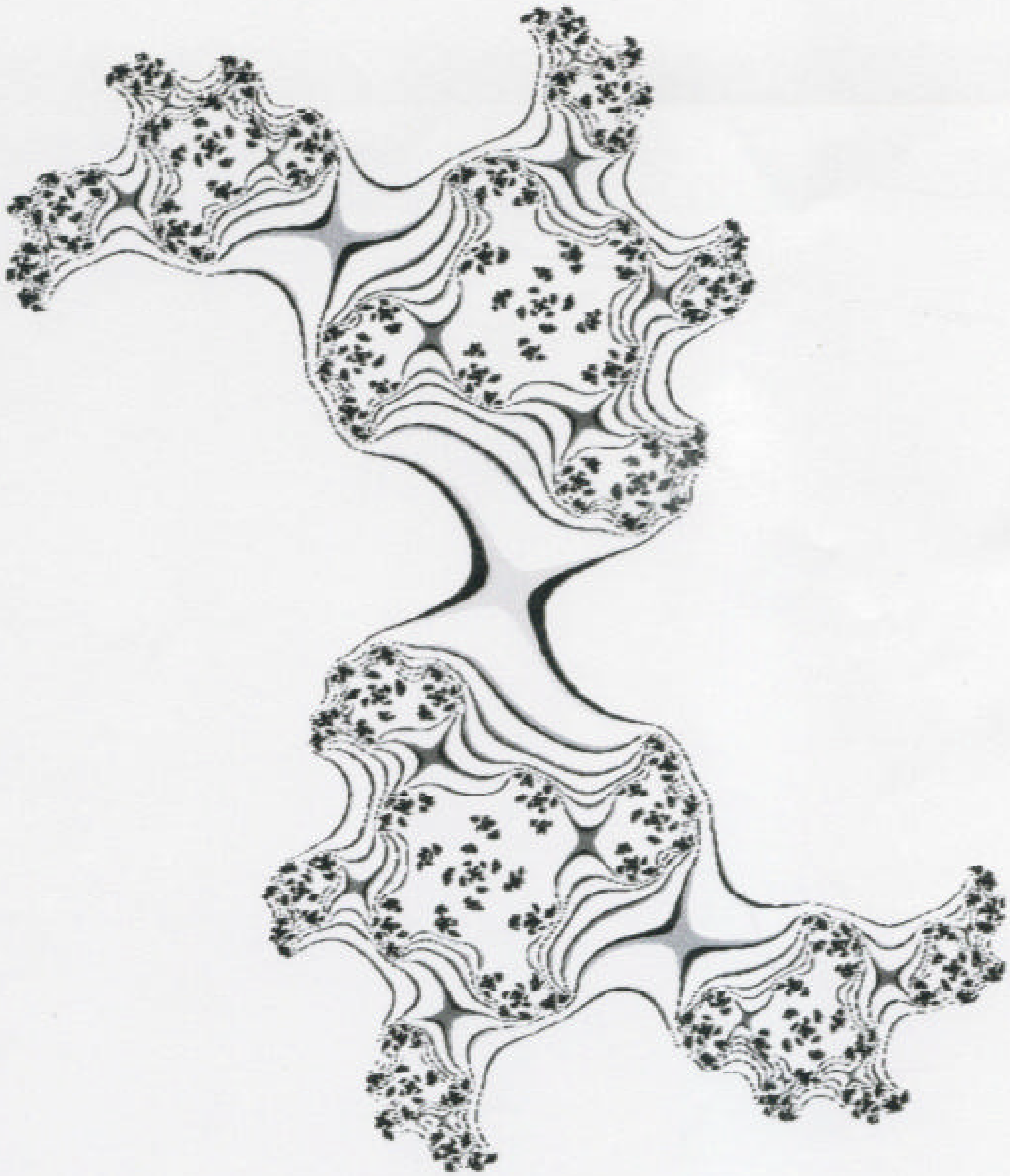
$$\mathrm{Dim}(K) \geq 1 + c_1 \varepsilon^2$$

(Harder, requires TSP)
Many natural examples.

32







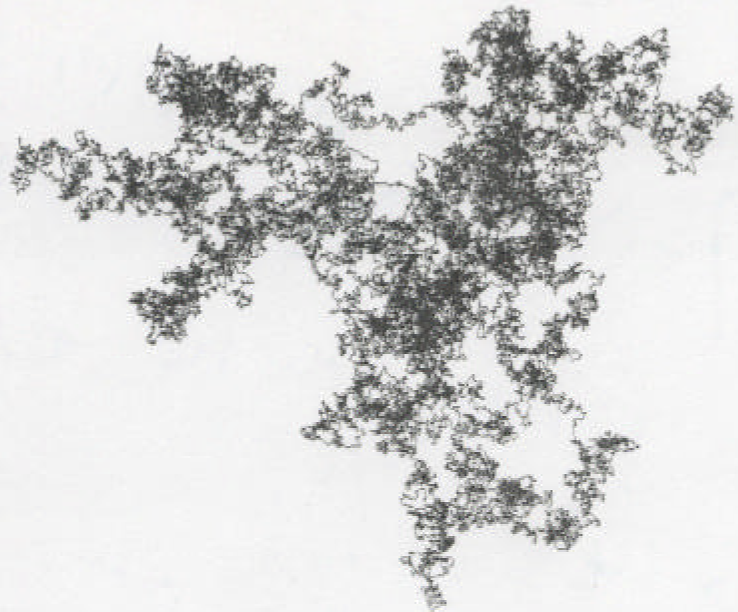


FIGURE 1. planar Brownian path

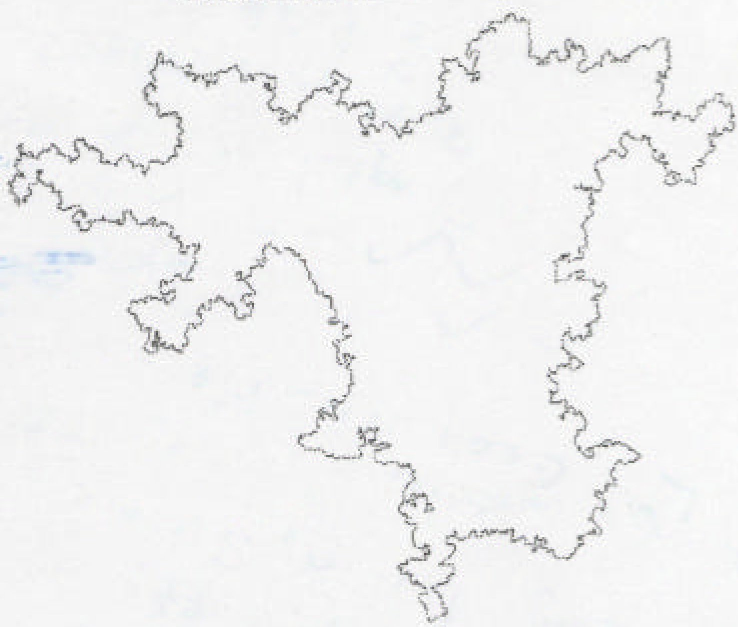


FIGURE 2. Brownian frontier

The Stochastic Loewner evolution (SLE)

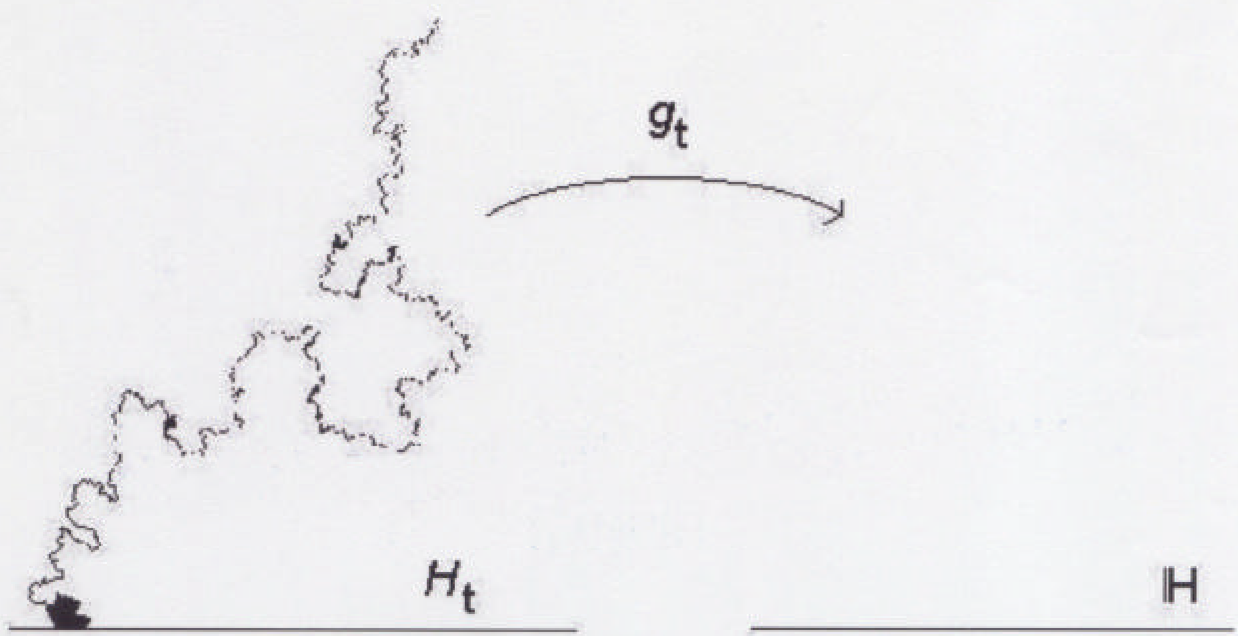
Chordal SLE_κ is a collection of random conformal maps

$$g_t : H_t \to \mathbb{H}$$

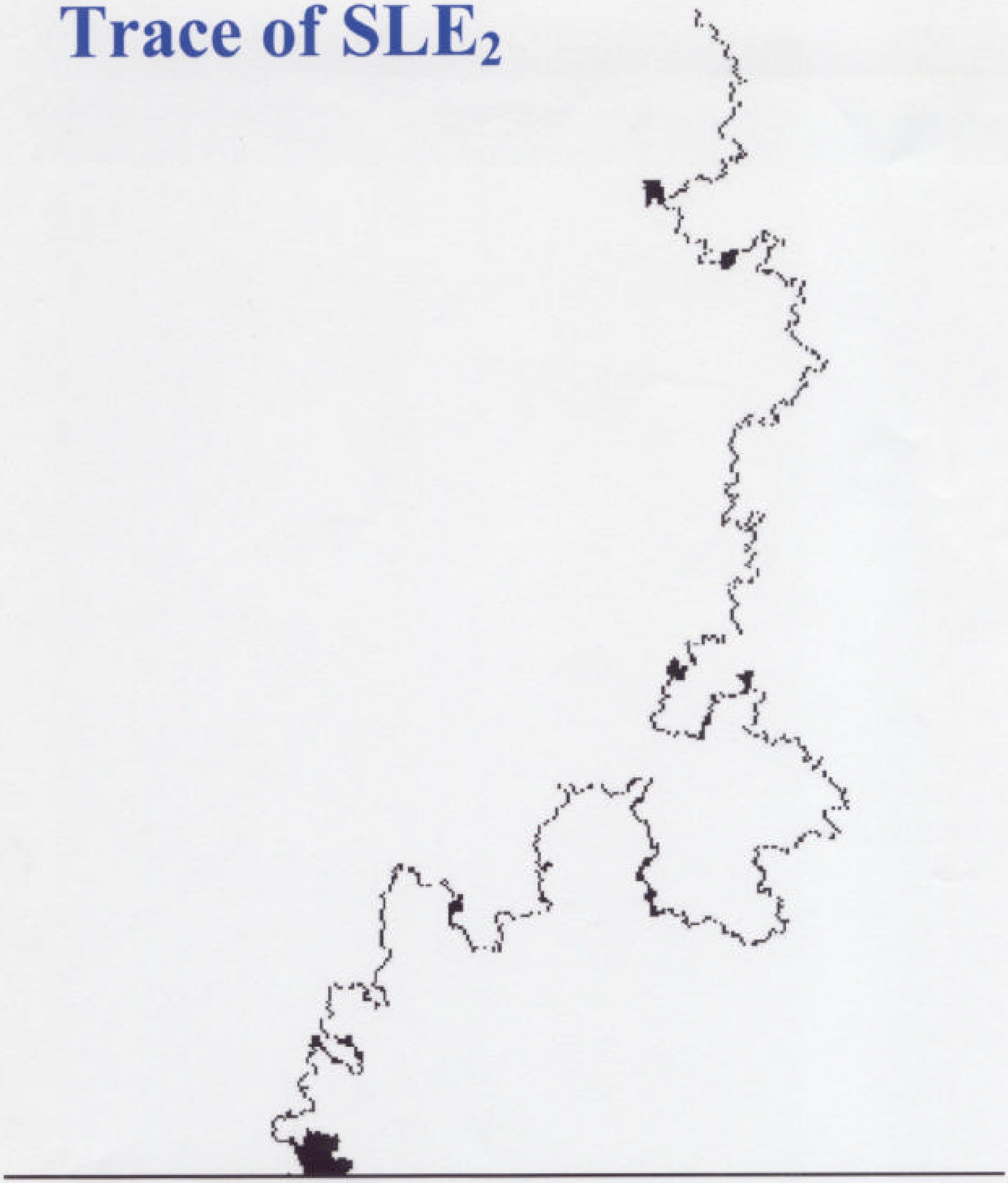
satisfying the differential equation

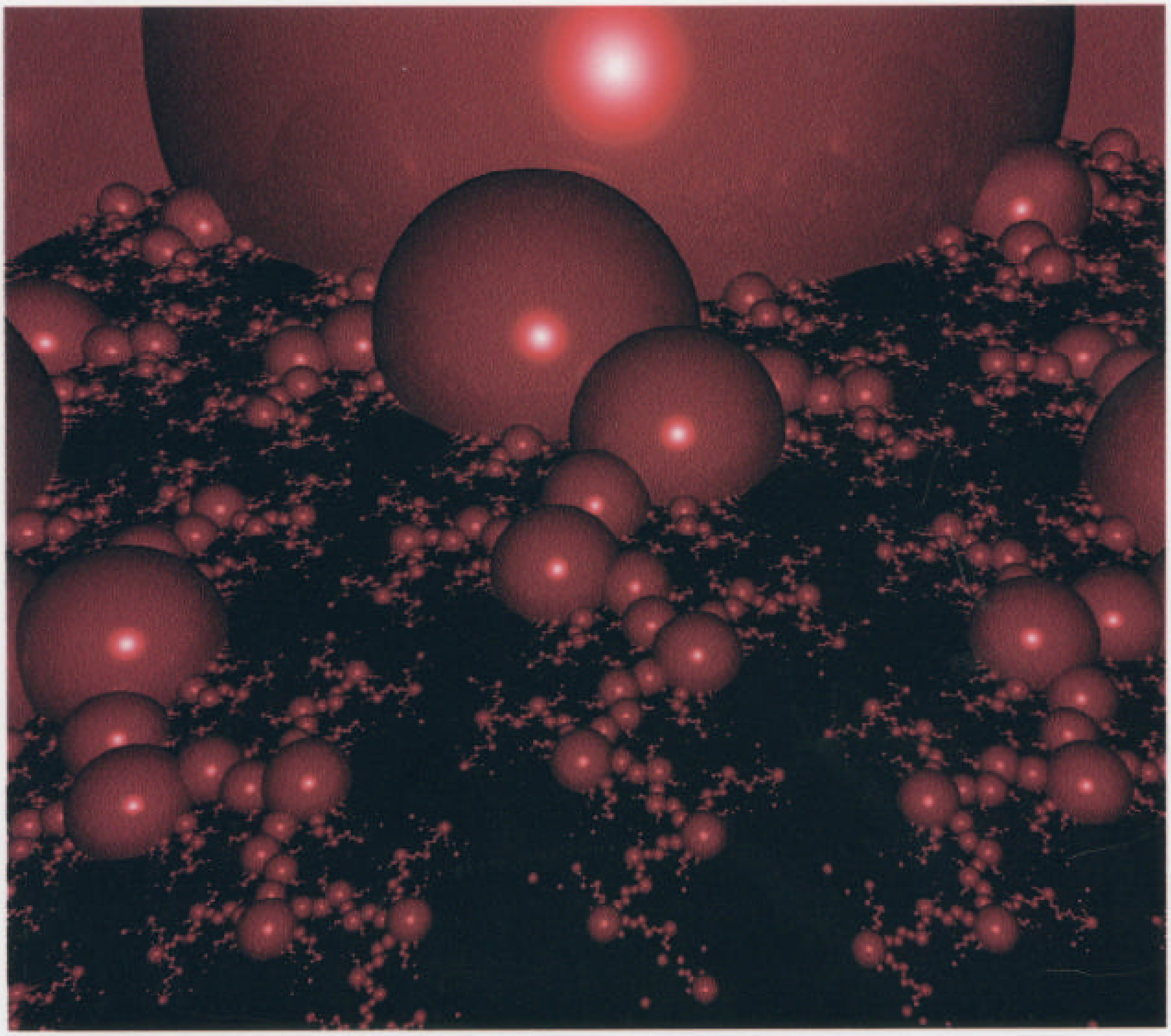
$$\partial_t g_t(z) = \frac{2}{g_t(z) - \sqrt{\kappa} B_t}, \quad g_0(z) = z,$$

where B_t is a Brownian motion.



Trace of SLE_2





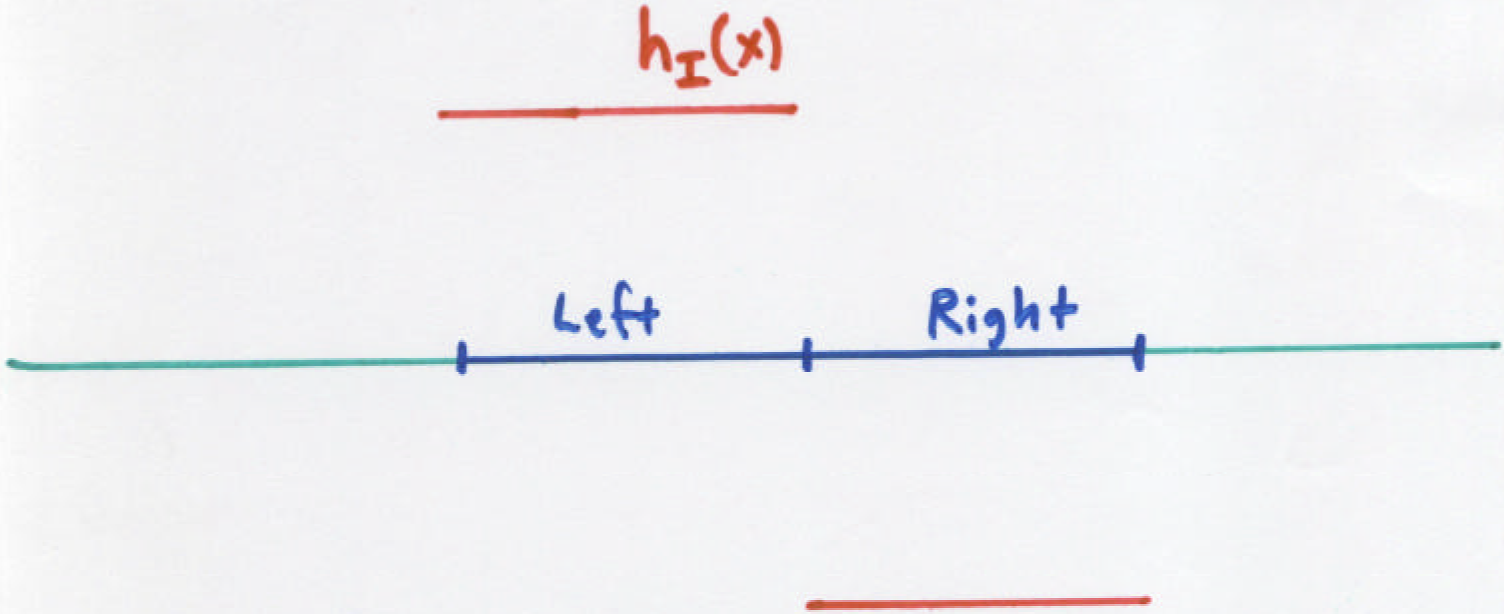
Jeff Brock

$$Hf(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(y)}{x-y}\,dy$$

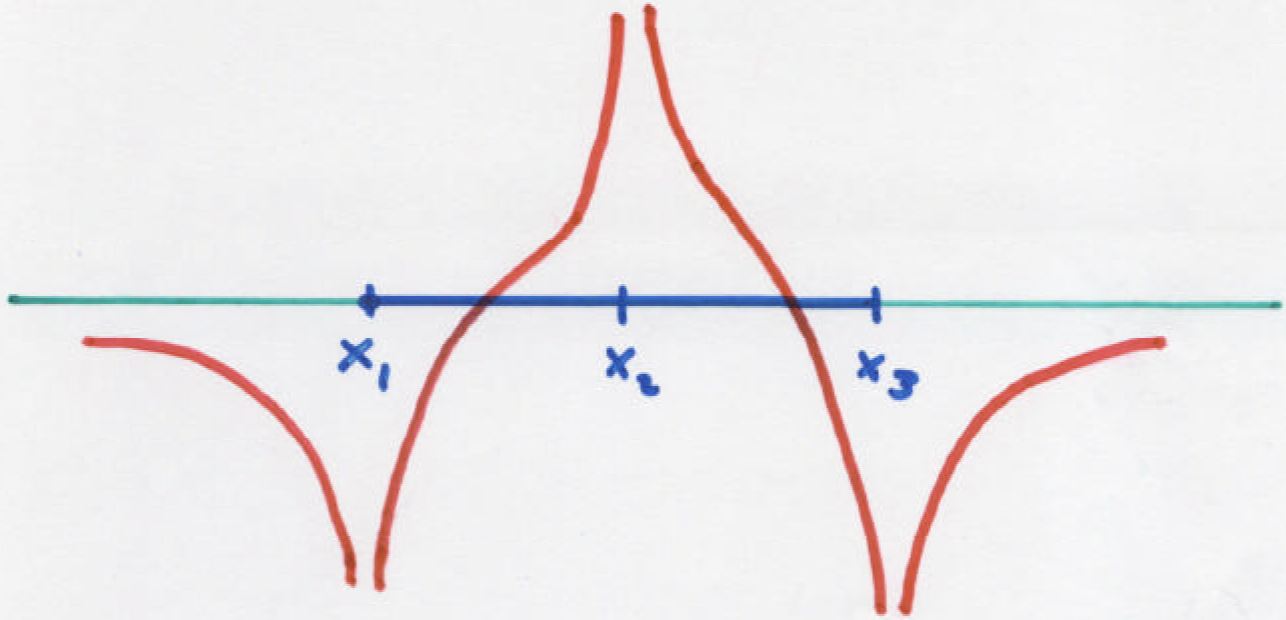
$$\Rightarrow \|Hf\|_{L^2} \leq c\|f\|_{L^2} \qquad \left(\stackrel{!}{=} \|f\|_{L^2}\right)$$

$$h_I(x) = \begin{cases} +\,\ell(I)^{-1/2} & \text{on Left} \\ -\,\ell(I)^{-1/2} & \text{on Right} \\ 0 & \text{off } I \end{cases}$$

(Haar)



$$\|f\|_{L^2}^2 = \sum_I |\langle f, h_I\rangle|^2$$



$$H\psi_I(x) = c\left(\log|x-x_1| + \log|x-x_3| - 2\log|x-x_2|\right)$$

By Calculus (!!)

$$\sum_J |\langle H\psi_I, \psi_J\rangle| \leq B \quad \text{rows}$$

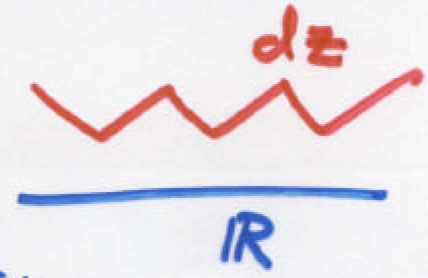
$$\sum_I |\langle H\psi_I, \psi_J\rangle| \leq B \quad \text{columns}$$

$$H \leftrightarrow (m_{I,J}) \text{ on } \ell^2$$

Schur's Lemma $\Longrightarrow$ Bounded on ℓ^2

Coifman, J, Semmes 1987

$\{h_I\} \longleftrightarrow dx$

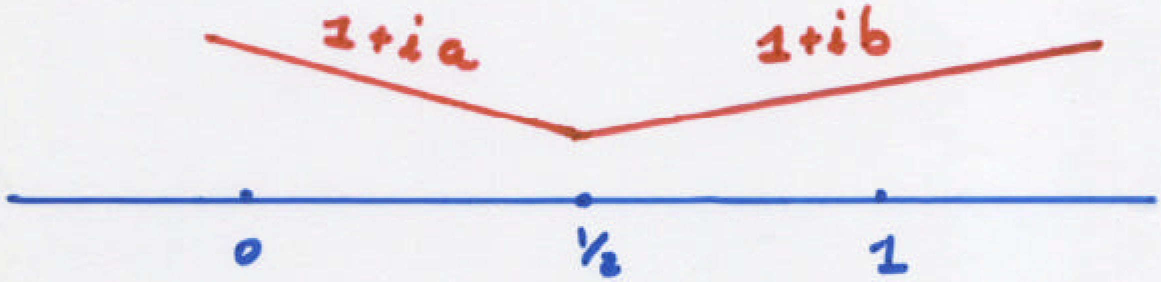


$\{\psi_I\} \longleftrightarrow (1+iA')dx$

$$\psi_I = \begin{cases} c_L & \text{on Left} \\ c_R & \text{on Right} \\ 0 & \text{off } I \end{cases}$$

$$\int \psi_I \, dz = 0 \quad \text{and} \quad \int \psi_I \psi_J \, dz = \delta_{I,J}$$

Example $I = [0,1]$



$$\psi = \left(\frac{1+ib}{1+ia}\right)^{1/2} \cdot c \quad \text{on } [0, 1/2]$$

$$\psi = -\left(\frac{1+ia}{1+ib}\right)^{1/2} \cdot c \quad \text{on } [1/2, 1]$$

$\{\psi_I\}$ is a FRAME

1. $$F = \sum_I \langle F, \psi_I \rangle_\Gamma \, \psi_I$$

$$\langle F, G \rangle_\Gamma = \int F G \, dz$$
no bar!

2. $$\|F\|_{L^2}^2 \sim \sum_I |\langle F, \psi_I \rangle|^2$$

And The Cauchy Integral is almost diagonal in this frame.

(Note B, C, R)

$$m_{I,J} = \langle \mathcal{C}(\psi_I), \psi_J \rangle_\Gamma \approx \delta_{I,J}$$

$$\mathcal{C} \longleftrightarrow (m_{I,J}) \qquad \ell^2 \to \ell^2$$

Row Sums: $\sum_J |m_{I,J}| \le B$

Column Sums: $\sum_I |m_{I,J}| \le B$

Schur's Lemma

Needs.

- Modeling and simulations of complex phenomena depending on many parameters require efficient computational representations (or transcriptions)of objects in high dimensions .
- A Mathematical/Algorithmic language for organization and structuring of complex natural phenomena,extending the traditional formulas.

(Coifman)

Geometry

Equals

Multi-Scale

S.V.D.

Needs.

- Modeling and simulations of complex phenomena depending on many parameters require efficient computational representations (or transcriptions)of objects in high dimensions .
- A Mathematical/Algorithmic language for organization and structuring of complex natural phenomena,extending the traditional formulas.

(Coifman)