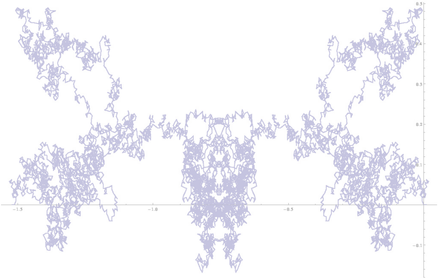
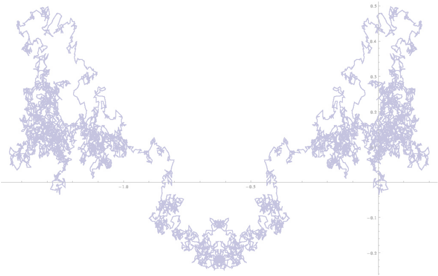
Random / not Random

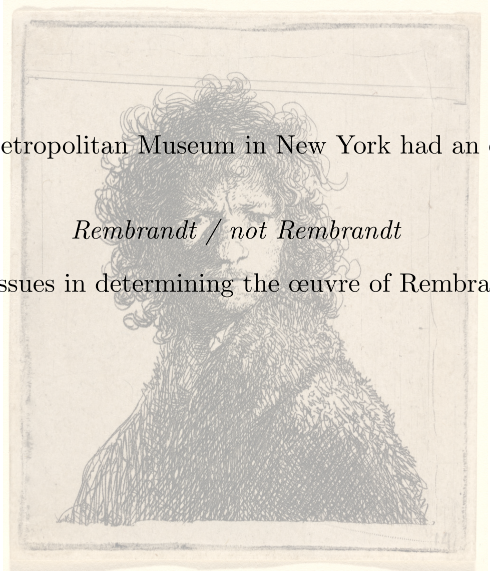




E. Kowalski
ETH Zürich

September 12, 2019
Heilbronn Conference

Random / not Random

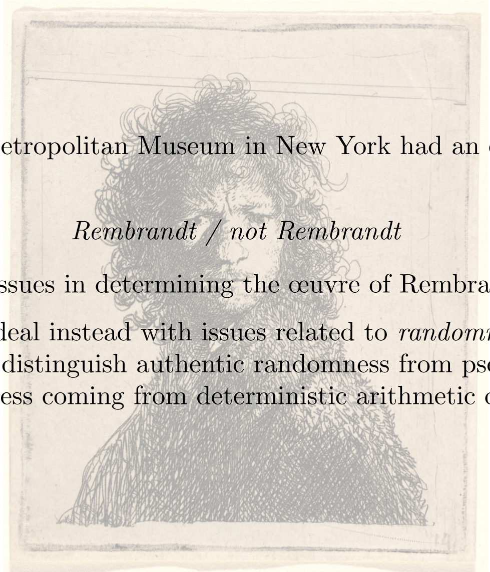


In 1995, the Metropolitan Museum in New York had an exhibition entitled

Rembrandt / not Rembrandt

exploring the issues in determining the œuvre of Rembrandt.

Random / not Random



In 1995, the Metropolitan Museum in New York had an exhibition entitled

Rembrandt / not Rembrandt

exploring the issues in determining the œuvre of Rembrandt.

This talk will deal instead with issues related to *randomness*, where we attempt to distinguish authentic randomness from pseudo- or quasi-randomness coming from deterministic arithmetic objects.

Values of the Euler function

Normalized Euler function: $$f(n) = \frac{\varphi(n)}{n} = \prod_{p|n}\Big(1 - \frac{1}{p}\Big).$$

Schoenberg (1928): if we look at integers $n \leq \mathrm{N}$ and let $\mathrm{N} \to +\infty$, the probability distribution of $f(n)$ converges in law to the infinite random product

$$\mathrm{F} = \prod_p \Big(1 - \frac{\mathrm{B}_p}{p}\Big)$$

where (B_p) are independent Bernoulli random variables with

$$\mathbf{P}(\mathrm{B}_p = 1) = \frac{1}{p}, \qquad \mathbf{P}(\mathrm{B}_p = 0) = 1 - \frac{1}{p}.$$

Values of the Euler function

Normalized Euler function: $$f(n) = \frac{\varphi(n)}{n} = \prod_{p|n}\Big(1 - \frac{1}{p}\Big).$$

Schoenberg (1928): if we look at integers $n \leq \mathrm{N}$ and let $\mathrm{N} \to +\infty$, the probability distribution of $f(n)$ converges in law to the infinite random product

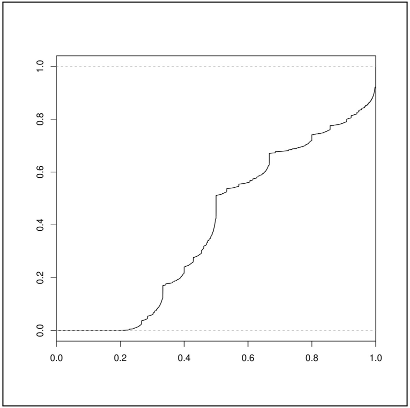
$$\mathrm{F} = \prod_{p}\Big(1 - \frac{\mathrm{B}_p}{p}\Big)$$

where (B_p) are independent Bernoulli random variables with

$$\mathbf{P}(\mathrm{B}_p = 1) = \frac{1}{p}, \qquad \mathbf{P}(\mathrm{B}_p = 0) = 1 - \frac{1}{p}.$$

In other words: $$\lim_{\mathrm{N}\to+\infty} \frac{1}{\mathrm{N}}|\{n \leq \mathrm{N} \mid \varphi(n) \leq \alpha n\}| = \mathbf{P}(\mathrm{F} \leq \alpha)$$

Values of the Euler function



Erdős (1939): the distribution is singular; $g(\alpha) = \mathbf{P}(\mathrm{F} \leq \alpha)$ is continuous, strictly increasing, and has $g'(\alpha) = 0$ for almost all α.

Kloosterman paths

Kloosterman sums: $$\mathrm{Kl}(a;p) = \frac{1}{\sqrt{p}} \sum_{1\leq x<p} \exp\Big(2i\pi \frac{ax+\bar{x}}{p}\Big)$$

(p prime, a coprime to p, $x\bar{x} \equiv 1 \pmod{p}$)

Kloosterman paths: continuous function $\mathcal{K}_p(a)\colon [0,1] \to \mathbf{C}$ linearly interpolating

$$\frac{j}{p-1} \mapsto \frac{1}{\sqrt{p}} \sum_{1\leq x\leq j} \exp\Big(2i\pi \frac{ax+\bar{x}}{p}\Big), \qquad 0 \leq j \leq p-1.$$

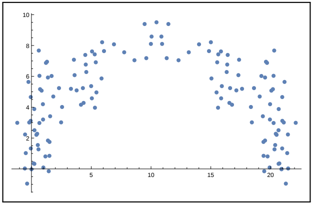
Kloosterman paths

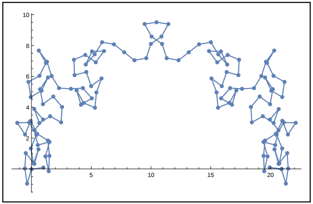
Kloosterman sums: $$\mathrm{Kl}(a;p) = \frac{1}{\sqrt{p}} \sum_{1\leq x<p} \exp\Big(2i\pi \frac{ax+\bar{x}}{p}\Big)$$

(p prime, a coprime to p, $x\bar{x} \equiv 1 \pmod{p}$)

Kloosterman paths: continuous function $\mathcal{K}_p(a)\colon [0,1] \to \mathbf{C}$ linearly interpolating

$$\frac{j}{p-1} \mapsto \frac{1}{\sqrt{p}} \sum_{1\leq x\leq j} \exp\Big(2i\pi \frac{ax+\bar{x}}{p}\Big), \qquad 0 \leq j \leq p-1.$$





Kloosterman paths

K. & Sawin (2016): as $p \to +\infty$, if we take a modulo p uniformly at random, the Kloosterman paths $\mathcal{K}_p(a)$ converge in law to the Fourier series

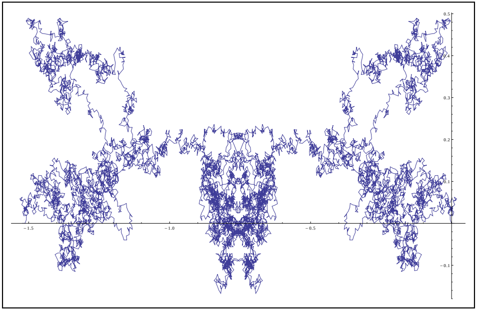
$$\mathrm{K}(t) = t\mathrm{X}_0 + \sum_{\substack{h \in \mathbf{Z} \\ h \neq 0}} \mathrm{X}_h \frac{\exp(2i\pi th) - 1}{2i\pi h}$$

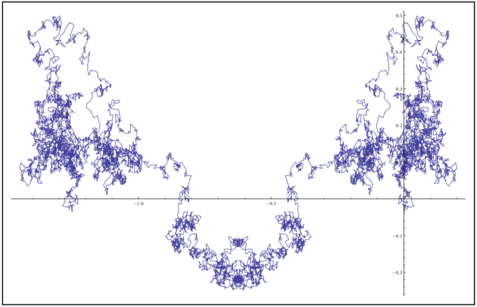
where $(\mathrm{X}_h)_{h \in \mathbf{Z}}$ are independent and distributed on $[-2, 2]$ according to the density

$$\frac{1}{\pi} \sqrt{1 - \frac{x^2}{4}} dx.$$

Picturesque randomness

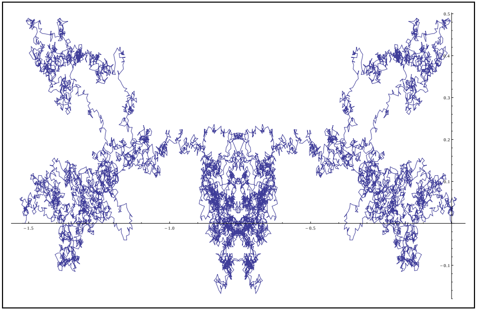
One of these two pictures is a sample of the “authentic” random Fourier series, the other one is a Kloosterman path.

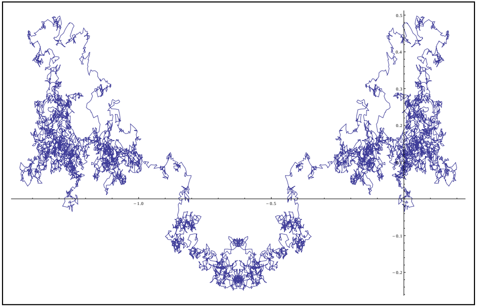




Picturesque randomness

One of these two pictures is a sample of the “authentic” random Fourier series, the other one is a Kloosterman path.





The “bat-like” shape (dixit Granville and Granville, *Prime suspects*) is due to the fact that the Fourier coefficients are purely imaginary.

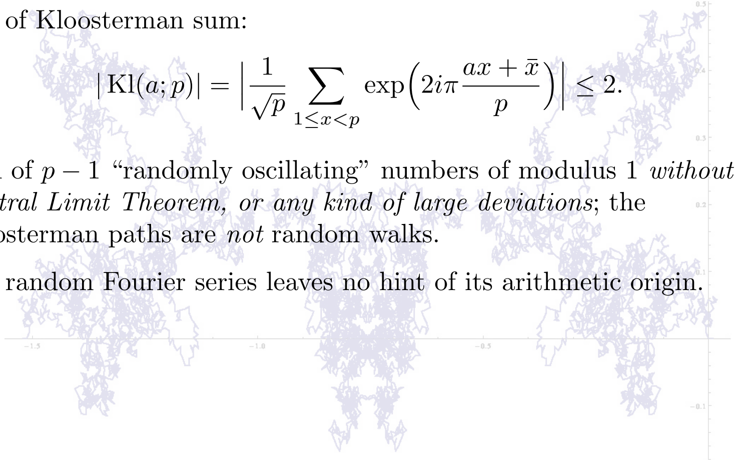
Kloosterman paths

Size of Kloosterman sum:

$$|\operatorname{Kl}(a;p)| = \Big|\frac{1}{\sqrt{p}}\sum_{1\leq x<p}\exp\Big(2i\pi\frac{ax+\bar{x}}{p}\Big)\Big| \leq 2.$$

Sum of $p-1$ "randomly oscillating" numbers of modulus 1 *without Central Limit Theorem, or any kind of large deviations*; the Kloosterman paths are *not* random walks.

The random Fourier series leaves no hint of its arithmetic origin.



Kloosterman paths

Size of Kloosterman sum:

$$|\,\mathrm{Kl}(a;p)| = \Big|\frac{1}{\sqrt{p}}\sum_{1\leq x<p}\exp\Big(2i\pi\frac{ax+\bar{x}}{p}\Big)\Big| \leq 2.$$

Sum of $p-1$ “randomly oscillating” numbers of modulus 1 *without Central Limit Theorem, or any kind of large deviations*; the Kloosterman paths are *not* random walks.

The random Fourier series leaves no hint of its arithmetic origin.

(Vexing) open question: does there exist a continuous function f in the support of K that is “space filling”?

Equivalently: is there a continuous space-filling curve f with Fourier coefficients of size $\mathrm{O}(1/h)$?

Where does the randomness come from?

Weil (1948): the full length Kloosterman sum $\mathrm{Kl}(a;p)$ is the *trace* of a matrix $\theta_{a,p} \in \mathrm{SU}_2(\mathbf{C})$; so $|\mathrm{Kl}(a;p)| \leq 2$.

Deligne / Katz (1988): these matrices are uniformly distributed, up to conjugacy, in $\mathrm{SU}_2(\mathbf{C})$.

These results are special cases of the general form of the Riemann Hypothesis over Finite Fields, due to Deligne (the most important result in number theory of the 20th century).

Where does the randomness come from?

Weil (1948): the full length Kloosterman sum $\mathrm{Kl}(a;p)$ is the *trace* of a matrix $\theta_{a,p} \in \mathrm{SU}_2(\mathbf{C})$; so $|\mathrm{Kl}(a;p)| \leq 2$.

Deligne / Katz (1988): these matrices are uniformly distributed, up to conjugacy, in $\mathrm{SU}_2(\mathbf{C})$.

These results are special cases of the general form of the Riemann Hypothesis over Finite Fields, due to Deligne (the most important result in number theory of the 20th century).

The (pseudo-)randomness implications of this remarkable result are certainly still very far from being exhausted.

A classical question of analysis

Consider $f\colon \mathbf{R} \to \mathbf{C}$ continuous with period 1.

Parseval formula: $$\sum_{n\in\mathbf{Z}} |c_n(f)|^2 = \int_0^1 |f(x)|^2 dx < +\infty$$

where

$$c_n(f) = \int_0^1 f(x)\exp(-2i\pi nx)dx.$$

Question: is the exponent 2 best possible? Could there exist $\delta > 0$ such that $\sum_{n\in\mathbf{Z}} |c_n(f)|^{2-\delta} < +\infty$ for any continuous f?

A classical question of analysis

Legendre symbol: p prime, $\left(\frac{n}{p}\right) = \begin{cases} 0 & p \mid n \\ 1 & n \text{ a square mod } p \\ -1 & n \text{ not a square mod } p \end{cases}$

Carleman (1917): $$f(x) = \sum_{k \geq 1} \frac{1}{k^2} f_{p_k}(x),$$

$$p_k \equiv 1 \pmod 4 \text{ such that } p_k > 4p_{k-1}^2,$$

$$f_p(x) = \frac{2}{p^{1/2}} \sum_{n=1}^{p-1} \left(1 - \frac{n}{p}\right) \left(\frac{n}{p}\right) \cos(2\pi n x).$$

Then $$\sum_{n \in \mathbf{Z}} |c_n(f)|^{2-\delta} = +\infty \text{ for any } \delta > 0.$$

A classical question of analysis

Legendre symbol: p prime, $\left(\frac{n}{p}\right) = \begin{cases} 0 & p \mid n \\ 1 & n \text{ a square mod } p \\ -1 & n \text{ not a square mod } p \end{cases}$

Carleman (1917): $$f(x) = \sum_{k \geq 1} \frac{1}{k^2} f_{p_k}(x),$$

$$p_k \equiv 1 \pmod 4 \text{ such that } p_k > 4p_{k-1}^2,$$
$$f_p(x) = \frac{2}{p^{1/2}} \sum_{n=1}^{p-1} \left(1 - \frac{n}{p}\right)\left(\frac{n}{p}\right) \cos(2\pi n x).$$

Then $$\sum_{n \in \mathbf{Z}} |c_n(f)|^{2-\delta} = +\infty \text{ for any } \delta > 0.$$

But also $$g(x) = \sum_{n \geq 2} \frac{e^{2i\pi n \log n}}{\sqrt{n}(\log n)^2} e^{2i\pi n x}$$ (cf Zygmund, p. 199).

A much more difficult question: ultraflat polynomials

$$f_{\mathrm{N}}(x) = \sum_{0 \le m \le \mathrm{N}} a(m) \exp(2i\pi m x), \qquad |a(m)| = 1$$

Question (Erdős 1957, Littlewood 1966): is it possible to find f_{N} such that $|f_{\mathrm{N}}(x)| = \sqrt{\mathrm{N}}(1 + o(1))$ for all x?

A much more difficult question: ultraflat polynomials

$$f_{\mathrm{N}}(x) = \sum_{0 \le m \le \mathrm{N}} a(m) \exp(2i\pi m x), \qquad |a(m)| = 1$$

Question (Erdős 1957, Littlewood 1966): is it possible to find f_{N} such that $|f_{\mathrm{N}}(x)| = \sqrt{\mathrm{N}}(1 + o(1))$ for all x?

Kahane (1980): probabilistic construction.

Bombieri–Bourgain (1999): explicit arithmetic construction with
$|f_{\mathrm{N}}(x)| = \sqrt{\mathrm{N}} + \mathrm{O}(\mathrm{N}^{1/2-1/18})$.

A much more difficult question: ultraflat polynomials

$$f_{\mathrm{N}}(x) = \sum_{0\leq m\leq \mathrm{N}} a(m)\exp(2i\pi mx), \qquad |a(m)| = 1$$

Question (Erdős 1957, Littlewood 1966): is it possible to find f_{N} such that $|f_{\mathrm{N}}(x)| = \sqrt{\mathrm{N}}(1+o(1))$ for all x?

Kahane (1980): probabilistic construction.

Bombieri–Bourgain (1999): explicit arithmetic construction with $|f_{\mathrm{N}}(x)| = \sqrt{\mathrm{N}} + \mathrm{O}(\mathrm{N}^{1/2-1/18})$.

The proof involves again the Riemann Hypothesis over Finite Fields.

Pseudo-random functions in the sense of Gowers

Gowers norms: p prime, $f\colon \mathbf{Z}/p\mathbf{Z} \to \mathbf{C}$, $k \geq 0$

$$\|f\|_0 = \frac{1}{p}\Big|\sum_{x\in\mathbf{Z}/p\mathbf{Z}} f(x)\Big|, \quad \|f\|_{k+1}^{2^{k+1}} = \frac{1}{p}\sum_{h\in\mathbf{Z}/p\mathbf{Z}} \Big\|\Big(x \mapsto f(x)\overline{f(x+h)}\Big)\Big\|_k^{2^k}.$$

Pseudo-random functions in the sense of Gowers

Gowers norms: p prime, $f\colon \mathbf{Z}/p\mathbf{Z} \to \mathbf{C}$, $k \geq 0$

$$\|f\|_0 = \frac{1}{p}\Big|\sum_{x \in \mathbf{Z}/p\mathbf{Z}} f(x)\Big|, \quad \|f\|_{k+1}^{2^{k+1}} = \frac{1}{p}\sum_{h \in \mathbf{Z}/p\mathbf{Z}} \Big\|\Big(x \mapsto f(x)\overline{f(x+h)}\Big)\Big\|_k^{2^k}.$$

Exercise **(Tao–Vu)**. If f is “random” then $\|f\|_k^{2^k} = \mathrm{O}(p^{-1})$.

Pseudo-random functions in the sense of Gowers

Gowers norms: p prime, $f\colon \mathbf{Z}/p\mathbf{Z} \to \mathbf{C}$, $k \geq 0$

$$\|f\|_0 = \frac{1}{p}\Big|\sum_{x\in\mathbf{Z}/p\mathbf{Z}} f(x)\Big|, \quad \|f\|_{k+1}^{2^{k+1}} = \frac{1}{p}\sum_{h\in\mathbf{Z}/p\mathbf{Z}} \Big\|\Big(x \mapsto f(x)\overline{f(x+h)}\Big)\Big\|_k^{2^k}.$$

Exercise **(Tao–Vu)**. If f is "random" then $\|f\|_k^{2^k} = \mathrm{O}(p^{-1})$.

Fouvry, K., Michel (2013): for $f(a) = \mathrm{Kl}(a;p)$, we have

$$\|f\|_k^{2^k} \leq 20^{(k+1)2^k}/p.$$

The approximation property

Question: does there exist f continuous on $[0,1]^2$ such that

$$\int_0^1 f(x,t)f(t,y)dt = 0, \qquad (x,y) \in [0,1]^2, \qquad \int_0^1 f(t,t)dt \neq 0 \quad ?$$

The approximation property

Question: does there exist f continuous on $[0,1]^2$ such that

$$\int_0^1 f(x,t)f(t,y)dt = 0, \qquad (x,y) \in [0,1]^2, \qquad \int_0^1 f(t,t)dt \neq 0 \quad ?$$

Grothendieck (1955): Yes $\Longleftrightarrow$ *some* Banach space E does not have the *approximation property* (the identity on E cannot be approximated uniformly on compact sets by finite rank operators).

Enflo (1973): constructs spaces without the approximation property, so the answer is *Yes*. Davie has a different probabilistic construction.

The approximation property

Question: does there exist f continuous on $[0,1]^2$ such that

$$\int_0^1 f(x,t)f(t,y)dt = 0, \qquad (x,y) \in [0,1]^2, \qquad \int_0^1 f(t,t)dt \neq 0 \quad ?$$

Grothendieck (1955): Yes $\Longleftrightarrow$ *some* Banach space E does not have the *approximation property* (the identity on E cannot be approximated uniformly on compact sets by finite rank operators).

Enflo (1973): constructs spaces without the approximation property, so the answer is *Yes*. Davie has a different probabilistic construction.

Key probabilistic requirement. For $k \geq 1$, find $(\alpha_i)_{1\leq i\leq 3\cdot 2^k}$ in $\{-2,1\}$ with sum 0 such that

$$\Big|\sum_i \alpha_i \chi(i)\Big| \leq \mathrm{C}(k+1)^{1/2} 2^{k/2}$$

for all characters of $\mathbf{Z}/3\cdot 2^k\mathbf{Z}$.

Expander graphs

(Barzdin–Kolmogorov 1967; Bassalygo–Pinsker 1973):
$(\Gamma_n)_{n\geq 1}$, sequence of finite d-regular graphs with $|\Gamma_n| \to +\infty$.

δ-expander: for all $n \geq 1$, all $\emptyset \neq \mathrm{X} \subset \Gamma_n$, we have

$$(\star) \qquad \frac{|\{\text{edges of } \Gamma_n \text{ joining X to } \Gamma_n - \mathrm{X}\}|}{\min(|\mathrm{X}|, |\Gamma_n - \mathrm{X}|)} \geq \delta > 0.$$

Expander graphs

(Barzdin–Kolmogorov 1967; Bassalygo–Pinsker 1973):
$(\Gamma_n)_{n\geq 1}$, sequence of finite d-regular graphs with $|\Gamma_n| \to +\infty$.

δ-expander: for all $n \geq 1$, all $\emptyset \neq \mathrm{X} \subset \Gamma_n$, we have

$$(\star) \qquad \frac{|\{\text{edges of } \Gamma_n \text{ joining X to } \Gamma_n - \mathrm{X}\}|}{\min(|\mathrm{X}|, |\Gamma_n - \mathrm{X}|)} \geq \delta > 0.$$

Existence first proved by probabilistic methods: a “random” d-regular graph with n vertices has probability at least $\geq \delta_d > 0$ of satisfying $(\star)$

Ramanujan graphs

Condition ($\star$) equivalent to

$$\frac{\sum_{x\sim y}|f(x)-f(y)|^2}{\sum_{x\in\Gamma_n}|f(x)|^2}\geq\delta'>0\qquad\Big(f\colon\Gamma_n\to\mathbf{C},\quad\sum_{x\in\Gamma_n}f(x)=0\Big)$$

(for some δ' depending on δ).

Alon–Boppana (1986): best possible δ' is $2\sqrt{d-1}$.

Ramanujan graph (**Lubotzky–Phillips–Sarnak**, 1988): a graph such that $\delta'=2\sqrt{d-1}$ is possible.

Ramanujan graphs

Condition ($\star$) equivalent to

$$\frac{\sum_{x\sim y} |f(x)-f(y)|^2}{\sum_{x\in\Gamma_n} |f(x)|^2} \geq \delta' > 0 \qquad \Big(f\colon \Gamma_n \to \mathbf{C}, \quad \sum_{x\in\Gamma_n} f(x) = 0\Big)$$

(for some δ' depending on δ).

Alon–Boppana (1986): best possible δ' is $2\sqrt{d-1}$.

Ramanujan graph (**Lubotzky–Phillips–Sarnak**, 1988): a graph such that $\delta' = 2\sqrt{d-1}$ is possible.

Lubotzky–Phillips–Sarnak construct explicit Ramanujan graphs for $d = p+1$; essential tools are results of Deligne (not only the Riemann Hypothesis).

Ramanujan graphs

Condition $(\star)$ equivalent to

$$\frac{\sum_{x\sim y}|f(x)-f(y)|^2}{\sum_{x\in\Gamma_n}|f(x)|^2}\geq\delta'>0\qquad\Big(f\colon\Gamma_n\to\mathbf{C},\quad\sum_{x\in\Gamma_n}f(x)=0\Big)$$

(for some δ' depending on δ).

Alon–Boppana (1986): best possible δ' is $2\sqrt{d-1}$.

Ramanujan graph (**Lubotzky–Phillips–Sarnak**, 1988): a graph such that $\delta'=2\sqrt{d-1}$ is possible.

Lubotzky–Phillips–Sarnak construct explicit Ramanujan graphs for $d=p+1$; essential tools are results of Deligne (not only the Riemann Hypothesis).

Marcus–Spielman–Srivastava (2015): probabilistic construction of bipartite Ramanuajan graphs (but for all $d\geq 3$).

The Banach–Mazur compact spaces

Banach–Mazur distance: $n \geq 1$ integer; E, F complex Banach spaces of dimension n; $\log d_{\mathrm{BM}}(\mathrm{E}, \mathrm{F}) = \min_{u\colon \mathrm{E} \simeq \mathrm{F}} \|u\| \, \|u^{-1}\|$.

The Banach–Mazur compact spaces

Banach–Mazur distance: $n \geq 1$ integer; E, F complex Banach spaces of dimension n; $\log d_{\mathrm{BM}}(\mathrm{E}, \mathrm{F}) = \min_{u\colon \mathrm{E} \simeq \mathrm{F}} \|u\| \, \|u^{-1}\|$.

Banach–Mazur spaces: BM_n = space of Banach spaces of dimension n, up to isometry, with distance d_{BM}; it is a compact metric space.

The Banach–Mazur compact spaces

Banach–Mazur distance: $n \geq 1$ integer; E, F complex Banach spaces of dimension n; $\log d_{\mathrm{BM}}(\mathrm{E}, \mathrm{F}) = \min_{u\colon \mathrm{E} \simeq \mathrm{F}} \|u\| \, \|u^{-1}\|$.

Banach–Mazur spaces: BM_n = space of Banach spaces of dimension n, up to isometry, with distance d_{BM}; it is a compact metric space.

Gluskin (1981): the diameter of BM_n is of order about n for $n \to +\infty$.

The Banach–Mazur compact spaces

Banach–Mazur distance: $n \geq 1$ integer; E, F complex Banach spaces of dimension n; $\log d_{\mathrm{BM}}(\mathrm{E}, \mathrm{F}) = \min_{u\colon \mathrm{E} \simeq \mathrm{F}} \|u\| \, \|u^{-1}\|$.

Banach–Mazur spaces: BM_n = space of Banach spaces of dimension n, up to isometry, with distance d_{BM}; it is a compact metric space.

Gluskin (1981): the diameter of BM_n is of order about n for $n \to +\infty$.

Probabilistic construction: for *random* vectors $\mathrm{X} = (\mathrm{X}_k)$ on the euclidean unit sphere of $\mathbf{C}^n$, define E_{X} by the norm

$$\|x\|_{\mathrm{X}} = \inf \Big\{ \sum_k |\lambda_k| \;\Big|\; x = \sum_k \lambda_k \mathrm{X}_k \Big\}.$$

With high probability, $d_{\mathrm{BM}}(\mathrm{E}_{\mathrm{X}}, \mathrm{E}_{\widetilde{\mathrm{X}}}) \approx n$.

A conjecture about Banach–Mazur spaces

Attempt at derandomization of Gluskin's construction: take $n = p$, identify $\mathbf{C}^p$ with functions $\mathbf{Z}/p\mathbf{Z} \to \mathbf{C}$, and take families X and $\widetilde{\mathrm{X}}$ of functions

$$f(x) = \exp\Big(2i\pi \frac{\mathrm{P}(x)}{p}\Big), \qquad \widetilde{f}(x) = \Big(\frac{\mathrm{Q}(x)}{p}\Big)$$

where P and Q are polynomials of bounded degree $d \geq 2$.

Question. Is is true that $d_{\mathrm{BM}}(\mathrm{E}_{\mathrm{X}}, \mathrm{E}_{\widetilde{\mathrm{X}}}) \approx p$?

A conjecture about Banach–Mazur spaces

Attempt at derandomization of Gluskin's construction: take $n = p$, identify $\mathbf{C}^p$ with functions $\mathbf{Z}/p\mathbf{Z} \to \mathbf{C}$, and take families X and $\widetilde{\mathrm{X}}$ of functions

$$f(x) = \exp\Big(2i\pi \frac{\mathrm{P}(x)}{p}\Big), \qquad \widetilde{f}(x) = \Big(\frac{\mathrm{Q}(x)}{p}\Big)$$

where P and Q are polynomials of bounded degree $d \geq 2$.

Question. Is is true that $d_{\mathrm{BM}}(\mathrm{E}_{\mathrm{X}}, \mathrm{E}_{\widetilde{\mathrm{X}}}) \approx p$?

These spaces appear in any case quite naturally in many results of Fouvry, K., Michel; do they have special properties as Banach spaces?