Mixing times and representation theory

Lucas Teyssier

PIMS-CNRS-Simons postdoc at UBC

PIMS PDF Seminar Series
6 December 2023

Example 1: riffle shuffle

Credit: Will Roya, Playingcarddeck.com

♠ : Motivation: different ways to mix a deck of cards

Example 1: riffle shuffle

Credit: Will Roya, Playingcarddeck.com

Example 2: smooshing

Credit: Carson Ford, medium.com

♠ : Motivation: different ways to mix a deck of cards

Example 1: riffle shuffle

Credit: Will Roya, Playingcarddeck.com

Example 2: smooshing

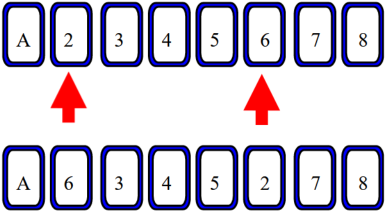
Credit: Carson Ford, medium.com

Question: how long does it take to mix?

♠ : The random transposition shuffle

Method :

- Pick two cards uniformly and independently;
- If different, interchange them;
- If they are the same card, do nothing.



Credit: Elchanan Mossel

Method :

- Pick two cards uniformly and independently;
- If different, interchange them;
- If they are the same card, do nothing.

Interpretation :

- Random walk on $\mathfrak{S}_n$ with

$$P(\sigma,\sigma\tau) = \mu_n(\tau) = \begin{cases} 1/n & \text{if } \tau = id \\ 2/n^2 & \text{if } \tau \text{ is a transp.} \end{cases}$$

P : transition matrix
μ_n : increment measure.

♠ : The random transposition shuffle

Method :

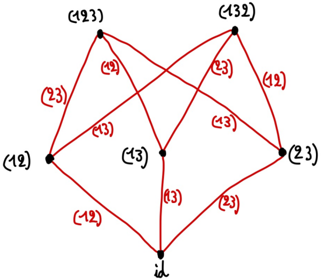
- Pick two cards uniformly and independently;
- If different, interchange them;
- If they are the same card, do nothing.

Interpretation :

- Random walk on $\mathfrak{S}_n$ with

$$P(\sigma, \sigma\tau) = \mu_n(\tau) = \begin{cases} 1/n & \text{if } \tau = id \\ 2/n^2 & \text{if } \tau \text{ is a transp.} \end{cases}$$

P : transition matrix
μ_n : increment measure.



Cayley graph for $n = 3$

♠ : Distance to stationarity

Question : in which sense do we converge to uniformity?

♠ : Distance to stationarity

Question : in which sense do we converge to uniformity?

$\nu_n(t)$: distribution of the walk after t steps.

Définition

Distance to stationarity after t steps :

$$\mathrm{d}_n(t) := \mathrm{d}_{\mathrm{TV}}(\nu_n(t), \mathrm{Unif}_n).$$

where for probability measures μ and ν on $\mathfrak{S}_n$,

$$\mathrm{d}_{\mathrm{TV}}(\mu,\nu) = \max_{A \subset \mathfrak{S}_n} \left|\mu(A) - \nu(A)\right| = \frac{1}{2}\,\mathrm{d}_1(\mu,\nu).$$

♠ : Distance to stationarity

Question : in which sense do we converge to uniformity?

$\nu_n(t)$: distribution of the walk after t steps.

Définition

Distance to stationarity after t steps :

$$\mathrm{d}_n(t) := \mathrm{d}_{\mathrm{TV}}(\nu_n(t), \mathrm{Unif}_n).$$

where for probability measures μ and ν on $\mathfrak{S}_n$,

$$\mathrm{d}_{\mathrm{TV}}(\mu,\nu) = \max_{A\subset\mathfrak{S}_n} \left|\mu(A) - \nu(A)\right| = \frac{1}{2}\,\mathrm{d}_1(\mu,\nu).$$

Question : How large should be take t so that $d_n(t) \approx 0$?

♠ : Cutoff for random transpositions

Theorem (Diaconis and Shahshahani, 1981)

It takes $\frac{1}{2}n\ln(n)$ *steps to mix a deck of n cards by random transpositions. For every* $0 < \epsilon < 1$*,*

$$d_n\left((1-\epsilon)\frac{1}{2}n\ln(n)\right) \xrightarrow[n\to+\infty]{} 1 \quad \& \quad d_n\left((1+\epsilon)\frac{1}{2}n\ln(n)\right) \xrightarrow[n\to+\infty]{} 0$$

That is what is called the **cutoff phenomenon**.

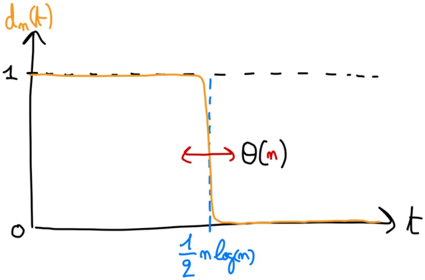
♠ : Cutoff for random transpositions

Theorem (Diaconis and Shahshahani, 1981)

It takes $\frac{1}{2} n \ln(n)$ *steps to mix a deck of n cards by random transpositions. For every* $0 < \epsilon < 1$,

$$d_n\left((1-\epsilon)\frac{1}{2}n\ln(n)\right) \xrightarrow[n\to+\infty]{} 1 \quad \& \quad d_n\left((1+\epsilon)\frac{1}{2}n\ln(n)\right) \xrightarrow[n\to+\infty]{} 0$$

That is what is called the **cutoff phenomenon**.
More precisely, it takes $\frac{1}{2} n \ln(n) + \Theta(n)$ steps to mix.



♠ : Mixing time of the riffle shuffle

Question: for $n = 52$ cards. How many shuffles do we need?

♠ : Mixing time of the riffle shuffle

Question: for $n = 52$ cards. How many shuffles do we need?

Answer: About 7.

TABLE 1
Total variation distance for m shuffles of 52 cards

m	1	2	3	4	5	6	7	8	9	10
$\|Q^m - U\|$	1.000	1.000	1.000	1.000	0.924	0.614	0.334	0.167	0.085	0.043

♠ : Mixing time of the riffle shuffle

Question: for $n = 52$ cards. How many shuffles do we need?

Answer: About 7.

296 D. BAYER AND P. DIACONIS

TABLE 1
Total variation distance for m shuffles of 52 cards

m	1	2	3	4	5	6	7	8	9	10
$\|Q^m - U\|$	1.000	1.000	1.000	1.000	0.924	0.614	0.334	0.167	0.085	0.043


Science Times
The New York Times
Giant Telescope Prepares to Take Its First Look Skyward
In Shuffling Cards, 7 Is Winning Number
Wide Beliefs On Depression In Women Contradicted

New York Times, 9 Jan. 1990

Question: and as the number of cards diverges?

♠ : Mixing time of the riffle shuffle

Question: and as the number of cards diverges?

Answer: Aldous 1986: Cutoff at $\frac{3}{2}\log_2 n$.

♠ : Mixing time of the riffle shuffle

Question: and as the number of cards diverges?

Answer: Aldous 1986: Cutoff at $\frac{3}{2}\log_2 n$.

Answer: Bayer–Diaconis 1992: Precise estimates for $n = 52$, and cutoff profile:

Theorem (Bayer–Diaconis, 1992)

For the riffle shuffle, we have for every $c \in \mathbb{R}$,

$$d_n\left(\frac{3}{2}\log_2(n) + c\right) \xrightarrow[n\to+\infty]{} p(c) := \mathrm{d}_{\mathrm{TV}}\left(\mathcal{N}(0,1), \mathcal{N}\left(\frac{2^{-c}}{2\sqrt{3}}, 1\right)\right).$$

(Written up to integer parts.)

♠ : Cutoff profile for random transpositions

Question: and for transpositions, can we find the profile?
Question asked by N. Berestycki at an AIM workshop in 2016.

♠ : Cutoff profile for random transpositions

Question: and for transpositions, can we find the profile?
Question asked by N. Berestycki at an AIM workshop in 2016.

Theorem (T., 2020)

For random transpositions, we have for every $c \in \mathbb{R}$,

$$d_n\left(\frac{1}{2}n\ln(n) + cn\right) \xrightarrow[n\to+\infty]{} p(c) := \mathrm{d_{TV}}\left(\mathrm{Poiss}\left(1+e^{-2c}\right), \mathrm{Poiss}(1)\right).$$

(Written up to integer parts.)

♠ : Cutoff profile for random transpositions

Question: and for transpositions, can we find the profile?
Question asked by N. Berestycki at an AIM workshop in 2016.

Theorem (T., 2020)

For random transpositions, we have for every $c \in \mathbb{R}$,

$$d_n\left(\frac{1}{2}n\ln(n) + cn\right) \xrightarrow[n\to+\infty]{} p(c) := \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Poiss}\left(1+e^{-2c}\right), \mathrm{Poiss}(1)\right).$$

(Written up to integer parts.)

- Several different types of profiles are known. For example with
  - **normal laws** for the riffle shuffle (Bayer–Diaconis, 1992), the random walk on $(\mathbb{Z}/2\mathbb{Z})^n$ (Diaconis–Graham–Morrison, 1990)), or simple excusion process on the circle (Lacoin 2016),
  - **Poisson laws** for k-cycles ($k = o(n)$, Nestoridi–Olesker-Taylor, 2022) or more generally all congugacy classes of the symmetric group (Olesker–Taylor T. 2024?),
  - **Tracy–Widom distributions** for the ASEP on a segment (Bufetov–Nejjar 2022),
  - **free Meixner laws** for the diffusion on O_N^+ (Freslon–T.–Wang, 2022).

♠ : Some results related to random transpositions

On random transpositions themselves :

Cutoff result : Diaconis, Shahshahani, 1981, *PTRF*
Precise lower bound : Matthews, 1988, *J. of Th. Prob.*
Phase transition result : N. Berestycki, Durrett, 2006, *PTRF*
More precise estimates on the cutoff window : Saloff-Coste–Zuniga, 2010, *AAP*
Probability of long cycles : Alon, Kozma, 2013, *Duke*
Strong stationary time : White, 2019
Cutoff profile : T., 2020, *Ann. Prob.*

Generalisations to other conjugacy classes :

Almost-precutoff for all conjugacy classes Roichman, 1996, *Invent. Math.*
Some conjugacy classes with few fixed points Lulov–Pak, 2002, *J. Alg. Comb.*
Precutoff for all conjugacy classes with few fixed points Larsen–Shalev, 2008, *Invent. Math.*
Cutoff for k-cycles : N. Berestycki, Schramm, Zeitouni, 2011, *Ann. Prob.*
Cutoff for conjugacy-invariant walks on $\mathfrak{S}_n$: N. Berestycki, Şengül, 2014, *PTRF*
Profile for k-cycles : Nestoridi, Olesker-Taylor, 2021, *PTRF*
Cutoff + profile for all conjugacy classes : Olesker-Taylor–T., 2024?

Some other generalisations :

Biaised random transpositions : Matheau-Raven, 2020
Quantum random transpositions : Freslon, T., Wang, 2021, *PTRF*
Star random transpositions : Nestoridi, 2021

♡ : The non-commutative Fourier transform

Using the **Fourier transform** : **key point** to study the walk.
Idea initialy due to **Diaconis and Shahshahani**.

♡ : The non-commutative Fourier transform

Using the **Fourier transform** : **key point** to study the walk.
Idea initialy due to **Diaconis and Shahshahani**.

Not the one on $\mathbb{R}$, where for $\xi \in \mathbb{R}$,

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-i\xi x} \mathrm{d}x,$$

♡ : The non-commutative Fourier transform

Using the **Fourier transform** : **key point** to study the walk.
Idea initialy due to **Diaconis and Shahshahani**.

Not the one on $\mathbb{R}$, where for $\xi \in \mathbb{R}$,

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-i\xi x} \mathrm{d}x,$$

but instead the one of **finite groups** G, where for $\lambda \in \widehat{G}$,

$$\hat{f}(\lambda) = \sum_{g \in G} f(g) \rho^{\lambda}(g).$$

♡ : The non-commutative Fourier transform

Using the **Fourier transform** : **key point** to study the walk.
Idea initialy due to **Diaconis and Shahshahani**.

Not the one on $\mathbb{R}$, where for $\xi \in \mathbb{R}$,

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-i\xi x} \mathrm{d}x,$$

but instead the one of **finite groups** G, where for $\lambda \in \widehat{G}$,

$$\hat{f}(\lambda) = \sum_{g \in G} f(g) \rho^{\lambda}(g).$$

Inverse Fourier transform, isometry between Hilbert spaces, Parseval identity.

Pierre-Loïc Méliot, *Representation Theory of Symmetric Groups*, chap. 1.

♡ : A method to find cutoff profiles

For transpositions, we then apply the **inverse Fourier transform** on $\mathfrak{S}_n$ to $f := \nu_n(t) - \mathrm{Unif}_n$, and use that μ_n is **constant on conjugacy classes** (so by Schur's lemma each $\widehat{f}(\lambda)$ is a multiple of the identity (as a matrix)), to get

$$2\mathrm{d}_n(t) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \left| \sum_{\lambda \in \widehat{\mathfrak{S}_n}^*} \mathrm{d}_\lambda s_\lambda^t \mathrm{ch}^\lambda(\sigma) \right| .$$

♡ : A method to find cutoff profiles

For transpositions, we then apply the **inverse Fourier transform** on $\mathfrak{S}_n$ to $f := \nu_n(t) - \mathrm{Unif}_n$, and use that μ_n is **constant on conjugacy classes** (so by Schur's lemma each $\widehat{f}(\lambda)$ is a multiple of the identity (as a matrix)), to get

$$2\mathrm{d}_n(t) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \left| \sum_{\lambda \in \widehat{\mathfrak{S}_n}^*} \mathrm{d}_\lambda s_\lambda^t \mathrm{ch}^\lambda(\sigma) \right|.$$

s_λ: eigenvalues of the (transition matrix of the) chain

d_λ: multiplicities

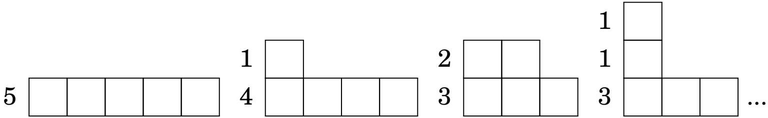
$\mathrm{ch}^\lambda(\sigma)$: "eigenvectors".

♡ : Representations of the symmetric group

- Irreducible representations λ of $\mathfrak{S}_n$ $\longleftrightarrow$ Young diagrams of size n.
  (i.e. partitions $\lambda = (\lambda_1, \lambda_2, \ldots)$ of the integer n.)

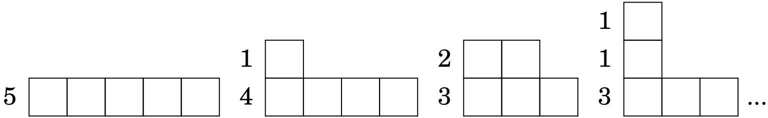
♡ : Representations of the symmetric group

- Irreducible representations λ of $\mathfrak{S}_n$ $\longleftrightarrow$ Young diagrams of size n.
  (i.e. partitions $\lambda = (\lambda_1, \lambda_2, \ldots)$ of the integer n.)
  Example: $n = 5 = 4 + 1 = 3 + 2 = \ldots.$



♡ : Representations of the symmetric group

- Irreducible representations λ of $\mathfrak{S}_n \longleftrightarrow$ Young diagrams of size n.
  (i.e. partitions $\lambda = (\lambda_1, \lambda_2, \ldots)$ of the integer n.)
  Example: $n = 5 = 4 + 1 = 3 + 2 = \ldots.$



- **Hook-length (of a box):** Example: $\lambda = [7, 5, 4, 1]$, u = orange box.
  The hook-length of u is $H(\lambda, u) = 5$.

♢ : The hook-length formula

The dimensions d_λ are easy to compute:

$$d_\lambda = |\mathrm{ST}(\lambda)| = \frac{|\lambda|!}{\prod_{u\in\lambda} H(\lambda,u)}.$$

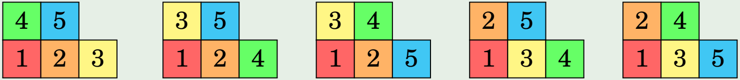
♢ : The hook-length formula

The dimensions d_λ are easy to compute:

$$d_\lambda = |\mathrm{ST}(\lambda)| = \frac{|\lambda|!}{\prod_{u \in \lambda} H(\lambda, u)}.$$

Example

The diagram $\lambda = [3,2]$ has 5 standard tableaux, so $d_\lambda = |\mathrm{ST}(\lambda)| = 5$.



.

♢ : The hook-length formula

The dimensions d_λ are easy to compute:

$$d_\lambda = |\mathrm{ST}(\lambda)| = \frac{|\lambda|!}{\prod_{u\in\lambda} H(\lambda,u)}.$$

Example

The diagram $\lambda = [3,2]$ has 5 standard tableaux, so $d_\lambda = |\mathrm{ST}(\lambda)| = 5$.

4	5	
1	2	3

3	5	
1	2	4

3	4	
1	2	5

2	5	
1	3	4

2	4	
1	3	5

.

Example

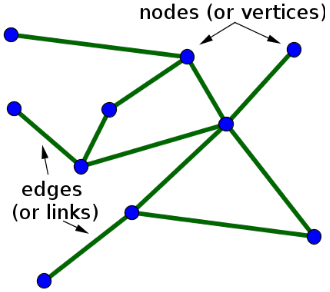
The hook-lengths of (the boxes of) $\lambda = [3,2]$ are $4,3,1,2,1$, so
$d_\lambda = \frac{|\lambda|!}{\prod_{u\in\lambda} H(\lambda,u)} = \frac{5!}{4\cdot3\cdot1\cdot2\cdot1} = 5.$

2	1	
4	3	1

.

$\diamondsuit$: Another type of problems: the cover time problem

$\Gamma = (V, E)$: graph (finite, connected).
$(X_t)_{t \geq 0}$: simple random walk on Γ.
T_x = "hitting time of the vertex x" = $\min\{t : X_t = x\}$.



Credits: mathinsight.org

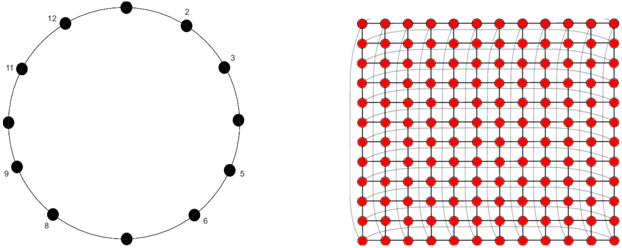
Cover time: $\tau_{\text{cov}} = \max_{x \in V} T_x$ = "first time at which all vertices have been visited".

Research on cover times: active since the works of Aldous in the 80's.

Question: $\mathbb{E}[\tau_{\text{cov}}]$? Is τ_{cov} concentrated?

♢ : Cover time of tori

Example: the d-dimensional torus $(\mathbb{Z}/m\mathbb{Z})^d$. $n := m^d$.



Credits: Ljupco Kocarev ($d = 1$), Markus Quade ($d = 2$), researchgate

- $d = 1$: $\tau_{\text{cov}} \asymp n^2$, not concentrated.
- $d = 2$: cover time cutoff: $\tau_{\text{cov}} \sim c_2 n (\log n)^2$
  (Dembo–Peres–Rosen–Zeitouni 2004, *Ann. Math.*)
- $d \geq 3$, cover time profile: $\tau_{\text{cov}} \approx c_d n (\log n + \chi)$, where $\chi \sim$ Gumbel, i.e. $\mathbb{P}(\chi \leq s) = e^{-e^{-s}}$.
  (Belius 2013, *Ann. Prob.* / De Prata 2012)

$\diamondsuit$: Cover time of vertex-transitive graphs

Our result: characterisation of Gumbel fluctuations.
$\mathrm{t_{hit}} = \mathrm{t_{hit}}(\Gamma) = \max_{x,y\in V} \mathbb{E}_x T_y$.
$\mathrm{Diam}(\Gamma)$ = diameter of Γ = $\max_{x,y\in\Gamma} d(x,y)$.
$n = n(\Gamma) := |V|$.

Theorem (N. Berestycki–Hermon–T. 2023+)

For vertex-transitive graphs of (uniformly) bouned degree, we have

$$\frac{\tau_{\mathrm{cov}}}{\mathrm{t_{hit}}} - \log n \xrightarrow[n\to\infty]{} \chi \quad \textit{if and only if} \quad \mathrm{Diam}(\Gamma)^2 \log n = o(n).$$

$\diamondsuit$: Cover time of vertex-transitive graphs

Our result: characterisation of Gumbel fluctuations.

$\mathrm{t_{hit}} = \mathrm{t_{hit}}(\Gamma) = \max_{x,y\in V} \mathbb{E}_x T_y$.

$\mathrm{Diam}(\Gamma)$ = diameter of Γ = $\max_{x,y\in\Gamma} d(x,y)$.

$n = n(\Gamma) := |V|$.

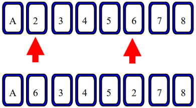
Theorem (N. Berestycki–Hermon–T. 2023+)

For vertex-transitive graphs of (uniformly) bouned degree, we have

$$\frac{\tau_{\mathrm{cov}}}{\mathrm{t_{hit}}} - \log n \xrightarrow[n\to\infty]{} \chi \quad \textit{if and only if} \quad \mathrm{Diam}(\Gamma)^2 \log n = o(n).$$

- "Gumbel iff Γ is a bit more than *2-dimensional*"
- Also iff the last points to be covered are "uniform"

$\diamondsuit$: Thank you for your attention!



$$\mathrm{d}_n\left(\frac{3}{2}\log_2(n)+c\right)\xrightarrow[n\to+\infty]{}\mathrm{d}_{\mathrm{TV}}\left(\mathcal{N}(0,1),\mathcal{N}\left(\frac{2^{-c}}{2\sqrt{3}},1\right)\right)$$

$$\mathrm{d}_n\left(\frac{1}{2}n\ln(n)+cn\right)\xrightarrow[n\to+\infty]{}\mathrm{d}_{\mathrm{TV}}\left(\mathrm{Poiss}\left(1+e^{-2c}\right),\mathrm{Poiss}(1)\right)$$

$$\hat{f}(\lambda)=\sum_{g\in G}f(g)\rho^{\lambda}(g)$$

$$\frac{\tau_{\mathrm{cov}}}{\mathrm{t}_{\mathrm{hit}}}-\log n\xrightarrow[n\to\infty]{}\chi$$