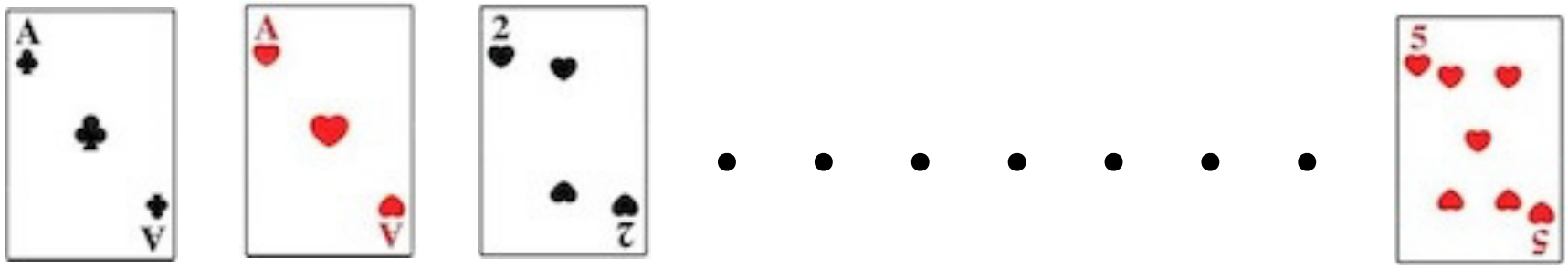
How many times should you mix a deck of cards?

Guillaume Conchon--Kerjan

Cumberland Lodge, 24th February 2024

Mixing a deck of cards

You have a deck of n distinct cards (think $n = 52$) and you know its order.



You decide to perform random shuffles to mix it. Ideally:

- you end up with every possible order having the same probability (uniform mixing)
- you do not have to do too many shuffles...

Questions:
Which random rule to choose for shuffling the cards?
Is it possible to reach a good mixing after enough shuffles?
What is a good mixing, by the way?
... If the mixing is bad, how to take advantage of it?

What is a good mixing?

We work on the set $S_n = \{\text{possible orderings of } n \text{ cards}\}$.
We have $|S_n| = n \times (n-1) \times \ldots \times 2 \times 1 = n!$ (n choices for first card, $n-1$ for the second, etc.)

A *mixing* is a probability distribution on S_n. Concretely, we have an additive map μ : subsets of $S_n \to [0,1]$, and

$$\mu(\{x\}) = \mathbb{P}(x \text{ is the ordering of the deck})$$

(μ is nonnegative, $\mu(S_n) = 1$). The uniform mixing μ_U is given by:

$$\mu_U(\{x\}) = \frac{1}{n!} \ \forall x \in S_n.$$

$\hookrightarrow$ The closer a mixing μ is to μ_U, the better!

What is a good mixing? - the total variation distance

For each $x \in S_n$, we can measure the difference $|\mu(\{x\}) - 1/n!|$. Summing over $x \in S_n$ gives the Total Variation distance:

$$\|\mu - \mu_U\|_{TV} = \frac{1}{2} \sum_{x \in S_n} = |\mu(\{x\}) - 1/n!| \quad \in [0, 1].$$

$\hookrightarrow$ the closer $\|\mu - \mu_U\|_{TV}$ is to 0, the better.

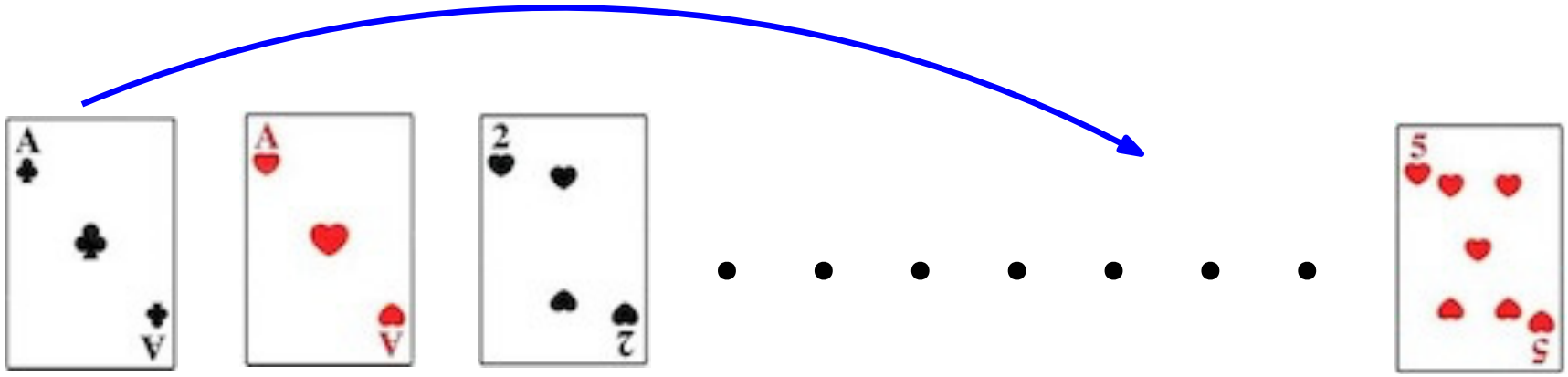
Quick check with an initial known ordering x_0 of the 52 cards: distribution $\mu_0(\{x\}) = 1_{x=x_0}$ (1 if $x = x_0$, 0 else), and

$$\|\mu - \mu_U\|_{TV} = \frac{1}{2} \left(|1 - 1/52!| + \sum_{x \in S_{52} \setminus \{x_0\}} |0 - 1/52!| \right) = 1 - 1/52! \simeq 1...$$

A non-mixed deck of cards is a very bad mixing indeed.
We will denote μ_t the distribution after t shuffles, $t \geq 0$.

Shuffling methods

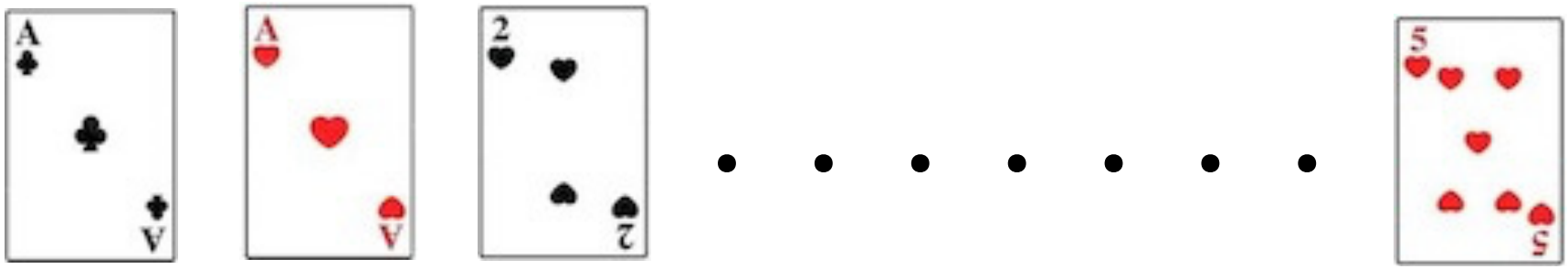
Shuffling 1: the Top-to-Random Shuffle (Aldous Diaconis '86).
Pick the top card, and place it uniformly at random in one of the 52 positions. Do it again, etc.



E.g. the top card is sent to 47th position.
(The top card could remain in the top position.)

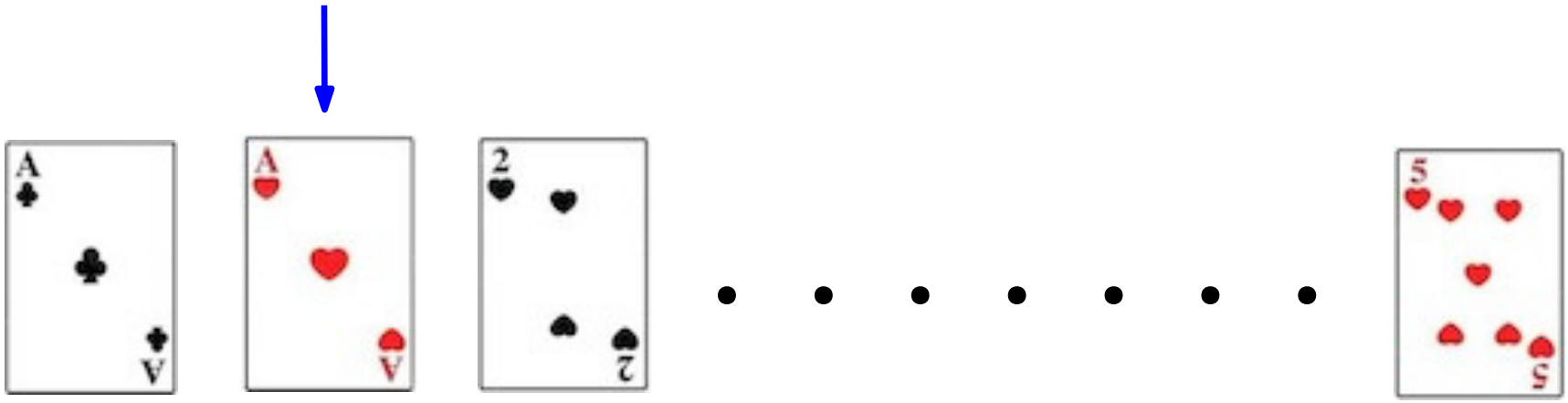
Shuffling methods

Shuffling 2: the Random Transposition (Diaconis Shahshahani '81).
Pick two cards uniformly at random, and switch them.



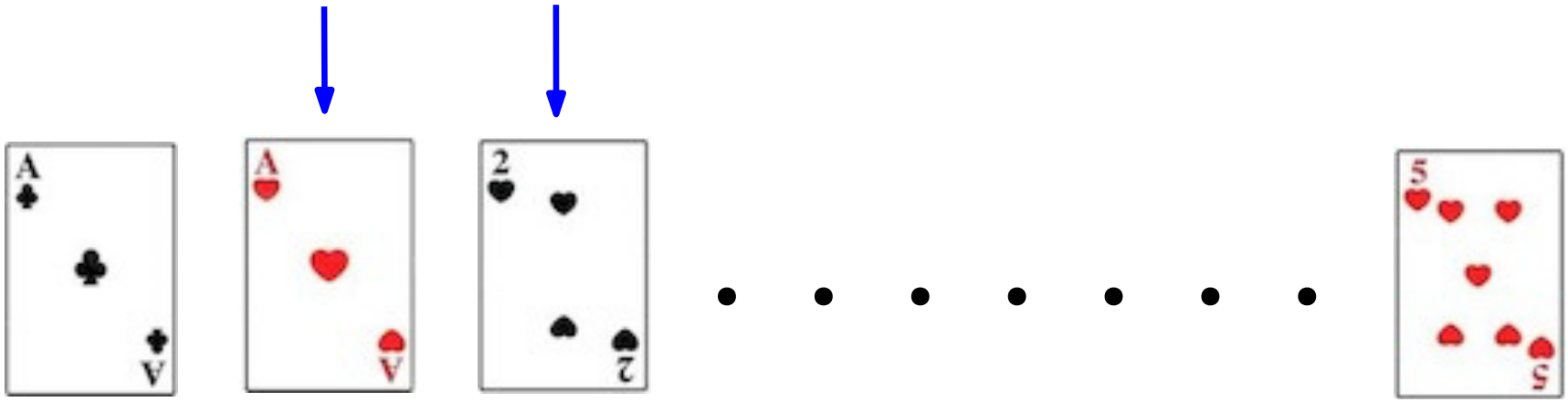
Shuffling methods

Shuffling 2: the Random Transposition.
Pick two cards uniformly at random, and switch them.



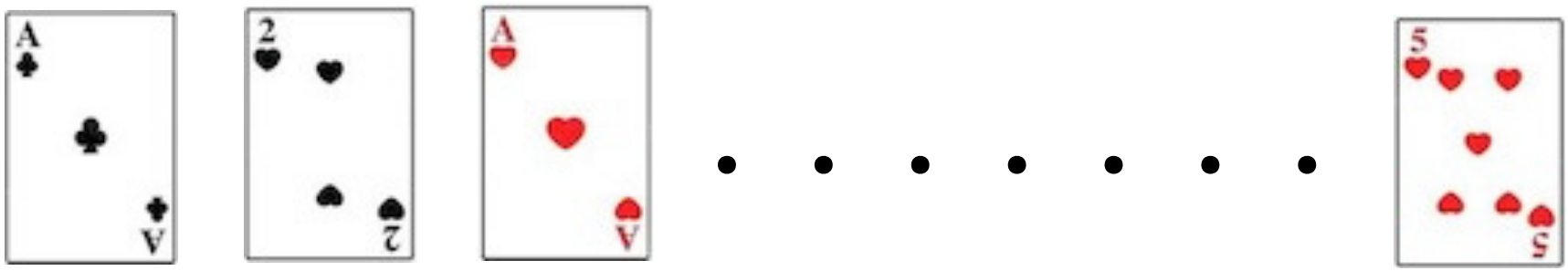
Shuffling methods

Shuffling 2: the Random Transposition.
Pick two cards uniformly at random, and switch them.



Shuffling methods

Shuffling 2: the Random Transposition.
Pick two cards uniformly at random, and switch them.

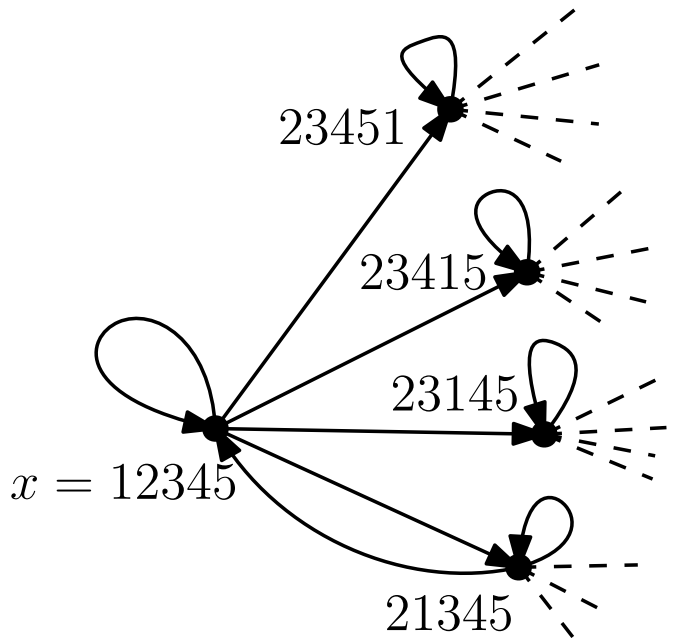


/!\ The mixing can never be perfect, because...

Modelling a shuffling method - Markov Chains

Formally, a shuffling method is defined as: for each $x \in S_n$, give transition probabilities $p(x, y)$ from x to each $y \in S_n$, i.e. the probability that shuffling once a deck in order x yields a deck in order y.

Example with top-to-random shuffle: from x, n options are equally likely (the top card can go to n positions with the same probability, $n = 5$ here, cards are 1,2,3,4,5).



Markov Chains and mixing times

Does $\|\mu_t - \mu_U\|_{TV} \to 0$ as $t \to \infty$? Is $\|\mu_t - \mu_U\|_{TV}$ at least non-increasing in t? (if it increases: more shuffles could 'undo' the mixing)

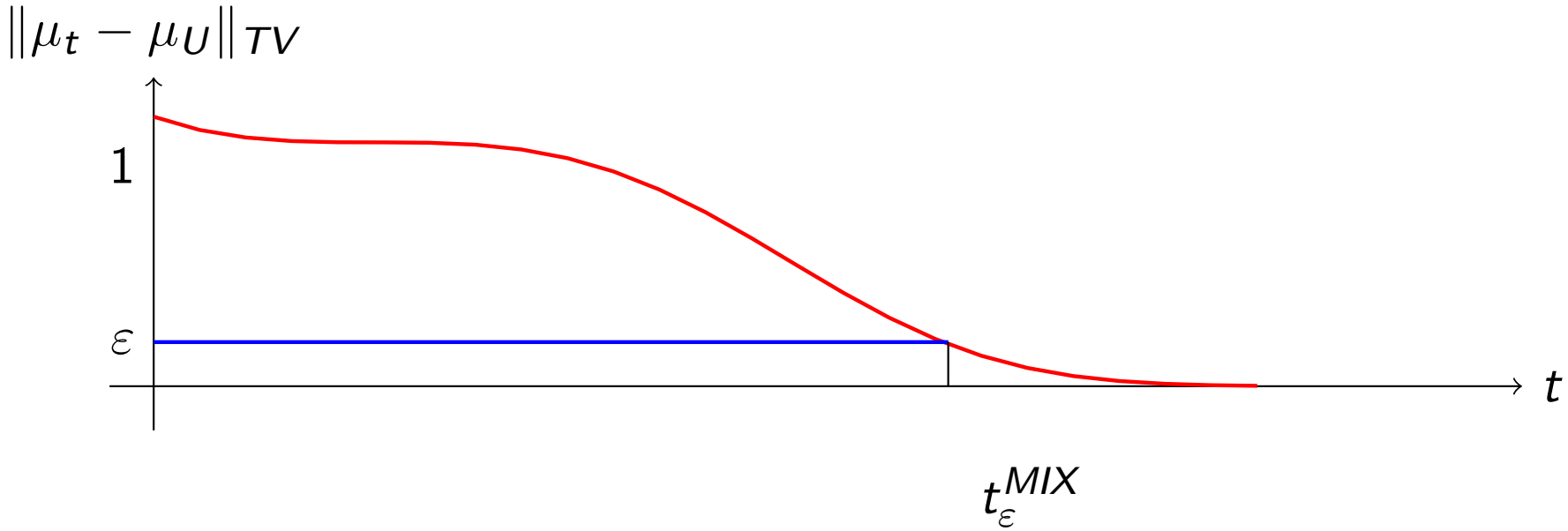
Convergence Theorem for Markov Chains

Yes, if p is aperiodic (not like Random Transposition Shuffle) and irreducible (from all x, there is a path to any y).
μ_t converges to a unique *invariant distribution* (which is μ_U for Shufflings 1 and 2: starting from a random deck of distribution μ_U, one shuffle yields a deck with the same distribution).

$\Rightarrow$ For any precision threshold $\varepsilon > 0$, we can define the ε-**mixing time** t_ε^{MIX} as the smallest t such that $\|\mu_t - \mu_U\|_{TV} \leq \varepsilon$.

"after at least t_ε^{MIX} shuffles, with probability $1 - \varepsilon$, the deck is well-mixed".

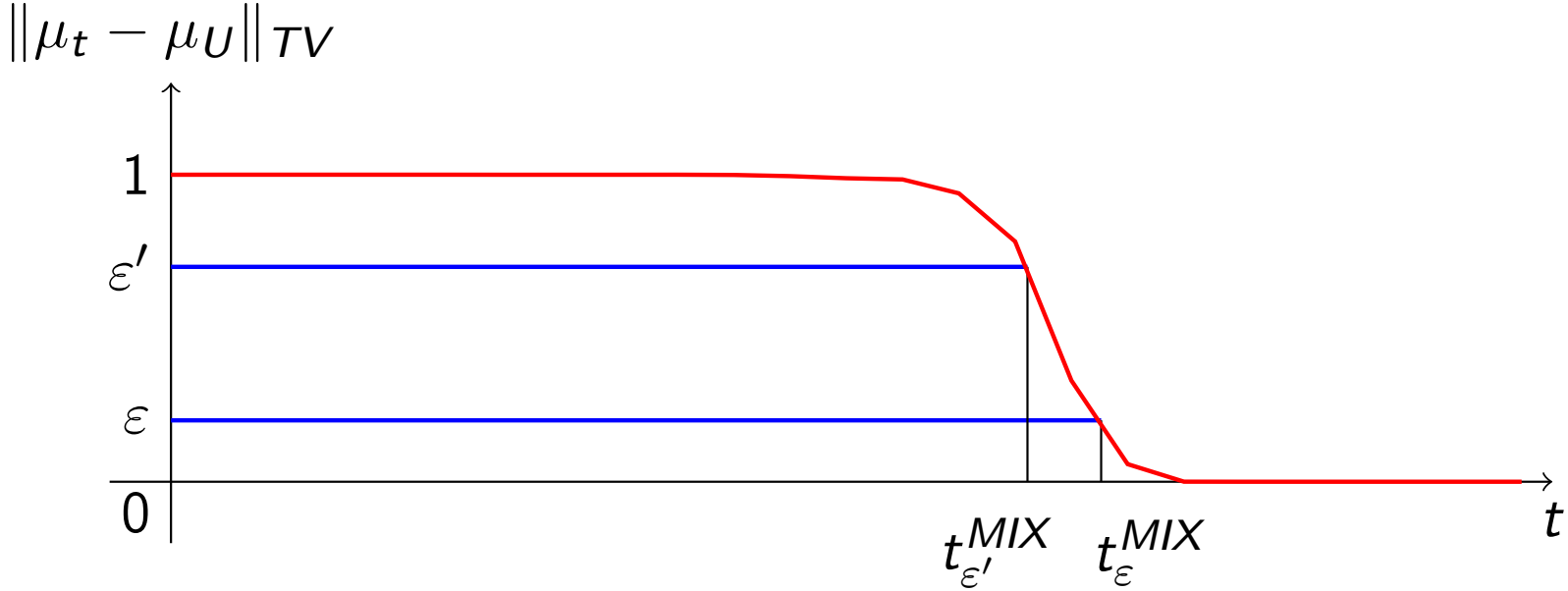
Markov Chains and mixing times



What is the shape of the curve (is the slope abrupt or smooth)? How does t_ε^{MIX} scale with n?

The cutoff phenomenon

For Shuffling i ($i = 1$ or 2), it turns out that $t_\varepsilon^{MIX} \simeq c_i n \log n$ for ANY $\varepsilon \in (0,1)$.
$\Rightarrow$ there is a crucial number of shuffles: below the mixing is very bad, above it is great!



How to prove this?

Studying a Markov chain - the matrix point of view

See p as a $|S_n| \times |S_n|$ matrix, and μ_t as a vector on S_n.
The key:

$$p^t(x,y) = \sum_{z_1,\dots,z_{t-1} \in S_n} \underbrace{p(x,z_1)\dots p(z_{t-1},y)}_{\text{any path of length } t \text{ from } x \text{ to } y} = \mathbb{P}(\text{go from } x \text{ to } y \text{ in } t \text{ shuffles})$$

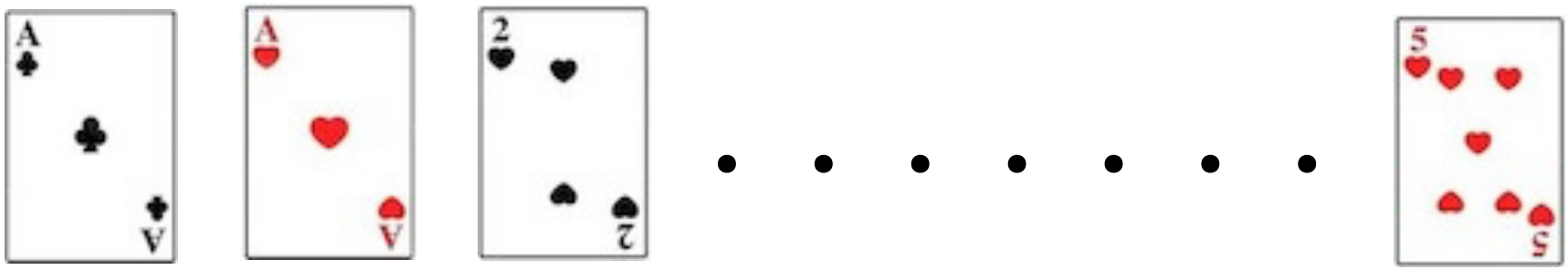
Issue: computing p^t for large t will be long...

+ Lots of structure coming from S_n, the permutation group $\to$ we have powerful algebraic tools.

Cool, but not doing this today.
Technical remark: entries λ_i of D are usually in $(-1,1)$ (except $\lambda_1 = 1$), need to wait time $t \sim \log |S_n|$ for $\sum_i \lambda_i^t$ to become negligible.

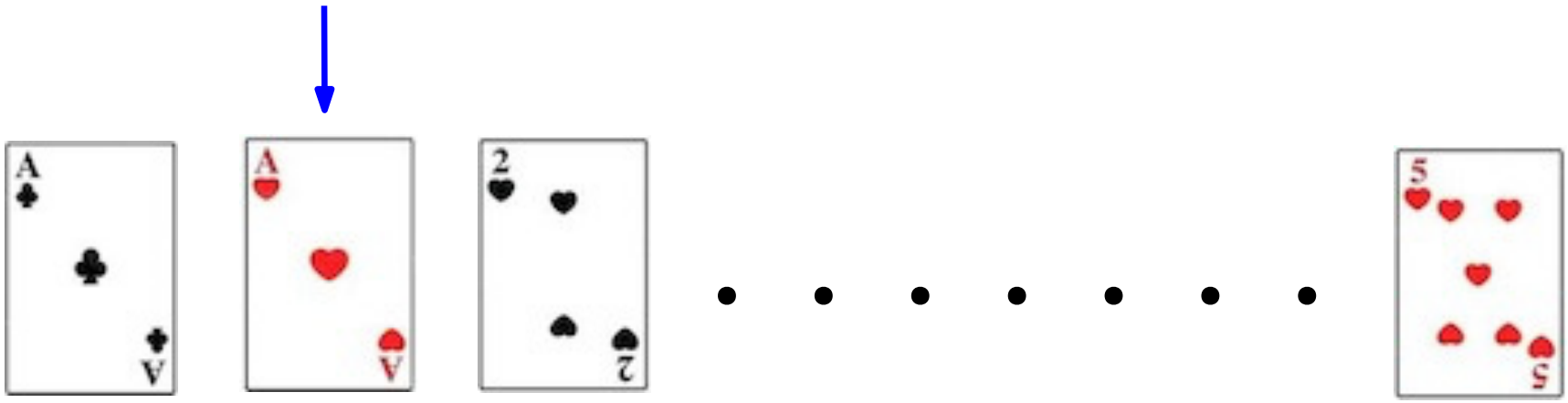
Shuffling 2: The coupon collector problem

Reminder of the Random Transposition Shuffle: pick two cards uniformly at random, and switch them.



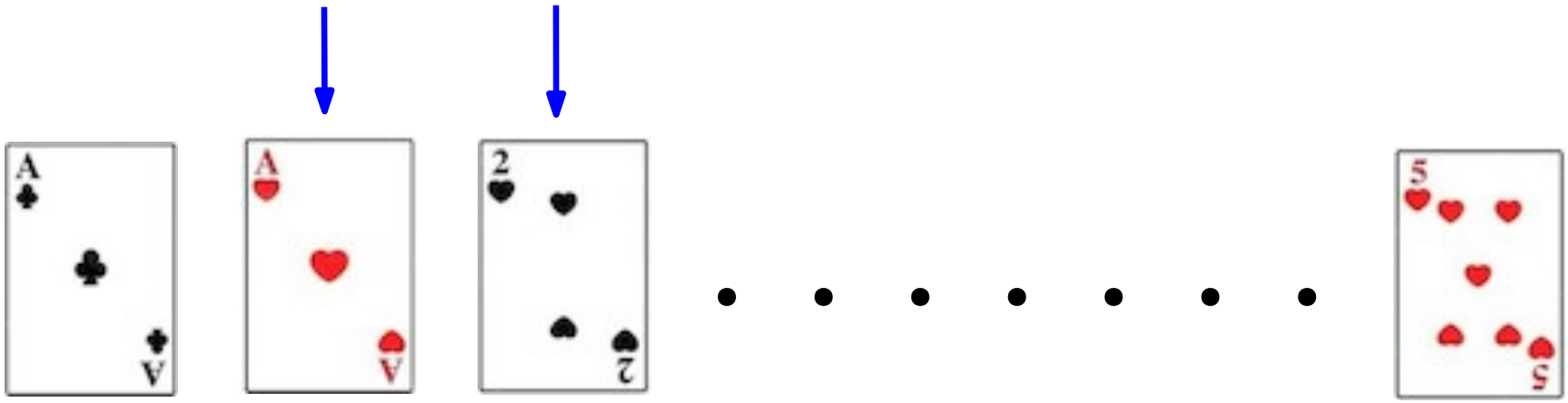
Shuffling 2: The coupon collector problem

Reminder of the Random Transposition Shuffle: pick two cards uniformly at random, and switch them.



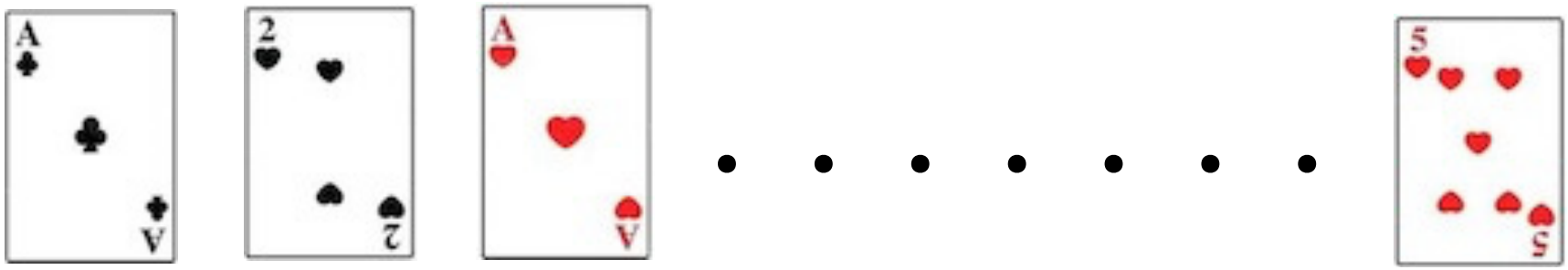
Shuffling 2: The coupon collector problem

Reminder of the Random Transposition Shuffle: pick two cards uniformly at random, and switch them.



Shuffling 2: The coupon collector problem

Reminder of the Random Transposition Shuffle: pick two cards uniformly at random, and switch them.



Shuffling 2: The coupon collector problem

Idea: if at time t there are still a few cards (say 6) that we have not sampled, they have the same relative order, which has μ_U-probability $1/6! = 1/720$, so $\|\mu_t - \mu_U\|_{TV} \gtrsim 1 - 1/720$...

By what time do we have a chance $\geq 1 - \varepsilon$ to have touched every card?

Almost the coupon collector problem: you have n distinct coupons to collect, each biscuit box contains one coupon (chosen uniformly at random). How many boxes will you buy until you have the full collection?

Shuffling 2: The coupon collector problem

The coupon collector: time $c_1 (= 1)$ to get a first coupon, then c_2 for a second (distinct), etc. Full collection by time $T_n = c_1 + \ldots + c_n$.

What is the law of c_i?

$$\Rightarrow \quad c_i \sim \text{Geometric}((n - i + 1)/n).$$

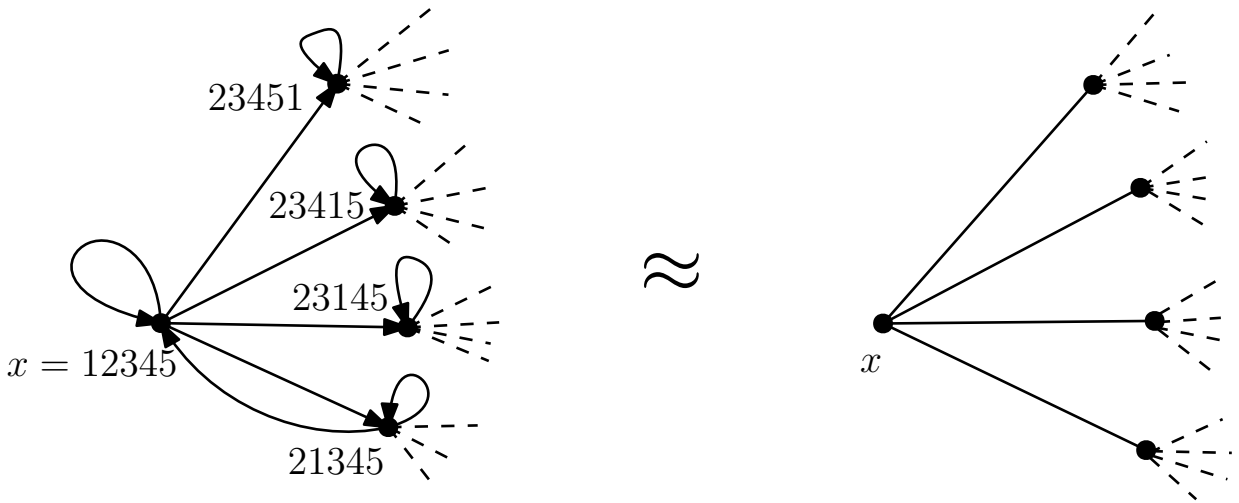
We can compute $\mathbb{E}[c_i] = n/(n - i + 1)$, so

$$\mathbb{E}[T_n] = \sum_i \mathbb{E}[c_i] = n \times (1 + 1/2 + \ldots + 1/n) \sim n \log n.$$

Gives the correct order for the mixing time. Can guess the right constant: $t_\varepsilon^{MIX} \sim (n \log n)/2$ (we sample two cards at each shuffle).

Mixing Times and Random Walks

See the successive shuffles as a random trajectory on a graph (vertex set S_n, edges given by possible shuffles).
At each step, see $n-1$ new possible vertices to visit (with Top to Random Shuffle). If few cycles: after k steps, $(n-1)^k$ potential orderings of the deck.

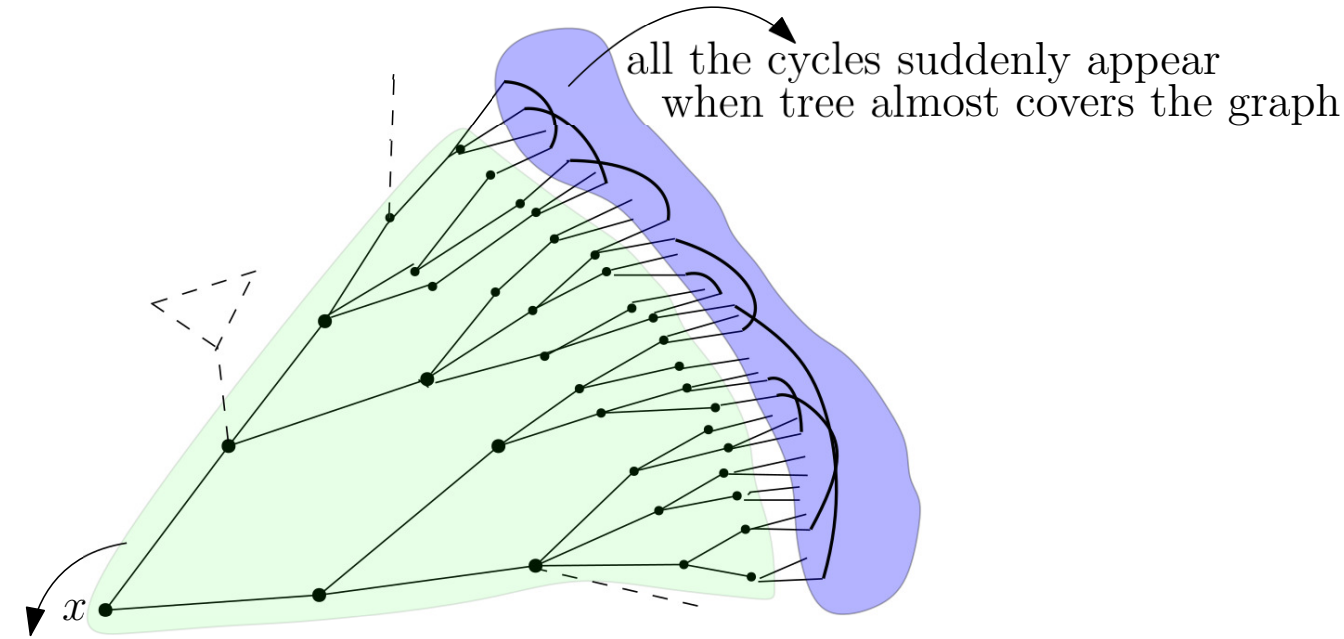


The mixing time could be of order $\sim log_{n-1}(|S_n|) \sim n$... not quite true, the graph approximation is too strong here.

Mixing Times and Random Walks

Notion of *entropy* $h > 0$ for a (rather) regular and *expanding* graph: quantifies the loss of information at each step. Cutoff at the entropic time

$$\log(\text{number of vertices})/h.$$



graph seen from x: like a tree with very few cycles.
exponential growth: will reach all vertices within logarithmic time

Turns out this is true for 'most' graphs with n vertices having each degree $d \geq 3$, as $n \to \infty$.

A more realistic shuffle - the Riffle Shuffle

For $n = 52$, $n \log n \simeq 205...$ is there a shorter shuffling?

Shuffling 3: the Riffle-Shuffle.
i) Cut the deck in two at the m-th card with $m \sim Binomial(n, 1/2)$ (more even cuts are more likely).
ii) Place the two decks next to each other and shuffle cards as follows: if the left deck has A cards, the right deck has B cards, the next card is from the left deck with probability $A/(A+B)$.

A more realistic shuffle - the Riffle Shuffle

For $n = 52$, $n \log n \simeq 205...$ is there a shorter shuffling?

Shuffling 3: the Riffle-Shuffle.
i) Cut the deck in two at the m-th card with $m \sim Binomial(n, 1/2)$ (more even cuts are more likely).
ii) Place the two decks next to each other and shuffle cards as follows: if the left deck has A cards, the right deck has B cards, the next card is from the left deck with probability $A/(A+B)$.

Bayer-Diaconis '92: by algebraic and combinatorial computations, $t_\varepsilon^{MIX} \sim \frac{3}{2} \log_2(n)$ (hence cutoff).

The 'magic seven': for $n = 52$, six shuffles give a very bad mixing, eight shuffles a very good one.

Badly mixed parts: consecutive sequences. If card A followed card B in the initial deck, when you see card A you should bet that card B is next.

Some references

P. Diaconis and M. Shahshahani, Generating a Random Permutation with Random Transpositions, *Probability Theory and Related Fields*, 1981 **(random transpositions)**

D. Aldous and P. Diaconis, Shuffling Cards and Stopping Times, *American Mathematical Monthly*, 1986 **(top to random)**

Bayer and P. Diaconis, Trailing the Dovetail Shuffle to its Lair, *Annals of Applied Probability*, 1992 **(riffle-shuffle)**

E. Lubetzky and A. Sly, Cutoff Phenomena for Random Walks on Random Regular Graphs, *Duke Mathematical Journal*, 2010 **(random walks on graphs)**

How a magician-mathematician revealed a casino loophole, S. Keating, *BBC online*, 2022

Thank you for your attention!