Upper bound for mixing times by coupling

Theorem (3.3)

Assume (X, Y) is, under $\mathbb{P}_{x,y}$, a coupling of Markov chains with transition kernel P on E, started respectively at δ_x, δ_y. Set $\tau_{\mathrm{couple}} = \inf\{t : X_t = Z_t\}$. Then

$$||P^t(x, \cdot) - P^t(y, \cdot)||_{TV} \leq \mathbb{P}_{x,y}(\tau_{\mathrm{couple}} > t).$$

Moreover $d(t) \leq \max_{(x,y)\in E^2} \mathbb{P}_{x,y}(\tau_{\mathrm{couple}} > t)$.

Upper bound for mixing times by coupling

Proof : Set $Z_t = Y_t \mathbb{1}_{\{t<\tau_{\text{couple}}\}} + X_t \mathbb{1}_{\{t\geq\tau_{\text{couple}}\}}$. Then, by Markov property at time τ_{couple}, (X, Z) remains a coupling of Markov chains started at δ_x, δ_y. We refer to (X, Z) as a *coalescent* coupling : we have indeed $X_s = Z_s \Rightarrow X_t = Z_t \ \forall t \geq s$. Of course $\widetilde{\tau_{\text{couple}}} := \inf\{t : X_t = Z_t\} = \tau_{\text{couple}}$.
Now (X_t, Z_t) is a coupling of $P^t(x, \cdot)$ and $P^t(y, \cdot)$, so that

$$||P^t(x, \cdot) - P^t(y, \cdot)||_{TV} \leq \mathbb{P}_{x,y}(X_t \neq Z_t) = \mathbb{P}(\tau_{\text{couple}} > t).$$

It remains to recall that
$d(t) \leq \overline{d}(t) = \max_{(x,y)\in E^2} ||P^t(x, \cdot) - P^t(y, \cdot)||_{TV}$ to conclude.

Exemple 1 : lazy SSRW on the discrete circle

Consider $(\varepsilon_t)_{t\in\mathbb{N}}, (\xi_t)_{t\in\mathbb{N}}$ i.i.d according to the Bernoulli distribution with parameter $1/2$. Set

$$X_{t+1} = X_t + \mathbb{1}_{\{\xi_{t+1}=\varepsilon_{t+1}=1\}} - \mathbb{1}_{\{\xi_{t+1}=0,\varepsilon_{t+1}=1\}} \quad [n],$$

$$Y_{t+1} = Y_t + \mathbb{1}_{\{\xi_{t+1}=1,\varepsilon_{t+1}=0\}} - \mathbb{1}_{\{\xi_{t+1}=\varepsilon_{t+1}=0\}} \quad [n].$$

It is easily seen that (X, Y) is a coupling of lazy symmetric simple random walks on the discrete circle $\frac{\mathbb{Z}}{n\mathbb{Z}}$. Between time t and $t+1$, each of the two walks may jump by $+1$ along the discrete circle with probability $1/4$, by -1 with probability $1/4$, and otherwise stays put.

In fact, the variable ε_{t+1} is used to determine which of the two walks jumps between t and $t+1$ while the other one stays put, and the variable ξ_{t+1} is used to determine wether the direction in which the moving walk makes its jump.

It is easily seen that $X - Y$ is the symmetric simple random walk on the discrete circle.

We could also have used (X, Z) the coalescent coupling deduced from (X, Y). Note that $S = X - Z$ is a simple random walk on the discrete circle absorbed at the origin. Now if S is started at k, it is known (this is the expected time of play in glamber's ruin problem, deduced from the fact that $(S_t(n - S_t) - t \wedge \tau_{\text{couple}})_{t \geq 0}$ is a martingale) that $\mathbb{E}[\tau_{\text{couple}}] = k(n - k)$. Hence, whatever the starting points for X, Y, $\mathbb{E}[\tau_{\text{couple}}] \leq \frac{n^2}{4}$, and by Markov inequality we deduce that

$$\max_{x,y \in \frac{\mathbb{Z}}{n\mathbb{Z}}} \mathbb{P}_{x,y}(\tau_{\text{couple}} > t) \leq \frac{n^2}{4t},$$

so we conclude by Theorem (3.3) that $t_{\text{mix}} \leq n^2$, and thus $t_{\text{mix}}(\varepsilon) \leq \lceil \log_2(1/\varepsilon) \rceil n^2$.

Example 1 : lazy SSRW on the discrete circle

This, in fact, provides the correct order of magnitude for t_{mix} as $n \to \infty$. Indeed, by CLT, $\frac{X_{cn^2}}{\sqrt{c/2n}} \to Z \sim \mathcal{N}(0,1)$. For $c = 1/16$ and n large enough we find that

$$\mathbb{P}(|X_{cn^2}| \geq n/4) \leq \mathbb{P}(|Z| \geq \sqrt{2}) \approx 0.157.$$

But then $P^{n^2/16}(0, B^c_{n/4}) \leq 0.16$, and since $\pi(B^c_{n/4}) = 1/2$ we have $d(n^2/4) \geq 1/2 - 0.16 > 1/4$ so $t_{\mathrm{mix}} \geq n^2/16$ for large enough n. We have shown that for large enough n,

$$\frac{n^2}{16} \leq t_{\mathrm{mix}} \leq n^2,$$

in other words, $t_{\mathrm{mix}} = \Theta(n^2)$.
The case of lazy asymmetric simple random walk on the discrete circle can be treated in the exact same way : simply note that $X - Y$ remains, in fact, symmetric.

Exercise III.13 : lazy SSRW on the discrete d-dimensional torus

The idea is similar, except we have to work with d coordinates. Add $(i_t)_{t\in\mathbb{N}}$ i.i.d according to uniform distribution on $\{1, ..., d\}$. Whenever $i_{t+1} = i$, perform a jump for one of the two walks along the ith coordinate, leaving the other idle. Now along each coordinate, we have a coupling of $(1 - 1/2d)$-lazy SSRW on the discrete circle.

The important observation is that we can accelerate coalescing of the two walks by letting each $X^i - Z^i$ be absorbed at the origin. It is easy that if ${\tau_{\text{couple}}}^i$ denotes the coupling time of the ith coordinate, for any starting points x, y we have $\mathbb{E}_{x,y}[{\tau_{\text{couple}}}^i] \leq \frac{dn^2}{4}$. Now $\tau_{\text{couple}} = \max_{1\leq i\leq d} {\tau_{\text{couple}}}^i$, but a simple bound provides $\mathbb{E}[\tau_{\text{couple}}] \leq \sum_{i=1}^{d} \mathbb{E}[{\tau_{\text{couple}}}^i] \leq \frac{d^2n^2}{4}$, so by Markov's inequality and Theorem (3.3) we conclude that $t_{\text{mix}} \leq d^2n^2$.

Exercise III.13 : lazy SSRW on the discrete d-dimensional torus

When d is fixed and $n \to \infty$ this obviously provides the right order of magnitude for t_{mix}. On the other hand when n is fixed and $d \to \infty$ our last bound is too naive.
In the case $n = 2$, τ_{couple} exactly has the distribution of the collection time in the coupon collector problem. Note that in that case τ_{couple} is remarkably concentrated around its mean value $d \log(d)$, which allows to show that $t_{\mathrm{mix}}(\varepsilon) \leq d \log(d) + C(\varepsilon)d$ whatever the choice of $\varepsilon \in (0, 1)$.
For more general n, this is a bit more involved, and the right estimate is $\mathbb{E}[\tau_{\mathrm{couple}}] \leq C(n)d \log(d)$, but one could at least easily establish that $\mathbb{E}[\tau_{\mathrm{couple}}] \leq C(n)d(\log(d))^2$.

Exercise III.14 : lazy SRW on binary tree with depth k

The idea is to first couple the depths of the walks until these match (it is enough that only one of the two jumps at each step until the depths are the same), then the walks (by performing a similar move in the tree for both walks so their depths remains the same) when they meet. This allows to ensure that the coupling time is bounded above by the time it takes to reach the leaves of the tree for the walk which starts closest to the root (by then the depths must match), plus the time it takes that walk to get back to the root, at which they must necessarily have met. Note this commute time is exactly the commute time between 0 and k for $(|X_t|, t \geq 0)$ which is an asymmetric simple random walk reflected at $0, k$.
It follows (e.g. using theorem for expected commute times and effective resistances for $|X|$) that $\mathbb{E}[\tau_{\text{couple}}] \leq 8 \cdot 2^k$, thus by Markov's inequality and Theorem (3.3) $t_{\text{mix}} \leq 32 \cdot 2^k$.

Exercise III.14 : lazy SRW on binary tree with depth k

Once again this coupling method gives the right order of magnitude for t_{mix} when $k \to \infty$. Indeed if the walk starts at a leaf on the right side of the tree, it must go through the root before it can visit any vertex in the left side of the tree. It remains to argue that for some constant c, starting from a leaf, the number of visits to the set of leaves before it reaches the root is greater than $c2^k$ with probability at least $1/4$. Such estimate is easy to establish by studying the (asymmetric) walk $(|X_t|, t \geq 0)$, and we conclude that $t_{\mathrm{mix}} = \Theta(2^k)$.

Coupling methods : further examples

Coupling methods are also useful in examples of walks on the permutation group $\mathfrak{S}_n$: see exercise III.16 for which jumps are transpositions.
Exercise III.17 focuses on a random walk on graph colourings, where the graph is fixed, and jumps are performed by changing the colour of only one vertex, as long as the new clour differs from that of neighbouring vertices. Provided there are enough colours (in particular there exists admissible colourings, the space of admissible colourings is a communicating class, and the stationary distribution is the uniform distribution on admissible colourings), a coupling argument can be used to bound above the mixing time. Since the walk can be started at any colouring (admissible or not), this suggests an algorithm to approximate the uniform drawing of an admissible colouring : run the walk for a time larger than that bound.