Cover times and generic chaining

Joseph Lehec

Cover times and generic chaining

Joseph Lehec*

January 7, 2015

Abstract

A recent result of Ding, Lee and Peres expresses the cover time of the random walk on a graph in terms of generic chaining for the commute distance. Their argument is based on the Ray-Knight isomorphism theorem. The purpose of this article is to present an alternative approach to this problem, based only on elementary hitting times estimates and chaining arguments.

1 Introduction

Let $(X_n)_{n\geq 0}$ be an irreducible Markov chain on some state space M. Given $A \subset M$ let

$$T(A) = \inf\{n \geq 0 \colon X_n \in A\}$$

be the first time the chain hits A and let

$$T_{cov}(A) = \sup_{x\in A} T(x)$$

be the first time the chain X has visited every point of A. The cover time of A is by definition

$$\mathrm{cov}(A) = \sup_{x\in A} \big(\mathsf{E}_x\, T_{cov}(A)\big),$$

where E_x stands for conditional expectation given $X_0 = x$ (similarly P_x shall stand for conditional probability given $X_0 = x$ in the sequel). To avoid trivial situations, the chain is assumed to be positive recurrent throughout so that $\mathrm{cov}(A) < +\infty$ if and only if A is finite.
Using the strong Markov property it is easily seen that given x, y, z in M

$$\mathsf{E}_x\, T(y) + \mathsf{E}_y\, T(z)$$

is the expectation (under P_x) of the first time the chain has visited y and z (in this order). This implies that

$$\mathsf{E}_x\, T(y) + \mathsf{E}_y\, T(z) \geq \mathsf{E}_x\, T(z).$$

*Université Paris-Dauphine

Therefore the commute time

$$d(x,y) = \mathsf{E}_x\, T(y) + \mathsf{E}_y\, T(x)$$

is a distance on M. This article deals with a problem dating back to Matthews' article [6] at least: Can $\mathrm{cov}(A)$ be estimated in terms of metric properties of (A,d)? An arguably definitive answer to this question has recently be given by Ding Lee and Peres [3]; their result is expressed in terms of generic chaining.

1.1 The generic chaining

The generic chaining is a tool designed by Talagrand to estimate suprema of Gaussian processes. Let us describe it briefly and refer to the book [7] for details.
Throughout we let $(N_n)_{n\geq 0}$ be the following sequence of integers:

$$N_0 = 1,\ N_n = 2^{2^n}, n \geq 1. \tag{1}$$

Given a set S, a sequence $(\mathcal{A}_n)_{n\geq 0}$ of partitions of S is called admissible if $\mathcal{A}_{n+1}$ is a refinement of $\mathcal{A}_n$ and if $|\mathcal{A}_n| \leq N_n$ for every $n \geq 0$, where $|\mathcal{A}_n|$ is just the cardinality of $\mathcal{A}_n$. The cardinality condition implies in particular that $\mathcal{A}_0 = \{S\}$. Given a sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ of S and $s \in S$ we let $A_n(s)$ be the only element of $\mathcal{A}_n$ containing s.

Definition 1. Let (S,d) be a metric space. Set

$$\gamma_2(S,d) = \inf\Big[\sup_{s\in S}\Big(\sum_{n=0}^{+\infty} 2^{n/2}\Delta(A_n(s),d)\Big)\Big],$$

where the infimum is taken over all admissible partitions $(\mathcal{A}_n)_{n\geq 0}$ of S, and $\Delta(A,d)$ denotes the diameter of A.

Recall that a Gaussian process is a family $(Y_s)_{s\in S}$ of random variables such that every linear combination of the variables Y_s is Gaussian. The process is said to be centered if $\mathsf{E}\, Y_s = 0$ for every s. The fundamental result of Talagrand reads as follows.

Theorem 2. *Let $(Y_s)_{s\in S}$ be a centered Gaussian process. Then*

$$\frac{1}{L}\gamma_2(S,d) \leq \mathsf{E}\sup_{s\in S} Y_s \leq L\gamma_2(S,d) \tag{2}$$

where L is a universal constant and d is the following distance on S

$$d(s,t) = \sqrt{\mathsf{E}(Y_s - Y_t)^2}. \tag{3}$$

The upper bound is not specific to Gaussian processes, it applies to any centered process $(Y_s)_{s\in S}$ satisfying

$$\mathsf{P}(Y_s - Y_t \geq u) \leq \mathrm{e}^{-u^2/2d(s,t)^2} \tag{4}$$

for all $s,t \in S$, for all $u > 0$ and for some distance d. Using a union bound it is not hard to see that a centered process satisfying (4) satisfy

$$\mathsf{E} \sup_{s\in A} Y_s \leq C\sqrt{\log|A|} \max_{s,t\in A} d(s,t), \tag{5}$$

for every finite subset A of S. The proof of the upper bound of (2) consists in applying this union bound repeatedly and at different scales.
The lower bound is another story, it is specific to Gaussian processes and much more difficult to prove. Roughly speaking the argument relies on two properties: the concentration of the Gaussian measure and the Sudakov inequality. Let us state the latter; if $(Y_s)_{s\in S}$ is a centered Gaussian process then for all finite subset A of S

$$\mathsf{E} \sup_{s\in A} Y_s \geq c\sqrt{\log|A|} \min_{s\neq t\in A} d(x,y) \tag{6}$$

where c is a universal constant and d is the L^2-distance (3).

1.2 The Ding, Lee and Peres theorem

Cover times satisfy inequalities analoguous to (5) and (6) due to Matthews [6]: for any finite subset A of M

$$\begin{aligned} \mathrm{cov}(A) &\leq (1+\log|A|) \max_{x,y\in A} \big(\mathsf{E}_x\, T(y)\big) \\ \mathrm{cov}(A) &\geq \log|A| \min_{x\neq y\in A} \big(\mathsf{E}_x\, T(y)\big). \end{aligned} \tag{7}$$

In view of these inequalities it seems natural to conjecture that the correct order of magnitude for $\mathrm{cov}(A)$ is

$$\gamma_1(A,d) = \inf\Big[\sup_{x\in A}\Big(\sum_{n=0}^{+\infty} 2^n \Delta(A_n(x),d)\Big)\Big],$$

rather than $\gamma_2(A,d)$ (recall that d is the commute distance $d(x,y) = \mathsf{E}_x\, T(y) + \mathsf{E}_y\, T(x)$). This is not quite correct. Here is the result of Ding, Lee and Peres.

Theorem 3. *If the Markov chain $(X_n)_{n\geq 0}$ is reversible (and if the state space M is finite) then*

$$\frac{1}{L}\big[\gamma_2(M,\sqrt{d})\big]^2 \leq \mathrm{cov}(M) \leq L\big[\gamma_2(M,\sqrt{d})\big]^2$$

for some universal constant L.

Remark. Actually the inequality remains valid when M is infinite. Indeed since $d(x,y) \geq 1$ when $x \neq y$, we then have $\gamma_2(M,\sqrt{d}) = +\infty$.

The correct order of magnitude $\gamma_2(M,\sqrt{d})^2$ is comparable to our wrong guess: clearly

$$\gamma_1(M,d) \leq \big[\gamma_2(M,\sqrt{d})\big]^2.$$

1.3 Purpose of the present article

The proof of Ding, Lee and Peres is very involved. In particular it relies on the Ray-Knight isomorphism theorem which makes a connection between local times of the chain and the Gaussian free field associated to the chain. It may be interesting to have a simpler proof relying only on elementary hitting times estimates and on Talagrand's generic chaining. The purpose of this article is to provide such a proof.
Unfortunately we fail to recover entirely Theorem 3, here is what we prove.

Theorem 4. *If $(X_n)_{n\geq 0}$ is irreducible and positive recurrent, then*

$$\mathrm{cov}(M) \leq L\big[\gamma_2(M, \sqrt{d})\big]^2. \tag{8}$$

for some universal L. More generally we have

$$\mathrm{cov}(A) \leq L\big[\gamma_2(A, \sqrt{d})\big]^2, \tag{9}$$

for every subset A of M.

Inequality (8) is slightly stronger than the upper bound of Theorem 3 since the chain is no longer assumed to be reversible. Besides, it is not clear whether the approach of Ding, Lee and Peres yields (9).

Theorem 5. *If in addition the chain $(X_n)_{n\geq 0}$ is reversible then*

$$\gamma_1(M, d) \leq L\,\mathrm{cov}(M), \tag{10}$$

where L is a universal constant. Again we actually have

$$\gamma_1(A, d) \leq L\,\mathrm{cov}(A),$$

for every $A \subset M$.

Remark. The reversibility assumption is necessary. Indeed, consider the discrete torus $\mathbb{Z}_N$ and the Markov kernel given by

$$P(x, x+1) = 1, \quad \forall x \in \mathbb{Z}_N.$$

Clearly $d(x, y) = N$ for all $x \neq y$, which implies that

$$\gamma_1(T, d) \approx N \log(N).$$

On the other hand $T_{cov}(\mathbb{Z}_N) = N$ almost surely (whatever the starting point).

Since $\gamma_1(M, d) \leq [\gamma_2(M, \sqrt{d})]^2$ inequality (10) is weaker than the lower bound of Theorem 3. Let us comment a little bit more on this. In order to compute $\gamma_1(M, d)$ one can restrict to partitions $(\mathcal{A}_n)_{n\geq 0}$ satisfying

$$\mathcal{A}_n = \{\{x\},\ x \in M\}$$

for $n \geq k$, where k is the only integer satisfying

$$N_{k-1} < |M| \leq N_k.$$

Then by convexity we get

$$\Big(\sum_{n=0}^{\infty} 2^{n/2}\sqrt{\Delta(A_n(x),d)}\Big)^2 = \Big(\sum_{n=0}^{k} 2^{n/2}\sqrt{\Delta(A_n(x),d)}\Big)^2$$
$$\leq (k+1)\sum_{n=0}^{\infty} 2^n \Delta(A_n(x),d)$$

for every $x \in M$, yielding

$$\big[\gamma_2(M,\sqrt{d})\big]^2 \leq C\log(\log|M|)\gamma_1(M,d)$$

for some universal C (provided $|M| \geq 3$). Therefore the estimate (10) is off the correct order of magnitude by at most a factor $\log(\log|M|)$. This is sharp; there is a Markov chain for which the gap is indeed $\log(\log|M|)$ (see the appendix).

2 The upper bound

Since $(X_n)_{n\geq 0}$ is an irreducible, positive recurrent Markov chain, there is a unique invariant probability measure which we denote by π. The purpose of this section is to bound

$$\mathsf{E}\sup_{x\in M} T(x)$$

through a chaining argument. Since no estimate such as (4) is available for hitting times, the chaining procedure will be different from Talagrand's, and is taken from the articles [2, 4].
We need a couple of additional notations. Let

$$T^0(x) = 0$$
$$T^k(x) = \inf(n \geq T^{k-1}(x)+1,\ X_n = x), \quad \forall k \geq 1.$$

When the chain starts from x, the variable $T^k(x)$ is just the k-th return time to x. Lastly, let

$$N_k = \sum_{n=0}^{k-1} \delta_{X_n}$$

be the empirical measure of the chain X. In other words $N_k(x)$ is the number of visits to x before time k.
The following deviation estimate is due to Kahn, Kim, Lovasz and Vu [4].

Lemma 6. *Let $x \neq y$ in M. Then for every $\epsilon > 0$ and for every integer k*

$$\mathsf{P}_x\Big(N_{T^k(x)}(y) \leq (1-\epsilon)\frac{k\pi(y)}{\pi(x)}\Big) \leq \exp\Big[-\frac{\epsilon^2 k}{4\pi(x)d(x,y)}\Big].$$

Let us sketch the argument. Because of the strong Markov property, under P_x the variables

$$\big(N_{T^i(x)}(y) - N_{T^{i-1}(x)}(y)\big)_{i\geq 1}$$

are independent and identically distributed. And it is a standard fact (see for instance [1, chapter 2]) that their law is geometric: for every integer r

$$\mathsf{P}_x\big(N_{T^1(x)}(y) \geq r\big) = p_{xy}(1-p_{yx})^r,$$

where

$$p_{xy} = \mathsf{P}_x(T(y) \leq T^1(x)) = \frac{1}{\pi(x)d(x,y)}.$$

The previous lemma is thus a Hoeffding type estimate for sums of independent geometric variables. We refer to [4] for the details.
Our next tool is taken from Barlow, Ding, Nachmias and Peres [2].

Lemma 7. *Let A be a finite subset of M, let $z \in A$ and let k be an integer. Then*

$$\begin{aligned}\mathsf{E}_z\, T_{cov}(A) &\leq \frac{\mathsf{E}_z\, T^k(z)}{\mathsf{P}_z\big(T_{cov}(A) \leq T^k(z)\big)} \\ &= \frac{k}{\pi(z)\,\mathsf{P}_z\big(T_{cov}(A) \leq T^k(z)\big)}\end{aligned}$$

Again we sketch the proof and refer to [2] for details. Let

$$N = \inf(n \geq 1,\ T_{cov}(A) \leq T^{nk}(z)).$$

Then by Wald's identity

$$\mathsf{E}_z\, T_{cov}(A) \leq \mathsf{E}_z\, T^{Nk}(z) = \mathsf{E}_z(N)\,\mathsf{E}_z\, T^k(z).$$

On the other hand if N is larger than n then the walk fails to cover A during any of the following intervals of time

$$[0, T^k(z)),\ [T^k(z), T^{2k}(z)), \ldots,\ [T^{(n-1)k}(z), T^{nk}(z))$$

so that

$$\mathsf{P}_x(N > n) \leq \mathsf{P}_z\big(T_{cov}(A) \geq T^k(z)\big)^n.$$

The result follows.
The authors of [2] combine these two lemmas with a nice chaining argument. Although it is not written this way, their result is essentially the Dudley version of Theorem 4:

$$\mathrm{cov}(M) \leq L\big(\sum_{n=0}^{\infty} \mathrm{e}_n(M, \sqrt{d})2^{n/2}\big)^2 \tag{11}$$

where

$$\mathrm{e}_n(M, \sqrt{d}) = \inf_A\big(\sup_{x\in M} \sqrt{d(x,A)}\big)$$

where the infimum is taken over all subsets A of M satisfying $|A| \leq N_n$. This is weaker than Theorem 4. Indeed swapping the sup and the sum in the definition of γ_2, it is easily seen that

$$\gamma_2(M, \sqrt{d}) \leq C \sum_{n=0}^{\infty} \mathrm{e}_n(M, \sqrt{d}) 2^{n/2},$$

for some universal constant C. We show that it is possible to modify BDLP's chaining argument to obtain Theorem 4.
Let z, x, y in M such that $x \neq y$ and let k, l be two integers larger than 1. Observe that

$$\begin{aligned} \mathsf{P}_z(T^l(y) > T^k(x)) &= \mathsf{P}_z(N_{T^k(x)}(y) \leq l-1) \\ &\leq \mathsf{P}_z(N_{T^k(x)}(y) - N_{T^1(x)}(y) \leq l-1) \\ &= \mathsf{P}_x(N_{T^{k-1}(x)}(y) \leq l-1). \end{aligned}$$

The latest equality being a consequence of the strong Markov property. If $(l-1)/\pi(y) < (k-1)/\pi(x)$, applying Lemma 6 to $k-1, l-1$ and

$$\epsilon = 1 - \frac{(l-1)\pi(x)}{(k-1)\pi(y)}$$

gives

$$\mathsf{P}_z(T^l(y) > T^k(x)) \leq \exp\left[- \frac{\left(\frac{k-1}{\pi(x)} - \frac{l-1}{\pi(y)}\right)^2}{4 d(x,y) \frac{k-1}{\pi(x)}} \right]. \tag{12}$$

This will be our key estimate. Lastly, we shall use the following elementary fact: if x and y are distinct elements of M then

$$\frac{1}{\pi(x)} = \mathsf{E}_x T^1(x) \leq \mathsf{E}_x T(y) + \mathsf{E}_y T(x) = d(x,y).$$

Let us reformulate Theorem 4.

Proposition 8. *Let $A \subset M$, let $z \in A$ and let $(\mathcal{A}_n)_{n \geq 0}$ be an admissible sequence of partitions of A. Then*

$$\mathsf{E}_z(T_{cov}(A)) \leq L \Big(\sup_{x \in A} \sum_{n=0}^{\infty} 2^{n/2} \sqrt{\Delta(A_n(x))} \Big)^2.$$

Recall that $A_n(x)$ denotes the only element of $\mathcal{A}_n$ containing x. Also Δ denotes the diameter with respect to the commute distance.

Proof. Let $t_0(A) = z$, and for each n and for each $B \in \mathcal{A}_n$ let $t_n(B)$ be an arbitrary element of B. Given $x \in A$, we let $x_n = t_n(A_n(x))$. We can assume that A is finite and that

$$\mathcal{A}_n = \{\{x\},\ x \in A\}$$

for n large enough (the right hand of the desired inequality equals $+\infty$ otherwise). Therefore $x_n = x$ eventually. Let

$$r_n(x) = \sup_{y \in A_n(x)} \sum_{k=n}^{+\infty} 2^{k/2} \sqrt{\Delta(A_k(y))}$$

and

$$k_n(x) = \lfloor 34 \cdot \pi(x_n) r_n(x) r_0(x) \rfloor + 1,$$

where $\lfloor r \rfloor$ denotes the integer part of r. Observe that $r_n(x)$ and $k_n(x)$ depend only on $A_n(x)$. In particular $k_0(x)$ depends on nothing. Also

$$\begin{aligned} r_n(x) - r_{n+1}(x) &\geq 2^{n/2} \sqrt{\Delta(A_n(x))} \\ &\geq 2^{n/2} \sqrt{d(x_n, x_{n+1})}. \end{aligned}$$

We claim that for every x and n

$$\mathsf{P}_z\big(T^{k_{n+1}(x)}(x_{n+1}) > T^{k_n(x)}(x_n)\big) \leq \mathrm{e}^{-2^{n+3}} \leq \frac{1}{N_{n+3}}. \tag{13}$$

Indeed, if $x_n = x_{n+1}$ then $k_{n+1}(x) \leq k_n(x)$ and the inequality is trivial. Otherwise write

$$\begin{aligned} \frac{k_n(x)-1}{\pi(x_n)} - \frac{k_{n+1}(x)-1}{\pi(x_{n+1})} &\geq 34 \cdot (r_n(x) - r_{n+1}(x)) r_0(x) - \frac{1}{\pi(x_n)} \\ &\geq 34 \cdot 2^{n/2} \sqrt{d(x_n, x_{n+1})} r_0(x) - \frac{1}{\pi(x_n)}. \end{aligned}$$

Since $x_n \neq x_{n+1}$ and $\sqrt{d(x_n, x_{n+1})} \leq r_0(x)$ we have

$$\frac{1}{\pi(x_n)} \leq \sqrt{d(x_n, x_{n+1})} r_0(x).$$

Therefore

$$\begin{aligned} \frac{k_n(x)-1}{\pi(x_n)} - \frac{k_{n+1}(x)-1}{\pi(x_{n+1})} &\geq (34 \cdot 2^{n/2} - 1) \sqrt{d(x_n, x_{n+1})} r_0(x) \\ &\geq 33 \cdot 2^{n/2} \sqrt{d(x_n, x_{n+1})} r_0(x). \end{aligned}$$

Also

$$\frac{k_n(x)-1}{\pi(x_n)} \leq 34 \cdot r_n(x) r_0(x) \leq 34 \cdot r_0(x)^2.$$

Since $33^2/(4 \cdot 34) \geq 2^3$, combining (12) with the last two inequalities yields (13).
The number of possible couples (x_n, x_{n-1}) is at most $N_n N_{n+1}$. Recall the definition (1) of N_n and observe that $N_n^2 \leq N_{n+1}$ for all n. A union bound shows that the probability that there exists x and n such that

$$T^{k_{n+1}(x)}(x_{n+1}) \geq T^{k_n(x)}(x_n)$$

is at most

$$\sum_{n\geq 0}\frac{N_nN_{n+1}}{N_{n+3}}\leq\sum_{n\geq 0}\frac{1}{N_{n+2}}\leq\sum_{n\geq 4}2^{-n}=\frac{1}{8}.$$

Therefore with probability at least $7/8$, we have

$$T^{k_{n+1}(x)}(x_{n+1})\leq T^{k_n(x)}(x_n)$$

for all x and n, hence

$$T^{k_n(x)}(x_n)\leq T^{k_0(x)}(x_0)=T^{k_0}(z).$$

Since $x_n=x$ for n large enough and $k_n(x)\geq 1$ we obtain

$$\forall x\in A,\ T(x)\leq T^{k_0}(z)$$

with probability $7/8$ at least. In other words

$$\mathsf{P}_z(T_{cov}(A)\leq T^{k_0}(z))\geq\frac{7}{8}.$$

Together with Lemma 7 we get

$$\mathsf{E}_z\,T_{cov}(A)\leq\frac{8k_0}{7\pi(z)}\leq\frac{8}{7}(34\cdot r_0^2+\frac{1}{\pi(z)}).$$

Unless $A=\{z\}$, in which case $\mathrm{cov}(A)=0$ and there is nothing to prove, we have $1/\pi(z)\leq\Delta(A)\leq r_0^2$. Therefore

$$\mathsf{E}_z\,T_{cov}(A)\leq\frac{8\cdot 35}{7}\Big(\sup_{x\in A}\sum_{n=0}^{\infty}2^{n/2}\sqrt{\Delta(A_n(x))}\Big)^2.$$ □

3 The lower bound

We start this section with another definition; given $A\subset M$ let

$$\mathrm{cov}_-(A)=\min_{x\in A}\mathsf{E}_x\,T_{cov}(A)$$
$$\mathrm{cov}_+(A)=\max_{x\in A}\mathsf{E}_x\,T_{cov}(A).$$

Note that the cover time of A, which was previously denoted by $\mathrm{cov}(A)$, is now denoted by $\mathrm{cov}_+(A)$ to avoid confusions with $\mathrm{cov}_-(A)$. In this section we prove the following

Proposition 9. *Let $(X_n)_{n\geq 0}$ be an irreducible, positive recurrent Markov chain on a discrete state space M. If the chain is reversible then for every finite subset A of M*

$$\gamma_1(A,d)\leq L(\mathrm{cov}_-(A)+\Delta(A,d)),$$

where L is a universal constant.

Remarks. 1. This yields Theorem 5 since clearly

$$\begin{aligned}\mathrm{cov}_-(A) &\le \mathrm{cov}_+(A)\\ \Delta(A,d) &\le \mathrm{cov}_+(A).\end{aligned}$$

2. The term $\Delta(A,d)$ cannot be removed from the inequality. Indeed if $M=\{0,1\}$ and the transitions are given by the matrix

$$\begin{pmatrix}\epsilon & 1-\epsilon\\ \epsilon & 1-\epsilon\end{pmatrix}$$

then

$$\gamma_1(M,d) \ge \Delta(M,d) = \frac{1}{\epsilon(1-\epsilon)}$$

whereas $\mathrm{cov}_-(M) = \min(\frac{1}{\epsilon}, \frac{1}{1-\epsilon})$.

3.1 Talagrand's growth condition

Recall the majoring measure theorem: if $(Y_s)_{s\in S}$ is a centered Gaussian process then

$$\gamma_2(S,d) \le L\,\mathsf{E}\sup_{s\in S} Y_s,$$

where d is the L^2 distance (3). The proof of Talagrand consists in showing (using Gaussian concentration and Sudakov's inequality) that the functional

$$A \mapsto \mathsf{E}\sup_{s\in A} Y_s$$

satisfies an abstract growth condition, and that such functionals dominate γ_2. Here is the definition of the growth condition adapted to the γ_1 situation (rather than γ_2).

Definition 10 (Growth condition)**.** Let (M,d) be a metric space. A functional $F\colon \mathcal{P}(M) \to \mathbb{R}_+$ is said to satisfy the growth condition with parameters $r>1$ and $\tau \in \mathbb{N}$ if for every step $n\in\mathbb{N}$ and every scale $a>0$ the followings holds. Let $m = N_{n+\tau}$, for every sequence $H_1,\dots,H_m$ of non-empty subsets of M satisfying

1. $\Delta(\cup_{i\le m} H_i) \le ra$,
2. $d(H_i,H_j) \ge a$ for all $i\neq j$,
3. $\Delta(H_i) \le a/r$ for all i,

we have

$$F(\cup_{i\le m} H_i) \ge a2^n + \min_{i\le m} F(H_i).$$

Theorem 11. *If F is a non-decreasing for the inclusion and satisfies the growth condition with parameters r and τ then*

$$\gamma_1(M,d) \le L2^\tau(\Delta(M,d) + rF(M)),$$

where L is a universal constant.

We refer to [7, Theorem ??] for a proof of this theorem. The purpose of the rest of this section is to show that the functional

$$A \mapsto \mathrm{cov}_-(A)$$

is non-decreasing and satisfies the growth condition on (M, d) (where d is the commute distance) with universal parameters τ and r.

Lemma 12. *The functional $A \mapsto \mathrm{cov}_-(A)$ is non-decreasing for the inclusion.*

Proof. We use the strong Markov property. The shift operator is denoted by σ: for every integer k

$$\sigma_k(X_0, X_1, \dots) = (X_k, X_{k+1}, \dots).$$

Let $A \subset B$ and let $x \in B$. Then

$$T_{cov}(B) \geq T(A) + T_{cov}(A) \circ \sigma_{T(A)}.$$

In words: at time $T(A)$ the chain has yet to visit every point of $A \backslash \{X_{T(A)}\}$. By the strong Markov property

$$\begin{aligned} \mathsf{E}_x\, T_{cov}(B) &\geq \mathsf{E}_x\, T(A) + \mathsf{E}_x\big[\mathsf{E}_{X_{T(A)}}\, T_{cov}(A)\big]. \\ &\geq \mathsf{E}_x\, T(A) + \mathrm{cov}_-(A) \\ &\geq \mathrm{cov}_-(A), \end{aligned}$$

which is the result. □

3.2 Variations on Matthews' bound

The following is due to Matthews [6].

Lemma 13. *Let A be a finite subset of M, let $a > 0$ and assume that $\mathsf{E}_x\, T(y) \geq a$ for every $x \neq y$ in A. Then*

$$\mathrm{cov}_-(A) \geq a \sum_{k=1}^{|A|-1} \frac{1}{k} \geq a \log(|A|).$$

Proof. Let $x \in A$. Assuming that $|A| \geq 2$ (otherwise the result is trivial) we have

$$\sum_{y \in A, y \neq x} \mathsf{P}_x\big(T_{cov}(A) = T(y)\big) = 1.$$

So there exists $y \in A$ such that

$$\mathsf{P}_x\big(T_{cov}(A) = T(y)\big) \geq \frac{1}{|A| - 1}. \tag{14}$$

Let $A' = A \backslash \{y\}$, let $S = T_{cov}(A')$ and let $T = T_{cov}(A)$. Clearly

$$T = S + (T(y) \circ \sigma_S) \mathbf{1}_{\{S < T(y)\}}.$$

By the strong Markov property

$$\mathsf{E}_x\, T = \mathsf{E}_x\, S + \mathsf{E}_x\big[\big(\mathsf{E}_{X_S}\, T(y)\big)\mathbf{1}_{\{S<T(y)\}}\big].$$

On the event $\{S < T(y)\}$ the point X_S is an element of A different from y. Therefore $\mathsf{E}_{X_S}\, T(y) \geq a$. Together with (14) we obtain

$$\mathsf{E}_x\, T_{cov}(A) \geq \mathsf{E}_x\, T_{cov}(A') + \frac{a}{|A|-1}.$$

An obvious induction on $|A|$ finishes the proof. □

The following lemma is proved the same way.

Lemma 14. *Let $H_1, \ldots, H_m$ be non-empty subsets of M satisfying*

$$\mathsf{E}_x\, T(y) \geq a, \forall (x,y) \in H_i \times H_j,\ \forall i \neq j.$$

Then for all $x \in \cup_{i\leq m} H_i$

$$\mathsf{E}_x \max_{i\leq m} T(H_i) \geq a \log(m).$$

An additional application of the strong Markov property yields the following refinement of the previous lemma.

Proposition 15. *Let $H_1, \ldots, H_m$ be non-empty subsets of M satisfying $\mathsf{E}_x\, T_y \geq a$ for all $(x,y) \in H_i \times H_j$, for all $i \neq j$. Then*

$$\mathrm{cov}_-\big(\bigcup_{i\leq m} H_i\big) \geq a\log(m) + \min_{i\leq m} \mathrm{cov}_-(H_i). \tag{15}$$

Proof. Let $x \in \cup_{i\leq m} H_i$. Let $S = \max_{i\leq m} T(H_i)$ and $T = T_{cov}(\cup_{i\leq m} H_i)$. If $S = T(H_i)$ then at time S the chain has yet to visit every point of $H_i \backslash \{X_S\}$. Therefore

$$T \geq S + \sum_{i=1}^{m} \big(T_{cov}(H_i)\circ\sigma_S\big)\mathbf{1}_{\{S=T(H_i)\}}$$

Using the strong Markov property, we get

$$\begin{aligned}\mathsf{E}_x\, T &\geq \mathsf{E}_x\, S + \sum_{i=1}^{m} \mathsf{E}_x\big[\big(\mathsf{E}_{X_S}\, T_{cov}(H_i)\big)\mathbf{1}_{\{S=T(H_i)\}}\big]\\ &\geq \mathsf{E}_x\, S + \min_{i\leq m} \mathrm{cov}_-(H_i).\end{aligned}$$

Together with the previous lemma we get the result. □

We are close to the desired growth condition. We would like to obtain the inequality (15) under the weaker hypothesis

$$d(x,y) = \mathsf{E}_x\, T(y) + \mathsf{E}_y\, T(x) \geq a,\ \forall x, y \in H_i \times H_j, i \neq j.$$

This is done in the next section. Roughly speaking, reversibility insures that for a reasonable proportion of x and y the hitting times $\mathsf{E}_x\, T(y)$ and $\mathsf{E}_y\, T(x)$ are of the same of order of magnitude.

3.3 Reversibility

Again this part of the argument is taken from Kahn, Kim, Lovasz and Vu's article [4]. We start with a simple lemma concerning directed graphs. Given a directed graph $G = (V, E)$, a path of G is a sequence $x_1, \dots, x_m$ of vertices satisfying $(x_i, x_{i+1}) \in E$ for $i \leq m$. The length of such a path is defined to be m. An independent set is a subset A of V satisfying $(x, y) \notin E$ for all x, y in A.

Lemma 16. *If every path of G has length at most m then G has an independent set of cardinality at least $|V|/m$.*

This is standard, but we still sketch the argument. It is easy to show by induction on m that G is then m-colorable: it is possible to map the vertices of G to $\{1, \dots, m\}$ in such a way that connected points have different images. Then by the pigeon hole principle, at least $|V|/m$ vertices have the same image, which is the result.
From now on the chain $(X_n)_{n\geq 0}$ is assumed to be reversible. Consequently, we have the following commuting property for hitting times.

Lemma 17. *For every sequence $x_1, \dots, x_m$ of elements of M we have*

$$\begin{aligned} &\mathsf{E}_{x_1} T(x_2) + \cdots + \mathsf{E}_{x_{m-1}} T(x_m) + \mathsf{E}_{x_m} T(x_1) \\ &\qquad = \mathsf{E}_{x_1} T(x_m) + \mathsf{E}_{x_m} T(x_{m-1}) + \cdots + \mathsf{E}_{x_2} T(x_1). \end{aligned} \tag{16}$$

We refer to [5, Lemma 10.10] for a proof.

Corollary 18. *Let A be a subset of M and $a > 0$. If $\Delta(A, d) \leq 16a$ and if $d(x, y) \geq a$ for all $x \neq y$ in A then there exists a subset A' of A satisfying*

- $|A'| \geq |A|/33$.
- $\mathsf{E}_x T(y) \geq a/4$ *for all* $x \neq y$ *in* A'.

Proof. We define a graph G with vertex set A by saying that the edge (x, y) is present if $x \neq y$ and $\mathsf{E}_x T(y) \leq a/4$. Let $x_1, \dots, x_m$ be a path of G. Then the inequalities

$$\begin{aligned} \mathsf{E}_{x_i} T(x_{i+1}) &\leq a/4 \\ \mathsf{E}_{x_{i+1}} T(x_i) &\geq 3a/4 \end{aligned}$$

and equation (16) give

$$\frac{(m-1)a}{4} + \mathsf{E}_{x_m} T(x_1) \geq \frac{3(m-1)a}{4} + \mathsf{E}_{x_1} T(x_m).$$

Together with the bound on the diameter of A we obtain $m - 1 \leq 32$. Therefore G has an independent set of cardinality at least $|A|/33$. This is our set A'. □

3.4 The growth condition for the cover time

Proposition 19. *The functional $A \mapsto \mathrm{cov}_-(A)$ satisfies the growth condition with parameters $r = 16$ and $\tau = 5$.*

Proof. Let $n \in \mathbb{N}$, let $a > 0$ and $m = N_{n+5}$. Let $H_1, \dots, H_m$ satisfy

1. $\Delta(\cup_{i \le m} H_i) \le 16a$.
2. $d(H_i, H_j) \ge a$, for all $i \ne j$.
3. $\Delta(H_i) \le a/16$ for all $i \le m$.

Let $x_1, \dots, x_m$ belong to $H_1, \dots, H_m$ respectively. By the first two properties and Corollary 18, there exists a subset I of $\{1, \dots, m\}$ satisfying

- $|I| \ge m/33$.
- $\mathsf{E}_{x_i} T(x_j) \ge a/4$ for every $i \ne j$ in I.

Let $i \ne j$ in I and let $(x, y) \in H_i \times H_j$. Then

$$\mathsf{E}_x T(y) \ge \mathsf{E}_{x_i} T(x_j) - \mathsf{E}_{x_i} T(x) - \mathsf{E}_y T(x_j) \ge \frac{a}{4} - \frac{a}{16} - \frac{a}{16} = \frac{a}{8}.$$

Proposition 15 gives

$$\operatorname{cov}_-\big(\bigcup_{i \in I} H_i\big) \ge \frac{a}{8} \log(|I|) + \min_{i \in I} \operatorname{cov}_-(H_i).$$

Since

$$|I| \ge N_{n+5}/33 \ge N_{n+5}/N_3 \ge N_{n+4} \ge \mathrm{e}^{8 \cdot 2^n}$$

we obtain

$$\operatorname{cov}_-\big(\bigcup_{i \le m} H_i\big) \ge a 2^n + \min_{i \le m} \operatorname{cov}_-(H_i),$$

which is the result. □

Then, by Theorem 11 we obtain

$$\gamma_1(M, d) \le L(\operatorname{cov}_-(M) + \Delta(M, d)).$$

Obviously we can replace M by any subset A of M in this inequality: if a functional F satisfy the growth condition on (M, d) then it also satisfies it on (A, d).

Appendix

We have seen in the introduction that for any metric space (M, d)

$$\big[\gamma_2(M, \sqrt{d})\big]^2 \le C \log(\log|M|) \gamma_1(M, d). \tag{17}$$

We show in this appendix that this is sharp and that the example saturating the inequality can be chosen to be the state space of a reversible Markov chain equipped with the commute distance. The example is taken from [4] and was pointed out to the author by James Lee.
Let M be a rooted tree of depth D (large enough) satisfying

- nodes at depth $i \leq D-1$ have $N_i + 1$ children,
- edges between depth i and depth $i+1$ have multiplicity 2^i,

and let X be the random walk on this graph. The stationary measure is given by $\pi(x) = d(x)/2E$ for every x, where $d(x)$ is the number of edges (counted with multiplicity) starting from x and E is the total number of edges. Besides π is reversible. Let us compute the commute distance d. Because of the tree structure it is easily seen that

$$d(x,y) = \sum_{i=0}^{n-1} d(x_i, x_{i+1}) \tag{18}$$

where $x_0, \ldots, x_n$ is the shortest path from x to y. Therefore it is enough to compute $d(x,y)$ when x and y are neighbors, in which case we use the formula (see [1])

$$\mathsf{P}_x(T(y) < T^1(x)) = \frac{1}{\pi(x)d(x,y)}.$$

Because of the tree structure again $\mathsf{P}_x(T(y) < T^1(x))$ is just the transition probability from x to y. We obtain

$$d(x,y) = 2E \cdot 2^{-i}$$

when (x,y) is an edge between depth i and depth $i+1$. When x and y are any two nodes of M, equality (18) then implies that

$$E \cdot 2^{-i+1} \leq d(x,y) \leq E \cdot 2^{-i+3} \tag{19}$$

where i is the depth of their closest common ancestor.

Proposition 20. *There is a universal constant C such that*

$$\frac{D \cdot E}{C} \leq \gamma_1(M,d) \leq C \cdot D \cdot E \tag{20}$$

$$\frac{D \cdot \sqrt{E}}{C} \leq \gamma_2(M, \sqrt{d}) \leq C \cdot D \cdot \sqrt{E}. \tag{21}$$

Since D is of the order of $\log(\log|M|)$, this shows that (17) is sharp (up to the constant).

Proof. Let us start with the upper bound of (20). It is more convenient to use the following definition for γ_1:

$$\gamma_1(M,d) = \inf \sup_{x \in M} \sum_{i=0}^{+\infty} 2^i d(x, M_i)$$

where the infimum is taken over every sequence $(M_i)_{i \in \mathbb{N}}$ of subsets of M satisfying the cardinality condition $|M_i| \leq N_i$ for every i. It is well known (see [7]) that this definition coincides with the one with partitions, up to a universal factor.

For $0 \leq i \leq D$ let S_i be the set of vertices of depth at most i. Using (19) we obtain $d(x, S_i) \leq E \cdot 2^{-i+3}$ for every $x \in M$. Therefore

$$\sup_{x \in M} \sum_{i=0}^{+\infty} 2^i d(x, S_i) \leq E \sum_{i=0}^{D} 2^i 2^{-i+3} = 8E \cdot (D+1).$$

Besides, it is easily shown that

$$|S_i| \leq N_{i+3}.$$

The sequence $(S_i)_{n \in \mathbb{N}}$ does not quite satisfies the right cardinality condition, but this is not a big deal. If we shift the sequence by letting $M_0 = M_1 = M_2 = S_0$ and $M_i = S_{i-3}$ for $i \geq 3$, we still have

$$\sup_{x \in M} \sum_{i=0}^{+\infty} 2^i d(x, M_i) \leq C \cdot E \cdot D$$

for some universal C, which proves the upper bound of (20).
To prove the lower bound we need to show that the previous sequence of approximations is essentially optimal. Let $(M_i)_{i \geq 0}$ be a sequence of subsets of M satisfying $|M_i| \leq N_i$ for every i. A vertex x of depth $i \leq D-1$ has $N_i + 1$ children. So at least one them, call it y, has the following property: neither y nor any of its offsprings belong to M_i. Using this observation, we can construct inductively a sequence $x_0, x_1, \ldots, x_D$, where x_0 is the root of M and such that

- x_{i+1} is a child of x_i,
- neither x_{i+1} nor any of its offsprings belong to M_i,

for every $i \leq D-1$. Let $i \leq D-1$ and let $x \in M_i$. Since x is not an offspring of x_{i+1} we have $d(x, x_D) \geq E \cdot 2^{-i+1}$. Thus

$$\sum_{i=0}^{\infty} 2^i d(x_D, M_i) \geq E \sum_{i=0}^{D-1} 2^i 2^{-i+1} = 2E \cdot D,$$

which proves the lower bound of (20).
Inequality (21) is proved exactly the same way. □

Acknowledgement

This work owes a lot to stimulating discussions the author had with Guillaume Aubrun.

References

[1] D. Aldous and J.A. Fill, *Reversible Markov chains and random walks on graphs*, in preparation, online version at the following web page.
`http://www.stat.berkeley.edu/users/aldous/RWG/book.html`

[2] M.T. Barlow, J. Ding, A. Nachmias and Y. Peres, The evolution of the cover time, Combin. Probab. Comput. 20 (3) (2011) 331–345.

[3] J. Ding, J.R. Lee and Y. Peres, Cover times, blanket times, and majorizing measures, Ann. of Math. 175 (3) (2012) 1409–1471.

[4] J. Kahn, J.H. Kim, L. Lovász and V.H. Vu, The cover time, the blanket time, and the Matthews bound, in *41st Annual Symposium on Foundations of Computer Science*, IEEE Comput. Soc. Press, Los Alamitos, 2000, pp. 467–475.

[5] D.A. Levin, Y. Peres and E.L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, 2009.

[6] P. Matthews, Covering problems for Brownian motion on spheres, Ann. Probab. 16 (1) (1988) 189–199.

[7] M. Talagrand, *The generic chaining*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2005.