Combinatorial Proof of the Identity
for Cover Times on Finite Graphs

Takuya OHWA[*]

Graduate School of Mathematics, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

We give a combinatorial proof and an extension of the identity for the probability density of the cover time of a random walk on a finite graph obtained in [10] by considering it as the Möbius inversion formula. In addition, we obtain a similar identity for the cover time of multiple random walks.

KEYWORDS: Markov chain, Cover time, Möbius inversion formula

1. Introduction

The cover time of a random walk on a finite graph is the minimal time taken to visit all the vertices of the graph. It has been extensively studied and many works have been devoted to obtaining the estimation of the expectation of the cover time. For instance, Feige showed in [6, 7] that the expected cover times of simple random walks for any connected graphs on N vertices are bounded between the fastest case $O(N \log N)$ and the slowest one $O(N^3)$. These bounds are optimal in the sense that the former is attained by the complete graph and the latter by the lollipop graph (a path graph on $\frac{N}{3}$ vertices which is connected to a complete graph on $\frac{2N}{3}$ vertices). On the other hand, Matthews showed another type of bound in [9] in terms of the minimal and maximal expected hitting times. It is much easier to compute hitting times than cover times, so this bound is applied to several problems (cf. [4]). Among many works concerning cover times other than the basic results above, there are few results for distributions of cover times comparing with expectations (cf. [1, 5]). In the previous paper [10], we gave a representation of the joint distribution of the cover time and the last visited point as the alternating sum of those of exit times and exit points, which was obtained by using the spectral theory for a reversible random walk. In the present paper, we give a simple proof of this representation by regarding it as Möbius inversion. In addition, this approach gives us the following: we succeed in dropping the condition of the reversibility of the random walk in [10] and in obtaining a similar representation for cover times of multiple random walks which is discussed in [2, 3].

Let $G = (V(G), E(G))$ be a finite undirected and connected graph with the set of vertices $V(G)$ and the set of edges $E(G)$. We simply write $V = V(G)$ and set the cardinality of V as $|V| = N$. We assume $N \geq 2$ throughout this paper. For each subset W of V, the subgraph of G induced by W is the graph with the set of vertices W and the set of edges $\{\{x, y\} \in E(G); x, y \in W\}$, and $N_G(W) = \{v \in W^c; \{v, w\} \in E(G) \text{ for some } w \in W\}$ is the set of neighbors of W in G where $W^c = V \setminus W$.

We fix a root $r \in V$ and mainly treat a sequence of real numbers $(\sigma_a)_{a \in V}$, a real valued function defined on V, which has the property below:

  (A) For every $v \in V$ there exists a walk from r to v in G, denoted by $v_1 \cdots v_n$ where $v_1 = r$ and $v_n = v$, such that $\sigma_{v_i} \leq \sigma_v$ for all $i = 1, \ldots, n$.

Here a walk in G is a sequence $a_1 \cdots a_n$ of vertices such that $\{a_i, a_{i+1}\} \in E(G)$ for all $i \in \{1, \ldots, n-1\}$.

The following example shows that this property **(A)** actually holds for cover times.

Example 1. Let $(w_t)_{t=0,1,\ldots}$ be a discrete time random walk on G starting from $r \in V$; when the random walk is at $v \in V$, the next position is chosen from the neighbors of v according to any given transition probability. The sequence of the first hitting time to $v \in V$, defined as

$$\sigma_v = \inf\{s \geq 0; w_s = v\},$$

has the property **(A)** (if σ's are finite). Then $\max_{a \in V} \sigma_a$ is merely the time taken by the random walk to cover all the vertices of G, and for a subset $\Lambda \subset V$, $\min_{b \notin \Lambda} \sigma_b = \inf\{s \geq 0; w_s \in \Lambda^c\}$ is the time at which w_t exits from Λ. The quantity $\max_{a \in V} \sigma_a$ is called the cover time of G and $\min_{b \notin \Lambda} \sigma_b$ the first exit time from Λ.

Let $\mathcal{C}$ be the totality of vertex sets of connected induced subgraphs of G, that is, $\mathcal{C} = \{\Lambda \subset V;$ the subgraph of G induced by Λ is connected$\}$. Formally, $\mathcal{C}$ includes the emptyset. We set $\mathcal{C}_r = \{\Lambda \in \mathcal{C}; \Lambda \ni r\} \cup \{\emptyset\}$ and regard it as a partially ordered set with inclusion. The following main theorem is obtained from the Möbius inversion formula for $\mathcal{C}_r$.

Theorem 1. *Let $(\sigma_a)_{a \in V}$ be a sequence of real numbers having the property* **(A)** *for a root $r \in V$ and set $\mathcal{D}_r = \{\Lambda \in \mathcal{C}_r; \Lambda \cup N_G(\Lambda) = V\} \setminus \{V\}$. Then,*

$$\varphi\left(\max_{a \in V} \sigma_a\right) = \sum_{\Lambda \in \mathcal{D}_r} (-1)^{N - |\Lambda| + 1} \varphi\left(\min_{b \notin \Lambda} \sigma_b\right)$$

for any function $\varphi : \mathbb{R} \to \mathbb{R}$.

As a corollary, we obtain an identity for the joint distribution of the cover time and the last visited point through the expression as the Laplace transform of this distribution by taking φ as a suitable function. A random walk on a state space X with a transition probability $(p(x, y))_{x,y \in X}$ is said to be irreducible if, for any states $x, y \in X$, there exists a sequence $z_1 \cdots z_n$ with $z_1 = x$ and $z_n = y$ such that $p(z_i, z_i + 1) > 0$ for $i = 1, \ldots, n - 1$. In order to obtain $\mathcal{D}_r$ from an irreducible transition probability $(p(x, y))_{x,y \in X}$, we construct a graph G as $V(G) = X$ and $E(G) = \{\{x, y\}; p(x, y) > 0$ or $p(y, x) > 0\}$. Here we remark that such a graph is connected since the transition probability is irreducible, that is, for any vertices $x, y \in X$, there exists a walk $v_1 \cdots v_n$ with $v_1 = x$ and $v_n = y$ such that $p(v_i, v_{i+1}) > 0$ for $i = 1, \ldots, n - 1$.

Theorem 2. *Let $(w_t)_{t=0,1,\ldots}$ be an irreducible random walk on a finite state space X starting from r. We denote the cover time of X by C_X defined as $C_X = \max_{a \in X} \sigma_a$ and the first exit time from a subset $\Lambda \subset X$ by T_Λ defined as $T_\Lambda = \min_{b \notin \Lambda} \sigma_b$, where σ_v is the first hitting time to $v \in X$. Then,*

$$P_r(C_X = t, w_{C_X} = y) = \sum_{\Lambda \in \mathcal{D}_r \setminus \mathcal{D}_y} (-1)^{|X| - |\Lambda| + 1} P_r(T_\Lambda = t, w_{T_\Lambda} = y)$$

for each $y \in X$, where P_r is the law of the random walk starting from r.

We remark that Theorem 2 generalizes the main theorem in [10] in the sence that the reversibility of random walks required in [10] is dropped in Theorem 2. In addition it holds for continuous time random walks under the obvious modification although it is stated only for discrete time ones here.

In order to show that the expression in Theorem 1 is useful in calculating cover times, we give the following examples of $\mathcal{D}_r$ for some graphs.

Example 2.
(i) If G is a path with end-vertices v_1 and v_2, then it holds that

$$\mathcal{D}_r = \begin{cases} \{\{v_2\}^c\}, & r = v_1, \\ \{\{v_1\}^c\}, & r = v_2, \\ \{\{v_1\}^c, \{v_2\}^c, \{v_1, v_2\}^c\}, & \text{otherwise.} \end{cases}$$

In general, if G is a tree, then

$$\mathcal{D}_r = \begin{cases} \{W^c; W \subset L, W \neq \emptyset\}, & r \notin L, \\ \{W^c; W \subset L, W \not\ni r, W \neq \emptyset\}, & r \in L, \end{cases}$$

where $L = \{v \in V; |N_G(\{v\})| = 1\}$ is the set of leaves of G.
(ii) If G is a cycle, then

$$\mathcal{D}_r = \{\{v\}^c; v \in V, v \neq r\} \cup \{\{v, w\}^c; \{v, w\} \in E(G), v \neq r, w \neq r\}.$$

(iii) If r is adjacent to any other vertices in G, then $\mathcal{D}_r = \{\Lambda \subsetneq V; \Lambda \ni r\}$.

By using Example 2, let us show the expectation of the cover time of a random walk on a cycle. We denote the set of vertices by $X = V = \{0, \ldots, N - 1\}$ and the set of edges by $E(G) = \{\{i, i + 1\}; i = 0, \ldots, N - 2\} \cup \{\{N - 1, 0\}\}$. Fix $0 < p < 1$ and we define the transition probability $(p(x, y))_{x,y \in X}$ of the random walk by

$$p(x, y) = \begin{cases} p, & y = x + 1 \bmod N, \\ 1 - p, & y = x - 1 \bmod N, \\ 0, & \text{otherwise.} \end{cases}$$

From Theorem 2 and (ii) in Example 2, we have

$$E_0[C_X] = \sum_{n=1}^{N-1} E_0[T_{\{n\}^c}] - \sum_{n=1}^{N-2} E_0[T_{\{n,n+1\}^c}],$$

where E_0 is the expectation with respect to P_0. We mention that $E_0[T_{\{n\}^c}]$ and $E_0[T_{\{n,n+1\}^c}]$ are easily obtained since they satisfy some recurrence equation from the strong Markov property (cf. [8]). If $p = \frac{1}{2}$, then we have

$$E_0[C_X] = \sum_{n=1}^{N-1} n(N-n) - \sum_{n=1}^{N-2} n(N-1-n) = \frac{N(N-1)}{2}.$$

If $p \neq \frac{1}{2}$, then we have

$$E_0[C_X] = \sum_{n=1}^{N-1} \frac{1}{p-q}\left(n - \frac{N(p^n q^{-n} - 1)}{p^N q^{-N} - 1}\right) - \sum_{n=1}^{N-2} \frac{1}{p-q}\left(n - \frac{(N-1)(p^n q^{-n} - 1)}{p^{N-1} q^{-(N-1)} - 1}\right)$$

$$= -\frac{1}{(p^N - q^N)(p^{N-1} - q^{N-1})}\left(N(pq)^{\frac{N-1}{2}} - \frac{p^N - q^N}{(p-q)}\right)^2 + \frac{N(p^{\frac{N-1}{2}} - q^{\frac{N-1}{2}})}{(p-q)(p^{\frac{N-1}{2}} + q^{\frac{N-1}{2}})} \tag{1.1}$$

where $q = 1 - p$. From (1.1), we can verify that $E_0[C_X] \to \frac{N(N-1)}{2}$ as $p \to \frac{1}{2}$ and $E_0[C_X] \sim \frac{N}{|p-q|}$ as $N \to \infty$, where $f(N) \sim g(N)$ means that $\lim_{N\to\infty} \frac{f(N)}{g(N)} = 1$.

2. Equivalence conditions for the property (A)

Without loss of generality, we may assume $\sigma_r = 0$. In this section, we provide two kinds of conditions equivalent to the property **(A)**. Before stating these, we give the following notations:

- For a tree T with a fixed root x, we define a partial order $\leq_T$ by $v \leq_T v'$ on T if $v \in V(xPv')$, where xPv' is the unique path between x and v' in T.
- Let $\ell : E(G) \to [0, \infty)$ be a non-negative weight on $E(G)$. We define a pseudo distance with this weight ℓ in G by

$$d_\ell(v, v') = \min_{vPv' \subset G} \sum_{e \in E(vPv')} \ell(e) \tag{2.1}$$

for each $v, v' \in V$.

A subgraph G' of G is said to be a spanning subgraph if $V(G') = V(G)$. In particular, if a spanning subgraph is a tree, then it is called a spanning tree. We claim the proposition below.

Proposition 1. *The following assertions are equivalent for a sequence $(\sigma_a)_{a \in V}$ of real numbers on a connected graph G with a fixed root r:*

(i) *The sequence $(\sigma_a)_{a \in V}$ satisfies the property **(A)** for the root r;*
(ii) *There exists a spanning tree T of G with the root r such that $\sigma_v \leq \sigma_{v'}$ if $v \leq_T v'$;*
(iii) *There exists a non-negative weight ℓ on $E(G)$ such that $\sigma_a = d_\ell(r, a)$, where $d_\ell(r, \cdot)$ is a pseudo distance from r defined by ℓ.*

Proof. Firstly, under the assumption that $(\sigma_a)_{a \in V}$ satisfies the property **(A)**, we will construct a spanning tree T satisfying the condition in (ii) in a similar way to the depth-first search. Starting from the root r, we trace the edges of G to a vertex of the neighbors of the current vertex, which has not yet been visited and whose value is not less than that of the current vertex. Whenever such a neighbor of the current vertex exists, we repeat this procedure. If there is no such vertex, we go back along the edge by which the current vertex was first reached; we repeat this tracing back until such a neighbor of the current vertex appears. Once such a neighbor appears, we pursue the previous procedure. We repeat these procedures. If there exist no such neighbors and the current vertex is r, then the traversal is terminated. Clearly the graph formed by all the traversed edges is a tree T of G. In addition, we can easily check that $\sigma_v \leq \sigma_{v'}$ if $v \leq_T v'$ for the partial order defined by this tree T and the root r. Since the property **(A)** guarantees such a T is the spanning subgraph of G, it is proved that the condition (i) implies (ii).

Secondly, we assume that a spanning tree T of G with the root r satisfies the condition (ii). Let us assign a non-negative weight as

$$\ell(\{v, v'\}) = \begin{cases} \sigma_{v'} - \sigma_v, & \text{if } \{v, v'\} \in E(T) \text{ and } v \leq_T v', \\ \max_{a \in V} \sigma_a + 1, & \text{otherwise}, \end{cases}$$

for all $\{v, v'\} \in E(G)$. Then, it is clear that $\sigma_a = d_\ell(r, a)$ holds for all $a \in V$.

Finally, let us assume that $(\sigma_a)_{a \in V}$ is expressed as $\sigma_a = d_\ell(r, a)$ for some non-negative weight ℓ. Then there exists a path $r\tilde{P}v$ between r and v in G such that $\sigma_v = \sum_{e \in E(r\tilde{P}v)} \ell(e)$. For each $v' \in V(r\tilde{P}v)$, let $r\tilde{P}v'$ be a path between r and v' in $r\tilde{P}v$. Then,

$$\sigma_v = \sum_{e \in E(r\tilde{P}v)} \ell(e) \geq \sum_{e \in E(r\tilde{P}v')} \ell(e) \geq \min_{rPv' \subset G} \sum_{e \in E(rPv')} \ell(e) = \sigma_{v'}.$$

We have thus proved the proposition. $\square$

3.   Möbius inversion

Let us recall the Möbius inversion formula (cf. [11]). Let $\mathscr{P}$ be a finite partially ordered set with a partial order $\preceq$ and f a function on $\mathscr{P}$. If a function g is given by

$$g(x) = \sum_{y \preceq x} f(y)$$

for each $x \in \mathscr{P}$, then f can be expressed as

$$f(y) = \sum_{x \preceq y} g(x) \mu_{\mathscr{P}}(x, y)$$

for all $y \in \mathscr{P}$, where $\mu_{\mathscr{P}}$ is called the Möbius function of $\mathscr{P}$ and $\mu_{\mathscr{P}}$ is given recursively by

$$\mu_{\mathscr{P}}(x, y) = \begin{cases} 1, & x = y, \\ -\sum_{x \preceq z \prec y} \mu_{\mathscr{P}}(x, z) = -\sum_{x \prec z \preceq y} \mu_{\mathscr{P}}(z, y), & x \prec y, \\ 0, & \text{otherwise}, \end{cases} \tag{3.1}$$

for $x, y \in \mathscr{P}$. For example, let $\mathfrak{P}$ be the power set of V with inclusion as a partial order. Then we have

$$\mu_{\mathfrak{P}}(A, B) = \begin{cases} (-1)^{|B|-|A|} & A \subset B, \\ 0, & \text{otherwise}, \end{cases} \tag{3.2}$$

which is nothing but the principle of inclusion-exclusion. We remark that for any $x, y \in \mathscr{P}$ the restriction to the interval $[x, y] = \{z \in \mathscr{P}; x \preceq z \preceq y\}$ of the Möbius function of $\mathscr{P}$ equals the Möbius function of $[x, y]$, that is, $\mu_{\mathscr{P}}|_{[x,y]} = \mu_{[x,y]}$.

First, we define a function f which may be applied to the proof of Theorem 1.

Proposition 2.   *Let $(\sigma_a)_{a \in V}$ be a sequence of real numbers and $\mathscr{P}$ be a family of subsets of V with inclusion as a partial order which satisfies the following two conditions:*
  (i) $\emptyset \in \mathscr{P}$;
  (ii) $\{v' \in V; \sigma_{v'} \le \sigma_v\} \in \mathscr{P}$ for all $v \in V$.
For a function $\varphi : \mathbb{R} \to \mathbb{R}$, we define the function f on $\mathscr{P}$ as

$$f(A) = \begin{cases} \varphi\left(\min_{b \in V} \sigma_b\right), & A = \emptyset, \\ \left(\varphi\left(\min_{b \notin A} \sigma_b\right) - \varphi\left(\max_{a \in A} \sigma_a\right)\right)\chi(A), & A \ne \emptyset, A \ne V, \\ 0, & A = V, \end{cases}$$

where

$$\chi(A) = \begin{cases} 1, & \min_{b \notin A} \sigma_b > \max_{a \in A} \sigma_a, \\ 0, & \text{otherwise}, \end{cases}$$

and $g(\Lambda) = \sum_{A \subset \Lambda} f(A)$ for $\Lambda \in \mathscr{P}$. Then we have

$$g(\Lambda) = \begin{cases} \varphi\left(\min_{b \notin \Lambda} \sigma_b\right), & \Lambda \ne V, \\ \varphi\left(\max_{a \in V} \sigma_a\right), & \Lambda = V, \end{cases}$$

for all $\Lambda \in \mathscr{P}$.

Proof.   For our convenience, let us set $V = \{1, \ldots, n\}$ such that $\sigma_1 \le \cdots \le \sigma_n$. The claim is trivially valid for $\sigma_1 = \sigma_n$, so we may assume that $\sigma_1 < \sigma_n$. We put

$$\widetilde{\mathscr{P}} = \{[i]; 1 \le i < n, \sigma_i < \sigma_{i+1}\},$$

where $[i] = \{1, \ldots, i\}$. Note that $\widetilde{\mathscr{P}} \subset \mathscr{P}$ by the the assumption of (ii) and $\chi(\Lambda) = 0$ if $\Lambda \notin \widetilde{\mathscr{P}}$. Moreover, we see that

$$f([i]) = \left(\varphi\left(\min_{b \notin [i]} \sigma_b\right) - \varphi\left(\max_{a \in [i]} \sigma_a\right)\right)\chi([i])$$

$$= \varphi(\sigma_{i+1}) - \varphi(\sigma_i)$$

for $[i] \in \widetilde{\mathscr{P}}$. Therefore

$$g(\Lambda) = \sum_{A \subset \Lambda} f(A) = \varphi(\sigma_1) + \sum_{i;[i] \subset \Lambda, [i] \in \widetilde{\mathcal{P}}} (\varphi(\sigma_{i+1}) - \varphi(\sigma_i))$$

for $\Lambda \in \mathcal{P}$, where the summation of the last term is understood to be 0 when $[i] \not\subset \Lambda$ for all $[i] \in \widetilde{\mathcal{P}}$. Note that $[i] \not\subset \Lambda$ for all $[i] \in \widetilde{\mathcal{P}}$ if and only if $\{i \in V; \sigma_i = \sigma_1\} \not\subset \Lambda$, that is, there exists an $i \in V$ such that $i \notin \Lambda$ and $\sigma_i = \sigma_1$. Hence, for $\Lambda \in \mathcal{P}$ with $\Lambda \not\supset \{i \in V; \sigma_i = \sigma_1\}$, we have

$$g(\Lambda) = \varphi(\sigma_1) = \varphi\left(\min_{b \notin \Lambda} \sigma_b\right).$$

For $\Lambda \in \mathcal{P}$ with $\Lambda \supset \{i \in V; \sigma_i = \sigma_1\}$, put $i_0 = \max\{i \in V; [i] \subset \Lambda, [i] \in \widetilde{\mathcal{P}}\}$. Then, $[i] \subset \Lambda$ for all $[i] \in \widetilde{\mathcal{P}}$ such that $i \leq i_0$. Thus

$$g(\Lambda) = \varphi(\sigma_1) + \sum_{i;[i] \subset \Lambda, [i] \in \widetilde{\mathcal{P}}} (\varphi(\sigma_{i+1}) - \varphi(\sigma_i)) = \varphi(\sigma_{i_0+1}).$$

Here the last equality follows from the assumption of (ii). Clearly if $\Lambda = V$, then $\sigma_{i_0+1} = \max_{a \in V} \sigma_a$ and hence

$$g(\Lambda) = \varphi(\sigma_{i_0+1}) = \varphi\left(\max_{a \in V} \sigma_a\right).$$

If $\Lambda \neq V$ and $\Lambda \supset \{i \in V; \sigma_i = \sigma_1\}$, then we see that

$$\min_{b \notin \Lambda} \sigma_b = \min_{b \notin [i_0]} \sigma_b = \sigma_{i_0+1}$$

and then we have

$$g(\Lambda) = \varphi(\sigma_{i_0+1}) = \varphi\left(\min_{b \notin \Lambda} \sigma_b\right).$$

$\square$

If $(\sigma_a)_{a \in V}$ satisfies the property (**A**) for a root $r \in V$, the subgraphs induced by $\{v' \in V; \sigma_{v'} \leq \sigma_v\}$ are connected for all $v \in V$, that is, $\{v' \in V; \sigma_{v'} \leq \sigma_v\} \in \mathcal{C}_r$. Next let us express the Möbius function for $\mathcal{C}_r$. Recall that $\mathcal{C}_r$ is a partially ordered set with inclusion.

Proposition 3. *Let $\mu_{\mathcal{C}_r}$ be the Möbius function for $\mathcal{C}_r$ and $A, B \in \mathcal{C}_r$. If $A \neq \emptyset$,*

$$\mu_{\mathcal{C}_r}(A, B) = \begin{cases} (-1)^{|B|-|A|}, & A \subset B \subset A \cup N_G(A), \\ 0, & \text{otherwise.} \end{cases}$$

If $A = \emptyset$,

$$\mu_{\mathcal{C}_r}(A, B) = \begin{cases} 1, & B = \emptyset, \\ -1, & B = \{r\}, \\ 0, & \text{otherwise.} \end{cases}$$

Proof. The statement is trivial for $A \not\subset B$ from (3.1); we give a proof for $A \subset B$ using (3.1).

Case 1. Let us assume $A = \emptyset$. The claim is obvious when $B = \emptyset$ or $B = \{r\}$. For the other case, we have

$$\mu_{\mathcal{C}_r}(\emptyset, B) = - \sum_{\emptyset \subsetneq \Lambda \subset B} \mu_{\mathcal{C}_r}(\Lambda, B) = -\left(\mu_{\mathcal{C}_r}(\{r\}, B) + \sum_{\{r\} \subsetneq \Lambda \subset B} \mu_{\mathcal{C}_r}(\Lambda, B)\right) = 0.$$

Here the first equality and the last equality follow from the second line of (3.1).

Case 2. Let us assume $A \neq \emptyset$ and $A \subset B \subset A \cup N_G(A)$. In this case, the interval $[A, B] = \{\Lambda \in \mathcal{C}_r; A \subset \Lambda \subset B\}$ is equal to $\{\Lambda \subset V; A \subset \Lambda \subset B\}$. This partially ordered subset of $\mathcal{C}_r$ is isomorphic to the power set of $B \setminus A$ with inclusion as a partial order, say $\mathfrak{P} = \{S; S \subset B \setminus A\}$: The bijection $\varphi : [A, B] \to \mathfrak{P}$ given by $\varphi(\Lambda) = \Lambda \setminus A$ is the isomorphism since $\Lambda \subset \Lambda'$ if and only if $\varphi(\Lambda) \subset \varphi(\Lambda')$. Thus we have

$$\mu_{\mathcal{C}_r}(A, B) = \mu_{[A,B]}(A, B) = \mu_{\mathfrak{P}}(\emptyset, B \setminus A) = (-1)^{|B \setminus A|} = (-1)^{|B|-|A|}.$$

Case 3. Let us assume $A \neq \emptyset$ and $A \subset B \not\subset A \cup N_G(A)$. Fix any $A \neq \emptyset$ and we will prove that $\mu_{\mathcal{P}}(A, B) = 0$ for all B such that $A \subset B \not\subset A \cup N_G(A)$ by induction on the cardinality of $|B \setminus A|$. We set

$$\mathcal{B} = \{B \in \mathcal{C}_r; A \subset B \not\subset A \cup N_G(A)\}$$
$$= \{B \in \mathcal{C}_r; A \subset B, \exists v \in B \text{ such that } v \notin A \cup N_G(A)\}. \tag{3.3}$$

We remark that $|B \setminus A| \geq 2$ for all $B \in \mathcal{B}$, which is easily verified from (3.3). For each $B \in \mathcal{B}$, we divide the summation as follows:

$$\mu_{\mathcal{C}_r}(A, B) = - \sum_{A \subset \Lambda \subsetneq B} \mu_{\mathcal{C}_r}(A, \Lambda)$$

$$= - \left(\sum_{A \subset \Lambda \subsetneq B, \Lambda \subset \widetilde{B}} + \sum_{A \subset \Lambda \subsetneq B, \Lambda \not\subset \widetilde{B}} \right) \mu_{\mathcal{C}_r}(A, \Lambda), \tag{3.4}$$

where

$$\widetilde{B} = B \cap (A \cup N_G(A)) = A \cup (B \cap N_G(A)).$$

We see that $A \subset \widetilde{B} \subsetneq B$ and $\widetilde{B} \in \mathcal{C}_r$ from the definition of $\widetilde{B}$ and (3.3). So the first summation of (3.4) is 0 as in Case 1;

$$\sum_{A \subset \Lambda \subsetneq B, \Lambda \subset \widetilde{B}} \mu_{\mathcal{C}_r}(A, \Lambda) = \sum_{A \subset \Lambda \subset \widetilde{B}} \mu_{\mathcal{C}_r}(A, \Lambda) = \mu_{\mathcal{C}_r}(A, \widetilde{B}) + \sum_{A \subset \Lambda \subsetneq \widetilde{B}} \mu_{\mathcal{C}_r}(A, \Lambda) = 0.$$

On the other hand, the second summation of (3.4) is obviously 0 for the base case $|B \setminus A| = 2$ because there is no $\Lambda \in \mathcal{C}_r$ such that $A \subset \Lambda \subsetneq B$ and $\Lambda \not\subset \widetilde{B}$. For the case $|B \setminus A| \geq 3$, using the induction hypothesis shows that

$$\mu_{\mathcal{C}_r}(A, \Lambda) = 0 \text{ for all } \Lambda \in \mathcal{C}_r \text{ such that } A \subset \Lambda \subsetneq B \text{ and } \Lambda \not\subset \widetilde{B}. \tag{3.5}$$

In order to use the induction hypothesis, we check $\Lambda \in \mathcal{B}$ and $|\Lambda \setminus A| < |B \setminus A|$. From $\Lambda \not\subset \widetilde{B}$, there exists a $v \in \Lambda$ such that $v \notin \widetilde{B} = A \cup (B \cap N_G(A))$. Moreover, such a v satisfies

$$v \notin A \cup (B \cap N_G(A)) \cup (B^c \cap N_G(A)) = A \cup N_G(A)$$

because $v \in \Lambda \subset B$. Hence $\Lambda \in \mathcal{B}$. It is clear that $|\Lambda \setminus A| < |B \setminus A|$ from $A \subset \Lambda \subsetneq B$. Thus, we can apply the induction hypothesis to all Λ such that $A \subset \Lambda \subsetneq B$ and $\Lambda \not\subset \widetilde{B}$, and obtain (3.5). $\square$

Proof of Theorem 1. From the Möbius inversion formula on $\mathcal{P} = \mathcal{C}_r$, the function f given in Proposition 2 is expressed as $f(A) = \sum_{\Lambda \subset A} g(\Lambda) \mu_{\mathcal{P}}(\Lambda, A)$ for all $A \in \mathcal{C}_r$. Especially, putting $A = V$, we have

$$0 = f(V) = \sum_{\Lambda \subset V} g(\Lambda) \mu_{\mathcal{C}_r}(\Lambda, V)$$

$$= \varphi \left(\max_{a \in V} \sigma_a \right) + \sum_{\Lambda \subsetneq V} \varphi \left(\min_{b \notin \Lambda} \sigma_b \right) \mu_{\mathcal{C}_r}(\Lambda, V),$$

and

$$\mu_{\mathcal{C}_r}(\Lambda, V) = \begin{cases} (-1)^{N-|\Lambda|}, & \text{if } \Lambda \cup N_G(\Lambda) = V, \\ 0, & \text{otherwise,} \end{cases}$$

from Proposition 3. Thus we have completed the proof of Theorem 1. $\square$

4. Application to cover times for single and multiple random walks

4.1 Single random walk

We prove Theorem 2 in this subsection.

Proof of Theorem 2. We consider an irreducible random walk $(w_t)_{t=0,1,\dots}$ starting from $r \in X$. As is seen in Example 1, we put σ_a as the first hitting time to $a \in X$. Since the state space is finite and the random walk is irreducible, σ_a is finite almost surely for all $a \in X$. So it is enough to consider the case where σ's are finite. Recall that $\max_{a \in X} \sigma_a = C_X$ is the cover time and $\min_{b \notin \Lambda} \sigma_b = T_\Lambda$ is the first exit time from $\Lambda \subset X$. Setting $|X| = N$, the identity

$$\varphi(C_X) = \sum_{\Lambda \in \mathcal{D}_r} (-1)^{N-|\Lambda|+1} \varphi(T_\Lambda) \tag{4.1}$$

holds for any function $\varphi : \mathbb{R} \to \mathbb{R}$ from Theorem 1. Then, for any fixed $\lambda > 0$ and $b \in \mathbb{R}$, we take the function φ as

$$\varphi(z) = \varphi_b(z) = e^{-\lambda z} \delta_b(z)$$

for $z \in \mathbb{R}$, where δ is Kronecker's delta, that is, $\delta_b(b) = 1$ and $\delta_b(z) = 0$ if $z \neq b$. Therefore the identity

$$e^{-\lambda C_X} \delta_{\sigma_y}(C_X) = \sum_{\Lambda \in \mathcal{D}_r} (-1)^{N-|\Lambda|+1} e^{-\lambda T_\Lambda} \delta_{\sigma_y}(T_\Lambda)$$

holds by taking $b = \sigma_y$ for $y \in X$. We can replace $\sum_{\Lambda \in \mathcal{D}_r}$ with $\sum_{\Lambda \in \mathcal{D}_r \setminus \mathcal{D}_y}$ since $\delta_{\sigma_y}(T_\Lambda) = 0$ if $\Lambda \ni y$. Thus we get the identity in the form of the Laplace transform of the joint distribution of the cover time and the last visited point by taking the expectation of the both sides. Noting that $\delta_{\sigma_y}(C_X) = 1$ if and only if $w_{C_X} = y$ and that $\delta_{\sigma_y}(T_\Lambda) = 1$ if and only if $w_{T_\Lambda} = y$, we have

$$E_r[e^{-\lambda C_X}; w_{C_X} = y] = \sum_{\Lambda \in \mathcal{D}_r \setminus \mathcal{D}_y} (-1)^{N-|\Lambda|+1} E_r[e^{-\lambda T_\Lambda}; w_{T_\Lambda} = y],$$

where E_r is the expectation with respect to P_r. This completes the proof. $\square$

4.2 Multiple random walks

We consider k-multiple random walks $(w_t^{(1)}, \ldots, w_t^{(k)})$ on a finite state space X with $|X| = N$ starting from $(r_1, \ldots, r_k)$, where each $(w_t^{(i)})_{t=0,1,\ldots}$ is a random walk on X. We do not assume that they are independent; we only assume that each random walk is irreducible. Let us construct a connected graph G as $V(G) = X$ and $E(G) = \bigcup_{i=1}^k E(G^{(i)})$, where $G^{(i)}$ is a graph constructed by $w_t^{(i)}$ as in Section 4.1. Let $\sigma_a^{(i)}$ be the first hitting time to $a \in V$ of $w_t^{(i)}$, $R = \{r_1, \ldots, r_k\}$ and $\sigma_a^R = \min_{i=1,\ldots,k} \sigma_a^{(i)} = \inf\{t \geq 0; \exists i \in \{1, \ldots, k\} \text{ s.t. } w_t^{(i)} = a\}$. Hence the cover time of multiple random walks $(w_t^{(1)}, \ldots, w_t^{(k)})$ can be defined by $C_X^R = \max_{a \in V} \sigma_a^R$. Let us set

$$\mathcal{C}_R = \left\{ \bigcup_{i=1}^k \Lambda^{(i)}; \Lambda^{(i)} \in \mathcal{C}_{r_i}, \Lambda^{(i)} \neq \emptyset \right\} \cup \{\emptyset\},$$

where $\mathcal{C}_{r_i}$ is given by $G^{(i)}$, that is, $\mathcal{C}_{r_i} = \{\Lambda \subset V; \text{the subgraph of } G^{(i)} \text{ induced by } \Lambda \text{ is connected}, \Lambda \ni r_i\} \cup \{\emptyset\}$. If the subgraph induced by R is connected, we see that $(\sigma_a^R)_{a \in V}$ has the property **(A)** for any $r = r_i$. Then Theorem 1 can be applied for $(\sigma_a^R)_{a \in V}$. On the other hand, if the subgraph induced by R is disconnected, $(\sigma_a^R)_{a \in V}$ may not have the property **(A)** for some $r = r_i$. However, for each $v \in V$, the subset $\{v' \in V; \sigma_{v'}^R \leq \sigma_v^R\}$ must be given as the union of $\Lambda^{(i)} \in \mathcal{C}_{r_i}$ for some i. Thus we have that Proposition 2 holds for the $(\sigma_a^R)_{a \in V}$ and $\mathcal{P} = \mathcal{C}_R$. We also obtain the Möbius function $\mu_{\mathcal{C}_R}$ for $\mathcal{C}_R$ in the same way as in the proof of Proposition 3: If $A \neq \emptyset$,

$$\mu_{\mathcal{C}_R}(A, B) = \begin{cases} (-1)^{|B|-|A|}, & A \subset B \subset A \cup N_G(A), \\ 0, & \text{otherwise}, \end{cases}$$

for all $B \in \mathcal{C}_R$; if $A = \emptyset$,

$$\mu_{\mathcal{C}_R}(A, B) = \begin{cases} 1, & B = \emptyset, \\ -1, & B = R, \\ 0, & \text{otherwise}, \end{cases}$$

for all $B \in \mathcal{C}_R$. Consequently, we obtain the following theorem for the $(\sigma_a^R)_{a \in V}$.

Theorem 3. *Set* $\mathcal{D}_R = \{\Lambda \in \mathcal{C}_R; \Lambda \cup N_G(\Lambda) = V\} \setminus \{\emptyset, V\}$. *Then, we have*

$$\varphi(C_X^R) = \varphi\left(\max_{a \in V} \sigma_a^R\right) = \sum_{\Lambda \in \mathcal{D}_R} (-1)^{N-|\Lambda|+1} \varphi\left(\min_{b \notin \Lambda} \sigma_b^R\right) \tag{4.2}$$

$$= \sum_{\Lambda \in \mathcal{D}_R} (-1)^{N-|\Lambda|+1} \varphi\left(\min_{i=1,\ldots,k} T_\Lambda^{(i)}\right)$$

for any function $\varphi : \mathbb{R} \to \mathbb{R}$, *where* $T_\Lambda^{(i)} = \min_{b \notin \Lambda} \sigma_b^{(i)}$ *is the first exit time of* $w^{(i)}$ *from* Λ. *In addition, we have*

$$P_R(C_X^R = t) = \sum_{\Lambda \in \mathcal{D}_R} (-1)^{N-|\Lambda|+1} P_R\left(\min_{i=1,\ldots,k} T_\Lambda^{(i)} = t\right),$$

where P_R *is the law of the random walks.*

We remark that the expression (4.2) in Theorem 3 generalizes Theorem 1. The following are examples demonstrating theorems stated above.

Example 3. We consider two random walks starting from (r_1, r_2) such that the graph constructed by the random walks is a path graph G of length $N - 1$. We denote the set of vertices by $V(G) = \{1, \ldots, N\}$, the set of edges by $E(G) = \{\{i, i + 1\}; i = 1, \ldots, N - 1\}$, and let $[i, j] = \{i, i + 1, \ldots, j\}$.

(i) If $r_1 = n$ and $r_2 = n + 1$ for some $n \in \{1, \ldots, N - 1\}$, we see that $\mathcal{C}_{\{n,n+1\}} = \{[i, j]; 1 \leq i \leq n, n + 1 \leq j \leq N\}$ and

$$\mathcal{D}_{\{n,n+1\}} = \begin{cases} \{[1, N - 1]\}, & n = 1, \\ \{[2, N]\}, & n = N - 1, \\ \{[1, N - 1], [2, N], [2, N - 1]\}, & \text{otherwise}. \end{cases}$$

Thus, if $2 \leq n \leq N - 1$, we have

$$\varphi\left(C_X^{\{n,n+1\}}\right) = \varphi\left(\min_{i=1,2} T_{[1,N-1]}^{(i)}\right) + \varphi\left(\min_{i=1,2} T_{[2,N]}^{(i)}\right) - \varphi\left(\min_{i=1,2} T_{[2,N-1]}^{(i)}\right).$$

(ii) If $r_1 = 1$ and $r_2 = N$, we see that $\mathcal{C}_{\{1,N\}} = \{[1,i] \cup [j,N]; 1 \le i, j \le N\}$ and

$$\mathcal{D}_{\{1,N\}} = \{\{n\}^c; 2 \le n \le N - 1\} \cup \{\{n, n+1\}^c; 2 \le n \le N - 2\}.$$

Thus we have

$$\varphi\left(C_X^{\{1,N\}}\right) = \sum_{n=2}^{N-1} \varphi\left(\min_{i=1,2} T_{\{n\}^c}^{(i)}\right) - \sum_{n=2}^{N-2} \varphi\left(\min_{i=1,2} T_{\{n,n+1\}^c}^{(i)}\right)$$
$$= \sum_{n=2}^{N-1} \varphi\left(\min\{\sigma_n^{(1)}, \sigma_n^{(N)}\}\right) - \sum_{n=2}^{N-2} \varphi\left(\min\{\sigma_n^{(1)}, \sigma_{n+1}^{(N)}\}\right).$$

Acknowledgement

REFERENCES

[1] Aldous, D. J., and Fill, J., Reversible Markov chains and random walks on graphs, available via: http://stat-www.berkeley.edu/ users/aldous/RWG/book.html (2001).

[2] Alon, N., Avin, C., Kouký, M., Kozma, G., Lotker, Z., and Tuttle, M., Many Random Walks Are Faster Than One, available via: arXiv:math/0705.0467v2. (2007).

[3] Broder, A., Karlin, A. R., Raghavan, P., and Upfal, E., "Trading space for time in undirected *s-t* connectivity," *SIAM J. Comput.*, **23**: 324–334 (1994).

[4] Levin, David A., Peres, Y., and Wilmer, Elizabeth L., Markov chains and mixing times, American Mathematical Society, Providence (2008).

[5] Dembo, A., "Favorite Points, Cover Times and Fractals," *Lec. Notes Math.*, **1869**: 3–101 (2005).

[6] Feige, U., "A tight upper bound for the cover time of random walks on graphs," *Random Structures and Algorithms.*, **6**: 51–54 (1995).

[7] Feige, U., "A tight lower bound for the cover time of random walks on graphs," *Random Structures and Algorithms.*, **6**: 433–438 (1995).

[8] Feller, W., An introduction to probability theory and its applications. Vol. II. Third edition John Wiley & Sons, Inc., New York-London-Sydney (1971).

[9] Matthews, P., "Covering problems for brownian motion on spheres," *Ann. Probab.*, **16**: 189–199 (1988).

[10] Ohwa, T., and Shirai, T., "Joint distribution of the cover time and the last visited point of finite Markov chains," *Kyushu J. Math.*, **62**: 281–292 (2008).

[11] Rota, G. C., "On the foundations of combinatorial theory. I. Theory of Möbius functions," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **2**: 340–368 (1964).