Elect. Comm. in Probab. **5** (2000) 39–50

ELECTRONIC COMMUNICATIONS in PROBABILITY

LOOP-ERASED RANDOM WALKS, SPANNING TREES AND HAMILTONIAN CYCLES

PHILIPPE MARCHAL
Laboratoire de Probabilits, UMR 7599, Université Paris 6,
4 Place Jussieu, 75252 Paris cedex 05, France.
email: marchal@clipper.ens.fr

submitted February 12, 1999; *Final version accepted* December 3, 1999

AMS 1991 Subject classification: 60J10
Loop-erased random walk, spanning tree, Wilson's algorithm.

Abstract
We establish a formula for the distribution of loop-erased random walks at certain random times. Several classical results on spanning trees, including Wilson's algorithm, follow easily, as well as a method to construct random Hamiltonian cycles.

1 Introduction

Associated to a random walk X on a finite or denumerable state space V is the loop-erased random walk Y constructed as follows. Start X at some point $u \in V$ and, for each time n, consider the whole path from u to X_n. Then Y_n is the self-avoiding path obtained by removing the cycles in the order they appear (a formal definition is given in the sequel).

Loop-erasing is a quite natural method to construct self-avoiding paths and is therefore an object of interest in the study of self-avoiding random walks [6]. Moreover, in recent years, random spanning trees and forests have been much considered in probability (see e.g. [2]) and Wilson's algorithm [8], which uses the loop-erasing procedure, has proved to be a useful tool. One of the main interests in that context is the planar case, where the problem of conformal invariance plays major role. Some recent results in that domain can be found in Kenyon [5], although the methods are different, using domino tilings.

We determine in this note the distribution of Y at certain random times: the first hitting time by X of some subset $U \subset V$, or an independent time with geometric distribution. Our main formula (Theorem 1), which refines a result by Lawler, provides an easy proof of Wilson's algorithm as well as classical results on spanning trees.

We shall concentrate on the case when the state space V is finite. However, our results can be expressed in terms of the eigenvalues of the discrete Laplacian, see Section 2.3 (here we mean by discrete Laplacian an approximation of the second derivative, and not the graph-theoretical Laplacian). This suggests a possible generalization for a denumerable state space, although we shall not deal with the question here.

The remainder of this note is organized as follows. The law of Y stopped at certain random times is determined in Section 2, and the application to spanning trees is the topic of Section 3. Using the same methods, we treat the case of Hamiltonian cycles in Section 4.

2 The distribution of loop-erased paths

2.1 Definitions and notations

Consider a Markov chain X on a finite set V, and fix a point $u \in V$. Let X start at u and denote by X' the path process, that is, for every integer n,

$$X'_n = (X_0, X_1, \dots, X_n)$$

The loop-erasure Y_n is constructed as follows: put $a_0 = u$. Then define by recurrence l_k to be the last time X visits a_k before time n, and $a_{k+1} = X_{l_k+1}$. This yields a finite sequence that stops at some integer m when $l_m = n$, and we put

$$Y_n = (a_0, a_1, \dots, a_m)$$

One easily checks that the previous definition is equivalent to the following: put $a_0 = u$ and Y_0 is the empty path. By recurrence, if $Y_n = (a_0, a_1, \dots, a_m)$ and X_{n+1} does not belong to the support of Y_n, put

$$Y_{n+1} = (a_0, a_1, \dots, a_m, X_{n+1})$$

Else if $X_{n+1} = a_k$ for some $k \leq m$,

$$Y_{n+1} = (a_0, a_1, \dots, a_k)$$

That is, the cycles are erased in the order they appear.

We adopt the following notations. The weight of a subset $S = \{e_1, \dots, e_n\} \subset V \times V$ of directed edges is

$$w(S) = \prod_{i=1}^{n} p_{e_i} \in \mathbb{R}$$

and its valuation is

$$v(S)(t) = \prod_{i=1}^{n} (p_{e_i} t) \in \mathbb{R}[t]$$

Note that the edges need not be adjacent, and we shall consider the case when S is a tree in Section 3.

Let M be the transition matrix of the chain X. For every subset $U \subset V$, we denote by T_U the hitting time (for X) of U, by $G(t, x, x; U)$ the Green function for the random walk killed at T_U:

$$G(t, x, x; U) = \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = x,\ T_U < n) t^n = [Id - tM_U]^{-1}_{x,x}$$

and by M_U the matrix M restricted to the coefficients in $(V - U) \times (V - U)$. In this setting, if γ is a path, we shall write γ instead of $supp(\gamma)$ to simplify the notations.

2.2 The results

Let $U \subset V$ with $u \notin U$. We first determine the law of Y at T_U. Let Γ_U be the set of self-avoiding paths γ of the form $\gamma = (a_0, a_1, \dots, a_m)$ where $a_0 = u$, $a_m \in U$ and for every $k \leq m-1$, $a_k \notin U$. For such a path $\gamma \in \Gamma_U$, put $V_k = U \cup \{a_0, \dots a_{k-1}\}$, $1 \leq k \leq m-1$ and $V_0 = U$. A last-exit decomposition easily leads to the following formula, whose proof can be found in Section 3.1 of [6]:

$$\mathbb{E}_u[t^{T_U}; Y_{T_U} = \gamma] = \left(\prod_{k=0}^{m-1} G(t, a_k, a_k; V_k) \right) v(\gamma)(t) \tag{1}$$

The key remark is the following: for every $k \leq m-1$,

$$G(t, a_k, a_k; V_k) = \frac{\det[Id - tM_{V_k \cup \{a_k\}}]}{\det[Id - tM_{V_k}]} = \frac{\det[Id - tM_{V_{k+1}}]}{\det[Id - tM_{V_k}]} \tag{2}$$

This is a straightforward application of the Cramer formula for the inversion of matrices. Together with (1), this implies the following:

Theorem 1 *Let γ be the self-avoiding path starting at v. Then*

$$\mathbb{E}_u(t^{T_U},\ Y_{T_U} = \gamma) = \frac{\det(Id - tM_{\gamma \cup U})}{\det(Id - tM_U)} v(\gamma)(t)$$

A similar result appears in [6] for $t = 1$, but the correcting factor $\det(Id - M_{\gamma \cup U}) / \det(Id - M_U)$ is not computed explicitely. We shall express this factor in terms of the eigenvalues of the discrete Laplacian in the next subsection. Our formula also enables us to compute the expected run time of Wilson's algorithm, which plays a key role in computer science (see Section 3).

We can also determine the probability that $Y_T = \gamma$ for an an independent random time T with geometric distribution with parameter p, i.e. $P(T = n) = (1-p)p^n$.
It suffices to add a cemetery point ∂ to the set of vertices V, and to consider the random walk X' where the transition probabilities $p_{i,j}$ are replaced by $pp_{i,j}$ for $i, j \in V$, and the transition probability from any vertex $u' \in V$ to ∂ equals $1-p$. We then look at the loop-erased path at the first hitting time of ∂. Theorem 1 then becomes:

Corollary 1 *With the same notations as in Theorem 1, we have, for an independent time T with geometric distribution of parameter p,*

$$\mathbb{E}_u(t^T,\ Y_T = \gamma) = (1-p) \frac{\det(Id - ptM_\gamma)}{\det(Id - ptM)} \prod_{i=1}^{m} (p_{a_{i-1}, a_i} pt)$$

Note that for $t = 1$, when $p \to 1$, the limit distribution is the invariant measure on the set of self-avoiding paths starting at v for the chain Y. According to Corollary 1, this invariant measure is proportional to the measure that assigns to each path γ the weight

$$\det(Id - M_\gamma) w(\gamma)$$

In particular, the restriction of the invariant measure to Γ_U, for any subset $U \in V$ such that $u \notin U$, is proportional to the probability measure computed in Theorem 1. For an explanation of this fact, see the proof of the Markov chain tree theorem in [7].

2.3 The correcting factor

The quantities $\det(Id - M_{\gamma\cup U})$ and $\det(Id - M_U)$ can be expressed in terms of the eigenvalues of the discrete Laplacian, an operator defined as follows: for a function f: $V \to \mathbb{R}$, we put

$$\Delta f(x) = \mathbb{E}_x\left[f(X_1)\right] - f(x)$$

We say that λ is an eigenvalue of the Laplacian with Dirichlet boundary condition on $U' \subset V$ if there exists a nonzero function f satisfying

$$\begin{aligned} f(x) &= 0 && \text{if } x \in U' \\ \Delta f(x) &= \lambda f(x) && \text{if } x \notin U' \end{aligned}$$

Moreover, the multiplicity of λ is the the dimension of the vector space constituted by the functions f such that for every $x \notin U'$ and for some integer n,

$$(\Delta - \lambda Id)^n f(x) = 0$$

Let $\lambda_1, \ldots, \lambda_k$ be the eigenvalues of the Laplacian with Dirichlet boundary condition on U, and $a_1, \ldots, a_k$ their respective multiplicities. Then we have

$$\det(Id - M_U) = \prod_{i=1}^{k} (-\lambda_i)^{a_i} \tag{3}$$

We can define likewise the eigenvalues $\mu_1, \ldots, \mu_l$ of the Laplacian with Dirichlet boundary condition on $\gamma \cup U$ and their respective multiplicities $b_1, \ldots, b_l$. Then

$$\det(Id - M_{\gamma\cup U}) = \prod_{i=1}^{l} (-\mu_i)^{b_i} \tag{4}$$

Alternatively, the ratio

$$\frac{\det(Id - M_{\gamma\cup U})}{\det(Id - M_U)}$$

can be expressed in the following way. Say that ν is a semi-eigenvalue if there exists a nonzero function f satisfying:

$$\begin{aligned} f(x) &= 0 && \text{if } x \in U \\ \Delta f(x) &= \nu f(x) && \text{if } x \in \gamma - U \\ \Delta f(x) &= 0 && \text{if } x \notin \gamma \cup U \end{aligned}$$

Define the multiplicity of ν as the dimension of the the vector space constituted by the functions f such that

$$\begin{aligned} (\Delta - \nu)^n f(x) &= 0 && \text{for some } n \text{ if } x \in supp(\gamma) - U \\ (\Delta)^n f(x) &= 0 && \text{for some } n \text{ if } x \notin (supp(\gamma) \cup U) \end{aligned}$$

Let $\nu_1, \ldots, \nu_p$ be the semi-eigenvalues and $c_1, \ldots, c_p$ their respective multiplicities. Then

$$\frac{\det(Id - M_{\gamma\cup U})}{\det(Id - M_U)} = \prod_{i=1}^{p}(-\nu_i)^{c_i} \tag{5}$$

The proof of (5) is an easy exercise of linear algebra. Write the matrix M_U where the last coefficients correspond to the vertices of the path γ, i.e.

$$M_U = \begin{pmatrix} M_{\gamma\cup U} & A \\ B & C \end{pmatrix}$$

Define the polynomial

$$P(\lambda) = \det \begin{pmatrix} M_{\gamma\cup U} & A \\ B & C - \lambda Id \end{pmatrix}$$

Let $P(\lambda) = a_k\lambda_k + \ldots + a_i\lambda + a_0$. Then the product of the roots of P is

$$\frac{a_0}{a_k} = \frac{\det(Id - M_U)}{\det(Id - M_{\gamma\cup U})}$$

On the other hand, the roots of P are the complex numbers λ such that there exists a vector

$$F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$$

satisfying

$$\begin{pmatrix} M_{\gamma\cup U} & A \\ B & C - \lambda Id \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

or, equivalently,

$$\begin{pmatrix} M_{\gamma\cup U} & A \\ B & C \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda F_2 \end{pmatrix}$$

That is, the roots of P are the semi-eigenvalues, which proves (5).

3 Application to spanning trees

A spanning tree of a graph is defined as a connected, acyclic subgraph containing all the vertices. Spanning trees are a classical problem in computer science, see the references in [8]. Their study in probability has been the the subject of interesting developements in recent years, see for instance [2] and the references therein.
A rooted spanning tree Γ on V can be viewed as a collection of directed edges where each edge points towards the root. Hence we can define its weight with respect to the random walk X, as in Section 2.1. Wilson's algorithm is a method for constructing a random spanning tree with probability proportional to the weight [8]. Fix the root R and chose any ordering on $V - \{R\}$. We define by recurrence a growing sequence of trees $T(n)$ by:

- $T(0) = \{R\}$
- If $T(n)$ spans V, we are done.

- Else, let v be the minimal vertex in the ordering that is not in $T(n)$, and perform the loop-erased random walk Y starting at v up to the first hitting time of $T(n)$. Add this loop-erased path to $T(n)$ to create $T(n+1)$.

If the chain is irreducible, it is clear that the sequence ends almost surely and that the final tree is a spanning tree. Furthermore we have:

Theorem 2 (Wilson) *Wilson's algorithm constructs a random spanning tree rooted at R with probability proportional to the weight w. In particular, it is independent of the ordering chosen on $V-\{R\}$.*

We want to derive Wilson's theorem from Theorem 1.
Let Γ be a spanning tree rooted at R, and choose an ordering on $V-\{R\}$. We begin by remarking that any sequence of paths, obtained by running Wilson's algorithm, that constructs Γ will in fact construct the same growing sequence of (deterministic) subtrees $(U(0), U(1), \ldots, U(n))$ of Γ. Wilson's theorem is a consequence of the following lemma:

Lemma 1 *For any integer $k \leq n$, denote by S_k the support of $U(k)$. Recall that $T(0), T(1), \ldots$ is the growing sequence of random trees constructed by Wilson's algorithm. Then we have for every k,*

$$\mathbb{P}(T(k)=U(k)) = \frac{\det(Id-M_{S_k})}{\det(Id-M_{\{R\}})} w(U(k))$$

Wilson's result follows by taking $k=n$.

Proof
We proceed by induction on k. For $k=1$, as $U(1)$ is a self-avoiding path from the first vertex in $V-\{R\}$ to the root, this is nothing else than Theorem 1.
Suppose that the result is established for $k-1$. Let v_k be the minimal vertex in $V-S_k$ and γ_k the path from v_k to S_{k-1} in the tree $T(k)$. Then the conditional probability

$$\mathbb{P}(T(k)=U(k) | T(k-1)=U(k-1))$$

is the probability that the random walk, started at v_k up to the hitting time of S_{k-1}, has loop-erasure γ_k. Using Theorem 1, we see that this probability equals

$$\mathbb{P}(T(k)=U(k) | T(k-1)=U(k-1)) = \frac{\det(Id-M_{S_k})}{\det(Id-M_{S_{k-1}})} w(\gamma_k)$$

which completes the proof.

□

An easy generalization of Lemma 2 is the following

Corollary 2 *Let Γ' be a tree rooted at R, that is not necessarily a spanning tree. Then the probability that a random spanning tree rooted at R contains Γ' is*

$$\frac{\det(Id-M_{\Gamma'})}{\det(Id-M_{\{R\}})} w(\Gamma')$$

Proof
The probability that a random spanning tree rooted at R contains Γ' does not depend on the ordering on $V-\{R\}$. Let $L_1,\dots,L_n$ be the leaves of Γ' and chose an ordering on $V-\{R\}$ where the first n vertices are $L_1,\dots,L_n$. Then the probability we are considering is exactly the probability that $T(n)=\Gamma'$, which is given by Lemma 2.

□

Remark. If the random walk is reversible, it is well-known that it is associated to an electric network. In that case, given any set B of directed edges, the probability that all the edges in B belong to the random spanning tree can be computed by the transfer current theorem, see [3]. Our result only enables us to handle the case when B is a tree rooted at R. Nevertheless, it also holds for non-reversible Markov chains.
Another connection of spanning trees to Markov chains is the Markov chain tree theorem [7], which can also be easily derived by our method:

Corollary 3 (Markov chain tree theorem) *The stationary distribution of the Markov chain X is proportional to the measure that assigns to every vertex u the mass*

$$\sum_{S\in\mathcal{T}_u} w(S)$$

where $\mathcal{T}_u$ is the set of spanning trees rooted at u.

Proof
As Wilson's theorem constructs with probability 1 a spanning tree rooted at u, taking $k=n$ in Lemma 2 yields:

$$\sum_{S\in\mathcal{T}_u} w(S)=\det(Id-M_{\{u\}}) \tag{6}$$

But it is well-known (see [9], chapter 1) that the measure

$$\sum_{u\in V}\det(Id-M_{\{u\}})\delta_u$$

is invariant.

□

Remark. Formula (6) can be seen as a generalization of Khirchoff's formula counting the number of spanning trees in a given graph [4].

As a last application of our method, let us compute the expected number of steps in Wilson's algorithm. This plays a key role in computer science, see [8]. Denote by T the number of steps of the random walk up to the complete construction of the spanning tree rooted at R. Then we have

Proposition 1 *Let n be the number of points in $V-\{R\}$. Then:*
(i) The distribution of T is given by:

$$\mathbb{E}(t^T)=\frac{\det(Id-M_{\{R\}})}{\det(Id-tM_{\{R\}})}t^n$$

(ii) In particular,

$$\mathbb{E}(T) = tr((Id - M_{\{R\}})^{-1}) \tag{7}$$

Proof
As a straightforward generalization of Lemma 2 we have

$$g(t) := \mathbb{E}(t^T) = \frac{\sum_{Z \in \mathcal{T}_R} w(Z) t^n}{\det(Id - tM_{\{R\}})}$$

where $\mathcal{T}_R$ denotes the set of spanning trees rooted at R. According to (6), $\sum_{Z \in \mathcal{T}_R} w(Z) = \det(Id - M_{\{R\}})$, which yields (i).

The computation of $\mathbb{E}(U) = g'(1)$ leads to

$$g'(1) = \frac{n - \frac{d}{dt}(\det(Id - tM_{\{R\}}))_{t=1}}{\det(Id - M_{\{R\}})}$$

But we have

$$\frac{d}{dt}(\det(Id - tM_{\{R\}}))_{t=1} = \frac{d}{dt}(\det(Id - (1+t)M_{\{R\}}))_{t=0}$$

and $\det(Id - (1+t)M_{\{R\}})$ is a polynomial in t whose coefficient of degree 1 is

$$\sum_{i,j} (-1)^{i+j} M_{i,j} \det((Id - M_{\{R\}})^{i,j})$$

where $M_{i,j}$ is the (i,j) coefficient of the matrix M and $(Id - M_{\{R\}})^{i,j}$ is the matrix $Id - M_{\{R\}}$ where we erase the i-th line and j-th column. Developping the determinant $\det(Id - M_{\{R\}})$ along the i-th line we get:

$$\sum_j M_{i,j}(-1)^{i+j} \det((Id - M_{\{R\}})^{i,j}) = \det(Id - M_{\{R\}}) - \det((Id - M_{\{R\}})^{i,i})$$

for every i. Summing on i yields

$$\mathbb{E}(T) = \frac{\sum_{i \in V - \{R\}} \det(Id - M_{\{R,i\}})}{\det(Id - M_{\{R\}})}$$

and this latter quantity is equal to $tr((Id - M_{\{R\}})^{-1})$, according to the Cramer formula for the inversion of matrices.

□

Remark. The expected number of steps is expressed in [8] as

$$\mathbb{E}(T) = \sum_i \pi(i)(t_{i,R} + t_{R,i}) \tag{8}$$

where $t_{x,y}$ is the expected time for the Markov chain to go from x to y, and π is the invariant probability measure. To see that (7) and (8) are equivalent, expand $(Id - M_{\{R\}})^{-1}$ as a power series

$$\sum_{k \geq 0} M_{\{R\}}^k$$

to see that the (i,i) coefficient in $(Id - M_{\{R\}})^{-1}$ is the expected number $G_R(i,i)$ of visits to i for the Markov chain started at i and absorbed at R. But it is classical that $G_R(i,i)$ divided by the commute time $(t_{i,R} + t_{R,i})$ equals the stationary distribution $\pi(i)$ (see [1]).
Therefore

$$((Id - M_{\{R\}})^{-1})_{i,i} = \pi(i)(t_{i,R} + t_{R,i})$$

Summing this over all the states $i \in V$ identifies (8) with (7).
This remark is due to Yuval Peres.

4 Hamiltonian cycles

The loop-erasing procedure also enables us to construct Hamiltonian cycles, i.e. simple cycles spanning all the vertices of V. Indeed, let the random walk start at some vertex $u \in V$, denote by $0 = T_0 < T_1 < T_2 \dots$ the successive return times of the random walk to u and put

$$n = \min\{m,\ supp(Y_{T_m - 1}) = V\}$$

Then if we add at the end of the path $Y_{T_n - 1}$ the point u, we get a Hamiltonian cycle. In other words, each time the random walk returns to u, the loop-erasing procedure yields a simple cycle containing u. We erase this cycle if it is not Hamiltonian and keep it if it is Hamiltonian.

Theorem 3 *Let H be the Hamiltonian cycle constructed by the algorithm described above and T the run time. Denote by $\mathcal{C}$ the set of Hamiltonian cycles, and consider a cycle $\gamma \in \mathcal{C}$. Then*

$$\mathbb{E}_u(t^T,\ H = \gamma) = \frac{v(\gamma)(t)}{\det(Id - tM) + \sum_{\gamma' \in \mathcal{C}} v(\gamma')(t)} \tag{9}$$

As a consequence, the algorithm yields a random Hamiltonian cycle with probability proportional to the weight.

Note that the joint law of (T, H) does not depend on the starting point u, which is not obvious by simple path transform arguments.

Proof
To prove this result, let us introduce some definition. An excursion will refer to a path $(a_0 = u, \dots, a_n = u)$ such that for $0 < j < n$, $a_j \neq u$. An excursion is Hamiltonian if the loop-erasure of the path $(a_0, \dots, a_{n-1})$ contains all the vertices of V. In that case, by convention, we say that the loop-erasure of the excursion is the loop-erasure of $(a_0, \dots, a_{n-1})$. We denote by $\mathcal{E}$ the set of Hamiltonian excursions and by E be the set of non-Hamiltonian excursions. Consider the Hamiltonian cycle

$$\gamma = (a_0 = u, a_1, \dots, a_{|V|} = u)$$

and denote by $\mathcal{E}_\gamma$ the set of Hamiltonian excursions whose loop-erasure is $(a_0 = u, a_1, \dots, a_{|V|-1})$. Of course, $\mathcal{E}_\gamma \subset \mathcal{E}$. We want to compute

$$\mathbb{E}_u(t^T,\ Y = \gamma) = \frac{1}{1 - \sum_{e \in E} v(e)(t)} \sum_{\delta \in \mathcal{E}_\gamma} v(\delta)(t)$$

If the random walk X starts at u, conditioning on X_1 and applying Theorem 1 we get:

$$\sum_{\delta\in\mathcal{E}_\gamma} v(\delta)(t) = \frac{v(\gamma)(t)}{\det(Id - tM_{\{u\}})}$$

The generating function of from u to u is given by:

$$\frac{1}{1-\sum_{e\in E} v(e)(t) - \sum_{\gamma'\in\mathcal{C}}\sum_{\delta\in\mathcal{E}_{\gamma'}} v(\delta)(t)} = \frac{\det(Id - tM_{\{u\}})}{\det(Id - tM)}$$

since any path from u to u is a sequence of excursions. Hence

$$\begin{aligned}\frac{1}{1-\sum_{e\in E} v(e)(t)} &= \frac{\det(Id - tM_{\{u\}})}{\det(Id - tM) + \det(Id - tM_{\{u\}})\sum_{\gamma'\in\mathcal{C}}\sum_{\delta\in\mathcal{E}_{\gamma'}} v(\delta)(t)} \\ &= \frac{\det(Id - tM_{\{u\}})}{\det(Id - tM) + \sum_{\gamma'\in\mathcal{C}} v(\gamma')(t)}\end{aligned}$$

Putting pieces together we recover (9) and the second assertion follows by taking $t = 1$.

$\square$

Note that one can run several times Wilson's algorithm rooted at u until the spanning tree constructed is a linear one, i.e. is constituted by a single path $(a_1, \ldots a_n = u)$, and then take the Hamiltonian cycle $(a_0 = u, a_1, \ldots a_n = u)$. In that case one constructs a random Hamiltonian cycle with probability proportional to

$$\prod_{i=2}^{|V|} p_{a_{i-1},a_i}$$

but this is different from the weight of the cycle, which is

$$\prod_{i=1}^{|V|} p_{a_{i-1},a_i}$$

The expected run time of the algorithm is given by the following:

Corollary 4 *The expected time to construct a random Hamiltonian cycle by the algorithm of Theorem 3 is:*

$$\mathbb{E}(T) = \frac{\sum_{t\in\mathcal{T}} w(t)}{\sum_{\gamma\in\mathcal{C}} w(\gamma)}$$

where $\mathcal{T}$ is the set of rooted spanning trees of V (rooted at any vertex) and $\mathcal{C}$ is the set of Hamiltonian cycles.

Proof
We want to compute $f'(1)$, f being the function defined by

$$f(t) = \mathbb{E}_u(t^T) = \frac{\sum_{\gamma\in\mathcal{C}} v(\gamma)(t)}{\det(Id - tM) + \sum_{\gamma\in\mathcal{C}} v(\gamma)(t)}$$

As $\det(Id - M) = 0$ we get:

$$f'(1) = \frac{\frac{d}{dt}(\det(Id - tM))_{t=1}}{\sum_{\gamma\in\mathcal{C}} w(\gamma)}$$

Computing $(\frac{d}{dt}\det(Id - tM))_{t=1}$ as in the proof of Proposition 1 we find

$$\left(\frac{d}{dt}\det(Id - tM)\right)_{t=1} = -|V|\det(Id - M) + \sum_{u\in V}\det(Id - M_{\{u\}})$$

Using (6) and the fact that $\det(Id - M) = 0$ we get the result.

□

A naive method to construct a random Hamiltonian cycle is the following. Start the random walk X at u and define the loop time L as the first time a loop appears:

$$L = \min\{n > 0,\ X_n \in \{X_0, \ldots, X_{n-1}\}\}$$

Then we stop X at L and keep the whole path if it is a Hamiltonian cycle (in which case $L = |V|$). Else we restart X at u, and so on. It is clear that this procedure yields a random Hamiltonian cycle with probability proportional to the weight. The number of steps T' of the algorithm has the following distribution:

$$\mathbb{E}_u(t^{T'}) = \frac{\mathbb{P}_u((X_0, \ldots, X_L) \in \mathcal{C})t^{|V|}}{1 - \sum_{n\leq|V|}\mathbb{P}_u(L = n, (X_0, \ldots, X_L) \notin \mathcal{C})t^n}$$

Remark that

$$\mathbb{P}_u((X_0, \ldots, X_L) \in \mathcal{C}) = \sum_{\gamma\in\mathcal{C}} w(\gamma)$$

Thus

$$\mathbb{E}_u(T') = \frac{\mathbb{E}_u(L)}{\sum_{\gamma\in\mathcal{C}} w(\gamma)}$$

In particular, the expected run time of the naive algorithm depends on the starting point. Also remark that $\mathbb{E}_u(T') \geq \mathbb{E}(T)$ since $\mathbb{E}_u(L) \geq 1 \geq \sum_{t\in\mathcal{T}} w(t)$.

Acknowledgements. I thank Yuval Peres for the remark at the end of Section 3, and the referee for some useful comments on the first draft.

References

[1] D. Aldous and J. Fill: *Reversible Markov chains.* To appear.

[2] I. Benjamini, R. Lyons, Y. Peres and O. Schramm: Uniform spanning forests. *To appear.*

[3] R. Burton and R. Pemantle: Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances. *Ann. Probab.* **21** (1993), no. 3, 1329–1371.

[4] G. Kirchhoff: Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird. *Ann. Phys. Chem.* **72** (1847) 497-508.

[5] R. Kenyon: The asymptotic determinant of the discrete Laplacian. *Preprint* (1999)

[6] Gregory F. Lawler, Loop-erased random walk, in *Perplexing Problems in Probability*, Birkhauser-Boston (1999), 197–217.

[7] R. Lyons and Y. Peres: *Probability on Trees and Networks.* To appear.

[8] J. Propp and D. Wilson: How to Get a Perfectly Random Sample From a Generic Markov Chain and Generate a Random Spanning Tree of a Directed Graph. *Journal of Algorithms* **27** (1998) 170-217.

[9] E. Seneta: *Non-negative matrices and Markov chains.* 2nd ed. (1981) Springer.